Optimal Quantization and Approximation in Source Coding and Stochastic Control

by

Naci Saldi

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

Queen's University
Kingston, Ontario, Canada
September 2015

Abstract

This thesis deals with non-standard optimal quantization and approximation problems in source coding and stochastic control.

The first part of the thesis considers randomized quantization. Adapted from stochastic control, a general representation of randomized quantizers that is probabilistically equivalent to common models in the literature is proposed via mixtures of joint probability measures induced by deterministic quantizers. Using this general model, we prove the existence of an optimal randomized quantizer for the generalized distribution preserving quantization problem. A Shannon theoretic version of this source coding problem is also considered, in which an optimal (minimum distortion) coding of stationary and memoryless source is studied under the requirement that the quantizer's output distribution also be stationary and memoryless possibly different than source distribution. We provide a characterization of the achievable rate region where the rate region includes both the coding rate and the rate of common randomness shared between the encoder and the decoder.

In the second part of the thesis, we consider the quantization problems in stochastic control from viewpoints of information transmission and computation. The first problem studies the finite-action approximation (via quantization of the action space) of deterministic stationary policies of a discrete time Markov decision process (MDP), while the second problem considers finite-state approximations (via quantization of the state space) of discrete time Markov decision process. Under certain continuity conditions on the components of the MDP, we establish that optimal policies for the finite models can approximate with arbitrary precision optimal deterministic stationary policies for the original MDP. Combining these results leads to a constructive scheme for obtaining near optimal solutions via well known algorithms developed for finite state/action MDPs. For both problems, we also obtain explicit bounds on the approximation error in terms of the number of representation points in the quantizer, under further conditions.

Acknowledgments

First, I would like to thank my advisors Professor Serdar Yüksel and Professor Tamás Linder for their generous support, encouragement and enlightening guidance throughout the development of this thesis. I have been really fortunate to work in person with such wonderful scholars. I have learned incredibly much from them during my PhD studies from which I will benefit during my entire research career. It would not have been possible for me to successfully finish my PhD without their insightful discussions, advices, and comments. Thank you very much!

I am also very thankful to Professor Abdol-Reza Mansouri for both accepting to be my thesis examiner and teaching me his wonderful control theory course, which is the best control theory course I have ever taken. I would like to thank Professor Glen Takahara for his excellent probability theory course from which I benefit a lot. Moreover, I would like to thank my supervisory and thesis committee members, Professor Bahman Gharesifard, Professor Maxim Raginsky, Professor Steven D. Blostein, and Professor Boris Levit for their time and effort. Finally, I thank our graduate secretary Jennifer Read for her enormous help starting from my first day in Canada up to the end of my studies at Queen's University.

I want to thank my mathematics teacher Orhan Varol from Zonguldak Science High School, from whom I learned what mathematics mean. I still keep the notebook of his mathematics course taught at the second year of the high school. Furthermore, I am deeply grateful to my elementary school teacher Nevin Ermiş, who have been a great resource of support and guidance during my early school years.

I am very thankful to Professor Peter Taylor for his continuous support to me and my wife starting the very beginning of our stay in Canada. He has been like our Canadian parent. I am really grateful to have met such a unique personality in my life time. I will always remember the homy dinners we had with you, Professor Taylor. Many thanks!

I had a great time in Kingston with very good friends. I want to thank all my friends, in particular, Saber Jafarpour, Marcos Vasconcelos and Shahab Asoodeh.

I would like to thank my parents Muzaffer Saldi and Şahinde Saldi. By buying me the book 'A Mathematician's Apology' when I was a teenager, they sparked my passion for math.

My brother Necdet Saldi's support during my expatriate years in Canada is priceless. I am really grateful to have such a wonderful person as my brother. I am deeply thankful and indebted to you, Necdet. Thank you very much!

I would also to thank my parents-in-law Ayhan Keskin and Ismail Keskin for sending their beloved daughter with me to Canada faraway from homeland. I thank my brother-in-law Nezih Keskin for adding humor to my life.

Last but definitely not least, I am deeply grateful to my wife Rana Keskin for her constant support, tolerance, guidance, and most importantly love. To allow me to do my PhD in Canada, she put aside her career without any complaint and decided to move with me to Canada in a heartbeat. I am and will always be deeply indebted to her all my life. Without her presence I could not have even started my doctoral

studies. From the early beginning of this process, her support to all my decisions is priceless. She has always been a source of inspiration in my life. I still remember the moment that a proof of one of the important results in this thesis occurred to me while we were chatting to relax. I do not think that words are enough to express my appreciation to you, my love. This thesis is humbly dedicated to yourself!

To my one and only love, Rana

whom I love

and

will love

till the end of my life...

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

Quantization is a method of representing a quantity (e.g., analog signal, random variable, etc.) taking values in a large set by a quantity taking values in a smaller discrete set (in general, a subset of the large set). The simplest example for quantization is analog-to-digital conversion in which a continuous quantity (usually amplitude of some signal) is converted to a digital number. As the conversion is not one-to-one, it unavoidably introduces some error, usually called the *quantization error*. The main goal in quantization is to reduce this error to the smallest value possible. The history of the quantization dates at least as far back as the 1948 paper of Oliver, et al. [70], quantization was used to transmit the amplitude of the sampled signal in pulse-code modulation systems.

After the introduction of the rate-distortion theory in Shannon's seminal paper [95], a commonly accepted approach to model a quantity to be quantized is to view it as an output of some random process, called the source. Such models arise, for instance, in communication systems. In this situation, the quality of quantization is usually evaluated in terms of its expected quantization error, defined as some function of the source and its quantized representation.

The classical optimal quantization problem involves the minimization of the expected quantization error given the number of representation points and the probabilistic law of the random process. In general, the random process is taken to be a single random variable or a stationary and/or memoryless random process. The fundamental assumption imposed on the random process is that it is realized by an external system which is operating independently of quantization, i.e., if the random process is stationary and memoryless with a given common law, it will remain the same regardless of the quantization applied. This property in general simplifies the quantization problem and allows one to obtain quite strong results, such as the calculation of the minimum achievable quantization error given the number of representation points, in addition to proving the existence of optimal quantization.

With the emergence of control applications such as the control of decentralized and distributed systems, digital control of automatic systems, and control over communication systems, there has been a growing need for the information theoretic treatment of the problems in control theory. In particular, the problem of optimal quantization for information transmission from a plant/sensor to a controller and from a controller to an actuator turns out to be crucial for the networked control applications, which has lead to a surge of research activity in this direction. We refer to [107] and references therein for a good introduction and an extensive literature review. As well, in the classical theory of stochastic optimal control, although structural properties of the optimal policies have been studied extensively in the literature, computing such policies is still a substantial problem for systems with a large number of states. Since a general approach is to construct a reduced model by quantizing the set of states of the system under consideration, this problem can be viewed, to a certain degree, as an optimal quantization problem.

While at first glance one might hope that these problems can be solved by applying readily available methods in information theory developed for optimal quantization problems, it turns out that systems and design objectives in control theory and in information theory have quite different characteristics.

One main difficulty that prevents the direct application of classical methods of optimal quantization to stochastic control is the characteristics of the quantization error. In stochastic control, the purpose of the decision maker is to minimize overall cost incurred at each time step by applying appropriate control policies. Since the application of the quantization to the states and/or to the actions will cause to a deviation from the probabilistic behaviour of the state and action processes at each time step, the effect of the quantization on the future must also be taken into account in these problems. Such deviations may cause drastic changes in the system performance such as instability or significant jump in the overall cost. However, in the classical optimal quantization theory the purpose is simply minimize the expectation of a single-letter distortion measure given the current and past observations. Hence, it is unlikely that the classical approach for optimal quantization problem will lead to an optimal solution to the quantization problem in stochastic control.

Another potentially important issue that differentiates classical quantization and quantization in control systems is the difference in objectives. In decentralized and distributed control applications, the decision makers are asked to perform additional tasks, such as coordination or seeking consensus, in addition to cost minimization.

In such cases, these requirements have to be included into the problem formulation in an appropriate way, and therefore, the design of quantization should take into account these additional requirements. To do this we need to abandon, to some extent, the classical approaches developed in information theory, and investigate new solution techniques for such non-standard quantization problems, especially from the stochastic control perspective.

Hence, there is a need to re-investigate and generalize, if necessary, some concepts in the theory of quantization and source coding in order that they can be applied to stochastic control. This investigation is also beneficial for the field of information theory itself (as it is demonstrated in Chapter 2 of this thesis).

In this thesis, we investigate quantization, in particular randomized quantization, from the stochastic control perspective and apply the results in the approximation problem for stochastic control. The first part of the thesis is devoted to randomized quantization. Here we propose a general model (adapted from stochastic control) which formalizes the definition of randomized quantization. This general model is, then, used to solve an optimal quantization problem with unconventional objectives; namely, we prove the existence of optimal randomized quantizer for a generalized distribution preserving quantization problem. A Shannon-theoretic version of this problem is also considered where a stationary and memoryless source is encoded subject to a distortion constraint and an additional requirement that the reproduction also be stationary and memoryless with a given distribution. We completely characterize the rate-distortion region, where the rate region measures both the coding rate and the rate of common randomness shared between the encoder and the decoder. In the second part of the thesis, we consider the quantization problem in stochastic control from both the information transmission point of view (i.e., quantization of actions) and the computational point of view (i.e., quantization of *states* and actions). The first part concerns the finite-action approximation of stationary policies of a discrete-time Markov decision process (MDP), while the second part studies the finite-state approximations of discrete time MDP. For both problems, it is shown that approximating models give policies which approximate optimal stationary policies of the original models with arbitrary precision. We also obtain explicit rates of convergence bounds quantifying how the approximation improves as the size of the approximating finite action and state spaces increases, under further conditions.

In the next two sections we present some background information and literature review on quantization, randomized quantization, and Markov decision processes (MDPs). The literature review on MDPs mainly focuses on the approximation of optimal policies rather than structural properties of optimal policies. These sections serve as an introduction material for the problems that will be dealt with in the subsequent chapters.

1.2 Quantization and Randomized Quantization

In quantization, a source random variable X taking values in an infinite set, or a finite set with large cardinality, is represented by an output Y taking values from a prespecified set with low cardinality. It is used to compress data in order to store and send it digitally in communication systems. The mapping that realizes the quantization process is called a *quantizer*. The set of source values is usually called the source alphabet and the set of output values are chosen from a set that is called the reproduction alphabet. Let X and Y denote the source and reproduction alphabets, respectively. Here X and Y can be finite or countable sets or $\mathbb{R}^n$ for some $n \geq 1$ or abstract measurable spaces. Very often, $\mathsf{X} = \mathsf{Y}$. An M-level quantizer (M is a positive integer) from the source alphabet X to the reconstruction alphabet Y is a mapping $q : \mathsf{X} \to \mathsf{Y}$ whose range $q(\mathsf{X}) = \{q(x) : x \in \mathsf{X}\}$ contains *at most* M points of Y. Hence, $Y = q(X)$. The mapping q is assumed to be measurable. When X and Y are metric spaces (or more generally topological spaces), measurability will refer to Borel measurability. If $\mathcal{Q}_M$ denotes the set of all M-level quantizers, then our definition implies $\mathcal{Q}_M \subset \mathcal{Q}_{M+1}$ for all $M \geq 1$. The rate of an M-level quantizer q is defined as

$$R(q) = \log_2 |q(\mathsf{X})|$$

which (approximately) gives the number of bits needed to represent $Y = q(X)$ using binary strings of a fixed length. Let $\rho : \mathsf{X} \times \mathsf{Y} \to \mathbb{R}$ be a nonnegative function, called the distortion measure. If the source X is an X-valued random variable with distribution μ, then the distortion associated with quantizer q is the expectation

$$D(q) = \mathbb{E}\big[\rho(X, Y)\big] = \mathbb{E}\big[\rho(X, q(X))\big] = \int_{\mathsf{X}} \rho(x, q(x))\mu(dx).$$

The performance of a quantizer q is usually characterized by its rate $R(q)$ and its distortion $D(q)$ which are obviously conflicting quantities. The balancing of these two quantities is usually called rate-distortion tradeoff. The goal is to find the optimal tradeoff between rate and distortion, i.e., given $R \geq 0$ or $d \geq 0$

$$\textbf{(D)}\ \text{minimize } D(q)$$
$$\text{subject to } R(q) \leq R.$$

or

$$\textbf{(R)}\ \text{minimize } R(q)$$
$$\text{subject to } D(q) \leq d.$$

Observe that the constraint in problem **(D)** can be written as $q \in \mathcal{Q}_M$ where $M = 2^R$. The existence of the optimal quantizers can be shown under various general conditions when the source and reproduction alphabets are finite dimensional Euclidean spaces; see [1, 2, 74] and references therein. The most general was by Abaya and Wise [1], who showed the existence of the optimal quantizer q^* in problem **(D)** for $\mathsf{X} = \mathsf{Y} = \mathbb{R}^n$ and arbitrary source distribution μ when the distortion measure ρ is given by $\rho(x, y) = C_0(\|x - y\|)$, where $C_0 : [0, \infty) \to [0, \infty)$ is a lower semi-continuous and nondecreasing function.

A more involved and operationally important problem than showing the existence of an optimal quantizer for a given rate level $R \geq 0$ is the evaluation of the minimum distortion

$$D(R) := \big\{D(q) : R(q) \leq R\big\}$$

achievable for source distribution μ. The same is true for the minimum achievable rate given some distortion level d. However, it is possible to obtain a single-letter expression for $D(R)$ in the limit of large block lengths. Consider a stationary and memoryless source $\{X_n\}_{n \geq 1}$ with common distribution μ. Let $x^n = (x_1, \ldots, x_n)$ and $y^n = (y_1, \ldots, y^n)$ denote generic elements of X^n and Y^n, respectively. Define the distortion between sequences x^n and y^n as

$$\rho_n(x_n, y_n) = \frac{1}{n} \sum_{i=1}^{n} \rho(x_i, y_i).$$

By an abuse of notation, for any $R \geq 0$ let $\mathcal{Q}_{n,R}$ denote the set of 2^{nR}-level quantizers mapping X^n into Y^n . Hence, the distortion of quantizer $q \in \mathcal{Q}_{n,R}$ is given by

$$D(q) = \mathbb{E}\big[\rho_n(X^n, q(X^n))\big].$$

Then

$$D_n(R) \coloneqq \inf\big\{D(q) : q \in \mathcal{Q}_{n,R}\big\}$$

is the minimum distortion achievable by quantizers in $\mathcal{Q}_{n,R}$. We also define

$$D^I(R) \coloneqq \inf\big\{\mathbb{E}[\rho(X,Y)] : X \sim \mu, I(X;Y) \leq R\big\},$$

where $I(X;Y)$ denotes the mutual information between random variables X and Y [25]. A classical source coding theorem [25, Chapter 10] states that

$$\lim_{n \to \infty} D_n(R) = D^I(R);$$

that is, in the limit of large block lengths n the minimum achievable distortion is close to $D^I(R)$ when the normalized (by block length n) rate is less than R. Analogously, if we define

$$R_n(D) \coloneqq \inf\bigl\{R : q \in \mathcal{Q}_{n,R} \text{ and } D(q) \leq D\bigr\},$$
$$R^I(D) \coloneqq \inf\bigl\{I(X;Y) : X \sim \mu, \mathbb{E}[\rho(X,Y)] \leq D\bigr\},$$

then the same theorem also implies that

$$\lim_{n\to\infty} R_n(D) = R^I(D).$$

Indeed, the functions $D^I(\,\cdot\,)$ and $R^I(\,\cdot\,)$ are inverses of each other. In information theory, they are called the distortion-rate and rate-distortion functions, respectively. They provide an achievable lower bound for the distortion (resp., rate) for any given admissible rate (resp., distortion) levels.

In randomized quantization, the quantizer is randomly selected from a given collection of quantizers and then used to map the input signal. In Chapter 2 we give three (probabilistically) equivalent definitions of randomized quantization. For the purposes of this chapter, we can think of an M-level randomized quantizer as a pair (q, ν) where $\mathsf{q} : \mathsf{X} \times \mathsf{Z} \to \mathsf{Y}$ is a measurable mapping such that $\mathsf{q}(\,\cdot\,, z)$ is an M-level quantizer for all $z \in \mathsf{Z}$ and ν is a distribution of the randomizing Z-valued random variable Z with independent of the source X. Hence, $Y = \mathsf{q}(X, Z)$. Analogous with the deterministic case, the distortion associated with a randomized quantizer (q, ν)

is the expectation

$$D(\mathsf{q},\nu) = \mathbb{E}\big[\rho(X,Y)\big] = \mathbb{E}\big[\rho(X,\mathsf{q}(X,Z))\big] = \int_{\mathsf{Z}}\int_{\mathsf{X}} \rho(x,\mathsf{q}(x,z))\mu(dx)\nu(dz)$$

and the rate of (q,ν) is given by

$$R(\mathsf{q},\nu) = \log_2(M).$$

Although allowing for randomization in the quantization procedure does not improve the optimal rate-distortion tradeoff at a fixed quantizer rate (i.e., problem **(D)**), other measures of performance may be improved by using randomized quantizers.

1.3 Markov Decision Processes

A discrete time Markov decision process (MDP) is a mathematical model for sequential decision making under stochastic uncertainty. It is useful in modelling a wide range of systems in engineering, economics, and biology (see [36], [54]). An MDP can be specified by the following components: (i) The state space X and the action space A, where X and A are Borel spaces (i.e., Borel subsets of complete and separable metric spaces), (ii) the transition probability $p(\,\cdot\,|x,a)$ on X given $\mathsf{X}\times\mathsf{A}$ which gives the probability of the next state given that the current state-action pair is (x,a), (iii) the one-stage cost functions $c_t : \mathsf{X}\times\mathsf{A} \to \mathbb{R}$, $t = 0,1,2,\ldots$ (in general $c_t = c$ for some $c : \mathsf{X}\times\mathsf{A}\to\mathbb{R}$), and (iv) the initial distribution μ on X.

If X_t and A_t denote the state and action variables at time step t, then with these definitions, we have

$$\mathsf{Pr}\big\{X_0 \in \cdot\big\} = \mu(\cdot)$$

$$\mathsf{Pr}\big\{X_{t+1} \in \cdot \,\big|X_{[0,t]}, A_{[0,t]}\big\} = \mathsf{Pr}\big\{X_{t+1} \in \cdot \,\big|X_t, A_t\big\} = p(\cdot\,|X_t, A_t), t = 1, 2, \ldots$$

where $X_{[0,t]} = (X_0, \ldots, X_t)$ and $A_{[0,t]} = (A_0, \ldots, A_t)$. In this model, at each time step t, the decision maker observes the state of the system X_t and chooses an action A_t, using a decision function (control policy) π_t, depending on the observation obtained up to that time $\big(X_0, A_0, X_1, \ldots, A_{t-1}, X_{t-1}, X_t\big)$. The action can be a selection of a point from the action set, i.e., $\pi_t(X_0, A_0, X_1, \ldots, A_{t-1}, X_{t-1}, X_t) = A_t$ (deterministic policy), or putting a probability distribution over an action set, i.e., $\pi_t(X_0, A_0, X_1, \ldots, A_{t-1}, X_{t-1}, X_t) = \mathsf{Pr}\big\{A_t \in \cdot\big\}$ (randomized policy). The effect of choosing an action at t is twofold: an immediate cost $c_t(X_t, A_t)$ is incurred and the state of the system evolves to a new state probabilistically according to the transition probability, i.e., $p(\cdot\,|X_t, A_t) = \mathsf{Pr}\big\{X_{t+1} \in \cdot\,|X_t, A_t\big\}$. The main objective is to choose an optimal control policy which minimizes the cost which usually is taken as one of the functions below. Here the expectations are taken with respect to the probability measure on the sequence space of states and control actions induced by the initial distribution μ and control policy $\pi = \{\pi_t\}$.

(i) Discounted Cost: $\mathbb{E}\Big[\sum_{t=0}^{\infty} \beta^t c(X_t, A_t)\Big]$ for some $\beta \in (0, 1)$.

(ii) Average Cost: $\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\Big[\sum_{t=0}^{T} c(X_t, A_t)\Big]$.

If we write $w(\pi, x)$ to denote the cost function (either (i) or (ii)) of policy π for initial point x ($\mu = \delta_x$ point mass at x), the optimal cost, called value function, of the control problem is defined as

$$w^*(x) \coloneqq \inf_{\pi} w(\pi, x).$$

A policy π^* is called optimal if

$$w(\pi^*, x) = w^*(x) \text{ for all } x \in \mathsf{X}.$$

In the theory of MDPs, a large body of research is devoted to studying the structural properties of optimal policies for various cost criteria. In particular, researchers have investigated conditions under which the optimal policy deterministically depends only on the current state. Such deterministic policies are usually called deterministic stationary policies in the literature [54] and are induced by measurable functions from the state space to the action space. The significance of deterministic stationary policies comes from the fact that it is the smallest structured set of control policies in which one can find globally optimal policy for a large class of infinite horizon discounted cost (see., e.g., [53, 54]) or average cost optimal control problems (see, e.g., [18, 66, 69]).

Although we have good characterizations for the existence of the optimal stationary policies, computing such policies is in general computationally infeasible for large (e.g., infinite) state and action spaces. Furthermore, in networked control, the transmission of such control actions to an actuator is not realistic when there is an information transmission constraint (physically limited by the presence of a communication channel) between a plant, a controller or an actuator; that is, actions of the stationary policy must be quantized in order to be reliably transmitted to an actuator.

Hence, there is a practical need to approximate optimal stationary policies by policies which are computable and transmittable. In the literature, various methods have been developed to tackle the approximation problem: approximate dynamic programming, approximate value or policy iteration, simulation based techniques, neuro-dynamic programming (or reinforcement learning), state aggregation, etc. We refer the reader to [37, 22, 12, 77, 72, 80, 100, 10, 33, 34, 35] and references therein for a rather complete survey of these techniques. A brief review of these methods will be given in Chapter 5. It is important to note that most of these works are for MDPs with discrete (i.e., finite or countable) state and action spaces (see, e.g., [80, 12, 77, 72, 37, 100, 22]).

For MDPs with uncountable state spaces, the approximation problem had not been studied as extensively as in the finite or countable cases (see, e.g., [102, 62, 10]). However, in recent years, there has been an interest in the approximation problem for MDPs with general Borel state and actions spaces for both classical optimal stochastic control problem and also for problems with constraints [33, 34, 35]. For instance, in [35], the authors adopt a simulation based approximation technique leading to probabilistic guarantees on the approximation error. Specifically, [35] considered Borel state and action spaces with a possibly unbounded one-stage cost function and developed a sequence of approximations for the MDP using the empirical distributions generated by a probability measure ψ with respect to which the transition probability of the MDP is absolutely continuous. By imposing Lipschitz type continuity conditions on the components of the control model, [35] obtained a concentration inequality type upper bound on the accuracy of the approximation based on the Wasserstein distance of order 1 between the probability measure ψ and its empirical estimate.

1.4 Contributions and Organization of the Thesis

1.4.1 Chapter 2

In this chapter, a general representation of randomized quantizers that includes the common models in the literature is introduced via appropriate mixtures of joint probability measures on the product of the source and reproduction alphabets. Then, using this representation and results from optimal transport theory, we study fixed-rate randomized vector quantization under the constraint that the quantizer's output has a given fixed probability distribution. This setup is the generalization of the so-called distribution preserving quantization problem. The existence of an optimal (minimum distortion) randomized quantizer having a given output distribution is established under various conditions. For sources with densities and the mean square distortion measure, it is shown that this optimum can be attained by randomizing quantizers having convex codecells. In the last part of this chapter, we study a Shannon-theoretic version of the generalized distribution preserving quantization problem where a stationary and memoryless source is encoded subject to a distortion constraint and the additional requirement that the reproduction also be stationary and memoryless with a given distribution. We provide a single-letter expression for the optimum distortion in the limit of large block-lengths. The results of Chapter 2 have appeared in part in [82, 87].

1.4.2 Chapter 3

In this chapter we generalize the rate-distortion result derived in Chapter 2 by introducing constraints on the rate of common randomness, shared between the encoder and the decoder, as an additional design parameter. In this setup the encoder and decoder are assumed to have access to independent *rate limited* common randomness unlike in the rate-distortion problem considered in Chapter 2, where unlimited common randomness is available between the encoder and the decoder. In our main result, we completely characterize the set of achievable coding and common randomness rate pairs at any distortion level, thereby providing the optimal tradeoff between these two rate quantities. We also consider two variations of this problem where we investigate the effect of relaxing the strict output distribution constraint and the role of 'private randomness' used by the decoder on the rate region. Our results have strong connections with Cuff's recent work on distributed channel synthesis [28, 29]. In particular, our achievability proof combines a coupling argument with the approach developed by Cuff, where instead of explicitly constructing the encoder-decoder pair, a joint distribution is constructed from which a desired encoder-decoder pair is established. We show, however, that for our problem, the separated solution of first finding an optimal channel and then synthesizing this channel results in a suboptimal rate region. The results of Chapter 3 have appeared in part in [85, 84].

1.4.3 Chapter 4

In this chapter, we study the finite action approximation of stationary policies for a discrete-time Markov decision process with Borel state and action spaces under strong and weak continuity assumptions on the transition probability, respectively.

We introduce a new family of policies, called deterministic stationary quantizer policies, and show that such policies can approximate optimal deterministic stationary policies with arbitrary precision, thus demonstrating that one can search for near optimal policies within the class of quantized control policies. We also derive explicit bounds on the approximation error in terms of the rate of the approximating quantizers. Under the weak continuity of the transition probability, we apply these results to a fully observed reduction of a partially observed Markov decision processes (POMDPs), and show that one can obtain near optimal policies even when there is an information transmission constraint from the controller to the plant. The results of Chapter 4 have appeared in part in [86, 81, 83, 88].

1.4.4 Chapter 5

In this chapter, we study finite-state approximations of discrete time Markov decision processes with discounted and average costs and Borel state and action spaces. Under certain continuity properties of the one-stage cost function and the transition probability, we show that the stationary policies obtained from the finite model, which is constructed by quantizing the state space of the original system on a finite grid, can approximate the optimal stationary policy with arbitrary precision. For compact-state MDPs, we obtain explicit rate of convergence bounds quantifying how the approximation improves as the number of the grid points used to quantize the state space increases. Using information theoretic arguments, the order optimality of the obtained rates of convergence rates is also established for a large class of problems. The results of Chapter 5 have appeared in part in [90, 89].

1.5 Notation and Conventions

The following notation will be used throughout the thesis. For a metric space E, we let $\mathcal{B}(\mathsf{E})$ and $\mathcal{P}(\mathsf{E})$ denote the Borel σ-algebra on E and the set of probability measures on $(\mathsf{E}, \mathcal{B}(\mathsf{E}))$, respectively. It will be tacitly assumed that any metric space is equipped with its Borel σ-algebra and all probability measures on such spaces will be Borel measures [17, Definition 7.1.1]. For any $\nu \in \mathcal{P}(\mathsf{E})$ and measurable real function g on E, we define $\nu(g) \coloneqq \int g d\nu$. $\delta_e \in \mathcal{P}(\mathsf{E})$ denotes the point mass at e: $\delta_e(A) = 1$ if $e \in A$ and $\delta_e(A) = 0$ if $e \notin A$ for any Borel set $A \subset \mathsf{E}$. The product of metric spaces will be equipped with the product Borel σ-algebra. Unless otherwise specified, the term "measurable" will refer to Borel measurability. We always equip $\mathcal{P}(\mathsf{E})$ with the Borel σ-algebra $\mathcal{B}(\mathcal{P}(\mathsf{E}))$ generated by the topology of weak convergence [15]. We will sometimes use $\mathcal{M}(\mathsf{E})$ in place of $\mathcal{B}(\mathcal{P}(\mathsf{E}))$. If E is a Borel space (i.e., Borel subset of complete and separable metric space), then $\mathcal{P}(\mathsf{E})$ is metrizable with the Prokhorov metric which makes $\mathcal{P}(\mathsf{E})$ into a Borel space [73]. $B(\mathsf{E})$ denotes the set of all bounded measurable real functions on E and $C_b(\mathsf{E})$ denotes the set of all bounded continuous real valued functions on E. For any $u \in C_b(\mathsf{E})$ or $u \in B(\mathsf{E})$, let $\|u\| \coloneqq \sup_{e \in \mathsf{E}} |u(e)|$ which turns $C_b(\mathsf{E})$ and $B(\mathsf{E})$ into Banach spaces. A sequence $\{\mu_n\}$ of measures on E is said to converge setwise to a measure μ if $\mu_n(g) \to \mu(g)$ for all $g \in B(\mathsf{E})$, or equivalently, $\mu_n(D) \to \mu(D)$ for all $D \in \mathcal{B}(\mathsf{E})$. Given any measurable function $w : \mathsf{E} \to [1, \infty)$ and any real valued measurable function u on E, we define the w-norm of u as

$$\|u\|_w \coloneqq \sup_{e \in \mathsf{E}} \frac{|u(e)|}{w(e)}.$$

$B_w(\mathsf{E})$ denotes the set of all real valued measurable functions on E with finite w-norm and $C_w(\mathsf{E})$ denotes the set of all real valued continuous functions on E with finite w-norm. Analogously, for any finite signed measure ϑ on E, we define the w-norm of ϑ as

$$\|\vartheta\|_w \coloneqq \sup_{\|g\|_w \leq 1} \left| \int_{\mathsf{E}} g(e)\vartheta(de) \right|.$$

When $w = 1$, $\|\vartheta\|_w$ reduces to the total variation norm, in which case we use the notation $\|\vartheta\|_{TV}$ instead of $\|\vartheta\|_w$. Let $\mathsf{E}^n = \prod_{i=1}^n \mathsf{E}_i$ $(2 \leq n \leq \infty)$ be a finite or a infinite product space. By an abuse of notation, any function g on $\prod_{j=i_1}^{i_n} \mathsf{E}_j$, where $\{i_1, \ldots, i_n\} \subseteq \{1, \ldots, n\}$, is also treated as a function on E^n by identifying it with its natural extension to E^n. For finite n, the elements of E^n are $e^n = (e_1, \ldots, e_n)$, $e_i \in \mathsf{E}$, $i = 1, \ldots, n$. A similar convention also applies to a sequence of random variables which will be denoted by upper case letters. For any positive real number R, we define $[2^{nR}] \coloneqq \{1, \ldots, \lceil 2^{nR} \rceil\}$, where $\lceil 2^{nR} \rceil$ is the smallest integer greater than or equal to 2^{nR}. For any triple (X, Y, U) of random variables or vectors, the notation $X - U - Y$ means that they form a Markov chain in this order. For any random vector $U^n = (U_1, \ldots, U_n)$, the random measure p_{U^n} denotes the empirical distribution of U^n, given by

$$p_{U^n}(\,\cdot\,) \coloneqq \frac{1}{n} \sum_{i=1}^n \delta_{U_i}(\,\cdot\,).$$

The notation $V \sim \nu$ means that random variable V has distribution ν. For any probability distribution ν on E, ν^n denotes the n-fold product distribution $\underbrace{\nu \otimes \cdots \otimes \nu}_{n\text{-times}}$ on E^n.

Part I

Randomized Quantization with Constraints

Chapter 2

Output Constrained Randomized Quantization

2.1 Introduction

Recall that a quantizer maps a source (input) alphabet into a finite collection of points (output levels) from a reproduction alphabet. A quantizer's performance is usually characterized by its rate, defined as the logarithm of the number of output levels, and its expected distortion when the input is a random variable. One usually talks about randomized quantization when the quantizer used to encode the input signal is randomly selected from a given collection of quantizers. Although introducing randomization in the quantization procedure does not improve the optimal rate-distortion tradeoff, randomized quantizers may have other advantages over their deterministic counterparts.

In what appears to be the first work explicitly dealing with randomized quantization, Roberts [79] found that adding random noise to an image before quantization and subtracting the noise before reconstruction may result in a perceptually more pleasing image. Schuchman [92] and Gray and Stockham [48] analyzed versions of such so called *dithered* scalar quantizers where random noise (dither) is added to the input signal prior to uniform quantization. If the dither is subtracted after the quantization operation, the procedure is called subtractive dithering; otherwise it is called non-subtractive dithering. Under certain conditions, dithering results in uniformly distributed quantization noise that is independent of the input [92, 48], which allows for a simple modeling of the quantization process by an additive noise channel. In the information theoretic literature the properties of entropy coded dithered lattice quantizers have been extensively studied. For example, such quantizers have been used to provide achievable bounds on the performance of universal lossy compression systems by Ziv [114] and Zamir and Feder [111, 112]. Recently Akyol and Rose [3], [4], introduced a class of randomized *nonuniform* scalar quantizers obtained via applying companding to a dithered uniform quantizer and investigated optimality conditions for the design of such quantizers. Random codes used to prove the achievability part of Shannon's rate-distortion theorem [94] can also be considered as randomized quantizers. One should also note that an analogous randomized code definition is also made in the channel coding literature where the encoder and the decoder are selected randomly from the family of deterministic encoder-decoder pairs [63, 27]. These type of codes are used to improve the performance when there is channel uncertainty.

Dithered uniform/lattice and companding quantizers, as well as random rate-distortion codes, pick a random quantizer from a "small" structured subset of all possible quantizers. Such special randomized quantizers may be suboptimal for certain tasks and one would like to be able to work with more general (or completely general) classes of randomized quantizers. For example, Li *et al.* [64] considered *distribution-preserving* dithered scalar quantization, where the quantizer output is restricted to have the same distribution as the source, to improve the perceptual quality of mean square optimal quantizers in audio and video coding. Dithered quantizers or other structured randomized quantizers classes likely do not provide optimal performance in this problem. In an unpublished work [65] the same authors considered more general distribution-preserving randomized vector quantizers and lower bounded the minimum distortion achievable by such schemes when the source is stationary and memoryless.

In this chapter we propose a general model which formalizes the notion of randomly picking a quantizer from the set of *all* quantizers with a given number of output levels. Note that this set is much more complex and less structured then, for example, the *parametric* family of all quantizers having a given number of convex codecells. Inspired by work in stochastic control (e.g., [21]) our model represents the set of all quantizers acting on a given source as a subset of all joint probability measures on the product of the source and reproduction alphabets. Then a randomized quantizer corresponds to a mixture of probability measures in this subset. The usefulness of the model is demonstrated by rigorously setting up a generalization of the distribution-preserving quantization problem where the goal is to find a randomized quantizer minimizing the distortion under the constraint that the output has a given distribution (not necessarily that of the source). We show that under quite general conditions an optimal solution (i.e., an optimal randomized quantizer) exists for this generalized problem. We also consider a relaxed version of the output distribution constraint where the output distribution is only required to belong to some neighborhood (in the weak topology) of a target distribution. For this problem we show the optimality of randomizing among finitely many quantizers. While for a fixed quantizer dimension we can only provide existence results, for stationary and memoryless source and output distributions we also develop a rate-distortion theorem which identifies the minimum distortion in the limit of large block lengths in terms of the so-called output-constrained distortion-rate function. This last result solves a general version of a problem that was left open in [65].

2.2 Models of Randomized Quantization

In this chapter X denotes the input alphabet and Y is the reconstruction (output) alphabet. Throughout we set $\mathsf{X} = \mathsf{Y} = \mathbb{R}^n$, the n-dimensional Euclidean space for some $n \geq 1$, although most of the results hold in more general settings; for example if the input and output alphabets are Borel spaces.

In what follows we define three models of randomized quantization; two that are commonly used in the source coding literature, and a third abstract model that will nevertheless prove very useful.

Model 1

One general model of M-level randomized quantization that is often used in the information theoretic literature is depicted in Fig. 2.1.

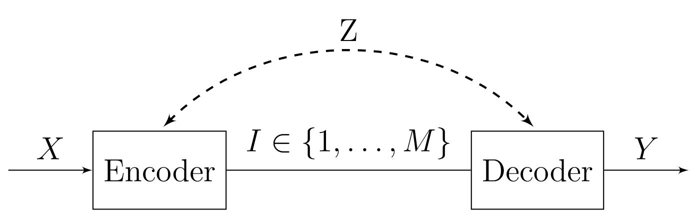

Figure 2.1: Randomized source code (quantizer).

Here X and Y are the source and output random variables taking values in X and Y, respectively. The index I takes values in $\{1, \ldots, M\}$, and Z is a $\mathsf{Z} = \mathbb{R}^m$-valued random variable which is independent of X and which is assumed to be available at both the encoder and the decoder. The encoder is a measurable function $e : \mathsf{X} \times \mathsf{Z} \to \{1, \ldots, M\}$ which maps (X, Z) to I, and the decoder is a measurable function $d : \{1, \ldots, M\} \times \mathsf{Z} \to \mathsf{Y}$ which maps (I, Z) to Y. For a given source distribution, in a probabilistic sense a Model 1 quantizer is determined by the triple (e, d, ν), where ν denotes the distribution of Z.

Note that codes used in the random coding proof of the forward part of Shannon's rate distortion theorem can be realized as Model 1 quantizers. In this case Z may be taken to be the random codebook consisting of $M = 2^{nR}$ code vectors of dimension n, each drawn independently from a given distribution. This Z can be represented as an $m = nM$-dimensional random vector that is independent of X. The encoder outputs the index I of the code vector Y in the codebook that best matches X (in distortion or in a joint-typicality sense) and the decoder can reconstruct this Y since it is a function of I and Z.

Remark 2.1. Although codes used in the random coding proof of the achievability part of Shannon's source coding theorem can be realized as Model 1 quantizers, the common randomness Z, in this case, is indeed not a part of the system design since it is used only to prove the existence of a (single) deterministic code that achieves the requirements. However, as will be seen in the sequel, the common randomness Z will be an important design parameter in problems that we are interested in this thesis. Therefore, it might be useful to name Model 1 randomized quantizers, for which common randomness is only used for establishing the existence of a certain deterministic code, as *random* codes rather than *randomized* codes. We refer reader to [63, p. 2151] for a discussion on distinction between random codes and randomized codes.

Model 2

Model 1 can be collapsed into a more tractable equivalent model. In this model, a randomized quantizer is a pair (q, ν), where $\mathsf{q} : \mathsf{X} \times \mathsf{Z} \to \mathsf{Y}$ is a measurable mapping such that $\mathsf{q}(\,\cdot\,, z)$ is an M-level quantizer for all $z \in \mathsf{Z}$ and ν is a probability measure on Z, the distribution of the randomizing random variable Z. Here q is the composition of the encoder and the decoder in Model 1: $\mathsf{q}(x, z) = d(e(x, z), z)$.

Model 2 quantizers include, as special cases, subtractive and non-subtractive dithering of M-level uniform quantizers, as well as the dithering of non-uniform quantizers. For example, if $n = m = 1$ and q_u denotes a uniform quantizer, then

$$\mathsf{q}(x, z) = q_u(x + z) - z$$

is a dithered uniform quantizer using subtractive dithering,

$$\mathsf{q}(x, z) = q_u(x + z)$$

is a dithered uniform quantizer with non-subtractive dithering, and with appropriate mappings g and h,

$$\mathsf{q}(x, z) = h\big(q_u(g(x) + z) - z\big).$$

is a dithered non-uniform quantizer (e.g., [64] and [4]). We note that dithered lattice quantizers [114, 111, 110] can also be considered as Model 2 type randomized quantizers when the source has a bounded support (so that with probability one only finitely many lattice points can occur as output points).

Let $\rho : \mathsf{X} \times \mathsf{Y} \to \mathbb{R}$ be a nonnegative measurable function, called the distortion measure. From now on we assume that the source X has distribution μ (denoted as $X \sim \mu$). The distortion associated with Model 2 quantizer (q, ν) or with Model 1 quantizer (e, d, ν), with $\mathsf{q}(x, z) = d(e(x, z), z)$, is the expectation

$$\begin{aligned} L(\mathsf{q}, \nu) &= \int_{\mathsf{Z}} \int_{\mathsf{X}} \rho(x, \mathsf{q}(x, z)) \, \mu(dx) \nu(dz) \\ &= \mathbb{E}\big[\rho(X, \mathsf{q}(X, Z))\big] \end{aligned} \tag{2.1}$$

where $Z \sim \nu$ is independent of X.

Model 3

In this model, instead of considering quantizers as functions that map X into a finite subset of Y, first we represent them as special probability measures on $\mathsf{X} \times \mathsf{Y}$ (see, e.g, [19],[108],[61],[45]). This leads to an alternative representation where a randomized quantizer is identified with a mixture of probability measures. In certain situations the space of these "mixing probabilities" representing *all* randomized quantizers will turn out to be more tractable than considering the quite unstructured space of all Model 1 triples (e, d, ν) or Model 2 pairs (q, ν).

Definition 2.1. *A* stochastic kernel *[54] (or* regular conditional probability *[32]) on* Y *given* X *is a function* $Q(dy|x)$ *such that for each* $x \in \mathsf{X}$, $Q(\,\cdot\,|x)$ *is a probability measure on* Y, *and for each Borel set* $B \subset \mathsf{Y}$, $Q(B|\,\cdot\,)$ *is a measurable function from* X *to* $[0, 1]$.

A quantizer q from X into Y can be represented as a stochastic kernel Q on Y

given X by letting [108], [19],

$$Q(dy|x) = \delta_{q(x)}(dy).$$

If we fix the distribution μ of the source X, we can also represent q by the probability measure $v(dx\,dy) = \mu(dx)\delta_{q(x)}(dy)$ on $\mathsf{X} \times \mathsf{Y}$. Thus we can identify the set $\mathcal{Q}_M$ of all M-level quantizers from X to Y with the following subset of $\mathcal{P}(\mathsf{X} \times \mathsf{Y})$:

$$\Gamma_\mu(M) = \big\{ v \in \mathcal{P}(\mathsf{X} \times \mathsf{Y}) : v(dx\,dy) = \mu(dx)\delta_{q(x)}(dy),\ q \in \mathcal{Q}_M \big\}. \tag{2.2}$$

Note that $q \mapsto \mu(dx)\delta_{q(x)}(dy)$ maps $\mathcal{Q}_M$ onto $\Gamma_\mu(M)$, but this mapping is one-to-one only if we consider equivalence classes of quantizers in $\mathcal{Q}_M$ that are equal μ almost everywhere (μ-a.e).

The following lemma is proved in the Section 2.7.1.

Lemma 2.1. *$\Gamma_\mu(M)$ is a Borel subset of $\mathcal{P}(\mathsf{X} \times \mathsf{Y})$.*

Now we are ready to introduce Model 3 for randomized quantization. Let P be a probability measure on $\mathcal{P}(\mathsf{X} \times \mathsf{Y})$ which is supported on $\Gamma_\mu(M)$, i.e., $P(\Gamma_\mu(M)) = 1$. Then P induces a "randomized quantizer" $v_P \in \mathcal{P}(\mathsf{X} \times \mathsf{Y})$ via

$$v_P(A \times B) = \int_{\Gamma_\mu(M)} v(A \times B)\, P(dv)$$

for Borel sets $A \subset \mathsf{X}$ and $B \subset \mathsf{Y}$, which we abbreviate to

$$v_P = \int_{\Gamma_\mu(M)} v\, P(dv). \tag{2.3}$$

Since each v in $\Gamma_\mu(M)$ corresponds to a quantizer with input distribution μ, P can be thought as a probability measure on the set of all M-level quantizers $\mathcal{Q}_M$.

Let $\mathcal{P}_0(\Gamma_\mu(M))$ denote the set of probability measures on $\mathcal{P}(\mathsf{X} \times \mathsf{Y})$ supported on $\Gamma_\mu(M)$. We define the set of M-level Model 3 randomized quantizers as

$$\Gamma_\mu^{\mathrm{R}}(M) = \Big\{ v_P \in \mathcal{P}(\mathsf{X} \times \mathsf{Y}) : v_P = \int_{\Gamma_\mu(M)} v P(dv),\ P \in \mathcal{P}_0(\Gamma_\mu(M)) \Big\}. \tag{2.4}$$

Note that if $v_P \in \Gamma_\mu^{\mathrm{R}}(M)$ is a Model 3 quantizer, then the X-marginal of v_P is equal to μ, and if X and Y are random vectors (defined on the same probability space) with joint distribution v_P, then they provide a stochastic representation of the random quantizer's input and output, respectively. Furthermore, the distortion associated with v_P is

$$\begin{aligned} L(v_P) &:= \int_{\mathsf{X}\times\mathsf{Y}} \rho(x,y) v_P(dx\,dy) \\ &= \int_{\Gamma_\mu(M)} \int_{\mathsf{X}\times\mathsf{Y}} \rho(x,y) v(dx\,dy) P(dv) \\ &= \mathbb{E}\big[\rho(X,Y)\big]. \end{aligned}$$

2.2.1 Equivalence of models

Here we show that the three models of randomized quantization are essentially equivalent. As before, we assume that the source distribution μ is fixed. The following two results are proved in Section 2.7.2 and Section 2.7.3, respectively.

Theorem 2.1. *For each Model 2 randomized quantizer (q, ν) there exists a Model 3 randomized quantizer $v_P \in \Gamma_\mu^{\mathrm{R}}(M)$ such that $(X,Y) = (X, \mathsf{q}(X,Z))$ has distribution v_P. Conversely, for any $v_P \in \Gamma_\mu^{\mathrm{R}}(M)$ there exists a Model 2 randomized quantizer*

(q, ν) *such that* $(X, \mathsf{q}(X, Z)) \sim \upsilon_P$.

Theorem 2.2. *Models 1 and 2 of randomized quantization are equivalent in the sense of Theorem 2.1.*

Remark 2.2.

(a) Clearly, any two equivalent randomized quantizers have the same distortion. The main result of this section is Theorem 2.1. Theorem 2.2 is intuitively obvious, but proving that any Model 2 quantizer can be decomposed into an equivalent Model 1 quantizer with measurable encoder and decoder is not quite trivial.

(b) Since the dimension m of the randomizing random vector Z was arbitrary, we can take $m = 1$ in Theorem 2.1. In fact, the proof also implies that any Model 2 or 3 randomized quantizer is equivalent (in the sense of Theorem 2.1) to a Model 2 quantizer (q, ν), where $\mathsf{q} : \mathsf{X} \times [0, 1] \to \mathsf{Y}$ and ν is the uniform distribution on $[0, 1]$.

(c) Assume that $(\mathsf{Z}, \mathcal{A}, \nu)$ is an *arbitrary* probability space. For any randomized quantizer $\mathsf{q} : \mathsf{X} \times \mathsf{Z} \to \mathsf{Y}$ in the form $\mathsf{q}(X, Z)$, where $Z \sim \nu$ is independent of X, there exists a Model 3 randomized quantizer υ_P such that $(X, \mathsf{q}(X, Z)) \sim \upsilon_P$. This can be proved by using the same proof method as in Theorem 2.1. In view of the previous remark and Theorem 2.1, this means that uniform randomization over the unit interval $[0, 1]$ suffices under the most general circumstances.

(d) All results in this section remain valid if the input and reproduction alphabets X and Y are arbitrary uncountable Borel spaces. In this case, uniform randomization over the unit interval still provides the most general model possible.

In the next two sections, Model 3 will be used to represent randomized quantizers because it is particularly suited to treating the optimal randomized quantization problem under an output distribution constraint.

2.3 Optimal Randomized Quantization with Fixed Output Distribution

Let ψ be a probability measure on Y and let $\Lambda(M,\psi)$ denote the set of all M-level Model 2 randomized quantizers (q,ν) such that the output $\mathsf{q}(X,Z)$ has distribution ψ. As before, we assume that $X \sim \mu$, $Z \sim \nu$, and Z and X are independent. We want to show the existence of a minimum-distortion randomized quantizer having output distribution ψ, i.e, the existence of $(\mathsf{q}^*,\nu^*) \in \Lambda(M,\psi)$ such that

$$L(\mathsf{q}^*,\nu^*) = \inf_{(\mathsf{q},\nu)\in\Lambda(M,\psi)} L(\mathsf{q},\nu).$$

If we set $\psi = \mu$, the above problem is reduced to showing the existence of a distribution-preserving randomized quantizer [64, 65] having minimum distortion.

The set of M-level randomized quantizers is a fairly general (nonparametric) set of functions and it seems difficult to investigate the existence of an optimum directly. On the other hand, Model 3 provides a tractable framework for establishing the existence of an optimal randomized quantizer under quite general conditions.

Let $\Gamma_{\mu,\psi}$ be the set of all joint distributions $v \in P(\mathsf{X}\times\mathsf{Y})$ having X-marginal μ and Y-marginal ψ. Then

$$\Gamma^{\mathrm{R}}_{\mu,\psi}(M) = \Gamma^{\mathrm{R}}_{\mu}(M) \cap \Gamma_{\mu,\psi} \tag{2.5}$$

is the subset of Model 3 randomized quantizers which corresponds to the class of output-distribution-constrained Model 2 randomized quantizers $\Lambda(M,\psi)$.

For any $v \in \mathcal{P}(\mathsf{X} \times \mathsf{Y})$ let

$$L(v) = \int_{\mathsf{X}\times\mathsf{Y}} \rho(x,y) v(dx\,dy).$$

Using these definitions, finding optimal randomized quantizers with a given output distribution can be posed as finding v in $\Gamma^{\mathrm{R}}_{\mu,\psi}(M)$ which minimizes $L(v)$, i.e.,

$$\begin{aligned} \textbf{(P1)} \text{ minimize } & L(v) \\ \text{subject to } & v \in \Gamma^{\mathrm{R}}_{\mu,\psi}(M). \end{aligned}$$

We can prove the existence of the minimizer for **(P1)** under either of the following assumptions. Here $\|x\|$ denotes the Euclidean norm of $x \in \mathbb{R}^n$.

Assumption 2.1. $\rho(x,y)$ is continuous and $\psi(B) = 1$ for some compact subset B of Y.

Assumption 2.2. $\rho(x,y) = \|x-y\|^2$.

Theorem 2.3. *Suppose* $\inf_{v\in\Gamma^{\mathrm{R}}_{\mu,\psi}(M)} L(v) < \infty$. *Then there exists a minimizer with finite cost for problem* **(P1)** *under either Assumption 2.1 or Assumption 2.2.*

The theorem is proved in Section 2.7.4 with the aid of optimal transport theory [99]. The optimal transport problem for marginals $\pi \in \mathcal{P}(\mathsf{X})$, $\lambda \in \mathcal{P}(\mathsf{Y})$ and cost function $c : \mathsf{X} \times \mathsf{Y} \to [0,\infty]$ is defined as

$$\begin{aligned} \text{minimize } & \int_{\mathsf{X}\times\mathsf{Y}} c(x,y) v(dx\,dy) \\ \text{subject to } & v \in \Gamma_{\pi,\lambda}. \end{aligned}$$

In the proof of Theorem 2.3 we set up a relaxed version of the optimization problem

(**P1**). We show that if the relaxed problem has a minimizer, then (**P1**) also has a minimizer, and then prove the existence of a minimizer for the relaxed problem using results from optimal transport theory.

Remark 2.3. Note that the product distribution $\mu \otimes \psi$ corresponds to a 1-level randomized quantizer (the equivalent Model 2 randomized quantizer is given by $\mathsf{q}(x, z) = z$ and $Z \sim \psi$). Hence $\mu \otimes \psi \in \Gamma^{\mathrm{R}}_{\mu,\psi}(M)$ for all $M \geq 1$, and if $L(\mu \otimes \psi) < \infty$, then the condition $\inf_{v \in \Gamma^{\mathrm{R}}_{\mu,\psi}(M)} L(v) < \infty$ holds. In particular, if both μ and ψ have finite second moments $\int \|x\|^2 \mu(dx) < \infty$ and $\int \|y\|^2 \psi(dy) < \infty$, and $\rho(x, y) = \|x - y\|^2$ (Assumption 2.2), then $\inf_{v \in \Gamma^{\mathrm{R}}_{\mu,\psi}(M)} L(v) < \infty$.

Optimal transport theory can also be used to show that, under some regularity conditions on the input distribution and the distortion measure, the randomization can be restricted to quantizers having a certain structure. Here we consider sources with densities and the mean square distortion. A quantizer $q : \mathsf{X} \to \mathsf{Y}$ with output points $q(\mathsf{X}) = \{y_1, \ldots, y_k\} \subset \mathsf{Y}$ is said to have *convex codecells* if $q^{-1}(y_i) = \{x : q(x) = y_i\}$ is a convex subset of $\mathsf{X} = \mathbb{R}^n$ for all $i = 1, \ldots, k$. Let $\mathcal{Q}_{M,\mathrm{c}}$ denote the set of all M-level quantizers having convex codecells. The proof of the following theorem is given in Section 2.7.5.

Theorem 2.4. *Suppose $\rho(x, y) = \|x - y\|^2$ and μ admits a probability density function. Then an optimal randomized quantizer in Theorem 2.3 can be obtained by randomizing over quantizers with convex cells. That is*

$$\min_{v \in \Gamma^{\mathrm{R}}_{\mu,\psi}(M)} L(v) = \min_{v \in \Gamma^{\mathrm{R,c}}_{\mu,\psi}(M)} L(v),$$

where $\Gamma^{\mathrm{R,c}}_{\mu,\psi}(M)$ represents the Model 3 quantizers with output distribution ψ that are *obtained by replacing $\mathcal{Q}_M$ with $\mathcal{Q}_{M,\mathrm{c}}$ in* (2.2)*.*

Remark 2.4. Each quantizer having M convex codecells can be described using $nM + (n+1)M(M-1)/2$ real parameters if μ has a density and any two quantizers that are μ-a.e. equal are considered equivalent. One obtains such a parametric description by specifying the M output points using nM real parameters, and specifying the M convex polytopal codecells by $M(M-1)/2$ hyperplanes separating pairs of distinct codecells using $(n+1)M(M-1)/2$ real parameters. Thus Theorem 2.4 replaces the *nonparametric* family of quantizers Q_M in Theorem 2.3 with the *parametric* family $Q_{M,\mathrm{c}}$.

2.4 Approximation with Finite Randomization

Since randomized quantizers require common randomness that must be shared between the encoder and the decoder, it is of interest to see how one can approximate the optimal cost by randomizing over finitely many quantizers. Clearly, if the target probability measure ψ on Y is not finitely supported, then no finite randomization exists with this output distribution. In this section we relax the fixed output distribution constraint and consider the problem where the output distribution belongs to some neighborhood (in the weak topology) of ψ. We show that one can always find a finitely randomized quantizer which is optimal (resp., ε-optimal) for this relaxed problem if the distortion measure is continuous and bounded (resp., arbitrary).

Let $B(\psi, \delta)$ denote the open ball in $\mathcal{P}(\mathsf{Y})$, with respect to the Prokhorov metric [15] (see also (2.22) in Section 2.7.6), having radius $\delta > 0$ and centered at the target input distribution ψ. Also, let $\mathcal{M}^{\delta}_{\mu,\psi}$ denote the set of all $v \in \Gamma^{\mathrm{R}}_{\mu}(M)$ whose Y marginal belongs to $B(\psi, \delta)$. That is, $\mathcal{M}^{\delta}_{\mu,\psi}$ represents all randomized quantizers in

$\Gamma_\mu^{\mathrm{R}}(M)$ whose output distribution is within distance δ of the target distribution ψ. We consider the following relaxed version of the minimization problem **(P1)**:

$$\textbf{(P3)}\ \text{minimize } L(v)$$
$$\text{subject to } v \in \mathcal{M}_{\mu,\psi}^{\delta}.$$

The set of *finitely randomized* quantizers in $\Gamma_\mu^{\mathrm{R}}(M)$ is obtained by taking finite mixtures of quantizers in $\Gamma_\mu(M)$, i.e.,

$$\Gamma_\mu^{\mathrm{FR}}(M) = \Big\{ v_P \in \Gamma_\mu^{\mathrm{R}}(M) : v_P = \int_{\Gamma_\mu(M)} v P(dv),\ |\operatorname{supp}(P)| < \infty \Big\}.$$

Theorem 2.5. *Assume the distortion measure ρ is continuous and bounded and let $v \in \mathcal{M}_{\mu,\psi}^{\delta}$ be arbitrary. Then there exists v_F in $\mathcal{M}_{\mu,\psi}^{\delta} \cap \Gamma_\mu^{\mathrm{FR}}(M)$ such that $L(v_F) \leq L(v)$.*

The proof is given in Section 2.7.6.

Although the minimum in **(P3)** may not be achieved by any $v \in \mathcal{M}_{\mu,\psi}^{\delta}$, the theorem implies that if the problem has a solution, it also has a solution in the set of finitely randomized quantizers.

Corollary 2.1. *Assume ρ is continuous and bounded and suppose there exists $v^* \in \mathcal{M}_{\mu,\psi}^{\delta}$ with $L(v^*) = \inf_{v \in \mathcal{M}_{\mu,\psi}^{\delta}} L(v)$. Then there exists $v_F \in \mathcal{M}_{\mu,\psi}^{\delta} \cap \Gamma_\mu^{\mathrm{FR}}(M)$ such that $L(v_F) = L(v^*)$.*

The continuity of L, implied by the boundedness and continuity of ρ is crucial in the proof of Theorem 2.5 and thus for Corollary 2.1. However, the next theorem shows that for an arbitrary ρ, any $\varepsilon > 0$, and $v \in \mathcal{M}_{\mu,\psi}^{\delta}$, there exists v_F in $\mathcal{M}_{\mu,\psi}^{\delta} \cap \Gamma_\mu^{\mathrm{FR}}(M)$

such that $L(v_F) \leq L(v) + \varepsilon$. That is, for any $\varepsilon > 0$ there exists an ε-optimal finitely randomized quantizer for **(P3)**. The theorem is proved in Section 2.7.7

Theorem 2.6. *Let ρ be an arbitrary distortion measure and assume* $\inf_{v \in \mathcal{M}^{\delta}_{\mu,\psi}} L(v) < \infty$. *Then,*

$$\inf_{v \in \mathcal{M}^{\delta}_{\mu,\psi} \cap \Gamma^{\mathrm{FR}}_{\mu}(M)} L(v) = \inf_{v \in \mathcal{M}^{\delta}_{\mu,\psi}} L(v).$$

Remark 2.5. The above results on finite randomization heavily depend on our use of the Prokhorov metric as a measure of "distance" between two probability measures. In particular, if one considers other measures of closeness, such as the Kullback-Leibler (KL) divergence or the total variation distance, then finite randomization may not suffice if the target output distribution is not discrete. In particular, if the target output distribution ψ has a density and $\tilde{\psi}$ denotes the (necessarily discrete) output distribution of any finitely randomized quantizer, then $\tilde{\psi}$ is not absolutely continuous with respect to ψ and for the KL divergence we have $D_{KL}(\tilde{\psi}\|\psi) = \infty$, while for the total variation distance we have $\|\tilde{\psi} - \psi\|_{TV} = 1$.

2.5 A Source Coding Theorem

After proving the existence of an optimum randomized quantizer in problem **(P1)** in Section 2.3, one would also like to evaluate the minimum distortion

$$L^* \coloneqq \min\{L(v) : v \in \Gamma^{\mathrm{R}}_{\mu,\psi}(M)\} \tag{2.6}$$

achievable for fixed source and output distributions μ and ψ and given number of quantization levels M. For any given blocklength n this seems to be a very hard problem in general. However, we are able to prove a rate-distortion type result that explicitly identifies L^* in the limit of large block lengths n if the source and output distributions correspond to two stationary and memoryless (i.e., i.i.d.) processes.

With a slight abuse of the notation used in previous sections, we let $\mathsf{X} = \mathsf{Y}$ and consider a sequence of problems **(P1)** with input and output alphabets $\mathsf{X}^n = \mathsf{Y}^n$, $n \geq 1$, and corresponding source and output distributions $\mu^n = \mu \otimes \cdots \otimes \mu$ and $\psi^n = \psi \otimes \cdots \otimes \psi$.

Assumption 2.3. We assume that $\mathsf{X} = \mathsf{Y}$ is a finite set or $\mathsf{X} = \mathsf{Y} = \mathbb{R}$. The distortion measure is given by $\rho(x, y) = d(x, y)^p$, where d is the metric on X. Here, $p > 0$ when X is finite and $p = 2$ when $\mathsf{X} = \mathbb{R}$, in which case we also assume that $d(x, y) = |x-y|$ (so that ρ is the squared error) and that the source distribution μ and the desired output distribution ψ have finite second moments, i.e., $\int x^2 \mu(dx) < \infty$, $\int y^2 \psi(dy) < \infty$.

For $R \geq 0$ let $\Gamma^{\mathrm{R}}_{\mu^n,\psi^n}(2^{nR})$ denote the set of n-dimensional Model 3 randomized quantizers defined in (2.5) having input distribution μ^n, output distribution ψ^n, and at most 2^{nR} levels (i.e., rate R). Then

$$L_n(\mu, \psi, R) \coloneqq \inf\bigl\{L(v) : v \in \Gamma^{\mathrm{R}}_{\mu^n,\psi^n}(2^{nR})\bigr\}$$

is the minimum distortion achievable by such quantizers.

We also define

$$D(\mu, \psi, R) = \inf\bigl\{\mathbb{E}[\rho(X, Y)] : X \sim \mu,\ Y \sim \psi,\ I(X; Y) \leq R\bigr\}, \tag{2.7}$$

where the infimum is taken over pairs of all joint distributions of real random variables X and Y such that X has distribution μ, Y has distribution ψ, and their mutual information $I(X; Y)$ is upper bounded by R.

One can trivially adapt the standard proof from rate-distortion theory to show that similar to the distortion-rate function, $D(\mu, \psi, R)$ is a convex and nonincreasing function of R. Note that $D(\mu, \psi, R)$ is finite for all $R \geq 0$ by the assumption that μ and ψ have finite second moments. The distortion-rate function $D(\mu, R)$ of the i.i.d. source μ, is obtained from $D(\mu, \psi, R)$ as

$$D(\mu, R) = \inf_{\psi \in \mathcal{P}(\mathsf{Y})} D(\mu, \psi, R).$$

By a standard argument one can easily show that the sequence $\{nL_n(\mu, \psi, R)\}_{n \geq 1}$ is subadditive and so $\inf_{n \geq 1} L_n(\mu, \psi, R) = \lim_{n \to \infty} L_n(\mu, \psi, R)$. Thus the limit represents the minimum distortion achievable with rate-R randomized quantizers for an i.i.d. source with marginal μ under the constraint that the output is i.i.d. with marginal ψ. The next result proves that this limit is equal to $D(\mu, \psi, R)$, which one could thus call the "output-constrained distortion-rate function."

Theorem 2.7. *We have*

$$\lim_{n \to \infty} L_n(\mu, \psi, R) = D(\mu, \psi, R). \tag{2.8}$$

Remark 2.6.

(a) As usual, the proof of the theorem consists of a converse and an achievability part. The converse (Lemma 2.2 below) directly follows from the usual proof of the converse part of the rate-distortion theorem. In fact, this was first noticed in [65] where the special case $\psi = \mu$ was considered and (in a different formulation) it was shown that for all n

$$L_n(\mu, \mu, R) \geq D(\mu, \mu, R).$$

Virtually the same argument implies that $L_n(\mu, \psi, R) \geq D(\mu, \psi, R)$ for all n and ψ. Nevertheless, we write out the proof in Section 2.7.8 since, strictly speaking, the proof in [65] is only valid if ψ is discrete with finite (Shannon) entropy or it has a density and finite differential entropy.

(b) The proof of the converse part (Lemma 2.2) is valid for any randomized quantizer whose output Y^n satisfies $Y_i \sim \psi$, $i = 1, \ldots, n$. Thus the theorem also holds if in the definition of $L_n(\mu, \psi, D)$, the randomized quantizers are required to have outputs with identically distributed (but not necessarily independent) components having common distribution ψ.

(c) In [65] it was left as an open problem if $D(\mu, \mu, R)$ can be asymptotically achieved by a sequence of distribution-preserving randomized quantizers. The authors presented an incomplete achievability proof for the special case of Gaussian μ using dithered lattice quantization. We prove the achievability of $D(\mu, \psi, R)$ for arbitrary μ and ψ using a fundamentally different (but essentially non-constructive) approach. In particular, our proof is based on random coding where the codewords are uniformly distributed on the type class of an n-type that well approximates the target distribution ψ, combined with optimal coupling from mass transport theory.

(d) With only minor changes in the proof, the theorem remains valid if $\mathsf{X} = \mathsf{Y}$ are arbitrary Polish spaces with metric d and $\rho(x, y) = d(x, y)^p$ for some $p \geq 1$. In this case the finite second moment conditions translate into $\int d(x,x_0)^p\,\mu(dx) < \infty$ and $\int d(y,y_0)^p\,\psi(dy) < \infty$ for some (and thus all) $x_0, y_0 \in \mathsf{X}$.

Proof of Theorem 2.7. In this proof we use Model 2 of randomized quantization which is more suitable here than Model 3. Also, it is easier to deal with the rate-distortion performance that with the distortion-rate performance. Thus, following the notation in [113], for $D \geq 0$ we define the *minimum mutual information with constraint output* ψ as

$$I_m(\mu\|\psi, D) = \inf\big\{I(X;Y) : X \sim \mu, Y \sim \psi, \mathbb{E}[\rho(X,Y)] \leq D\big\}, \tag{2.9}$$

where the infimum is taken over pairs of all joint distributions of X with marginal μ and Y with marginal ψ such that $\mathbb{E}[\rho(X,Y)] \leq D$. If this set of joint distributions is empty, we let $I_m(\mu\|\psi, D) = \infty$. Clearly, the extended real valued functions $I_m(\mu\|\psi, \,\cdot\,)$ and $D(R,\mu,\,\cdot\,)$ are inverses of each other. Hence $I_m(\mu\|\psi,\, D)$ is a nonincreasing, convex function of D. Analogous with $D(R,\mu,\,\cdot\,)$, for each $D \geq 0$, $I_m(\mu\|\psi,\, D)$ can be interpreted as minimum achievable coding rate given distortion level D. Therefore, any R satisfying $R \geq I(X;Y)$ with $X \sim \mu$, $Y \sim \psi$, and $\mathbb{E}[\rho(X,Y)] \leq D$ is achievable for D; that is, for any $\varepsilon > 0$ and all sufficiently large n, there exists a randomized source code $v \in \Gamma^R_{\mu^n,\psi^n}(2^{nR})$ such that $L(v) \leq D + \varepsilon$.

The converse part of the theorem, i.e., the statement $L_n(\mu,\psi,R) \geq D(R,\mu,\psi)$ for all $n \geq 1$, is directly implied by the following lemma. The proofs of all lemmas in this section are given in Section 2.7.8.

Lemma 2.2. *For all $n \geq 1$ if a randomized quantizer has input distribution μ^n,* *output distribution ψ^n, and distortion D, then its rate is lower bounded as*

$$R \geq I_m(\mu \| \psi, D).$$

In the rest of the proof we show the achievability of $D(R, \mu, \psi)$. We first prove this for finite alphabets and then generalize to continuous alphabets.

For each n let ψ_n be a closest n-type [25, Chapter 11] to ψ in the l_1-distance which is absolutely continuous with respect to ψ, i.e., $\psi_n(y) = 0$ whenever $\psi(y) = 0$. Let D be such that $I_m(\mu \| \psi, D) < \infty$, let $\varepsilon > 0$ be arbitrary, and set $R = I_m(\mu \| \psi, D) + \varepsilon$. Assume $X^n \sim \mu^n$ for $n \geq 1$. For each n generate 2^{nR} codewords uniformly and independently drawn from $T_n(\psi_n)$, the *type class* of ψ_n [25], i.e., independently (of each other and of X^n) generate random codewords $U^n(1), \ldots, U^n(2^{nR})$ such that $U^n(i) \sim \psi_n^{(n)}$, where

$$\psi_n^{(n)}(y^n) = \begin{cases} \frac{1}{|T_n(\psi_n)|}, & \text{if } y^n \in T_n(\psi_n) \\ 0, & \text{otherwise.} \end{cases}$$

(As usual, for simplicity we assume that 2^{nR} is an integer.) Let $\hat{X}^n$ denote the output of the nearest neighborhood encoder: $\hat{X}^n = \arg\min_{1 \leq i \leq 2^{nR}} \rho_n(X^n, U^n(i))$. In case of ties, we choose $U^n(i)$ with the smallest index i. The next lemma states the intuitively clear fact that $\hat{X}^n$ is uniformly distributed on $T_n(\psi_n)$.

Lemma 2.3. $\hat{X}^n \sim \psi_n^{(n)}$.

The idea for this random coding scheme comes from [113] where an infinite i.i.d. codebook $\{U^n(i)\}_{i=1}^{\infty}$ was considered and the coding rate was defined as $(1/n) \log N_n$, where N_n is the smallest index i such that $\rho_n(X^n, U^n(i)) \leq D$. If the $U^n(i)$ are uniformly chosen from the type class $T_n(\psi_n)$, then by Theorem 1 and Appendix A and B of [113], $(1/n)\log N_n - I_m(\mu\|\psi_n, D) \to 0$ in probability.

Our scheme converts this variable-length random coding scheme into a fixed-rate scheme by considering, for each blocklength n, the finite codebook $\{U^n(i)\}_{i=1}^{2^{nR}}$. Letting $\rho_{\max} = \max_{x,y}\rho(x,y)$, the expected distortion of our scheme is bounded as

$$\mathbb{E}[\rho_n(X^n, \hat{X}^n)] \le D + \rho_{\max} \Pr\Big\{ \frac{1}{n}\log N_n > R \Big\}.$$

Since $I_m(\mu\|\psi_n, D) \to I_m(\mu\|\psi, D)$ by the continuity of $I_m(\mu\|\psi, D)$ in ψ (see [113, Appendix A]), we have $R \ge I_m(\mu\|\psi_n, D) + \delta$ for some $\delta > 0$ if n is large enough. Thus the above bound implies

$$\limsup_{n\to\infty} \mathbb{E}[\rho_n(X^n, \hat{X}^n)] \le D. \tag{2.10}$$

Hence our random coding scheme has the desired rate and distortion as $n \to \infty$. However, its output $\hat{X}^n$ has distribution $\psi_n^{(n)}$ instead of the required ψ^n. The next lemma shows that the normalized Kullback-Leibler divergence (relative entropy, [25]) between $\psi_n^{(n)}$ and ψ^n asymptotically vanishes.

Lemma 2.4. $\frac{1}{n}\mathcal{D}(\psi_n^{(n)}\|\psi^n) \to 0$ *as* $n \to \infty$.

Let $\pi, \lambda \in \mathcal{P}(\mathsf{X})$. The optimal transportation cost $\hat{T}_n(\pi, \lambda)$ between π and λ (see, e.g., [99]) with cost function ρ_n is defined by

$$\hat{T}_n(\pi, \lambda) = \inf\big\{\mathbb{E}[\rho_n(U^n, V^n)] : U^n \sim \pi,\ V^n \sim \lambda\big\}, \tag{2.11}$$

where the infimum is taken over all joint distribution of pairs of random vectors

(U^n, V^n) satisfying the given marginal distribution constraints. The joint distribution achieving $\hat{T}_n(\pi, \lambda)$ as well as the resulting pair (U^n, V^n) are both called an optimal coupling of π and λ. Optimal couplings exist when X is finite or $\mathsf{X} = \mathbb{R}^n$, $\rho(x, y) = (x - y)^2$, and both π and λ both have finite second moments [99].

Now consider an optimal coupling $(\hat{X}^n, Y^n)$ of $\psi_n^{(n)}$ and ψ^n. If Z_1 and Z_2 are uniform random variables on $[0, 1]$ such that $Z = (Z_1, Z_2)$ is independent of X^n, then the random code and optimal coupling can be "realized" as $(U^n(1), \ldots, U^n(2^{nR})) = f_n(Z_1)$, $\hat{X}^n = \hat{f}_n(X^n, Z_1)$, and $Y^n = g_n(\hat{X}^n, Z_2)$, where f_n, $\hat{f}_n$, and g_n are suitable (measurable) functions. Combining random coding with optimal coupling this way gives rise to a randomized quantizer of type Model 2 whose output has the desired distribution ψ^n (see Fig. 2.2).

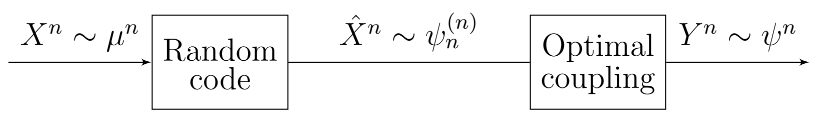

Figure 2.2: $D(R, \mu, \psi)$ achieving randomized quantizer scheme.

The next lemma uses Marton's inequality [67] to show that the extra distortion introduced by the coupling step asymptotically vanishes.

Lemma 2.5. *We have*

$$\lim_{n \to \infty} \hat{T}_n(\psi_n^{(n)}, \psi^n) = 0$$

and consequently

$$\limsup_{n \to \infty} \mathbb{E}\big[\rho_n(X^n, Y^n)\big] \le D.$$

In summary, we have shown that there exists a sequence of Model 2 randomized quantizers having rate $R = I_m(\mu \| \psi, D) + \varepsilon$ and asymptotic distortion upper bounded by D which satisfy the output distribution constraint $Y^n \sim \psi^n$. Since $\varepsilon > 0$

is arbitrary, this completes the proof of the achievability of $I_m(\mu\|\psi, D)$ (and the achievability of $D(\mu, \psi, R)$) for finite source and reproduction alphabets.

Remark 2.7. We note that an obvious approach to achievability would be to generate a codebook where the codewords have i.i.d. components drawn according to ψ. However, the output distribution of the resulting the scheme would be *too far* from the desired ψ^n. In particular, such a scheme produces output $\hat{X}^n$ whose empirical distribution (type) converges to a "favorite type" which is typically different from ψ [113, Theorem 4]. As well, the rate achievable with this scheme at distortion level D is [106, Theorem 2]

$$R = \min_{\psi' \in \mathcal{P}(\mathsf{Y})} \big(I_m(\mu\|\psi', D) + \mathcal{D}(\psi'\|\psi) \big)$$

which is typically strictly less than $I_m(\mu\|\psi, D)$.

Now let $\mathsf{X} = \mathsf{Y} = \mathbb{R}$, $\rho(x, y) = (x - y)^2$, and assume that μ and ψ have finite second moments. We make use of the final alphabet case to prove achievability for this continuous case. The following lemma provides the necessary link between the two cases.

Lemma 2.6. *There exist a sequence $\{A_k\}$ of finite subsets of $\mathbb{R}$ and sequences of probability measures $\{\mu_k\}$ and $\{\psi_k\}$, both supported on A_k, such that*

(i) $\hat{T}_1(\mu, \mu_k) \to 0$, $\hat{T}_1(\psi, \psi_k) \to 0$ *as* $k \to \infty$*;*

(ii) For any $\varepsilon > 0$ *and* $D \geq 0$ *such that* $I_m(\mu\|\psi, D) < \infty$*, we have* $I_m(\mu_k\|\psi_k, D + \varepsilon) \leq I_m(\mu\|\psi, D)$ *for all* k *large enough.*

Let μ_k^n and ψ_k^n denote the n-fold products of μ_k and ψ_k, respectively. Definition (2.11) of optimal coupling implies that $\hat{T}_n(\mu^n, \mu_k^n) \leq \hat{T}_1(\mu, \mu_k)$ and $\hat{T}_n(\psi^n, \psi_k^n) \leq$

$\hat{T}_1(\psi, \psi_k)$. Hence for any given $\varepsilon > 0$ by Lemma 2.6 we can choose k large enough such that for all n,

$$\hat{T}_n(\mu^n, \mu_k^n) \leq \varepsilon \text{ and } \hat{T}_n(\psi^n, \psi_k^n) \leq \varepsilon, \tag{2.12}$$

and also $I_m(\mu_k \| \psi_k, D + \varepsilon) \leq I_m(\mu \| \psi, D)$.

Now for each n define the following randomized quantizer:

(a) Realize the optimal coupling between μ^n and μ_k^n.

(b) Apply the randomized quantizer scheme for the finite alphabet case with common source and output alphabet A_k, source distribution μ_k^n, and output distribution ψ_k^n. Set the rate of the quantizer to $R = I_m(\mu \| \psi, D) + \varepsilon$.

(c) Realize the optimal coupling between ψ_k^n and ψ^n.

In particular, the optimal couplings are realized as follows: in (a) the source $X^n \sim \mu^n$ is mapped to $X^n(k) \sim \mu_k^n$, which serves as the source in (b), via $X^n(k) = \hat{f}_{n,k}(X^n, Z_3)$, and in (c) the output $Y^n(k) \sim \psi_k^n$ of the scheme in (b) is mapped to $Y^n \sim \psi^n$ via $Y^n = \hat{g}_{n,k}(Y^n(k), Z_4)$, where Z_3 and Z_4 are uniform randomization variables that are independent of X^n. Thus the composition of these three steps is a valid Model 2 randomized quantizer.

Since $R = I_m(\mu \| \psi, D) + \varepsilon$, in step (b) the asymptotic (in n) distortion $D + \varepsilon$ can be achieved by Lemma 2.6(ii). Using (2.12) and the triangle inequality for the norm $\|V^n\|_2 := \left(\sum_{i=1}^n E[V_i^2]\right)^{1/2}$ on $\mathbb{R}^n$-valued random vectors having finite second moments, it is straightforward to show that the asymptotic distortion of the overall scheme is upper bounded by $D + l(\varepsilon)$, where $l(\varepsilon) \to 0$ as $\varepsilon \to 0$. Since $\varepsilon > 0$

can be taken to be arbitrarily small by choosing k large enough, this completes the achievability proof for the case $\mathsf{X} = \mathsf{Y} = \mathbb{R}$ □

2.6 Conclusion

We investigated a general abstract model for randomized quantization that provides a more suitable framework for certain optimal quantization problems than the ones usually considered in the source coding literature. In particular, our model formalizes the notion of randomly picking a quantizer from the set all *all* quantizers with a given number of output levels. Using this model, we proved the existence of an optimal randomized vector quantizer under the constraint that the quantizer output has a given distribution.

For stationary and memoryless source and output distributions, a rate-distortion theorem was proved, characterizing the minimum achievable distortion (or coding rate) at a given coding rate (or distortion) in the limit of large blocklengths, where unlimited common randomness is available between the encoder and the decoder (i.e., random variable Z in Models 1 and 2). In the next chapter, we consider the general case where the available common randomness may be rate limited. There, we completely characterize the set of achievable coding and common randomness rate pairs at any distortion level.

2.7 Proofs

2.7.1 Proof of Lemma 2.1

For a fixed probability measure μ on X define

$$\Delta_\mu = \big\{ v \in \mathcal{P}(\mathsf{X} \times \mathsf{Y}) : v(\,\cdot\, \times \mathsf{Y}) = \mu \big\}$$

(Δ_μ is the set of all probability measures in $\mathcal{P}(\mathsf{X} \times \mathsf{Y})$ whose X-marginal is μ). The following proposition, due to Borkar [21, Lemma 2.2], gives a characterization of the extreme points of Δ_μ.

Proposition 2.1. *Δ_μ is closed and convex, and its set of extreme points $\Delta_{\mu,e}$ is a Borel set in $\mathcal{P}(\mathsf{X} \times \mathsf{Y})$. Furthermore, $v \in \Delta_{\mu,e}$ if and only if $v(dx\,dy)$ can be disintegrated as*

$$v(dx\,dy) = Q(dy|x)\mu(dx)$$

where $Q(\,\cdot\,|x)$ is a Dirac measure for μ-a.e. x, i.e., there exists a measurable function $f : \mathsf{X} \to \mathsf{Y}$ such that $Q(\,\cdot\,|x) = \delta_{f(x)}(\,\cdot\,)$ for μ-a.e. x.

In fact, Borkar did not explicitly state Borel measurability of $\Delta_{\mu,e}$ in [21], but the proof of [21, Lemma 2.3] clearly implies this.

By Proposition 2.1 it is clear that $v \in \Gamma_\mu(M)$ if and only if $v \in \Delta_{\mu,e}$ and its marginal on Y is supported on a set having at most M elements, i.e., for some $L \le M$ and $\{y_1, \ldots, y_L\} \subset \mathsf{Y}$,

$$v(\mathsf{X} \times \{y_1, \ldots, y_L\}) = 1.$$

Let $\{y_n\}_{n\geq 1}$ be a countable dense subset of Y and define following subsets of $\Delta_{\mu,e}$:

$$\Omega_k = \bigcup_{n_1\geq 1,\dots,n_M\geq 1} \left\{ v \in \Delta_{\mu,e} : v\Big(\mathsf{X} \times \bigcup_{i=1}^{M} B(y_{n_i}, 1/k)\Big) = 1 \right\}$$

and

$$\Sigma = \bigcap_{k=1}^{\infty} \Omega_k$$

where $B(y,r)$ denotes the open ball in Y centered at y having radius r. Sets of the form

$$\left\{ v \in \mathcal{P}(\mathsf{X}\times\mathsf{Y}) : v\Big(\mathsf{X} \times \bigcup_{i=1}^{M} B(y_{n_i}, 1/k)\Big) = 1 \right\}$$

are Borel sets by [9, Proposition 7.25]. Since $\Delta_{\mu,e}$ is a Borel set, Ω_k is a Borel set for all k. Thus Σ is a Borel set in $\mathcal{P}(\mathsf{X}\times\mathsf{Y})$. We will prove that $\Sigma = \Gamma_\mu(M)$.

Since $\{y_n\}_{n\geq 1}$ is dense in Y, for any $v \in \Gamma_\mu(M)$ and $k \geq 1$ there exist $\tilde{n}_1, \dots, \tilde{n}_M$ such that $\mathrm{supp}(v(\mathsf{X}\times\,\cdot\,)) \subset \bigcup_{i=1}^{M} B(y_{\tilde{n}_i}, 1/k)$. Thus $\Gamma_\mu(M) \subset \Omega_k$ for all k, implying $\Gamma_\mu(M) \subset \Sigma$.

To prove the inclusion $\Sigma \subset \Gamma_\mu(M)$, let $v \in \Sigma$ and notice that for all k there exist $n_1^k, n_2^k, \dots, n_M^k$ such that

$$v\Big(\mathsf{X} \times \bigcup_{i=1}^{M} B(y_{n_i^k}, 1/k)\Big) = 1.$$

Let us define $K_n = \mathsf{X} \times \bigcap_{k=1}^{n} \bigcup_{i=1}^{M} B(y_{n_i^k}, 1/k)$. Clearly, $K_{n+1} \subset K_n$ and $v(K_n) = 1$, for all n. Letting

$$G = \bigcap_{k=1}^{\infty} \bigcup_{i=1}^{M} B(y_{n_i^k}, 1/k),$$

we have $v(\mathsf{X} \times G) = 1$. If we can prove that G has at most M distinct elements, then $v \in \Gamma_\mu(M)$. Assuming the contrary, there must exist distinct $\{\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_M, \hat{y}_{M+1}\} \subset G$. Let $\varepsilon = \min\{\|\hat{y}_i - \hat{y}_j\| : i,j = 1, \ldots, M+1, i \neq j\}$. Clearly, for $\frac{1}{k} < \frac{\varepsilon}{4}$, $\bigcup_{i=1}^{M} B(y_{n_i^k}, 1/k)$ cannot contain $\{\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_M, \hat{y}_{M+1}\}$, a contradiction. Thus G has at most M elements and we obtain $\Sigma = \Gamma_\mu(M)$. □

2.7.2 Proof of Theorem 2.1

We will need the following result which gives a necessary and sufficient condition for the measurability of a mapping from a measurable space to $\mathcal{P}(\mathsf{E})$, where E is a Borel space. It is proved for compact E in [31, Theorem 2.1] and for noncompact E it is the corollary of [9, Proposition 7.25].

Theorem 2.8. *Let $(\Omega, \mathcal{F})$ be a measurable space and let E be a Borel space. A mapping $h : \Omega \to \mathcal{P}(\mathsf{E})$ is measurable if and only if the real valued functions $\omega \mapsto h(\omega)(A)$ from Ω to $[0,1]$ are measurable for all $A \in \mathcal{B}(\mathsf{E})$.*

For any (q, ν) define $f : \mathbb{R}^m \to \Gamma_\mu(M)$ by $f(z) = \delta_{\mathsf{q}(x,z)}(dy)\mu(dx)$. By Theorem 2.8, f is measurable if and only if the mappings $z \mapsto \int \delta_{\mathsf{q}(x,z)}(C_x)\mu(dx)$ are measurable for all $C \in \mathcal{B}(\mathsf{X} \times \mathsf{Y})$, where $C_x = \{y : (x,y) \in C\}$. Observe that $\delta_{\mathsf{q}(x,z)}(C_x)$ is a measurable function of (x,z) because $\{(x,z) \in \mathsf{X} \times \mathsf{Z} : \delta_{\mathsf{q}(x,z)}(C_x) = 1\} = \{(x,z) \in \mathsf{X} \times \mathsf{Z} : (x, \mathsf{q}(x,z)) \in C\}$. By [14, Theorem 18.3] $\int \delta_{\mathsf{q}(x,z)}(C_x)\mu(dx)$ is measurable as well. Hence f is measurable.

Thus we can define the probability measure P supported on $\Gamma_\mu(M)$ by $P = \nu \circ f^{-1}$ (i.e., $P(B) = \nu(f^{-1}(B))$ for any Borel set $B \subset \Gamma_\mu(M)$). Then, for the corresponding v_P we have $(X,Y) \sim v_P$, i.e., for $C \in \mathcal{B}(\mathsf{X} \times \mathsf{Y})$,

$$\begin{aligned}
\Pr\big\{\big(X, \mathsf{q}(X,Z)\big) \in C\big\} &= \int_{\mathsf{Z}} \int_{\mathsf{X}} \delta_{\mathsf{q}(x,z)}(C_x)\mu(dx)\nu(dz) \\
&= \int_{\mathsf{Z}} f(z)(C)\nu(dz) \\
&= \int_{\Gamma_\mu(M)} v(C)P(dv) \\
&= v_P(C).
\end{aligned}$$

Conversely, let v_P be defined as in (2.3) with P supported on $\Gamma_\mu(M)$, i.e., $v_P = \int_{\Gamma_\mu(M)} vP(dv)$. Define the mapping $\Gamma_\mu(M) \ni v \mapsto q_v$, where q_v is the μ-a.e. defined quantizer in $\mathcal{Q}_M$, giving $v(dx\,dy) = \mu(dx)\delta_{q_v(x)}(dy)$. Since $\Gamma_\mu(M)$ is an uncountable Borel space, there is a measurable bijection (Borel isomorphism) $g : \mathbb{R}^m \to \Gamma_\mu(M)$ between $\mathbb{R}^m$ and $\Gamma_\mu(M)$ [32]. Now define q by $\mathsf{q}(x,z) = q_{g(z)}(x)$ and let $\nu = P \circ g$. Then for all z, $\mathsf{q}(\,\cdot\,, z)$ is a μ-a.e. defined M-level quantizer. However, it is not clear whether $\mathsf{q}(x,z)$ is measurable. Therefore we will construct another measurable function $\tilde{\mathsf{q}}(x,z)$ such that $\tilde{\mathsf{q}}(\,\cdot\,,z)$ is an M-level quantizer and $\tilde{\mathsf{q}}(\,\cdot\,,z) = \mathsf{q}(\,\cdot\,,z)$ μ-a.e., for all z. Then we will prove that $(X,Y) = (X, \tilde{\mathsf{q}}(X,Z)) \sim v_p$ where $Z \sim \nu$. Define the stochastic kernel on $\mathsf{X} \times \mathsf{Y}$ given $\Gamma_\mu(M)$ as

$$\gamma(dx\,dy|v) = v(dx\,dy).$$

Clearly, γ is well defined because $\Gamma_\mu(M)$ is a Borel subset of $\mathcal{P}(\mathsf{X} \times \mathsf{Y})$. Observe that for each $v \in \Gamma_\mu(M)$, we have

$$\gamma(C|v) = \int_{\mathsf{X}} \delta_{q_v(x)}(C_x)\mu(dx) \tag{2.13}$$

for $C \in \mathcal{B}(\mathsf{X} \times \mathsf{Y})$. Furthermore, by [9, Proposition 7.27] there exists a stochastic kernel $\eta(dy|x,v)$ on Y given $\mathsf{X} \times \Gamma_\mu(M)$ which satisfies for all $C \in \mathcal{B}(\mathsf{X} \times \mathsf{Y})$ and $v \in \Gamma_\mu(M)$,

$$\gamma(C|v) = \int_{\mathsf{X}} \eta(C_x|x,v)\mu(dx). \tag{2.14}$$

Since $\mathcal{B}(\mathsf{Y})$ is countably generated by the separability of Y, for any $v \in \Gamma_\mu(M)$ we have $\eta(\,\cdot\,|x,v) = \delta_{q_v(x)}(\,\cdot\,)$ μ-a.e. by (2.13) and (2.14). Since η is a stochastic kernel, it can be represented as a measurable function from $\mathsf{X} \times \Gamma_\mu(M)$ to $\mathcal{P}(\mathsf{Y})$, i.e.,

$$\eta : \mathsf{X} \times \Gamma_\mu(M) \to \mathcal{P}(\mathsf{Y}).$$

Define $\mathcal{P}_1(\mathsf{Y}) = \{\psi \in \mathcal{P}(\mathsf{Y}) : \psi(\{y\}) = 1 \text{ for some } y \in \mathsf{Y}\}$. $\mathcal{P}_1(\mathsf{Y})$ is a closed (thus measurable) subset of $\mathcal{P}(\mathsf{Y})$ by [73, Lemma 6.2]. Hence, $\mathsf{M} := \eta^{-1}(\mathcal{P}_1(\mathsf{Y}))$ is also measurable. Observe that for any $v \in \Gamma_\mu(M)$ we have $\mathsf{M}_v := \{x \in \mathsf{X} : (x,v) \in \mathsf{M}\} \supset \{x \in \mathsf{X} : \eta(\,\cdot\,|x,v) = \delta_{q_v(x)}(\,\cdot\,)\}$. Thus $\mu(\mathsf{M}_v) = 1$ for all $v \in \Gamma_\mu(M)$, which implies $\mu \otimes P\big(\mathsf{M}\big) = 1$. Define the function $\tilde{q}_v$ from $\mathsf{X} \times \Gamma_\mu(M)$ to Y as

$$\tilde{q}_v(x) = \begin{cases} \tilde{y}, & \text{if } (x,v) \in \mathsf{M}, \text{ where } \eta(\{\tilde{y}\}|x,v) = 1, \\ y, & \text{otherwise,} \end{cases}$$

where y is fixed. By construction, $\tilde{q}_v(x) = q_v(x)$ μ-a.e., for all $v \in \Gamma_\mu(M)$. For any

$C \in \mathcal{B}(\mathsf{Y})$ we have

$$\begin{aligned}\tilde{q}_v^{-1}(C) &= \{(x,v) \in \mathsf{X} \times \Gamma_\mu(M) : \tilde{q}_v(x) \in C\} \\ &= \{(x,v) \in \mathsf{M} : \tilde{q}_v(x) \in C\} \cup \{(x,v) \in \mathsf{M}^c : \tilde{q}_v(x) \in C\}.\end{aligned}$$

Clearly $\{(x,v) \in \mathsf{M}^c : \tilde{q}_v(x) \in C\} = \mathsf{M}^c$ or $\emptyset$ depending on whether or not y is an element of C. Hence, $\tilde{q}_v^{-1}(C) \in \mathcal{B}(\mathsf{X} \times \Gamma_\mu(M))$ if $\{(x,v) \in \mathsf{M} : \tilde{q}_v(x) \in C\} \in \mathcal{B}(\mathsf{X} \times \Gamma_\mu(M))$. But $\{(x,v) \in \mathsf{M} : \tilde{q}_v(x) \in C\} = \{(x,v) \in \mathsf{M} : \eta(C|x,v) = 1\}$ which is in $\mathcal{B}(\mathsf{X} \times \Gamma_\mu(M))$ by the measurability of $\eta(C|\,\cdot\,,\,\cdot\,)$. Thus, $\tilde{q}$ is a measurable function from $\mathsf{X} \times \Gamma_\mu(M)$ to Y.

Let us define $\tilde{\mathsf{q}}$ as $\tilde{\mathsf{q}}(x,z) = \tilde{q}_{g(z)}(x)$. By the measurability of g it is clear that $\tilde{\mathsf{q}}$ is measurable. In addition, for any $z \in \mathsf{Z}$ $\tilde{\mathsf{q}}(\,\cdot\,,z)$ is an M-level quantizer which is μ-a.e. equal to $\mathsf{q}(\,\cdot\,,z)$. Finally, if $Z \sim \nu$ is independent of X and $Y = \tilde{\mathsf{q}}(X,Z)$, then $(X,Y) \sim v_P$, i.e.,

$$\begin{aligned}\Pr\Big\{\big(X, \tilde{\mathsf{q}}(X,Z)\big) \in C\Big\} &= \int_{\mathsf{Z}} \int_{\mathsf{X}} \delta_{\tilde{\mathsf{q}}(x,z)}(C_x)\mu(dx)\nu(dz) \\ &= \int_{\Gamma_\mu(M)} \int_{\mathsf{X}} \delta_{\tilde{q}_v(x)}(C_x)\mu(dx)P(dv) \\ &= \int_{\Gamma_\mu(M)} \int_{\mathsf{M}_v} \eta(C_x|x,v)\mu(dx)P(dv) \\ &= \int_{\Gamma_\mu(M)} \gamma(C|v)P(dv) \\ &= \int_{\Gamma_\mu(M)} v(C)P(dv) \\ &= v_p(C). \quad \square\end{aligned}$$

2.7.3 Proof of Theorem 2.2

If (e, d, ν) is a Model 1 randomized quantizer, then setting $\mathsf{q}(x, z) = d(e(x, z), z)$ defines a Model 2 randomized quantizer (q, ν) such that the joint distributions of their inputs and outputs coincide.

Conversely, let (q, ν) be a Model 2 randomized quantizer. It is obvious that q can be decomposed into an encoder $e : \mathsf{X} \times \mathsf{Z} \to \{1, \ldots, M\}$ and decoder $d : \{1, \ldots, M\} \times \mathsf{Z} \to \mathsf{Y}$ such that $d(e(x, z), z) = \mathsf{q}(x, z)$ for all x and z. The difficulty lies in showing that this can be done so that the resulting e and d are measurable. In fact, we instead construct measurable e and d whose composition is $\mu \otimes \nu$-a.e. equal to q, which is sufficient to imply the theorem.

Let (q, ν) be a Model 2 randomized quantizer. Since $\mathbb{R}^n$ and $[0, 1]$ are both uncountable Borel spaces, there exists a Borel isomorphism $f : \mathbb{R}^n \to [0, 1]$ [32]. Define $\hat{\mathsf{q}} : \mathsf{X} \times \mathsf{Z} \to [0, 1]$ by $\hat{\mathsf{q}} = f \circ \mathsf{q}$. Hence, $\hat{\mathsf{q}}$ is measurable and, for any fixed z, $\hat{\mathsf{q}}(\cdot, z)$ is an M-level quantizer from X to $[0, 1]$. Also note that $\mathsf{q} = f^{-1} \circ \hat{\mathsf{q}}$.

Now for any fixed $z \in \mathsf{Z}$ consider only those output points of $\hat{\mathsf{q}}(\cdot, z)$ that occur with *positive* μ probability and order these according to their magnitude from the smallest to the largest. For $i = 1, \ldots, M$ let the function $f_i(z)$ take the value of the ith smallest such output point. If there is no such value, let $f_i(z) = 1$. We first prove that all the f_i are measurable and then define the encoder and the decoder in terms of these functions.

Observe that for any $a \in [0, 1]$, by definition

$$\{z \in \mathsf{Z} : f_1(z) \leq a\} = \Big\{ z \in \mathsf{Z} : \int_{\mathsf{X}} \delta_{\hat{\mathsf{q}}(x,z)}([0, a]) \mu(dx) > 0 \Big\},$$

where the set on the right hand side is a Borel set by Fubini's theorem. Hence, f_1 is a measurable function. Define $E_1 = \{(x, z) \in \mathsf{X} \times \mathsf{Z} : \hat{\mathsf{q}}(x, z) - f_1(z) = 0\}$, a Borel set. Letting $E_{1,z} = \{x \in \mathsf{X} : (x, z) \in E_1\}$ denote the z-section of E_1, for any $a \in [0, 1)$ we have

$$\{z \in \mathsf{Z} : f_2(z) \leq a\} = \Big\{z \in \mathsf{Z} : \int_{\mathsf{X} \setminus E_{1,z}} \delta_{\hat{\mathsf{q}}(x,z)}([0, a])\mu(dx) > 0\Big\},$$

and thus f_2 is measurable. Continuing in this fashion, we define the Borel sets $E_i = \{(x, z) : \hat{\mathsf{q}}(x, z) - f_i(z) = 0\}$ and write, for any $a \in [0, 1)$,

$$\{z \in \mathsf{Z} : f_i(z) \leq a\} = \Big\{z \in \mathsf{Z} : \int_{\mathsf{X} \setminus \bigcup_{j=1}^{i-1} E_{i,z}} \delta_{\hat{\mathsf{q}}(x,z)}([0, a])\mu(dx) > 0\Big\},$$

proving that f_i is measurable for all $i = 1, \dots, M$.

Define

$$\begin{aligned} N &= \big\{(x, z) \in \mathsf{X} \times \mathsf{Z} : \hat{\mathsf{q}}(x, z) \neq f_i(z) \text{ for all } i = 1, \dots, M\big\} \\ &= \mathsf{X} \times \mathsf{Z} \setminus \bigcup_{i=1}^{M} E_i. \end{aligned}$$

Clearly, N is a Borel set and $\mu \otimes \nu(N) = 0$ by Fubini's theorem and the definition of $f_1, \dots, f_M$. Now we can define

$$e(x, z) = \sum_{i=1}^{M} i\, 1_{\{\hat{\mathsf{q}}(x,z) = f_i(z)\}} + M\, 1_N(x, z)$$

and

$$d(i,z)=\sum_{j=1}^{M} f^{-1}\circ f_j(z)1_{\{i=j\}},$$

where 1_B denotes the indicator of event (or set) B. The measurability of $\hat{\mathsf{q}}$ and f, $f_1,\ldots,f_M$ implies that e and d are measurable. Since $d(e(x,z),z)=\hat{\mathsf{q}}(x,z)$ $\mu\otimes\nu$-a.e. by construction, this completes the proof. □

2.7.4 Proof of Theorem 2.3

I) ***Proof under Assumption 1***

To simplify the notation we redefine the reconstruction alphabet as $\mathsf{Y}=B$, so that Y is a compact subset of $\mathbb{R}^n$. It follows from the continuity of ρ that L is lower semicontinuous on $\mathcal{P}(\mathsf{X}\times\mathsf{Y})$ for the weak topology (see, e.g., [99, Lemma 4.3]). Hence, to show the existence of a minimizer for problem **(P1)** it would suffice to prove that $\Gamma^{\mathrm{R}}_{\mu,\psi}(M)=\Gamma^{\mathrm{R}}_{\mu}(M)\cap\Gamma_{\mu,\psi}$ is compact. It is known that $\Gamma_{\mu,\psi}$ is compact [99, Chapter 4], but unfortunately $\Gamma_{\mu}(M)$ is not closed [108] and it seems doubtful that $\Gamma^{\mathrm{R}}_{\mu}(M)$ is compact. Hence, we will develop a different argument which is based on optimal transport theory. We will first give the proof under Assumption 1; the proof under Assumption 2 then follows via a one-point compactification argument.

Let $\mathcal{P}_M(\mathsf{Y})=\{\psi_0\in\mathcal{P}(\mathsf{Y}):|\operatorname{supp}(\psi_0)|\leq M\}$ be the set of discrete distributions with M atoms or less on Y.

Lemma 2.7. *$\mathcal{P}_M(\mathsf{Y})$ is compact in $\mathcal{P}(\mathsf{Y})$.*

Proof. Let $\{\psi_n\}$ be an arbitrary sequence in $\mathcal{P}_M(\mathsf{Y})$. Each ψ_n can be represented by points $(y_1^n,\ldots,y_M^n)=y^n\in\mathsf{Y}^M$ and $(p_1^n,\ldots,p_M^n)=p^n\in K_s$, where $K_s=$

$\{(p_1, \ldots, p_M) \in \mathbb{R}^M : \sum_{i=1}^M p_i = 1, p_i \geq 0\}$ is the probability simplex in $\mathbb{R}^M$. Let $w_n = (y^n, p^n)$. Since $\mathsf{Y}^M \times K_s$ is compact, there exists a subsequence $\{w^{n_k}\}$ converging to some w in $\mathsf{Y}^M \times K_s$. Let ψ be the probability measure in $\mathcal{P}_M(\mathsf{Y})$ which is represented by w. It straightforward to show that ψ is a weak limit of $\{\psi^{n_k}\}$. This completes the proof. □

Define

$$\hat{\Gamma}_\mu(M) = \bigcup_{\psi_0 \in \mathcal{P}_M(\mathsf{Y})} \big\{\hat{v} \in \Gamma_{\mu,\psi_0} : L(\hat{v}) = \min_{v \in \Gamma_{\mu,\psi_0}} L(v)\big\}.$$

The elements of $\hat{\Gamma}_\mu(M)$ are the probability measures which solve the optimal transport problem (see, e.g., [99]) for fixed input marginal μ and some output marginal ψ_0 in $\mathcal{P}_M(\mathsf{Y})$. At the end of this proof Lemma 2.11 shows that $\hat{\Gamma}_\mu(M)$ is a Borel set. Let $\hat{\Gamma}_\mu^{\mathrm{R}}(M)$ be the randomization of $\hat{\Gamma}_\mu(M)$, obtained by replacing $\Gamma_\mu(M)$ with $\hat{\Gamma}_\mu(M)$ in (2.4). Define the optimization problem **(P2)** as

$$\begin{aligned} \textbf{(P2)}\ & \text{minimize } L(v) \\ & \text{subject to } v \in \hat{\Gamma}^{\mathrm{R}}_{\mu,\psi}(M), \end{aligned}$$

where $\hat{\Gamma}^{\mathrm{R}}_{\mu,\psi}(M) = \hat{\Gamma}^{\mathrm{R}}_\mu(M) \cap \Gamma_{\mu,\psi}$.

Proposition 2.2. *For any $v^* \in \Gamma^{\mathrm{R}}_{\mu,\psi}(M)$ there exists $\hat{v} \in \hat{\Gamma}^{\mathrm{R}}_{\mu,\psi}(M)$ such that $L(v^*) \geq L(\hat{v})$. Hence, the distortion of any minimizer in* **(P2)** *is less than or equal to the distortion of a minimizer in* **(P1)**.

To prove Proposition 2.2 we need the following lemma.

Lemma 2.8. *Let P be a probability measure on $\Gamma_\mu(M)$. Then there exists a measurable mapping $f : \Gamma_\mu(M) \to \hat{\Gamma}_\mu(M)$ such that $v(\mathsf{X} \times \cdot) = f(v)(\mathsf{X} \times \cdot)$ and*

$L(v) \geq L(f(v))$, *P-a.e.*

Proof. Define the projections $f_1 : \Gamma_\mu(M) \to \mathcal{P}_M(\mathsf{Y})$ and $f_2 : \hat{\Gamma}_\mu(M) \to \mathcal{P}_M(\mathsf{Y})$ by $f_1(v) = v(\mathsf{X} \times \cdot\,)$, $f_2(v) = v(\mathsf{X} \times \cdot\,)$. Note that f_1 is continuous and f_2 is continuous and onto. Define $\tilde{P} = P \circ f_1^{-1}$ on $\mathcal{P}_M(\mathsf{Y})$. By Yankov's lemma [36, Appendix 3] there exists a mapping g from $\mathcal{P}_M(\mathsf{Y})$ to $\hat{\Gamma}_\mu(M)$ such that $f_2(g(\psi)) = \psi$ $\tilde{P}$-a.e. Then, it is straightforward to show that $f = g \circ f_1$ satisfies conditions $v(\mathsf{X} \times \cdot\,) = f(v)(\mathsf{X} \times \cdot\,)$ and $L(v) \geq L(f(v))$, P-a.e. □

Proof of Proposition 2.2. Let $v^* \in \Gamma^{\mathrm{R}}_{\mu,\psi}(M)$, i.e.,

$$v^* = \int_{\Gamma_\mu(M)} v P(dv) \text{ and } v^*(\mathsf{X} \times \cdot\,) = \psi.$$

By Lemma 2.8 there exists $f : \Gamma_\mu(M) \to \hat{\Gamma}_\mu(M)$ such that $v(\mathsf{X}\times\cdot\,) = f(v)(\mathsf{X}\times\cdot\,)$ and $L(v) \geq L(f(v))$, P-a.e. Define $\tilde{P} = P \circ f^{-1} \in \mathcal{P}(\hat{\Gamma}_\mu(M))$ and $\hat{v} = \int_{\hat{\Gamma}_\mu(M)} v\tilde{P}(dv) \in \hat{\Gamma}^{\mathrm{R}}_\mu(M)$. We have

$$\begin{aligned} L(v^*) &= \int_{\Gamma_\mu(M)} L(v)P(dv) \geq \int_{\Gamma_\mu(M)} L(f(v))P(dv) \\ &= \int_{\hat{\Gamma}_\mu(M)} L(v)\tilde{P}(dv) = L(\hat{v}) \end{aligned}$$

as well as

$$\begin{aligned} v^*(\mathsf{X} \times \cdot\,) &= \int_{\Gamma_\mu(M)} v(\mathsf{X} \times \cdot\,)P(dv) \\ &= \int_{\Gamma_\mu(M)} f(v)(\mathsf{X} \times \cdot\,)P(dv) \\ &= \int_{\hat{\Gamma}_\mu(M)} v(\mathsf{X} \times \cdot\,)\tilde{P}(dv) = \hat{v}(\mathsf{X} \times \cdot\,). \end{aligned}$$

This completes the proof. $\square$

Recall the set Δ_μ and its set of its extreme points $\Delta_{\mu,e}$ from Proposition 2.1. It is proved in [21] and [20] that any $\tilde{v} \in \Delta_\mu$ can be written as $\tilde{v} = \int_{\Delta_{\mu,e}} vP(dv)$ for some $P \in \mathcal{P}(\Delta_{\mu,e})$. By Proposition 2.1 we also have $\Gamma_\mu(M) \subset \Delta_{\mu,e}$. The following lemma is based on these two facts.

Lemma 2.9. *Let $\tilde{v} \in \Delta_\mu$ which is represented as $\tilde{v} = \int_{\Delta_{\mu,e}} vP(dv)$. If $\tilde{v}(\mathsf{X} \times \cdot) \in \mathcal{P}_M(\mathsf{Y})$, then $P(\Gamma_\mu(M)) = 1$.*

Proof. Since $\tilde{v}(\mathsf{X} \times \cdot) \in \mathcal{P}_M(\mathsf{Y})$, there exist a finite set $B \subset Y$ having $M' \leq M$ elements such that $\tilde{v}(\mathsf{X} \times B) = 1$. We have

$$\begin{aligned}\tilde{v}(\mathsf{X} \times B) &= \int_{\Delta_{\mu,e}} v(\mathsf{X} \times B)P(dv) \\ &= \int_{\Delta_{\mu,e} \setminus \Gamma_\mu(M)} v(\mathsf{X} \times B)P(dv) \\ &\quad + \int_{\Gamma_\mu(M)} v(\mathsf{X} \times B)P(dv).\end{aligned}$$

Since $v(\mathsf{X} \times B) < 1$ for all $v \in \Delta_{\mu,e} \setminus \Gamma_\mu(M)$, we obtain $P(\Gamma_\mu(M)) = 1$. $\square$

Lemma 2.9 implies $\hat{\Gamma}_\mu(M) \subset \Gamma_\mu^{\mathrm{R}}(M)$ because $v(\mathsf{X} \times \cdot) \in \mathcal{P}_M(\mathsf{Y})$ when $v \in \hat{\Gamma}_\mu(M)$. Define $h : \mathcal{P}(\Gamma_\mu(M)) \to \Delta_\mu$ as follows:

$$h(P)(\,\cdot\,) = \int_{\Gamma_\mu(M)} v(\,\cdot\,)P(dv). \tag{2.15}$$

It is clear that the range of h is $\Gamma_\mu^{\mathrm{R}}(M) \subset \Delta_\mu$.

Lemma 2.10. *h is continuous.*

Proof. Assume $\{P_n\}$ converges weakly to P in $\mathcal{P}(\Gamma_\mu(M))$. Then, for any continuous and bounded real function f on $\mathsf{X} \times \mathsf{Y}$

$$\lim_{n\to\infty} \int_{\Gamma_\mu(M)} \int_{\mathsf{X}\times\mathsf{Y}} f(x,y)v(dx\,dy)P_n(dv) \\ = \int_{\Gamma_\mu(M)} \int_{\mathsf{X}\times\mathsf{Y}} f(x,y)v(dx\,dy)P(dv)$$

if the mapping $v \mapsto \int_{\mathsf{X}\times\mathsf{Y}} f(x,y)v(dx\,dy)$ is continuous and bounded on $\Gamma_\mu(M)$. Clearly this mapping is continuous by the definition of weak convergence and bounded by the boundedness of f. Thus

$$\int_{\Gamma_\mu(M)} vP_n(dv) \to \int_{\Gamma_\mu(M)} vP(dv)$$

weakly, completing the proof. □

Since $\hat{\Gamma}_\mu(M) \subset \Gamma_\mu^{\mathrm{R}}(M)$, we have $\mathcal{P}^{\mathrm{opt}}(\Gamma_\mu(M)) := h^{-1}(\hat{\Gamma}_\mu(M)) \subset \mathcal{P}(\Gamma_\mu(M))$, which is measurable by the measurability of $\hat{\Gamma}_\mu(M)$ and h. Let $g : \mathcal{P}^{\mathrm{opt}}(\Gamma_\mu(M)) \to \hat{\Gamma}_\mu(M)$ be the restriction of h to $\mathcal{P}^{\mathrm{opt}}(\Gamma_\mu(M))$. Clearly g is measurable and onto. By Yankov's lemma [36] for any probability measure P on $\hat{\Gamma}_\mu(M)$ there exists a measurable mapping $\varphi : \hat{\Gamma}_\mu(M) \to \mathcal{P}^{\mathrm{opt}}(\Gamma_\mu(M))$ such that $g(\varphi(\hat{v})) = \hat{v}$ P-a.e. In addition, since $\varphi(\hat{v}) \in g^{-1}(\hat{v})$ P-a.e., we have

$$L(\hat{v}) = \int_{\Gamma_\mu(M)} L(v)\varphi(\hat{v})(dv) \tag{2.16}$$

and

$$\hat{v}(\mathsf{X} \times \,\cdot\,) = \int_{\Gamma_\mu(M)} v(\mathsf{X} \times \,\cdot\,)\varphi(\hat{v})(dv) \tag{2.17}$$

P-a.e. Define the stochastic kernel $\Pi(dv|\hat{v})$ on $\Gamma_\mu(M)$ given $\hat{\Gamma}_\mu(M)$ as

$$\Pi(dv|\hat{v}) = \varphi(\hat{v})(dv). \tag{2.18}$$

Since φ is measurable, $\Pi(dv|\hat{v})$ is well defined. Observe that both φ and $\Pi(dv|\hat{v})$ depend on the probability measure $P \in \hat{\Gamma}_\mu(M)$.

Proposition 2.3. *If* **(P2)** *has a minimizer v^*, then we can find $\bar{v} \in \Gamma^{\mathrm{R}}_{\mu,\psi}(M)$ such that $L(\bar{v}) = L(v^*)$, implying that $\bar{v}$ is a minimizer for* **(P1)***.*

Proof. v^* can be written as $v^* = \int_{\hat{\Gamma}_\mu(M)} \hat{v} P(d\hat{v})$. Consider the stochastic kernel $\Pi(dv|\hat{v})$ defined in (2.18). Composing P with Π we obtain a probability measure Λ on $\hat{\Gamma}_\mu(M) \times \Gamma_\mu(M)$ given by

$$\Lambda(d\hat{v}\, dv) = P(d\hat{v})\Pi(dv|\hat{v}). \tag{2.19}$$

Let $\tilde{P} = \Lambda(\hat{\Gamma}_\mu(M) \times \,\cdot\,) \in \mathcal{P}(\Gamma_\mu(M))$. Define the randomized quantizer $\bar{v} \in \Gamma^{\mathrm{R}}_\mu(M)$ as $\bar{v} = \int_{\Gamma_\mu(M)} v \tilde{P}(dv)$. We show that $L(v^*) = L(\bar{v})$ and $v^*(\mathsf{X} \times \,\cdot\,) = \bar{v}(\mathsf{X} \times \,\cdot\,)$ which will complete the proof. We have

$$\begin{aligned}
L(v^*) &= \int_{\hat{\Gamma}_\mu(M)} L(\hat{v}) P(d\hat{v}) \\
&= \int_{\hat{\Gamma}_\mu(M)} \int_{\Gamma_\mu(M)} L(v)\varphi(\hat{v})(dv) P(d\hat{v}) \quad \text{(by (2.16))} \\
&= \int_{\hat{\Gamma}_\mu(M) \times \Gamma_\mu(M)} L(v) \Lambda(d\hat{v}\, dv) \quad \text{(by (2.18))} \\
&= \int_{\Gamma_\mu(M)} L(v) \tilde{P}(dv) = L(\bar{v}).
\end{aligned}$$

Similarly,

$$
\begin{aligned}
v^*(\mathsf{X}\times\ \cdot\ ) &= \int_{\hat{\Gamma}_\mu(M)} \hat{v}(\mathsf{X}\times\ \cdot\ )P(d\hat{v}) \\
&= \int_{\hat{\Gamma}_\mu(M)} \int_{\Gamma_\mu(M)} v(\mathsf{X}\times\ \cdot\ )\varphi(\hat{v})(dv)P(d\hat{v}) \text{ (by (2.17))} \\
&= \int_{\hat{\Gamma}_\mu(M)\times\Gamma_\mu(M)} v(\mathsf{X}\times\ \cdot\ )\Lambda(d\hat{v}\,dv) \quad \text{(by (2.18))} \\
&= \int_{\Gamma_\mu(M)} v(\mathsf{X}\times\ \cdot\ )\tilde{P}(dv) = \bar{v}(\mathsf{X}\times\ \cdot\ ).
\end{aligned}
$$

By Proposition 2.2, $\bar{v}$ is a minimizer for **(P1)**. □

Hence, to prove the existence of a minimizer for **(P1)** it is enough prove the existence of a minimizer for **(P2)**. Before proceeding to the proof we need to define the optimal transport problem. Optimal transport problem for marginals $\pi \in \mathcal{P}(\mathsf{X})$, $\lambda \in \mathcal{P}(\mathsf{Y})$ and cost function $c : \mathsf{X}\times\mathsf{Y} \to [0,\infty]$ is defined as:

$$
\begin{aligned}
&\text{minimize } \int_{\mathsf{X}\times\mathsf{Y}} c(x,y)v(dx\,dy) \\
&\text{subject to } v \in \Gamma_{\pi,\lambda}.
\end{aligned} \tag{2.20}
$$

The following result is about the structure of the optimal v in (2.20). It uses the concept of c-cyclically monotone sets [99, Definition 5.1]. A set $B \subset \mathsf{X}\times\mathsf{Y}$ is said to be c-cyclically monotone if for any $N \geq 1$ and pairs $(x_1,y_1),\ldots,(x_N,y_N)$ in B, the following inequality holds:

$$
\sum_{i=1}^{N} c(x_i,y_i) \leq \sum_{i=1}^{N} c(x_i,y_{i+1}),
$$

where $y_{N+1} \coloneqq y_1$.

Informally, when $v \in \Gamma_{\pi,\lambda}$ is concentrated on a c-cyclically monotone set, then its cost cannot be improved by local perturbations; see the discussion in [99, Chapter 5]. The following result shows that an optimal v must concentrate on a c-cyclically monotone set.

Proposition 2.4 ([75, Theorem 1.2], [99, Theorem 5.10])**.** *Let $c : \mathsf{X} \times \mathsf{Y} \to [0, \infty]$ be continuous. If $v \in \Gamma_{\pi,\lambda}$ is a solution to the optimal transport problem (2.20) and $\int_{\mathsf{X}\times\mathsf{Y}} c(x, y)v(dx\, dy) < \infty$, then v is concentrated on some c-cyclically monotone set.*

For any $K \subset \mathcal{P}(\mathsf{X})$ and $S \subset \mathcal{P}(\mathsf{Y})$ define $\Xi_{K,S} \subset \mathcal{P}(\mathsf{X} \times \mathsf{Y})$ as the set of probability measures which are concentrated on some c-cyclically monotone set and solve (2.20) for some $\pi \in K$, $\lambda \in S$. The following result is a slight modification of [99, Corollary 5.21].

Proposition 2.5. *If K and S are compact, then $\Xi_{K,S}$ is compact.*

Proof. Let $\{v_n\}$ be a sequence in $\Xi_{K,S}$. It can be shown that there exists a subsequence $\{v_{n_k}\}$ converging to v whose marginals belong to K and S [99, Lemma 4.4]. Since each v_{n_k} is concentrated on a c-cyclically monotone set by assumption, it can be shown by using the continuity of c that v is also concentrated on a c-cyclically monotone set (see proof of Theorem 5.20 in [99]). Then v is also an element of $\Xi_{K,S}$ by [75, Theorem B]. □

Since $\{\mu\}$ and $\mathcal{P}_M(\mathsf{Y})$ are both compact, we obtain that $\Xi_{\{\mu\},\mathcal{P}_M(\mathsf{Y})}$ is compact. Thus it follows that $\mathcal{P}(\Xi_{\{\mu\},\mathcal{P}_M(\mathsf{Y})})$ is also compact. Furthermore, by Proposition 2.4 we have $\Xi_{\{\mu\},\mathcal{P}_M(\mathsf{Y})} \supset \{v \in \hat{\Gamma}_\mu(M) : L(v) < \infty\}$. Hence the randomization can be restricted to $\Xi_{\{\mu\},\mathcal{P}_M(\mathsf{Y})}$ when defining $\hat{\Gamma}^{\mathrm{R}}_\mu(M)$ for **(P2)**. Let $\Xi^{\mathrm{R}}_{\{\mu\},\mathcal{P}_M(\mathsf{Y})}$ be the randomization of $\Xi_{\{\mu\},\mathcal{P}_M(\mathsf{Y})}$ obtained by replacing $\Gamma_\mu(M)$ with $\Xi_{\{\mu\},\mathcal{P}_M(\mathsf{Y})}$ in (2.4).

One can show that the mapping $\mathcal{P}(\Xi_{\{\mu\},\mathcal{P}_M(\mathsf{Y})}) \ni P \mapsto v_P \in \Xi^{\mathrm{R}}_{\{\mu\},\mathcal{P}_M(\mathsf{Y})}$ is continuous by using the same proof as in Lemma 2.10. Thus $\Xi^{\mathrm{R}}_{\{\mu\},\mathcal{P}_M(\mathsf{Y})}$ is the continuous image of a compact set, and thus it is also compact. This, together with the compactness of $\Gamma_{\mu,\psi}$ and the lower semicontinuity of L, implies the existence of the minimizer for **(P2)** under Assumption 2.1.

To tie up a loose end, we still have to show that $\hat{\Gamma}_\mu(M)$ is measurable, which will complete the proof under Assumption 2.1.

Lemma 2.11. $\hat{\Gamma}_\mu(M)$ *is a Borel set.*

Proof. Let us define $\hat{\Gamma}^{\mathrm{f}}_\mu(M) := \{v \in \hat{\Gamma}_\mu(M) : L(v) < \infty\}$ and $\hat{\Gamma}^{\infty}_\mu(M) = \hat{\Gamma}_\mu(M) \setminus \hat{\Gamma}^{f}_\mu(M)$. Since solutions to the optimal transport problem having finite costs must concentrate on c-cyclically monotone sets by Proposition 2.4, we have $\hat{\Gamma}^{\mathrm{f}}_\mu(M) = \{v \in \Xi_{\{\mu\},\mathcal{P}_M(\mathsf{Y})} : L(v) < \infty\}$. Hence, $\hat{\Gamma}^{f}_\mu(M)$ is a Borel set since $\Xi_{\{\mu\},\mathcal{P}_M(\mathsf{Y})}$ is compact and L is lower semi-continuous. Recall the continuous mapping f_2 in the proof of Lemma 2.8. Since $\Xi_{\{\mu\},\mathcal{P}_M(\mathsf{Y})}$ is compact, $\{v \in \Xi_{\{\mu\},\mathcal{P}_M(\mathsf{Y})} : L(v) \leq N\}$ is also compact for all $N \geq 0$. Hence, $f_2\big(\hat{\Gamma}^{\mathrm{f}}_\mu(M)\big) = \bigcup_{N=0}^{\infty} f_2\big(\{v \in \Xi_{\{\mu\},\mathcal{P}_M(\mathsf{Y})} : L(v) \leq N\}\big)$ is σ-compact, so a Borel set, in $\mathcal{P}_M(\mathsf{Y})$. Since $f_2\big(\hat{\Gamma}^{\infty}_\mu(M)\big) = \mathcal{P}_M(\mathsf{Y}) \setminus f_2\big(\hat{\Gamma}^{f}_\mu(M)\big)$, $f_2\big(\hat{\Gamma}^{\infty}_\mu(M)\big)$ is also a Borel set. Note that for any $v \in \hat{\Gamma}^{\infty}_\mu(M)$ we have $L(v) = \infty$, which means that all $\tilde{v}$ with the same marginals as v are also in $\hat{\Gamma}^{\infty}_\mu(M)$. This implies $\hat{\Gamma}^{\infty}_\mu(M) = f_2^{-1}\big(f_2\big(\hat{\Gamma}^{\infty}_\mu(M)\big)\big)$. Hence, $\hat{\Gamma}^{\infty}_\mu(M)$ is a Borel set. □

II) ***Proof under Assumption 2***

It is easy to check that the proof under Assumption 2.1 remains valid if X and Y are arbitrary uncountable Polish spaces such that Y is compact, and the distortion measure ρ is an extended real valued function (no steps exploited the special structure of $\mathbb{R}^n$). Let Y be the one-point compactification of $\mathbb{R}^n$ [32]. Y is clearly an uncountable

Polish space. Define the extended real valued distortion measure $\rho : \mathsf{X} \times \mathsf{Y} \to [0, \infty]$ by

$$\rho(x,y) = \begin{cases} \|x-y\|^2, & \text{if } y \in \mathbb{R}^n \\ \infty, & \text{if } y = \infty. \end{cases} \tag{2.21}$$

It is straightforward to check that ρ is continuous. Define L on $\mathcal{P}(\mathsf{X}\times\mathsf{Y})$ as before, but with this new distortion measure ρ. The proof under Assumption 2.1 gives a minimizer $v^* = \int_{\Gamma_\mu(M)} v P(dv)$ for **(P1)**. Define $\tilde{\Gamma}_\mu(M) = \{v \in \Gamma_\mu(M) : v(\mathsf{X} \times \{\infty\}) = 0\}$. Since $L(v^*) < \infty$ by assumption, $P(\tilde{\Gamma}_\mu(M)) = 1$. This implies that v^* is also a minimizer for the problem **(P1)** when $\mathsf{X} = \mathsf{Y} = \mathbb{R}^n$ and $\rho = \|x-y\|^2$. □

2.7.5 Proof of Theorem 2.4

From the proof of Theorem 2.3 recall the set $\hat{\Gamma}_\mu(M)$ of probability measures which solve the optimal mass transport problem for fixed input marginal μ and some output marginal ψ_0 in $\mathcal{P}_M(\mathsf{Y})$. It is known that if μ admits a density and $\rho(x,y) = \|x-y\|^2$, then each $v \in \hat{\Gamma}_\mu(M)$ is in the form $v(dx\,dy) = \mu(dx)\delta_{q(x)}(dy)$ for some $q \in \mathcal{Q}_{M,\mathrm{c}}$ (see, e.g. [68, Theorem 1]). Thus in this case $\hat{\Gamma}_\mu(M) \subset \Gamma_\mu(M)$, which implies that $\hat{\Gamma}^{\mathrm{R}}_{\mu,\psi}(M) \subset \Gamma^{\mathrm{R,c}}_{\mu,\psi}(M) \subset \Gamma^{\mathrm{R}}_{\mu,\psi}(M)$. Recall the problem **(P2)** in the proof of Theorem 2.3. It was shown that **(P2)** has a minimizer v^*. It is clear from the previous discussion that v^* is obtained by randomizing over the set of quantizers having convex codecells represented by $\hat{\Gamma}_\mu(M)$. On the other hand, v^* is also a minimizer for the problem **(P1)** by Proposition 2.2 in the proof of Theorem 2.3. □

2.7.6 Proof of Theorem 2.5

Recall the continuous mapping $h : \mathcal{P}(\Gamma_\mu(M)) \to \Gamma_\mu^{\mathrm{R}}(M)$ defined in (2.15). Let $\mathcal{P}_F(\Gamma_\mu(M))$ denote the set of probability measures on $\Gamma_\mu(M)$ having finite support. Clearly $h(\mathcal{P}_F(\Gamma_\mu(M))) = \Gamma_\mu^{\mathrm{FR}}(M)$.

Lemma 2.12. *$\Gamma_\mu^{\mathrm{FR}}(M)$ is dense in $\Gamma_\mu^{\mathrm{R}}(M)$.*

Proof. Since $\Gamma_\mu(M)$ is a separable metric space, $\mathcal{P}_F(\Gamma_\mu(M))$ is dense in $\mathcal{P}(\Gamma_\mu(M))$ by [73, Theorem 6.3]. Since $\Gamma_\mu^{\mathrm{FR}}(M)$ is the image of a $\mathcal{P}_F(\Gamma_\mu(M))$ under the continuous function h which maps $\mathcal{P}(\Gamma_\mu(M))$ onto $\Gamma_\mu^{\mathrm{R}}(M)$, it is dense in $\Gamma_\mu^{\mathrm{R}}(M)$. □

Recall that the Prokhorov metric on $\mathcal{P}(\mathsf{E})$, where (E, d) is a metric space, is defined as [15]

$$d_P(v,\nu) = \inf\left\{\alpha : v(A) \le \nu(A^\alpha) + \alpha, \nu(A) \le v(A^\alpha) + \alpha \text{ for all } A \in \mathcal{B}(\mathsf{E})\right\} \quad (2.22)$$

where

$$A^\alpha = \left\{e \in \mathsf{E} : \inf_{e' \in A} d(e,e') < \alpha\right\}.$$

Hence for $v, \nu \in \mathcal{P}(\mathsf{X} \times \mathsf{Y})$,

$$\begin{aligned} d_P(v,\nu) &\ge \inf\left\{\alpha : v(\mathsf{X} \times B) \le \nu((\mathsf{X} \times B)^\alpha) + \alpha,\right. \\ &\qquad \left.\nu(\mathsf{X} \times B) \le v((\mathsf{X} \times B)^\alpha) + \alpha, B \in \mathcal{B}(\mathsf{Y})\right\} \\ &= d_P\big(v(\mathsf{X} \times \,\cdot\,), \nu(\mathsf{X} \times \,\cdot\,)\big) \end{aligned}$$

(note that $(\mathsf{X} \times B)^\alpha = \mathsf{X} \times B^\alpha$). This implies

$$G_\psi^\alpha := \{v \in \mathcal{P}(\mathsf{X} \times \mathsf{Y}) : v(\mathsf{X} \times \,\cdot\,) \in B(\psi, \alpha)\}$$

$$\supset \quad \{v \in \mathcal{P}(\mathsf{X} \times \mathsf{Y}) : d_P(\hat{v}, v) < \alpha\}, \tag{2.23}$$

where $\hat{v}$ is such that $\hat{v}(\mathsf{X} \times \cdot\,) = \psi$ and $\alpha > 0$. Recall that given a metric space E and $A \subset \mathsf{E}$, a set $B \subset A$ is relatively open in A if $B = A \cap U$ for some open set $U \subset \mathsf{E}$.

Lemma 2.13. *$\mathcal{M}^{\delta}_{\mu,\psi}$ is relatively open in $\Gamma^{\mathrm{R}}_{\mu}(M)$.*

Proof. Since $\mathcal{M}^{\delta}_{\mu,\psi} = G^{\delta}_{\psi} \cap \Gamma^{\mathrm{R}}_{\mu}(M)$, it is enough to prove that G^{δ}_{ψ} is open in $\mathcal{P}(\mathsf{X} \times \mathsf{Y})$. Let $\tilde{v} \in G^{\delta}_{\psi}$. Then $\tilde{v}(\mathsf{X} \times \cdot\,) \in B(\psi, \delta)$ by definition, and there exists $\delta_0 > 0$ such that $B(\tilde{v}(\mathsf{X} \times \cdot\,), \delta_0) \subset B(\psi, \delta)$. By (2.23) we have

$$\big\{v \in \mathcal{P}(\mathsf{X} \times \mathsf{Y}) : d_P(\tilde{v}, v) < \delta_0\big\} \subset G^{\delta_0}_{v(\mathsf{X} \times \cdot\,)}\,.$$

We also have $G^{\delta_0}_{v(\mathsf{X} \times \cdot\,)} \subset G^{\delta}_{\psi}$ since $B(\tilde{v}(\mathsf{X} \times \cdot\,), \delta_0) \subset B(\psi, \delta)$. This implies that G^{δ}_{ψ} is open in $\mathcal{P}(\mathsf{X} \times \mathsf{Y})$. □

I) ***Case 1***

First we treat the case $L(v) > \inf_{v' \in \Gamma_{\mu}(M)} L(v')$. If ρ is continuous and bounded, then L is continuous. Hence, $\{v' \in \Gamma^{\mathrm{R}}_{\mu}(M) : L(v') < L(v)\}$ is relatively open in $\Gamma^{\mathrm{R}}_{\mu}(M)$. Define $F := \{v' \in \Gamma^{\mathrm{R}}_{\mu}(M) : L(v') < L(v)\}$.

Lemma 2.14. *$F \cap \mathcal{M}^{\delta}_{\mu,\psi}$ is* nonempty *and relatively open in $\Gamma^{\mathrm{R}}_{\mu}(M)$.*

Proof. By Lemma 2.13 and the above discussion the intersection is clearly relatively open in $\Gamma^{\mathrm{R}}_{\mu}(M)$, so we need to show that it is not empty. Since $L(v) > \inf_{v' \in \Gamma_{\mu}(M)} L(v')$, there exists $\tilde{v} \in \Gamma_{\mu}(M)$ such that $L(\tilde{v}) < L(v)$. Define the sequence of randomized quantizers $\{v_n\} \in \Gamma^{\mathrm{R}}_{\mu}(M)$ by letting $v_n = \frac{1}{n}\tilde{v} + (1 - \frac{1}{n})v$. Then,

$v_n \to v$ weakly because for any continuous and bounded real function f on $\mathsf{X} \times \mathsf{Y}$

$$\lim_{n\to\infty} \left| \int_{\mathsf{X}\times\mathsf{Y}} f dv_n - \int_{\mathsf{X}\times\mathsf{Y}} f dv \right| = \lim_{n\to\infty} \frac{1}{n} \left| \int_{\mathsf{X}\times\mathsf{Y}} f d\tilde{v} - \int_{\mathsf{X}\times\mathsf{Y}} f dv \right| = 0.$$

Hence there exists n_0 such that $v_n \in M^{\delta}_{\mu,\psi}$ for all $n \geq n_0$. On the other hand, for any n

$$\begin{aligned} L(v_n) &= L\left(\frac{1}{n}\tilde{v} + \Big(1 - \frac{1}{n}\Big)v\right) \\ &= \frac{1}{n}L(\tilde{v}) + \Big(1 - \frac{1}{n}\Big)L(v) \\ &< L(v). \end{aligned}$$

This implies $v_n \in \mathcal{M}^{\delta}_{\mu,\psi} \cap F$ for all $n \geq n_0$, completing the proof. □

Hence, we can conclude that there exists finitely randomized quantizer $v_F \in F \cap M^{\delta}_{\mu,\psi}$ by Lemmas 2.12 and 2.14. By the definition of F we also have $L(v_F) < L(v)$. This completes the proof of the theorem for this case.

II) ***Case 2***

The case $L(v) = \inf_{v' \in \Gamma_\mu(M)} L(v') := L^*$ is handled similarly. Define the subset of $\Gamma_\mu(M)$ whose elements correspond to optimal quantizers:

$$\Gamma_{\mu,\mathrm{opt}}(M) = \{v^{'} \in \Gamma_\mu(M) : L(v^{'}) = L^*\}.$$

Define $\Gamma_{\mu,\mathrm{opt}}(M) = L^{-1}(L^*) \cap \Gamma_\mu(M)$ and let $\Gamma^{\mathrm{R}}_{\mu,\mathrm{opt}}(M)$ be the randomization of $\Gamma_{\mu,\mathrm{opt}}(M)$, obtained by replacing $\Gamma_\mu(M)$ with $\Gamma_{\mu,\mathrm{opt}}(M)$ in (2.4). Note that if $L(v) = L^*$, then v is obtained by randomizing over the set $\Gamma_{\mu,\mathrm{opt}}(M)$, i.e., $v \in \Gamma^{\mathrm{R}}_{\mu,\mathrm{opt}}(M)$. Let $\Gamma^{\mathrm{FR}}_{\mu,\mathrm{opt}}(M)$ denote the set obtained by the finite randomization of $\Gamma_{\mu,\mathrm{opt}}(M)$. By using the same proof method as in Lemma 2.12 we can prove that $\Gamma^{\mathrm{FR}}_{\mu,\mathrm{opt}}(M)$ is dense in $\Gamma^{\mathrm{R}}_{\mu,\mathrm{opt}}(M)$. In addition, $\mathcal{M}^{\delta}_{\mu,\psi}$ is relatively open in $\Gamma^{\mathrm{R}}_{\mu,\mathrm{opt}}(M)$ by Lemma 2.13. Thus, there exists finitely randomized quantizer $v_F \in \mathcal{M}^{\delta}_{\mu,\psi} \cap \Gamma^{\mathrm{R}}_{\mu,\mathrm{opt}}(M)$ with $L(v_F) = L(v) = L^*$. This completes the proof of Theorem 2.5. □

2.7.7 Proof of Theorem 2.6

Let $\hat{v} \in \mathcal{M}^{\delta}_{\mu,\psi}$ be such that $L(\hat{v}) < \inf_{v \in \mathcal{M}^{\delta}_{\mu,\psi}} L(v) + \varepsilon/2$. Let $\hat{P}$ be the probability measure on $\Gamma_{\mu}(M)$ that induces $\hat{v}$, i.e., $\hat{v} = \int_{\Gamma_{\mu}(M)} v\hat{P}(dv)$. Consider a sequence of independent and identically distributed (i.i.d.) random variables $X_1, X_1, \ldots, X_n, \ldots$ defined on some probability space $(\Omega, \mathcal{F}, \gamma)$ which take values in $\big(\Gamma_{\mu}(M), \mathcal{B}(\Gamma_{\mu}(M))\big)$ and have common distribution $\hat{P}$. Then $L(X_1), L(X_2), \ldots$ are i.i.d. $\mathbb{R}$-valued random variables with distribution $\hat{P} \circ L^{-1}$. Thus we have

$$\begin{aligned}\int_{\Omega} L(X_i(\omega))\gamma(d\omega) &= \int_{\Gamma_{\mu}(M)} L(v)\hat{P}(dv) = L(\hat{v}) \\ &< \inf_{v \in M^{\delta}_{\mu,\psi}} L(v) + \frac{\varepsilon}{2}\end{aligned}$$

by assumption. The empirical measures P^{ω}_n on $\Gamma_{\mu}(M)$ corresponding to $X_1, \ldots, X_n$ are

$$P^{\omega}_n(\,\cdot\,) := \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i(\omega)}(\,\cdot\,).$$

By the strong law of large numbers

$$\frac{1}{n}\sum_{i=1}^{n} L(X_i) = \int_{\Gamma_{\mu}(M)} L(v)P^{\omega}_n(dv) \to \int_{\Gamma_{\mu}(M)} L(v)\hat{P}(dv) = L(\hat{v}) \tag{2.24}$$

γ-a.s. As a subset of $\mathcal{P}(\mathsf{X}\times\mathsf{Y})$, $\Gamma_\mu(M)$ with the Prokhorov metric is a separable metric space, and thus by [32, Theorem 11.4.1] we also have the almost sure convergence of empirical measures, i.e., $P_n^\omega \to \hat{P}$ weakly γ-a.s. Thus there exists $\hat{\omega} \in \Omega$ for which both convergence results hold. Define the sequence of finitely randomized quantizers $\{v_n\}$ by $v_n = \int_{\Gamma_\mu(M)} v P_n^{\hat{\omega}}(dv)$. By (2.24) $L(v_n) \to L(\hat{v})$ and by Lemma 2.10 in the proof of Theorem 2.3 $v_n \to \hat{v}$ weakly. Since $\mathcal{M}_{\mu,\psi}^\delta$ is a relatively open neighborhood of $\hat{v}$ in $\Gamma_\mu^{\mathrm{R}}(M)$, we can find sufficiently large n such that $v_n \in \mathcal{M}_{\mu,\psi}^\delta$ and $L(v_n) < L(\hat{v}) + \frac{\varepsilon}{2}$. Hence, for any $\varepsilon > 0$ there exists an ε-optimal finitely randomized quantizer for **(P3)**. $\square$

2.7.8 Proofs for Section 2.5

Proof of Lemma 2.2. The proof uses standard notation for information quantities [25]. Let $X^n \sim \mu^n$, $Z \sim \nu$, and $Y^n = \mathsf{q}(X^n, Z) \sim \psi^n$, where (q, ν) is an arbitrary Model 2 randomized quantizer with at most 2^{nR} levels (Z is independent of X^n). Let $D_i = E[\rho(X_i, Y_i)]$ and $D = \frac{1}{n}\sum_{i=1}^n D_i = E[\rho_n(X^n, Y^n)]$. Since $\mathsf{q}(\,\cdot\,, z)$ has at most 2^{nR} levels for each z,

$$
\begin{aligned}
nR \geq H(Y^n|Z) &\geq I(X^n; Y^n|Z) \\
&\geq I(X^n; Y^n) && (2.25) \\
&\geq \sum_{i=1}^n I(X_i; Y_i) && (2.26) \\
&\geq \sum_{i=1}^n I_m(\mu\|\psi, D_i) \\
&\geq n I_m(\mu\|\psi, D)
\end{aligned}
$$

where in the last two inequalities follow since $Y_i \sim \psi$, $i = 1, \ldots, n$ and $I_m(\mu \| \psi, D)$ is convex in D [113, Appendix A]. Inequalities (2.25) and (2.26) follow from the chain rule for mutual information (Kolmogorov's formula) [46, Corollary 7.14], which in particular implies that $I(U; V|W) \geq I(U; V)$ for general random variables U, V, and W, defined on the same probability space, such that U and W are independent. This proves that $R \geq I_m(\mu \| \psi, D)$. $\square$

Proof of Lemma 2.3. Let $U^{2^{nR}} = \big(U^n(1), \ldots, U^n(2^{nR})\big)$ which is a $n2^{nR}$-vector. Then, we can write

$$\hat{X}^n = g(X^n, U^{2^{nR}})$$

for a function g from $\mathsf{Y}^{n(2^{nR}+1)}$ to Y^n. Observe the following:

(i) For any permutation σ of $\{1, \ldots, n\}$, X^n and $X^n_\sigma = \big(X_{\sigma(1)}, \ldots, X_{\sigma(n)}\big)$ have the same distribution. The same issue is true for $U^n(i)$ and $U^n(i)_\sigma$ for all i because for any $u^n \in T_n(\psi_n)$, $u^n_\sigma \in T_n(\psi_n)$ and this mapping is a bijection on $T_n(\psi_n)$. It follows from the independence of X^n and $U^n(i)$ that (X^n, U^{nR}) and $(X^n_\sigma, U^{2^{nR}}_\sigma)$ have the same distribution, where $U^{2^{nR}}_\sigma := \big(U^n(1)_\sigma, \ldots, U^n(2^{nR})_\sigma\big)$. Thus, $g(X^n, U^{2^{nR}})$ and $g(X^n_\sigma, U^{2^{nR}}_\sigma)$ have the same distribution.

(ii) For any $x^n \in \mathsf{X}^n$ and $y^n \in \mathsf{Y}^n$, $\rho_n(x^n, y^n) = \rho_n(x^n_\sigma, y^n_\sigma)$. Thus, if g outputs $u^n(i)$ for inputs $x^n, u^n(1), \ldots, u^n(2^{nR})$, then g outputs $u^n(i)_\sigma$ for inputs $x^n_\sigma, u^n(1)_\sigma, \ldots, u^n(2^{nR})_\sigma$. It follows that

$$g(X^n_\sigma, U^{2^{nR}}_\sigma) = g(X^n, U^{2^{nR}})_\sigma.$$

Together with $i)$ this implies that $\hat{X}^n$ and $\hat{X}^n_\sigma$ have the same distribution.

Let u^n and $v^n \in T_n(\psi_n^{(n)})$ and so $u^n = v^n_\sigma$ for some permutation σ. Then *(ii)* implies

$$\Pr\{\hat{X}^n = u^n\} = \Pr\{\hat{X}^n_\sigma = u^n\}.$$

Since $\Pr\{\hat{X}^n = v^n\} = \Pr\{\hat{X}^n_\sigma = v^n_\sigma\}$ and $v^n_\sigma = u^n$, we obtain

$$\Pr\{\hat{X}^n = u^n\} = \Pr\{\hat{X}^n = v^n\}$$

proving that $\hat{X}^n$ is uniform on $T_n(\psi_n^{(n)})$. □

Proof of Lemma 2.4. By [25, Theorem 11.1.2] we have

$$\begin{aligned}\frac{1}{n}\mathcal{D}(\psi_n^{(n)}\|\psi^n) &= \frac{1}{n}\sum_{y^n \in T_n(\psi_n)} \psi_n^{(n)}(y^n)\log\frac{\psi_n^{(n)}(y^n)}{\psi^n(y^n)} \\ &= \frac{1}{n}\log\frac{2^{n(H(\psi_n)+\mathcal{D}(\psi_n\|\psi))}}{|T_n(\psi_n)|}.\end{aligned}$$

From [25, Theorem 11.1.3],

$$\frac{1}{(n+1)^{|\mathsf{X}|}}2^{nH(\psi_n)} \le |T_n(\psi_n)| \le 2^{nH(\psi_n)}$$

and thus $\frac{1}{n}\mathcal{D}(\psi_n^{(n)}\|\psi^n)$ is sandwiched between $\mathcal{D}(\psi_n\|\psi)$ and $\frac{|\mathsf{X}|}{n}\log(n+1)+\mathcal{D}(\psi_n\|\psi)$. Thus

$$\lim_{n\to\infty}\frac{1}{n}\mathcal{D}(\psi_n^{(n)}\|\psi^n) = \lim_{n\to\infty}\mathcal{D}(\psi_n\|\psi) = 0$$

where the second limit holds since X is a finite set and $\psi_n \to \psi$ in the l_1-distance. □

Proof of Lemma 2.5. Let ρ^H denote the Hamming distortion and let $\rho^H_n(x^n, y^n) =$

$(1/n)\sum_{i=1}^n \rho^H(x_i, y_i)$. Since $\rho(x,x) = 0$ for all $x \in \mathsf{X}$, we have

$$\rho_n(x^n, y^n) \le \rho_{\max}\, \rho_n^H(x^n, y^n).$$

Let $T_n^H(\psi_n^{(n)}, \psi^n)$ be the distortion of the optimal coupling between $\psi_n^{(n)}$ and ψ^n when the cost function is ρ_n^H. Then the above inequality gives

$$\hat{T}_n(\psi_n^{(n)}, \psi^n) \le \rho_{\max}\, T_n^H(\psi_n^{(n)}, \psi^n).$$

On the other hand, by Marton's inequality [67, Proposition 1]

$$T_n^H(\psi_n^{(n)}, \psi^n) \le \sqrt{\frac{1}{2n}\mathcal{D}(\psi_n^{(n)} \| \psi^n)}.$$

Combining these bounds with $\frac{1}{n}D(\psi_n^{(n)} \| \psi^n) \to 0$ (Lemma 2.4), we obtain

$$\lim_{n\to\infty} \hat{T}_n(\psi_n^{(n)}, \psi^n) = 0 \tag{2.27}$$

which is the first statement of the lemma.

Recall that $\rho(x,y) = d(x,y)^p$ for some $p > 0$, where d is a metric. Let $q = \max\{1, p\}$. If $p \ge 1$, then $\|V^n\|_p := \big(E\big[\sum_{i=1}^n |V_i|^p\big]\big)^{1/q}$ is a norm on $\mathbb{R}^n$-valued random vectors whose components have finite pth moments, and if $1 < p < 0$, we still have $\|U^n + V^n\|_p \le \|U^n\|_p + \|V^n\|_p$. Thus we can upper bound $E[\rho_n(X^n, Y^n)]$ as follows:

$$\left(E\left[\frac{1}{n}\sum_{i=1}^n \rho(X_i, Y_i)\right]\right)^{1/q} = \left(E\left[\frac{1}{n}\sum_{i=1}^n d(X_i, Y_i)^p\right]\right)^{1/q}$$

$$\leq \left(E\left[\frac{1}{n} \sum_{i=1}^{n} d(X_i, \hat{X}_i)^p \right] \right)^{1/q} + \left(E\left[\frac{1}{n} \sum_{i=1}^{n} d(\hat{X}_i, Y_i)^p \right] \right)^{1/q}$$

$$= \left(E[\rho_n(X^n, \hat{X}^n)] \right)^{1/q} + \hat{T}_n(\psi_n^{(n)}, \psi^n)^{1/q}.$$

Hence (2.10) and (2.27) imply

$$\limsup_{n\to\infty} E[\rho_n(X^n, Y^n)] \leq D$$

as claimed. □

Proof of Lemma 2.6. Let $X \sim \mu$ and $Y \sim \psi$ such that $I(X;Y)$ achieves $I_m(\mu\|\psi, D) < \infty$ at distortion level D (the existence of such pair follows from an analogous statements for rate-distortion functions [26]) . Let q_k denote the uniform quantizer on the interval $[-k, k]$ having 2^k levels, where we extend q_k to the real line by using the nearest neighborhood encoding rule. Let $X(k) = q_k(X)$ and $Y(k) = q_k(Y)$. We clearly have

$$E[(X - X(k))^2] \to 0,\ E[(Y - Y(k))^2] \to 0 \text{ as } k \to \infty. \tag{2.28}$$

Let μ_k and ψ_k denote the distributions of $X(k)$ and $Y(k)$, respectively. Then by [99, Theorem 6.9] it follows that $\hat{T}_1(\mu_k, \mu) \to 0$ and $\hat{T}_1(\psi_k, \psi) \to 0$ as $k \to \infty$ since $\mu_k \to \mu$, $\psi_k \to \psi$ weakly, and $E[X(k)^2] \to E[X^2]$, $E[Y(k)^2] \to E[Y^2]$.

By the data processing inequality, we have for all k,

$$I(X(k); Y(k)) \leq I(X;Y). \tag{2.29}$$

Also note that (2.28) implies

$$\limsup_{k\to\infty} E\big[\rho_1(X(k), Y(k))\big] = \limsup_{k\to\infty} E\big[\big(X(k) - Y(k)\big)^2\big] \le D.$$

Thus, for given $\varepsilon > 0$, if k is large we have $I_m(\mu_k \| \psi_k, D + \varepsilon) \le I_m(\mu \| \psi, D)$ as claimed. $\square$

Chapter 3

Source Coding with Limited Common Randomness

3.1 Introduction

In Section 2.5, a rate distortion theorem was obtained for stationary and memoryless sources under the assumption that the output must also be a stationary and memoryless process and common randomness (in the form of a random variable uniformly distributed on the unit interval [0,1]) is shared by the encoder and the decoder.

In this chapter, we aim to characterize the achievable rate distortion region for the same setup, where, however, the rate region measures both the coding rate and the rate of common randomness shared between the encoder and the decoder. To give a more precise definition of the problem, analogous to the communication system in Fig. 2.1, consider the communication system in Fig. 3.1. We note that in this chapter we use Model 1 of randomized quantization which is evidently more suitable here than Models 2 and 3.

The source block $X^n = (X_1, \ldots, X_n)$ consists of n independent drawings of a random variable X which takes values in a set X and has distribution μ. The stochastic encoder takes the source and the common randomness, which is available at rate R_c

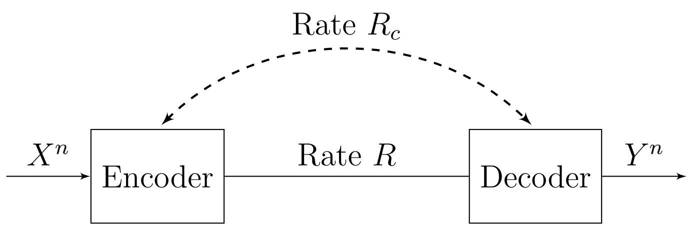

Figure 3.1: Randomized source coding with limited common randomness

bits per source symbol, as its inputs and produces an output at a rate R bits per source symbol. Observing the output of the encoder and the common randomness, the decoder (stochastically) generates the output (reconstruction) which takes values from a reproduction alphabet Y. Here $\mathsf{X} = \mathsf{Y}$ is either a finite set or the real line. The common randomness is assumed to be independent of the source. As usual, the fidelity of the reconstruction is characterized by the expected distortion

$$\mathbb{E}\biggl[\frac{1}{n}\sum_{i=1}^{n}\rho(X_i, Y_i)\biggr],$$

where $\rho : \mathsf{X} \times \mathsf{Y} \to [0, \infty)$ is a distortion measure. Analogous to the rate distortion problem in Section 2.5, we require that the output $Y^n = (Y_1, \ldots, Y_n)$ be a sequence of independent and identically distributed (i.i.d.) random variables with a given common distribution ψ.

For $D \geq 0$, a rate pair (R, R_c) is said to be *achievable* at distortion level D if, for any $\varepsilon > 0$ and all n large enough, there exists a system as in Fig. 3.1 with coding rate R and common randomness rate R_c, such that the distortion of the system is less than $D + \varepsilon$ and the output distribution constraint for Y^n holds. The main problem considered in this chapter is finding the set of all achievable rate pairs, denoted by $\mathcal{R}(D)$.

We recall that Theorem 2.7 showed for both finite and continuous source and reproduction alphabets that the minimum coding rate for unlimited common randomness at distortion D is the so-called "minimum mutual information with constrained output ψ" $I(\mu\|\psi, D)$ given in (2.9). Thus the set of achievable coding rates for unlimited common randomness $R_c = \infty$, here denoted by $\mathcal{R}(D, \infty)$, is

$$\mathcal{R}(D, \infty) = \{R \in \mathbb{R} : R \geq I(X;Y),\ P_{X,Y} \in \mathcal{G}(D)\},$$

where $\mathcal{G}(D)$ is the set of probability distributions $P_{X,Y}$ of $\mathsf{X} \times \mathsf{Y}$-valued random variables (X, Y) defined as

$$\mathcal{G}(D) \coloneqq \{P_{X,Y} : P_X = \mu, P_Y = \psi, \mathbb{E}[\rho(X, Y)] \leq D\}.$$

In this chapter, we generalize the above rate distortion result by studying the optimal tradeoff between the coding rate R and common randomness rate R_c for the system in Fig. 3.1. In particular, we find a single-letter characterization of the entire achievable rate region $\mathcal{R}(D)$ of pairs (R, R_c). Apart from the theoretical appeal of obtaining a computable characterization of the rate region via information theoretic quantities, this investigation is also motivated by the fact that the common randomness rate R_c has a direct affect on the complexity of the system since each possible value of the common randomization picks a different (stochastic) encoder and decoder pair from a finite set whose size is proportional to 2^{nR_c}. We also consider two variations of the problem, in which we investigate the effect of relaxing the strict output distribution constraint and the role of private randomness used by the decoder on the rate region. For both of these problems, we give the complete characterizations of the achievable rate pairs.

It is important to point out that the block diagram in Fig. 3.1 depicting the generalized distribution preserving quantization problem has the same structure as the system studied by Cuff [28, 29] to synthesize memoryless channels up to vanishing total variation error. Although many other problems in information theory share a similar representation, the connection with Cuff's work is more than formal. The distortion and output distribution constraints in our problem replaces the requirement in [29] that the joint distribution of the input X^n and output Y^n should arbitrarily well approximate (in total variation) the joint distribution obtained by feeding the input X^n to a given memoryless channel. Using the main result [29, Theorem II.1] one can obtain an inner bound, albeit a loose one, for our problem. A good part of our proof consists of tailoring Cuff's arguments in [29] to our setup to obtain a tight achievable rate region. We also note that unlike in the distributed channel synthesis problem in [29], our results also allow for continuous source and reproduction alphabets.

In the remainder of this chapter, we suppose that Assumption 2.3 in Section 2.5 holds.

3.2 Rate Distortion Region

Let $\{X_n\}_{n\geq 1}$ be a stationary and memoryless source (sequence of i.i.d. random variables) with common distribution μ on source alphabet X, and let K be a random variable uniformly distributed over $[2^{nR_c}]$ which is independent of X^n. Here K represents the common randomness that is shared between the encoder and the decoder.

In the remainder of this chapter, any randomization that is separately used by the encoder and/or the decoder, and is independent of the source $\{X_n\}_{n\geq 1}$ and the common randomness K is called *private randomization*.

In this setting since common randomness, shared between the encoder and the decoder, is rate limited, the encoder and the decoder are allowed to use private randomization. Namely, for a positive integer n and nonnegative numbers R and R_c, a (n, R, R_c) *randomized source code* is defined by an encoder $E = E_{J|X^n,K}$ and the decoder $F_{Y^n|J,K}$, where E is a regular conditional probability (see [32]) on $[2^{nR}]$ given $\mathsf{X}^n \times [2^{nR_c}]$ and F is a regular conditional probability on Y^n given $[2^{nR}] \times [2^{nR_c}]$. Hence, letting J and Y^n be the output of the encoder and the decoder, respectively, the joint distribution of (K, X^n, J, Y^n) is given, in a somewhat informal notation, by

$$(K, X^n, J, Y^n) \sim F_{Y^n|J,K} E_{J|X^n,K} P_K P_{X^n}. \tag{3.1}$$

The distortion of the code is $\mathbb{E}[\rho_n(X^n, Y^n)]$, where $\rho_n(x^n, y^n) := \frac{1}{n}\sum_{i=1}^{n} \rho(x_i, y_i)$.

Remark 3.1. If K_1 and K_2 are uniform random variables on $[0, 1]$ such that they are independent of each other and the pair (X^n, K), then a (n, R, R_c) randomized source code can be realized as $J = e_n(X^n, K_1, K)$ and $Y^n = d_n(J, K_2, K)$, where e_n and d_n are suitable (measurable) functions. Hence, (d_n, g_n, ν) is a randomized quantizer of Model 1 where $Z = (K_1, K_2, K) \sim \nu$.

Definition 3.1. *For any nonnegative real number D and desired output distribution ψ, the pair (R, R_c) is said to be ψ-*achievable *if, for any $\varepsilon > 0$ and all sufficiently large n, there exists a randomized (n, R, R_c) source code such that*

$$\mathbb{E}[\rho_n(X^n, Y^n)] \leq D + \varepsilon$$

$$Y^n \sim \psi^n.$$

In the rest of this chapter ψ will be kept fixed, so we drop referring to ψ and simply write that (R, R_c) is *achievable*. For $D \geq 0$ we let $\mathcal{R}(D)$ denote the set of all achievable (R, R_c) pairs. The following theorem, which is the main result in this chapter, characterizes the closure of this region in terms of an auxiliary random variable U on alphabet U.

Theorem 3.1. *For any $D \geq 0$ the closure $\mathrm{cl}\,\mathcal{R}(D)$ of $\mathcal{R}(D)$ is given by*

$$\begin{aligned} \mathrm{cl}\,\mathcal{R}(D) &= \mathcal{L}(D) \\ &\coloneqq \left\{ \begin{array}{rcl} (R, R_c) \in \mathbb{R}^2 & : & \exists P_{X,Y,U} \in \mathcal{M}(D) \ s.t. \\ R & \geq & I(X;U), \\ R + R_c & \geq & I(Y;U) \end{array} \right\}, \end{aligned} \tag{3.2}$$

where, for $\mathsf{X} = \mathsf{Y}$ finite,

$$\mathcal{M}(D) \coloneqq \left\{ \begin{array}{ll} P_{X,Y,U} : & P_X = \mu, P_Y = \psi, \\ & \mathbb{E}[\rho(X,Y)] \leq D, \ X - U - Y, \\ & |\mathsf{U}| \leq |\mathsf{X}| + |\mathsf{Y}| + 1 \end{array} \right\}. \tag{3.3}$$

When $\mathsf{X} = \mathsf{Y} = \mathbb{R}$, the cardinality bound for U in (3.3) *is replaced by $\mathsf{U} = \mathbb{R}$.*

3.2.1 Connections with Distributed Channel Synthesis

As mentioned before, Cuff's work on distributed channel synthesis [29] is intrinsically related to above problem. The main objective of [29] is to simulate a memoryless channel by a system as in Fig. 3.1. To be more precise, let $Q(y|x)$ denote a given discrete memoryless channel with input alphabet X and output alphabet Y to be simulated (synthesized) for input X having distribution μ. Let $\pi = \mu Q$ be the joint distribution of the resulting input-output pair (X, Y).

Definition 3.2 ([29]). *The pair (R, R_c) is said to be* achievable *for synthesizing a memoryless channel Q with input distribution μ if there exists a sequence of (n, R, R_c) randomized source codes such that*

$$\lim_{n\to\infty} \|P_{X^n,Y^n} - \pi^n\|_{TV} = 0, \tag{3.4}$$

where $X^n \sim \mu^n$ is the memoryless source, Y^n is the output of the decoder, π^n is the n-fold product of $\pi = \mu Q = P_X Q$, and $\|\cdot\|_{TV}$ is the total variation distance for probability measures: $\|\gamma - \nu\|_{TV} := \frac{1}{2}\sum_v |\gamma(v) - \nu(v)|$.

Theorem 3.2. *[29, Theorem II.1] The closure $\mathcal{C}$ of the set of all achievable (R, R_c) pairs is given by*

$$\mathcal{C} = \mathcal{S} := \left\{ \begin{array}{rcl} (R, R_c) \in \mathbb{R}^2 & : & \exists P_{X,Y,U} \in \mathcal{D} \text{ s.t.} \\ R & \geq & I(X;U), \\ R + R_c & \geq & I(X,Y;U) \end{array} \right\}, \tag{3.5}$$

where

$$\mathcal{D} := \{P_{X,Y,U} : P_{X,Y} = \pi, X - U - Y, |\mathsf{U}| \leq |\mathsf{X}||\mathsf{Y}| + 1\}.$$

Moreover, the total variation error goes to zero exponentially fast with respect to n in the interior of $\mathcal{C}$.

This result can be used to obtain an achievable rate region (inner bound) for our problem as follows: Let $\pi = P_{X,Y}$ be such that $P_X = \mu$, $P_Y = \psi$, and $\mathbb{E}[\rho(X, Y] \leq D$.

Applying Theorem 3.2 with this input distribution and the channel induced by $P_{X,Y}$, consider an achievable rate pair (R, R_c) in (3.5). Using basic results from optimal transport theory [99] one can show that (3.4) and the fact that $\mathbb{E}[\rho(X,Y)] \le D$ imply the existence of a sequence of channels, to be used at the decoder side, that when fed with Y^n, produces output $\hat{Y}^n$ which has the exact distribution ψ^n and which additionally satisfies

$$\limsup_{n\to\infty} \mathbb{E}[\rho_n(X^n, \hat{Y}^n)] \le D.$$

Augmenting the channel synthesis code with these channels at the decoder side thus produces a sequence of valid codes for our problem, implying that the rate pair (R, R_c) is achievable by our Definition 3.1.

Using the above argument, one can easily show that Cuff's result directly implies (without resorting to Theorem 3.1) the following inner bound for $\mathcal{R}(D)$. The proof is given in Section 3.8.1.

Corollary 3.1. *For any $D \ge 0$,*

$$\operatorname{cl}\mathcal{R}(D) \supset \mathcal{S}(D) := \left\{ \begin{array}{rcl} (R, R_c) \in \mathbb{R}^2 & : & \exists P_{X,Y,U} \in \mathcal{H}(D) \ s.t. \\ R & \ge & I(X;U), \\ R + R_c & \ge & I(X,Y;U) \end{array} \right\}, \tag{3.6}$$

where

$$\mathcal{H}(D) := \left\{ \begin{array}{ll} P_{X,Y,U} : & P_X = \mu, P_Y = \psi, \\ & \mathbb{E}[\rho(X,Y)] \le D, X - U - Y, \\ & |\mathsf{U}| \le |\mathsf{X}||\mathsf{Y}| + 1 \end{array} \right\}. \tag{3.7}$$

In general, this inner bound is loose. For example, for $R_c = 0$, only the constraint

$R \geq I(X,Y;U)$ is active in (3.6) since $I(X,Y;U) \geq I(X;U)$ always holds. Hence, letting $\mathcal{S}(D,0)$ denote the set of Rs such that $(R,0) \in \mathcal{S}(D)$, we obtain

$$\mathcal{S}(D,0) = \{R \in \mathbb{R} : \exists P_{X,Y,U} \in \mathcal{H}(D) \text{ s.t. } R \geq I(X,Y;U)\}.$$

The minimum of $\mathcal{S}(D,0)$ can be written as

$$\min\{R \in \mathcal{S}(D,0)\} = \min\{C(X;Y) : P_{X,Y} \in \mathcal{G}(D)\} =: C_0(\mu\|\psi, D),$$

where $C(X;Y)$ is Wyner's common information [104] defined for a given joint distribution $P_{X,Y}$ by

$$C(X;Y) \coloneqq \inf_{U:X-U-Y} I(X,Y;U), \tag{3.8}$$

where the infimum is taken over all joint distributions $P_{X,Y,U}$ such that U has a finite alphabet and $X - U - Y$. However, the resulting rate $C_0(\mu\|\psi, D)$ is not optimal as Example 3.1 in Section 3.3.2 will show.

The suboptimality of $C_0(\mu\|\psi, D)$ implies that a 'separated' solution which first finds an 'optimal' channel and then synthesizes this channel is not optimal for the constrained rate distortion problem we consider.

3.3 Special Cases

The extreme points at $R_c = \infty$ and $R_c = 0$ of the rate region $\mathcal{L}(D)$ in our Theorem 3.1 are of particular interest. Let $\mathcal{L}(D, R_c)$ be the set of coding rates R such that $(R, R_c) \in \mathcal{L}(D)$.

3.3.1 Unlimited Common Randomness

This was the situation studied in Section 2.5 where it was assumed that the common randomness is of the form of a real-valued random variable that is uniformly distributed on the interval $[0, 1]$. If $R_c = \infty$, then the effective constraint in (3.2) is $R \geq I(X; U)$. Since $I(X; U) \geq I(X; Y)$ by the data processing inequality and the condition $X - Y - Y$, we can set $U = Y$ to obtain $\min\{R \in \mathcal{L}(D, \infty)\} = I(\mu \| \psi, D)$, recovering Theorem 2.7 (see also (2.9)). Furthermore, for the finite alphabet case whenever $R_c \geq H(Y|X)$, we have from (3.2) that $R + R_c \geq I(X; U) + H(Y|X) \geq I(X; Y) + H(Y|X) = H(Y) \geq I(Y; U)$, so the effective constraint is again $R \geq I(X; U)$. Considering (X, Y) such that $P_{X,Y}$ achieves the minimum in (2.9) and letting $U = Y$, we have

$$\min\{R \in \mathcal{L}(D, R_c)\} = I(\mu \| \psi, D) \tag{3.9}$$

or equivalently

$$\mathcal{L}(D, R_c) = \mathcal{L}(D, \infty). \tag{3.10}$$

Hence, $H(Y|X)$ is a sufficient common randomness rate above which the minimum communication rate does not decrease. In fact, letting

$$R_c^{\min} = \min\{R_c : \mathcal{L}(D, R_c) = \mathcal{L}(D, \infty)\}$$

we can determine $R_c^{\min}$ in terms of the so-called *necessary conditional entropy* [29], defined for a joint distribution $P_{X,Y}$ as

$$H(Y\dagger X) \coloneqq \min_{f:X-f(Y)-Y} H(f(Y)|X)$$

where minimum is taken over all functions $f : \mathsf{Y} \to \mathsf{Y}$ such that $X - f(Y) - Y$. Using the discussion in [30, Section VII-C] one can verify that $R_c^{\min}$ is the minimum of $H(Y\dagger X)$ over all joint distribution of (X, Y) achieving the minimum in (2.9).

3.3.2 No Common Randomness

Setting $R_c = 0$ means that no common randomness is available. In this case (3.2) gives $R \geq \max\bigl(I(X;U), I(Y;U)\bigr)$. Hence the minimum communication rate at distortion D is given by

$$\min\{R \in \mathcal{L}(D, 0)\} = I_0(\mu\|\psi, D),$$

where

$$I_0(\mu\|\psi, D) \coloneqq \min\bigl\{\max\bigl(I(X;U), I(Y;U)\bigr) : P_{X,Y,U} \in \mathcal{M}(D)\bigr\}. \tag{3.11}$$

Note that the minimum achievable coding rate $I_0(\mu\|\psi, D)$ is *symmetric* with respect to μ and ψ, i.e., $I_0(\mu\|\psi, D) = I_0(\psi\|\mu, D)$. This is clear from the definition (3.11), but can also be deduced from the operational meaning of $I_0(\mu\|\psi, D)$ since in the absence of the common randomness K, the encoder-decoder structure is fully reversible. In general such symmetry no longer holds for $\min\{R \in \mathcal{R}(D, R_c)\}$ when $R_c > 0$.

The following lemma states that $I_0(\mu\|\psi, D)$ is convex in D. The proof simply follows from a time-sharing argument and the operational meaning of $I_0(\mu\|\psi, D)$ implied by Theorem 3.1. It is given in the Section 3.8.2.

Lemma 3.1. *$I_0(\mu\|\psi, D)$ is a convex function of D.*

An upper bound for $I_0(\mu\|\psi, D)$ can be given in terms of Wyner's common information. Since $\max\big(I(X;U), I(Y;U)\big) \le I(X,Y;U)$, we have $I_0(\mu\|\psi, D) \le \min\{I(X,Y;U) : P_{X,Y,U} \in \mathcal{M}(D)\}$. The latter expression can also be written as

$$\min\{C(X;Y) : P_{X,Y} \in \mathcal{G}(D)\} =: C_0(\mu\|\psi, D). \tag{3.12}$$

However, the resulting upper bound $I_0(\mu\|\psi, D) \le C_0(\mu\|\psi, D)$ is not tight in general as the next example shows.

Example 3.1. Let $\mathsf{X} = \mathsf{Y} = \{0,1\}$, and let $\mu = \psi = \mathsf{Bernoulli}(1/2)$, i.e., $\mu(0) = \mu(1) = \frac{1}{2}$. Assume the distortion measure ρ is the Hamming distance $\rho(x,y) = 1_{\{x \neq y\}}$. If $X \sim \mu$ and $Y \sim \psi$, then the channel $P_{Y|X}$ from X to Y must be Binary Symmetric Channel (BSC) with some crossover probability a_0, i.e.,

$$P_{Y|X}(\,\cdot\,|0) = 1 - P_{Y|X}(\,\cdot\,|1) = \mathsf{Bernoulli}(a_0).$$

Wyner in [104, Section 3] showed that when $a_0 \in [0, 1/2]$,

$$C(X;Y) = 1 + h(a_0) - 2h(a_1),$$

where $a_1 = \frac{1}{2}(1 - \sqrt{1 - 2a_0})$, and $h(\lambda) = -\lambda\log(\lambda) - (1-\lambda)\log(1-\lambda)$. Define $C(a_0) := 1 + h(a_0) - 2h(a_1)$ which is decreasing and strictly concave in $[0, 1/2]$. Notice that $\mathbb{E}[\rho(X,Y)] = a_0$ when $P_{Y|X} = \mathsf{BSC}(a_0)$. Hence, for any $D \in [0, 1/2]$, we have

$$
\begin{aligned}
&C_0(\mu\|\psi, D)\\
&= \min\{C(X;Y) : P_{X,Y} \in \mathcal{G}(D)\}\\
&= \min\{C(X;Y) : P_X = \mu, P_{Y|X} = \mathsf{BSC}(a_0), a_0 \le D\}\\
&= \min_{a_0 \le D} C(a_0) = C(D)
\end{aligned}
$$

implying that $C_0(\mu\|\psi, D)$ is strictly concave for $D \in [0, 1/2]$. This, together with Lemma 3.1 and the easy-to-check facts that $C_0(\mu\|\psi, 0) = I_0(\mu\|\psi, 0) = 1$ and $C_0(\mu\|\psi, 1/2) = I_0(\mu\|\psi, 1/2) = 0$ implies that

$$
I_0(\mu\|\psi, D) < C_0(\mu\|\psi, D), \quad D \in (0, 1/2).
$$

3.4 Examples

In general determining the entire rate region $\mathcal{L}(D)$ in Theorem 3.1 seems to be difficult even for simple cases. In this section we obtain possibly suboptimal achievable rate regions (inner bounds) for two setups by restricting the channels $P_{U|X}$ and $P_{Y|U}$ so that the resulting optimization problem becomes manageable.

3.4.1 Doubly Symmetric Binary Source

In this section we obtain an inner bound for the setup in Example 3.1 (i.e., when $\mathsf{X} = \mathsf{Y} = \{0, 1\}$, $\mu = \psi = \mathsf{Bernoulli}(1/2)$, and ρ the Hamming distance) by restricting the auxiliary random variable U to be $\mathsf{Bernoulli}(1/2)$. Since $P_X = P_U = P_Y = \mathsf{Bernoulli}(1/2)$, for any $P_{X,Y,U} \in \mathcal{M}_s(D)$, the channels $P_{U|X}$ and $P_{Y|U}$ must be $\mathsf{BSC}(a_1)$

and $\mathsf{BSC}(a_2)$, respectively, for some $a_1, a_2 \in [0,1]$. Hence, since $\mathbb{E}[\rho(X,Y)] = a$ when $P_{X|Y} = \mathsf{BSC}(a)$, the resulting achievable rate region is

$$\mathcal{L}_s(D) = \left\{ \begin{array}{rcl} (R, R_c) \in \mathbb{R}^2 & : & (a_1, a_2) \in \Phi(D) \text{ s.t.} \\ R & \geq & 1 - h(a_1), \\ R + R_c & \geq & 1 - h(a_2). \end{array} \right\},$$

where

$$\Phi(D) \coloneqq \{(a_1, a_2) \in [0,1]^2 : a_1 + a_2 - 2a_1a_2 \leq D\}.$$

Let us define $\varphi(a_1, a_2) = a_1 + a_2 - 2a_1a_2$. Note that since $\varphi(\frac{1}{2}+r, \frac{1}{2}+m) = \frac{1}{2} - 2rm$ and $h(\frac{1}{2}-r) = h(\frac{1}{2}+r)$ for any $r, m \in [\frac{-1}{2}, \frac{1}{2}]$; we may assume without loss of generality that $a_1, a_2 \in [0, \frac{1}{2}]$ in the definition of $\Phi(D)$. Furthermore, since $\varphi(a_1, a_2) > D$ when $D < a_1 < \frac{1}{2}$ or $D < a_2 < \frac{1}{2}$, we can refine the definition of $\mathcal{L}_s(D)$ for $0 \leq D < \frac{1}{2}$ as

$$\mathcal{L}_s(D) = \left\{ \begin{array}{rcl} (R, R_c) \in \mathbb{R}^2 & : & (a_1, a_2) \in \Phi_r(D) \text{ s.t.} \\ R & \geq & 1 - h(a_1), \\ R + R_c & \geq & 1 - h(a_2). \end{array} \right\},$$

where

$$\Phi_r(D) \coloneqq \{(a_1, a_2) \in [0,D]^2 : a_1 + a_2 - 2a_1a_2 \leq D\}.$$

Notice that for any fixed a_1, $(a_1, a_2) \in \Phi_r(D)$ if and only if $a_2 \leq \frac{D-a_1}{1-2a_1}$, where the expression on the righthand side of the inequality is a concave function of a_1. Hence, $\Phi_r(D)$ is a convex region. In the remainder of this section we characterize the boundary $\bigcup_{R_c} \min\{R : (R, R_c) \in \mathcal{L}_s(D)\} \times \{R_c\}$ of $\mathcal{L}_s(D)$.

If $R_c = \infty$, then $(R, \infty) \in \mathcal{L}_s(D) \Leftrightarrow R \geq 1 - h(a_1)$ where $a_1 \in [0, D]$. Hence, the minimum R is equal to $1 - h(D)$ for $R_c = \infty$. Moreover, if $R = 1 - h(D)$ or equivalently $a_1 = D$, then $(R, R_c) \in \mathcal{L}_s(D) \Leftrightarrow R_c + 1 - h(D) \geq 1 - h(a_2) = 1 - h(0) = 1$ since $(D, a_2) \in \Phi_r(D)$ only if $a_2 = 0$. Hence, if $R_c \geq h(D)$, then

$$\min\{R : (R, R_c) \in \mathcal{L}_s(D)\} = 1 - h(D).$$

Recall that for an arbitrary $0 \leq R_c < h(D)$, $(R, R_c) \in \mathcal{L}_s(D) \Leftrightarrow R \geq \max\{1 - h(a_1), 1 - h(a_2) - R_c\}$ where $(a_1, a_2) \in \Phi_r(D)$. We now prove that the minimum R is attained when $1 - h(a_1) = 1 - h(a_2) - R_c$ and $a_1 + a_2 - 2a_1 a_2 = D$. The second equality is clear since the binary entropy function h is increasing in $[0, D]$. To prove the first claim by contradiction, let us assume (without loss of generality) that the minimum is achieved when $1 - h(a_1) > 1 - h(a_2) - R_c$ $\big($so $\min\{R : (R, R_c) \in \mathcal{L}_s(D)\} = 1 - h(a_1)\big)$. Since h is increasing and continuous and $\Phi_r(D)$ is a convex region in the lower-left corner of the square $[0, D]^2$ having nonempty interior, there exist $\varepsilon_1, \varepsilon_2 > 0$ such that $(a_1 + \varepsilon_1, a_2 - \varepsilon_2) \in \Phi_r(D)$ and $1 - h(a_1 + \varepsilon_1) \geq 1 - h(a_2 - \varepsilon_2) - R_c$. But $\min\{R : (R, R_c) \in \mathcal{L}_s(D)\} = 1 - h(a_1) > 1 - h(a_1 + \varepsilon_1)$, which is a contradiction.

Hence, for all $D \in (0, \frac{1}{2})$ the minimum coding rate when $0 \leq R_c < h(D)$ is given by

$$\min\{R : (R, R_c) \in \mathcal{L}_s(D)\} = \min\{1 - h(a_1) : (a_1, a_2) \in \Pi(D, R_c)\}$$

where

$$\Pi(D, R_c) := \left\{ \begin{array}{r} (a_1, a_2) \in \Phi_r(D) : 1 - h(a_1) = 1 - h(a_2) - R_c \\ \text{and } a_1 + a_2 - 2a_1a_2 = D \end{array} \right\}.$$

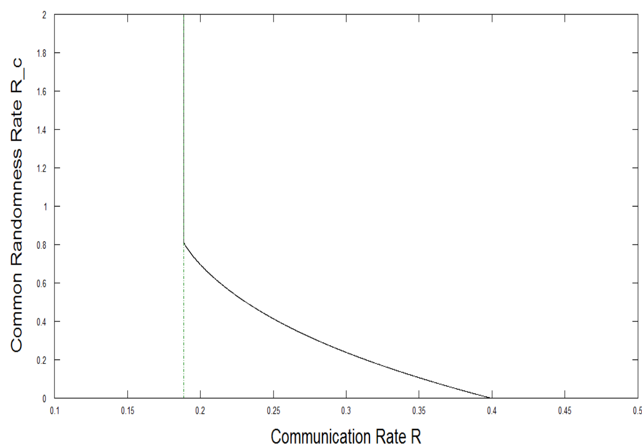

Figure 3.2: $\mathcal{L}_s(D)$ for binary symmetric source for $D = 0.25$

Figure 3.2 shows the rate region $\mathcal{L}_s(D)$ for $D = 0.25$. At the boundary of $\mathcal{L}_s(D)$, the coding rate R ranges from $1 - h(a^*) = 0.39$ bits $\big(a^* = \frac{1}{2}(1 - \sqrt{1 - 2D}) = 0.15\big)$ to $h(D) = 0.19$ bits while the common randomness rate R_c ranges from 0 to $1 - h(D) = 0.81$ bits.

3.4.2 Gaussian Source

Let $\mathrm{N}(m, \sigma)$ denote a Gaussian random variable with mean m and variance σ^2 (similar notation will be used for the vector case). In this section, we obtain an inner bound for the case $\mathsf{X} = \mathsf{Y} = \mathbb{R}$, $\mu = \mathrm{N}(0, \sigma_X)$, $\psi = \mathrm{N}(0, \sigma_Y)$, and ρ is the squared error distortion (i.e., $\rho(x, y) = |x - y|^2$) by restricting (X, U, Y) to be Gaussian $\big($or, equivalently, restricting (X, U) and (U, Y) to be Gaussian since $X - U - Y\big)$.

Remark 3.2. Recall that for $R_c = \infty$, the minimum coding rate is given by (2.9). However if $X \sim \mathrm{N}(0, \sigma_X)$ and $Y \sim \mathrm{N}(0, \sigma_Y)$, then for any $P_{X,Y} \in \mathcal{G}(D)$, one has the lower bound

$$\begin{aligned} I(X;Y) &= h(X) + h(Y) - h(X,Y) \\ &\geq \frac{1}{2}\log(2\pi e \sigma_X^2) + \frac{1}{2}\log(2\pi e \sigma_Y^2) - \log(2\pi e \det(C)^{\frac{1}{2}}), \end{aligned}$$

where C is the covariance matrix of (X, Y). The equality is achieved when (X, Y) is jointly Gaussian [25, Theorem 8.6.5]. Hence, we can restrict (X, Y) to be Gaussian in the definition of $I(\mu\|\psi, D)$, i.e.,

$$I(\mu\|\psi, D) \coloneqq \min\{I(X, Y) : P_{X,Y} \in \mathcal{G}_g(D)\},$$

where

$$\mathcal{G}_g(D) \coloneqq \{P_{X,Y} \in \mathcal{G}(D) : P_{X,Y} = \mathrm{N}(0, C) \text{ for some } C\}.$$

This implies that the inner bound we obtain in this section is tight for $R_c = \infty$ $\big($i.e., $\mathcal{L}_s(D, \infty) = \mathcal{L}(D, \infty)\big)$. $\mathcal{L}(D, \infty)$ for the case $\mu = \psi = \mathrm{N}(0, \sigma)$ was derived in [65, Proposition 2].

Note that without any loss of generality we can take U to have zero mean and unit variance. Indeed, let $\tilde{U} = (U - \delta_U)/\sigma_U$. Then $\tilde{U} \sim \mathrm{N}(0, 1)$, $X - \tilde{U} - Y$, and $(X, \tilde{U}, Y)$ is Gaussian with $I(X; U) = I(X; \tilde{U})$ and $I(Y; U) = I(Y; \tilde{U})$. Hence, in the remainder of this section, we assume $U \sim \mathrm{N}(0, 1)$.

Let us write $U = aX + V$ and $Y = bU + W$, where $a, b \in \mathbb{R}$, and $V \sim \mathrm{N}(0, \sigma_V)$, $W \sim$

$N(0, \sigma_W)$, and (X, V, W) are independent. With this representation, the constraints in the definition of the achievable rate region become

$$
\begin{aligned}
1 &= a^2\sigma_X^2 + \sigma_V^2, \\
\sigma_Y^2 &= b^2 + \sigma_W^2, \\
(1-ab)^2\sigma_X^2 + b^2\sigma_V^2 + \sigma_W^2 &\leq D,
\end{aligned}
$$

Then, if we substitute $\sigma_V^2 = 1 - a^2\sigma_X^2 \geq 0$ and $\sigma_W^2 = \sigma_Y^2 - b^2 \geq 0$ into the last equation, we can write the distortion constraint as

$$
\sigma_X^2 + \sigma_Y^2 - 2ab\sigma_X^2 \leq D.
$$

Since

$$
\begin{aligned}
I(X;U) &= H(X) + H(U) - H(X,U) \\
&= \frac{1}{2}\log(2\pi\,\mathrm{e}\,\sigma_X^2) + \frac{1}{2}\log(2\pi\,\mathrm{e}) - \log(2\pi\,\mathrm{e}\det(C_X)^{\frac{1}{2}}) \\
&= \frac{1}{2}\log\Big(\frac{1}{(1-a^2\sigma_X^2)}\Big)
\end{aligned}
$$

and

$$
\begin{aligned}
I(Y;U) &= H(Y) + H(U) - H(Y,U) \\
&= \frac{1}{2}\log(2\pi\,\mathrm{e}\,\sigma_Y^2) + \frac{1}{2}\log(2\pi\,\mathrm{e}) - \log(2\pi\,\mathrm{e}\det(C_Y)^{\frac{1}{2}}) \\
&= \frac{1}{2}\log\Big(\frac{\sigma_Y^2}{(\sigma_Y^2 - b^2)}\Big),
\end{aligned}
$$

where C_X is the covariance matrix of (X, U) and C_Y is the covariance matrix of

(Y, U), the resulting achievable rate region can be written as

$$\mathcal{L}_s(D) = \left\{ \begin{array}{rcl} (R, R_c) \in \mathbb{R}^2 & : & (a,b) \in \Psi(D) \text{ s.t.} \\ R & \geq & \frac{1}{2}\log\big(\frac{1}{(1-a^2\sigma_X^2)}\big), \\ R + R_c & \geq & \frac{1}{2}\log\big(\frac{\sigma_Y^2}{(\sigma_Y^2 - b^2)}\big). \end{array} \right\},$$

where

$$\Psi(D) \coloneqq \{(a,b) \in [0, \sigma_X^{-1}] \times [0, \sigma_Y] : \sigma_X^2 + \sigma_Y^2 - 2ab\sigma_X^2 \leq D\}.$$

Note that the region $\Psi(D)$ is convex. Let us define $I_1(a) = \log\big(\frac{1}{(1-a^2\sigma_X^2)}\big)$ and $I_2(b) = \log\big(\frac{\sigma_Y^2}{(\sigma_Y^2-b^2)}\big)$; then I_1 and I_2 are increasing functions. As in Section 3.4.1, we characterize the boundary $\bigcup_{R_c} \min\{R : (R, R_c) \in \mathcal{L}_s(D)\} \times \{R_c\}$ of $\mathcal{L}_s(D)$.

If $R_c = \infty$, then $(R, \infty) \in \mathcal{L}_s(D) \Leftrightarrow R \geq I_1(a)$ where $(a, b) \in [0, \sigma_X^{-1}] \times [0, \sigma_Y]$ and $\sigma_X^2 + \sigma_Y^2 - 2ab\sigma_X^2 \leq D$. Using the monotonicity of I_1 and the distortion constraint, it is straightforward to show that

$$\min\{R : (R, \infty) \in \mathcal{L}_s(D)\} = I_1\big(\frac{\sigma_X^2 + \sigma_Y^2 - D}{2\sigma_X^2\sigma_Y}\big).$$

By Remark 3.2, this is the minimum coding rate (i.e., rate-distortion function) for $R_c = \infty$.

When $0 \leq R_c < \infty$ is arbitrary, we can use the same technique as in Section 3.4.1 to prove that the minimum of R is attained when $I_1(a) = I_2(b) - R_c$ and $\sigma_X^2 + \sigma_Y^2 - 2ab\sigma_X^2 = D$ (I_1 and I_2 are increasing continuous functions and $\Psi(D)$ is a convex region with nonempty interior in the upper-right corner of the rectangle $[0, \sigma_X^{-1}] \times [0, \sigma_Y]$). As a consequence, we can describe the minimum coding rate when $0 \leq R_c < \infty$ as follows:

$$\min\{R : (R, R_c) \in \mathcal{L}_s(D)\} = \min\{I_1(a) : (a, b) \in \Lambda(D, R_c)\}$$

where

$$\Lambda(D, R_c) \coloneqq \left\{ \begin{array}{rl} (a, b) \in \Psi(D) : & I_1(a) = I_2(b) - R_c \text{ and} \\ & \sigma_X^2 + \sigma_Y^2 - 2ab\sigma_X^2 = D \end{array} \right\}.$$

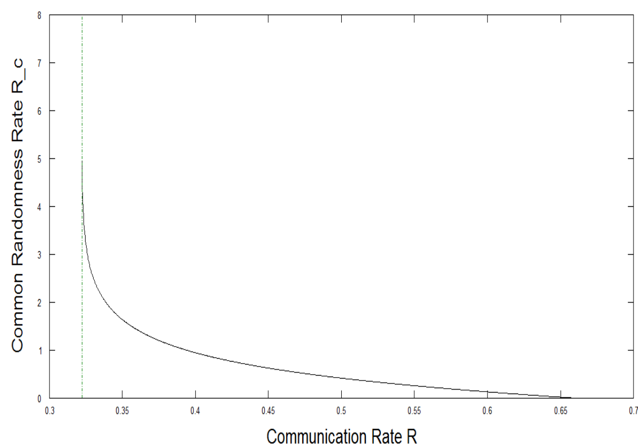

Figure 3.3: $\mathcal{L}_s(D)$ for Gaussian source for $D = 0.8$

Figure 3.3 shows the rate region $\mathcal{L}_s(D)$ for $\sigma_X = \sigma_Y = 1$ and $D = 0.8$. At the boundary of $\mathcal{L}_s(D)$, the coding rate R ranges from $I_1(\sqrt{\frac{2-D}{2}}) = 0.65$ bits to $I_1(\frac{2-D}{2}) = 0.32$ bits while the common randomness rate R_c ranges from 0 to infinity.

3.5 Two Variations

In this section we consider two variations of the rate-distortion problem defined in Section 3.2. Throughout this section we assume that the source alphabet X and the reproduction alphabet Y are finite.

3.5.1 Rate Region with Empirical Distribution Constraint

First, we investigate the effect on the achievable rate region of relaxing the strict output distribution constraint on Y^n and requiring only that the empirical output distribution p_{Y^n} converges to the distribution ψ.

Definition 3.3. *For any positive real number D and desired output distribution ψ, the pair (R, R_c) is said to be* empirically achievable *if there exists a sequence of (n, R, R_c) randomized source codes such that*

$$\limsup_{n\to\infty} \mathbb{E}[\rho_n(X^n, Y^n)] \le D,$$

$$\|p_{Y^n} - \psi\|_{TV} \to 0 \text{ in probability as } n \to \infty.$$

For any $D \ge 0$ we let $\mathcal{R}_e(D)$ denote the set of all empirically achievable rate pairs (R, R_c), and define $\mathcal{R}_e(D, R_c)$ as the set of coding rates R such that $(R, R_c) \in \mathcal{R}_e(D)$.

This setup is motivated by the work of Cuff *et. al.* [30, Section II] on empirical coordination. The main objective of [30, Section II] is to empirically simulate a memoryless channel by a system as in Fig. 3.1. To be more precise, let $Q(y|x)$ denote a given discrete memoryless channel with input alphabet X and output alphabet Y to be simulated (synthesized) for input X having distribution μ. Let $\pi = \mu Q$ be the joint distribution of the resulting input-output pair (X, Y).

Definition 3.4. *The pair (R, R_c) is said to be* achievable *for empirically synthesizing a memoryless channel Q with input distribution μ if there exists a sequence of*

(n, R, R_c) *randomized source codes such that*

$$\lim_{n\to\infty} \|p_{X^n,Y^n} - \pi\|_{TV} = 0 \text{ in probability.} \tag{3.13}$$

Let $\mathcal{C}_e$ denote the the set of all achievable (R, R_c) pairs and let $\mathcal{C}_e(R_c)$ denote the set of all rates R such that $(R, R_c) \in \mathcal{C}_e$. The following theorem, which is a combination of [30, Theorems 2 and 3], characterizes the entire set $\mathcal{C}_e$.

Theorem 3.3. *The set* $\mathcal{C}_e$ *of all achievable* (R, R_c) *is given by*

$$\mathcal{C}_e = \left\{ \begin{array}{rcl} (R, R_c) \in \mathbb{R}^2 & : & \exists P_{X,Y} \in \mathcal{G} \text{ s.t.} \\ R & \geq & I(X;Y) \end{array} \right\},$$

where

$$\mathcal{G} := \{P_{X,Y} : P_{X,Y} = \pi\}.$$

Hence, $\mathcal{C}_e(R_c) = \mathcal{C}_e(0)$ *for any* R_c.

Using the above theorem and the arguments in [30, Section VII], one can show that the set of empirically achievable rate pairs (R, R_c) at the distortion level D can be described as:

Theorem 3.4. *For any* $D \geq 0$ *we have*

$$\begin{aligned} \mathcal{R}_e(D, 0) &= \mathcal{L}(D, \infty), \\ \mathcal{R}_e(D, R_c) &= \mathcal{R}_e(D, 0) \text{ for all } R_c. \end{aligned} \tag{3.14}$$

In other words, $\mathcal{R}_e(D) = \mathcal{L}(D, \infty) \times [0, \infty)$.

The proof of Theorem 3.4 is given in the Section 3.8.3. Note that (3.14) states that unlike in the original problem defined in Section 3.2, here common randomness cannot decrease the necessary coding rate.

3.5.2 Feedback Rate Region

In this section we investigate the effect on the rate region of private randomness used by the decoder. Namely, we determine the achievable rate region for a randomized source code having no (private) randomness at the decoder, i.e., when the decoder F is a deterministic function of random variables J and K. In this setup, since the encoder can reconstruct the output Y^n of the decoder by reading off J and K, the common randomness K may be interpreted as feedback from the output of the decoder to the encoder [103, p. 5]. Hence, we call such a code a *randomized source code with feedback.*

Definition 3.5. *For any positive real number* D *and desired output distribution* ψ*, the pair* (R, R_c) *is said to be* feedback achievable *if there exists a sequence of* (n, R, R_c) *randomized source codes with feedback such that*

$$\limsup_{n\to\infty} \mathbb{E}[\rho_n(X^n, Y^n)] \leq D,$$
$$\lim_{n\to\infty} \|P_{Y^n} - \psi^n\|_{TV} = 0. \tag{3.15}$$

Note that here we relax the strict i.i.d. output distribution constraint, because without private randomness at the decoder, such a distribution in general cannot exactly be achieved for reasonably finite rates (R, R_c) when there is a distortion constraint. Indeed, this is evident in the achievability proof of Theorem 3.1.

For any $D \geq 0$ we let $\mathcal{R}_f(D)$ denote the set of all feedback achievable (R, R_c) pairs. The following theorem, proved in the Section 3.8.4, characterizes the closure of this set.

Theorem 3.5. *For any $D \geq 0$,*

$$\operatorname{cl}\mathcal{R}_f(D) = \left\{ \begin{array}{rcl} (R, R_c) \in \mathbb{R}^2 & : & \exists P_{X,Y} \in \mathcal{G}(D) \ s.t. \\ R & \geq & I(X;Y), \\ R + R_c & \geq & H(Y) \end{array} \right\}, \tag{3.16}$$

or equivalently,

$$\operatorname{cl}\mathcal{R}_f(D) = \left\{ \begin{array}{rcl} (R, R_c) \in \mathbb{R}^2 & : & \exists P_{X,Y,U} \in \mathcal{M}(D) \ s.t. \\ R & \geq & I(X;U), \\ R + R_c & \geq & H(Y) \end{array} \right\}. \tag{3.17}$$

Remark 3.3.

(a) It is important to note that if we allow the decoder to use private randomness while preserving the output distribution constraint (3.15), one can prove that the resulting achievable rate region is $\mathcal{L}(D)$. In this case, the only part to prove is the converse, since the achievability is obvious. However, the converse can be proven by using a similar technique as in [29, Section VI]. Hence, if we allow the decoder to use private randomness, replacing the strict output distribution constraint in the Definition 3.1 with (3.15) does not change the achievable rate region.

(b) Since $\mathcal{L}(D) \supset \operatorname{cl}\mathcal{R}_f(D)$, where the inclusion is strict in general, in addition to common randomness, private randomness used by decoder is also useful to decrease the necessary coding rate, which is interesting because private randomness in general increases the distortion.

3.6 Proof of Theorem 3.1

Our proof relies on techniques developed by Cuff in [29]. In particular, in the achievability part, we apply the 'likelihood encoder' of [30, 29] which is an elegant alternative to the standard random coding argument. The converse part of the proof is an appropriately modified version of the converse argument in [29]; however, in our setup this technique also works in the continuous alphabet case, while in [29] the finite alphabet assumption seem quite difficult to relax.

3.6.1 Achievability for Discrete Alphabets

Assume that (R, R_c) is in the interior of $\mathcal{L}(D)$. Then there exists $P_{X,Y,U} \in \mathcal{M}(D)$ such that $R > I(X;U)$ and $R + R_c > I(Y;U)$. The method used in this part of the proof comes from [29, Section V] where instead of explicitly constructing the encoder-decoder pair, a joint distribution was constructed from which the desired encoder-decoder behavior is established.

In this section, distributions which depend on realizations of some random variable (e.g., random codebook) will be denoted as bold upper case letters, but without referring to the corresponding realization for notational simplicity.

For each n, generate a random 'codebook' $\mathcal{C}_n := \{U^n(j,k)\}$ of u^n sequences independently drawn from P_U^n and indexed by $(j,k) \in [2^{nR}] \times [2^{nR_c}]$. For each realization

$\{u^n(j,k)\}$ of $\mathcal{C}_n$, define a distribution $\mathbf{\Gamma}_{X^n,Y^n,J,K}$ such that (J,K) is uniformly distributed on $[2^{nR}]\times[2^{nR_c}]$ and (X^n,Y^n) is the output of the stationary and memoryless channel $P^n_{X,Y|U}$ when we feed it with $u^n(J,K)$, i.e.,

$$\mathbf{\Gamma}_{X^n,Y^n,J,K}(x^n,y^n,j,k) \coloneqq \frac{1}{2^{n(R+R_c)}} P^n_{X,Y|U}(x^n,y^n|u^n(j,k)). \tag{3.18}$$

$\{\mathbf{\Gamma}_{X^n,Y^n,J,K}\}_{n\geq 1}$ are the distributions from which we derive a sequence of encoder-decoder pairs which for all n large enough *almost* meet the requirements in Definition 3.1.

Lemma 3.2 (Soft covering lemma [29, Lemma IV.1])**.** *Let $P_{V,W} = P_V P_{W|V}$ be the joint distribution of some random vector (V,W) on $\mathsf{V}\times\mathsf{W}$, where P_V is the marginal on V and $P_{W|V}$ is the conditional probability on W given V. For each n, generate the set $\mathcal{B}_n = \{V^n(i)\}$ of v^n sequences independently drawn from P^n_V and indexed by $i\in[2^{nR}]$. Let us define a random measure on W^n as*

$$\mathbf{P}_{W^n}(w^n) \coloneqq \frac{1}{|\mathcal{B}_n|}\sum_{i=1}^{|\mathcal{B}_n|} P_{W^n|V^n}(w^n|V^n(i)),$$

where $P_{W^n|V^n} = \prod_{i=1}^n P_{W|V}$. If $R \geq I(V;W)$, then we have

$$\mathbb{E}_{\mathcal{B}_n}\big[\|\mathbf{P}_{W^n} - P^n_W\|_{TV}\big] \leq \frac{3}{2}\exp\{-\kappa n\},$$

for some $\kappa > 0$.

Since $R + R_c > I(Y;U)$, by the soft covering lemma

$$\mathbb{E}_{\mathcal{C}_n}\big[\|\mathbf{\Gamma}_{Y^n} - P^n_Y\|_{TV}\big] \leq \frac{3}{2}\exp\{-cn\}, \tag{3.19}$$

where $c > 0$ and $\mathbb{E}_{\mathcal{C}_n}$ denotes expectation with respect to the distribution of $\mathcal{C}_n$. Note that for any fixed k, the collection $\mathcal{C}_n(k) := \{U^n(j,k)\}_j$ is a random codebook of size 2^{nR}. Since $R > I(X;U)$, the soft covering lemma again gives

$$\mathbb{E}_{\mathcal{C}_n(k)}\left[\|\mathbf{\Gamma}_{X^n|K=k} - P_X^n\|_{TV}\right] \leq \frac{3}{2}\exp\{-dn\}, \tag{3.20}$$

where $d > 0$ (same for all k) and $\mathbb{E}_{\mathcal{C}_n(k)}$ denotes expectation with respect to the distribution of $\mathcal{C}_n(k)$. Then, by the definition of total variation, we have

$$\begin{aligned}\mathbb{E}_{\mathcal{C}_n}\left[\|\mathbf{\Gamma}_{X^n,K} - \frac{1}{2^{nR_c}}P_X^n\|_{TV}\right] &:= \mathbb{E}_{\mathcal{C}_n}\left[\frac{1}{2}\sum_{x^n,k}\left|\mathbf{\Gamma}_{X^n,K}(x^n,k) - \frac{1}{2^{nR_c}}P_X^n(x^n)\right|\right]\\ &= \frac{1}{2^{nR_c}}\mathbb{E}_{\mathcal{C}_n}\left[\frac{1}{2}\sum_{x^n,k}\left|\mathbf{\Gamma}_{X^n|K}(x^n|k) - P_X^n(x^n)\right|\right]\\ &= \frac{1}{2^{nR_c}}\sum_k \mathbb{E}_{\mathcal{C}_n(k)}\left[\|\mathbf{\Gamma}_{X^n|K=k} - P_X^n\|_{TV}\right]\\ &\leq \frac{3}{2}\exp\{-dn\}. \end{aligned} \tag{3.21}$$

Furthermore, the expected value (taken with respect to the distribution of $\mathcal{C}_n$) of the distortion induced by $\mathbf{\Gamma}_{X^n,Y^n}$ is upper bounded by D as a result of the symmetry in the construction of $\mathcal{C}_n$, i.e.,

$$\begin{aligned}\mathbb{E}_{\mathcal{C}_n}\left[\sum_{x^n,y^n}\rho_n(x^n,y^n)\mathbf{\Gamma}_{X^n,Y^n}(x^n,y^n)\right] &= \mathbb{E}_{\mathcal{C}_n}\left[\sum_{j,k}\sum_{x^n,y^n}\rho_n(x^n,y^n)\mathbf{\Gamma}_{X^n,Y^n,J,K}(x^n,y^n,j,k)\right]\\ &= \sum_{x^n,y^n}\rho_n(x^n,y^n)\sum_{j,k}\mathbb{E}_{\mathcal{C}_n}\left[\mathbf{\Gamma}_{X^n,Y^n,J,K}(x^n,y^n,j,k)\right]\\ &= \sum_{x^n,y^n}\rho_n(x^n,y^n)P_{X,Y}^n(x^n,y^n) \leq D, \end{aligned} \tag{3.22}$$

where the last equality follows from the symmetry and the independence in the codebook construction, and the last inequality follows from the definition of $\mathcal{M}(D)$.

Now, since $\mathbf{\Gamma}_{Y^n,J|X^n,K} = \mathbf{\Gamma}_{J|X^n,K}\mathbf{\Gamma}_{Y^n|J,K}$, we define a randomized (n, R, R_c) source code such that it has the encoder-decoder pair $(\mathbf{\Gamma}_{J|X^n,K}, \mathbf{\Gamma}_{Y^n|J,K})$. Hence, (n, R, R_c) depends on the realization of $\mathcal{C}_n$. Let $\mathbf{P}_{X^n,Y^n,J,K}$ denote the distribution induced by (n, R, R_c), i.e.,

$$\mathbf{P}_{X^n,Y^n,J,K}(x^n, y^n, j, k) := \frac{1}{2^{nR_c}} P_X^n(x^n)\mathbf{\Gamma}_{Y^n,J|X^n,K}(y^n, j|x^n, k).$$

If two distributions are passed through the same channel, then the total variation between the joint distributions is the same as the total variation between the input distributions [29, Lemma V.2]. Hence, by (3.21)

$$\mathbb{E}_{\mathcal{C}_n}\left[\|\mathbf{\Gamma}_{X^n,Y^n,K,J} - \mathbf{P}_{X^n,Y^n,K,J}\|_{TV}\right] \leq \frac{3}{2}\exp\{-dn\}. \tag{3.23}$$

Then, (3.22) and (3.23) give

$$\mathbb{E}_{\mathcal{C}_n}\left[\sum_{x^n,y^n} \rho_n(x^n, y^n)\mathbf{P}_{X^n,Y^n}(x^n, y^n)\right] \leq D + \alpha\exp\{-dn\}, \tag{3.24}$$

where $\alpha = \rho_{\max}\frac{3}{2}$. By virtue of the properties of total variation distance, (3.19) and (3.23) also imply

$$\begin{aligned}\mathbb{E}_{\mathcal{C}_n}\left[\|\mathbf{P}_{Y^n} - P_Y^n\|_{TV}\right] &\leq \mathbb{E}_{\mathcal{C}_n}\left[\|\mathbf{P}_{Y^n} - \mathbf{\Gamma}_{Y^n}\|_{TV}\right] + \mathbb{E}_{\mathcal{C}_n}\left[\|\mathbf{\Gamma}_{Y^n} - P_Y^n\|_{TV}\right] \\ &\leq \frac{3}{2}\exp\{-dn\} + \frac{3}{2}\exp\{-cn\} \\ &= \alpha_n\exp\{-dn\}, \end{aligned} \tag{3.25}$$

where (without any loss of generality) we assumed $d < c$ and where $\alpha_n \coloneqq \frac{3}{2}\big(1 + \exp\{-(c-d)n\}\big) \leq 2$ if n is large enough.

Define the following functions of the random codebook $\mathcal{C}_n$:

$$D(\mathcal{C}_n) \coloneqq \sum_{x^n, y^n} \rho_n(x^n, y^n) \mathbf{P}_{X^n,Y^n}(x^n, y^n),$$

$$G(\mathcal{C}_n) \coloneqq \|\mathbf{P}_{Y^n} - P_Y^n\|.$$

Thus, the expectations of $D(\mathcal{C}_n)$ and $G(\mathcal{C}_n)$ satisfy (3.24) and (3.25), respectively. For any $\delta \in (0, d)$, Markov's inequality gives

$$\mathsf{Pr}\Big\{G(\mathcal{C}_n) \leq \exp\{-\delta n\}\Big\} \geq 1 - \frac{\alpha_n \exp\{-dn\}}{\exp\{-\delta n\}}, \tag{3.26}$$

$$\mathsf{Pr}\Big\{D(\mathcal{C}_n) \leq D + \delta\Big\} \geq 1 - \frac{D + \alpha \exp\{-dn\}}{D + \delta}. \tag{3.27}$$

Since

$$\lim_{n\to\infty}\left(2 - \frac{\alpha_n \exp\{-dn\}}{\exp\{-\delta n\}} - \frac{D + \beta \exp\{-dn\}}{D + \delta}\right) = 2 - \frac{D}{D+\delta} > 1,$$

there exists a positive $N(\delta)$ such that for $n \geq N(\delta)$, we have

$$\mathsf{Pr}\Big\{\Big(D(\mathcal{C}_n) \leq D + \delta\Big) \bigcap \Big(G(\mathcal{C}_n) \leq \exp\{-\delta n\}\Big)\Big\} > 0.$$

This means that for each $n \geq N(\delta)$, there is a realization of $\mathcal{C}_n$ which gives

$$\sum_{x^n, y^n} \rho_n(x^n, y^n) \mathbf{P}_{X^n, Y^n}(x^n, y^n) \leq D + \delta \tag{3.28}$$

$$\|\mathbf{P}_{Y^n} - P_Y^n\| \leq \exp\{-\delta n\}. \tag{3.29}$$

Hence, the sequence of (n, R, R_c) randomized source codes corresponding to these realizations almost satisfies the achievability constraints. Next we can slightly modify this coding scheme so that the code exactly satisfies the i.i.d. output distribution constraint $Y^n = \psi^n = P_Y^n$ while having distortion upper bounded by $D + \delta$.

Before presenting this modification, we recall the definition of optimal coupling and the optimal transportation cost from Section 2.5. Let π, λ be probability measures over finite or continuous alphabets W and V, respectively. The optimal transportation cost $\hat{T}(\pi, \lambda)$ between π and λ (see, e.g., [99]) with respect to a cost function $c : \mathsf{V} \times \mathsf{W} \to [0, \infty)$ is defined by

$$\hat{T}(\pi, \lambda) = \inf\big\{\mathbb{E}[c(V, W)] : V \sim \pi, W \sim \lambda\big\}. \tag{3.30}$$

The distribution achieving $\hat{T}(\pi, \lambda)$ is called an optimal coupling of π and λ. In this section, somewhat informally, we also call the corresponding conditional probability on W given V an optimal coupling.

Consider the (n, R, R_c) randomized source code depicted in Fig. 3.4 which is obtained by augmenting (n, R, R_c) with the optimal coupling $T_{\hat{Y}^n|Y^n}$ between $\mathbf{P}_{Y^n}$ and ψ^n with transportation cost $\hat{T}(\mathbf{P}_{Y^n}, \psi^n)$ when the cost function is $\rho_n(x^n, y^n) = \sum_{i=1}^n d(x_i, y_i)^p$, where d is a metric on X. Using [99, Theorem 6.15] one can show

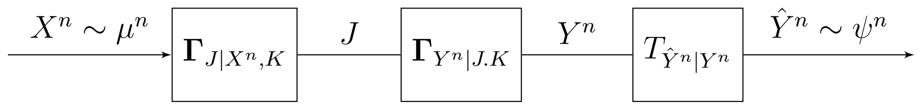

Figure 3.4: Code achieving optimal rate region for discrete alphabets.

$$
\begin{aligned}
\hat{T}_n(\mathbf{P}_{Y^n}, \psi^n) &= \frac{1}{n} W_q^q(\mathbf{P}_{Y^n}, \psi^n) \\
&\leq \frac{1}{n} 2^{\frac{q}{r}} \sum_{y^n} d_n(y_0^n, y^n)^q \big| \mathbf{P}_{Y^n}(y^n) - \psi^n(y^n) \big| \\
&\leq 2^{\frac{q}{r}} \rho_{\max} \| \mathbf{P}_{Y^n} - \psi^n \|_{TV} \\
&\leq 2^{\frac{q}{r}} \rho_{\max} \exp\{-\delta n\}, \text{ by (3.29)}
\end{aligned} \tag{3.31}
$$

where $y_0^n \in \mathsf{Y}^n$ is arbitrary, $q = \max\{1, p\}$, W_q denotes the Wasserstein distance of order q [99, Definition 6.1], $\frac{1}{q} + \frac{1}{r} = 1$, and d_n is the product metric defined as $d_n(x^n, y^n) = \left(\sum_{i=1}^n d(x_i, y_i)^p\right)^{1/q}$.

Recall that $\rho(x, y) = d(x, y)^p$ for some $p > 0$. Recall also that if $p \geq 1$, then $\|V^n\|_p := \left(E\big[\sum_{i=1}^n |V_i|^p\big]\right)^{1/p}$ is a norm on $\mathbb{R}^n$-valued random vectors whose components have finite pth moments, and if $1 < p < 0$, we still have $\|U^n + V^n\|_p \leq \|U^n\|_p + \|V^n\|_p$. Thus we can upper bound the distortion $\mathbb{E}[\rho_n(X^n, \hat{Y}^n)]$ of the code in Fig. 3.4 as follows:

$$
\begin{aligned}
\left(\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n \rho(X_i, \hat{Y}_i)\right]\right)^{1/q} &= \left(\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n d(X_i, \hat{Y}_i)^p\right]\right)^{1/q} \\
&\leq \left(\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n d(X_i, Y_i)^p\right]\right)^{1/q} + \left(\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n d(Y_i, \hat{Y}_i)^p\right]\right)^{1/q} \\
&= \Big(\mathbb{E}[\rho_n(X^n, Y^n)]\Big)^{1/q} + \hat{T}_n(P_{Y^n}, \psi^n)^{1/q},
\end{aligned}
$$

Hence, by (3.28) and (3.31) we obtain

$$\limsup_{n\to\infty} \mathbb{E}[\rho_n(X^n, \hat{Y}^n)] \leq D + \delta,$$

which completes the proof.

3.6.2 Achievability for Continuous Alphabets

In this section, we let $\mathsf{X} = \mathsf{Y} = \mathbb{R}$, $\rho(x,y) = (x-y)^2$, and assume that μ and ψ have finite second moments. Analogous to the achievability proof of Theorem 2.7 for continuous alphabets, we make use of the discrete case to prove the achievability for the continuous case.

Assume that (R, R_c) is in the interior of $\mathcal{L}(D)$. Then there exists $P_{X,Y,U} \in \mathcal{M}(D)$ such that $R > I(X;U)$ and $R+R_c > I(Y;U)$. Let q_k denote the uniform quantizer on the interval $[-k,k]$ having 2^k levels, the collection of which is denoted by L_k. Extend q_k to the entire real line by using the nearest neighborhood encoding rule. Define $X(k) \coloneqq q_k(X)$ and $Y(k) \coloneqq q_k(Y)$. Let μ_k and ψ_k denote the distributions of $X(k)$ and $Y(k)$, respectively. It is clear that

$$\mathbb{E}[(X - X(k))^2] \to 0, \text{ and } \mathbb{E}[(Y - Y(k))^2] \to 0 \text{ as } k \to \infty. \tag{3.32}$$

Moreover, by [99, Theorem 6.9] it follows that $\hat{T}(\mu_k, \mu) \to 0$ and $\hat{T}(\psi_k, \psi) \to 0$ as $k \to \infty$ since $\mu_k \to \mu$, $\psi_k \to \psi$ weakly [15], and $\mathbb{E}[X(k)^2] \to \mathbb{E}[X^2]$, $\mathbb{E}[Y(k)^2] \to \mathbb{E}[Y^2]$. For each k define $D_k \coloneqq \mathbb{E}[(X(k) - Y(k))^2]$. Then by (3.32)

$$\lim_{k\to\infty} D_k = \mathbb{E}[(X-Y)^2] \leq D.$$

For any k, let $\mathcal{M}_k(D_k)$ be the set of distributions obtained by replacing μ, ψ, and $\mathsf{X} = \mathsf{Y}$ with μ_k, ψ_k, and $\mathsf{X}_k = \mathsf{Y}_k = L_k$, respectively, in (3.3). Note that $X(k) - U - Y(k)$ and

$$I(X(k);U) \leq I(X;U) \text{ and } I(Y(k);U) \leq I(Y;U) \tag{3.33}$$

by data processing inequality which implies $R > I(X(k);U)$ and $R + R_c > I(Y(k);U)$. Hence, $P_{X(k),Y(k),U} \in \mathcal{M}_k(D_k)$. Then, using the achievability result for discrete alphabets, for any k, one can find a sequence of $(n, R, R_c)^k$ randomized source codes for common source and reproduction alphabet L_k, source distribution μ_k, and desired output distribution ψ_k such that the upper limit of the distortions of these codes is upper bounded by D_k.

For each k and n, consider the randomized source codes defined in Fig. 3.5. We

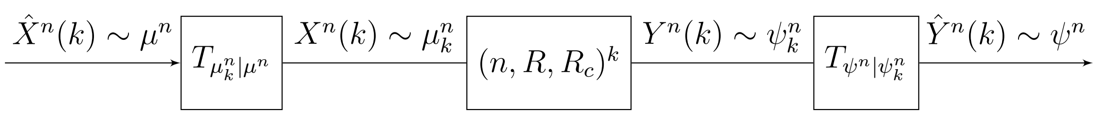

Figure 3.5: Code achieving optimal rate region for continuous alphabets.

note that the definition of the optimal transportation cost implies that $\hat{T}(\mu_k^n, \mu^n) \leq \hat{T}(\mu_k, \mu)$ and $\hat{T}(\psi_k^n, \psi^n) \leq \hat{T}(\psi_k, \psi)$. Hence, using the triangle inequality for the norm $\|V^n\|_2 \coloneqq \left(\sum_{i=1}^n E[V_i^2]\right)^{1/2}$ on $\mathbb{R}^n$-valued random vectors having finite second moments, for all k, we have

$$\limsup_{n\to\infty} \mathbb{E}\left[\left(\hat{X}^n(k) - \hat{Y}^n(k)\right)^2\right]^{1/2}$$
$$\leq \limsup_{n\to\infty} \left(\hat{T}(\mu_k^n, \mu^n)^{1/2} + \mathbb{E}\left[\left(X^n(k) - Y^n(k)\right)^2\right]^{1/2} + \hat{T}(\psi_k^n, \psi^n)^{1/2}\right)$$

$$\leq \hat{T}(\mu_k,\mu)^{1/2} + \hat{T}(\psi_k,\psi)^{1/2} + \limsup_{n\to\infty} \mathbb{E}\left[\left(X^n(k) - Y^n(k)\right)^2\right]^{1/2}$$
$$\leq \hat{T}(\mu_k,\mu)^{1/2} + \hat{T}(\psi_k,\psi)^{1/2} + D_k^{1/2}.$$

By choosing k large enough we can make the last term arbitrarily close to D, which completes the proof.

3.6.3 Cardinality Bound

In this section, we show that for any discrete distribution $\Lambda_{X,Y,W}$ forming a Markov chain $X - W - Y$, there exists a discrete distribution $\Gamma_{X,Y,U}$ forming another Markov chain $X - U - Y$ such that

$$|\mathsf{U}| \leq |\mathsf{X}| + |\mathsf{Y}| + 1,$$
$$\Gamma_X = \Lambda_X$$
$$\Gamma_Y = \Lambda_Y,$$
$$\mathbb{E}_\Gamma[\rho(X,Y)] = \mathbb{E}_\Lambda[\rho(X,Y)],$$
$$I_\Gamma(X;U) = I_\Lambda(X;W),$$
$$I_\Gamma(Y;U) = I_\Lambda(Y;W),$$

where $I_P(X;U)$ denotes the mutual information computed with respect to the distribution P. Let $\mathcal{P}(\mathsf{X}) \times \mathcal{P}(\mathsf{Y})$ denote the product of probability simplices $\mathcal{P}(\mathsf{X})$ and $\mathcal{P}(\mathsf{Y})$ representing the set of all distributions of independent random variables over $\mathsf{X} \times \mathsf{Y}$. This set is compact and connected when viewed as a subset of $\mathbb{R}^{|\mathsf{X}|+|\mathsf{Y}|}$. Without loss of generality $\mathsf{X} = \{1,\ldots,|\mathsf{X}|\}$ and $\mathsf{Y} = \{1,\ldots,|\mathsf{Y}|\}$. Since $H(X)$ is fixed in $I(X;W) = H(X) - H(X|W)$ (similarly $H(Y)$ is fixed in $I(Y;W) = H(Y) - H(Y|W)$), we define the following real valued continuous functions on $\mathcal{P}(\mathsf{X}) \times \mathcal{P}(\mathsf{Y})$:

$$g_j(\nu) = \begin{cases} \nu_x(j), & \text{if } j = 1, \ldots, |\mathsf{X}| - 1 \\ \nu_y(j), & \text{if } j = |\mathsf{X}|, \ldots, |\mathsf{X}| + |\mathsf{Y}| - 2 \\ \mathbb{E}_\nu[\rho(X, Y)], & \text{if } j = |\mathsf{X}| + |\mathsf{Y}| - 1 \\ H(\nu_x), & \text{if } j = |\mathsf{X}| + |\mathsf{Y}| \\ H(\nu_y), & \text{if } j = |\mathsf{X}| + |\mathsf{Y}| + 1, \end{cases}$$

where $\nu = \nu_x \otimes \nu_y$ and $H(P)$ denotes the entropy of the distribution P. By so-called 'support lemma' [43, Appendix C], there exists a random variable $U \sim \Gamma_U$, taking values in U with $|\mathsf{U}| \leq |\mathsf{X}| + |\mathsf{Y}| + 1$, and a conditional probability $\Gamma_{X|U}\Gamma_{Y|U}$ on $\mathsf{X} \times \mathsf{Y}$ given U such that for $j = 1, \ldots, |\mathsf{X}| + |\mathsf{Y}| + 1$,

$$\sum_w g_j(\Lambda_{X|W=w}\Lambda_{Y|W=w})\Lambda_W(w) = \sum_u g_j(\Gamma_{X|U=u}\Gamma_{Y|U=u})\Gamma_U(u),$$

which completes the proof.

3.6.4 Converse

We use the standard approach to prove the converse in Theorem 3.1, i.e., that $\operatorname{cl}\mathcal{R}(D) \subset \mathcal{L}(D)$ for any $D \geq 0$. We note that this proof holds both for finite alphabets and continuous alphabets.

For each R_c, define the minimum coding rate R at distortion level D as

$$\min\{R \in \mathcal{R}(D, R_c)\} =: I_{R_c}(\mu\|\psi, D).$$

Using a time-sharing argument and the operational meaning of $I_{R_c}(\mu\|\psi, D)$, one can prove that $I_{R_c}(\mu\|\psi, D)$ is convex in D, and therefore, continuous in D, $0 < D < \infty$ (see the proof of Lemma 3.1). Since $I_{R_c}(\mu\|\psi, D)$ is nonincreasing in D, we have $I_{R_c}(\mu\|\psi, 0) \geq \lim_{D\to 0} I_{R_c}(\mu\|\psi, D)$. But by the definition of $\mathcal{R}(0, R_c)$, we also have $\lim_{D\to 0} I_{R_c}(\mu\|\psi, D) \in \mathcal{R}(0, R_c)$, so that $I_{R_c}(\mu\|\psi, 0) = \lim_{D\to 0} I_{R_c}(\mu\|\psi, D)$. Hence, $I_{R_c}(\mu\|\psi, D)$ is also continuous at $D = 0$. Let us define $\mathcal{R}^*(D) = \{(R, R_c) \in \mathbb{R}^2 : R > I_{R_c}(\mu\|\psi, D)\}$ and let $(R, R_c) \in \mathcal{R}^*(D)$. Since $I_{R_c}(\mu\|\psi, D)$ is continuous in D, there exists $\varepsilon > 0$ such that $R > I_{R_c}(\mu\|\psi, D - \varepsilon)$. Hence, there exists, for all sufficiently large n, a (n, R, R_c) randomized source code such that

$$\begin{aligned}
\mathbb{E}[\rho_n(X^n, Y^n)] &\leq D, \\
Y^n &\sim \psi^n.
\end{aligned}$$

For each n, define the random variable $Q_n \sim \mathsf{Unif}\{1, \ldots, n\}$ which is independent of (X^n, Y^n, J, K), associated with the n^{th} randomized source code. Since $J \in [2^{nR}]$,

$$\begin{aligned}
nR \geq H(J) \geq H(J|K) &\geq I(X^n; J|K) \\
&\overset{(a)}{=} I(X^n; J, K) \\
&= \sum_{i=1}^{n} I(X_i; J, K|X^{i-1}) \\
&\overset{(b)}{=} \sum_{i=1}^{n} I(X_i; J, K, X^{i-1}) \\
&\geq \sum_{i=1}^{n} I(X_i; J, K) \\
&= nI(X_{Q_n}; J, K|Q_n)
\end{aligned}$$

$$\stackrel{(c)}{=} nI(X_{Q_n}; J, K, Q_n),$$

where (a) follows from the independence of X^n and K, (b) follows from i.i.d. nature of the source X^n and (c) follows from the independence of X_{Q_n} and Q_n. Similarly, for the sum rate we have

$$\begin{aligned}
n(R+R_c) \geq H(J,K) &\geq I(Y^n; J, K) \\
&= \sum_{i=1}^{n} I(Y_i; J, K|Y^{i-1}) \\
&\stackrel{(a)}{=} \sum_{i=1}^{n} I(Y_i; J, K, Y^{i-1}) \\
&\geq \sum_{i=1}^{n} I(Y_i; J, K) \\
&= nI(Y_{Q_n}; J, K|Q_n) \\
&\stackrel{(b)}{=} nI(Y_{Q_n}; J, K, Q_n),
\end{aligned}$$

where (a) follows from i.i.d. nature of the output Y^n and (b) follows from the independence of Y_{Q_n} and Q_n. Notice that $X_{Q_n} \sim \mu$, $Y_{Q_n} \sim \psi$, and $X_{Q_n} - (J, K, Q_n) - Y_{Q_n}$. We also have

$$\begin{aligned}
\mathbb{E}[\rho(X_{Q_n}, Y_{Q_n})] &= \mathbb{E}\Big[\mathbb{E}\big[\rho(X_{Q_n}, Y_{Q_n})|Q_n\big]\Big] \\
&= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\big[\rho(X_{Q_n}, Y_{Q_n})|Q_n = i\big] \\
&= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\big[\rho(X_i, Y_i)\big] \\
&= \mathbb{E}\big[\rho_n(X^n, Y^n)\big] \leq D.
\end{aligned}$$

Define $U = (J, K, Q_n)$ and denote by $P_{X,Y,U}$ the distribution of (X_{Q_n}, Y_{Q_n}, U). Hence, $P_{X,Y,U} \in \mathcal{M}(D)$ which implies that $(R, R_c) \in \mathcal{L}(D)$. Hence, $\mathcal{R}^*(D) \subset \mathcal{L}(D)$. But, since $\mathcal{L}(D)$ is closed in $\mathbb{R}^2$, we also have $\operatorname{cl} \mathcal{R}^*(D) = \operatorname{cl} \mathcal{R}(D) \subset \mathcal{L}(D)$.

3.7 Conclusion

By generalizing the practically motivated distribution preserving quantization problem, we have derived the rate distortion region for randomized source coding of a stationary and memoryless source, where the output of the code is restricted to be also stationary and memoryless with some specified distribution. For a given distortion level, the rate region consists of coding and common randomness rate pairs, where the common randomness is independent of the source and shared between the encoder and the decoder. Unlike in classical rate distortion theory, here shared independent randomness can decrease the necessary coding rate communicated between the encoder and decoder.

3.8 Proofs

3.8.1 Proof of Corollary 3.1

Assume that (R, R_c) is in the interior of $\mathcal{S}(D)$. Then there exists $P_{X,Y,U} \in \mathcal{H}(D)$ such that $R > I(X;U)$ and $R + R_c > I(X,Y;U)$. Let $\pi = P_{X,Y}$. By Theorem 3.2 there exists a sequence of (n, R, R_c) randomized source codes such that

$$\lim_{n\to\infty} \|P_{X^n,Y^n} - \pi^n\| = 0, \tag{3.34}$$

where (X^n, Y^n) denotes the input-output of the n^{th} code. Since ρ_n is bounded, we have

$$\begin{aligned}\limsup_{n\to\infty}\big|\mathbb{E}[\rho_n(X^n,Y^n)] - D\big| &= \limsup_{n\to\infty}\big|\mathbb{E}[\rho_n(X^n,Y^n)] - \mathbb{E}_{\pi^n}[\rho_n(X^n,Y^n)]\big| \\ &\leq \limsup_{n\to\infty}\|P_{X^n,Y^n} - \pi^n\|_{TV}\,\rho_{\max} = 0, \end{aligned} \tag{3.35}$$

where $\mathbb{E}_{\pi^n}$ denotes the expectation with respect to π^n. Let $T_{\hat{Y}^n|Y^n}$ be the optimal coupling (i.e., conditional probability) between P_{Y^n} and ψ^n with the transportation cost $\hat{T}(P_{Y^n}, \psi^n)$ with cost function ρ_n. By [99, Theorem 6.15] and (3.34) one can prove that $\limsup_{n\to\infty}\hat{T}(P_{Y^n}, \psi^n) = 0$ as in (3.31).

For each n, let us define the following encoder-decoder pair (see Fig. 3.6)

$$\tilde{E}^n_{J|X^n,K} := E^n_{J|X^n,K} \tag{3.36}$$

$$\tilde{F}^n_{\hat{Y}^n|J,K} := T_{\hat{Y}^n|Y^n} \circ F^n_{Y^n|J,K}, \tag{3.37}$$

where (E^n, F^n) is the encoder-decoder pair of the n^{th} code. Note that the randomized

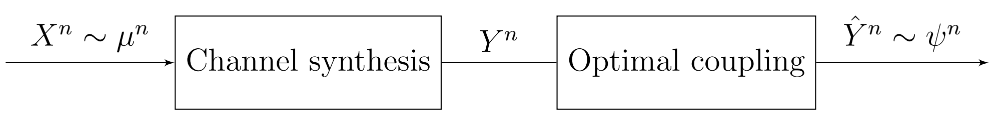

Figure 3.6: Code achieving suboptimal rate region

source code defined in (3.36) and (3.37) has rates (R, R_c) and output distribution ψ^n. Furthermore, using the triangle inequality as in Section 3.6.1 one can prove that

$$\limsup_{n\to\infty}\mathbb{E}[\rho_n(X^n, \hat{Y}^n)] \leq D$$

using (3.35) and the fact that $\limsup_{n\to\infty} \hat{T}(P_{Y^n}, \psi^n) = 0$. This completes the proof.

3.8.2 Proof of Lemma 3.1

Let D_1 and D_2 be two distinct positive real numbers and choose $\alpha \in (0,1)$. Fix any $\varepsilon > 0$. Let δ be a small positive number which will be specified later. By the definition of $I_0(\mu\|\psi, D)$ and by Theorem 3.1 there exist positive real numbers R_1 and R_2 such that

$$R_i \leq I_0(\mu\|\psi, D_i) + \delta, i = 1, 2,$$

and such that for all sufficiently large n there exist randomized $(n, R_1, 0)$ and $(n, R_2, 0)$ source codes having output distribution ψ^n which satisfy

$$\mathbb{E}\left[\rho_n\left(X^n, F^{(1)}\big(E^{(1)}(X^n)\big)\right)\right] \leq D_1 + \delta, i = 1, 2,$$

where $(E^{(1)}, F^{(1)})$ and $(E^{(2)}, F^{(2)})$ are the encoder-decoder pairs for these codes. Let $\{k_M\}_{M\geq 1}$ be a sequence of positive integers such that $\lim_{M\to\infty} \frac{k_M}{M} = \alpha$. Let N be a positive integer which will be specified later. For the source block X^{nN} define the following randomized source code:

$$E := \big(\underbrace{E^{(1)}, \ldots, E^{(1)}}_{k_N\text{-times}}, \underbrace{E^{(2)}, \ldots, E^{(2)}}_{N-k_N\text{-times}}\big),$$
$$F := \big(\underbrace{F^{(1)}, \ldots, F^{(1)}}_{k_N\text{-times}}, \underbrace{F^{(2)}, \ldots, F^{(2)}}_{N-k_N\text{-times}}\big).$$

Note that the output distribution for this randomized source code is ψ^{nN}, and its rate R and distortion D satisfy the following

$$
\begin{aligned}
R &= \frac{1}{nN}\big(k_N n R_1 + (N-k_N) n R_2\big) \\
&\leq \frac{k_N}{N} I_0(\mu\|\psi, D_1) + \frac{N-k_N}{N} I_0(\mu\|\psi, D_2) + \delta,
\end{aligned}
$$

and

$$
D = \mathbb{E}\big[\rho_{nN}(X^{nN}, Y^{nN})\big] \leq \frac{k_N}{N} D_1 + \frac{N-k_N}{N} D_2 + \delta.
$$

Since $\lim_{M\to\infty} \frac{k_M}{M} = \alpha$, one can choose N and δ such that R is upper bounded by $\alpha I_0(\mu\|\psi, D_1) + (1-\alpha) I_0(\mu\|\psi, D_2) + \varepsilon$ and D is upper bounded by $\alpha D_1 + (1-\alpha) D_2 + \varepsilon$. By Definition 3.1, this yields

$$
\begin{aligned}
&I_0\big(\mu\|\psi, \alpha D_1 + (1-\alpha) D_2\big) \\
&\qquad\qquad \leq \alpha I_0(\mu\|\psi, D_1) + (1-\alpha) I_0(\mu\|\psi, D_2) + \varepsilon.
\end{aligned}
$$

Since ε is arbitrary, this completes the proof.

3.8.3 Proof of Theorem 3.4

Since $\mathcal{R}_e(D, R_c) \supset \mathcal{R}_e(D, 0)$ for all R_c, it is enough to prove that

$$
\begin{aligned}
\mathcal{R}_e(D, 0) &\supset \mathcal{L}(D, \infty), \\
\mathcal{R}_e(D, R_c) &\subset \mathcal{L}(D, \infty).
\end{aligned}
$$

Recall that

$$\mathcal{L}(D,\infty) = \{R \in \mathbb{R} : \exists P_{X,Y} \in \mathcal{G}(D) \text{ s.t. } R \geq I(X;Y)\}.$$

Let us assume that $R \in \mathcal{L}(D,\infty)$. Then, there exists $P_{X,Y} =: \pi \in \mathcal{G}(D)$ such that $R \geq I(X;Y)$. Fix any $\varepsilon > 0$. By Theorem 3.3 there exists a sequence of (n, R, ∞) randomized source codes such that

$$\lim_{n\to\infty} \|p_{X^n,Y^n} - \pi\|_{TV} = 0 \text{ in probability,} \tag{3.38}$$

which implies

$$\lim_{n\to\infty} \|p_{Y^n} - \psi\|_{TV} = 0 \text{ in probability.}$$

Hence, this sequence of codes satisfies the second constraint in Definition 3.3. To show that the codes satisfy the distortion constraint, we use the same steps in [30, Section VII-D]. We have

$$\begin{aligned}
\rho_n(X^n, Y^n) &= \frac{1}{n}\sum_{i=1}^{n} \rho(X_i, Y_i) \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{x,y} 1_{\{X_i=x, Y_i=y\}}\rho(x,y) \\
&= \sum_{x,y} \rho(x,y)\frac{1}{n}\sum_{i=1}^{n} 1_{\{X_i=x, Y_i=y\}} \\
&= \mathbb{E}_{p_{X^n,Y^n}}[\rho(X,Y)],
\end{aligned}$$

where 1_B denotes the indicator of event B and $\mathbb{E}_{p_{X^n,Y^n}}$ denotes the expectation with respect to the empirical distribution p_{X^n,Y^n} of (X^n, Y^n). For any $\varepsilon_1 > 0$, by (3.38) we have

$$\mathsf{Pr}\left\{ \|p_{X^n,Y^n} - \pi\|_{TV} > \varepsilon_1 \right\} < \varepsilon_1,$$

for all sufficiently large n. Define the event $B_{\varepsilon_1} \coloneqq \left\{ \|p_{X^n,Y^n} - \pi\|_{TV} \le \varepsilon_1 \right\}$. Then, for all sufficiently large n, we obtain

$$\begin{aligned}
\mathbb{E}[\rho_n(X^n, Y^n)] &= \mathbb{E}\left[\mathbb{E}_{p_{X^n,Y^n}}\left[\rho(X,Y)\right] \right] \\
&= \mathbb{E}\left[\mathbb{E}_{p_{X^n,Y^n}}\left[\rho(X,Y)\right] 1_{B_{\varepsilon_1}} \right] + \mathbb{E}\left[\mathbb{E}_{p_{X^n,Y^n}}\left[\rho(X,Y)\right] 1_{B^c_{\varepsilon_1}} \right] \\
&\le \mathbb{E}\left[\mathbb{E}_{p_{X^n,Y^n}}\left[\rho(X,Y)\right] 1_{B_{\varepsilon_1}} \right] + \rho_{\max}\varepsilon_1 \\
&\le \mathbb{E}_\pi\left[\rho(X,Y)\right] + 2\varepsilon_1\rho_{\max} \\
&\le D + 2\varepsilon_1\rho_{\max}.
\end{aligned}$$

By choosing ε_1 such that $2\varepsilon_1\rho_{\max} < \varepsilon$, we obtain $\mathcal{R}_e(D, 0) \supset \mathcal{L}(D, \infty)$.

To prove $\mathcal{R}_e(D, R_c) \subset \mathcal{L}(D, \infty)$, we use the same arguments as in [30, Section VII-B]. Let us choose $R \in \mathcal{R}_e(D, R_c)$ with the corresponding sequence of (n, R, R_c) randomized source codes satisfying constraints in Definition 3.3. For each n, define the random variable $Q_n \sim \mathsf{Unif}\{1, \ldots, n\}$ which is independent of the input-output (X^n, Y^n) of the code (n, R, R_c). Then, we have

$$\begin{aligned}
nR &\ge H(J) \\
&\ge I(X^n; Y^n)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} I(X_i; Y^n | X^{i-1}) \\
&= \sum_{i=1}^{n} I(X_i; Y^n, X^{i-1}) \\
&\geq \sum_{i=1}^{n} I(X_i; Y_i) \\
&= nI(X_{Q_n}; Y_{Q_n} | Q_n) \\
&\overset{(a)}{=} nI(X_{Q_n}; Y_{Q_n}, Q_n) \\
&\geq nI(X_{Q_n}; Y_{Q_n}),
\end{aligned} \tag{3.39}
$$

where (a) follows from the independence of X_{Q_n} and Q_n. We also have

$$
\begin{aligned}
\mathbb{E}[\rho(X_{Q_n}, Y_{Q_n})] &= \mathbb{E}\Big[\mathbb{E}\big[\rho(X_{Q_n}, Y_{Q_n}) | Q_n\big]\Big] \\
&= \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\big[\rho(X_{Q_n}, Y_{Q_n}) | Q_n = i\big] \\
&= \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\big[\rho(X_i, Y_i)\big] \\
&= \mathbb{E}\big[\rho_n(X^n, Y^n)\big].
\end{aligned} \tag{3.40}
$$

One can prove $P_{Y_{Q_n}} \to \psi$ in total variation (see, e.g., [30, Section VII-B-3]). Since the set of probability distributions over $\mathsf{X} \times \mathsf{Y}$ is compact with respect to the total variation distance, we can find a subsequence $\{(X_{Q_{n_k}}, Y_{Q_{n_k}})\}$ of $\{(X_{Q_n}, Y_{Q_n})\}$ such that

$$
P_{X_{Q_{n_k}}, Y_{Q_{n_k}}} \to P_{\hat{X}, \hat{Y}}
$$

in total variation for some $P_{\hat{X},\hat{Y}}$. But, since $P_{X_{Q_{n_k}}} = \mu$ for all k and $P_{Y_{Q_n}} \to \psi$ in total variation, we must have $P_{\hat{X}} = \mu$ and $P_{\hat{Y}} = \psi$. Now, taking the limit of (3.39) and (3.40) through this subsequence, we obtain

$$R \geq \lim_{k\to\infty} I(X_{Q_{n_k}}; Y_{Q_{n_k}}) = I(\hat{X}; \hat{Y})$$

and

$$\begin{aligned}\mathbb{E}[\rho(\hat{X}, \hat{Y})] &= \lim_{k\to\infty} \mathbb{E}[\rho(X_{Q_{n_k}}, Y_{Q_{n_k}})] \\ &= \lim_{k\to\infty} \mathbb{E}[\rho_{n_k}(X^{n_k}, Y^{n_k})] \leq D.\end{aligned}$$

Hence, $R \in \mathcal{L}(D, \infty)$ which completes the proof.

3.8.4 Proof of Theorem 3.5

Achievability: Assume (R, R_c) is in the interior of $\operatorname{cl} R_f(D)$. Then there exists $P_{X,Y} =: \pi \in \mathcal{G}(D)$ such that $R > I(X;Y)$ and $R + R_c > H(Y)$. By [8, Theorem 1] or [29, Section III-E], there exists a sequence of (n, R, R_c) randomized source codes with feedback such that

$$\|P_{X^n,Y^n} - \pi^n\|_{TV} \to 0.$$

Hence, $\|P_{Y^n} - \psi^n\|_{TV} \to 0$ and

$$\lim_{n\to\infty} \mathbb{E}[\rho_n(X^n, Y^n)] = \lim_{n\to\infty} \mathbb{E}_{\pi^n}[\rho_n(X^n, Y^n)] \leq D$$

completing the proof.

Converse: Let $(R, R_c) \in \mathrm{cl}\,\mathcal{R}_f(D)$. Using a similar argument as in Section 3.8.1, one can show that

$$nR \geq nI(X_{Q_n}; Y_{Q_n}), \tag{3.41}$$

and

$$\mathbb{E}[\rho(X_{Q_n}, Y_{Q_n})] = \mathbb{E}[\rho_n(X^n, Y^n)], \tag{3.42}$$

where $Q_n \sim \mathsf{Unif}\{1, \ldots, n\}$ is independent of input-output (X^n, Y^n) of the corresponding randomized source code, and $P_{Y_{Q_n}} \to \psi$ in total variation. Also, there is a subsequence $\{(X_{Q_{n_k}}, Y_{Q_{n_k}})\}$ such that $P_{X_{Q_{n_k}}, Y_{Q_{n_k}}} \to P_{\hat{X}, \hat{Y}}$ in total variation for some $P_{\hat{X}, \hat{Y}}$ with $P_{\hat{X}} = \mu$ and $P_{\hat{Y}} = \psi$. By taking the limit of (3.41) and (3.42) through this subsequence we obtain

$$R \geq I(\hat{X}; \hat{Y}), \tag{3.43}$$

$$\mathbb{E}[\rho(\hat{X}, \hat{Y})] \leq D. \tag{3.44}$$

Hence, the first inequality in (3.16) is satisfied. To show the second inequality, let $\varepsilon > 0$ and define

$$S_\varepsilon^{(n)} := \left\{ y^n \in \mathsf{Y}^n : 2^{-n(H(\psi)+\varepsilon)} \leq P_{Y^n}(y^n) \leq 2^{-n(H(\psi)-\varepsilon)} \right\}.$$

Since $\|P_{Y^n} - \psi^n\| \to 0$, by the weak AEP [25, Theorem 3.1.2-4]

$$|S_\varepsilon^{(n)}| \geq (1 - \varepsilon')2^{n(H(\psi)-\varepsilon')}$$

for all sufficiently large n, where $\varepsilon' \to 0$ as $\varepsilon \to 0$. Note that for each n, the number of y^ns with positive probability at the output of the decoder is at most $2^{n(R+R_c)}$ (since the decoder is deterministic function of J and K). Hence, we have

$$\begin{aligned} 2^{n(R+R_c)} &\geq |S_\varepsilon^{(n)}| \\ &\geq (1-\varepsilon')2^{n(H(\psi)-\varepsilon')}. \end{aligned}$$

Since ε and n are arbitrary, this yields $R + R_c \geq H(\psi) = H(Y)$.

Part II

Quantization in Approximation Problems for Stochastic Control

Chapter 4

Quantization of the Action Space and Asymptotic Optimality of Quantized Policies

4.1 Introduction

In this chapter, we study the finite-action approximation of optimal control policies for a discrete time Markov decision processes (MDPs) with Borel state and action spaces, under discounted and average cost criteria. Various stochastic control problems may benefit from such an investigation.

The optimal information transmission problem in networked control systems is one such example. In many applications to networked control, the perfect transmission of the control actions to an actuator is infeasible when there is a communication channel of finite capacity between a controller and an actuator. Hence, the actions of the controller must be quantized to facilitate reliable transmission to an actuator. Although the problem of optimal information transmission from a plant/sensor to a controller has been studied extensively (see, e.g., [107] and references therein), much less is known about the problem of transmitting actions from a controller to an actuator. Such transmission schemes usually require a simple encoding/decoding rule since an actuator does not have the computational/intelligence capability of a controller to use complex algorithms. Therefore, time-invariant uniform quantization is a practically useful encoding rule for controller-actuator communication.

The investigation of the finite-action approximation problem is also useful in computing near optimal policies and learning algorithms for MDPs. In the next chapter, we will consider the development of *finite-state* approximations for obtaining near optimal policies. However, to establish constructive control schemes, one needs to quantize the action space as well. Thus, results on approximate optimality of finite-action models pave the way for practical computation algorithms which are commonly used for finite-state/action MDPs. These include deterministic algorithms such as value iteration and policy iteration [54], as well as stochastic algorithms such as Q-learning [96], among other algorithms [24, 60]. One other application regarding approximation problems is on learning a controlled Markov chain using simulations. If one can ensure that learning a control model with only finitely many control actions is sufficient for approximate optimality, then it is easier to develop efficient learning algorithms which allow for the approximate computation of finitely many transition probabilities. In particular, results developed in the learning and information theory literature for conditional kernel estimations [49] (with control-free models) can be applied to transition probability estimation for MDPs.

Motivated as above, in this chapter we investigate the following approximation problem: For uncountable Borel state and action spaces, under what conditions can the optimal performance (achieved by some optimal stationary policy) be arbitrarily well approximated if the controller action set is restricted to be finite?

This problem will be treated for two cases: (i) MDPs with strongly continuous transition probability and (ii) MDPs with weakly continuous transition probability. Under further assumptions, we also obtain explicit performance bounds on the approximation error in terms of the cardinality of the finite action set.

Various approximation results, which are somewhat related to our work in this chapter, have been established for MDPs with Borel state and action spaces in the literature along the theme of computation of near optimal policies. With the exception of [62], these works assume in general more restrictive continuity conditions on the transition probability than our conditions. In [62], the authors considered an approximation problem in which all the components of the original model are *allowed* to vary in the reduced model (varying only the action space corresponds to the setup considered in this chapter). Under weak continuity of the transition probability, [62] established the convergence of the reduced models to the original model for the discounted cost when the one-stage cost function is bounded. In this chapter we study the approximation problem for two different continuity assumptions (strong and weak continuity) on the transition probability and allow the one-stage cost function to be unbounded under weak continuity assumption. In addition, we also consider the approximation problem for the challenging average cost case. Hence, our results can be applied to a wider range of stochastic systems.

To prove the approximation result under the weak continuity assumption, we approximate the optimality operators associated with the dynamic programming equations characterizing optimality of deterministic stationary policies, which is a standard technique in approximate dynamic programming. However, for the strong continuity case, we develop a novel technique in which we approximate strategic measures induced by policies on the infinite product of state and action spaces. One advantage of this technique is that it can be used to study approximation problem for constrained Markov decision processes for which the dynamic programming principle no longer hold.

In the following section, we first review the definition of discrete time Markov decision processes.

4.2 Formal Definition of Markov Decision Process

A discrete-time Markov decision process (MDP) can be described by a five-tuple

$$\big(\mathsf{X}, \mathsf{A}, \{\mathsf{A}(x) : x \in \mathsf{X}\}, p, c\big),$$

where Borel spaces X and A denote the *state* and *action* spaces, respectively. The collection $\{\mathsf{A}(x) : x \in \mathsf{X}\}$ is a family of nonempty subsets $\mathsf{A}(x)$ of A, which gives the admissible actions for the state $x \in \mathsf{X}$. The *stochastic kernel* $p(\,\cdot\,|x,a)$ denotes the *transition probability* of the next state given that previous state-action pair is (x,a) [54]. Hence, it satisfies: (i) $p(\,\cdot\,|x,a)$ is an element of $\mathcal{P}(\mathsf{X})$ for all (x,a), and (ii) $p(D|\,\cdot\,,\,\cdot\,)$ is a measurable function from $\mathsf{X} \times \mathsf{A}$ to $[0,1]$ for each $D \in \mathcal{B}(\mathsf{X})$. The *one-stage cost* function c is a measurable function from $\mathsf{X} \times \mathsf{A}$ to $[0,\infty)$. In the remainder of this thesis, it is assumed that $\mathsf{A}(x) = \mathsf{A}$ for all $x \in \mathsf{X}$.

Define the history spaces $\mathsf{H}_0 = \mathsf{X}$ and $\mathsf{H}_t = (\mathsf{X} \times \mathsf{A})^t \times \mathsf{X}$, $t = 1, 2, \ldots$ endowed with their product Borel σ-algebras generated by $\mathcal{B}(\mathsf{X})$ and $\mathcal{B}(\mathsf{A})$. A *policy* is a sequence $\pi = \{\pi_t\}$ of stochastic kernels on A given H_t. The set of all policies is denoted by Π. Let Φ denote the set of stochastic kernels φ on A given X, and let $\mathbb{F}$ denote the set of all measurable functions f from X to A. A *randomized Markov* policy is a sequence $\pi = \{\pi_t\}$ of stochastic kernels on A given X. A *deterministic Markov* policy is a sequence of stochastic kernels $\pi = \{\pi_t\}$ on A given X such that $\pi_t(\,\cdot\,|x) = \delta_{f_t(x)}(\,\cdot\,)$ for some $f_t \in \mathbb{F}$. The set of randomized and deterministic Markov policies are denoted by RM and M, respectively. A *randomized stationary* policy is a constant sequence $\pi = \{\pi_t\}$ of stochastic kernels on A given X such that $\pi_t(\,\cdot\,|x) = \varphi(\,\cdot\,|x)$ for all t for some $\varphi \in \Phi$. A *deterministic stationary* policy is a constant sequence of stochastic kernels $\pi = \{\pi_t\}$ on A given X such that $\pi_t(\,\cdot\,|x) = \delta_{f(x)}(\,\cdot\,)$ for all t for some $f \in \mathbb{F}$. The set of randomized and deterministic stationary policies are identified with the sets Φ and $\mathbb{F}$, respectively.

According to the Ionescu Tulcea theorem [54], an initial distribution μ on X and a policy π define a unique probability measure P_μ^π on $\mathsf{H}_\infty = (\mathsf{X} \times \mathsf{A})^\infty$, which is called a *strategic measure* [38]. Hence, for any μ and any policy π we can think of MDP as a stochastic process $\{X_t, A_t\}_{t \geq 0}$ defined on a probability space $\big(\mathsf{H}_\infty, \mathcal{B}(\mathsf{H}_\infty), P_\mu^\pi\big)$ where X_t is a X-valued random variable, A_t is a A-valued random variable and P_μ^π-almost surely they satisfy

$$\begin{aligned} P_\mu^\pi(X_0 \in \,\cdot\,) &= \mu(\,\cdot\,) \\ P_\mu^\pi(X_t \in \,\cdot\,|X_{[0,t-1]}, A_{[0,t-1]}) &= P_\mu^\pi(X_t \in \,\cdot\,|X_{t-1}, A_{t-1}) = p(\,\cdot\,|X_{t-1}, A_{t-1}) \\ P_\mu^\pi(A_t \in \,\cdot\,|X_{[0,t]}, A_{[0,t-1]}) &= \pi_t(\,\cdot\,|X_{[0,t]}, A_{[0,t-1]}) \end{aligned}$$

where $X_{[0,t]} = (X_0, \ldots, X_t)$ and $A_{[0,t]} = (A_0, \ldots, A_t)$ $(t \geq 1)$. In this chapter, we sometimes use the following notation. For any π and initial distribution μ, we let $\lambda_t^{\pi,\mu}$, $\lambda_{(t)}^{\pi,\mu}$ and $\gamma_t^{\pi,\mu}$ to denote the law of X_t, $(X_0, \ldots, X_t)$ and (X_t, A_t) for all $t \geq 0$, respectively, i.e.,

$$X_t \sim \lambda_t^{\pi,\mu}$$

$$
\begin{aligned}
(X_0, \ldots, X_t) &\sim \lambda_{(t)}^{\pi,\mu} \\
(X_t, A_t) &\sim \gamma_t^{\pi,\mu}.
\end{aligned}
$$

If $\mu = \delta_x$, we replace μ with x in $\lambda_t^{\pi,\mu}$, $\lambda_{(t)}^{\pi,\mu}$ and $\gamma_t^{\pi,\mu}$.

The expectation with respect to P_μ^π is denoted by $\mathbb{E}_\mu^\pi$. If $\mu = \delta_x$, we write P_x^π and $\mathbb{E}_x^\pi$ instead of $P_{\delta_x}^\pi$ and $\mathbb{E}_{\delta_x}^\pi$. The cost functions to be minimized in this thesis are the β-discounted cost and the average cost, respectively given by

$$
\begin{aligned}
J(\pi, x) &= \mathbb{E}_x^\pi \biggl[\sum_{t=0}^{\infty} \beta^t c(X_t, A_t) \biggr], \\
V(\pi, x) &= \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_x^\pi \biggl[\sum_{t=0}^{T-1} c(X_t, A_t) \biggr].
\end{aligned}
$$

With this notation, the discounted and average value functions of the control problem are defined as

$$
\begin{aligned}
J^*(x) &\coloneqq \inf_{\pi \in \Pi} J(\pi, x), \\
V^*(x) &\coloneqq \inf_{\pi \in \Pi} V(\pi, x).
\end{aligned}
$$

A policy π^* is said to be optimal if $J(\pi^*, x) = J^*(x)$ (or $V(\pi^*, x) = V^*(x)$ for the average cost) for all $x \in \mathsf{X}$. Under fairly mild conditions, the set $\mathbb{F}$ of deterministic stationary policies contains an optimal policy for discounted cost (see, e.g., [54, 39]) and average cost optimal control problems (under somewhat stronger continuity/recurrence conditions, see, e.g., [39]).

The optimality of deterministic stationary policies is usually characterized by equalities or inequalities defined using optimality operators. For any real function

u on X, define the optimality operator T_β for $\beta \in (0,1]$ as

$$T_\beta u(x) := \min_{a \in \mathsf{A}} \left[c(x,a) + \beta \int_{\mathsf{X}} u(y) p(dy|x,a) \right] \tag{4.1}$$

In the literature T_β is called the *Bellman optimality operator*.

Using T_β, the discounted cost optimality equation (DCOE) for $\beta \in (0,1)$ is given by

$$J^*(x) = T_\beta J^*(x) \ \text{ for all } x \in \mathsf{X}; \tag{4.2}$$

that is, value function J^* of β-discounted cost is a fixed point of the optimality operator T_β. This fixed point equation, if the right hand side is well defined (i.e., measurable), always holds as a result of the principle of dynamic programming. A stationary policy $f^* : \mathsf{X} \to \mathsf{A}$ is discounted cost optimal if it attains the minimum in (4.2), i.e.,

$$T_\beta J^*(x) = c(x, f^*(x)) + \beta \int_{\mathsf{X}} J^*(y) p(dy|x, f^*(x)).$$

Under mild technical conditions, an optimal stationary policy f^* exists and attains the minimum in DCOE.

For the average cost criterion, the average cost optimality equation (ACOE) and inequality (ACOI) are respectively given by

$$\rho + h(x) = T_1 h(x) \tag{4.3}$$

$$\rho + h(x) \geq T_1 h(x), \tag{4.4}$$

where ρ is a scalar and h is a real function on X. A stationary policy f^* is average cost optimal if and only if it attains the minimum in either (4.3) or (4.4):

$$\rho + h(x) = c(x, f^*(x)) + \int_{\mathsf{X}} h(y)p(dy|x, f^*(x))$$
$$\rho + h(x) \geq c(x, f^*(x)) + \int_{\mathsf{X}} h(y)p(dy|x, f^*(x))$$

provided that $\lim_{t\to\infty} \mathbb{E}_x^\pi\big[h(X_t)|X_0 = x\big]/t = 0$ for all π and $x \in \mathsf{X}$. In this case we have $V^*(x) = \rho$ for all $x \in \mathsf{X}$; that is, the average value function equals to the scalar ρ for all initial points x. In the literature, the triplet (ρ, h, f^*) is called a *canonical triplet*. The existence of ρ and h satisfying ACOE or ACOI is in general established by the so-called 'vanishing discount' approach, in which the limiting behaviour (as $\beta \to 1$) of the discounted value function is used to show the existence of ρ and h.

4.2.1 The Problem of Quantizing the Action Space

To give a precise definition of the problem we study in this chapter, we first give the definition of a quantizer from the state to the action space.

Definition 4.1. *A measurable function* $q : \mathsf{X} \to \mathsf{A}$ *is called a* quantizer *from* X *to* A *if the range of* q*, i.e.,* $q(\mathsf{X}) = \{q(x) \in \mathsf{A} : x \in \mathsf{X}\}$*, is finite.*

The elements of $q(\mathsf{X})$ (the possible values of q) are called the *levels* of q. The rate $R = \log_2 |q(\mathsf{X})|$ of a quantizer q (approximately) represents the number of bits needed to losslessly encode the output levels of q using binary codewords of equal length. Let $\mathcal{Q}$ denote the set of all quantizers from X to A. A *deterministic stationary quantizer* policy is a constant sequence $\pi = \{\pi_t\}$ of stochastic kernels on A given X such that $\pi_t(\,\cdot\,|x) = \delta_{q(x)}(\,\cdot\,)$ for all t for some $q \in \mathcal{Q}$. For any finite set $\Lambda \subset \mathsf{A}$, let $\mathcal{Q}(\Lambda)$

denote the set of all elements in $\mathcal{Q}$ having range Λ. Analogous with $\mathbb{F}$, the set of all deterministic stationary quantizer policies induced by $\mathcal{Q}(\Lambda)$ will be identified with the set $\mathcal{Q}(\Lambda)$.

Our main objective in this chapter is to find conditions on the components of the MDP under which there exists a sequence of finite subsets $\{\Lambda_n\}_{n\geq 1}$ of A for which the following holds:

(P) For any initial point x, we have $\lim_{n\to\infty} \inf_{q\in\mathcal{Q}(\Lambda_n)} J(q,x) = \inf_{f\in\mathbb{F}} J(f,x)$ (or $\lim_{n\to\infty} \inf_{q\in\mathcal{Q}(\Lambda_n)} V(q,x) = \inf_{f\in\mathbb{F}} V(f,x)$ for the average cost), provided that the set $\mathbb{F}$ of deterministic stationary policies is an optimal class for the MDP.

In other words, if for each n, MDP_n is defined as the Markov decision process having the components $\{\mathsf{X},\Lambda_n,p,c\}$, then **(P)** is equivalent to stating that value function of MDP_n converges to the value function of the original MDP.

4.3 Near Optimality of Quantized Policies Under Strong Continuity

In this section we consider the problem **(P)** for the MDPs with strongly continuous transition probability. We impose the assumptions below on the components of the Markov decision process; additional assumptions will be made for the average cost problem in Section 4.3.2.

Assumption 4.1.

(a) The one stage cost function c is nonnegative and bounded satisfying $c(x,\,\cdot\,)\in C_b(\mathsf{A})$ for all $x\in\mathsf{X}$.

(b) The stochastic kernel $p(\,\cdot\,|x,a)$ is setwise continuous in $a\in\mathsf{A}$, i.e., if $a_k\to a$, then $p(\,\cdot\,|x,a_k)\to p(\,\cdot\,|x,a)$ setwise for all $x\in\mathsf{X}$.

(c) A is compact.

Remark 4.1. Note that if X is countable, then $B(\mathsf{X}) = C_b(\mathsf{X})$ (X is endowed with the discrete topology) which implies the equivalence of setwise convergence and weak convergence. Hence, results developed in this section are applicable to the MDPs having weakly continuous, in the action variable, transition probabilities when the state space is countable.

Remark 4.2. Note that any MDP can be modeled by a discrete time dynamical system of the form $X_{t+1} = F(X_t, A_t, V_t)$, where the $\{V_t\}$ is a sequence of independent and identically distributed (i.i.d.) random variables with values in some space V and common distribution ν. In many applications, the function F has a well behaved structure and is in the form $F(x, a, v) = H(x, a)G(v)$ or $F(x, a, v) = H(x, a) + G(v)$, e.g., the *fisheries management model* [54, p. 5], the *cash balance model* [41], and the *Pacific halibut fisheries management model* [35]. In these systems, Assumption 4.1-(a) holds for common noise processes. For instance, if ν admits a continuous density, which is often the case in practice, then Assumption 4.1-(a) usually holds. We refer the reader to [41, Section 4] for a discussion on the relevance of the setwise continuity assumption on inventory control problems. In addition, the widely studied and practically important case of the additive noise system in our Example 4.1 in the next section also satisfies Assumption 4.1-(a).

We now define the ws^∞ topology on $\mathcal{P}(\mathsf{H}_\infty)$ which was first introduced by Schäl in [91]. Recall that for any $\nu \in \mathcal{P}(\mathsf{E})$ and measurable real function g on E, we define $\nu(g) := \int g d\nu$, where E is some metric space. Let $\mathcal{C}(\mathsf{H}_0) = B(\mathsf{X})$ and let $\mathcal{C}(\mathsf{H}_t)$ $(t \geq 1)$ be the set of real valued functions g on H_t such that $g \in B(\mathsf{H}_t)$ and $g(x_0, \cdot, x_1, \cdot, \ldots, x_{t-1}, \cdot, x_t) \in C_b(\mathsf{A}^t)$ for all $(x_0, \ldots, x_t) \in \mathsf{X}^{t+1}$. The ws^∞ topology on $\mathcal{P}(\mathsf{H}_\infty)$ is defined as the smallest topology which renders all mappings $P \mapsto P(g)$,

$g \in \bigcup_{t=0}^{\infty} \mathcal{C}(\mathsf{H}_t)$, continuous.

Let d_{A} denote the metric on A. Since the action space A is compact and thus totally bounded, one can find a sequence of finite sets $\Lambda = \{a_{n,1}, \ldots, a_{n,k_n}\} \subset \mathsf{A}$ such that for all n,

$$\min_{i \in \{1,\ldots,k_n\}} d_{\mathsf{A}}(a, a_{n,i}) < 1/n \text{ for all } a \in \mathsf{A}.$$

In other words, Λ_n is a $1/n$-net in A. In the rest of this chapter, we assume that the sequence $\{\Lambda_n\}_{n\geq 1}$ is fixed. To ease the notation in the sequel, let us define the mapping $\Upsilon_n : \mathbb{F} \to \mathcal{Q}(\Lambda_n)$ as

$$\Upsilon_n(f)(x) = \arg\min_{a \in \Lambda_n} d_{\mathsf{A}}(f(x), a), \tag{4.5}$$

where ties are broken so that $\Upsilon_n(f)(x)$ is measurable. Hence, for all $f \in \mathbb{F}$, we have

$$\sup_{x \in \mathsf{X}} d_{\mathsf{A}}\big(\Upsilon_n(f)(x), f(x)\big) < 1/n; \tag{4.6}$$

that is $\Upsilon_n(f)$ converges *uniformly* to f as $n \to \infty$.

4.3.1 Discounted Cost

In this section we consider the problem **(P)** for the discounted cost with a discount factor $\beta \in (0,1)$ under the Assumption 4.1. Since the one-stage cost function c is bounded, the discounted cost satisfies the following

$$\sup_{f \in \mathbb{F}} \sum_{t=T+1}^{\infty} \beta^t \gamma_t^{f,x}(c) \to 0 \quad \text{as} \quad T \to \infty. \tag{4.7}$$

The following theorem is the main result of this section which states that for any

$f \in \mathbb{F}$, the discounted cost function of $\Upsilon_n(f) \in \mathcal{Q}(\Lambda_n)$ converges to the discounted cost function of f as $n \to \infty$ which implies that the discounted value function of the MDP$_n$ converges to the discounted value function of the original MDP.

Theorem 4.1. *Let $f \in \mathbb{F}$ and $\{\Upsilon_n(f)\}$ be the quantized approximations of f. Then, $J(\Upsilon_n(f), x) \to J(f, x)$ as $n \to \infty$, for all $x \in \mathsf{X}$.*

The proof of Theorem 4.1 requires the following proposition which is proved in Section 4.7.1.

Proposition 4.1. *Suppose Assumption 4.1-(b),(c) hold. Then for any $f \in \mathbb{F}$, the strategic measures $\{P_x^{\Upsilon_n(f)}\}$ induced by the quantized approximations $\{\Upsilon_n(f)\}$ of f converge to the strategic measure P_x^f of f in the ws^∞ topology, for all $x \in \mathsf{X}$.*

Proof of Theorem 4.1. Let $\Upsilon_n(f) = q_n$. Proposition 4.1 implies that $\gamma_t^{q_n,x}(c) \to \gamma_t^{f,x}(c)$ as $n \to \infty$ for all t. Then, we have

$$\limsup_{n\to\infty} |J(q_n, x) - J(f, x)| \leq \limsup_{n\to\infty} \sum_{t=0}^{\infty} |\gamma_t^{q_n,x}(c) - \gamma_t^{f,x}(c)|$$
$$\leq \lim_{n\to\infty} \sum_{t=0}^{T} |\gamma_t^{q_n,x}(c) - \gamma_t^{f,x}(c)| + 2 \sup_{f' \in \mathbb{F}} \sum_{t=T+1}^{\infty} \gamma_t^{f',x}(c)$$

Since the first and second terms in the last expression converge to zero as $T \to \infty$ by Proposition 4.1 and (4.7), respectively, the proof is complete. □

Remark 4.3. Notice that this proof implicitly shows that J is sequentially continuous with respect to the strategic measures in the ws^∞ topology.

The following is a generic example frequently considered in the theory of Markov decision processes (see [58]).

Example 4.1. Let us consider an additive-noise system given by

$$X_{t+1} = F(X_t, A_t) + V_t,\ t = 0, 1, 2, \ldots$$

where $\mathsf{X} = \mathbb{R}^m$ and the $\{V_t\}$ is a sequence of independent and identically distributed (i.i.d.) random vectors whose common distribution has a continuous, bounded, and strictly positive probability density function. A non-degenerate Gaussian distribution satisfies this condition. We assume that the action space A is a compact subset of $\mathbb{R}^d$ for some $d \geq 1$, the one stage cost function c satisfies Assumption 4.1-(a), and $F(x, \cdot)$ is continuous for all $x \in \mathsf{X}$. It is straightforward to show that Assumption 4.1-(b) holds under these conditions. Hence, Theorem 4.1 holds for this system.

4.3.2 Average Cost

In contrast to the discounted cost criterion, the expected average cost is in general not sequentially continuous with respect to strategic measures for the ws^∞ topology under practical assumptions. Hence, in this section we develop an approach based on the convergence of the sequence of invariant probability measures under quantized stationary policies to solve **(P)** for the average cost criterion.

Observe that any deterministic stationary policy f defines a stochastic kernel on X given X via

$$Q_f(\,\cdot\,|x) \coloneqq \lambda_1^{f,x}(\,\cdot\,) = p(\,\cdot\,|x, f(x)). \tag{4.8}$$

Let us write $Q_f g(x) \coloneqq \lambda_1^{f,x}(g)$. If Q_f admits an ergodic invariant probability measure ν_f, then by [56, Theorem 2.3.4 and Proposition 2.4.2], there exists an invariant set with full ν_f measure such that for all x in that set we have

$$
\begin{aligned}
J(f,x) &= \limsup_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \gamma_t^{f,x}(c) \\
&= \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \lambda_t^{f,x}(c_f) = \nu_f(c_f), \qquad (4.9)
\end{aligned}
$$

where $c_f(x) \coloneqq c(x, f(x))$. Let $\mathsf{M}_f \in \mathcal{B}(\mathsf{X})$ be the set of all $x \in \mathsf{X}$ such that convergence in (4.9) holds. Hence, $\nu_f(\mathsf{M}_f) = 1$ if ν_f exists. The following assumptions will be imposed in this section.

Assumption 4.2. Suppose Assumption 4.1 holds. In addition, we have

(e) For any $f \in \mathbb{F}$, Q_f has a unique invariant probability measure ν_f.

(f1) The set $\Gamma_{\mathbb{F}} \coloneqq \{\nu \in \mathcal{P}(\mathsf{X}) : \nu Q_f = \nu \text{ for some } f \in \mathbb{F}\}$ is relatively sequentially compact in the setwise topology.

(f2) There exists $x \in \mathsf{X}$ such that for all $B \in \mathcal{B}(\mathsf{X})$, $\lambda_t^{f,x}(B) \to \nu_f(B)$ uniformly in $f \in \mathbb{F}$.

(g) $\mathsf{M} \coloneqq \bigcap_{f\in\mathbb{F}} \mathsf{M}_f \neq \emptyset$.

The following theorem is the main result of this section. It states that for any $f \in \mathbb{F}$, the average cost function of $\Upsilon_n(f) \in \mathcal{Q}(\Lambda_n)$ converges to the average cost function of f as $n \to \infty$. In other words, the average value function of MDP_n converges to the average value function of the original MDP.

Theorem 4.2. *Let $x \in \mathsf{M}$. Then, we have $V(\Upsilon_n(f), x) \to V(f, x)$ as $n \to \infty$, under the Assumption 4.2 with (f1) or (f2).*

Proof. See Section 4.7.2. □

In the rest of this section we will derive conditions under which the conditions in Assumption 4.2 hold. To begin with, Assumption 4.2-(e),(f2),(g) are satisfied under any of the conditions Ri, $i \in \{0, 1, 1(a), 1(b), 2, \ldots, 6\}$ in [57]. Moreover, $\mathsf{M} = \mathsf{X}$ in Assumption 4.2-(g) if at least one of the above conditions holds. The next step is to find sufficient conditions for Assumption 4.2-(e),(f2),(g) to hold.

Observe that the stochastic kernel p on X given $\mathsf{X} \times \mathsf{A}$ can be written as a measurable mapping from $\mathsf{X} \times \mathsf{A}$ to $\mathcal{P}(\mathsf{X})$ if $\mathcal{P}(\mathsf{X})$ is equipped with its Borel σ-algebra generated by the weak topology [56], i.e., $p(\,\cdot\,|x, a) : \mathsf{X} \times \mathsf{A} \to \mathcal{P}(\mathsf{X})$. We impose the following assumption:

(e1) $p(\,\cdot\,|x, a) \leq \zeta(\,\cdot\,)$ for all $x \in \mathsf{X}$, $a \in \mathsf{A}$ for some finite measure ζ on X.

Proposition 4.2. *Suppose (e1) holds. Then, for any $f \in \mathbb{F}$, Q_f has an invariant probability measure ν_f. Furthermore, $\Gamma_{\mathbb{F}}$ is sequentially relatively compact in the setwise topology. Hence, (e1) implies Assumption 4.2-(f1). In addition, if these invariant measures are unique, then Assumption 4.2-(e),(g) also hold with $\mathsf{M} = \mathsf{X}$ in Assumption 4.2-(g).*

Proof. For any $f \in \mathbb{F}$, define $Q_{f,x}^{(T)}(\,\cdot\,) \coloneqq \frac{1}{T} \sum_{t=0}^{T-1} \lambda_t^{f,x}(\,\cdot\,)$ for some $x \in \mathsf{X}$. Clearly, $Q_{f,x}^{(T)} \leq \zeta$ for all T. Hence, by [56, Corollary 1.4.5] there exists a subsequence $\{Q_{f,x}^{(T_k)}\}$ which converges to some probability measure ν_f setwise. Following the same steps in [50, Theorem 4.17] one can show that $\nu_f(g) = \nu_f(Q_f g)$, for all $g \in B(\mathsf{X})$. Hence, ν_f is an invariant probability measure for Q_f.

Furthermore, (e1) implies that $\nu_f \leq \zeta$ for all $\nu_f \in \Gamma_{\mathbb{F}}$. Thus, $\Gamma_{\mathbb{F}}$ is relatively sequentially compact in the setwise topology by again [56, Corollary 1.4.5].

Finally, for any f, if the invariant measure ν_f is unique, then every setwise convergent subsequence of the relatively sequentially compact sequence $\{Q_{f,x}^{(T)}\}$ must converge to ν_f. Hence, $Q_{f,x}^{(T)} \to \nu_f$ setwise which implies that $V(f,x) = \lim_{T\to\infty} Q_{f,x}^{(T)}(c_f) = \nu_f(c_f)$ for all $x \in \mathsf{X}$ since $c_f(z) \coloneqq c(z, f(z)) \in B(\mathsf{X})$. Thus, $\mathsf{M} = \mathsf{X}$ in Assumption 4.2-(g). $\square$

Example 4.2. Let us consider an additive-noise system in Example 4.1 with the same assumptions. Furthermore, we assume F is bounded. Observe that for any $f \in \mathbb{F}$, if Q_f has an invariant probability measure, then it has to be unique [56, Lemma 2.2.3] since there cannot exist disjoint invariant sets due to the positivity of g. Since this system satisfies (e1) and $R1(a)$ in [57] due to the boundedness of F, Assumption 4.2-(e),(f1),(f2),(g) hold with $\mathsf{M} = \mathsf{X}$. This means that Theorem 4.2 holds for an additive noise system under the above conditions.

4.4 Near Optimality of Quantized Policies Under Weak Continuity Conditions

In this section we consider **(P)** for the MDPs with weakly continuous transition probability. Specifically, we will show that the value function of MDP$_n$ converges to the value function of the original MDP, which is equivalent to **(P)**.

An important motivation for replacing strong continuity with the weak continuity comes from the fact that for the fully observed reduction of a partially observed MDP (POMDP), the setwise continuity of the transition probability in the action variable is a prohibitive condition even for simple systems such as the one described below. We refer the reader to [52, Chapter 4] and Section 4.4.3 of this chapter for the basics of POMDPs.

Example 4.3. Consider the system dynamics

$$X_{t+1} = X_t + A_t,$$
$$Y_t = X_t + V_t,$$

where $X_t \in \mathsf{X}$, $Y_t \in \mathsf{Y}$ and $A_t \in \mathsf{A}$, and where X, Y and A are the state, observation and action spaces, respectively. We assume that $\mathsf{X} = \mathsf{Y} = \mathsf{A} = \mathbb{R}_+$ and the 'noise' $\{V_t\}$ is a sequence of i.i.d. random variables uniformly distributed on $[0, 1]$. Here, the stochastic kernel $\mathsf{Pr}\{Y_t \in \cdot\,|X_t = x\} \coloneqq r(\cdot\,|x)$ is called the observation channel. It is easy to see that the transition probability is weakly continuous (with respect to state-action variables) and the observation channel is continuous in total variation (with respect to state variable) for this POMDP. Hence, by [40, Theorem 3.7] the transition probability, denoted as η, of the fully observed reduction of the POMDP is weakly continuous in the state-action variables. However, the same conclusion cannot be drawn for the setwise continuity of η with respect to the action variable as shown below.

Let z denote the generic state variable for the fully observed MDP, where the state variables are elements of $\mathcal{P}(\mathsf{X})$ which is equipped with the Borel σ-algebra generated by the topology of weak convergence. If we define the function $F(z, a, y) \coloneqq \Pr\{X_{t+1} \in \cdot\,|Z_t = z, A_t = a, Y_{t+1} = y\}$ from $\mathcal{P}(\mathsf{X}) \times \mathsf{A} \times \mathsf{Y}$ to $\mathcal{P}(\mathsf{X})$ and the stochastic kernel $H(\cdot\,|z, a) \coloneqq \Pr\{Y_{t+1} \in \cdot\,|Z_t = z, A_t = a\}$ on Y given $\mathcal{P}(\mathsf{X}) \times \mathsf{A}$, then η can be written as

$$\eta(\cdot\,|z, a) = \int_{\mathsf{Y}} 1_{\{F(z,a,y) \in \cdot\}} H(dy|z, a),$$

where $Z_t \in \mathcal{P}(\mathsf{X})$ denotes the posterior distribution of the state X_t given the past observations, i.e.,

$$Z_t(\,\cdot\,) \coloneqq \Pr\{X_t \in \,\cdot\,|Y_0, \ldots, Y_t, A_0, \ldots, A_{t-1}\}.$$

Let us set $z = \delta_0$ (point mass at $0 \in \mathsf{X}$), $\{a_k\} = \{\frac{1}{k}\}$, and $a = 0$. Hence, $a_k \to a$. We show that $\eta(\,\cdot\,|z, a_k) \nrightarrow \eta(\,\cdot\,|z, a)$ setwise.

Observe that for all k and $y \in \mathsf{Y}$, we have $F(z, a_k, y) = \delta_{\frac{1}{k}}$ and $F(z, a, y) = \delta_0$. Define the open set O with respect to the weak topology in $\mathcal{P}(\mathsf{X})$ as

$$O \coloneqq \Big\{z \in \mathcal{P}(\mathsf{X}) : \Big|\int_{\mathsf{X}} g(x)\delta_1(dx) - \int_{\mathsf{X}} g(x)z(dx)\Big| < 1\Big\},$$

where g is the symmetric triangular function between $[-1, 1]$ with $g(0) = 1$. Observe that we have $F(z, a_k, y) \in O$ for all k and y, but $F(z, a, y) \notin O$ for all y. Hence,

$$\eta(O|z, a_k) \coloneqq \int_{\mathsf{Y}} 1_{\{F(z,a_k,y)\in O\}} H(dy|z, a_k) = 1,$$

but

$$\eta(O|z, a) \coloneqq \int_{\mathsf{Y}} 1_{\{F(z,a,y)\in O\}} H(dy|z, a) = 0.$$

This means that $\eta(\,\cdot\,|z, a_k) \nrightarrow \eta(\,\cdot\,|z, a)$ setwise. Hence, η does not satisfy the setwise continuity assumption.

4.4.1 Discounted Cost

In this section we consider the problem **(P)** for the discounted cost with a discount factor $\beta \in (0,1)$ for MDPs with weakly continuous transition probability. The following assumptions will be imposed for both the discounted cost and the average cost. These assumptions are used in the literature for studying discounted Markov decision processes with unbounded one-stage cost and weakly continuous transition probability.

Assumption 4.3.

(a) The one stage cost function c is nonnegative and continuous.

(b) The stochastic kernel $p(\,\cdot\,|x,a)$ is weakly continuous in $(x,a) \in \mathsf{X} \times \mathsf{A}$, i.e., if $(x_k, a_k) \to (x,a)$, then $p(\,\cdot\,|x_k,a_k) \to p(\,\cdot\,|x,a)$ weakly.

(c) A is compact.

(d) There exist nonnegative real numbers M and $\alpha \in [1, \frac{1}{\beta})$, and a continuous weight function $w : \mathsf{X} \to [1,\infty)$ such that for each $x \in \mathsf{X}$, we have

$$\sup_{a \in \mathsf{A}} c(x,a) \leq Mw(x), \tag{4.10}$$

$$\sup_{a \in \mathsf{A}} \int_{\mathsf{X}} w(y) p(dy|x,a) \leq \alpha w(x), \tag{4.11}$$

and $\int_{\mathsf{X}} w(y)p(dy|x,a)$ is continuous in (x,a).

Recall the Bellman optimality operator T_β defined in (4.1). Since β is fixed here, we will write T instead of T_β. Therefore, for any real-valued measurable function u

on X, $Tu : \mathsf{X} \to \mathbb{R}$ is given by

$$Tu(x) \coloneqq \min_{a \in \mathsf{A}} \left[c(x,a) + \beta \int_{\mathsf{X}} u(y) p(dy|x,a) \right]. \tag{4.12}$$

Recall also that $C_w(\mathsf{X})$ and $B_w(\mathsf{X})$ denote the set of all real valued continuous and measurable functions on X with finite w-norm, respectively.

Lemma 4.1. *For any $u \in C_w(\mathsf{X})$ the function $l_u(x,a) \coloneqq \int_{\mathsf{X}} u(y) p(dy|x,a)$ is continuous in (x,a).*

Proof. For any nonnegative continuous function g on X, the function $l_g(x,a) = \int_{\mathsf{X}} g(y) p(dy|x,a)$ is lower semi-continuous in (x,a), if p is weakly continuous (see, e.g., [54, Proposition E.2]). Define the nonnegative continuous function g by letting $g = bw + u$, where $b = \|u\|_w$. Then l_g is lower semi-continuous. Since $l_u = l_g - bl_w$ and l_w is continuous by Assumption 4.3-(d), l_u is lower semi-continuous. Analogously, define the nonnegative continuous function v by letting $v = -u + bw$. Then l_v is lower semi-continuous. Since $l_u = bl_w - l_v$ and l_w is continuous by Assumption 4.3-(d), l_u is also upper semi-continuous. Therefore, l_u is continuous. □

Lemma 4.2. *Let Y be any of the compact sets A or Λ_n. Define the operator T_{Y} on $B_w(\mathsf{X})$ by letting*

$$T_{\mathsf{Y}} u(x) \coloneqq \min_{a \in \mathsf{Y}} \left[c(x,a) + \beta \int_{\mathsf{X}} u(y) p(dy|x,a) \right].$$

Then T_{Y} maps $C_w(\mathsf{X})$ into itself. Moreover, $C_w(\mathsf{X})$ is closed with respect to the w-norm.

Proof. Note that $T_{\mathsf{Y}} u(x) = \min_{a \in \mathsf{Y}} \big(c(x,a) + \beta l_u(x,a) \big)$. The function l_u is continuous by Lemma 4.1, and therefore, $T_{\mathsf{Y}}u$ is also continuous by [9, Proposition 7.32]. Since T_{Y} maps $B_w(\mathsf{X})$ into itself, $T_{\mathsf{Y}}u \in C_w(\mathsf{X})$.

For the second statement, let u_n converge to u in w-norm in $C_w(\mathsf{X})$. It is sufficient to prove that u is continuous. Let $x_k \to x_0$. Since $B \coloneqq \{x_0, x_1, x_2, \ldots\}$ is compact, w is bounded on B. Therefore, $u_n \to u$ uniformly on B which implies that $\lim_{k\to\infty} u(x_k) = u(x_0)$. This completes the proof. □

Lemma 4.2 implies that T maps $C_w(\mathsf{X})$ into itself. It can also be proved that T is a contraction operator with modulus $\sigma \coloneqq \beta\alpha$ (see [55, Lemma 8.5.5]); that is,

$$\|Tu - Tv\|_w \leq \sigma \|u - v\|_w \text{ for all } u, v \in C_w(\mathsf{X}).$$

The following theorem is a known result in the theory of Markov decision processes (see e.g., [55, Section 8.5, p.65]).

Theorem 4.3. *Suppose Assumption 4.3 holds. Then, the discounted value function J^* is the unique fixed point in $C_w(\mathsf{X})$ of the contraction operator T, i.e.,*

$$J^* = TJ^*. \tag{4.13}$$

Furthermore, a deterministic stationary policy f^* is optimal if and only if

$$J^*(x) = c(x, f^*(x)) + \beta \int_{\mathsf{X}} J^*(y) p(dy|x, f^*(x)). \tag{4.14}$$

Finally, there exists a deterministic stationary policy f^* which is optimal, so it satisfies (4.14).

Define, for all $n \geq 1$, the operator T_n (which will be used to approximate T) by

$$T_n u(x) \coloneqq \min_{a \in \Lambda_n} \biggl[c(x,a) + \beta \int_{\mathsf{X}} u(y) p(dy|x,a) \biggr]. \tag{4.15}$$

Note that T_n is the Bellman optimality operator for MDP_n having components $\{\mathsf{X}, \Lambda_n, p, c\}$. Analogous with T, it can be shown that T_n is a contraction operator with modulus $\sigma = \alpha\beta$ mapping $C_w(\mathsf{X})$ into itself. Let $J_n^* \in C_w(\mathsf{X})$ (discounted value function of MDP_n) denote the fixed point of T_n.

The following theorem is the main result of this section which states that the discounted value function of MDP_n converges to the discounted value function of the original MDP.

Theorem 4.4. *For any compact set $K \subset \mathsf{X}$ we have*

$$\lim_{n \to \infty} \sup_{x \in K} |J_n^*(x) - J^*(x)| = 0. \tag{4.16}$$

Therefore,

$$\lim_{n \to \infty} |J_n^*(x) - J^*(x)| = 0 \ \ \textit{for all } x \in \mathsf{X}.$$

To prove Theorem 4.4, we need following results. They are proved in Section 4.7.3 and in Section 4.7.4, respectively.

Lemma 4.3. *For any compact subset K of X and for any $\varepsilon > 0$, there exists a compact subset K_ε of X such that*

$$\sup_{(x,a) \in K \times \mathsf{A}} \int_{K_\varepsilon^c} w(y) p(dy|x,a) < \varepsilon. \tag{4.17}$$

Lemma 4.4. *Let $\{u_n\}$ be a sequence in $C_w(\mathsf{X})$ with $\sup_n \|u_n\|_w := L < \infty$. If u_n converges to $u \in C_w(\mathsf{X})$ uniformly on each compact subset of X, then for any $f \in \mathbb{F}$ and compact subset K of X we have*

$$\lim_{n\to\infty} \sup_{x\in K} \biggl| \int_{\mathsf{X}} u_n(y) p(dy|x, f_n(x)) - \int_{\mathsf{X}} u(y) p(dy|x, f(x)) \biggr| = 0,$$

where $f_n = \Upsilon_n(f)$.

Let us define $v^0 = v_n^0 = 0$, and $v^{t+1} = T v^t$ and $v_n^{t+1} = T_n v_n^t$ for $t \geq 1$; that is, $\{v^t\}_{t\geq 1}$ and $\{v_n^t\}_{t\geq 1}$ are successive approximations to the discounted value functions of the original MDP and MDP$_n$, respectively. Lemma 4.2 implies that v^t and v_n^t are in $C_w(\mathsf{X})$ for all t and n. By [55, Theorem 8.3.6, p. 47], [55, (8.3.34), p. 52] and [55, Section 8.5, p. 65] we have

$$v^t(x) \leq J^*(x) \leq M \frac{w(x)}{1-\sigma}, \tag{4.18}$$

$$\|v^t - J^*\|_w \leq M \frac{\sigma^t}{1-\sigma}, \tag{4.19}$$

and

$$v_n^t(x) \leq J_n^*(x) \leq M \frac{w(x)}{1-\sigma}, \tag{4.20}$$

$$\|v_n^t - J_n^*\|_w \leq M \frac{\sigma^t}{1-\sigma}. \tag{4.21}$$

Since for each n and u, $Tu \leq T_n u$, we also have $v^t \leq v_n^t$ for all $t \geq 1$ and $J^* \leq J_n^*$.

Lemma 4.5. *For any compact set $K \subset \mathsf{X}$ and $t \geq 1$, we have*

$$\lim_{n\to\infty} \sup_{x\in K} |v_n^t(x) - v^t(x)| = 0. \tag{4.22}$$

Proof. We prove (4.22) by induction. For $t = 1$, the claim holds since $v^0 = v_n^0 = 0$, and c is uniformly continuous on $K \times \mathsf{A}$ for any compact K subset of X. Assume the claim is true for $t \geq 1$. We fix any compact set K. Let f_t^* denote the selector of $Tv^t = v^{t+1}$; that is,

$$v^{t+1}(x) = Tv^t(x) = c(x, f_t^*(x)) + \beta \int_{\mathsf{X}} v^t(y) p(dy|x, f_t^*(x)),$$

and let $f_{t,n}^* \coloneqq \Upsilon_n(f_t^*)$ (see (4.5)). By (4.18) and (4.20) we have

$$v^t(x) \leq M \frac{w(x)}{1-\sigma} \tag{4.23}$$

$$v_n^t(x) \leq M \frac{w(x)}{1-\sigma}, \tag{4.24}$$

for all t and n. For each $n \geq 1$, we have

$$\begin{aligned}
&\sup_{x\in K} \bigl(v_n^{t+1}(x) - v^{t+1}(x)\bigr) \quad (\text{as } v^{t+1} \leq v_n^{t+1}) \\
&= \sup_{x\in K} \biggl(\min_{\Lambda_n} \Bigl[c(x,a) + \beta \int_{\mathsf{X}} v_n^t(y) p(dy|x,a) \Bigr] - \min_{\mathsf{A}} \Bigl[c(x,a) + \beta \int_{\mathsf{X}} v^t(y) p(dy|x,a) \Bigr] \biggr) \\
&\leq \sup_{x\in K} \biggl(\Bigl[c(x, f_{t,n}^*(x)) + \beta \int_{\mathsf{X}} v_n^t(y) p(dy|x, f_{t,n}^*(x)) \Bigr] \\
&\qquad\qquad - \Bigl[c(x, f_t^*(x)) + \beta \int_{\mathsf{X}} v^t(y) p(dy|x, f_t^*(x)) \Bigr] \biggr) \\
&\leq \sup_{x\in K} \bigl| c(x, f_{t,n}^*(x)) - c(x, f_t^*(x)) \bigr|
\end{aligned}$$

$$+ \beta \sup_{x\in K} \biggl| \int_{\mathsf{X}} v_n^t(y) p(dy|x, f_{t,n}^*(x)) - \int_{\mathsf{X}} v^t(y) p(dy|x, f_t^*(x)) \biggr|$$

Note that in the last expression, as $n \to \infty$, the first term goes to zero since c is uniformly continuous on $K \times \mathsf{A}$ and $f_{t,n}^* \to f_t^*$ uniformly, and the second term goes to zero by Lemma 4.4, (4.23), and (4.24). □

Now, using Lemma 4.5 we prove Theorem 4.4.

Proof of Theorem 4.4. Let us fix any compact set $K \subset \mathsf{X}$. Since w is bounded on K, it is enough to prove $\lim_{n\to\infty} \sup_{x\in K} \frac{|J_n^*(x) - J^*(x)|}{w(x)}$. We have

$$\sup_{x\in K} \frac{|J_n^*(x) - J^*(x)|}{w(x)} \leq \sup_{x\in K} \frac{|J_n^*(x) - v_n^t(x)|}{w(x)} + \sup_{x\in K} \frac{|v_n^t(x) - v^t(x)|}{w(x)} + \sup_{x\in K} \frac{|v^t(x) - J^*(x)|}{w(x)}$$
$$\leq 2M \frac{\sigma^t}{1-\sigma} + \sup_{x\in K} \frac{|v_n^t(x) - v^t(x)|}{w(x)} \quad \text{(by (4.19) and (4.21))}.$$

Since $w \geq 1$, $\sup_{x\in K} \frac{|v_n^t(x) - v^t(x)|}{w(x)} \to 0$ as $n \to \infty$ for all t by Lemma 4.5. Hence, the last expression can be made arbitrarily small. Since $t \geq 1$ is arbitrary and $\sigma \in (0,1)$, this completes the proof. □

4.4.2 Average Cost

In this section we consider the problem **(P)** for the average cost criterion for MDPs with weakly continuous transition probability. We prove an approximation result analogous to Theorem 4.4. To do this, some new assumptions are needed on the components of the original MDP in addition to the conditions in Assumption 4.3. A version of these assumptions were used in [98] and [44] to study the existence of the solution to the Average Cost Optimality Equality (ACOE) and Inequality (ACOI).

Assumption 4.4. Suppose Assumption 4.3 holds with (4.11) replaced by condition (e) below. Moreover, suppose there exist a probability measure λ on X and a continuous function $\phi : \mathsf{X} \times \mathsf{A} \to [0, \infty)$ such that

(e) $\int_{\mathsf{X}} w(y) p(dy|x,a) \leq \alpha w(x) + \lambda(w)\phi(x,a)$ for all $(x,a) \in \mathsf{X} \times \mathsf{A}$, where $\alpha \in (0,1)$.

(f) $p(D|x,a) \geq \lambda(D)\phi(x,a)$ for all $(x,a) \in \mathsf{X} \times \mathsf{A}$ and $D \in \mathcal{B}(\mathsf{X})$.

(g) The weight function w is μ-integrable.

(h) $\int_{\mathsf{X}} \phi(x, f(x))\lambda(dx) > 0$ for all $f \in \mathbb{F}$.

Recall that any $f \in \mathbb{F}$ defines a time-homogenous Markov chain $\{X_t\}_{t=1}^{\infty}$ (state process) with the transition probability Q_f on X given X (see (4.8)). For any $t \geq 1$, let $Q_f^t(\,\cdot\,|x)$ denote the t-step transition probability of this Markov chain given the initial point x. Hence, $Q_f^t(\,\cdot\,|x)$ is recursively given by

$$Q_f^{t+1}(\,\cdot\,|x) = \int_{\mathsf{X}} Q_f(\,\cdot\,|y) Q_f^t(dy|x).$$

The following theorem is a consequence of [98, Theorems 3.3 and 3.6].

Theorem 4.5. *Under Assumption 4.4 the following holds.*

(i) For each $f \in \mathbb{F}$, the stochastic kernel $Q_f(\,\cdot\,|x)$ is positive Harris recurrent with unique invariant probability measure ν_f. Furthermore, w is ν_f-integrable, and therefore, $\rho_f \coloneqq \int_{\mathsf{X}} c(x, f(x))\nu_f(dx) < \infty$.

(ii) There exist $f^* \in \mathbb{F}$ and $h^* \in C_w(\mathsf{X})$ such that the triplet (h^*, f^*, ρ_{f^*}) satisfies the average cost optimality equality (ACOE), i.e.,

$$\rho_{f^*} + h^*(x) = \min_{a \in \mathsf{A}} \biggl[c(x,a) + \int_{\mathsf{X}} h^*(y) p(dy|x,a) \biggr]$$

$$= c(x, f^*(x)) + \int_{\mathsf{X}} h^*(y) p(dy|x, f^*(x)),$$

and therefore,

$$\inf_{\pi \in \Pi} V(\pi, x) \coloneqq V^*(x) = \rho_{f^*},$$

for all $x \in \mathsf{X}$.

Proof. The only statement that does not directly follow from [98, Theorems 3.3 and 3.6] is the fact: $h^* \in C_w(\mathsf{X})$. Hence, we only prove this.

By [98, Theorem 3.5], h^* is the unique fixed point of the following contraction operator with modulus α

$$Fu(x) \coloneqq \min_{a \in \mathsf{A}} \biggl[c(x, a) + \int_{\mathsf{X}} u(y) p(dy|x, a) - \lambda(u) \phi(x, a) \biggr].$$

Since ϕ is continuous, by Lemma 4.1 the function inside the minimization is continuous in (x, a) if $u \in C_w(\mathsf{X})$. Then by Lemma 4.2, F maps $C_w(\mathsf{X})$ into itself. Therefore, $h^* \in C_w(\mathsf{X})$. □

This theorem implies that for each $f \in \mathbb{F}$, the average cost is given by $V(f, x) = \int_{\mathsf{X}} c(y, f(y)) \nu_f(dy)$ for all $x \in \mathsf{X}$ (instead of ν_f-a.e.).

Note that all the statements in Theorem 4.5 are also valid for MDP_n with an optimal policy f_n^* and a canonical triplet $(h_n^*, f_n^*, \rho_{f_n^*})$. Analogous with F, define the contraction operator F_n (with modulus α) corresponding to MDP_n as

$$F_n u(x) \coloneqq \min_{a \in \Lambda_n} \biggl[c(x, a) + \int_{\mathsf{X}} u(y) p(dy|x, a) - \lambda(u) \phi(x, a) \biggr],$$

and therefore, $h_n^* \in C_w(\mathsf{X})$ is its fixed point.

The next theorem is the main result of this section which states that the average cost value function, denoted as V_n^*, of MDP_n converges to the average cost value function V^* of the original MDP.

Theorem 4.6. *We have*

$$\lim_{n\to\infty} |V_n^* - V^*| = 0,$$

where V_n^* and V^* are both constants.

Let us define $u^0 = u_n^0 = 0$, and $u^{t+1} = Fu^t$ and $u_n^{t+1} = F_n u_n^t$ for $t \geq 1$; that is, $\{u^t\}_{t\geq 1}$ and $\{u_n^t\}_{t\geq 1}$ are successive approximations to h^* and h_n^*, respectively. Lemma 4.2 implies that u^t and u_n^t are in $C_w(\mathsf{X})$ for all t and n.

Lemma 4.6. *For all $u, v \in C_w(\mathsf{X})$ and $n \geq 1$, the following results hold: (i) if $u \leq v$, then $Fu \leq Fv$ and $F_n u \leq F_n v$; (ii) $Fu \leq F_n u$.*

Proof. Define the sub-stochastic kernel $\hat{p}$ by letting

$$\hat{p}(\,\cdot\,|x,a) \coloneqq p(\,\cdot\,|x,a) - \lambda(\,\cdot\,)\phi(x,a).$$

Using $\hat{p}$, F and F_n can be written as

$$\begin{aligned} Fu(x) &\coloneqq \min_{a\in\mathsf{A}} \biggl[c(x,a) + \int_{\mathsf{X}} u(y)\hat{p}(dy|x,a) \biggr], \\ F_n u(x) &\coloneqq \min_{a\in\Lambda_n} \biggl[c(x,a) + \int_{\mathsf{X}} u(y)\hat{p}(dy|x,a) \biggr]. \end{aligned}$$

Then the results follow from the fact that $\hat{p}(\,\cdot\,|x,a) \geq 0$ by Assumption 4.4-(f). □

Lemma 4.6 implies that $u^0 \le u^1 \le u^2 \le \ldots \le h^*$ and $u_n^0 \le u_n^1 \le u_n^2 \le \ldots \le h_n^*$. Note that $\|u^1\|_w, \|u_n^1\|_w \le M$ by Assumption 4.3-(d). Since

$$\|h^*\|_w \le \|h^* - u^1\|_w + \|u^1\|_w = \|Fh^* - Fu^0\|_w + \|u^1\|_w \le \alpha\|h^*\|_w + \|u^1\|_w$$

$$\|h_n^*\|_w \le \|h_n^* - u_n^1\|_w + \|u_n^1\|_w = \|F_n h_n^* - F_n u_n^0\|_w + \|u_n^1\|_w \le \alpha\|h_n^*\|_w + \|u_n^1\|_w,$$

we have

$$u^t(x) \le h^*(x) \le M\frac{w(x)}{1-\alpha},$$

and

$$u_n^t(x) \le h_n^*(x) \le M\frac{w(x)}{1-\alpha}.$$

By inequalities above and the facts $\|u^t - h^*\|_w \le \alpha^t\|h\|_w$ and $\|u_n^t - h_n^*\|_w \le \alpha^t\|h_n\|_w$, we also have

$$\|u^t - h^*\|_w \le M\frac{\alpha^t}{1-\alpha},$$

and

$$\|u_n^t - h_n^*\|_w \le M\frac{\alpha^t}{1-\alpha}.$$

By Lemma 4.6, for each n and v, we have $Fv \le F_n v$. Therefore, by the monotonicity of F and the fact $u^0 = u_n^0 = 0$, we have

$$\begin{aligned} u^t &\leq u_n^t \\ h^* &\leq h_n^*, \end{aligned} \tag{4.25}$$

for all t and n.

Lemma 4.7. *For any compact set $K \subset \mathsf{X}$ and $t \geq 1$, we have*

$$\lim_{n\to\infty} \sup_{x\in K} |u_n^t(x) - u^t(x)| = 0. \tag{4.26}$$

Proof. Note that for each $t \geq 1$, by dominated convergence theorem and $\lambda(w) < \infty$, we have $\lambda(u_n^t) \to \lambda(u^t)$ if $u_n^t \to u^t$ pointwise. Then, the rest of the proof can be done with the same arguments used to prove Lemma 4.5 and so we omit the details. □

Lemma 4.8. *For any compact set $K \subset \mathsf{X}$, we have*

$$\lim_{n\to\infty} \sup_{x\in K} |h_n^*(x) - h^*(x)| = 0.$$

Proof. The lemma can be proved using the same arguments as in the proof of Theorem 4.4. □

Now, using Lemma 4.8 we prove Theorem 4.6

Proof of Theorem 4.6. Recall that $V^* = \rho_{f^*}$ and $V_n^* = \rho_{f_n^*}$, and they satisfy the following ACOEs:

$$h^*(x) + \rho_{f^*} = \min_{a\in\mathsf{A}} \biggl[c(x,a) + \int_{\mathsf{X}} h^*(y)p(dy|x,a)\biggr] = c(x,f^*(x)) + \int_{\mathsf{X}} h^*(y)p(dy|x,f^*(x))$$

$$h_n^*(x) + \rho_{f_n^*} = \min_{a \in \Lambda_n} \left[c(x,a) + \int_{\mathsf{X}} h_n^*(y) p(dy|x,a) \right] = c(x, f_n^*(x)) + \int_{\mathsf{X}} h_n^*(y) p(dy|x, f_n^*(x)).$$

Note that $h_n^* \geq h^*$ (see (4.25)) and $\rho_{f_n^*} \geq \rho_{f^*}$. For each n, let $f_n \coloneqq \Upsilon_n(f^*)$. Then for any $x \in \mathsf{X}$ we have

$$\begin{aligned}
\limsup_{n\to\infty} \big(h_n^*(x) + \rho_{f_n^*}\big) &= \limsup_{n\to\infty} \left(\min_{a \in \Lambda_n} \left[c(x,a) + \int_{\mathsf{X}} h_n^*(y) p(dy|x,a) \right] \right) \\
&= \limsup_{n\to\infty} \left(c(x, f_n^*(x)) + \int_{\mathsf{X}} h_n^*(y) p(dy|x, f_n^*(x)) \right) \\
&\leq \limsup_{n\to\infty} \left(c(x, f_n(x)) + \int_{\mathsf{X}} h_n^*(y) p(dy|x, f_n(x)) \right) \\
&= c(x, f^*(x)) + \int_{\mathsf{X}} h^*(y) p(dy|x, f^*(x)) \qquad (4.27) \\
&= h^*(x) + \rho_{f^*} \\
&\leq \liminf_{n\to\infty} \big(h_n^*(x) + \rho_{f_n^*}\big),
\end{aligned}$$

where (4.27) follows from Lemma 4.4 and the fact that h_n^* converges to h^* uniformly on each compact subset K of X and $\sup_n \|h_n^*\|_w \leq \frac{M}{1-\alpha}$. Since $\lim_{n\to\infty} h_n^*(x) = h^*(x)$ by Lemma 4.8, we have $\lim_{n\to\infty} \rho_{f_n^*} = \rho_{f^*}$. This completes the proof. □

4.4.3 Application to Partially Observed MDPs

In this section we apply the result obtained in Section 4.4.1 to partially observed Markov decision processes (POMDPs). Consider a discrete time POMDP with state space X, action space A, and observation space Y, all Borel spaces. Let $p(\,\cdot\,|x,a)$ denote the transition probability of the next state given the current state-action pair is (x,a), and let $r(\,\cdot\,|x)$ denote the transition probability of the current observation given the current state variable x. The one-stage cost function, denoted by $\tilde{c}$, is again a measurable function from $\mathsf{X} \times \mathsf{A}$ to $[0, \infty)$.

Define the history spaces $\tilde{\mathsf{H}}_t = (\mathsf{Y} \times \mathsf{A})^t \times \mathsf{Y}$, $t = 0, 1, 2, \ldots$ endowed with their product Borel σ-algebras generated by $\mathcal{B}(\mathsf{Y})$ and $\mathcal{B}(\mathsf{A})$. A *policy* $\pi = \{\pi_t\}$ is a sequence of stochastic kernels on A given $\tilde{\mathsf{H}}_t$. We denote by Π the set of all policies. Hence, for any initial distribution μ and policy π we can think of POMDP as a stochastic process $\{X_t, Y_t, A_t\}_{t \geq 0}$ defined on a probability space $(\Omega, \mathcal{B}(\Omega), P_\mu^\pi)$ where $\Omega = \tilde{\mathsf{H}}_\infty \times \mathsf{X}^\infty$, X_t is a X-valued random variable, Y_t is a Y-valued random variable, A_t is a A-valued random variable, and P_μ^π-almost surely they satisfy

$$
\begin{aligned}
P_\mu^\pi(X_0 \in \,\cdot\,) &= \mu(\,\cdot\,) \\
P_\mu^\pi(X_t \in \,\cdot\,|X_{[0,t-1]}, Y_{[0,t-1]}, A_{[0,t-1]}) &= P_\mu^\pi(X_t \in \,\cdot\,|X_{t-1}, A_{t-1}) = p(\,\cdot\,|X_{t-1}, A_{t-1}) \\
P_\mu^\pi(Y_t \in \,\cdot\,|X_{[0,t]}, Y_{[0,t-1]}, A_{[0,t-1]}) &= P_\mu^\pi(Y_t \in \,\cdot\,|X_t) = r(\,\cdot\,|X_t) \\
P_\mu^\pi(A_t \in \,\cdot\,|X_{[0,t]}, Y_{[0,t]}, A_{[0,t-1]}) &= \pi_t(\,\cdot\,|Y_{[0,t]}, A_{[0,t-1]})
\end{aligned}
$$

where $X_{[0,t]} = (X_0, \ldots, X_t)$, $Y_{[0,t]} = (Y_0, \ldots, Y_t)$, and $A_{[0,t]} = (A_0, \ldots, A_t)$ $(t \geq 1)$. Let $\tilde{J}(\pi, \mu)$ denote the discounted cost function of the policy $\pi \in \Pi$ with initial distribution μ of the POMDP.

It is known that any POMDP can be reduced to a (completely observable) MDP [109], [78], whose states are the posterior state distributions or beliefs of the observer; that is, the state at time t is

$$
\mathsf{Pr}\{X_t \in \,\cdot\,|Y_0, \ldots, Y_t, A_0, \ldots, A_{t-1}\} \in \mathcal{P}(\mathsf{X}).
$$

We call this equivalent MDP the belief-MDP. The belief-MDP has state space $\mathsf{Z} = \mathcal{P}(\mathsf{X})$ and action space A. The transition probability η of the belief-MDP can be constructed as in Example 4.3 (see also [52])

$$\eta(\,\cdot\,|z,a) = \int_{\mathsf{Y}} 1_{\{F(z,a,y)\in\,\cdot\,\}} H(dy|z,a),$$

where $F(z,a,y) \coloneqq \Pr\{X_{t+1} \in \,\cdot\,|Z_t = z, A_t = a, Y_{t+1} = y\}$, $H(\,\cdot\,|z,a) \coloneqq \Pr\{Y_{t+1} \in \,\cdot\,|Z_t = z, A_t = a\}$, and Z_t denotes the posterior distribution of the state X_t given the past observations. The one-stage cost function c of the belief-MDP is given by

$$c(z,a) \coloneqq \int_{\mathsf{X}} \tilde{c}(x,a) z(dx). \tag{4.28}$$

Hence, the belief-MDP is a Markov decision process with the components $(\mathsf{Z}, \mathsf{A}, \eta, c)$.

For the belief-MDP define the history spaces $\mathsf{H}_t = (\mathsf{Z} \times \mathsf{A})^t \times \mathsf{Z}$, $t = 0, 1, 2, \ldots$. Φ denotes the set of all policies for the belief-MDP, where the policies are defined in an usual manner. Let $J(\varphi, \xi)$ denote the discounted cost function of policy $\varphi \in \Phi$ for initial distribution ξ of the belief-MDP.

Notice that any history vector $h_t = (z_0, \ldots, z_t, a_0, \ldots, a_{t-1})$ of the belief-MDP is a function of the history vector $\tilde{h}_t = (y_0, \ldots, y_t, a_0, \ldots, a_{t-1})$ of the POMDP. Let us write this relation as $i(\tilde{h}_t) = h_t$. Hence, for a policy $\varphi = \{\varphi_t\} \in \Phi$, we can define a policy $\pi^{\varphi} = \{\pi_t^{\varphi}\} \in \Pi$ as

$$\pi_t^{\varphi}(\,\cdot\,|\tilde{h}_t) \coloneqq \varphi_t(\,\cdot\,|i(\tilde{h}_t)).$$

Let us write this as a mapping from Φ to Π: $\Phi \ni \varphi \mapsto i(\varphi) = \pi^{\varphi} \in \Pi$. It is straightforward to show that the cost functions $J(\varphi, \xi)$ and $\tilde{J}(\pi^{\varphi}, \mu)$ are the same.

One can also prove that (see [109], [78])

$$\inf_{\varphi \in \Phi} J(\varphi, \xi) = \inf_{\pi \in \Pi} \tilde{J}(\pi, \mu) \tag{4.29}$$

and furthermore, that if φ is an optimal policy for belief-MDP, then π^{φ} is optimal for the POMDP as well. Hence, the POMDP and the corresponding belief-MDP are equivalent in the sense of cost minimization. We will impose the following assumptions on the components of the original POMDP.

Assumption 4.5.

(a) The one stage cost function $\tilde{c}$ is continuous and bounded.

(b) The stochastic kernel $p(\,\cdot\,|x,a)$ is weakly continuous in $(x,a) \in \mathsf{X} \times \mathsf{A}$.

(c) The stochastic kernel $r(\,\cdot\,|x)$ is continuous in total variation, i.e., if $x_k \to x$, then $r(\,\cdot\,|x_k) \to r(\,\cdot\,|x)$ in total variation.

(d) A is compact.

Note that by [9, Proposition 7.30], the one stage cost function c, which is defined in (4.28), is in $C_b(\mathsf{Z} \times \mathsf{A})$ under Assumption 4.5-(a),(b). Hence, the belief-MDP satisfies the conditions in Theorem 4.4 for $w = 1$ if η is weakly continuous. The following theorem is a consequence of [40, Theorem 3.7, Example 4.1] and Example 4.3.

Theorem 4.7.

(i) Under Assumption 4.5-(b),(c), the stochastic kernel η for belief-MDP is weakly continuous in (z,a).

(ii) If we relax the continuity of the observation channel in total variation to setwise or weak continuity, then η may not be weakly continuous even if the transition

probability p of POMDP is continuous in total variation.

(iii) Finally, η may not be setwise continuous in a, even if the observation channel is continuous in total variation.

Part(i) of Theorem 4.7 implies that belief-MDP satisfies conditions in Theorem 4.4. However, note that continuity of the observation channel in total variation in Assumption 4.5 cannot be relaxed to weak or setwise continuity. On the other hand, the continuity of the observation channel in total variation is not enough for the setwise continuity of η. Hence, results in Section 4.3 cannot be applied to the POMDP we consider even though we put a fairly strong condition on the observation channel.

Theorem 4.8. *Suppose Assumption 4.5 holds for the POMDP. Then we have*

$$\lim_{n\to\infty} |J_n^*(z) - J^*(z)| = 0 \quad \text{for all } z \in \mathsf{Z},$$

where J_n^* is the discounted value function of the belief-MDP with the components $\{\mathsf{Z}, \Lambda_n, \eta, c\}$ and J^* is the discounted value function of the belief-MDP with the components $\{\mathsf{Z}, \mathsf{A}, \eta, c\}$.

The significance of Theorem 4.8 is reinforced by the following observation. If we define $D\Pi\mathcal{Q}(\Lambda_n)$ as the set of deterministic policies in Π taking values in Λ_n, then the above theorem implies that for any given $\varepsilon > 0$ there exists $n \geq 1$ and $\pi^* \in D\Pi\mathcal{Q}(\Lambda_n)$ such that

$$\tilde{J}(\pi^*, \mu) < \min_{\pi\in\Pi} \tilde{J}(\pi, \mu) + \varepsilon,$$

where $\pi^* = \pi^{\varphi^*}$. This means that even when is an information transmission constraint from the controller to the plant, one might get ε-close to the value function for any small ε by quantizing the controller's actions and sending the encoded levels.

4.5 Rates of Convergence

In Sections 4.3 and 4.4 we consider the convergence of the finite-action models MDP_n to the original model. In this section we obtain performance bounds on the approximation errors due to quantization of the action space in terms of the number of points used to discretize action space. Namely, we study the following problem.

(Pr) For any $f \in \mathbb{F}$ and initial point x, the approximating sequence $\{\Upsilon_n(f)\}$ in **(P)** is such that $|W(f,x) - W(\Upsilon_n(f),x)|$ can be explicitly upper bounded by a term depending on the cardinality of Λ_n, where $W \in \{J, V\}$.

Thus **(Pr)** implies that the approximation error in **(P)** can be explicitly controlled by the number of points used to discretize the action space. We will impose a new set of assumptions in this section:

Assumption 4.6.

(h) A is infinite compact subset of $\mathbb{R}^d$ for some $d \geq 1$.

(j) c is bounded and $|c(x,\tilde{a}) - c(x,a)| \leq K_1 d_{\mathsf{A}}(\tilde{a},a)$ for all x, and some $K_1 \geq 0$.

(k) $\|p(\,\cdot\,|x,\tilde{a}) - p(\,\cdot\,|x,a)\|_{TV} \leq K_2 d_{\mathsf{A}}(\tilde{a},a)$ for all x, and some $K_2 \geq 0$.

(l) There exists positive constants C and $\beta \in (0,1)$ such that for all $f \in \mathbb{F}$, there is a (necessarily unique) probability measure $\nu_f \in \mathcal{P}(\mathsf{X})$ satisfying $\|\lambda_t^{f,x} - \nu_f\|_{TV} \leq C\kappa^t$ for all $x \in \mathsf{X}$ and $t \geq 1$.

Assumption 4.6-(l) implies that for any policy $f \in \mathbb{F}$, the stochastic kernel Q_f, defined in (4.8), has a unique invariant probability measure ν_f and satisfies *geometric* *ergodicity* [56]. Note that Assumption 4.6-(l) holds under any of the conditions Ri, $i \in \{0, 1, 1(a), 1(b), 2, \ldots, 5\}$ in [57]. Moreover, one can explicitly compute the constants C and κ for certain systems. For instance, consider an additive-noise system in Example 4.1 with Gaussian noise. Let $\mathsf{X} = \mathbb{R}$. Assume F has a bounded range so that $F(\mathbb{R}) \subset [-L, L]$ for some $L > 0$. Then, Assumption 4.6-(l) holds with $C = 2$ and $\kappa = 1 - \varepsilon L$, where $\varepsilon = \frac{1}{\sigma\sqrt{2\pi}} \exp^{-(2L)^2/2\sigma^2}$. For further conditions that imply Assumption 4.6-(l) we refer [57], [56].

The following example gives the sufficient conditions for the additive noise system under which Assumption 4.6-(j),(k),(l) hold.

Example 4.4. Consider the additive-noise system in Example 4.1. In addition to the assumptions there, suppose $F(x, \cdot)$ is Lipschitz uniformly in $x \in \mathsf{X}$ and the common density g of the V_t is Lipschitz on all compact subsets of X. Note that a Gaussian density has these properties. Let $c(x, a) := \|x - a\|^2$. Under these conditions, Assumption 4.6-(j),(k) hold for the additive noise system. If we further assume that F is bounded, then Assumption 4.6-(l) holds as well.

The following result is a consequence of the fact that if A is a compact subset of $\mathbb{R}^d$ then there exist a constant $\alpha > 0$ and finite subsets $\Lambda_n \subset \mathsf{A}$ with cardinality $|\Lambda_n| = n$ such that $\max_{x \in \mathsf{A}} \min_{y \in \Lambda_n} d_{\mathsf{A}}(x, y) \leq \alpha(1/n)^{1/d}$ for all n, where d_{A} is the Euclidean distance on A inherited from $\mathbb{R}^d$.

Lemma 4.9. *Let $\mathsf{A} \subset \mathbb{R}^d$ be compact. Then for any $f \in \mathbb{F}$ the sequence $\{\Upsilon_n(f)\}$ satisfies $\sup_{x \in \mathsf{X}} d_{\mathsf{A}}(\Upsilon_n(f)(x), f(x)) \leq \alpha(1/n)^{1/d}$ for some constant α.*

The following proposition is the key result in this section. It is proved in Section 4.7.5

Proposition 4.3. *Let $f \in \mathbb{F}$ and $\{q_n\}$ be the quantized approximations of f, i.e., $\Upsilon_n(f) = q_n$. For any initial point x, we have*

$$\|\lambda_t^{f,x} - \lambda_t^{q_n,x}\|_{TV} \leq \alpha K_2 (2t-1)(1/n)^{1/d} \tag{4.30}$$

for all $t \geq 1$ under Assumption 4.6-(h),(j),(k).

4.5.1 Discounted Cost

The following result solves **(Pr)** for the discounted cost criterion. The proof of it essentially follows from Proposition 4.3.

Theorem 4.9. *Let $f \in \mathbb{F}$ and $\{q_n\}$ be the quantized approximations of f, i.e., $\Upsilon_n(f) = q_n$. For any initial point x, we have*

$$|J(f,x) - J(q_n,x)| \leq K(1/n)^{1/d}, \tag{4.31}$$

where $K = \frac{\alpha}{1-\beta}(K_1 - \beta K_2 M + \frac{2\beta M K_2}{1-\beta})$ with $M := \|c\|$, under Assumption 4.6-(h),(j),(k).

Proof. Let $c_f(x) = c(x, f(x))$ and $c_{q_n}(x) = c(x, q_n(x))$. For any fixed n we have

$$\begin{aligned}
|J(f,x) - J(q_n,x)| &= \left| \sum_{t=0}^{\infty} \beta^t \lambda_t^{f,x}(c_f) - \sum_{t=0}^{\infty} \beta^t \lambda_t^{q_n,x}(c_{q_n}) \right| \\
&\leq \sum_{t=0}^{\infty} \beta^t \big(\, |\lambda_t^{f,x}(c_f) - \lambda_t^{f,x}(c_{q_n})| + |\lambda_t^{f,x}(c_{q_n}) - \lambda_t^{q_n,x}(c_{q_n})| \, \big) \\
&\leq \sum_{t=0}^{\infty} \beta^t \big(\, \|c_f - c_{q_n}\| + \|\lambda_t^{f,x} - \lambda_t^{q_n,x}\|_{TV} M \, \big) \\
&\leq \sum_{t=0}^{\infty} \beta^t \big(\sup_{x \in \mathsf{X}} d_{\mathsf{A}}(f(x), q_n(x)) K_1 + \|\lambda_t^{f,x} - \lambda_t^{f,x}\|_{TV} M \big)
\end{aligned}$$

$$\leq \sum_{t=0}^{\infty} \beta^t \big((1/n)^{1/d} \alpha K_1\big) + \sum_{t=1}^{\infty} \beta^t \big((1/n)^{1/d} (2t-1) K_2 \alpha M\big) \quad (4.32)$$
$$= (1/n)^{1/d} \frac{\alpha}{1-\beta} (K_1 - \beta K_2 M + \frac{2\beta M K_2}{1-\beta}).$$

Here (4.32) follows from Assumption 4.6-(j), Proposition 4.3, and Lemma 4.9, completing the proof. □

4.5.2 Average Cost

Note that for any $f \in \mathbb{F}$, Assumption 4.6-(l) implies that ν_f is an unique invariant probability measure for Q_f and that $V(f,x) = \nu_f(c_f)$ for all x, where $c_f(x) = c(x, f(x))$. The following theorem basically follows from Proposition 4.3 and the Assumption 4.6-(l). It solves **(Pr)** for the average cost criterion.

Theorem 4.10. *Let $f \in \mathbb{F}$ and $\{q_n\}$ be the quantized approximations of f, i.e., $\Upsilon_n(f) = q_n$. Then, under Assumption 4.6, for any $x \in \mathsf{X}$ we have*

$$|V(f,x) - V(q_n,x)| \leq 2MC\kappa^t + K_t (1/n)^{1/d} \quad (4.33)$$

for all $t \geq 0$, where $K_t = \big((2t-1)K_2 \alpha M + K_1 \alpha\big)$ and $M \coloneqq \|c\|$.

Observe that depending on the values of C and κ, we can first make the first term in the upper bound small enough by choosing sufficiently large t, and then for this t we can choose n large enough such that the second term in the upper bound is small.

Proof. For any n and $x \in \mathsf{X}$, we have

$$|V(f,x) - V(q_n,x)| = |\nu_f(c_f) - \nu_{q_n}(c_{q_n})| \leq |\nu_f(c_f) - \nu_f(c_{q_n})| + |\nu_f(c_{q_n}) - \nu_{q_n}(c_{q_n})|$$

$$
\begin{aligned}
&\leq \|c_f - c_{q_n}\| + \|\nu_f - \nu_{q_n}\|_{TV} M \\
&\leq \sup_{x \in \mathsf{X}} K_1 d_{\mathsf{A}}(f(x), q_n(x)) + \|\nu_f - \nu_{q_n}\|_{TV} M \text{ (by Assumption 4.6-(j))} \\
&\leq (1/n)^{1/d} K_1 \alpha + \big(\|\nu_f - \lambda_t^{f,x}\|_{TV} + \|\lambda_t^{f,x} - \lambda_t^{q_n,x}\|_{TV} + \|\lambda_t^{q_n,x} - \nu_{q_n}\|_{TV} \big) M \\
&\leq (1/n)^{1/d} K_1 \alpha + \big(2C\kappa^t + (1/n)^{1/d}(2t-1) K_2 \alpha \big) M \qquad (4.34) \\
&= 2MC\kappa^t + \big((2t-1) K_2 \alpha M + K_1 \alpha \big) (1/n)^{1/d},
\end{aligned}
$$

where (4.34) follows from Assumption 4.6-(l) and Proposition 4.3. □

4.5.3 Order Optimality

The following example demonstrates that the order of approximation errors in Theorems 4.9 and 4.10 cannot be better than $O((\frac{1}{n})^{\frac{1}{d}})$. More precisely, we exhibit a simple standard example where we can lower bound the approximation errors for the optimal stationary policy by $L(1/n)^{1/d}$, for some positive constant L.

In what follows $h(\,\cdot\,)$ and $h(\,\cdot\,|\,\cdot\,)$ denote differential and conditional differential entropies, respectively [25].

Consider the additive-noise system:

$$
X_{t+1} = F(X_t, A_t) + V_t, \;\; t = 0, 1, 2, \ldots,
$$

where $\mathsf{X} = \mathsf{A} \subset \mathbb{R}^d$ and the $\{V_t\}$ is a sequence i.i.d. random vectors whose common distribution has density g supported on some compact subset V of $\mathbb{R}^d$. We assume that $\sup_{(x,a) \in \mathbb{R}^d \times \mathbb{R}^d} \frac{\|F(x,a)\|}{\|x\|+\|a\|} < 1/2$. We choose V such that $\mathsf{X} = \mathsf{A}$ can restricted to be a compact subset of $\mathbb{R}^d$. For simplicity suppose that the initial distribution μ has the same density g. It is assumed that the differential entropy $h(g) := -\int_{\mathsf{X}} g(x) \log g(x) dx$

is finite. Let the one stage cost function be $c(x,a) \coloneqq \|x-a\|$. Clearly, the optimal stationary policy f^* is the identity $f^*(x)=x$, having the optimal cost $W(f,\mu)=0$, where $W \in \{J,W\}$. Let q_n be the quantized approximations of f^*. Fix any n and define $D_t \coloneqq \mathbb{E}_{\mu}^{q_n}\big[c(X_t,A_t)\big]$ for all t. Then, by the Shannon lower bound (SLB) [105, p. 12] we have for $n \geq 1$

$$\begin{aligned} \log n \geq R(D_t) &\geq h(X_t) + \theta(D_t) \\ &= h(F(X_{t-1},A_{t-1}) + V_{t-1}) + \theta(D_t) \\ &\geq h(F(X_{t-1},A_{t-1}) + V_{t-1}|X_{t-1},A_{t-1}) + \theta(D_t) \\ &= h(V_{t-1}) + \theta(D_t), \end{aligned} \tag{4.35}$$

where $\theta(D_t) = -d + \log\left(\frac{1}{dV_d\Gamma(d)}\big(\frac{d}{D_t}\big)^d\right)$, $R(D_t)$ is the rate-distortion function of X_t, V_d is the volume of the unit sphere $S_d = \{x : \|x\| \leq 1\}$, and Γ is the gamma function. Here, (4.35) follows from the independence of V_{t-1} and the pair (X_{t-1},A_{t-1}). Note that $h(V_{t-1}) = h(g)$ for all t. Hence, we obtain $D_t \geq L(1/n)^{1/d}$, where $L \coloneqq \frac{d}{2}\big(\frac{2^{h(g)}}{dV_d\Gamma(d)}\big)^{1/d}$. This gives $|J(f^*,\mu) - J(q_n,\mu)| \geq \frac{L}{1-\beta}(1/n)^{1/d}$ and $|V(f^*,\mu) - V(q_n,\mu)| \geq L(1/n)^{1/d}$.

4.6 Conclusion

In this chapter, we considered the finite-action approximation of stationary policies for a discrete-time Markov decision process with discounted and average costs. Under mild technical assumptions, it was shown that if one uses a sufficiently large number of points to discretize the action space, then the resulting finite-action MDP can approximate the original model with arbitrary precision. Under the Lipschitz continuity of the transition probability and the one-stage cost function explicit bounds were obtained on the performance loss due to quantization.

4.7 Proofs

4.7.1 Proof of Proposition 4.1

To ease the notation let $\Upsilon_n(f) = q_n$. Suppose $g \in \mathcal{C}(\mathsf{H}_t)$ for some t. Then we have $P_x^{q_n}(g) = \lambda_{(t)}^{q_n,x}(g_{q_n})$ and $P_x^f(g) = \lambda_{(t)}^{f,x}(g_f)$, where $g_{q_n} = g(x_0, q_n(x_0), \ldots, q_n(x_{t-1}), x_t)$ and $g_f = g(x_0, f(x_0), \ldots, f(x_{t-1}), x_t)$. Since g is continuous in the "a" terms by definition and q_n converges to f, we have $g_{q_n} \to g_f$. Hence, by [93, Theorem 2.4] it is enough to prove that $\lambda_{(t)}^{q_n,x} \to \lambda_{(t)}^{f,x}$ setwise as $n \to \infty$.

We will prove this by induction. Clearly, $\lambda_{(1)}^{q_n,x} \to \lambda_{(1)}^{f,x}$ setwise by Assumption 4.1-(b). Assume the claim is true for some $t \geq 1$. For any $h \in B(\mathsf{X}^{t+2})$ we can write $\lambda_{(t+1)}^{q_n,x}(h) = \lambda_{(t)}^{q_n,x}\big(\lambda_1^{q_n,x_t}(h)\big)$ and $\lambda_{(t+1)}^{f,x}(h) = \lambda_{(t)}^{f,x}\big(\lambda_1^{f,x_t}(h)\big)$. Since $\lambda_1^{q_n,x_t}(h) \to \lambda_1^{f,x_t}(h)$ for all $(x_0, \ldots, x_t) \in \mathsf{X}^{t+1}$ by Assumption 4.1-(b) and $\lambda_{(t)}^{q_n,x} \to \lambda_{(t)}^{f,x}$ setwise, we have $\lambda_{(t+1)}^{q_n,x}(h) \to \lambda_{(t+1)}^{f,x}(h)$ by again [93, Theorem 2.4] which completes the proof.

4.7.2 Proof of Theorem 4.2

Let $\Upsilon_n(f) = q_n$. Let Q_f and Q_{q_n} be the stochastic kernels, respectively, for f and q_n defined in (4.8). By Assumption 4.2-(e), Q_f and Q_{q_n} ($n \geq 1$) have unique, and so ergodic, invariant probability measures ν_f and ν_{q_n}, respectively. Since $x \in \mathsf{M}$, we have $V(q_n, x) = \nu_{q_n}(c_{q_n})$ and $V(f, x) = \nu_f(c_f)$, where $c_{q_n}(x) = c(x, q_n(x))$ and $c_f(x) = c(x, f(x))$. Observe that $c_{q_n}(x) \to c_f(x)$ for all x by Assumption 4.2-(a). Hence, if we prove $\nu_{q_n} \to \nu_f$ setwise, then by [93, Theorem 2.4] we have $V(q_n, x) \to V(f, x)$. We prove this first under (f1) and then under (f2).

I) Proof under assumption (f1)

We show that every setwise convergent subsequence $\{\nu_{q_{n_l}}\}$ of $\{\nu_{q_n}\}$ must converge to ν_f. Then, since $\Gamma_{\mathbb{F}}$ is relatively sequentially compact in the setwise topology, there is at least one setwise convergent subsequence $\{\nu_{q_{n_l}}\}$ of $\{\nu_{q_n}\}$, which implies the result.

Let $\nu_{q_{n_l}} \to \nu$ setwise for some $\nu \in \mathcal{P}(\mathsf{X})$. We will show that $\nu = \nu_f$ or equivalently ν is an invariant probability measure of Q_f. For simplicity, we write $\{\nu_{q_l}\}$ instead of $\{\nu_{q_{n_l}}\}$. Let $g \in B(\mathsf{X})$. Then by Assumption 4.2-(e) we have $\nu_{q_l}(g) = \nu_{q_l}(Q_{q_l}g)$. Since $Q_{q_l}g(x) \to Q_f g(x)$ for all x by Assumption 4.2-(b) and $\nu_{q_l} \to \nu$ setwise, we have $\nu_{q_l}(Q_{q_l}g) \to \nu(Q_f g)$ by [93, Theorem 2.4]. On the other hand since $\nu_{q_l} \to \nu$ setwise we have $\nu_{q_l}(g) \to \nu(g)$. Thus $\nu(g) = \nu(Q_f g)$. Since g is arbitrary, ν is an invariant probability measure for Q_f.

II) Proof under assumption (f2)

Observe that for all $x \in \mathsf{X}$ and t, $\lambda_t^{q_n,x} \to \lambda_t^{f,x}$ setwise as $n \to \infty$ since $P_x^{q_n} \to P_x^f$ in the ws^∞ topology (see Proposition 4.1). Let $B \in \mathcal{B}(\mathsf{X})$ be given and fix some $\varepsilon > 0$. By Assumption 4.2-(f2) we can choose T large enough such that $|\lambda_T^{\tilde{f},x}(B) - \nu_{\tilde{f}}(B)| < \varepsilon/3$ for all $\tilde{f} \in \{f, q_1, q_2, \cdots\}$. For this T, choose N large enough such that $|\lambda_T^{q_n,x}(B) - \lambda_T^{f,x}(B)| < \varepsilon/3$ for all $n \geq N$. Thus, for all $n \geq N$ we have $|\nu_{q_n}(B) - \nu_f(B)| \leq |\nu_{q_n}(B) - \lambda_T^{q_n,x}(B)| + |\lambda_T^{q_n,x}(B) - \lambda_T^{f,x}(B)| + |\lambda_T^{f,x}(B) - \nu_f(B)| < \varepsilon$. Since ε is arbitrary, we obtain $\nu_{q_n}(B) \to \nu_f(B)$, which completes the proof.

4.7.3 Proof of Lemma 4.3

Let us define the set of measures Ξ on X as

$$\Xi := \left\{ Q(\,\cdot\,|x,a) : Q(D|x,a) = \int_D w(y)p(dy|x,a),\ (x,a) \in K \times \mathsf{A} \right\}.$$

Note that Ξ is uniformly bounded since

$$\sup_{(x,a)\in K\times\mathsf{A}} \int_{\mathsf{X}} w(y)p(dy|x,a) \leq \alpha \sup_{x\in K} w(x) < \infty.$$

If the mapping $Q : K \times \mathsf{A} \ni (x,a) \mapsto Q(\,\cdot\,|x,a) \in \Xi$ is continuous with respect to the weak topology on Ξ, then Ξ (being a continuous image of the compact set $K \times \mathsf{A}$) is compact with respect to the weak topology. Then, by Prohorov's theorem [17, Theorem 8.6.2], Ξ is tight, completing the proof. Hence, we only need to prove the continuity of the mapping Q.

By Lemma 4.1, for any $u \in C_w(\mathsf{X})$, $\int_{\mathsf{X}} u(y)p(dy|x,a)$ is continuous in (x,a). Let $(x_k,a_k) \to (x,a)$ in $K \times \mathsf{A}$. Note that for any $g \in C_b(\mathsf{X})$, $gw \in C_w(\mathsf{X})$. Therefore, we have

$$\begin{aligned}\lim_{k\to\infty} \int_{\mathsf{X}} g(y)Q(dy|x_k,a_k) &= \lim_{k\to\infty} \int_{\mathsf{X}} g(y)w(y)p(dy|x_k,a_k) \\ &= \int_{\mathsf{X}} g(y)w(y)p(dy|x,a) = \int_{\mathsf{X}} g(y)Q(dy|x,a)\end{aligned}$$

proving that $Q(\,\cdot\,|x_k,a_k) \to Q(\,\cdot\,|x,a)$ weakly.

4.7.4 Proof of Lemma 4.4

Fix any compact subset K of X. We have

$$\begin{aligned}
&\sup_{x\in K}\bigg|\int_{\mathsf{X}} u_n(y)p(dy|x,f_n(x)) - \int_{\mathsf{X}} u(y)p(dy|x,f(x))\bigg| \\
&\quad\le \sup_{x\in K}\bigg|\int_{\mathsf{X}} u_n(y)p(dy|x,f_n(x)) - \int_{\mathsf{X}} u(y)p(dy|x,f_n(x))\bigg| \\
&\qquad\qquad + \sup_{x\in K}\bigg|\int_{\mathsf{X}} u(y)p(dy|x,f_n(x)) - \int_{\mathsf{X}} u(y)p(dy|x,f(x))\bigg| \\
&\quad\le \sup_{x\in K}\bigg|\int_{K_\varepsilon} u_n(y)p(dy|x,f_n(x)) - \int_{K_\varepsilon} u(y)p(dy|x,f_n(x))\bigg| \\
&\qquad\qquad + \sup_{x\in K}\bigg|\int_{K_\varepsilon^c} u_n(y)p(dy|x,f_n(x)) - \int_{K_\varepsilon^c} u(y)p(dy|x,f_n(x))\bigg| \\
&\qquad\qquad\qquad + \sup_{x\in K}\bigg|\int_{\mathsf{X}} u(y)p(dy|x,f_n(x)) - \int_{\mathsf{X}} u(y)p(dy|x,f(x))\bigg| \\
&\quad\le \sup_{y\in K_\varepsilon}|u_n(y)-u(y)| + (L+\|u\|_w)\varepsilon \quad \text{(by Lemma 4.3)} \\
&\qquad\qquad + \sup_{x\in K}\bigg|\int_{\mathsf{X}} u(y)p(dy|x,f_n(x)) - \int_{\mathsf{X}} u(y)p(dy|x,f(x))\bigg|
\end{aligned}$$

Let us define $l(x,a) \coloneqq \int_{\mathsf{X}} u(y)p(dy|x,a)$. Since $u(y) \in C_w(\mathsf{X})$, by Lemma 4.1 l is continuous, and therefore, uniformly continuous on $K \times \mathsf{A}$. Note that in the last expression as $n \to \infty$: (i) the first term goes to zero since $u_n \to u$ uniformly on K_ε and (ii) the last term goes to zero since l is uniformly continuous on $K \times \mathsf{A}$ and $f_n \to f$ uniformly. Then the result follows by observing that ε is arbitrary.

4.7.5 Proof of Proposition 4.3

We will prove this result by induction. Let x be an arbitrary initial point and fix n. For $t = 1$ the claim holds by the following argument:

$$\begin{aligned}\|\lambda_1^{f,x} - \lambda_1^{q_n,x}\|_{TV} &\leq K_2 d_{\mathsf{A}}(f(x), q_n(x)) \text{ (by Assumption 4.6-(k))} \\ &\leq (1/n)^{1/d} K_2 \alpha \text{ (by Lemma 4.9)}.\end{aligned}$$

Observe that the bound $\alpha K_2(2t-1)(1/n)^{1/d}$ is independent of the choice of initial point x for $t = 1$. Assume the claim is true for $t \geq 1$. Then we have

$$\begin{aligned}\|\lambda_{t+1}^{f,x} - \lambda_{n+1}^{q_n,x}\|_{TV} &= 2 \sup_{B \in \mathcal{B}(\mathsf{X})} \left|\lambda_1^{f,x}(\lambda_t^{f,x_1}(B)) - \lambda_1^{q_n,x}(\lambda_t^{q_n,x_1}(B))\right| \\ &= 2 \sup_{B \in \mathcal{B}(\mathsf{X})} \left|\lambda_1^{f,x}(\lambda_t^{f,x_1}(B)) - \lambda_1^{f,x}(\lambda_t^{q_n,x_1}(B)) + \lambda_1^{f,x}(\lambda_t^{q_n,x_1}(B)) - \lambda_1^{q_n,x}(\lambda_t^{q_n,x_1}(B))\right| \\ &\leq \lambda_1^{f,x}(\|\lambda_t^{f,x_1} - \lambda_t^{q_n,x_1}\|_{TV}) + 2\|\lambda_1^{f,x} - \lambda_1^{q_n,x}\|_{TV} \qquad (4.36) \\ &\leq (1/n)^{1/d}(2t-1)K_2\alpha + 2(1/n)^{1/d}K_2\alpha \qquad (4.37) \\ &= \alpha K_2(2(t+1)-1)(1/n)^{1/d}\alpha.\end{aligned}$$

Here (4.36) follows since $|\mu(h) - \eta(h)| \leq \|\mu - \eta\|_{TV}\|h\|$ and (4.37) follows since the bound $\lambda K_2(2t-1)(1/n)^{1/d}$ is independent of the initial distribution.

Chapter 5

Quantization of the State Space and Asymptotic Optimality of Finite Models for Borel MDPs

5.1 Introduction

In this chapter our aim is to study the finite-state approximation problem for computing near optimal policies for discrete time Markov decision processes (MDPs) with Borel state and action spaces, under discounted and average costs criteria.

In the theory of stochastic optimal control, although existence and structural properties of optimal policies have been studied extensively in the literature, computing such policies is generally a challenging problem for systems with uncountable state spaces. This situation also arises in the fully observed reduction of a partially observed Markov decision processes even when the original system has finite state and action spaces.

In this chapter we show that one way to compute approximately optimal solutions for such MDPs is to construct a reduced model with a new transition probability and a one-stage cost function by quantizing the state/action spaces, i.e., by discretizing them on a finite grid. It is reasonable to expect that when the one-stage cost function and the transition probability of the original model satisfy certain continuity properties, the optimal policy for the approximating finite model applied to the original model has cost that converges to the optimal cost, as the discretization becomes finer. Moreover, under additional continuity conditions on the transition probability and the one stage cost function one may also obtain bounds for a rate of approximation in terms of the number of points used to discretize the state space, thereby providing a tradeoff between the computation cost and the performance loss in the system. In this chapter, we examine such continuity conditions. In particular, we study the following two problems.

(Q1) Under what conditions on the components of the MDP do the true cost functions of the policies obtained from finite models converge to the optimal value function as the number of grid points goes to infinity? Here, we are only concerned with the convergence of the approximation; that is, we do not establish bounds for a rate of approximation.

(Q2) Can we obtain bounds on the performance loss due to discretization in terms of the number of grid points if we strengthen the conditions sufficient in **(Q1)**?

As mentioned in Section 1.3, various methods have been developed to compute near optimal policies in the literature: approximate dynamic programming, approximate value or policy iteration, simulation-based techniques, neuro-dynamic programming (or reinforcement learning), state aggregation, etc. [37, 22, 12, 77, 72, 80, 100, 10, 33, 34, 35]. Below, we summarize a portion of the existing techniques developed in the literature for this problem.

The Approximate Value Iteration (AVI) and Approximate Policy Iteration (API) algorithms are two powerful methods to approximate an optimal (deterministic stationary) policy for an MDP (see [37], [22], [12], [97] and references therein). In AVI, the idea is to compute approximately the value iteration function, using some prescribed basis functions, in each step of the value iteration algorithm. This way one can both approximately find the optimal value function and construct an approximately optimal policy. In API, at each step, first an approximate value function for a given policy is computed, again using prescribed basis functions, and then, an improved policy is generated using the approximate value function. The main drawback of these algorithms is the accumulation of the approximation error in each step.

Another well-known method for approximating an optimal policy is *state aggregation*. In this method, first similar states (e.g., with respect to cost and transition probabilities) are aggregated to form meta-states of a reduced MDP, and then an optimal policy is calculated according to the reduced MDP (see [77, 7, 72] and references therein). The basic issue with this method is how to efficiently aggregate states and construct a reduced MDP from the original one.

References [42, 100, 101, 23, 51] used the technique of truncating the state space when evaluating the value function in the value iteration algorithm. In these schemes, in each step the state space is truncated and the corresponding value function is calculated; the latter is proved to converge to the true value function. Then, using the truncated value function, approximately optimal policies are constructed.

With the exception of [35, 72], prior works that considered approximation problem in general study either a finite horizon cost or a discounted infinite horizon cost. As well, the majority of these results are for MDPs with discrete (i.e., finite or countable) state and action spaces, or a bounded one-stage cost function (see, e.g., [80, 12, 77, 72, 10, 37, 100, 22]). Those that consider general state and action spaces (see, e.g., [33, 34, 35, 10]) assume in general Lipschitz type continuity conditions on the components of the control model, in order to provide a rate of convergence analysis for the approximation error.

Our work differs from these results in the following ways: (i) we consider a general setup, where the state and action spaces are Borel, and the one-stage cost function is possibly unbounded, (ii) since we do not aim to provide rate of convergence result in the first problem **(Q1)**, the continuity assumptions we impose on the components of the control model are weaker than the conditions imposed in prior works that considered general state and action spaces, (iii) we also consider the challenging average cost criterion under reasonable assumptions. The price we pay for imposing weaker assumptions in **(Q1)** is that we cannot obtain explicit performance bounds in terms of the number of grid points used in the approximations. However, such bounds can be obtained under further assumptions on the transition probability and the one-stage cost functions; this is considered in problem **(Q2)** for compact-state MDPs.

Our approach to solve problem **(Q1)** can be summarized as follows: (i) first, we obtain approximation results for the compact-state case, (ii) we find conditions under which a compact representation leads to near optimality for non-compact state MDPs, (iii) we prove the convergence of the finite-state models to non-compact models. A by-product of this analysis, we obtain *compact-state-space approximation*s for an MDP with non-compact Borel state space. In particular, our findings directly lead to finite models if the state space is countable; similar problems in the countable context have been studied in the literature for the discounted cost [76, Section 6.10.2].

Combined with Chapter 4, where we investigated the asymptotic optimality of the quantization of action sets, the results in this chapter lead to a constructive algorithm for obtaining approximately optimal solutions. First the action space is quantized with small error, and then the state space is quantized with small error, which results in a finite model that well approximates the original MDP. When the state space is compact, we can also obtain rates of convergence for both approximations, and using information theoretic tools we can establish that the obtained rates of convergence are order-optimal for a given class of MDPs.

We note that the proposed method for solving the approximation problem for compact-state MDPs with the discounted cost is partly inspired by [80]. Specifically, we generalize the operator proposed for an approximate value iteration algorithm in [80] to uncountable state spaces. Then, unlike in [80], we use this operator as a transition step between the original optimality operator and the optimality operator of the approximate model. In [72], a similar construction was given for finite state-action MDPs. Our method to obtain finite-state MDPs from the compact-state model can be seen as a generalization of this construction.

5.2 Finite State Approximation of Compact State MDPs

In this section we consider **(Q1)** for the MDPs with compact state space. To distinguish compact-state MDPs from non-compact ones, the state space of the compact-state MDPs will be denoted by Z instead of X. We impose the assumptions below on the components of the Markov decision process; additional new assumptions will be made for the average cost problem in Section 5.2.2.

Assumption 5.1.

(a) The one-stage cost function c is in $C_b(\mathsf{Z} \times \mathsf{A})$.

(b) The stochastic kernel $p(\,\cdot\,|z,a)$ is weakly continuous in (z,a) and setwise continuous in a, i.e., for all z and a, $p(\,\cdot\,|z_k,a_k) \to p(\,\cdot\,|z,a)$ weakly when $(z_k,a_k) \to (z,a)$ and $p(\,\cdot\,|z,a_k) \to p(\,\cdot\,|z,a)$ setwise when $a_k \to a$.

(c) Z and A are compact.

Analogous with the construction of finite action models in Section 4.3, we first describe the procedure used to obtain finite-state models. Let d_{Z} denote the metric on Z. Since the state space Z is assumed to be compact and thus totally bounded, one can find a sequence $\big(\{z_{n,i}\}_{i=1}^{k_n}\big)_{n\geq 1}$ of finite grids in Z such that for all n,

$$\min_{i\in\{1,\ldots,k_n\}} d_{\mathsf{Z}}(z,z_{n,i}) < 1/n \text{ for all } z \in \mathsf{Z}.$$

The finite grid $\{z_{n,i}\}_{i=1}^{k_n}$ is called an $1/n$-net in Z. Let $\mathsf{Z}_n \coloneqq \{z_{n,1},\ldots,z_{n,k_n}\}$ and define function Q_n mapping Z to Z_n by

$$Q_n(z) \coloneqq \mathop{\arg\min}_{z_{n,i}\in\mathsf{Z}_n} d_{\mathsf{Z}}(z,z_{n,i}),$$

where ties are broken so that Q_n is measurable. In the literature, Q_n is often called a nearest neighborhood quantizer with respect to distortion measure d_{Z} [47]. For each n, Q_n induces a partition $\{\mathcal{S}_{n,i}\}_{i=1}^{k_n}$ of the state space Z given by

$$\mathcal{S}_{n,i} = \{z \in \mathsf{Z} : Q_n(z) = z_{n,i}\},$$

with diameter $\mathrm{diam}(\mathcal{S}_{n,i}) \coloneqq \sup_{z,y\in\mathcal{S}_{n,i}} d_{\mathsf{Z}}(z,y) < 2/n$. Let ν be a probability measure on Z satisfying

$$\nu(\mathcal{S}_{n,i}) > 0 \text{ for all } i, n. \tag{5.1}$$

The existence of such a probability measure follows from the separability of Z and the fact that $\operatorname{int} \mathcal{S}_{n,i} \neq \emptyset$ for all i, n. For example, if $\{z_n\}_{n\geq 1}$ is a countable dense subset of Z one can define ν as

$$\nu = \sum_{n=1}^{\infty} 2^{-n} \delta_{z_n}.$$

Then, $\nu(B) > 0$ for any nonempty open set B and therefore (5.1) holds.

We let $\nu_{n,i}$ be the restriction of ν to $\mathcal{S}_{n,i}$ defined by

$$\nu_{n,i}(\,\cdot\,) \coloneqq \frac{\nu(\,\cdot\,)}{\nu(\mathcal{S}_{n,i})}.$$

The measures $\nu_{n,i}$ will be used to define a sequence of finite-state MDPs, denoted as MDP_n ($n \geq 1$), to approximate the original model. To this end, for each n define the one-stage cost function $c_n : \mathsf{Z}_n \times \mathsf{A} \to [0, \infty)$ and the transition probability p_n on Z_n given $\mathsf{Z}_n \times \mathsf{A}$ by

$$c_n(z_{n,i}, a) \coloneqq \int_{\mathcal{S}_{n,i}} c(z, a) \nu_{n,i}(dz),$$
$$p_n(\,\cdot\,|z_{n,i}, a) \coloneqq \int_{\mathcal{S}_{n,i}} Q_n * p(\,\cdot\,|z, a) \nu_{n,i}(dz),$$

where $Q_n * p(\,\cdot\,|z, a) \in \mathcal{P}(\mathsf{Z}_n)$ is the pushforward of the measure $p(\,\cdot\,|z, a)$ with respect to Q_n; that is,

$$Q_n * p(z_{n,j}|z,a) = p\big(\{z \in \mathsf{Z} : Q_n(z) = z_{n,j}\}|z,a\big),$$

for all $z_{n,j} \in \mathsf{Z}_n$. For each n, we define MDP_n as a Markov decision process with the following components: Z_n is the state space, A is the action space, p_n is the transition probability and c_n is the one-stage cost function. History spaces, policies and cost functions are defined in a similar way as in the original model.

Remark 5.1. To avoid measurability problems associated with the operators that will be defined in the sequel, it is necessary to impose the setwise continuity of the transition probability with respect to the action variable. However, for the purposes of this section, without loss of generality, we can only assume the weak continuity of the transition probability in Assumption 5.1-(b). The reason is that any MDP with compact action space can be approximated with arbitrary precision by an MDP having a finite action space under weak continuity assumption as shown in Section 4.4, and in the case of finite action space, weak continuity and Assumption 5.1-(b) are equivalent. However, for the sake of completeness we use Assumption 5.1-(b) as it appears above.

5.2.1 Discounted Cost

Here we consider **(Q1)** for the discounted cost criterion with a discount factor $\beta \in (0,1)$. Recall the Bellman optimality operator T defined in (4.12). It can be proved that under Assumption 5.1-(a)(b), T is a contraction operator with modulus β mapping $C_b(\mathsf{Z})$ into itself (see [52, Theorem 2.8]); that is, $Tu \in C_b(\mathsf{Z})$ for all $u \in C_b(\mathsf{Z})$

and

$$\|Tu - Tv\| \leq \beta \|u - v\| \text{ for all } u, v \in C_b(\mathsf{Z}).$$

In this case, Theorem 4.3 in Section 4.4.1 implies that the value function J^* is the unique fixed point in $C_b(\mathsf{Z})$ of the contraction operator T, i.e.,

$$J^* = TJ^*.$$

Furthermore, a deterministic stationary policy f^* is optimal if and only if it satisfies the optimality equation, i.e.,

$$J^*(z) = c(z, f^*(z)) + \beta \int_{\mathsf{Z}} J^*(y) p(dy|z, f^*(z)). \tag{5.2}$$

Finally, there exists a deterministic stationary policy f^* which is optimal, so it satisfies (5.2).

Define, for all $n \geq 1$, the operator T_n, which is the Bellman optimality operator for MDP_n, by

$$T_n u(z_{n,i}) \coloneqq \min_{a \in \mathsf{A}} \biggl[c_n(z_{n,i}, a) + \beta \sum_{j=1}^{k_n} u(z_{n,j}) p_n(z_{n,j}|z_{n,i}, a) \biggr],$$

or equivalently,

$$T_n u(z_{n,i}) = \min_{a \in \mathsf{A}} \int_{\mathcal{S}_{n,i}} \biggl[c(z, a) + \beta \int_{\mathsf{Z}} \hat{u}(y) p(dy|z, a) \biggr] \nu_{n,i}(dz),$$

where $u : \mathsf{Z}_n \to \mathbb{R}$ and $\hat{u}$ is the piecewise constant extension of u to Z given by

$\hat{u}(z) = u \circ Q_n(z)$. For each n, under Assumption 5.1, [52, Lemma 2.5 and Theorem 2.2] implies the following: (i) T_n is a contraction operator with modulus β mapping $B(\mathsf{Z}_n)$ $\bigl(= C_b(\mathsf{Z}_n)\bigr)$ into itself, (ii) the fixed point of T_n is the value function J_n^* of MDP$_n$, and (iii) there exists an optimal stationary policy f_n^* for MDP$_n$, which therefore satisfies the optimality equation. Hence, we have

$$J_n^* = T_n J_n^* = T_n J_n(f_n^*, \cdot\,) = J_n(f_n^*, \cdot\,),$$

where J_n denotes the discounted cost for MDP$_n$. Let us extend the optimal policy f_n^* for MDP$_n$ to Z by letting $\hat{f}_n(z) = f_n^* \circ Q_n(z) \in \mathbb{F}$.

The following theorem is the main result of this section. It states that the cost function of the policy $\hat{f}_n$ converges to the value function J^* as $n \to \infty$.

Theorem 5.1. *The discounted cost of the policy $\hat{f}_n$, obtained by extending the optimal policy f_n^* of MDP$_n$ to Z, converges to the optimal value function J^* of the original MDP*

$$\lim_{n\to\infty} \|J(\hat{f}_n, \cdot\,) - J^*\| = 0.$$

Hence, to find a near optimal policy for the original MDP, it is sufficient to compute the optimal policy of MDP$_n$ for sufficiently large n, and then extend this policy to the original state space.

To prove Theorem 5.1 we need a series of technical results. We first define an operator $\hat{T}_n$ on $B(\mathsf{Z})$ by extending T_n to $B(\mathsf{Z})$:

$$\hat{T}_n u(z) := \min_{a\in\mathsf{A}} \int_{\mathcal{S}_{n,i_n(z)}} \biggl[c(x,a) + \beta \int_{\mathsf{Z}} u(y) p(dy|x,a) \biggr] \nu_{n,i_n(z)}(dx), \tag{5.3}$$

where $i_n : \mathsf{Z} \to \{1, \ldots, k_n\}$ maps z to the index of the partition $\{\mathcal{S}_{n,i}\}$ it belongs to.

Remark 5.2. In the rest of this chapter, when we take the integral of any function with respect to $\nu_{n,i_n(z)}$, it is tacitly assumed that the integral is taken over all set $\mathcal{S}_{n,i_n(z)}$. Hence, we can drop $\mathcal{S}_{n,i_n(z)}$ in the integral for the ease of notation.

Since the expression inside the minimization in (5.3) is a continuous function of a by Assumption 5.1-(b), $\hat{T}_n$ maps $B(\mathsf{Z})$ into itself by [54, Proposition D.5]. Furthermore, it is a contraction operator with modulus β which can be shown using [52, Proposition A.2]. Hence, it has a unique fixed point $\hat{J}_n^*$ that belongs to $B(\mathsf{Z})$, and this fixed point must be constant over the sets $\mathcal{S}_{n,i}$ because of the averaging operation on each $\mathcal{S}_{n,i}$. Furthermore, since $\hat{T}_n(u \circ Q_n) = (T_n u) \circ Q_n$ for all $u \in B(\mathsf{Z}_n)$, we have

$$\hat{T}_n(J_n^* \circ Q_n) = (T_n J_n^*) \circ Q_n = J_n^* \circ Q_n.$$

Hence, the fixed point of $\hat{T}_n$ is the piecewise constant extension of the fixed point of T_n, i.e.,

$$\hat{J}_n^* = J_n^* \circ Q_n.$$

Remark 5.3. This is the point where we need the setwise continuity of the transition probability p with respect to a, because if we only assume that the stochastic kernel p is weakly continuous, then it will be no longer true that $\hat{T}_n$ maps $B(\mathsf{Z})$ into itself (see [54, Proposition D.5]).

We now define another operator F_n on $B(\mathsf{Z})$ by simply interchanging the order of the minimum and the integral in (5.3), i.e.,

$$\begin{aligned} F_n u(z) &:= \int \min_{a \in \mathsf{A}} \left[c(x,a) + \beta \int_{\mathsf{Z}} u(y) p(dy|x,a) \right] \nu_{n,i_n(z)}(dx) \\ &= \Gamma_n T u(z), \end{aligned}$$

where

$$\Gamma_n u(z) := \int u(x) \nu_{n,i_n(z)}(dx).$$

We note that F_n is the extension (to infinite state spaces) of the operator defined in [80, p. 236] for the proposed approximate value iteration algorithm. However, unlike in [80], F_n will serve here as an intermediate point between T and $\hat{T}_n$ (or T_n) to solve **(Q1)** for the discounted cost. To this end, we first note that F_n is a contraction operator on $B(\mathsf{Z})$ with modulus β. Indeed it is clear that F_n maps $B(\mathsf{Z})$ into itself. Furthermore, for any $u, v \in B(\mathsf{Z})$, we clearly have $\|\Gamma_n u - \Gamma_n v\| \leq \|u - v\|$. Hence, since T is a contraction operator on $B(\mathsf{Z})$ with modulus β, F_n is also a contraction operator on $B(\mathsf{Z})$ with modulus β.

The following theorem states that the fixed point, say u_n^*, of F_n converges to the fixed point J^* (i.e., the value function) of T as n goes to infinity.

Theorem 5.2. *If u_n^* is the unique fixed point of F_n, then* $\lim_{n \to \infty} \|u_n^* - J^*\| = 0$.

The proof of Theorem 5.2 requires two lemmas.

Lemma 5.1. *For any $u \in B(\mathsf{Z})$, we have*

$$\|u - \Gamma_n u\| \leq 2 \inf_{r \in \mathsf{Z}^{k_n}} \|u - \Phi_r\|,$$

where $\Phi_r(z) = \Sigma_{i=1}^{k_n} r_i 1_{S_{n,i}}(z)$, $r = (r_1, \cdots, r_{k_n})$.

Proof. Fix any $r \in \mathsf{Z}^{k_n}$. Then, using the identity $\Gamma_n \Phi_r = \Phi_r$, we obtain

$$\begin{aligned}\|u - \Gamma_n u\| &\leq \|u - \Phi_r\| + \|\Phi_r - \Gamma_n u\| \\ &= \|u - \Phi_r\| + \|\Gamma_n \Phi_r - \Gamma_n u\| \\ &\leq \|u - \Phi_r\| + \|\Phi_r - u\|.\end{aligned}$$

Since r is arbitrary, this completes the proof. □

Notice that because of the operator Γ_n, the fixed point u_n^* of F_n must be constant over the sets $\mathcal{S}_{n,i}$. We use this property to prove the next lemma.

Lemma 5.2. *We have*

$$\|u_n^* - J^*\| \leq \frac{2}{1-\beta} \inf_{r \in \mathsf{Z}^{k_n}} \|J^* - \Phi_r\|.$$

Proof. Note that $\Gamma_n u_n^* = u_n^*$ since u_n^* is constant over the sets $\mathcal{S}_{n,i}$. Then, we have

$$\begin{aligned}\|u_n^* - J^*\| &\leq \|u_n^* - \Gamma_n J^*\| + \|\Gamma_n J^* - J^*\| \\ &= \|F_n u_n^* - \Gamma_n T J^*\| + \|\Gamma_n J^* - J^*\| \\ &= \|\Gamma_n T u_n^* - \Gamma_n T J^*\| + \|\Gamma_n J^* - J^*\| \text{ (by the definition of } F_n) \\ &\leq \|T u_n^* - T J^*\| + \|\Gamma_n J^* - J^*\| \text{ (since } \|\Gamma_n u - \Gamma_n v\| \leq \|u - v\|) \\ &\leq \beta \|u_n^* - J^*\| + \|\Gamma_n J^* - J^*\|.\end{aligned}$$

Hence, we obtain $\|u_n^* - J^*\| \leq \frac{1}{1-\beta} \|\Gamma_n J^* - J^*\|$. The result now follows from Lemma 5.1. □

Proof. (Proof of Theorem 5.2) Recall that since Z is compact, the function J^* is uniformly continuous and $\mathrm{diam}(\mathcal{S}_{n,i}) < 2/n$ for all $i = 1, \ldots, k_n$. Hence, $\inf_{r \in \mathsf{Z}^{k_n}} \|J^* - \Phi_r\| \to 0$ as $n \to \infty$ which completes the proof in view of Lemma 5.2. □

The next step is to show that the fixed point $\hat{J}_n^*$ of $\hat{T}_n$ converges to the fixed point J^* of T. To this end, we first prove the following result.

Lemma 5.3. *For any $u \in C_b(\mathsf{Z})$, $\|\hat{T}_n u - F_n u\| \to 0$ as $n \to \infty$.*

Proof. Note that since $\int_{\mathsf{Z}} u(x)p(dx|y,a)$ is continuous as a function of (y,a) by Assumption 5.1-(b), it is sufficient to prove that for any $l \in C_b(\mathsf{Z} \times \mathsf{A})$

$$\begin{aligned}\Big\|\min_a \int l(y,a)\nu_{n,i_n(z)}(dy) - \int \min_a l(y,a)\nu_{n,i_n(z)}(dy)\Big\| \\ \coloneqq \sup_{z\in\mathsf{Z}} \Big|\min_a \int l(y,a)\nu_{n,i_n(z)}(dy) - \int \min_a l(y,a)\nu_{n,i_n(z)}(dy)\Big| \to 0\end{aligned}$$

as $n \to \infty$. Fix any $\varepsilon > 0$. Define $\{z_i\}_{i=1}^\infty \coloneqq \bigcup_n \mathsf{Z}_n$ and let $\{a_i\}_{i=1}^\infty$ be a sequence in A such that $\min_{a\in\mathsf{A}} l(z_i,a) = l(z_i,a_i)$; such a_i exists for each z_i because $l(z_i,\,\cdot\,)$ is continuous and A is compact. Define $g(y) \coloneqq \min_{a\in\mathsf{A}} l(y,a)$, which can be proved to be continuous, and therefore uniformly continuous since Z is compact. Thus by the uniform continuity of l, there exists $\delta > 0$ such that $d_{\mathsf{Z}\times\mathsf{A}}\big((y,a),(y',a')\big) < \delta$ implies $|g(y) - g(y')| < \varepsilon/2$ and $|l(y,a) - l(y',a')| < \varepsilon/2$. Choose n_0 such that $2/n_0 < \delta$. Then for all $n \geq n_0$, $\max_{i\in\{1,\ldots,k_n\}} \mathrm{diam}(\mathcal{S}_{n,i}) < 2/n < \delta$. Hence, for all $y \in \mathcal{S}_{n,i}$ we have $|l(y,a_i) - \min_{a\in\mathsf{A}} l(y,a)| \leq |l(y,a_i) - l(z_i,a_i)| + |\min_{a\in\mathsf{A}} l(z_i,a) - \min_{a\in\mathsf{A}} l(y,a)| = |l(y,a_i) - l(z_i,a_i)| + |g(z_i) - g(y)| < \varepsilon$. This implies

$$\Big\|\min_a \int l(y,a)\nu_{n,i_n(z)}(dy) - \int \min_a l(y,a)\nu_{n,i_n(z)}(dy)\Big\|$$

$$\leq \left\| \int l(y, a_i)\nu_{n,i_n(z)}(dy) - \int \min_a l(y,a)\nu_{n,i_n(z)}(dy) \right\|$$
$$\leq \sup_{z \in \mathsf{Z}} \int \sup_{y \in \mathcal{S}_{n,i_n(z)}} \left| l(y,a_i) - \min_a l(y,a) \right| \nu_{n,i_n(z)}(dy) < \varepsilon.$$

This completes the proof. □

Theorem 5.3. *The fixed point $\hat{J}_n^*$ of $\hat{T}_n$ converges to the fixed point J^* of T.*

Proof. We have

$$\|\hat{J}_n^* - J^*\| \leq \|\hat{T}_n \hat{J}_n^* - \hat{T}_n J^*\| + \|\hat{T}_n J^* - F_n J^*\| + \|F_n J^* - F_n u_n^*\| + \|F_n u_n^* - J^*\|$$
$$\leq \beta \|\hat{J}_n^* - J^*\| + \|\hat{T}_n J^* - F_n J^*\| + \beta \|J^* - u_n^*\| + \|u_n^* - J^*\|.$$

Hence

$$\|\hat{J}_n^* - J^*\| \leq \frac{\|\hat{T}_n J^* - F_n J^*\| + (1+\beta)\|J^* - u_n^*\|}{1-\beta}.$$

The theorem now follows from Theorem 5.2 and Lemma 5.3. □

Recall the optimal stationary policy f_n^* for MDP_n and its extension $\hat{f}_n(z) = f_n^* \circ Q_n(z)$ to Z. Since $\hat{J}_n^* = J_n^* \circ Q_n$, it is straightforward to prove that $\hat{f}_n$ is the optimal selector of $\hat{T}_n \hat{J}_n^*$; that is,

$$\hat{T}_n \hat{J}_n^* = \hat{J}_n^* = \hat{T}_{\hat{f}_n} \hat{J}_n^*,$$

where $\hat{T}_{\hat{f}_n}$ is defined as

$$\hat{T}_{\hat{f}_n} u(z) := \int \left[c(x, \hat{f}_n(x)) + \beta \int_{\mathsf{Z}} u(y) p(dy|x, \hat{f}_n(x)) \right] \nu_{n,i_n(z)}(dx).$$

Define analogously

$$T_{\hat{f}_n} u(z) := c(z, \hat{f}_n(z)) + \beta \int_{\mathsf{Z}} u(y) p(dy|z, \hat{f}_n(z)).$$

It can be proved that both $\hat{T}_{\hat{f}_n}$ and $T_{\hat{f}_n}$ are contraction operators on $B(\mathsf{Z})$ with modulus β, and it is known that the fixed point of $T_{\hat{f}_n}$ is the true cost function of the stationary policy $\hat{f}_n$ (i.e., $J(\hat{f}_n, z)$).

Lemma 5.4. *$\|\hat{T}_{\hat{f}_n} u - T_{\hat{f}_n} u\| \to 0$ as $n \to \infty$, for any $u \in C_b(\mathsf{Z})$.*

Proof. The statement follows from the uniform continuity of the function $c(z, a) + \beta \int_{\mathsf{Z}} u(y) p(dy|z, a)$ and the fact that $\hat{f}_n$ is constant over the sets $\mathcal{S}_{n,i}$. □

Now, we prove the main result of this section.

Proof. (Proof of Theorem 5.1) We have

$$\begin{aligned} \|J(\hat{f}_n, \cdot\,) - J^*\| &\leq \|T_{\hat{f}_n} J(\hat{f}_n, \cdot\,) - T_{\hat{f}_n} J^*\| + \|T_{\hat{f}_n} J^* - \hat{T}_{\hat{f}_n} J^*\| \\ &\qquad + \|\hat{T}_{\hat{f}_n} J^* - \hat{T}_{\hat{f}_n} \hat{J}_n^*\| + \|\hat{J}_n^* - J^*\| \\ &\leq \beta \|J(\hat{f}_n, \cdot\,) - J^*\| + \|T_{\hat{f}_n} J^* - \hat{T}_{\hat{f}_n} J^*\| + \beta \|J^* - \hat{J}_n^*\| + \|\hat{J}_n^* - J^*\|. \end{aligned}$$

Hence, we obtain

$$\|J(\hat{f}_n, \cdot\,) - J^*\| \leq \frac{\|T_{\hat{f}_n} J^* - \hat{T}_{\hat{f}_n} J^*\| + (1 + \beta) \|\hat{J}_n^* - J^*\|}{1 - \beta}.$$

The result follows from Lemma 5.4 and Theorem 5.3. □

5.2.2 Average Cost

In this section we impose some new conditions on the components of the original MDP in addition to Assumption 5.1 to solve **(Q1)** for the average cost. A version of first two conditions were imposed in [98] to show the existence of the solution to the Average Cost Optimality Equation (ACOE) and the optimal stationary policy by using the fixed point approach.

Assumption 5.2. Suppose Assumption 5.1 holds with item (b) replaced by condition (f) below. In addition, there exist a non-trivial finite measure ζ on Z, a nonnegative measurable function θ on $\mathsf{Z}\times\mathsf{A}$, and a constant $\lambda \in (0,1)$ such that for all $(z,a) \in \mathsf{Z}\times\mathsf{A}$

(d) $p(B|z,a) \geq \zeta(B)\theta(z,a)$ for all $B \in \mathcal{B}(\mathsf{Z})$,

(e) $\frac{1-\lambda}{\zeta(\mathsf{Z})} \leq \theta(z,a)$,

(f) The stochastic kernel $p(\,\cdot\,|z,a)$ is continuous in (z,a) with respect to the total variation distance.

Recall that any deterministic stationary policy f defines a stochastic kernel $p(\,\cdot\,|z,f(z))$ on Z given Z which is the transition probability of the Markov chain $\{Z_t\}_{t=1}^{\infty}$ (state process) induced by f. In this chapter, instead of the notation introduced in Section 4.3.2, we write $p^t(\,\cdot\,|z,f(z))$ to denote the t-step ($t \geq 1$) transition probability of this Markov chain given the initial point z. The reason for using this notation is that we need to distinguish $p^t(\,\cdot\,|z,f(z))$ from $p^t(\,\cdot\,|z,f(y))$ for $z \neq y$ in the sequel.

The following theorem is a consequence of [44, Lemma 3.4 and Theorem 2.6] and [98, Theorems 3.3], which also holds with Assumption 5.2-(f) replaced by Assumption 5.1-(b).

Theorem 5.4. *Under Assumptions 5.2 the following holds.*

(i) For each $f \in \mathbb{F}$, the stochastic kernel $p(\,\cdot\,|z, f(z))$ is positive Harris recurrent with unique invariant probability measure μ_f. Therefore, we have

$$V(f, z) = \int_{\mathsf{Z}} c(z, f(z))\mu_f(dz) =: \rho_f.$$

(ii) There exist positive real numbers R and $\kappa < 1$ such that for every $z \in \mathsf{Z}$

$$\sup_{f \in \mathbb{F}} \|p^t(\,\cdot\,|z, f(z)) - \mu_f\|_{TV} \leq R\kappa^t,$$

where R and κ continuously depend on $\zeta(\mathsf{Z})$ and λ.

(iii) There exist $f^* \in \mathbb{F}$ and $h^* \in B(\mathsf{Z})$ such that the triplet (h^*, f^*, ρ_{f^*}) satisfies the average cost optimality inequality (ACOI), i.e.,

$$\begin{aligned} \rho_{f^*} + h^*(z) &\geq \min_{a \in \mathsf{A}} \left[c(z, a) + \int_{\mathsf{Z}} h^*(y) p(dy|z, a) \right] \\ &= c(z, f^*(z)) + \int_{\mathsf{Z}} h^*(y) p(dy|z, f^*(z)), \end{aligned}$$

and therefore,

$$\inf_{\pi \in \Pi} V(\pi, z) =: V^*(z) = \rho_{f^*}.$$

For each n, define the one-stage cost function $b_n : \mathsf{Z} \times \mathsf{A} \to [0, \infty)$ and the stochastic kernel q_n on Z given $\mathsf{Z} \times \mathsf{A}$ as

$$b_n(z, a) := \int c(x, a) \nu_{n, i_n(z)}(dx),$$

$$q_n(\,\cdot\,|z,a) \coloneqq \int p(\,\cdot\,|x,a)\nu_{n,i_n(z)}(dx).$$

Observe that c_n (i.e., the one stage cost function of MDP_n) is the restriction of b_n to Z_n, and p_n (i.e., the stochastic kernel of MDP_n) is the pushforward of the measure q_n with respect to Q_n; that is, $c_n(z_{n,i},a) = b_n(z_{n,i},a)$ for all $i = 1,\ldots,k_n$ and $p_n(\,\cdot\,|z_{n,i},a) = Q_n * q_n(\,\cdot\,|z_{n,i},a)$.

For each n, let $\widehat{\mathrm{MDP}}_n$ be defined as a Markov decision process with the following components: Z is the state space, A is the action space, q_n is the transition probability, and c is the one-stage cost function. Similarly, let $\widetilde{\mathrm{MDP}}_n$ be defined as a Markov decision process with the following components: Z is the state space, A is the action space, q_n is the transition probability, and b_n is the one-stage cost function. History spaces, policies and cost functions are defined in a similar way as before. The models $\widehat{\mathrm{MDP}}_n$ and $\widetilde{\mathrm{MDP}}_n$ are used as transitions between the original MDP and MDP_n in a similar way as the operators F_n and $\hat{T}_n$ were used as transitions between T and T_n for the discounted cost. We note that a similar technique was used in the proof of [72, Theorem 2], which studied the approximation problem for finite state-action MDPs. In [72] the one-stage cost function is first perturbed and then the transition probability is perturbed. We first perturb the transition probability and then the cost function. However, our proof method is otherwise quite different from that of [72, Theorem 2] since [72] assumes finite state and action spaces.

We note that a careful analysis of $\widetilde{\mathrm{MDP}}_n$ reveals that its Bellman optimality operator is essentially the operator $\hat{T}_n$. Hence, the value function of $\widetilde{\mathrm{MDP}}_n$ is the piecewise constant extension of the value function of MDP_n for the discounted cost. A similar conclusion will be made for the average cost in Lemma 5.5.

First, notice that if we define

$$\theta_n(z,a) \coloneqq \int \theta(y,a)\nu_{n,i_n(z)}(dy),$$

$$\zeta_n \coloneqq Q_n * \zeta \text{ (i.e., pushforward of } \zeta \text{ with respect to } Q_n),$$

then it is straightforward to prove that for all n, both $\widehat{\text{MDP}}_n$ and $\widetilde{\text{MDP}}_n$ satisfy Assumption 5.2-(d),(e) when θ is replaced by θ_n, and Assumption 5.2-(d),(e) is true for MDP_n when θ and ζ are replaced by the restriction of θ_n to Z_n and ζ_n, respectively.

Hence, Theorem 5.4 holds (with the same R and κ) for $\widehat{\text{MDP}}_n$, $\widetilde{\text{MDP}}_n$, and MDP_n for all n. Therefore, we denote by $\hat{f}_n^*$, $\tilde{f}_n^*$ and f_n^* the optimal stationary policies of $\widehat{\text{MDP}}_n$, $\widetilde{\text{MDP}}_n$, and MDP_n with the corresponding average costs $\hat{\rho}^n_{\hat{f}_n^*}$, $\tilde{\rho}^n_{\tilde{f}_n^*}$ and $\rho^n_{f_n^*}$, respectively.

Furthermore, we also write $\hat{\rho}^n_f$, $\tilde{\rho}^n_f$, and ρ^n_f to denote the average cost of any stationary policy f for $\widehat{\text{MDP}}_n$, $\widetilde{\text{MDP}}_n$, and MDP_n, respectively. The corresponding invariant probability measures are also denoted in a same manner, with μ replacing ρ.

The following lemma essentially says that MDP_n and $\widetilde{\text{MDP}}_n$ are not very different.

Lemma 5.5. *The stationary policy given by the piecewise constant extension of the optimal policy f_n^* of MDP_n to Z (i.e., $f_n^* \circ Q_n$) is optimal for $\widetilde{MDP}_n$ with the same cost function $\rho^n_{f_n^*}$. Hence, $\tilde{f}_n^* = f_n^* \circ Q_n$ and $\tilde{\rho}^n_{\tilde{f}_n^*} = \rho^n_{f_n^*}$.*

Proof. Note that by Theorem 5.4 there exists $h_n^* \in B(\mathsf{Z}_n)$ such that the triplet $(h_n^*, f_n^*, \rho^n_{f_n^*})$ satisfies the ACOI for MDP_n. But it is straightforward to show that the triplet $(h_n^* \circ Q_n, f_n^* \circ Q_n, \rho^n_{f_n^*})$ satisfies the ACOI for $\widetilde{\text{MDP}}_n$. By [44, Theorem 2.6 and Lemma 5.2], this implies that $f_n^* \circ Q_n$ is an optimal stationary policy for $\widetilde{\text{MDP}}_n$

with cost function $\rho^n_{f^*_n}$. Hence $\tilde{f}^*_n = f^*_n \circ Q_n$ and $\tilde{\rho}^n_{\tilde{f}^*_n} = \rho^n_{f^*_n}$. $\square$

The following theorem is the main result of this section. It states that if one applies the piecewise constant extension of the optimal stationary policy of MDP_n to the original MDP, the resulting cost function will converge to the value function of the original MDP.

Theorem 5.5. *The average cost of the optimal policy $\tilde{f}^*_n$ for $\widetilde{MDP}_n$, obtained by extending the optimal policy f^*_n of MDP_n to Z, converges to the optimal value function $J^* = \rho_{f^*}$ of the original MDP, i.e.,*

$$\lim_{n\to\infty} |\rho_{\tilde{f}^*_n} - \rho_{f^*}| = 0.$$

Hence, to find a near optimal policy for the original MDP, it is sufficient to compute the optimal policy of MDP_n for sufficiently large n, and then extend this policy to the original state space.

To show the statement of Theorem 5.5 we will prove a series of auxiliary results.

Lemma 5.6. *For all $t \geq 1$ we have*

$$\lim_{n\to\infty} \sup_{(y,f)\in\mathsf{Z}\times\mathbb{F}} \big\| p^t(\,\cdot\,|y, f(y)) - q^t_n(\,\cdot\,|y, f(y)) \big\|_{TV} = 0.$$

Proof. See Section 5.8.1. $\square$

Using Lemma 5.6 we prove the following result.

Lemma 5.7. *We have $\sup_{f\in\mathbb{F}} |\hat{\rho}^n_f - \rho_f| \to 0$ as $n \to \infty$, where $\hat{\rho}^n_f$ is the cost function of the policy f for $\widehat{MDP}_n$ and ρ_f is the cost function of the policy f for the original MDP.*

Proof. For any $t \geq 1$ and $y \in \mathsf{Z}$ we have

$$\begin{aligned}
\sup_{f \in \mathbb{F}} |\hat{\rho}_f^n - \rho_f| &= \sup_{f \in \mathbb{F}} \left| \int_{\mathsf{Z}} c(z, f(z)) \hat{\mu}_f^n(dz) - \int_{\mathsf{Z}} c(z, f(z)) \mu_f(dz) \right| \\
&\leq \sup_{f \in \mathbb{F}} \left| \int_{\mathsf{Z}} c(z, f(z)) \hat{\mu}_f^n(dz) - \int_{\mathsf{Z}} c(z, f(z)) q_n^t(dz|y, f(y)) \right| \\
&\quad + \sup_{f \in \mathbb{F}} \left| \int_{\mathsf{Z}} c(z, f(z)) q_n^t(dz|y, f(y)) - \int_{\mathsf{Z}} c(z, f(z)) p^t(dz|y, f(y)) \right| \\
&\quad + \sup_{f \in \mathbb{F}} \left| \int_{\mathsf{Z}} c(z, f(z)) p^t(dz|y, f(y)) - \int_{\mathsf{Z}} c(z, f(z)) \mu_f(dz) \right| \\
&\leq 2R\kappa^t \|c\| + \|c\| \sup_{(y,f) \in \mathsf{Z} \times \mathbb{F}} \left\| q_n^t(\,\cdot\,|y, f(y)) - p^t(\,\cdot\,|y, f(y)) \right\|_{TV} \text{ (by Theorem 5.4-(ii))},
\end{aligned}$$

where R and κ are the constants in Theorem 5.4. Then, the result follows from Lemma 5.6. □

The following theorem states that the value function of $\widehat{\text{MDP}}_n$ converges to the value function of the original MDP.

Lemma 5.8. *We have* $|\hat{\rho}^n_{\hat{f}_n^*} - \rho_{f^*}| \to 0$ *as* $n \to \infty$.

Proof. Notice that

$$\begin{aligned}
|\hat{\rho}^n_{\hat{f}_n^*} - \rho_{f^*}| &= \max(\hat{\rho}^n_{\hat{f}_n^*} - \rho_{f^*}, \rho_{f^*} - \hat{\rho}^n_{\hat{f}_n^*}) \\
&\leq \max(\hat{\rho}^n_{f^*} - \rho_{f^*}, \rho_{\hat{f}_n^*} - \hat{\rho}^n_{\hat{f}_n^*}) \\
&\leq \sup_f |\hat{\rho}^n_f - \rho_f|.
\end{aligned}$$

Then, the result follows from Lemma 5.7. □

Lemma 5.9. *We have* $\sup_{f \in \mathbb{F}} |\tilde{\rho}^n_f - \hat{\rho}^n_f| \to 0$ *as* $n \to \infty$.

Proof. It is straightforward to show that $b_n \to c$ uniformly. Since the probabilistic structure of $\widetilde{\text{MDP}}_n$ and $\widehat{\text{MDP}}_n$ are the same (i.e., $\hat{\mu}_f^n = \tilde{\mu}_f^n$ for all f), we have

$$\begin{aligned}\sup_{f \in \mathbb{F}} |\tilde{\rho}_f^n - \hat{\rho}_f^n| &= \sup_{f \in \mathbb{F}} \left| \int_{\mathsf{Z}} b_n(z, f(z)) \hat{\mu}_f^n(dz) - \int_{\mathsf{Z}} c(z, f(z)) \hat{\mu}_f^n(dz) \right| \\ &\leq \sup_{f \in \mathbb{F}} \int_{\mathsf{Z}} |b_n(z, f(z)) - c(z, f(z))| \hat{\mu}_f^n(dz) \\ &\leq \|b_n - c\|.\end{aligned}$$

This completes the proof. □

The next lemma states that the difference between the value functions of $\widetilde{\text{MDP}}_n$ and $\widehat{\text{MDP}}_n$ converges to zero.

Lemma 5.10. *We have* $|\tilde{\rho}_{\tilde{f}_n^*}^n - \hat{\rho}_{\hat{f}_n^*}^n| \to 0$ *as* $n \to \infty$.

Proof. See the proof of Lemma 5.8. □

The following result states that if we apply the optimal policy of $\widetilde{\text{MDP}}_n$ to $\widehat{\text{MDP}}_n$, then the resulting cost converges to the value function of $\widehat{\text{MDP}}_n$.

Lemma 5.11. *We have* $|\hat{\rho}_{\tilde{f}_n^*}^n - \hat{\rho}_{\hat{f}_n^*}^n| \to 0$ *as* $n \to \infty$.

Proof. Since $|\hat{\rho}_{\tilde{f}_n^*}^n - \hat{\rho}_{\hat{f}_n^*}^n| \leq |\hat{\rho}_{\tilde{f}_n^*}^n - \tilde{\rho}_{\tilde{f}_n^*}^n| + |\tilde{\rho}_{\tilde{f}_n^*}^n - \hat{\rho}_{\hat{f}_n^*}^n|$, then the result follows from Lemmas 5.9 and 5.10. □

Now, we are ready to prove the main result of this section.

Proof. (Proof of Theorem 5.5) We have $|\rho_{\tilde{f}_n^*} - \rho_{f^*}| \leq |\rho_{\tilde{f}_n^*} - \hat{\rho}_{\tilde{f}_n^*}^n| + |\hat{\rho}_{\tilde{f}_n^*}^n - \hat{\rho}_{\hat{f}_n^*}^n| + |\hat{\rho}_{\hat{f}_n^*}^n - \rho_{f^*}^n|$. The result now follows from Lemmas 5.7, 5.11 and 5.8. □

5.3 Finite State Approximation of Non-Compact State MDPs

In this section we consider **(Q1)** for noncompact state MDPs with unbounded one-stage cost. To solve **(Q1)**, we use the following strategy: (i) first, we define a sequence of compact-state MDPs to approximate the original MDP, (ii) we use Theorems 5.1 and 5.5 to approximate the compact-state MDPs by finite-state models, and (iii) we prove the convergence of the finite-state models to the original model. In fact, steps (ii) and (iii) will be accomplished simultaneously.

We impose the assumptions below on the components of the Markov decision process; additional assumptions will be imposed for the average cost problem.

Assumption 5.3.

(a) The one-stage cost function c is nonnegative and continuous.

(b) The stochastic kernel $p(\,\cdot\,|x,a)$ is weakly continuous in (x,a) and setwise continuous in a.

(c) X is locally compact and A is compact.

(d) There exist nonnegative real numbers M and $\alpha \in [1, \frac{1}{\beta})$, and a continuous weight function $w : \mathsf{X} \to [1,\infty)$ such that for each $x \in \mathsf{X}$, we have

$$\sup_{a \in \mathsf{A}} c(x,a) \leq M w(x), \tag{5.4}$$

$$\sup_{a \in \mathsf{A}} \int_{\mathsf{X}} w(y) p(dy|x,a) \leq \alpha w(x), \tag{5.5}$$

and $\int_{\mathsf{X}} w(y) p(dy|x,a)$ is continuous in (x,a).

Since X is locally compact separable metric space, there exists a nested sequence of compact sets $\{K_n\}$ such that $K_n \subset \operatorname{int} K_{n+1}$ and $\mathsf{X} = \bigcup_{n=1}^{\infty} K_n$ [5, Lemma 2.76].

Let $\{\nu_n\}$ be a sequence of probability measures such that for each $n \geq 1$, $\nu_n \in \mathcal{P}(K_n^c)$ and

$$\gamma_n \coloneqq \int_{K_n^c} w(x)\nu_n(dx) < \infty, \tag{5.6}$$

$$\gamma = \sup_n \tau_n \coloneqq \sup_n \max\biggl\{0, \sup_{(x,a)\in \mathsf{X}\times\mathsf{A}} \int_{K_n^c} \bigl(\gamma_n - w(y)\bigr)\, p(dy|x,a)\biggr\} < \infty. \tag{5.7}$$

For example, such probability measures can be constructed by choosing $x_n \in K_n^c$ such that $w(x_n) < \inf_{x \in K_n^c} w(x) + \frac{1}{n}$ and letting $\nu_n(\,\cdot\,) = \delta_{x_n}(\,\cdot\,)$.

Similar to the finite-state MDP construction in Section 5.2, we define a sequence of compact-state MDPs, denoted as c-MDP$_n$, to approximate the original model. To this end, for each n let $\mathsf{X}_n = K_n \cup \{\Delta_n\}$, where $\Delta_n \in K_n^c$ is a so-called pseudo-state. We define the transition probability p_n on X_n given $\mathsf{X}_n \times \mathsf{A}$ and the one-stage cost function $c_n : \mathsf{X}_n \times \mathsf{A} \to [0,\infty)$ by

$$p_n(\,\cdot\,|x,a) = \begin{cases} p\bigl(\,\cdot\, \cap K_n|x,a\bigr) + p\bigl(K_n^c|x,a\bigr)\delta_{\Delta_n}, & \text{if } x \in K_n \\ \int_{K_n^c} \Bigl(p\bigl(\,\cdot\, \cap K_n|z,a\bigr) + p\bigl(K_n^c|z,a\bigr)\delta_{\Delta_n}\Bigr)\nu_n(dz), & \text{if } x = \Delta_n, \end{cases}$$

$$c_n(x,a) = \begin{cases} c(x,a), & \text{if } x \in K_n \\ \int_{K_n^c} c(z,a)\nu_n(dz), & \text{if } x = \Delta_n. \end{cases}$$

With these definitions, c-MDP$_n$ is defined as a Markov decision process with the components $\bigl(\mathsf{X}_n, \mathsf{A}, p_n, c_n\bigr)$. History spaces, policies, and cost functions are defined in a similar way as in the original model. Let Π_n, Φ_n, and $\mathbb{F}_n$ denote the set of all policies, randomized stationary policies and deterministic stationary policies of c-MDP$_n$, respectively. For each policy $\pi \in \Pi_n$ and initial distribution $\mu \in \mathcal{P}(\mathsf{X}_n)$, we denote the cost functions for c-MDP$_n$ by $J_n(\pi,\mu)$ and $V_n(\pi,\mu)$.

To obtain the main result of this section, we introduce, for each n, another MDP, denoted by $\overline{\text{MDP}}_n$, with the components $\bigl(\mathsf{X},\mathsf{A},q_n,b_n\bigr)$ where

$$q_n(\,\cdot\,|x,a) = \begin{cases} p(\,\cdot\,|x,a), & \text{if } x \in K_n \\ \int_{K_n^c} p\bigl(\,\cdot\,|z,a\bigr)\nu_n(dz), & \text{if } x \in K_n^c, \end{cases}$$

$$b_n(x,a) = \begin{cases} c(x,a), & \text{if } x \in K_n \\ \int_{K_n^c} c(z,a)\nu_n(dz), & \text{if } x \in K_n^c. \end{cases}$$

For each policy $\pi \in \Pi$ and initial distribution $\mu \in \mathcal{P}(\mathsf{X})$, we denote the cost functions for $\overline{\text{MDP}}_n$ by $\bar{J}_n(\pi,\mu)$ and $\bar{V}_n(\pi,\mu)$.

5.3.1 Discounted Cost

In this section we consider **(Q1)** for the discounted cost criterion with a discount factor $\beta \in (0,1)$. The following result states that c-MDP$_n$ and $\overline{\text{MDP}}_n$ are equivalent for the discounted cost.

Lemma 5.12. *We have*

$$\bar{J}_n^*(x) = \begin{cases} J_n^*(x), & \text{if } x \in K_n \\ J_n^*(\Delta_n), & \text{if } x \in K_n^c, \end{cases} \tag{5.8}$$

where $\bar{J}_n^*$ is the discounted value function of $\overline{MDP}_n$ and J_n^* is the discounted value *function of c-MDP$_n$, provided that there exist optimal deterministic stationary policies for $\overline{\text{MDP}}_n$ and c-MDP$_n$. Furthermore, if, for any deterministic stationary policy $f \in \mathbb{F}_n$, we define $\bar{f}(x) = f(x)$ on K_n and $\bar{f}(x) = f(\Delta_n)$ on K_n^c, then*

$$\bar{J}_n(\bar{f}, x) = \begin{cases} J_n(f, x), & \text{if } x \in K_n \\ J_n(f, \Delta_n), & \text{if } x \in K_n^c. \end{cases} \tag{5.9}$$

In particular, if the deterministic stationary policy $f_n^* \in \mathbb{F}_n$ is optimal for c-MDP$_n$, then its extension $\bar{f}_n^*$ to X is also optimal for $\overline{\text{MDP}}_n$.

Proof. The proof of (5.9) is a consequence of the following facts: $b_n(x, a) = b_n(y, a)$ and $q_n(\,\cdot\,|x, a) = q_n(\,\cdot\,|y, a)$ for all $x, y \in K_n^c$ and $a \in \mathsf{A}$. In other words, K_n^c in $\overline{\text{MDP}}_n$ behaves like the pseudo state Δ_n in c-MDP$_n$ when $\bar{f}$ is applied to $\overline{\text{MDP}}_n$.

Let $\overline{\mathbb{F}}_n$ denote the set of all deterministic stationary policies in $\mathbb{F}$ which are obtained by extending policies in $\mathbb{F}_n$ to X. If we can prove that $\min_{f \in \mathbb{F}} \bar{J}_n(f, x) = \min_{f \in \overline{\mathbb{F}}_n} \bar{J}_n(f, x)$ for all $x \in \mathsf{X}$, then (5.8) follows from (5.9). Let $f \in \mathbb{F} \setminus \overline{\mathbb{F}}_n$. We have two cases: (i) $\bar{J}_n(f, z) = \bar{J}_n(f, y)$ for all $z, y \in K_n^c$ or (ii) there exists $z, y \in K_n^c$ such that $\bar{J}_n(f, z) < \bar{J}_n(f, y)$.

For the case (i), if we define the deterministic Markov policy π^0 as $\pi^0 = \{f_0, f, f, \ldots\}$, where $f_0(x) = f(z)$ on K_n^c for some fixed $z \in K_n^c$ and $f_0(x) = f(x)$ on K_n, then using the expression

$$\bar{J}_n(\pi^0, x) = b_n(x, f_0(x)) + \beta \int_{\mathsf{X}} \bar{J}_n(f, x') q_n(dx'|x, f_0(x)), \tag{5.10}$$

it is straightforward to show that $\bar{J}_n(\pi^0, x) = \bar{J}_n(f, x)$ on K_n and $\bar{J}_n(\pi^0, x) = \bar{J}_n(f, z)$ on K_n^c. Therefore, $\bar{J}_n(\pi^0, x) = \bar{J}_n(f, x)$ for all $x \in \mathsf{X}$ since $\bar{J}_n(f, x) = \bar{J}_n(f, z)$ for all $x \in K_n^c$. For all $t \geq 1$ define the deterministic Markov policy π^t as $\pi^t = \{f_0, \pi^{t-1}\}$. Analogously, one can prove that $\bar{J}_n(\pi^t, x) = \bar{J}_n(\pi^{t+1}, x)$ for all $x \in \mathsf{X}$. Since $\bar{J}_n(\pi^t, x) \to \bar{J}_n(f_0, x)$ as $t \to \infty$, we have $\bar{J}_n(f_0, x) = \bar{J}_n(f, x)$ for all $x \in \mathsf{X}$, where $f_0 \in \overline{\mathbb{F}}_n$.

For the second case, if we again consider the deterministic Markov policy $\pi^0 = \{f_0, f, f, \ldots\}$, then by (5.10) we have $\bar{J}_n(\pi^0, y) = \bar{J}_n(f, z) < \bar{J}_n(f, y)$. Since

$$\min_{f \in \mathbb{F}} \bar{J}_n(f, y) \leq \bar{J}_n(\pi^0, y),$$

this completes the proof. □

For each n, let us define w_n by letting $w_n(x) = w(x)$ on K_n and $w_n(x) = \int_{K_n^c} w(z)\nu_n(dz) =: \gamma_n$ on K_n^c. Hence, $w_n \in B(\mathsf{X})$ by (5.6).

Lemma 5.13. *For all n and $x \in \mathsf{X}$, the components of $\overline{MDP}_n$ satisfy the following:*

$$\sup_{a \in \mathsf{A}} b_n(x, a) \leq M w_n(x) \tag{5.11}$$

$$\sup_{a \in \mathsf{A}} \int_{\mathsf{X}} w_n(y) q_n(dy|x, a) \leq \alpha w_n(x) + \gamma, \tag{5.12}$$

where γ is the constant in (5.7).

Proof. See Section 5.8.2. □

Note that if we define $c_{n,0}(x) = 1 + \sup_{a \in \mathsf{A}} b_n(x, a)$ and

$$c_{n,t}(x) = \sup_{a \in \mathsf{A}} \int_{\mathsf{X}} c_{n,t-1}(y) q_n(dy|x, a),$$

by (5.11) and (5.12), and an induction argument, we obtain (see [55, p. 46])

$$c_{n,t}(x) \leq L w_n(x)\alpha^t + L\gamma \sum_{j=0}^{t-1} \alpha^j \quad \text{for all } x \in \mathsf{X}, \tag{5.13}$$

where $L = 1 + M$. Let $\beta_0 > \beta$ be such that $\alpha\beta_0 < 1$ and let $C_n : \mathsf{X} \to [1, \infty)$ be defined by

$$C_n(x) = \sum_{t=0}^{\infty} \beta_0^t c_{n,t}(x).$$

Then, for all $x \in \mathsf{X}$, by (5.13) we have

$$\begin{aligned} C_n(x) \coloneqq \sum_{t=0}^{\infty} \beta_0^t c_{n,t}(x) &\leq \frac{L}{1-\beta_0\alpha} w_n(x) + \frac{L\beta_0}{(1-\beta_0)(1-\beta_0\alpha)}\gamma \\ &\coloneqq L_1 w_n(x) + L_2. \end{aligned} \tag{5.14}$$

Hence $C_n \in B(\mathsf{X})$ as $w_n \in B(\mathsf{X})$. Moreover, for all $(x, a) \in \mathsf{X} \times \mathsf{A}$, C_n satisfies (see [55, p. 45])

$$\begin{aligned} \int_{\mathsf{X}} C_n(y) q_n(dy|x,a) &= \sum_{t=0}^{\infty} \beta_0^t \int_{\mathsf{X}} c_{n,t}(y) q_n(dy|x,a) \\ &\leq \sum_{t=0}^{\infty} \beta_0^t c_{n,t+1}(x) \\ &\leq \frac{1}{\beta_0} \sum_{t=0}^{\infty} \beta_0^t c_{n,t}(x) = \alpha_0 C_n(x), \end{aligned}$$

where $\alpha_0 \coloneqq \frac{1}{\beta_0}$ and $\alpha_0\beta < 1$ since $\beta_0 > \beta$. Therefore, for all $x \in \mathsf{X}$, components of

$\overline{\text{MDP}}_n$ satisfy

$$\sup_{a \in \mathsf{A}} b_n(x,a) \leq C_n(x) \tag{5.15}$$

$$\sup_{a \in \mathsf{A}} \int_{\mathsf{X}} C_n(y) q_n(dy|x,a) \leq \alpha_0 C_n(x). \tag{5.16}$$

Since (i) $b_n(x,a)$ is continuous in a for all $x \in \mathsf{X}$, (ii) $q_n(\,\cdot\,|x,a)$ is setwise continuous in a for all $x \in \mathsf{X}$, (iii) $C_n \in B(\mathsf{X})$, and (iv) $\alpha_0 \beta < 1$, $\overline{\text{MDP}}_n$ satisfies the assumptions in [55, Theorem 8.3.6, p. 47].

Let us define the Bellman optimality operator $\overline{T}_n : B(\mathsf{X}) \to B(\mathsf{X})$ (note that $B_{C_n}(\mathsf{X}) = B(\mathsf{X})$) for $\overline{\text{MDP}}_n$ by

$$\begin{aligned}
\overline{T}_n u(x) &= \min_{a \in \mathsf{A}} \biggl[b_n(x,a) + \beta \int_{\mathsf{X}} u(y) q_n(dy|x,a) \biggr] \\
&= \begin{cases} \min_{a \in \mathsf{A}} \bigl[c(x,a) + \beta \int_{\mathsf{X}} u(y) p(dy|x,a) \bigr], & \text{if } x \in K_n \\ \min_{a \in \mathsf{A}} \int_{K_n^c} \bigl[c(z,a) + \beta \int_{\mathsf{X}} u(y) p(dy|z,a) \bigr] \nu_n(dz), & \text{if } x \in K_n^c. \end{cases}
\end{aligned}$$

Then successive approximations to the discounted value function of $\overline{\text{MDP}}_n$ are given by $v_n^0 = 0$ and $v_n^{t+1} = \overline{T}_n v_n^t$ ($t \geq 1$). By [55, Theorem 8.3.6, p. 47] and [55, (8.3.34), p. 52] we have

$$v_n^t(x) \leq \bar{J}_n^*(x) \leq \frac{C_n(x)}{1 - \sigma_0} \text{ for all } x, \tag{5.17}$$

$$\|v_n^t - \bar{J}_n^*\|_{C_n} \leq \frac{\sigma_0^t}{1 - \sigma_0^t}, \tag{5.18}$$

where $\sigma_0 = \beta \alpha_0 < 1$.

Similar to v_n^t, let us define $v^0 = 0$ and $v^{t+1} = T v^t$, where $T : B_w(\mathsf{X}) \to B_w(\mathsf{X})$, the

Bellman optimality operator for the original MDP, is given by

$$Tu(x) = \min_{a \in \mathsf{A}} \biggl[c(x,a) + \beta \int_{\mathsf{X}} u(y) p(dy|x,a) \biggr].$$

Then, again by [55, Theorem 8.3.6, p. 47] and [55, (8.3.34), p. 52] we have

$$v^t(x) \leq J^*(x) \leq M \frac{w(x)}{1-\sigma} \text{ for all } x, \tag{5.19}$$

$$\|v^t - J^*\|_w \leq M \frac{\sigma^t}{1-\sigma}, \tag{5.20}$$

where $\sigma = \beta\alpha < 1$.

Lemma 5.14. *For any compact set $K \subset \mathsf{X}$, we have*

$$\lim_{n\to\infty} \sup_{x \in K} |v_n^t(x) - v^t(x)| = 0. \tag{5.21}$$

Proof. We prove (5.21) by induction on t. For $t = 1$, the claim trivially holds since any compact set $K \subset \mathsf{X}$ is inside K_n for sufficiently large n, and therefore, $b_n = c$ on K for sufficiently large n (recall $v_n^0 = v^0 = 0$). Assume the claim is true for $t \geq 1$. Fix any compact set K. Recall the definition of compact subsets K_ε of X in Lemma 4.3. By definition of q_n, b_n, and w_n, there exists $n_0 \geq 1$ such that for all $n \geq n_0$, $q_n = p$, $b_n = c$, and $w_n = w$ on K. With these observations, for each $n \geq n_0$ we have

$$\begin{aligned}
&\sup_{x\in K} |v_n^{t+1}(x) - v^{t+1}(x)| \\
&= \sup_{x \in K} \biggl| \min_{\mathsf{A}} \biggl[c(x,a) + \beta \int_{\mathsf{X}} v_n^t(y) p(dy|x,a) \biggr] - \min_{\mathsf{A}} \biggl[c(x,a) + \beta \int_{\mathsf{X}} v^t(y) p(dy|x,a) \biggr] \biggr| \\
&\leq \beta \sup_{(x,a) \in K \times \mathsf{A}} \biggl| \int_{\mathsf{X}} v_n^t(y) p(dy|x,a) - \int_{\mathsf{X}} v^t(y) p(dy|x,a) \biggr|
\end{aligned}$$

$$= \beta \sup_{(x,a)\in K\times \mathsf{A}} \left| \int_{K_\varepsilon} \big(v_n^t(y) - v^t(y)\big)\, p(dy|x,a) + \int_{K_\varepsilon^c} \big(v_n^t(y) - v^t(y)\big)\, p(dy|x,a) \right|$$
$$\leq \beta \bigg\{ \sup_{x\in K_\varepsilon} |v_n^t(x) - v^t(x)| + \sup_{(x,a)\in K\times \mathsf{A}} \left| \int_{K_\varepsilon^c} \big(v_n^t(y) - v^t(y)\big)\, p(dy|x,a) \right| \bigg\}$$

Note that we have $v^t \leq M \frac{w}{1-\sigma}$ by (5.19). Since $w_n \leq \gamma_{\max} w$, where $\gamma_{\max} := \max\{1, \gamma\}$, we also have $v_n^t \leq \frac{L_1 \gamma_{\max} w + L_2}{1-\sigma_0} \leq \frac{(L_1 \gamma_{\max} + L_2) w}{1-\sigma_0}$ by (5.14) and (5.17) (as $w \geq 1$). Let us define

$$R := \frac{L_1 \gamma_{\max} + L_2}{1-\sigma_0} + \frac{M}{1-\sigma}.$$

Then by Lemma 4.3 we have

$$\sup_{x\in K} |v_n^{t+1}(x) - v^{t+1}(x)| \leq \beta \sup_{x\in K_\varepsilon} |v_n^t(x) - v^t(x)| + \beta R \varepsilon.$$

Since the first term converges to zero as $n \to \infty$ by the induction hypothesis, and ε is arbitrary, the claim is true for $t+1$. This completes the proof. □

The following theorem states that the discounted value function of $\overline{\mathrm{MDP}}_n$ converges to the discounted value function of the original MDP uniformly on each compact set $K \subset \mathsf{X}$.

Theorem 5.6. *For any compact set $K \subset \mathsf{X}$ we have*

$$\lim_{n\to\infty} \sup_{x\in K} |\bar{J}_n^*(x) - J^*(x)| = 0. \tag{5.22}$$

Proof. Fix any compact set $K \subset \mathsf{X}$. Since w is continuous and therefore bounded on K, it is sufficient to prove $\lim_{n\to\infty} \sup_{x\in K} \frac{|\bar{J}_n^*(x) - J^*(x)|}{w(x)}$. Let n be chosen such that

$K \subset K_n$, and so, $w_n = w$ on K. Then we have

$$
\begin{aligned}
&\sup_{x\in K} \frac{|\bar{J}_n^*(x) - J^*(x)|}{w(x)} \\
&\leq \sup_{x\in K} \frac{|\bar{J}_n^*(x) - v_n^t(x)|}{w(x)} + \sup_{x\in K} \frac{|v_n^t(x) - v^t(x)|}{w(x)} + \sup_{x\in K} \frac{|v^t(x) - J^*(x)|}{w(x)} \\
&\leq \sup_{x\in K} \frac{|\bar{J}_n^*(x) - v_n^t(x)|}{C_n(x)} \frac{C_n(x)}{w(x)} + \sup_{x\in K} \frac{|v_n^t(x) - v^t(x)|}{w(x)} + M \frac{\sigma^t}{1-\sigma^t} \quad \text{(by (5.20))} \\
&\leq \sup_{x\in K} \frac{|\bar{J}_n^*(x) - v_n^t(x)|}{C_n(x)} \frac{(L_1 w_n(x) + L_2)}{w(x)} + \sup_{x\in K} \frac{|v_n^t(x) - v^t(x)|}{w(x)} + \frac{M\sigma^t}{1-\sigma^t} \quad \text{(by (5.14))} \\
&\leq (L_1 + L_2) \sup_{x\in K} \frac{|\bar{J}_n^*(x) - v_n^t(x)|}{C_n(x)} + \sup_{x\in K} \frac{|v_n^t(x) - v^t(x)|}{w(x)} + \frac{M\sigma^t}{1-\sigma^t} \quad (w_n = w \text{ on } K) \\
&\leq (L_1 + L_2) \frac{\sigma_0^t}{1-\sigma_0} + \sup_{x\in K} \frac{|v_n^t(x) - v^t(x)|}{w(x)} + \frac{M\sigma^t}{1-\sigma^t} \quad \text{(by (5.18))}.
\end{aligned}
$$

Since $w \geq 1$ on X, $\sup_{x\in K} \frac{|v_n^t(x) - v^t(x)|}{w(x)} \to 0$ as $n \to \infty$ for all t by Lemma 5.14. Hence, the last expression can be made arbitrarily small. This completes the proof. □

In the remainder of this section, we use the above results and Theorem 5.1 to compute a near optimal policy for the original MDP. It is straightforward to check that for each n, c-MDP$_n$ satisfies the assumptions in Theorem 5.1. Let $\{\varepsilon_n\}$ be a sequence of positive real numbers such that $\lim_{n\to\infty} \varepsilon_n = 0$.

By Theorem 5.1, for each $n \geq 1$, there exists a deterministic stationary policy $f_n \in \mathbb{F}_n$, obtained from the finite state approximations of c-MDP$_n$, such that

$$\sup_{x\in \mathsf{X}_n} |J_n(f_n, x) - J_n^*(x)| \leq \varepsilon_n,$$

where for each n, finite-state models are constructed replacing $\big(\mathsf{Z}, \mathsf{A}, p, c\big)$ with the components $\big(\mathsf{X}_n, \mathsf{A}, p_n, c_n\big)$ of c-MDP$_n$ in Section 5.2. By Lemma 5.12, for each $n \geq 1$

we also have

$$\sup_{x \in \mathsf{X}} |\bar{J}_n(f_n, x) - \bar{J}_n^*(x)| \leq \varepsilon_n, \tag{5.23}$$

where, with an abuse of notation, we also denote the extended (to X) policy by f_n. Let us define operators $\bar{R}_n : B_{C_n}(\mathsf{X}) \to B_{C_n}(\mathsf{X})$ and $R_n : B_w(\mathsf{X}) \to B_w(\mathsf{X})$ by

$$\bar{R}_n u(x) = \begin{cases} c(x, f_n(x)) + \beta \int_{\mathsf{X}} u(y) p(dy|x, f_n(x)), & \text{if } x \in K_n \\ \int_{K_n^c} \big[c(z, f_n(z)) + \beta \int_{\mathsf{X}} u(y) p(dy|z, f_n(z)) \big] \nu_n(dz), & \text{if } x \in K_n^c, \end{cases}$$
$$R_n u(x) = c(x, f_n(x)) + \beta \int_{\mathsf{X}} u(y) p(dy|x, f_n(x)).$$

By [55, Remark 8.3.10, p. 54], $\bar{R}_n$ is a contraction operator with modulus σ_0 and R_n is a contraction operator with modulus σ. Furthermore, the fixed point of $\bar{R}_n$ is $\bar{J}_n(f_n, x)$ and the fixed point of R_n is $J(f_n, x)$. For each $n \geq 1$, let us define $\bar{u}_n^0 = u_n^0 = 0$ and $\bar{u}_n^{t+1} = \bar{R}_n \bar{u}_n^t$, $u_n^{t+1} = R_n u_n^t$ $(t \geq 1)$. One can prove that (see the proof of [55, Theorem 8.3.6])

$$\bar{u}_n^t(x) \leq \bar{J}_n(f_n, x) \leq \frac{C_n(x)}{1 - \sigma_0}$$
$$\|\bar{u}_n^t - \bar{J}_n(f_n, \,\cdot\,)\|_{C_n} \leq \frac{\sigma_0^t}{1 - \sigma_0}$$
$$u_n^t(x) \leq J(f_n, x) \leq M \frac{w(x)}{1 - \sigma}$$
$$\|u_n^t - J(f_n, \,\cdot\,)\|_w \leq M \frac{\sigma^t}{1 - \sigma}.$$

Lemma 5.15. *For any compact set $K \subset \mathsf{X}$, we have*

$$\lim_{n\to\infty} \sup_{x\in K} |\bar{u}_n^t(x) - u_n^t(x)| = 0.$$

Proof. The lemma can be proved using the same arguments as in the proof of Lemma 5.14 and so we omit the details. □

Lemma 5.16. *For any compact set $K \subset \mathsf{X}$, we have*

$$\lim_{n\to\infty} \sup_{x\in K} |\bar{J}_n(f_n, x) - J(f_n, x)| = 0. \tag{5.24}$$

Indeed, this is true for all sequences of policies in $\mathbb{F}$.

Proof. The lemma can be proved using the same arguments as in the proof of Theorem 5.6. □

The following theorem is the main result of this section which states that the true cost functions of the policies obtained from finite state models converge to the value function of the original MDP. Hence, to obtain a near optimal policy for the original MDP, it is sufficient to compute the optimal policy for the finite state model that has sufficiently large number of grid points.

Theorem 5.7. *For any compact set $K \subset \mathsf{X}$, we have*

$$\lim_{n\to\infty} \sup_{x\in K} |J(f_n, x) - J^*(x)| = 0.$$

Therefore,

$$\lim_{n\to\infty} |J(f_n, x) - J^*(x)| = 0 \quad \textit{for all } x \in \mathsf{X}.$$

Proof. The result follows from (5.22), (5.23), and (5.24). □

5.3.2 Average Cost

In this section we obtain approximation results, analogous to Theorems 5.6 and 5.7, for the average cost criterion. To do this, we impose some new assumptions on the components of the original MDP in addition to Assumption 5.3. These assumptions are the unbounded counterpart of Assumption 5.2. With the exception of Assumption 5.4-(j), they are very similar to Assumption 4.4.

Assumption 5.4. Suppose Assumption 5.3 holds with item (b) and (5.5) replaced by conditions (j) and (e) below, respectively. In addition, there exist a probability measure η on X and a positive measurable function $\phi : \mathsf{X} \times \mathsf{A} \to (0, \infty)$ such that for all $(x, a) \in \mathsf{X} \times \mathsf{A}$

(e) $\int_{\mathsf{X}} w(y) p(dy|x, a) \leq \alpha w(x) + \eta(w)\phi(x, a)$, where $\alpha \in (0, 1)$.

(f) $p(D|x, a) \geq \eta(D)\phi(x, a)$ for all $D \in \mathcal{B}(\mathsf{X})$.

(g) The weight function w is η-integrable, i.e., $\eta(w) < \infty$.

(h) For each $n \geq 1$, $\inf_{(x,a) \in K_n \times \mathsf{A}} \phi(x, a) > 0$.

(j) The stochastic kernel $p(\cdot |x, a)$ is continuous in (x, a) with respect to the w-norm.

Analogous with Theorems 4.5 and 5.4, the following theorem is a consequence of [98, Theorems 3.3] and [44, Lemma 3.4 and Theorem 2.6] (see also [55, Proposition 10.2.5]), which also holds with Assumption 5.4-(j) replaced by Assumption 5.3-(b).

Theorem 5.8. *Under Assumption 5.4 the following hold.*

(i) For each $f \in \mathbb{F}$, the stochastic kernel $p(\,\cdot\,|x,f)$ is positive Harris recurrent with unique invariant probability measure μ_f. Furthermore, w is μ_f-integrable, and therefore, $\rho_f := \int_{\mathsf{X}} c(x,f)\mu_f(dx) < \infty$.

(ii) There exist positive real numbers R and $\kappa < 1$ such that

$$\sup_{f \in \mathbb{F}} \|p^t(\,\cdot\,|x,f(x)) - \mu_f\|_w \leq R\kappa^t \tag{5.25}$$

for all $x \in \mathsf{X}$, where R and κ continuously depend on α, $\eta(w)$, and $\inf_{f \in \mathbb{F}} \eta(\phi(y,f(y)))$.

(iii) There exist $f^* \in \mathbb{F}$ and $h^* \in B_w(\mathsf{X})$ such that the triplet (h^*, f^*, ρ_{f^*}) satisfies the average cost optimality inequality (ACOI), and therefore,

$$\inf_{\pi \in \Pi} V(\pi,x) := V^*(x) = \rho_{f^*},$$

for all $x \in \mathsf{X}$.

Note that this theorem implies that for each $f \in \mathbb{F}$, the average cost is given by $V(f,x) = \int_{\mathsf{X}} c(y,f(y))\mu_f(dy)$ for all $x \in \mathsf{X}$ (instead of μ_f-a.e.).

Remark 5.4. We note that if we further assume the continuity of $\phi(x,\,\cdot\,)$ on A for each $x \in \mathsf{X}$ in Assumption 5.4, then by [98, Theorem 3.6] the ACOE holds in Theorem 5.8-(iii) instead of the ACOI.

Recall that V_n and $\bar{V}_n$ denote the average costs of c-MDP$_n$ and $\overline{\text{MDP}}_n$, respectively. The value functions for average cost are denoted analogously to the discounted cost case. Similar to Lemma 5.12, the following result states that MDP$_n$ and $\overline{\text{MDP}}_n$ are not too different for the average cost.

Lemma 5.17. *Suppose Theorem 5.8 holds for $\overline{MDP}_n$ and Theorem 5.4 holds for MDP_n. Then we have*

$$\bar{V}_n^*(x) = \begin{cases} V_n^*(x), & \text{if } x \in K_n \\ V_n^*(\Delta_n), & \text{if } x \in K_n^c. \end{cases} \tag{5.26}$$

Furthermore, if, for any deterministic stationary policy $f \in \mathbb{F}_n$, we define $\bar{f}(x) = f(x)$ on K_n and $\bar{f}(x) = f(\Delta_n)$ on K_n^c, then

$$\bar{V}_n(\bar{f}, x) = \begin{cases} V_n(f, x), & \text{if } x \in K_n \\ V_n(f, \Delta_n), & \text{if } x \in K_n^c. \end{cases} \tag{5.27}$$

In particular, if the deterministic stationary policy $f_n^* \in \mathbb{F}_n$ is optimal for MDP_n, then its extension $\bar{f}_n^*$ to X is also optimal for $\overline{MDP}_n$.

Proof. Let the triplet $(h_n^*, f_n^*, \rho_{f_n^*}^n)$ satisfy the ACOI for c-MDP$_n$, so that f_n^* is an optimal policy and $\rho_{f_n^*}^n$ is the average value function for c-MDP$_n$. It is straightforward to show that the triplet $(\bar{h}_n^*, \bar{f}_n^*, \rho_{f_n^*}^n)$ satisfies the ACOI for $\widetilde{\text{MDP}}_n$, where

$$\bar{h}_n^*(x) = \begin{cases} h_n^*(x), & \text{if } x \in K_n \\ h_n^*(\Delta_n), & \text{if } x \in K_n^c, \end{cases}$$

and

$$\bar{f}_n^*(x) = \begin{cases} f_n^*(x), & \text{if } x \in K_n \\ f_n^*(\Delta_n), & \text{if } x \in K_n^c. \end{cases}$$

By [44, Theorem 2.6 and Lemma 5.2], this implies that $\bar{f}_n^*$ is an optimal stationary policy for $\overline{\text{MDP}}_n$ with cost function $\rho^n_{f_n^*}$. This completes the proof of the first part.

For the second part, let $f \in \mathbb{F}_n$ with an unique invariant probability measure $\mu_f \in \mathcal{P}(\mathsf{X}_n)$ and let $\bar{f} \in \mathbb{F}$ denote its extension to X with an unique invariant probability measure $\mu_{\bar{f}}$. It can be proved that

$$\mu_f(\,\cdot\,) = \mu_{\bar{f}}(\,\cdot\,\cap K_n) + \mu_{\bar{f}}(K_n^c)\delta_{\Delta_n}(\,\cdot\,).$$

Then we have

$$\begin{aligned}
\bar{V}_n(f,x) &= \int_{\mathsf{X}} b_n(x,\bar{f}(x))\mu_{\bar{f}}(dx) \\
&= \int_{K_n} c_n(x,\bar{f}(x))\mu_{\bar{f}}(dx) + \mu_{\bar{f}}(K_n^c)c_n(\Delta_n,\bar{f}(\Delta_n)) \\
&= \int_{\mathsf{X}_n} c_n(x,f(x))\mu_f(dx) \\
&= V_n(f,x).
\end{aligned}$$

This completes the proof. □

By Lemma 5.17, in the remainder of this section we need only consider $\overline{\text{MDP}}_n$ in place of MDP_n. Later we will show that Theorem 5.8 holds for $\overline{\text{MDP}}_n$ for n sufficiently large and that Theorem 5.4 holds for c-MDP_n for all n.

Recall the definition of constants γ_n and τ_n from (5.6) and (5.7). For each $n \geq 1$, we define $\phi_n : \mathsf{X} \times \mathsf{A} \to (0,\infty)$ and $\varsigma_n \in \mathbb{R}$ as

$$\phi_n(x,a) \coloneqq \begin{cases} \phi(x,a), & \text{if } x \in K_n \\ \int_{K_n^c} \phi(y,a)\nu_n(dy), & \text{if } x \in K_n^c, \end{cases}$$

$$\varsigma_n := \int_{K_n^c} w(y)\eta(dy).$$

Since $\eta(w) < \infty$ and τ_n can be made arbitrarily small by properly choosing ν_n, we assume, without loss of generality, the following.

Assumption 5.5. The sequence of probability measures $\{\nu_n\}$ is chosen such that the following holds

$$\lim_{n\to\infty}(\tau_n + \varsigma_n) = 0. \tag{5.28}$$

Let $\alpha_n := \alpha + \varsigma_n + \tau_n$.

Lemma 5.18. *For all n and $(x,a) \in \mathsf{X} \times \mathsf{A}$, the components of $\overline{MDP}_n$ satisfy the following:*

$$\begin{aligned} \sup_{a\in\mathsf{A}} b_n(x,a) &\leq M w_n(x) \\ \int_{\mathsf{X}} w_n(y) q_n(dy|x,a) &\leq \alpha_n w_n(x) + \eta(w_n)\phi_n(x,a), \\ q_n(D|x,a) &\geq \eta(D)\phi_n(x,a) \quad \text{for all } D \in \mathcal{B}(\mathsf{X}). \end{aligned} \tag{5.29}$$

Proof. See Section 5.8.3. □

We note that by (5.28), there exists $n_0 \geq 1$ such that $\alpha_n < 1$ for $n \geq n_0$. Hence, for each $n \geq n_0$, Theorem 5.8 holds for $\overline{\text{MDP}}_n$ with w replaced by w_n for some $R_n > 0$, and $\kappa_n \in (0,1)$, and we have $R_{\max} := \sup_{n\geq n_0} R_n < \infty$ and $\kappa_{\max} := \sup_{n\geq n_0} \kappa_n < 1$.

In the remainder of this section, it is assumed that $n \geq n_0$.

Lemma 5.19. *Let $g : \mathsf{X}\times\mathsf{A} \to \mathbb{R}$ be any measurable function such that* $\sup_{a\in\mathsf{A}} |g(x,a)| \leq$

$M_g w(x)$ for some $M_g \in \mathbb{R}$. Then, for all $t \geq 1$ and any compact set $K \subset \mathsf{X}$ we have

$$\sup_{(y,f)\in K\times\mathbb{F}} \left| \int_{\mathsf{X}} g_n(x, f(x)) q_n^t(dx|y, f(y)) - \int_{\mathsf{X}} g(x, f(x)) p^t(dx|y, f(y)) \right| \to 0$$

as $n \to \infty$, where $g_n(x,a) = g(x,a)$ on $K_n \times \mathsf{A}$ and $g_n(x,a) = \int_{K_n^c} g(z,a)\nu_n(dz)$ on $K_n^c \times \mathsf{A}$.

Proof. See Section 5.8.4. □

In the remainder of this section the above results are used to compute a near optimal policy for the original MDP. Let $\{\varepsilon_n\}$ be a sequence of positive real numbers converging to zero.

For each $f \in \mathbb{F}$, let μ_f^n denote the unique invariant probability measure of the transition kernel $q_n(\,\cdot\,|x, f(x))$ and let ρ_f^n denote the associated average cost; that is, $\rho_f^n \coloneqq \bar{V}_n(f, x) = \int_{\mathsf{X}} b_n(y, f(y))\mu_f^n(dy)$ for all initial points $x \in \mathsf{X}$. Therefore, the value function of $\overline{\text{MDP}}_n$, denoted by $\bar{V}_n^*$, is given by $V_n^*(x) = \inf_{f\in\mathbb{F}} \rho_f^n$, i.e., it is constant on X.

Before making the connection with Theorem 5.5, we prove the following result.

Lemma 5.20. *The transition probability p_n of c-MDP$_n$ is continuous in (x, a) with respect to the total variation distance.*

Proof. See Section 5.8.5. □

Thus we obtain that for each $n \geq 1$, c-MDP$_n$ satisfies the assumption in Theorem 5.5 for

$$\zeta(\,\cdot\,) = \eta(\,\cdot\, \cap K_n) + \eta(K_n^c)\delta_{\Delta_n}(\,\cdot\,),$$

$$\theta(x,a) = \begin{cases} \phi(x,a), & \text{if } x \in K_n \\ \int_{K_n^c} \phi(y,a)\nu_n(dy), & \text{if } x = \Delta_n, \end{cases}$$

and some $\lambda \in (0,1)$, where the existence of λ follows from Assumption 5.4-(h) and the fact that $\phi > 0$.

Consequently, there exists a deterministic stationary policy $f_n \in \mathbb{F}_n$, obtained from the finite state approximations of c-MDP$_n$, such that

$$\sup_{x \in \mathsf{X}_n} |V_n(f_n, x) - V_n^*(x)| \leq \varepsilon_n, \tag{5.30}$$

where finite-state models are constructed replacing $\big(\mathsf{Z}, \mathsf{A}, p, c\big)$ with the components $\big(\mathsf{X}_n, \mathsf{A}, p_n, c_n\big)$ of c-MDP$_n$ in Section 5.2. By Lemma 5.17, we also have

$$|\rho_{f_n}^n - \bar{V}_n^*| \leq \varepsilon_n, \tag{5.31}$$

where, by an abuse of notation, we also denote the policy extended to X by f_n.

Lemma 5.21. *We have*

$$\sup_{f \in \mathbb{F}} |\rho_f^n - \rho_f| \to 0 \tag{5.32}$$

as $n \to \infty$*.*

Proof. Fix any compact set $K \subset \mathsf{X}$. For any $t \geq 1$ and $y \in K$, we have

$$\begin{aligned}
\sup_{f \in \mathbb{F}} |\rho_f^n - \rho_f| &= \sup_{f \in \mathbb{F}} \biggl| \int_{\mathsf{X}} b_n(x, f(x)) \mu_f^n(dx) - \int_{\mathsf{X}} c(x, f(x)) \mu_f(dx) \biggr| \\
&\leq \sup_{f \in \mathbb{F}} \biggl| \int_{\mathsf{X}} b_n(x, f(x)) \mu_f^n(dx) - \int_{\mathsf{X}} b_n(x, f(x)) q_n^t(dx|y, f(y)) \biggr| \\
&\quad + \sup_{f \in \mathbb{F}} \biggl| \int_{\mathsf{X}} b_n(x, f(x)) q_n^t(dx|y, f(y)) - \int_{\mathsf{X}} c(x, f(x)) p^t(dx|y, f(y)) \biggr| \\
&\quad + \sup_{f \in \mathbb{F}} \biggl| \int_{\mathsf{X}} c(x, f(x)) p^t(dx|y, f(y)) - \int_{\mathsf{X}} c(x, f(x)) \mu_f(dx) \biggr| \\
&\leq M R_{\max} \kappa_{\max}^t + M R \kappa^t \\
&\quad + \sup_{(y,f) \in K \times \mathbb{F}} \biggl| \int_{\mathsf{X}} b_n(x, f(x)) q_n^t(dx|y, f(y)) - \int_{\mathsf{X}} c(x, f(x)) p^t(dx|y, f(y)) \biggr|,
\end{aligned}$$

where the last inequality follows from Theorem 5.8-(ii) and (5.4) in Assumption 5.3. The result follows from Lemma 5.19. □

Theorem 5.9. *The value function of $\overline{MDP}_n$ converges to the value function of the original MDP, i.e.,*

$$|\bar{V}_n^* - V^*| \to 0$$

as $n \to \infty$.

Proof. Since

$$\begin{aligned}
|\bar{V}_n^* - V^*| &= |\inf_{f \in \mathbb{F}} \rho_f^n - \inf_{f \in \mathbb{F}} \rho_f| \\
&\leq \sup_{f \in \mathbb{F}} |\rho_f^n - \rho_f|,
\end{aligned}$$

the result follows from Lemma 5.21 □

The following is the main result of this section which states that the true average cost of the policies f_n obtained from finite state approximations of c-MDP$_n$ converges to the average value function V^* of the original MDP.

Theorem 5.10. *We have*

$$|\rho_{f_n} - V^*| \to 0$$

as $n \to \infty$.

Proof. We have

$$\begin{aligned} |\rho_{f_n} - V^*| &\leq |\rho_{f_n} - \rho_{f_n}^n| + |\rho_{f_n}^n - \bar{V}_n^*| + |\bar{V}_n^* - V^*| \\ &\leq \sup_{f \in \mathbb{F}} |\rho_f - \rho_f^n| + \varepsilon_n + |\bar{V}_n^* - V^*| \quad \text{(by (5.31))} \end{aligned}$$

The result follows from Lemma 5.21 and Theorem 5.9. □

5.4 Discretization of the Action Space

For computing near optimal policies using well known algorithms, such as value iteration, policy iteration, and Q-learning, the action space must be finite. In this section, using results from Chapter 4 we show that, as a pre-processing step, the action space can taken to be finite if it has sufficiently large number of points for accurate approximation.

It was shown in Theorems 4.1 and 4.2 that any MDP with (infinite) compact action space and with bounded one-stage cost function can be well approximated by an MDP with finite action space under assumptions that are satisfied by c-MDP$_n$ for each n, for both the discounted cost and the average cost cases. Recall the sequence of finite subsets $\{\Lambda_k\}$ of A from Section 4.3. We define c-MDP$_{n,k}$ as the Markov decision process having the components $\bigl\{\mathsf{X}_n, \Lambda_k, p_n, c_n\bigr\}$ and we let $\mathbb{F}_n(\Lambda_k)$ denote the set of all deterministic stationary policies for c-MDP$_{n,k}$. Note that $\mathbb{F}_n(\Lambda_k)$ is the set of policies in $\mathbb{F}_n$ taking values only in Λ_k. Therefore, in a sense, c-MDP$_{n,k}$ and c-MDP$_n$ can be viewed as the same MDP, where the former has constraints on the set of policies. For each n and k, by an abuse of notation, let f_n^* and $f_{n,k}^*$ denote the optimal stationary policies of c-MDP$_n$ and c-MDP$_{n,k}$, respectively, for both the discounted and average costs. Then Theorems 4.1 and 4.2 show that for all n, we have

$$\lim_{k\to\infty} J_n(f_{n,k}^*, x) = J_n(f_n^*, x) \coloneqq J_n^*(x)$$
$$\lim_{k\to\infty} V_n(f_{n,k}^*, x) = V_n(f_n^*, x), \coloneqq V_n^*(x)$$

for all $x \in \mathsf{X}_n$. In other words, the discounted and average value functions of c-MDP$_{n,k}$ converge to the discounted and average value functions of c-MDP$_n$ as $k \to \infty$.

Let us fix $x \in \mathsf{X}$. For n sufficiently large (so $x \in K_n$), we choose k_n such that $|J_n(f_{n,k_n}^*, x) - J_n(f_n^*, x)| < 1/n$ (or $|V_n(f_{n,k_n}^*, x) - V_n(f_n^*, x)| < 1/n$ for the average cost). We note that if A is a compact subset of a finite dimensional Euclidean space, then by using Theorems 4.9 and 4.10 one can obtain an explicit expression for k_n in terms of n under further continuity conditions on c and p. By Lemmas 5.16 and 5.21, we have $|\bar{J}_n(f_{n,k_n}^*, x) - J(f_{n,k_n}^*, x)| \to 0$ and $|\bar{V}_n(f_{n,k_n}^*, x) - V(f_{n,k_n}^*, x)| \to 0$ as $n \to \infty$, where again by an abuse of notation, the policies extended to X are also denoted by f_{n,k_n}^*. Since $\bar{J}_n(f_{n,k_n}^*, x) = J_n(f_{n,k_n}^*, x)$ and $\bar{V}_n(f_{n,k_n}^*, x) = V_n(f_{n,k_n}^*, x)$, using Theorems 5.6 and 5.9 one can immediately obtain

$$\lim_{n\to\infty} J(f^*_{n,k_n}, x) = J^*(x)$$
$$\lim_{n\to\infty} V(f^*_{n,k_n}, x) = V^*(x).$$

Therefore, before discretizing the state space to compute the near optimal policies, one can discretize, without loss of generality, the action space A in advance on a finite grid using sufficiently large number of grid points.

5.5 Rates of Convergence for Compact-State MDPs

In this section we consider **(Q2)** for MDPs with compact state space; that is, we derive an upper bound on the performance loss due to discretization in terms of the cardinality of the set Z_n (i.e., number of grid points) . To do this, we will impose some new assumptions on the components of the MDP in addition to Assumptions 5.1 and 5.2. First, we present some definitions that are needed in the development.

For each $g \in C_b(\mathsf{Z})$, let

$$\|g\|_{\mathrm{Lip}} \coloneqq \sup_{(z,y)\in \mathsf{Z}\times\mathsf{Z}} \frac{|g(z)-g(y)|}{d_{\mathsf{Z}}(z,y)}.$$

If $\|g\|_{\mathrm{Lip}}$ is finite, then g is called Lipschitz continuous with Lipschitz constant $\|g\|_{\mathrm{Lip}}$. $\mathrm{Lip}(\mathsf{Z})$ denotes the set of all Lipschitz continuous functions on Z, i.e.,

$$\mathrm{Lip}(\mathsf{Z}) \coloneqq \{g \in C_b(\mathsf{Z}) : \|g\|_{\mathrm{Lip}} < \infty\}$$

and $\mathrm{Lip}(\mathsf{Z}, K)$ denotes the set of all $g \in \mathrm{Lip}(\mathsf{Z})$ with $\|g\|_{\mathrm{Lip}} \leq K$. The *Wasserstein* *distance of order* 1 [99, p. 95] between two probability measures ζ and ξ over Z is defined as

$$W_1(\zeta,\xi) \coloneqq \sup\biggl\{\biggl|\int_{\mathsf{Z}} g d\zeta - \int_{\mathsf{Z}} g d\xi\biggr| : g \in \mathrm{Lip}(\mathsf{Z},1)\biggr\}.$$

W_1 is also called the *Kantorovich-Rubinstein distance.* It is known that if Z is compact, then $W_1(\zeta,\xi) \leq \mathrm{diam}(\mathsf{Z})\|\zeta-\xi\|_{TV}$ [99, Theorem 6.13]. For compact Z, the Wasserstein distance of order 1 is weaker than total variation distance. Furthermore, for compact Z, the Wasserstein distance of order 1 metrizes the weak topology on the set of probability measures $\mathcal{P}(\mathsf{Z})$ [99, Corollary 6.11] which also implies that convergence in this sense is weaker than setwise convergence.

In this section we impose the following supplementary assumptions in addition to Assumption 5.1 and Assumption 5.2.

Assumption 5.6.

(g) The one-stage cost function c satisfies $c(\,\cdot\,,a) \in \mathrm{Lip}(\mathsf{Z},K_1)$ for all $a \in \mathsf{A}$ for some K_1.

(h) The stochastic kernel p satisfies $W_1\bigl(p(\,\cdot\,|z,a),p(\,\cdot\,|y,a)\bigr) \leq K_2 d_{\mathsf{Z}}(z,y)$ for all $a \in \mathsf{A}$ for some K_2.

(h') The stochastic kernel p satisfies: $\|p(\,\cdot\,|z,a)-p(\,\cdot\,|y,a)\|_{TV} \leq K_2 d_{\mathsf{Z}}(z,y)$ for all $a \in \mathsf{A}$ and for some K_2.

(j) Z is an infinite compact subset of $\mathbb{R}^d$ for some $d \geq 1$, equipped with the Euclidean norm.

We note that Assumption 5.6-(j) implies the existence of a constant $\alpha > 0$ and finite subsets $\mathsf{Z}_n \subset \mathsf{Z}$ with cardinality n such that

$$\max_{z \in \mathsf{Z}} \min_{y \in \mathsf{Z}_n} d_{\mathsf{Z}}(z, y) \leq \alpha (1/n)^{1/d} \tag{5.33}$$

for all n, where d_{Z} is the Euclidean distance on Z. In the remainder of this section, we replace Z_n defined in Section 5.2 with Z_n satisfying (5.33) in order to derive *explicit* bounds on the approximation error in terms of the cardinality of Z_n.

5.5.1 Discounted Cost

Assumptions 5.1 and 5.6 (without Assumption 5.6-(h')) are imposed throughout this section. Additionally, we assume that $K_2\beta < 1$. The last assumption is the key to prove the next result which states that the value function J^* of the original MDP for the discounted cost is in $\mathrm{Lip}(\mathsf{Z})$. Although this result is known in the literature [59], we give a short proof for the sake of completeness using a simple application of the value iteration algorithm.

Theorem 5.11. *Suppose Assumptions 5.1, 5.6 (without Assumption 5.6-(h')) and $K_2\beta < 1$ hold. Then the value function J^* for the discounted cost is in $\mathrm{Lip}(\mathsf{Z}, K)$, where $K = K_1 \frac{1}{1 - \beta K_2}$.*

Proof. Let $u \in \mathrm{Lip}(\mathsf{Z}, K)$ for some $K > 1$. Then $g = \frac{u}{K} \in \mathrm{Lip}(\mathsf{Z}, 1)$ and therefore, for all $a \in \mathsf{A}$ and $z, y \in \mathsf{Z}$ we have

$$\begin{aligned}\left| \int_{\mathsf{Z}} u(x) p(dx|z,a) - \int_{\mathsf{Z}} u(x) p(dx|y,a) \right| &= K \left| \int_{\mathsf{Z}} g(x) p(dx|z,a) - \int_{\mathsf{Z}} g(x) p(dx|y,a) \right| \\ &\leq K W_1\big(p(\,\cdot\,|z,a), p(\,\cdot\,|y,a)\big) \leq K K_2 d_{\mathsf{Z}}(z,y),\end{aligned}$$

by Assumption 5.6-(h). Hence, the Bellman optimality operator T of the MDP maps $u \in \mathrm{Lip}(\mathsf{Z}, K)$ to $Tu \in \mathrm{Lip}(\mathsf{Z}, K_1 + \beta K K_2)$, since, for all $z, y \in \mathsf{Z}$

$$\begin{aligned}
&|Tu(z) - Tu(y)| \\
&\qquad \leq \max_{a \in \mathsf{A}} \biggl\{ |c(z,a) - c(y,a)| + \beta \biggl| \int_{\mathsf{Z}} u(x) p(dx|z,a) - \int_{\mathsf{Z}} u(x) p(dx|y,a) \biggr| \biggr\} \\
&\qquad \leq K_1 d_{\mathsf{Z}}(z,y) + \beta K K_2 d_{\mathsf{Z}}(z,y) = \bigl(K_1 + \beta K K_2\bigr) d_{\mathsf{Z}}(z,y).
\end{aligned}$$

Now we apply T recursively to obtain the sequence $\{T^n u\}$ by letting $T^n u = T(T^{n-1}u)$, which converges to the value function J^* by the Banach fixed point theorem. Clearly, by induction we have for all $n \geq 1$

$$T^n u \in \mathrm{Lip}(\mathsf{Z}, K_n),$$

where $K_n = K_1 \sum_{i=0}^{n-1} (\beta K_2)^i + K (\beta K_2)^n$. If we choose $K < K_1$, then $K_n \leq K_{n+1}$ for all n and therefore, $K_n \uparrow K_1 \frac{1}{1-\beta K_2}$ since $K_2 \beta < 1$. Hence, $T^n u \in \mathrm{Lip}(\mathsf{Z}, K_1 \frac{1}{1-\beta K_2})$ for all n, and therefore, $J^* \in \mathrm{Lip}(\mathsf{Z}, K_1 \frac{1}{1-\beta K_2})$ since $\mathrm{Lip}(\mathsf{Z}, K_1 \frac{1}{1-\beta K_2})$ is closed with respect to the sup-norm $\| \cdot \|$. □

The following theorem is the main result of this section. Recall that the policy $\hat{f}_n \in \mathbb{F}$ is obtained by extending the optimal policy f_n^* of MDP_n to Z.

Theorem 5.12. *We have*

$$\|J(\hat{f}_n, \cdot) - J^*\| \leq \frac{\tau(\beta, K_2) K_1 \frac{1}{1-\beta K_2} + \frac{2K_1}{1-\beta}}{1-\beta} 2\alpha (1/n)^{1/d},$$

where $\tau(\beta, K_2) = (2+\beta)\beta K_2 + \frac{\beta^2 + 4\beta + 2}{(1-\beta)^2}$ *and* α *is the coefficient in (5.33).*

Proof. To prove the theorem, we obtain upper bounds on the expressions derived in Section 5.2.1 in terms of the cardinality n of Z_n. The proof of Theorem 5.1 gives

$$\|J(\hat{f}_n, \cdot) - J^*\| \leq \frac{\|T_{\hat{f}_n} J^* - \hat{T}_{\hat{f}_n} J^*\| + (1+\beta)\|\hat{J}_n^* - J^*\|}{1-\beta}.$$

To prove the theorem we upper bound $\|T_{\hat{f}_n} J^* - \hat{T}_{\hat{f}_n} J^*\|$ and $\|\hat{J}_n^* - J^*\|$ in terms n. For the first term we have

$$\begin{aligned}
&\|T_{\hat{f}_n} J^* - \hat{T}_{\hat{f}_n} J^*\| = \sup_{z \in \mathsf{Z}} \big| T_{\hat{f}_n} J^*(z) - \hat{T}_{\hat{f}_n} J^*(z) \big| \\
&\leq \sup_{z \in \mathsf{Z}} \int \bigg| c(z, \hat{f}_n(z)) + \beta \int_{\mathsf{Z}} J^*(y) p(dy|z, \hat{f}_n(z)) - c(x, \hat{f}_n(x)) \\
&\qquad\qquad - \beta \int_{\mathsf{Z}} J^*(y) p(dy|x, \hat{f}_n(x)) \bigg| \nu_{n, i_n(z)}(dx) \\
&\leq \sup_{z \in \mathsf{Z}} \int \bigg(K_1 d_{\mathsf{Z}}(x,z) + \beta \bigg| \int_{\mathsf{Z}} J^*(y) p(dy|z, \hat{f}_n(z)) - \int_{\mathsf{Z}} J^*(y) p(dy|x, \hat{f}_n(z)) \bigg| \bigg) \nu_{n, i_n(z)}(dx) \\
&\qquad\qquad (\text{since } \hat{f}_n(x) = \hat{f}_n(z) \text{ for all } x \in \mathcal{S}_{n, i_n(z)}) \\
&\leq \sup_{z \in \mathsf{Z}} \int (K_1 + \beta \|J^*\|_{\mathrm{Lip}} K_2) d_{\mathsf{Z}}(x,z) \nu_{n, i_n(z)}(dx) \\
&\leq (K_1 + \beta \|J^*\|_{\mathrm{Lip}} K_2) \max_{i \in \{1, \ldots, n\}} \operatorname{diam}(\mathcal{S}_{n,i}) \\
&\leq (K_1 + \beta \|J^*\|_{\mathrm{Lip}} K_2) 2\alpha (1/n)^{1/d}.
\end{aligned} \tag{5.34}$$

For the second term, the proof of Theorem 5.3 gives

$$\|\hat{J}_n^* - J^*\| \leq \frac{\|\hat{T}_n J^* - F_n J^*\| + (1+\beta)\|J^* - u_n^*\|}{1-\beta}.$$

First consider $\|\hat{T}_n J^* - F_n J^*\|$. Define

$$l(z,a) \coloneqq c(z,a) + \beta \int_{\mathsf{X}} J^*(y) p(dy|z,a),$$

so that

$$J^*(z) = \min_{a \in \mathsf{A}} l(z,a).$$

It is straightforward to show that $l(\,\cdot\,, a) \in \mathrm{Lip}(\mathsf{Z}, K_l)$ for all $a \in \mathsf{A}$, where $K_l = K_1 + \beta \|J^*\|_{\mathrm{Lip}} K_2$. By adapting the proof of Lemma 5.3 to the value function J^*, we obtain

$$\begin{aligned}
\|\hat{T}_n J^* - F_n J^*\| &= \sup_{z \in \mathsf{Z}} \left| \min_{a \in \mathsf{A}} \int l(x,a) \nu_{n,i_n(z)}(dx) - \int \min_{a \in \mathsf{A}} l(x,a) \nu_{n,i_n(z)}(dx) \right| \\
&\leq \sup_{z \in \mathsf{Z}} \int \sup_{y \in \mathcal{S}_{n,i_n(z)}} \big| l(y,a_i) - J^*(y) \big| \nu_{n,i_n(z)}(dy) \\
&\leq \max_{i \in \{1,\ldots,n\}} \int \sup_{y \in \mathcal{S}_{n,i}} \big\{ |l(y,a_i) - l(z_i,a_i)| + |J^*(z_i) - J^*(y)| \big\} \nu_{n,i}(dy) \\
&\leq \max_{i \in \{1,\ldots,n\}} \int \sup_{y \in \mathcal{S}_{n,i_n}} \big\{ K_l d_{\mathsf{Z}}(y,z_i) + \|J^*\|_{\mathrm{Lip}} d_{\mathsf{Z}}(z_i,y) \big\} \nu_{n,i}(dy) \\
&\leq (K_l + \|J^*\|_{\mathrm{Lip}}) \max_{i \in \{1,\ldots,n\}} \mathrm{diam}(\mathcal{S}_{n,i}) \\
&\leq (K_l + \|J^*\|_{\mathrm{Lip}}) 2\alpha (1/n)^{1/d}. \qquad (5.35)
\end{aligned}$$

For the expression $\|J^* - u_n^*\|$, by Lemma 5.2 we have

$$\|u_n^* - J^*\| \leq \frac{2}{1-\beta} \inf_{r \in \mathsf{Z}^{k_n}} \|J^* - \Phi_r\|,$$

where $\Phi_r(z) = \Sigma_{i=1}^{k_n} r_i 1_{S_{n,i}}(z)$, $r = (r_1, \ldots, r_{k_n})$. Since $\|J^*\|_{\mathrm{Lip}} < \infty$, we have

$\inf_{r\in\mathsf{Z}^{k_n}} \|J^* - \Phi_r\| \leq \|J^*\|_{\mathrm{Lip}} \max_{i\in\{1,\ldots,n\}} \operatorname{diam}(\mathcal{S}_{n,i}) \leq \|J^*\|_{\mathrm{Lip}} 2\alpha(1/n)^{1/d}$. Hence

$$\|u_n^* - J^*\| \leq \frac{2}{1-\beta}\|J^*\|_{\mathrm{Lip}} 2\alpha(1/n)^{1/d}. \tag{5.36}$$

Hence, by (5.35) and (5.36) we obtain

$$\|\hat{J}_n^* - J^*\| \leq \biggl(\bigl(\beta K_2 + \frac{\beta+3}{(1-\beta)^2}\bigr)\|J^*\|_{\mathrm{Lip}} + \frac{K_1}{1-\beta}\biggr) 2\alpha(1/n)^{1/d}. \tag{5.37}$$

Then, the result follows from (5.34) and (5.37), and the fact $\|J^*\|_{\mathrm{Lip}} \leq K_1 \frac{1}{1-\beta K_2}$. □

Remark 5.5. It is important to point out that if we replace Assumption 5.6-(h) with (h'), then Theorem 5.12 remains valid (with possibly different constants in front of the term $(1/n)^{1/d}$). However, in this case, we do not need the assumptions $K_2\beta < 1$.

5.5.2 Average Cost

In this section, we suppose that Assumptions 5.2 and 5.6 (without Assumption 5.6-(h)) hold. To obtain convergence rates for the average cost, we first prove a rate of convergence result for Lemma 5.6.

Lemma 5.22. *For all $t \geq 1$, we have*

$$\sup_{(y,f)\in\mathsf{Z}\times\mathbb{F}} \|p^t(\,\cdot\,|y,f(y)) - q_n^t(\,\cdot\,|y,f(y))\|_{TV} \leq K_2\alpha(1/n)^{1/d}\bigl(2^{t+1}-2\bigr),$$

where α is the coefficient in (5.33).

Proof. See Section 5.8.6. □

The following theorem is the main result of this section. Recall that the policy

$\tilde{f}_n^*$, the optimal policy of $\widetilde{\mathrm{MDP}}_n$, is obtained by extending the optimal policy f_n^* of MDP_n to Z.

Theorem 5.13. *For all $t \geq 1$, we have*

$$|\rho_{\tilde{f}_n^*} - \rho_{f^*}| \leq 4\|c\|R\kappa^t + 4K_1\alpha(1/n)^{1/d} + 2\|c\|K_2\alpha(1/n)^{1/d}\big(2^{t+1} - 2\big).$$

Proof. The proof of Theorem 5.5 gives

$$|\rho_{\tilde{f}_n^*} - \rho_{f^*}| \leq |\rho_{\tilde{f}_n^*} - \hat{\rho}^n_{\tilde{f}_n^*}| + |\hat{\rho}^n_{\tilde{f}_n^*} - \hat{\rho}^n_{\hat{f}_n^*}| + |\hat{\rho}^n_{\hat{f}_n^*} - \rho_{f^*}|.$$

Hence, to prove the theorem we obtain an upper bounds for $|\rho_{\tilde{f}_n^*} - \hat{\rho}^n_{\tilde{f}_n^*}|$, $|\hat{\rho}^n_{\tilde{f}_n^*} - \hat{\rho}^n_{\hat{f}_n^*}|$ and $|\hat{\rho}^n_{\hat{f}_n^*} - \rho_{f^*}|$ in terms of n. Consider the first term (recall the proof of Lemma 5.7)

$$\begin{aligned}
|\rho_{\tilde{f}_n^*} - \hat{\rho}^n_{\tilde{f}_n^*}| &\leq \sup_{f \in \mathbb{F}} |\hat{\rho}^n_f - \rho_f| \\
&\leq 2R\kappa^t\|c\| + \|c\| \sup_{(y,f) \in \mathsf{Z} \times \mathbb{F}} \|q_n^t(\,\cdot\,|y, f(y)) - p^t(\,\cdot\,|y, f(y))\|_{TV} \\
&\leq 2R\kappa^t\|c\| + \|c\|K_2\alpha(1/n)^{1/d} \sum_{i=1}^{t} 2^i \text{ (by Lemma 5.22).} \qquad (5.38)
\end{aligned}$$

For the second term, the proof of Lemma 5.11 gives

$$\begin{aligned}
|\hat{\rho}^n_{\tilde{f}_n^*} - \hat{\rho}^n_{\hat{f}_n^*}| &\leq |\hat{\rho}^n_{\tilde{f}_n^*} - \tilde{\rho}^n_{\tilde{f}_n^*}| + |\tilde{\rho}^n_{\tilde{f}_n^*} - \hat{\rho}^n_{\hat{f}_n^*}| \\
&\leq \sup_{f \in \mathbb{F}} |\hat{\rho}^n_f - \tilde{\rho}^n_f| + |\inf_{f \in \mathbb{F}} \tilde{\rho}^n_f - \inf_{f \in \mathbb{F}} \hat{\rho}^n_f| \\
&\leq 2 \sup_{f \in \mathbb{F}} |\hat{\rho}^n_f - \tilde{\rho}^n_f| \\
&\leq 2\|b_n - c\| \text{ (see the proof of Lemma 5.9)}
\end{aligned}$$

$$
\begin{aligned}
&\leq 2 \sup_{(z,a)\in \mathsf{Z}\times\mathsf{A}} \int |c(x,a) - c(z,a)| \nu_{n,i_n(z)}(dx) \\
&\leq 2 \sup_{z\in\mathsf{Z}} \int K_1 d_{\mathsf{Z}}(x,z) \nu_{n,i_n(z)}(dx) \\
&\leq 2K_1 \max_{i\in\{1,\ldots,n\}} \operatorname{diam}(\mathcal{S}_{n,i}) \\
&\leq 4K_1 \alpha (1/n)^{1/d}. && (5.39)
\end{aligned}
$$

For the last term, we have

$$
\begin{aligned}
|\hat{\rho}^n_{\hat{f}^*_n} - \rho_{f^*}| &= |\inf_{f\in\mathbb{F}} \hat{\rho}^n_f - \inf_{f\in\mathbb{F}} \rho_f| \\
&\leq \sup_{f\in\mathbb{F}} |\hat{\rho}^n_f - \rho_f| \\
&\leq 2R\kappa^t \|c\| + \|c\| K_2 \alpha (1/n)^{1/d} \sum_{i=1}^{t} 2^i \text{ (by (5.38))}. && (5.40)
\end{aligned}
$$

Hence, for any $t \geq 1$, by combining (5.38), (5.39), and (5.40) we obtain the result. □

To obtain a proper rate of convergence result (i.e., an upper bound that only depends on n) the t term in the upper bound in Theorem 5.13 has to be written as a function of n. This can be done by (approximately) minimizing the upper bound in Theorem 5.13 with respect to t for each n. Let us define the constants $I_1 \coloneqq 4\|c\|R$, $I_2 \coloneqq 4K_1\alpha$, and $I_3 \coloneqq 2\|c\|K_2\alpha$, and therefore, the upper bound in Theorem 5.13 becomes

$$
I_1\kappa^t + I_2(1/n)^{1/d} + I_3(1/n)^{1/d}\big(2^{t+1} - 2\big). \tag{5.41}
$$

For each n, it is straightforward to compute that the real number

$$t'(n) \coloneqq \ln\Big(\frac{n^{1/d}}{I_4}\Big)\frac{1}{\ln(\frac{2}{\kappa})}$$

makes the derivative of (5.41) zero, where $I_4 \coloneqq \big(\frac{I_1}{2I_3\ln(\frac{1}{\kappa})\ln(2)}\big)^{-1}$. Therefore, for n sufficiently large, (5.41) takes the minimum value at $t = t'(n)$.

Corollary 5.1. *We have*

$$|\rho_{\tilde{f}_n^*} - \rho_{f^*}| \leq \big(I_1 I_4^{\varrho_1} + 2I_3 I_4^{\varrho_1 - 1}\big)(1/n)^{\varrho_1/d} + \big(I_2 - 2I_3\big)(1/n)^{1/d},$$

where $\varrho_1 \coloneqq -\frac{\ln(\kappa)}{\ln(2/\kappa)}$.

5.5.3 Order Optimality

The following example demonstrates that the order of the performance losses in Theorems 5.12 and 5.13 cannot be better than $O((\frac{1}{n})^{\frac{1}{d}})$. More precisely, we exhibit a simple standard example where we can lower bound the performance loss by $L(1/n)^{1/d}$, for some positive constant L. A similar result was obtained in Section 4.5.3 for the case of quantization of action space. Therefore, when both state and action spaces are quantized, then the resulting construction is order optimal in the above sense as the approximation error, in this case, is bounded by the sum of the approximation errors in quantization of state space and quantization of action space.

In what follows $h(\,\cdot\,)$ and $h(\,\cdot\,|\,\cdot\,)$ denote differential and conditional differential entropies, respectively [25].

Consider the additive-noise system:

$$Z_{t+1} = F(Z_t, A_t) + V_t, t = 0, 1, 2, \ldots,$$

where $Z_t, A_t, V_t \in \mathbb{R}^d$. We assume that $\sup_{(z,a)\in\mathbb{R}^d\times\mathbb{R}^d} \frac{\|F(z,a)\|}{\|z\|+\|a\|} < 1/2$. The noise process $\{V_t\}$ is a sequence of i.i.d. random vectors whose common distribution has density g supported on some compact subset V of $\mathbb{R}^d$. We choose V such that $\mathsf{Z} = \mathsf{A}$ can be taken to be compact subsets of $\mathbb{R}^d$. For simplicity suppose that the initial distribution μ has the same density g. It is assumed that the differential entropy $h(g) \coloneqq -\int_{\mathsf{Z}} g(z) \log g(z) dz$ is finite. Let the one stage cost function be $c(z, a) \coloneqq \|z - a\|$. Clearly, the optimal stationary policy f^* is induced by the identity $f^*(z) = z$, having the optimal cost $J(f^*, \mu) = 0$ and $V(f^*, \mu) = 0$. Let $\hat{f}_n$ be the piece-wise constant extension of the optimal policy f_n^* of the MDP$_n$ to the set Z. Fix $n \geq 1$ and define $D_t \coloneqq \mathbb{E}_\mu^{\hat{f}_n}\big[c(Z_t, A_t)\big]$ for all t. Then, since $A_t = \hat{f}_n(Z_t)$ can take at most n values in A, by the Shannon lower bound (SLB) [105, p. 12] we have for $t \geq 1$

$$\begin{aligned}
\log n \geq R(D_t) &\geq h(Z_t) + \theta(D_t) \\
&= h(F(Z_{t-1}, A_{t-1}) + V_{t-1}) + \theta(D_t) \\
&\geq h(F(Z_{t-1}, A_{t-1}) + V_{t-1} | Z_{t-1}, A_{t-1}) + \theta(D_t) \\
&= h(V_{t-1}) + \theta(D_t), \qquad (5.42)
\end{aligned}$$

where $\theta(D_t) = -d + \log\bigg(\frac{1}{dV_d\Gamma(d)}\big(\frac{d}{D_t}\big)^d\bigg)$, $R(D_t)$ is the rate-distortion function of Z_t, V_d is the volume of the unit sphere $S_d = \{z : \|z\| \leq 1\}$, and Γ is the gamma function. Here, (5.42) follows from the independence of V_{t-1} and the pair (Z_{t-1}, A_{t-1}). Note that $h(V_{t-1}) = h(g)$ for all t. Thus, $D_t \geq L(1/n)^{1/d}$, where $L \coloneqq \frac{d}{2}\big(\frac{2^{h(g)}}{dV_d\Gamma(d)}\big)^{1/d}$.

Since we have obtained stage-wise error bounds, these give $|J(f^*,\mu) - J(\hat{f}_n,\mu)| \geq \frac{L}{1-\beta}(1/n)^{1/d}$ and $|V(f^*,\mu) - V(\hat{f}_n,\mu)| \geq L(1/n)^{1/d}$.

5.6 Numerical Examples

In this section, we consider two examples, the additive noise model and fisheries management problem, in order to illustrate our results numerically. Since computing true costs of the policies obtained from the finite models is intractable, we only compute the value functions of the finite models and illustrate their converge to the value function of the original MDP as $n \to \infty$.

Before proceeding to the examples, we note that all results in this chapter apply with straightforward modifications for the case of maximizing reward instead of minimizing cost.

5.6.1 Additive Noise System

In this example, the additive noise system is given by

$$X_{t+1} = F(X_t, A_t) + V_t, \; t = 0, 1, 2, \ldots$$

where $X_t, A_t, V_t \in \mathbb{R}$ and $\mathsf{X} = \mathbb{R}$. The noise process $\{V_t\}$ is a sequence of $\mathbb{R}$-valued i.i.d. random variables with common density g. Hence, the transition probability $p(\,\cdot\,|x,a)$ is given by

$$p(D|x,a) = \int_D g(v - F(x,a)) m(dv) \;\; \text{for all } D \in \mathcal{B}(\mathbb{R}),$$

where m is the Lebesgue measure. The one-stage cost function is $c(x,a) = (x-a)^2$, the action space is $\mathsf{A} = [-L, L]$ for some $L > 0$, and the cost function to be minimized is the discounted cost.

We assume that (i) g is a Gaussian probability density function with zero mean and variance σ^2, (ii) $\sup_{a \in \mathsf{A}} |F(x,a)|^2 \leq k_1 x^2 + k_2$ for some $k_1, k_2 \in \mathbb{R}_+$, (ii) $\beta < 1/\alpha$ for some $\alpha \geq k_1$, and (iv) F is continuous. Hence, Assumption 5.3 holds for this model with $w(x) = k + x^2$ and $M = 4\big(\frac{L^2}{k} + x^2\big)$, for some $k \in \mathbb{R}_+$.

For the numerical results, we use the following parameters: $F(x,a) = x + a$, $\beta = 0.3$, $L = 0.5$, and $\sigma = 0.1$.

We selected a sequence $\big\{[-l_n, l_n]\big\}_{n=1}^{15}$ of nested closed intervals, where $l_n = 0.5 + 0.25n$, to approximate $\mathbb{R}$. Each interval is uniformly discretized using $\lceil 2k_{\lceil \frac{n}{3} \rceil} l_n \rceil$ grid points, where $k_m = 5m$ for $m = 1, \ldots, 5$ and $\lceil q \rceil$ denotes the smallest integer greater than or equal to $q \in \mathbb{R}$. Therefore, the discretization is gradually refined. For each n, the finite state space is given by $\{x_{n,i}\}_{i=1}^{k_n} \cup \{\Delta_n\}$, where $\{x_{n,i}\}_{i=1}^{k_n}$ are the representation points in the uniform quantization of the closed interval $[-l_n, l_n]$ and Δ_n is a pseudo state. We also uniformly discretize the action space $\mathsf{A} = [-0.5, 0.5]$ by using $2k_{\lceil \frac{n}{3} \rceil}$ grid points. For each n, the finite state models are constructed as in Section 5.2 by replacing Z with $[-l_n, l_n]$ and by setting $\nu(\,\cdot\,) = \frac{1}{2} m_n(\,\cdot\,) + \frac{1}{2}\delta_{\Delta_n}(\,\cdot\,)$, where m_n is the Lebesgue measure normalized over $[-l_n, l_n]$.

We use the value iteration algorithm to compute the value functions of the finite models. Figure 5.1 displays the graph of these value functions corresponding to the different values for the number of grid points, when the initial state is $x = 0.7$. The figure illustrates that the value functions of the finite models converge to the value function of the original model.

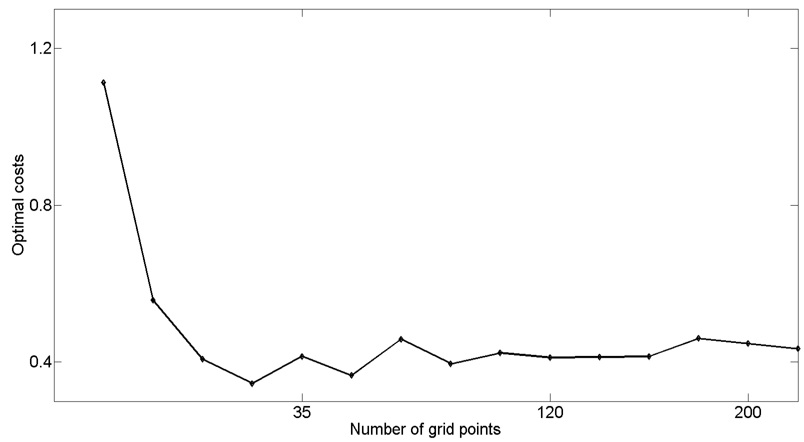

Figure 5.1: Optimal costs of the finite models when the initial state is $x = 0.7$

5.6.2 Fisheries Management Problem

In this example we consider the following population growth model, called a Ricker model, see [54, Section 1.3]:

$$X_{t+1} = \theta_1 A_t \exp\{-\theta_2 A_t + V_t\}, \ t = 0, 1, 2, \dots \tag{5.43}$$

where $\theta_1, \theta_2 \in \mathbb{R}_+$, X_t is the population size in season t, and A_t is the population to be left for spawning for the next season, or in other words, $X_t - A_t$ is the amount of fish captured in the season t. The one-stage 'reward' function is $u(x - a)$, where u is some utility function. In this model, the goal is to maximize the average reward.

The state and action spaces are $\mathsf{X} = \mathsf{A} = [\kappa_{\min}, \kappa_{\max}]$, for some $\kappa_{\min}, \kappa_{\max} \in \mathbb{R}_+$. Since the population left for spawning cannot be greater than the total population, for each $x \in \mathsf{X}$, the set of admissible actions is $\mathsf{A}(x) = [\kappa_{\min}, x]$ which is not consistent with our assumptions. However, we can (equivalently) reformulate above problem so that the admissible actions $\mathsf{A}(x)$ will become A for all $x \in \mathsf{X}$. In this case, instead of dynamics in equation (5.43) we have

$$X_{t+1} = \theta_1 \min(A_t, X_t) \exp\{-\theta_2 \min(A_t, X_t) + V_t\}, \ t = 0, 1, 2, \ldots$$

and $\mathsf{A}(x) = [\kappa_{\min}, \kappa_{\max}]$ for all $x \in \mathsf{X}$. The one-stage reward function is $u(x-a)1_{\{x \geq a\}}$.

Since X is already compact, it is sufficient to discretize $[\kappa_{\min}, \kappa_{\max}]$. The noise process $\{V_t\}$ is a sequence of independent and identically distributed (i.i.d.) random variables which have common density g supported on $[0, \lambda]$. Therefore, the transition probability $p(\,\cdot\,|x, a)$ is given by

$$\begin{aligned} p\big(D|x,a\big) &= \Pr\bigg\{ X_{t+1} \in D \bigg| X_t = x, A_t = a \bigg\} \\ &= \Pr\bigg\{ \theta_1 \min(a,x) \exp\{-\theta_2 \min(a,x) + v\} \in D \bigg\} \\ &= \int_D g\bigg(\log(v) - \log(\theta_1 \min(a,x)) + \theta_2 \min(a,x) \bigg) \frac{1}{v} m(dv), \end{aligned}$$

for all $D \in \mathcal{B}(\mathbb{R})$. To make the model consistent, we must have $\theta_1 y \exp\{-\theta_2 y + v\} \in [\kappa_{\min}, \kappa_{\max}]$ for all $(y, v) \in [\kappa_{\min}, \kappa_{\max}] \times [0, \lambda]$.

We assume that (i) $g > \epsilon$ for some $\epsilon \in \mathbb{R}_+$ on $[0, \lambda]$, (ii) g is continuous on $[0, \lambda]$, and (iii) the utility function u is continuous. Define $h(v, x, a) \coloneqq g\big(\log(v) - \log(\theta_1 \min(a, x)) + \theta_2 \min(a, x)\big)\frac{1}{v}$, and for each $(x, a) \in \mathsf{X} \times \mathsf{A}$, let $S_{x,a}$ denote the support of $h(\,\cdot\,, x, a)$. Then, Assumption 5.2 holds for this model with $\theta(x, a) = \inf_{v \in S_a} h(v, x, a)$ (provided that it is measurable), $\zeta = m_\kappa$ (Lebesgue measure restricted on $[\kappa_{\min}, \kappa_{\max}]$), and for some $\lambda \in (0, 1)$.

For the numerical results, we use the following values of the parameters:

$$\theta_1 = 1.1,\ \theta_2 = 0.1,\ \kappa_{\max} = 7,\ \kappa_{\min} = 0.005,\ \lambda = 0.5.$$

We assume that the noise process is distributed uniformly over $[0, 0.5]$. Hence, $g \equiv 1$ on $[0, 0.5]$ and otherwise zero. The utility function u is taken to be the shifted isoelastic utility function [33, Section 4.1]

$$u(z) = 3\big((z + 0.5)^{1/3} - (0.5)^{1/3}\big).$$

We selected 25 different values for the number n of grid points to discretize the state space: $n = 10, 20, 30, \ldots, 250$. The grid points are chosen uniformly over the interval $[\kappa_{\min}, \kappa_{\max}]$. We also uniformly discretize the action space A by using the following number of grid points: $5n = 50, 100, 150, \ldots, 1250$.

We use the relative value iteration algorithm [11, Chapter 4.3.1] to compute the value functions of the finite models. For each n, the finite state models are constructed as in Section 5.2 by replacing Z with $[\kappa_{\min}, \kappa_{\max}]$ and by setting $\nu(\,\cdot\,) = m_{\kappa}(\,\cdot\,)$.

Figure 5.2 shows the graph of the value functions of the finite models corresponding to the different values of n (number of grid points), when the initial state is $x = 2$. It can be seen that the value functions converge (to the value function of the original model).

5.7 Conclusion

In this chapter, the approximation of a discrete time MDP by finite-state MDPs was considered for discounted and average costs for both compact and non-compact state

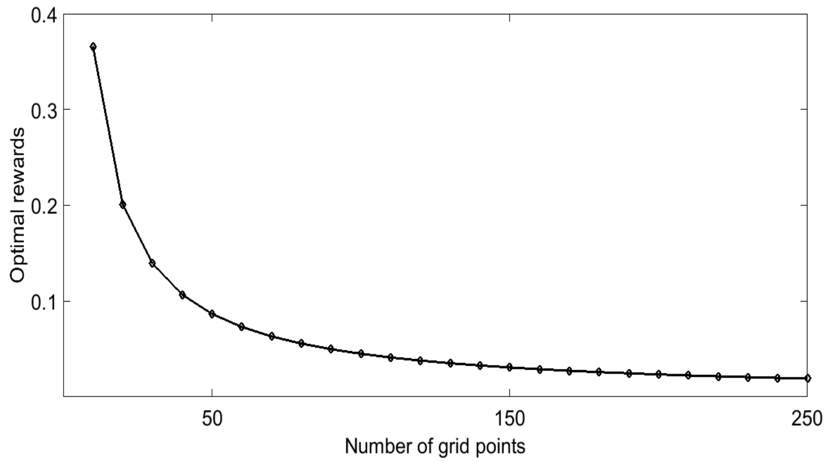

Figure 5.2: Optimal rewards of the finite models when the initial state is $x = 2$

spaces. Under usual conditions imposed for studying Markov decision processes, it was shown that if one uses a sufficiently large number of grid points to discretize the state space, then the resulting finite-state MDP yields a near optimal policy. Under the Lipschitz continuity of the transition probability and the one-stage cost function, explicit bounds were derived on the performance loss due to discretization in terms of the number of grid points for the compact state case. These results were then illustrated numerically by considering two different MDP models.

5.8 Proofs

5.8.1 Proof of Lemma 5.6

We will prove the lemma by induction. Note that if one views the stochastic kernel $p(\,\cdot\,|z,a)$ as a mapping from $\mathsf{Z} \times \mathsf{A}$ to $\mathcal{P}(\mathsf{Z})$, then Assumption 5.2-(f) implies that this mapping is continuous, and therefore uniformly continuous, when $\mathcal{P}(\mathsf{Z})$ is equipped with the metric induced by the total variation distance.

For $t = 1$ the claim holds by the following argument:

$$\sup_{(y,f)\in \mathsf{Z}\times\mathbb{F}} \big\|p(\,\cdot\,|y,f(y)) - q_n(\,\cdot\,|y,f(y))\big\|_{TV}$$

$$\begin{aligned}
&:= 2 \sup_{(y,f)\in \mathsf{Z}\times\mathbb{F}} \sup_{D\in\mathcal{B}(\mathsf{Z})} \big|p(D|y,f(y)) - q_n(D|y,f(y))\big| \\
&\leq 2 \sup_{(y,f)\in \mathsf{Z}\times\mathbb{F}} \sup_{D\in\mathcal{B}(\mathsf{Z})} \int \big|p(D|y,f(y)) - p(D|z,f(y))\big| \, \nu_{n,i_n(y)}(dz) \\
&\leq \sup_{(y,f)\in \mathsf{Z}\times\mathbb{F}} \int \big\|p(\,\cdot\,|y,f(y)) - p(\,\cdot\,|z,f(y))\big\|_{TV} \nu_{n,i_n(y)}(dz) \\
&\leq \sup_{y\in\mathsf{Z}} \sup_{(z,a)\in\mathcal{S}_{n,i_n(y)}\times\mathsf{A}} \big\|p(\,\cdot\,|y,a) - p(\,\cdot\,|z,a)\big\|_{TV}.
\end{aligned}$$

As the mapping $p(\,\cdot\,|z,a) : \mathsf{Z}\times\mathsf{A} \to \mathcal{P}(\mathsf{Z})$ is uniformly continuous with respect to the total variation distance and $\max_{n,i} \operatorname{diam}(\mathcal{S}_{n,i}) \to 0$ as $n \to \infty$, the result follows. Assume the claim is true for $t \geq 1$. Then we have

$$\begin{aligned}
&\sup_{(y,f)\in \mathsf{Z}\times\mathbb{F}} \big\|p^{t+1}(\,\cdot\,|y,f(y)) - q_n^{t+1}(\,\cdot\,|y,f(y))\big\|_{TV} \\
&:= 2 \sup_{(y,f)\in \mathsf{Z}\times\mathbb{F}} \sup_{D\in\mathcal{B}(\mathsf{Z})} \big|p^{t+1}(D|y,f(y)) - q_n^{t+1}(D|y,f(y))\big| \\
&\leq 2 \sup_{(y,f)\in \mathsf{Z}\times\mathbb{F}} \bigg(\sup_{D\in\mathcal{B}(\mathsf{Z})} \bigg| \int_{\mathsf{Z}} p(D|z,f(z)) p^t(dz|y,f(y)) - \int_{\mathsf{Z}} p(D|z,f(z)) q_n^t(dz|y,f(y)) \bigg| \\
&\qquad + 2 \sup_{D\in\mathcal{B}(\mathsf{Z})} \bigg| \int_{\mathsf{Z}} p(D|z,f(z)) q_n^t(dz|y,f(y)) - \int_{\mathsf{Z}} q_n(D|z,f(z)) q_n^t(dz|y,f(y)) \bigg| \bigg) \\
&\leq 2 \sup_{(y,f)\in \mathsf{Z}\times\mathbb{F}} \big\|p^t(\,\cdot\,|y,f(y)) - q_n^t(\,\cdot\,|y,f(y))\big\|_{TV} + \sup_{(z,f)\in \mathsf{Z}\times\mathbb{F}} \big\|p(\,\cdot\,|z,f(z)) - q_n(\,\cdot\,|z,f(z))\big\|_{TV}
\end{aligned}$$

where the last inequality follows from the following property of the total variation distance: for any $h \in \mathcal{B}(\mathsf{Z})$ and $\mu, \nu \in \mathcal{P}(\mathsf{Z})$ we have $\big|\int_{\mathsf{Z}} h(z)\mu(dz) - \int_{\mathsf{Z}} h(z)\nu(dz)\big| \leq \|h\| \|\mu - \nu\|_{TV}$. By the first step of the proof and the induction hypothesis, the last term converges to zero as $n \to \infty$. This completes the proof.

5.8.2 Proof Lemma 5.13

It is straightforward to prove (5.11) by using the definitions of b_n and w_n, and the equation (5.4). To prove (5.12), we have to consider two cases: $x \in K_n$ and $x \in K_n^c$. For the first case, $q_n(\,\cdot\,|x,a) = p(\,\cdot\,|x,a)$, and therefore, we have

$$\begin{aligned}
\sup_{a \in \mathsf{A}} \int_{\mathsf{X}} w_n(y) p(dy|x,a) &= \sup_{a \in \mathsf{A}} \biggl\{ \int_{\mathsf{X}} w(y) p(dy|x,a) + \int_{K_n^c} \bigl(\gamma_n - w(y)\bigr)\, p(dy|x,a) \biggr\} \\
&\leq \sup_{a \in \mathsf{A}} \int_{\mathsf{X}} w(y) p(dy|x,a) + \gamma \quad \text{(by (5.7))} \\
&\leq \alpha w(x) + \gamma = \alpha w_n(x) + \gamma \quad (\text{as } w_n = w \text{ on } K_n).
\end{aligned}$$

For $x \in K_n^c$, we have

$$\begin{aligned}
\sup_{a \in \mathsf{A}} \int_{\mathsf{X}} w_n(y) q_n(dy|x,a) &= \sup_{a \in \mathsf{A}} \int_{K_n^c} \biggl(\int_{\mathsf{X}} w_n(y) p(dy|z,a) \biggr) \nu_n(dz) \\
&\leq \int_{K_n^c} \biggl(\sup_{a \in \mathsf{A}} \int_{\mathsf{X}} w_n(y) p(dy|z,a) \biggr) \nu_n(dz) \\
&\leq \int_{K_n^c} \bigl(\alpha w(z) + \gamma\bigr)\, \nu_n(dz) \qquad (5.44) \\
&= \alpha w_n(x) + \gamma,
\end{aligned}$$

where (5.44) can be proved following the same arguments as for the case $x \in K_n$. This completes the proof.

5.8.3 Proof of Lemma 5.18

The proof of the first inequality follows from Assumption 5.4 and definitions of b_n and w_n. To prove the remaining two inequalities, we have to consider the cases $x \in K_n$ and $x \in K_n^c$ separately.

Let $x \in K_n$, and therefore, $q_n(\,\cdot\,|x,a) = p(\,\cdot\,|x,a)$. The second inequality holds since

$$\begin{aligned}
\int_{\mathsf{X}} w_n(y) p(dy|x,a) &= \int_{\mathsf{X}} w(y) p(dy|x,a) + \int_{K_n^c} \big(\gamma_n - w(y)\big)\, p(dy|x,a) \\
&\leq \int_{\mathsf{X}} w(y) p(dy|x,a) + \tau_n \\
&\leq \alpha w(x) + \eta(w)\phi(x,a) + \tau_n \\
&\leq \alpha w_n(x) + \eta(w_n)\phi_n(x,a) + \varsigma_n \phi_n(x,a) + \tau_n \ \ (\text{as } w_n = w \text{ and } \phi_n = \phi \text{ on } K_n) \\
&\leq \alpha_n w_n(x) + \eta(w_n)\phi_n(x,a), \ \ (\text{as } \phi_n \leq 1 \text{ and } w_n \geq 1).
\end{aligned}$$

For the last inequality, for all $D \in \mathcal{B}(\mathsf{X})$, we have

$$q_n(D|x,a) = p(D|x,a) \geq \eta(D)\phi(x,a) = \eta(D)\phi_n(x,a) \ \ (\text{as } \phi_n = \phi \text{ on } K_n).$$

Hence, inequalities hold for $x \in K_n$.

For $x \in K_n^c$, we have

$$\begin{aligned}
\int_{\mathsf{X}} w_n(y) q_n(dy|x,a) &= \int_{K_n^c} \bigg(\int_{\mathsf{X}} w_n(y) p(dy|z,a) \bigg) \nu_n(dz) \\
&\leq \int_{K_n^c} \big(\alpha w(z) + \eta(w_n)\phi(x,a) + \varsigma_n \phi(x,a) + \tau_n \big)\, \nu_n(dz) \qquad (5.45) \\
&= \alpha w_n(x) + \eta(w_n)\phi_n(x,a) + \varsigma_n \phi_n(x,a) + \tau_n \\
&\leq \alpha_n w_n(x) + \eta(w_n)\phi_n(x,a), \ \ (\text{since } \phi_n \leq 1 \text{ and } w_n \geq 1)
\end{aligned}$$

where (5.45) can be obtained following the same arguments as for the case $x \in K_n$.

The last inequality holds for $x \in K_n^c$ since

$$\begin{aligned} q_n(D|x,a) &= \int_{K_n^c} p(D|z,a)\nu_n(dz) \\ &\geq \int_{K_n^c} \eta(D)\phi(z,a)\nu_n(dz) \\ &= \eta(D)\phi_n(x,a). \end{aligned}$$

This completes the proof.

5.8.4 Proof of Lemma 5.19

We will prove the lemma by induction. Fix any compact set $K \subset \mathsf{X}$. We note that in the inequalities below, we repeatedly use the fact $\phi, \phi_n \leq 1$ without explicitly referring to this fact. Recall the definition of the compact subsets K_ε of X in Lemma 4.3 and the constant $\gamma_{\max} = \max\{1, \gamma\}$. Note that $\sup_{a \in \mathsf{A}} |g_n(x,a)| \leq M_g w_n(x) \leq M_g \gamma_{\max} w(x)$ for all $x \in \mathsf{X}$.

The claim holds for $t = 1$ by the following argument:

$$\begin{aligned} &\sup_{(y,f) \in K \times \mathbb{F}} \left| \int_{\mathsf{X}} g_n(x, f(x)) q_n(dx|y, f(y)) - \int_{\mathsf{X}} g(x, f(x)) p(dx|y, f(y)) \right| \\ &\quad = \sup_{(y,f) \in K \times \mathbb{F}} \left| \int_{\mathsf{X}} g_n(x, f(x)) p(dx|y, f(y)) - \int_{\mathsf{X}} g(x, f(x)) p(dx|y, f(y)) \right| \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(for } n \text{ sufficiently large)} \\ &\quad = \sup_{(y,f) \in K \times \mathbb{F}} \left| \int_{K_\varepsilon^c} g_n(x, f(x)) p(dx|y, f(y)) - \int_{K_\varepsilon^c} g(x, f(x)) p(dx|y, f(y)) \right| \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(for } n \text{ sufficiently large)} \\ &\quad \leq M_g (1 + \gamma_{\max}) \varepsilon, \end{aligned}$$

where the last inequality follows from Lemma 4.3. Since ε is arbitrary, the result follows.

Assume the claim is true for $t \geq 1$. Let us define $l_f(z) := \int_{\mathsf{X}} g(x, f(x)) p^t(dx|z, f(z))$ and $l_f^n(z) := \int_{\mathsf{X}} g_n(x, f(x)) q_n^t(dx|z, f(z))$. By recursively applying the inequalities in Assumption 5.4-(e) and in (5.29) we obtain

$$\sup_{f \in \mathbb{F}} |l_f(z)| \leq M_g \alpha^t w(z) + M_g \eta(w) \sum_{j=0}^{t-1} \alpha^j$$

and

$$\begin{aligned} \sup_{f \in \mathbb{F}} |l_f^n(z)| &\leq M_g \alpha_n^t w_n(z) + M_g \eta(w_n) \sum_{j=0}^{t-1} \alpha_n^j \\ &\leq M_g \alpha_{\max}^t \gamma_{\max} w(z) + M_g \eta(w) \gamma_{\max} \sum_{j=0}^{t-1} \alpha_{\max}^j, \end{aligned}$$

where $\alpha_{\max} := \sup_{n \geq n_0} \alpha_n < 1$. Then we have

$$\begin{aligned} &\sup_{(y,f) \in K \times \mathbb{F}} \left| \int_{\mathsf{X}} g_n(x, f(x)) q_n^{t+1}(dx|y, f(y)) - \int_{\mathsf{X}} g(x, f(x)) p^{t+1}(dx|y, f(y)) \right| \\ &= \sup_{(y,f) \in K \times \mathbb{F}} \left| \int_{\mathsf{X}} l_f^n(z) q_n(dz|y, f(y)) - \int_{\mathsf{X}} l_f(z) p(dz|y, f(y)) \right| \\ &= \sup_{(y,f) \in K \times \mathbb{F}} \left| \int_{\mathsf{X}} l_f^n(z) p(dz|y, f(y)) - \int_{\mathsf{X}} l_f(z) p(dz|y, f(y)) \right| \text{ (for } n \text{ sufficiently large)} \\ &\leq \sup_{(y,f) \in K \times \mathbb{F}} \left| \int_{K_\varepsilon^c} l_f^n(z) p(dz|y, f(y)) - \int_{K_\varepsilon^c} l_f(z) p(dz|y, f(y)) \right| + \sup_{(z,f) \in K_\varepsilon \times \mathbb{F}} |l_f^n(z) - l_f(z)| \\ &\leq R\varepsilon + \sup_{(z,f) \in K_\varepsilon \times \mathbb{F}} |l_f^n(z) - l_f(z)|, \end{aligned} \tag{5.46}$$

where R is given by

$$R \coloneqq M_g\bigg(\alpha^t + \alpha_{\max}^t\gamma_{\max} + \eta(w)\sum_{j=0}^{t-1}\alpha^j + \eta(w)\gamma_{\max}\sum_{j=0}^{t-1}\alpha_{\max}^j\bigg)$$

and the last inequality follows from Lemma 4.3. Since the claim holds for t and K_ε, the second term in (5.46) goes to zero as $n \to \infty$. Since ε is arbitrary, the result follows.

5.8.5 Proof of Lemma 5.20

To ease the notation, we define $M(\mathsf{X}_n)$, $M(\mathsf{X})$, and $M_w(\mathsf{X})$ as the subsets of $B(\mathsf{X}_n)$, $B(\mathsf{X})$, and $B_w(\mathsf{X})$, respectively, whose elements have (corresponding) norm less than one. Let $(x_k, a_k) \to (x, a)$ in $\mathsf{X}_n \times \mathsf{A}$. Since the pseudo state Δ_n is isolated and K_n is compact, we have two cases: (i) $x_k = x = \Delta_n$ for all k large enough, or (ii) $x_k \to x$ in K_n.

For the first case we have

$$\begin{aligned}
&\|p_n(\,\cdot\,|\Delta_n, a_k) - p_n(\,\cdot\,|\Delta_n, a)\|_{TV} \\
&\qquad = \sup_{g\in M(\mathsf{X}_n)}\bigg|\int_{\mathsf{X}_n} g(y)p_n(dy|\Delta_n, a_k) - \int_{\mathsf{X}_n} g(y)p_n(dy|\Delta_n, a)\bigg| \\
&\qquad \leq \sup_{g\in M(\mathsf{X})}\bigg|\int_{\mathsf{X}} g(y)q_n(dy|\Delta_n, a_k) - \int_{\mathsf{X}} g(y)q_n(dy|\Delta_n, a)\bigg| \qquad (5.47) \\
&\qquad = \sup_{g\in M(\mathsf{X})}\bigg|\int_{K_n^c}\bigg(\int_{\mathsf{X}} g(y)p(dy|z, a_k) - \int_{\mathsf{X}} g(y)p(dy|z, a)\bigg)\nu_n(dz)\bigg| \\
&\qquad \leq \int_{K_n^c}\sup_{g\in M(\mathsf{X})}\bigg|\int_{\mathsf{X}} g(y)p(dy|z, a_k) - \int_{\mathsf{X}} g(y)p(dy|z, a)\bigg|\nu_n(dz) \\
&\qquad \leq \int_{K_n^c}\sup_{g\in M_w(\mathsf{X})}\bigg|\int_{\mathsf{X}} g(y)p(dy|z, a_k) - \int_{\mathsf{X}} g(y)p(dy|z, a)\bigg|\nu_n(dz)
\end{aligned}$$

$$= \int_{K_n^c} \|p(\,\cdot\,|z,a_k) - p(\,\cdot\,|z,a)\|_w \nu_n(dz), \tag{5.48}$$

where (5.47) follows since if for any $g \in M(\mathsf{X}_n)$ we define $\bar{g} = g$ on K_n and $\bar{g} = g(\Delta_n)$ on K_n^c, then we have $\bar{g} \in M(\mathsf{X})$ and $\int_{\mathsf{X}_n} g(y) p_n(dy|x,a) = \int_{\mathsf{X}} \bar{g}(y) q_n(dy|x,a)$ for all $(x,a) \in \mathsf{X}_n \times \mathsf{A}$. Note that we have

$$\begin{aligned}
&\sup_{g \in M_w(\mathsf{X})} \left| \int_{\mathsf{X}} g(y) p(dy|z,a_k) - \int_{\mathsf{X}} g(y) p(dy|z,a) \right| \\
&\qquad\qquad \leq \int_{\mathsf{X}} w(y) p(dy|z,a_k) + \int_{\mathsf{X}} w(y) p(dy|z,a) \\
&\qquad\qquad \leq 2\big(\alpha + \eta(w)\big) w(z) \;\; \text{(by Assumption 5.4-(e), } \phi \leq 1 \text{, and } w \geq 1\text{).}
\end{aligned}$$

Since w (restricted to K_n^c) is ν_n-integrable, by the dominated convergence theorem (5.48) goes to zero as $k \to \infty$.

For the second case we have

$$\begin{aligned}
\|p_n(\,\cdot\,|x_k,a_k) - p_n(\,\cdot\,|x,a)\|_{TV} &= \sup_{g \in M(\mathsf{X}_n)} \left| \int_{\mathsf{X}_n} g(y) p_n(dy|x_k,a_k) - \int_{\mathsf{X}_n} g(y) p_n(dy|x,a) \right| \\
&\leq \sup_{g \in M(\mathsf{X})} \left| \int_{\mathsf{X}} g(y) q_n(dy|x_k,a_k) - \int_{\mathsf{X}} g(y) q_n(dy|x,a) \right| \\
&= \sup_{g \in M(\mathsf{X})} \left| \int_{\mathsf{X}} g(y) p(dy|x_k,a_k) - \int_{\mathsf{X}} g(y) p(dy|x,a) \right| \;\; \text{(since } x_k, x \in K_n\text{)} \\
&\leq \sup_{g \in M_w(\mathsf{X})} \left| \int_{\mathsf{X}} g(y) p(dy|x_k,a_k) - \int_{\mathsf{X}} g(y) p(dy|x,a) \right| \\
&= \|p(\,\cdot\,|x_k,a_k) - p(\,\cdot\,|x,a)\|_w.
\end{aligned}$$

By Assumption 5.4-(j) the last term goes to zero as $k \to \infty$.

5.8.6 Proof of Theorem 5.22

Similar to the proof of Lemma 5.6, we use induction. For $t = 1$, recalling the proof of Lemma 5.6, the claim holds by the following argument:

$$\begin{aligned}
\sup_{(y,f)\in \mathsf{Z}\times\mathbb{F}} \|p(\,\cdot\,|y,f(y)) - q_n(\,\cdot\,|y,f(y))\|_{TV} &\leq \sup_{y\in\mathsf{Z}} \sup_{(x,a)\in\mathcal{S}_{n,i_n(y)}\times\mathsf{A}} \|p(\,\cdot\,|y,a) - p(\,\cdot\,|x,a)\|_{TV} \\
&\leq K_2 \sup_{y\in\mathsf{Z}} \sup_{x\in\mathcal{S}_{n,i_n(y)}} d_{\mathsf{Z}}(x,y) \\
&\leq K_2 \max_{i\in\{1,\ldots,n\}} \mathrm{diam}(\mathcal{S}_{n,i}) \\
&\leq 2K_2\alpha(1/n)^{1/d}.
\end{aligned}$$

Now, assume the claim is true for $t \geq 1$. Again recalling the proof of Lemma 5.6, we have

$$\begin{aligned}
&\sup_{(y,f)\in \mathsf{Z}\times\mathbb{F}} \|p^{t+1}(\,\cdot\,|y,f(y)) - q_n^{t+1}(\,\cdot\,|y,f(y))\|_{TV} \\
&\leq 2 \sup_{(y,f)\in \mathsf{Z}\times\mathbb{F}} \left\|p^{t}(\,\cdot\,|y,f(y)) - q_n^{t}(\,\cdot\,|y,f(y))\right\|_{TV} + \sup_{(z,f)\in \mathsf{Z}\times\mathbb{F}} \left\|p(\,\cdot\,|z,f(z)) - q_n(\,\cdot\,|z,f(z))\right\|_{TV} \\
&\leq 2K_2\alpha(1/n)^{1/d} \sum_{i=1}^{t} 2^i + 2K_2\alpha(1/n)^{1/d} = 2K_2\alpha(1/n)^{1/d} \sum_{i=1}^{t+1} 2^i.
\end{aligned}$$

This completes the proof.

Chapter 6

Summary

In this thesis, we introduced and solved non-standard but operationally important quantization problems in source coding and stochastic control.

In Part I, the focus was on randomized quantization. We proposed three probabilistically equivalent general models randomized quantization (Models I, II, and III), which include all standard models in the literature. The usefulness of representing randomized quantization in three equivalent ways was demonstrated by studying a generalization of the distribution-preserving quantization problem from viewpoints of optimization and source coding. We first considered the existence of an optimal (minimum distortion) randomized quantizer subject to a constraint that the output of the quantizer has a given some fixed distribution. In this problem, we made use of Model III which was the most suitable for analysis. In addition to proving the existence of the optimal randomized quantizer, we also obtained structural result showing that optimal randomized quantizer can be constructed by randomizing only the set of quantizers with convex codecells. We then studied a Shannon theoretic version of the same problem using Models I and II of randomized quantization. Here, memoryless and stationary source is encoded to minimize the distortion under the constraint that the output of the coding scheme was also stationary and memoryless with a given distribution. We first characterized the set of achievable coding rate when the common randomness, shared between the encoder and the decoder, is unlimited and then characterized the set of achievable coding and common randomness rates when the common randomness is rate limited. We also considered variations of this coding problem where the effects of relaxing the fixed output distribution constraint and the 'private randomization' used by the decoder on the rate region were investigated and we exactly characterized the achievable rate regions.

Part II of the thesis was devoted to study approximation problems in stochastic control. Motivated mainly by the information transmission problem in stochastic control, we first considered the approximation of a discrete time Markov decision process by a reduced model having a finite set of actions. We established that finite action models can approximate the original model with arbitrary precision under two sets of assumptions. In the first set of assumpations it was assumed that the transition probability of the MDP is strongly continuous in the action variable while in the second one it was assumed that the transition probability of the MDP is weakly continuous in both state and action variables. A rate of convergence analysis was also established in which we derived an expression for the upper bound on the approximation error in terms of the number of representation points of the quantizer. In the second problem of Part II, finite state approximation of MDPs was considered. Analogous to the finite action approximation problem, we proved that the original model can be approximated with arbitrary precision by finite state models, where finite state models were obtained via quantization of the state space of the original

MDP. This result was established for MDPs with both compact state space and non-compact state space under different assumptions. The latter case was proved by approximating non-compact MDPs by compact ones. For MDPs with compact state space, we also obtained explicit rate of convergence bounds on the approximation error. Combining these two results led to a constructive algorithm for computing near optimal policies. The results were illustrated numerically by considering two different MDP models.

A unifying theme in above problems has been the interaction of control and information theory. In addition, the notion of 'approximate optimality' has been a practical design goal in both formulations.

6.1 Future Work

A possible research direction is to extend the results in Part II of the thesis to partially observed multi-agent decision problems for both the static and the dynamic case. Since the dynamic programming principle is in general not applicable to these type decision problems (i.e., optimal policies in general use the entire history), solving the approximation problem in this case is much more difficult and requires more sophisticated methods. Analogous to the belief MDP construction, a possible solution method is to transform the decision problem to a fully observed single-agent decision problem via expanding the state space, and then, applying results from Part II of this thesis to this equivalent problem. However, one drawback of this approach is that the mapping between the policies of the original model and the policies of the reduced model is not explicit in general. Another difficulty arises in quantizing the state space of the reduced model (i.e., the set of probability measures on some set) since the quantization of the set of probability measures is a difficult problem in itself even for finite probability spaces.

Another research direction related to Part II of the thesis is to study the near optimality of limited memory policies in partially observed multi-agent decision problems where a limited memory policy only uses the current M observations for some finite M. This problem is difficult to solve even in the single-agent case (i.e., POMDP) and can be thought of, to a certain degree, as a quantization problem with infinitely many representation points. Solving this problem would result in a considerable reduction of the size of the admissible policy set since the set of observations with limited memory is non-expanding over time. Therefore, one may compute the near optimal policy by approximation or, if applicable, exhaustive search methods. One way to approach this problem is again to transform the original model to a fully observed single-agent problem and examine the changes in the state variable of the reduced model when the limited observations are used. In the partially observed single-agent case, this corresponds to examining difference between the posterior distribution of the original state given entire observations and the posterior distribution of the original state given limited observations.

Regarding some future directions on Part I of the thesis, it would be interesting to generalize the results in Section 2.5 and Chapter 3 to setups where the target distribution is not memoryless but stationary and ergodic. Such problems also have connections with ergodic theory and the isomorphism theorem which investigates the existence of measurable bijective maps between two discrete stationary processes with identical entropies [71, 13]; in our case one would be considering a lossy version for such a problem.

Finally, we would like to explore the benefit of common randomness in decentralized control problems. It is known that randomness does not improve the system performance in a stochastic control system, or a team problem (see, e.g., [16, 107]), however common randomness is quite useful for coordination of actions, when there are multiple optimal policies and solving constrained optimization problems. In addition, in game theory, the presence of common randomness can lead to correlated equilibria as well as interesting consequences [6].

Bibliography

[1] E.F. Abaya and G.L. Wise. On the existence of optimal quantizers. *IEEE Trans. Inf. Theory*, 28(6):937–940, Nov. 1982.

[2] E.F. Abaya and G.L. Wise. Some remarks on the existence of optimal quantizers. *Statistics and Probability Letters*, 2(6):349–351, 1984.

[3] E. Akyol and K. Rose. On constrained randomized quantization. In *Proc. Data Compress. Conf.*, pages 72–81, Snowbird, Utah, USA, Apr. 2012.

[4] E. Akyol and K. Rose. On constrained randomized quantization. *IEEE Trans. Signal Processing*, 61(13):3291–3302, 2013.

[5] C.D. Aliprantis and K.C. Border. *Infinite Dimensional Analysis*. Springer, 2006.

[6] V. Anantharam and V. Borkar. Common randomness and distributed control: A counterexample. *Systems and Control Letters*, 56:568–572, 2007.

[7] C.L. Beck, S. Lall, T. Liang, and M. West. Model reduction, optimal prediction, and the Mori-Zwanzig representation of Markov chains. In *CDC 2009*, pages 3282–3287, Shanghai, Dec. 2009.

[8] C.H. Bennett, I. Devetak, A.W. Harrow, P.W. Shor, and A. Winter. The quantum reverse Shannon theorem and resource tradeoffs for simulating quantum channels. arXiv:0912.5537v5, 2013.

[9] D. P. Bertsekas and S. E. Shreve. *Stochastic optimal control: The discrete time case*. Academic Press New York, 1978.

[10] D.P. Bertsekas. Convergence of discretization procedures in dynamic programming. *IEEE Trans. Autom. Control*, 20(3):415–419, Jun. 1975.

[11] D.P. Bertsekas. *Dynamic Programming and Optimal Control: Volume II*. Athena Scientific, 1995.

[12] D.P. Bertsekas and J.N. Tsitsiklis. *Neuro-Dynammic Programming*. Athena Scientific, 1996.

[13] P. Billingsley. *Ergodic Theory and Information*. Wiley, 1965.

[14] P. Billingsley. *Probability and Measure*. Wiley, 3rd edition, 1995.

[15] P. Billingsley. *Convergence of probability measures*. New York: Wiley, 2nd edition, 1999.

[16] D. Blackwell. Memoryless strategies in finite-stage dynamic programing. *Ann. Math. Statist.*, pages 863–865, 1964.

[17] V.I. Bogachev. *Measure Theory: Volume II*. Springer, 2007.

[18] V. Borkar. Convex analytic methods in Markov decision processes. In E.A. Feinberg and A. Shwartz, editors, *Handbook of Markov Decision Processes*. Kluwer Academic Publisher, 2002.

[19] V. Borkar, S. Mitter, A. Sahai, and S. Tatikonda. Sequential source coding: An optimization viewpoint. In *Proc. IEEE Conference on Decision and Control*, pages 1035–1042, Seville, Spain, Dec. 2005.

[20] V.S. Borkar. On extremal solutions to stochastic control problems. *Appl. Math. Optim.*, 24(1):317–330, 1991.

[21] V.S. Borkar. White-noise representations in stochastic realization theory. *SIAM J. Control Optim.*, 31(5):1093–1102, 1993.

[22] L. Busoniu, D. Ernst, B..D. Schutter, and R. Babuska. Approximate dynamic programming with a fuzzy parametrization. *Automatica*, 46:804–814, 2010.

[23] R. Cavazos-Cadena. Finite-state approximations for denumerable state discounted Markov decision processes. *Appl. Math. Optim.*, 14:1–26, 1986.

[24] H.S. Chang, M.C. Fu, J. Hu, and S.I. Marcus. A survey of some simulation-based methods in Markov decision processes. *Communications in Information System*, 7:59–92, 2007.

[25] T.M. Cover and J.A. Thomas. *Elements of Information Theory.* Wiley, 2nd edition, 2006.

[26] I. Csiszár. On an extremum problem of information theory. *Studia Scientiatrum Mathematicarum Hungarica*, 9:57–70, 1974.

[27] I. Csiszar and J. Korner. *Information Theory: Coding Theorems for Discrete Memoryless Channels.* Akademiai Kiado, Budapest, 1981.

[28] P. Cuff. Communications requirements for generating correlated random variables. In *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2008.

[29] P. Cuff. Distributed channel synthesis. *IEEE Trans. Inf. Theory*, 59(11):7071–7096, Nov. 2013.

[30] P.W. Cuff, H.H. Permuter, and T.M. Cover. Coordination capacity. *IEEE Trans. Inf. Theory*, 56(9):4181–42205, Sep. 2010.

[31] L. Dubins and D. Freedman. Measurable sets of measures. *Pacific J. Math*, 14(4):1211–1222, 1964.

[32] R. M. Dudley. *Real Analysis and Probability.* New York: Chapman and Hall, 1989.

[33] F. Dufour and T. Prieto-Rumeau. Approximation of Markov decision processes with general state space. *J. Math. Anal. Appl.*, 388:1254–1267, 2012.

[34] F. Dufour and T. Prieto-Rumeau. Finite linear programming approximations of constrained discounted Markov decision processes. *SIAM J. Control Optim.*, 51(2):1298–1324, 2013.

[35] F. Dufour and T. Prieto-Rumeau. Approximation of average cost Markov decision processes using empirical distributions and concentration inequalities. *Stochastics*, pages 1–35, 2014.

[36] E. B. Dykin. *Controlled Markov Processes.* Berlin, New York: Springer-Verlag, 1979.

[37] A.M. Farahmand, R. Munos, and C. Szepesvari. Error propagation for approximate policy and value iteration. *Advances in Neural Information Processing Systems*, pages 568–576, 2010.

[38] E.A. Feinberg. On measurability and representation of strategic measures in Markov decision processes. *Statistics, Probability and Game Theory*, 30:29–43, 1996.

[39] E.A. Feinberg, P.O. Kasyanov, and N.V. Zadioanchuk. Average cost Markov decision processes with weakly continuous transition probabilities. *Math. Oper. Res.*, 37(4):591–607, Nov. 2012.

[40] E.A. Feinberg, P.O. Kasyanov, and M.Z. Zgurovsky. Partially observable total-cost Markov decision process with weakly continuous transition probabilities. arXiv:1401.2168, 2014.

[41] E.A. Feinberg and M.E. Lewis. Optimality inequalities for average cost Markov decision processes and the stochastic cash balance problem. *Math. Oper. Res.*, 32(4):769–783, Nov. 2007.

[42] B.L. Fox. Finite-state approximations to denumerable state dynamic programs. *J. Math. Anal. Appl.*, 34:665–670, 1971.

[43] A. El Gamal and Y.H. Kim. *Network Information Theory.* Cambridge, 2011.

[44] E. Gordienko and O. Hernandez-Lerma. Average cost Markov control processes with weighted norms: Existence of canonical policies. *Appl. Math.*, 23(2):199–218, 1995.

[45] S. Graf and H. Luschgy. *Foundations of Quantization for Probability Distributions*. Springer, 2000.

[46] R.M. Gray. *Entropy and Information Theory*. Springer, 2011.

[47] R.M. Gray and D.L. Neuhoff. Quantization. *IEEE Trans. Inf. Theory*, 44:2325–2383, Oct. 1998.

[48] R.M. Gray and T.G. Stockham. Dithered quantizers. *IEEE Trans. Inf. Theory*, 39(3):805–812, May 1993.

[49] L. Györfi and M. Kohler. Nonparametric estimation of conditional distributions. *IEEE Trans. Inf. Theory*, 53:1872–1878, May 2007.

[50] M. Hairer. Ergodic properties of Markov processes. *Lecture Notes*, 2006.

[51] O. Hernández-Lerma. Finite-state approximations for denumerable multidimensional state discounted Markov decision processes. *J. Math. Anal. Appl.*, 113:382–388, 1986.

[52] O. Hernández-Lerma. *Adaptive Markov Control Processes*. Springer-Verlag, 1989.

[53] O. Hernández-Lerma and J.B. Lasserre. Weak conditions for average optimality in Markov control processes. *Systems Control Lett.*, 22:287–291, 1994.

[54] O. Hernández-Lerma and J.B. Lasserre. *Discrete-Time Markov Control Processes: Basic Optimality Criteria*. Springer, 1996.

[55] O. Hernández-Lerma and J.B. Lasserre. *Further Topics on Discrete-Time Markov Control Processes*. Springer, 1999.

[56] O. Hernández-Lerma and J.B. Lasserre. *Markov Chains and Invariant Probabilities*. Birkhauser, 2003.

[57] O. Hernández-Lerma, R. Montes-De-Oca, and R. Cavazos-Cadena. Recurrence conditions for Markov decision processes with Borel state space: a survey. *Ann. Oper. Res.*, 28(1):29–46, 1991.

[58] O. Hernández-Lerma and R. Romera. Limiting discounted-cost control of partially observable stochastic systems. *SIAM J. Control Optim.*, 40(2):348–369, 2001.

[59] K. Hinderer. Lipshitz continuity of value functions in Markovian desision processes. *Math. Meth. Oper. Res.*, 62:3–22, 2005.

[60] R. Jain and P.P. Varaiya. Simulation-based uniform value function estimates of Markov decision processes. *SIAM J. Control Optim.*, 45(5):1633–1656, Nov 2006.

[61] W. Kreitmeier. Optimal vector quantization in terms of Wasserstein distance. *J. Multivariate Anal.*, 102(8):1225–1239, 2011.

[62] H. Langen. Convergence of dynamic programming models. *Math. Oper. Res.*, 6(4):493–512, 1981.

[63] A. Lapidoth and P. Narayan. Reliable communication under channel uncertainity. *IEEE Trans. Inf. Theory*, 44(6):2148–2177, Oct. 1998.

[64] M. Li, J. Klejsa, and W.B. Kleijn. Distribution preserving quantization with dithering and transformation. *IEEE Signal Processing Letters*, 17(12):1014–1017, 2010.

[65] M. Li, J. Klejsa, and W.B. Kleijn. On distribution preserving quantization. *arXiv preprint*, 2011.

[66] A.S. Manne. Linear programming and sequential decisions. *Management Sciences*, 6(3):259–267, 1980.

[67] K. Marton. Bounding $\bar{d}$-distance by informational divergence: a method to prove measure concentration. *Ann. Probab.*, 24(2):857–866, 1996.

[68] M. McAsey and L. Mou. Optimal locations and the mass transport problem. *Contemp. Math.*, 226:131–148, 1999.

[69] S. P. Meyn. *Control Techniques for Complex Networks*. Cambridge University Press, 2007.

[70] B.M. Oliver, J. Pierce, and C.E. Shannon. The philosophy of PCM. *Proc. IRE*, 36:1324–1331, Nov. 1948.

[71] D. Ornstein. Bernoulli shifts with the same entropy are isomorphic. *Advances in Math.*, 4:337–352, 1970.

[72] R. Ortner. Pseudometrics for state aggregation in average reward Markov decision processes. In *Algorithmic Learning Theory*. Springer-Verlag, 2007.

[73] K.R. Parthasarathy. *Probability Measures on Metric Spaces*. AMS Bookstore, 1967.

[74] D. Pollard. Quantization and the method of k-means. *IEEE Trans. Inf. Theory*, 28(2):199–205, Mar. 1982.

[75] A. Pratelli. On the sufficiency of c-cyclical monotonicity for optimality of transport plans. *Math. Z.*, 258(3):677–690, 2008.

[76] M.L. Puterman. *Markov Decision Processes*. Wiley-Interscience, 2005.

[77] Z. Ren and B.H. Krogh. State aggregation in Markov decision processes. In *CDC 2002*, Las Vegas, December 2002.

[78] D. Rhenius. Incomplete information in Markovian decision models. *Ann. Statist.*, 2:1327–1334, 1974.

[79] L. Roberts. Picture coding using pseudo-random noise. *IEEE Trans. Inf. Theory*, 8(2):145–154, Feb. 1962.

[80] B.V. Roy. Performance loss bounds for approximate value iteration with state aggregation. *Math. Oper. Res.*, 31(2):234–244, May 2006.

[81] N. Saldi, T. Linder, and S. Yüksel. Approximation of stationary control policies by quantized control in Markov decision processes. In *Communication, Control, and Computing, 52nd Annual Allerton Conference on*, pages 78–84, Monticello, IL, Oct. 2013.

[82] N. Saldi, T. Linder, and S. Yuksel. Randomized quantizers and optimal design with marginal constraints. In *ISIT 2013*, Istanbul, July 2013.

[83] N. Saldi, T. Linder, and S. Yüksel. Asymptotic optimality of quantized policies in stochastic control under weak continuity conditions. In *IEEE Conf. Decision Control*, Los Angeles, December 2014.

[84] N. Saldi, T. Linder, and S. Yüksel. Output constrained lossy source coding with limited common randomness. arXiv:1411.5767, 2014.

[85] N. Saldi, T. Linder, and S. Yüksel. Randomized source coding with limited common randomness. In *Communication, Control, and Computing, 52nd Annual Allerton Conference on*, Monticello, IL, Oct. 2014.

[86] N. Saldi, T. Linder, and S. Yüksel. Asymtotic optimality and rates of convergence of quantized stationary policies in stochastic control. *IEEE Trans. Autom. Control*, 60(2):553–558, Feb. 2015.

[87] N. Saldi, T. Linder, and S. Yuksel. Randomized quantization and source coding with constrained output distribution. *IEEE Trans. Inf. Theory*, 61(1):91–106, Jan. 2015.

[88] N. Saldi, S. Yüksel, and T. Linder. Near optimality of quantized policies in stochastic control under weak continuity conditions. arXiv:1410.6985, 2014.

[89] N. Saldi, S. Yüksel, and T. Linder. Asymptotic optimality of finite approximations to Markov decision processes with Borel spaces. arXiv:1503.02244, 2015.

[90] N. Saldi, S. Yüksel, and T. Linder. On finite state-action approximation of Markov decision processes with general state and action spaces. to appear in American Control Conference (ACC), 2015.

[91] M. Schäl. On dynamic programming: compactness of the space of policies. *Stochastic Process. Appl.*, 3(4):345–364, 1975.

[92] L. Schucman. Dither signals and their effect on quantization noise. *IEEE Trans. Commun.*, 12(4):162–165, Dec. 1964.

[93] R. Serfozo. Convergence of Lebesgue integrals with varying measures. *Sankhya Ser.A*, pages 380–402, 1982.

[94] C. E. Shannon. Coding theorems for a discrete source with a fidelity criterion. *IRE Nat. Conv. Rec.*, part 4:138–143, 1959.

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

[96] J.N. Tsitsiklis. Asynchronous stochastic approximation and Q-learning. *Machine Learning*, 16:185–202, 1994.

[97] J.N. Tsitsiklis and B.V. Roy. Feature-based methods for large scale dynamic programming. *Machine Learning*, 22:59–94, 1996.

[98] O. Vega-Amaya. The average cost optimality equation: a fixed point approach. *Bol. Soc. Mat. Mexicana*, 9(3):185–195, 2003.

[99] C. Villani. *Optimal transport: old and new.* Springer, 2009.

[100] D.J. White. Finite-state approximations for denumerable state infinite horizon discounted Markov decision processes. *J. Math. Anal. Appl.*, 74:292–295, 1980.

[101] D.J. White. Finite-state approximations for denumerable state infinite horizon discounted Markov decision processes with unbounded rewards. *J. Math. Anal. Appl.*, 186:292–306, 1982.

[102] W. Whitt. Approximations of dynamic programs I. *Math. Oper. Res.*, 3(3):231–243, 1978.

[103] A. Winter. Compression of sources of probability distributions and density operators. arXiv:quant-ph/0208131, 2002.

[104] A.D. Wyner. The common information of two dependent random variables. *IEEE Trans. Inf. Theory*, 21(2):163–179, Mar. 1975.

[105] Y. Yamada, S. Tazaki, and R.M. Gray. Asymptotic performance of block quantizers with difference distortion measures. *IEEE Trans. Inf. Theory*, 26:6–14, Jan. 1980.

[106] E.-H. Yang and J.C. Kieffer. On the performance of data compression algorithms based upon string matching. *IEEE Trans. Inf. Theory*, 44(1):47–65, Jan. 1998.

[107] S Yuksel and T. Basar. *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*. Birkhäuser, 2013.

[108] S. Yüksel and T. Linder. Optimization and convergence of observation channels in stochastic control. *SIAM J. Control Optim.*, 50(2):864–887, 2012.

[109] A.A. Yushkevich. Reduction of a controlled Markov model with incomplete data to a problem with complete information in the case of Borel state and control spaces. *Theory Prob. Appl.*, 21:153–158, 1976.

[110] R. Zamir. *Lattice Coding for Signals and Networks*. Oxford University Press (to be published), 2014.

[111] R. Zamir and M. Feder. On universal quantization by randomized uniform/lattice quantizers. *IEEE Trans. Inf. Theory*, 38(2):428–436, Mar. 1992.

[112] R. Zamir and M. Feder. Information rates of pre/post-filtered dithered quantizers. *IEEE Trans. Inf. Theory*, 42(5):1340–1353, Sep. 1996.

[113] R. Zamir and K. Rose. Natural type selection in adaptive lossy compression. *IEEE Trans. Inf. Theory*, 47(1):99–111, Jan. 2001.

[114] J. Ziv. On universal quantization. *IEEE Trans. Inf. Theory*, 31(3):344–347, May 1985.