Department of Electrical and Computer Engineering

Extremum Problems with Total Variation Distance Metric on the Space of Probability Measures and Applications

by

Ioannis Tzortzis

A DISSERTATION SUBMITTED TO THE DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING OF UNIVERSITY OF CYPRUS
IN PARTIAL FULFILLMENT FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

January 2015

Declaration of Authorship

The present doctoral dissertation was submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy of the University of Cyprus. It is a product of original work of my own, unless otherwise mentioned through references, notes, or any other statements.

....

Ioannis Tzortzis
Department of Electrical and Computer Engineering
University of Cyprus
January, 2015

Abstract

Extremum problems with total variation distance metric on the space of probability measures are of fundamental importance in stochastic optimal control, information theory and communication, mathematical finance, decision theory and in statistics and probability. Among others, the investigation of such problems utilizes concepts from measure and probability theory and function space optimization, and has applications in minimax stochastic control via dynamic programming, approximation of high-dimensional probability distributions by lower-dimensional, model reduction, etc. In this thesis, the formulation of extremum problems involving total variation distance metric, their extremum solutions, their discussion in terms of applications, and their application to the areas of minimax stochastic control and Markov process approximation, are investigated.

The first part of the thesis deals with the formulation of extremum problems, in which systems are represented by probability distributions on abstract spaces, and pay-offs are represented by total variation distance metric defined on the space of probability measures, subject to linear functional constraints on the space of probability measures, and vice-versa; that is with the roles of total variation metric and linear functional interchanged. By utilizing concepts from signed measures, the extremum solutions of such problems are obtained in closed form, and an associated emerging water-filling property of the partitioning of the alphabet spaces of the extremum solutions is elaborated. The results are derived for abstract spaces, while the high level ideas are also discussed for denumerable spaces endowed with the discrete topology.

The second part of the thesis addresses optimality of stochastic control strategies on a finite and on an infinite horizon, via dynamic programming subject to total variation distance ambiguity on the conditional distribution of the controlled process. The stochastic control problem is formulated using minimax theory, in which the control minimizes the pay-off while the conditional distribution, from the total variation set, maximizes it. By employing certain results of the first part, new dynamic programming recursions are derived which, in addition to the standard terms, include the oscillator semi-norm of the value function that codify the level of ambiguity with respect to total variation distance ball. In addition, for the infinite horizon case new policy iteration algorithms are presented to compute the optimal

strategies.

The third part of the thesis deals with the problem of approximating a finite state Markov process with a large number of states by a lower-dimensional process, with respect to certain measures of discriminating or approximating the distribution of the high-dimensional Markov process by a reduced one. By drawing upon the results of the first part, the approximation problem of finite state Markov process by another non-necessarily Markov process with reduced state space, is formulated as an optimization problem, with respect to a certain pay-off subject to a fidelity criterion defined by the total variation distance metric. The water-filling behaviour of the transition probabilities of the approximated process and the proposed recursive algorithms are new, and applicable to a variety of approximation problems spanning from optimal state reduction to optimal state aggregation.

Πρόλογος

Η βελτιστοποίηση προβλήματων με ακρότατα που χρησιμοποιούν ως μετρική απόστασης την ολική κύμανση στο χώρο των μέτρων πιθανοτήτων είναι θεμελιώδους σημασίας στον στοχαστικό έλεγχο, στη θεωρία πληροφορίας και της επικοινωνίας, στη μαθηματική χρηματοοικονομία, στη θεωρία αποφάσεων και στατιστικής και στην θεωρία πιθανοτήτων. Μεταξύ άλλων, η έρευνα τέτοιων προβλημάτων χρησιμοποιεί έννοιες απο την θεωρία μέτρου και πιθανοτήτων, απο την θεωρία βελτιστοποίησης συναρτήσεων, και έχει εφαρμογές στο στοχαστικό έλεγχο ελαχιστομεγίστων μέσω δυναμικού προγραμματισμού, στην προσέγγιση κατανομών πιθανοτήτων μεγάλων διαστάσεων απο κατανομές πιθανοτήτων μικρότερων διαστάσεων, κλπ. Στην παρούσα διατριβή, η διατύπωση και η επίλυση προβλημάτων βελτιστοποίησης με μετρική απόστασης την ολική κύμανση, η συζήτηση πιθανών εφαρμογών τους, καθώς και η εφαρμογή τους στον στοχαστικό έλεγχο ελαχιστομεγίστων και στην προσέγγιση διαδικασιών Markov, ερευνώνται.

Το πρώτο μέρος της διατριβής ασχολείται με την διατύπωση και επίλυση προβλημάτων με ακρότατα, στα οποία τα συστήματα ελέγχου εκπροσωπούνται απο κατανομές πιθανοτήτων σε αφηρημένους χώρους, και τα κριτήρια κόστους εκπροσωπούνται απο την μετρική της ολικής κύμανσης στο χώρο των μέτρων πιθανοτήτας και υπόκεινται σε γραμμικούς συναρτησιακούς περιορισμούς στο χώρο των μέτρων πιθανοτήτας και αντίστροφα, δηλαδή, με τους ρόλους της μετρικής ολικής κύμανσης και των γραμμικών συναρτησιακών να εναλλάσσονται. Χρησιμοποιώντας έννοιες των προσημασμένων μέτρων, οι βέλτιστες λύσεις τέτοιων προβλημάτων λαμβάνονται σε κλειστή μορφή και οι ιδιότητες τους ερευνώνται και συζητούνται. Τα αποτελέσματα προκύπτουν χρησιμοποιώντας αφηρημένους χώρους, ενώ οι ιδέες τους συζητούνται επίσης για αριθμήσιμους χώρους εφοδιασμένους με την διακριτή τοπολογία.

Το δεύτερο μέρος της διατριβής ασχολείται με την βελτιστοποίηση στοχαστικών στρατηγικών ελέγχου σε πεπερασμένο και σε άπειρο χρονικό ορίζοντα, μέσω δυναμικού προγραμματισμού χρησιμοποιώντας ως απόσταση αβεβαιότητας την ολική κύμανση της δεσμευμένης κατανομής της ελεγχόμενης διαδικασίας. Το στοχαστικό πρόβλημα ελέγχου διατυπώνεται χρησιμοποιώντας την θεωρία ελαχιστομεγίστων, κατά την οποία η διαδικασία ελέγχου ελαχιστοποιεί το κριτήριο κόστους ενώ η δεσμευμένη κατανομή απο το σύνολο της ολικής κύμανσης, το μεγιστοποιεί. Χρησιμοποιώντας συγκεκριμένα αποτελέσματα του πρώτου μέρους, νέες εξισώσεις δυναμικού προγραμματισμού εξάγονται, οι οποίες εκτός απο τους κλασικούς όρους συμπεριλαμβάνουν επίσης και τον ταλαντωτή ημι-νόρμα που κωδικοποιεί το επίπεδο της αβεβαιότητας. Επιπρόσθετα, όσον αφορά τα στοχαστικά προβλήματα με άπειρο χρονικό ορίζοντα, νέοι αλγόριθμοι παρουσιάζονται για τον υπολογισμό των βέλτιστων στρατηγικών ελέγχου.

Το τρίτο μέρος της διατριβής ασχολείται με το πρόβλημα προσέγγισης μιας πεπερασμένης διαδικασίας Markov με ένα μεγάλο αριθμό καταστάσεων απο μία διαδικασία χαμηλότερων διαστάσεων, ως προς ορισμένα μέτρα προσέγγισης της κατανομής της διαδικασίας Markov από κατανομές χαμηλότερων διαστάσεων. Αντλώντας αποτελέσματα απο το πρώτο μέρος, το πρόβλημα προσέγγισης μιας πεπερασμένης διαδικασίας Markov απο μια άλλη διαδικασία (μη-απαραίτητα Markov) με λιγότερες καταστάσεις διατυπώνεται ως ένα πρόβλημα βελτιστοποίησης, με το κριτήριο κόστους να υπόκειται σε κριτήρια ακρίβειας ορισμένα απο το μέτρο απόστασης της ολικής κύμανσης. Η συμπεριφορά των πιθανοτήτων μετάβασης της προσεγγιζόμενης διαδικασίας και οι προτεινόμενοι αναδρομικοί αλγόριθμοι έιναι καινούργιοι, και μπόρουν να εφαρμοστούν σε μια σειρά προβλημάτων που εκτείνονται απο την βέλτιστη αναγωγή καταστάσεων έως και την βέλτιστη συνάθροιση καταστάσεων.

Acknowledgements

First and foremost, i would like to express my deepest appreciation and thanks to my advisor, Professor Charalambos D. Charalambous, who has been a great advisor for me and his commitment to excellence has inspired me. Thank you so much for your excellent guidance, constant encouragement, patience and unlimited support during my postgraduate studies at the University of Cyprus. It has been a great pleasure and a privilege to have the opportunity to work with you the last seven years.

Special thanks go to Dr. Themistoklis Charalambous, who apart from being a brilliant collaborator, he is also a good friend. I sincerely appreciate your invaluable contribution, support and enthusiasm about my research. Thank you very much for your brilliant comments, remarkable advices and friendship during my PhD studies.

I would like to thank Professor Jan H. van Schuppen for the interesting discussions we had during his visits at the University of Cyprus. I am deeply indebted to you and thank you for your kind concern about my research work. I benefited by your advice and your willingness to give me constructive comments which shaped my work better.

For this dissertation I would like to thank my committee members Professor Christoforos Hadjicostis and Dr. Ioannis Krikidis, for their time, interest and helpful comments.

I would also like to thank my schoolmates and friends Christo Kourtellari, Photi Stavrou and Ioanna Ioannou, for their help in various ways. I am very grateful for the time we spent learning together and for all the fun we have had in the last seven years.

Finally, I would like to thank my family. Words are not enough to express my love to my mother Maria, my father Andreas, my sister Anastasia and my brother George. Thank you for your love, support, understanding and encouragement in every possible way. Last but not least, my deepest love, respect and gratitude goes to my lovely wife Katerina and my son Andreas. Andrea your birth gave a new meaning to my life. I can't wait seeing you beside me at my PhD graduation ceremony. Katerina your unconditional love, selfless support, patience and encouragement have motivated me to proceed through the doctoral program and complete my dissertation. Thank you for sharing all the joy and hardships during my long academic journey.

Contents

List of Figures

List of Tables

Abbreviations

TV	**T**otal **V**ariation
KL	**K**ullback **L**eibler
RV	**R**andom **V**ariable
DP	**D**ynamic **P**rogramming
MC	**M**arkov **C**hain
MP	**M**arkov **P**rocess
FSM	**F**inite **S**tate **M**arkov
MCM	**M**arkov **C**ontrol **M**odel
D-FCM	**D**iscounted-**F**eedback **C**ontrol **M**odel
D-MCM	**D**iscounted-**M**arkov **C**ontrol **M**odel

Notation

$\mathbb{Z}$	set of the integers
$\mathbb{Z}_+ = \{1, 2, \dots\}$	set of the positive integers
$\mathbb{Z}^n = \{1, 2, \dots, n\}$	set of the first n integers
$\mathbb{N} = \{0, 1, 2, \dots\}$	set of the natural numbers
$\mathbb{N}^n = \{0, 1, 2, \dots, n\}$	set of the first $n + 1$ natural numbers
$\mathbb{R}$	set of the real numbers
$\mathbb{R}_+ = [0, \infty)$	set of the positive real numbers
$\mathbb{R}_+^{n \times m}$	set of matrices with entries from $\mathbb{R}_+$ and with n rows and m columns
$\mathcal{M}_{sm}$	set of the finite signed measures
$\mathcal{M}_{sm}^0$	set of the finite signed measures which integrate to zero
$(\Omega, \mathcal{F})$	measurable space
$(\Omega, \mathcal{F}, P)$	probability space
$\mathbb{E}\{\cdot\}$	expectation operator
$\mathbb{E}\{\cdot \mid \mathcal{G}\}$	conditional expectation operator with respect to σ-algebra $\mathcal{G}$
$\perp$	singularity of two probability measures
$\overline{\Sigma}$	closure of a set Σ
ξ^+	$\max\{\xi, 0\}$
ξ^-	$\max\{-\xi, 0\}$
$\|\| \cdot \|\|_{TV}$	total variation norm
$\| \cdot \|$	absolute value or cardinality of a set
$\|\| \cdot \|\|_p$	p-norm defined on the set of Lebesgue measurable real-valued functions
$L_p(A)$	space of Lebesgue measurable real-valued functions defined on A with norm $\|\| \cdot \|\|_p$
A_{ij}	the $\{i, j\}^{th}$ component of matrix A

$A_{i\bullet}$	the i^{th} row of matrix A
$A_{\bullet j}$	the j^{th} column of matrix A
A^T	transpose of matrix A
I_n	$n \times n$ matrix with ones on its diagonal and zero elsewhere

1 Introduction

Extremum problems with total variation distance metric on the space of probability measures are of fundamental importance in stochastic optimal control, information theory and communication, mathematical finance, decision theory and in statistics and probability. Among others, the investigation of such problems utilizes concepts from measure and probability theory and function space optimization, and has applications in minimax stochastic control via dynamic programming, approximation of high-dimensional probability distributions by lower-dimensional, model reduction, etc. In this thesis, the formulation of extremum problems involving total variation distance metric, their extremum solutions, their discussion in terms of applications, and their application to the areas of minimax stochastic control and Markov process approximation, are investigated.

In the first part of the thesis (Chapter 3), our aim is to investigate extremum problems with pay-off being the total variation distance metric defined on the space of probability measures, subject to linear functional constraints on the space of probability measures, and vice-versa; that is, with the roles of total variation metric and linear functional interchanged. Utilizing concepts from signed measures, the extremum probability measures of such problems are obtained in closed form, by identifying the partition of the support set and the mass of these extremum measures on the partition. The results are derived for abstract spaces (specifically, complete separable metric spaces known as Polish spaces), while the high level ideas are also discussed for denumerable spaces endowed with the discrete topology. The main results of this part include:

(i) characterization of the properties of the extremum problems under investigation;

(ii) characterization of extremum measures on abstract spaces, and closed form solutions of the extremum measures for finite alphabet spaces;

(iii) convexity and concavity properties of extremum solutions.

In the second part of the thesis (Chapters 4 and 5), our aim is to address optimality of stochastic control strategies on a finite and on an infinite horizon, via dynamic programming subject to total variation distance ambiguity on the conditional distribution of the controlled process. We formulate the stochastic control problem using minimax theory, in which the minimizer-controller chooses its policy from some admissible set to minimize the pay-off, while the maximizer-conditional distribution chooses its distribution from a set described by the total variation distance to maximize the pay-off. First, we employ certain results from the first part, in particular, the maximization of a linear functional on the space of probability measures on abstract spaces, among those probability measures which are within a total variation distance from a nominal probability measure. Then we utilize the solution of the maximization to solve minimax stochastic control with deterministic control strategies, under a Markovian and a non-Markovian assumption, on the conditional distributions of the controlled process. The main results of this part include:

(i) minimax optimization subject to total variation distance ambiguity constraint;

(ii) new dynamic programming recursions, which involve the oscillator seminorm of the value function, in addition to the standard terms;

(iii) new infinite horizon discounted dynamic programming equation, the associated contractive property, and a new policy iteration algorithm;

(iv) new infinite horizon average dynamic programming equations, and new policy iteration algorithms.

Our aim in the third part of the thesis (Chapters 6), is to approximate a finite-state Markov process by another process with fewer states, called the approximating process. The approximation problem is formulated using two different methods. The first method, utilizes the total variation distance to discriminate the transition probabilities of a high-dimensional Markov process and a reduced order Markov process. The approximation is obtained by maximizing a linear functional defined in terms of transition probabilities of the reduced order Markov process over a total variation distance constraint. The transition probabilities of the approximated Markov process are given by a water-filling solution. The second method, utilizes total variation distance as a measure of discriminating the invariant probability of a Markov process by the approximating process. The approximation is obtained via two alternative formulations: (a) maximizing a linear functional of the occupancy distribution of the Markov process, and (b) maximizing the entropy of the approximating process invariant probability. For both formulations, once the reduced invariant probability is obtained, which does not necessarily correspond to a Markov process, a further approximation by a Markov process is proposed which minimizes the Kullback-Leibler divergence. The approximation is given by a water-filling like solution. Finally, the theoretical results of both methods are applied to specific examples to illustrate the methodology, and the water-filling behavior of the approximations. The main results of this part, based on the first method, include:

(i) a direct method for Markov by Markov approximation based on the transition probabilities of the original FSM process and the reduced one, exhibiting a water-filling behavior,

and, based on the second method:

(i) extremum measures which exhibit a water-filling behavior, and solve the approximation problems;

(ii) optimal partition functions which aggregate the original finite-state Markov process to form the reduced order finite-state Markov process;

(iii) iterative algorithms to compute the invariant distribution of the approximating process.

Next, we give an outline of the thesis motivation and objectives, and details which will follow in later chapters.

1.1. Extremum Problems with Total Variation Distance

Total variation distance metric on the space of probability measures is a fundamental quantity in statistics and probability, which over the years appeared in many diverse applications. In information theory it is used to define strong typicality and asymptotic equipartition of sequences generated by sampling from a given distribution [17]. In decision problems, it arises naturally when discriminating the results of observation of two statistical hypotheses [17]. In studying the ergodicity of Markov Chains, it is used to define the Dobrushin coefficient and establish the contraction property of transition probability distributions [41]. Moreover, distance in total variation of probability measures is related via upper and lower bounds to an anthology of distances and distance metrics [30]. The measure of distance in total variation of probability measures is a strong form of closeness of probability measures, and, convergence with respect to total variation of probability measures implies their convergence with respect to other distances and distance metrics.

In Chapter 3, we formulate and solve several extremum problems involving the total variation distance metric and we discuss their applications in the areas of control, communication and statistics. The main problems investigated are the following.

(a) Extremum problems of linear functionals on the space of measures subject to a total variation distance metric constraint defined on the space of measures.

(b) Extremum problems of total variation distance metric on the space of measures subject to linear functionals on the space of measures.

(c) Applications of these extremum problems, and their relations to other problems.

The formulation of these extremum problems, and their discussion in terms of applications are developed at the abstract level, in which systems are represented by probability distributions on abstract spaces (complete separable metric space, known as Polish spaces [23]), pay-offs are represented by linear functionals on the space of probability measures or by distance in variation of probability measures, and constraints by linear functionals or distance in variation of probability measures. We consider Polish spaces since they are general enough to handle various models of practical interest, such as stochastic control problems on Borel spaces.

Utilizing concepts from signed measures, closed form expressions of the probability measures are derived which achieve the extremum of these problems. The construction of the extremum measures involves the identification of the partition of their support set, and their mass defined on these partitions. Throughout the derivations we make extensive use of lower and upper bounds of pay-offs which are achievable, and convexity and concavity properties (i.e., these are convex optimization problems on the space of probability measures). Several simulations are carried out to illustrate the different features of the extremum solution of the various problems. An interesting observation concerning one of the extremum problems is its equivalent formulation as an extremum problem involving the oscillator semi-norm of the pay-off functional. The formulation and results obtained for these problems at the abstract level are discussed throughout the chapter in the context of various applications, often assuming denumerable spaces endowed with the discrete topology. Some specific applications discussed are listed below.

(i) Dynamic Programming Under Uncertainty in Distribution of the Controlled Process: To deal with uncertainty of transition probability distributions, via minimax theory, with total variation distance metric uncertainty constraints to codify the impact of incorrect distribution models on performance of the optimal strategies [16]. This formulation is applicable to Markov and non-Markov decision problems subject to uncertainty in distribution of the controlled process.

(ii) Approximation of Probability Distributions with Total Variation Distance Metric: To approximate a given high-dimensional probability distribution μ on a measurable space $(\Sigma, \mathcal{B}(\Sigma))$ by another lower-dimensional distribution ν on $(\overline{\Sigma}, \mathcal{B}(\overline{\Sigma}))$, $\overline{\Sigma} \subseteq \Sigma$, via minimization of the total variation distance metric between them subject to linear functional constraints. Model and graph reduction can be handled via such approximations. Graphs, for example, constitute the foundation of many real-world datasets. However, the size of the graph can become prohibitive to understand essential information that they contain. The reduction of graph-based models is significant in a wide variety of applications, such as placement of autonomous sensors, modeling Central Processing Unit (CPU) and database demands in web-based software engineering, and identifying the evolution in clusters within massive dynamic datasets in database research.

(iii) Maximization or Minimization of Entropy Subject to Total Variation Distance Metric Constraints: To invoke insufficient reasoning based on maximizing the entropy $H(\nu)$ of an unknown probability distribution ν on denumerable space Σ subject to a constraint on the total variation distance metric. This problem can be also associated with limited-length code word design that is useful in communication between distributed systems that aim at minimizing communication delays.

1.2. Dynamic Programming on Finite and Infinite Horizon

Dynamic programming recursions are often employed in optimal control and decision theory to establish existence of optimal strategies, to derive necessary and sufficient optimality conditions, and to compute the optimal strategies either in closed form or via algorithms [13, 39, 54]. The cost-to-go and the corresponding dynamic programming recursion, in their general form, are functionals of the conditional distribution of the underlying controlled process given the past and present values of the control and controlled processes [13]. Thus, any ambiguity of the controlled process conditional distribution will affect the optimality of the strategies. The term "ambiguity" is used to differentiate from the term "uncertainty" often used in control nomenclature to account for situations in which the true and nominal distribution (induced by models) are absolutely continuous, and hence they are defined on the same state space. This distinction is often omitted from various robust deterministic and stochastic control/filtering approaches, including minimax and risk-sensitive formulations [1, 3, 6, 8, 14, 15, 24, 33, 38, 44, 48, 56, 59]. The class of models is described by a ball with respect to the total variation distance between the nominal distribution and the true distribution, hence it admits distributions which are singular with respect to the nominal distribution.

The main objective in Chapters 4 and 5, is to investigate the effect on the cost-to-go and dynamic programming recursion of the ambiguity in the controlled process conditional distribution, and hence on the optimal decision strategies. Specifically, we quantify the conditional distribution ambiguity of the controlled process by a ball with respect to the total variation distance metric, centered at a nominal conditional distribution, and then we derive a new dynamic programming using minimax theory, with two players: player I the control process and player II the conditional distribution (controlled process), opposing each others actions. In this minimax game formulation, player's I objective is to minimize the cost-to-go, while player's II objective is to maximize it. The maximization over the total variation distance ball of player II is addressed by employing certain results of Chapter 3, related to the maximization of linear functionals on a subset of the space of signed measures. Utilizing these results, a new dynamic programming recursion is presented which, in addition to the standard terms, includes additional terms that codify the level of ambiguity allowed by player II with respect to the total variation distance ball. Thus, the effect of player I, the control process, is to minimize, in addition to the classical terms, the difference between the maximum and minimum values of the cost-to-go, scaled by the radius of the total variation distance ambiguity set. We treat in a unified way the finite horizon case, under both the Markovian and non-Markovian nominal controlled processes, and the infinite horizon case. For the infinite horizon case we consider both discounted and average pay-offs. For the infinite horizon case with discounted pay-off we show that the operator associated with the resulting dynamic programming equation under total variation distance ambiguity is contractive, and consequently, we derive a new policy iteration algorithm to compute the optimal strategies. For the infinite horizon with average pay off we derive new dynamic programming equations under certain irreducibility/reducibility conditions, and we present new policy iteration algorithms. Finally, we provide examples for the finite and for the infinite horizon cases.

Previous related work on optimization of stochastic systems subject to total variation distance ambiguity is found in [50] for continuous time controlled diffusion processes described by Itô differential equations. However, the solution method employed in [50] is fundamentally different; it approaches the maximization problem indirectly, by employing Large Deviations concepts to derive the maximizing measure as a convex combination of a tilted probability measure and the nominal measure, under restrictions on the class of measures considered. The dynamic programming equation derived in [50] is limited by the assumption that the maximizing measure is absolutely continuous with respect to the nominal measure.

In Chapters 4 and 5, our focus is to understand the effect of total variation distance ambiguity of the conditional distribution on dynamic programming, from a different point of view, utilizing concepts from signed measures. Consequently, we derive new dynamic programming recursions which depends explicitly on the radius of the total variation distance, the closed form expression of the maximizing measure, or the oscillator seminorm of the value function. One of the fundamental properties of the maximizing conditional distribution is that, as the ambiguity radius increases, the maximizing conditional distribution becomes singular with respect to the nominal distribution. The point to be made here is that the total variation distance ambiguity set admits controlled process distributions which are not necessarily defined on the same state space as the nominal controlled process distribution. In terms of robustness of the optimal policies, this additional feature is very attractive compared to minimax techniques based on relative entropy uncertainty or risk-sensitive payoffs [1,3,6,8,14,15,24,33,38,44,48,56,59], because often the true controlled distribution lies on a higher-dimensional state space compared to the nominal controlled process distribution.

1.3. Approximation of Markov Processes by Lower Dimensional Processes

Finite-State Markov (FSM) processes are often employed to model physical phenomena in many diverse areas, such as machine learning, information theory (lossy compression), networked control systems, telecommunications, speech processing, systems biology, etc. In many of these applications the state-space of the Markov process is prohibitively large, in performing simulations. One approach often pursue to overcome the large number of states is to approximate the Markov process by a lower-dimensional Markov process, with respect to certain measures of discriminating or approximating the distribution of the high-dimensional Markov process by a reduced one. Such methods are described using relative entropy as a measure of approximation in [19, 55, 58, 61] (and references therein). Further discussion of model reduction methods for Markov chains can be found in [7]. In general, approximating a Markov process by another process subject to a fidelity of reproduction is not necessarily Markov, but a hidden Markov process. This is a well known result of Information Theory [17], dealing with lossy compression of Markov sources. Model reduction of hidden Markov models via aggregation can be found in [18, 20, 58]. Specifically, in [20] the aggregated hidden Markov model is expressed as a function of a partition function and a recursive learning algorithm is proposed, which solves the optimal partition problem.

In this chapter, the approximation problem of a FSM process by another process (FSM or not) with reduced state-space is formulated as an optimization problem, with respect to a certain pay-off subject to a fidelity constraint defined by the total variation distance metric, using two different methods which are described below.

Method 1: Approximate the transition probabilities of a FSM process by another FSM process with reduced transition probabilities. This approximation problem is formulated as a maximization of a linear functional of the transition probabilities of the reduced FSM process, subject to a fidelity criterion defined by the total variation distance between the transition probabilities of the high and low FSM process.

Method 2: Approximate a FSM process by another process with lower-dimensional state-space, without imposing the assumption that the approximating process is also a Markov process. The following two formulations are investigated.

(a) Maximize an average pay-off, described in terms of the occupation measure of the high-dimensional Markov process, subject to a fidelity criterion defined by the total variation distance metric, between the invariant distribution of the higher-dimensional Markov process and the invariant distribution of the lower-dimensional process.

(b) Maximize the entropy (Jayne's maximum entropy [34]) of the invariant distribution of the lower-dimensional process, subject to a fidelity criterion defined by the total variation distance metric, between the invariant distribution of the higher-dimensional Markov process and the invariant distribution of the lower-dimensional process.

For both formulations, the resulting approximated process is not necessarily Markov. The crux of the approach considered lies in finding an optimal partition function which aggregates states of the original FSM process to form the reduced order process. A Markov process approximation is obtained by minimizing the Kullback-Leibler divergence between the transition probability matrices of the high and low-dimensional FSM process.

1.4. Organization of Thesis

We now briefly indicate the content of the remaining of the thesis.

In Chapter 2 we briefly summarize mathematical theory that will be used in this thesis. Thus, in Section 2.1, we review basic concepts of probability theory following a measure-theoretic approach, and we summarize results from signed measures which are particularly relevant for the characterization of the extremum measures. In Section 2.2, we introduce total variation distance metric and we review some of its basic properties. In Section 2.3, we briefly state relations of the total variation distance to other distance metrics, and we give some of its applications.

In Chapter 3 we investigate extremum problems with pay-off being the total variation distance metric defined on the space of probability measures, subject to linear functional constraints on the space of probability measures, and vice-versa. In Section 3.1, we introduce the precise definitions of the extremum problems under investigation, while several related problems are discussed together with their applications. In Section 3.2, some of the properties of the extremum problems are discussed, and signed measures are utilized to convert the extremum problems into equivalent ones, and to characterize the extremum measures on abstract spaces. In Section 3.3, closed form expressions of the extremum measures are derived for finite alphabet spaces, by identifying the support sets and the extremum measures on these sets. The results are also extended to the countable alphabet case. In Section 3.4, several examples illustrate how the optimal distribution of the extremum problems behaves, for different scenarios of the support set of the distribution, and an application to the area of information theory is presented. Finally, Section 3.5 concludes by discussing the most important results obtained in this chapter.

In Chapter 4, we address optimality of stochastic control strategies on a finite horizon, via dynamic programming subject to total variation distance ambiguity on the conditional distribution of the controlled process. Thus, in Section 4.1, first we introduce the definition of finite horizon discounted Markov control model with deterministic strategies, and then we describe the abstract formulation of the minimax problem under total variation distance ambiguity. In Section 4.2, we introduce the general definition of finite horizon discounted feedback control model with randomized and deterministic control policies, under total variation distance uncertainty, and then we apply the characterization of the maximizing distribution to the dynamic programming recursion. In Section 4.3, we apply the abstract setup to both the feedback control model and to the Markov control model, and we derive new dynamic programming recursions which characterize the optimality of minimax strategies. In Section 4.4, we illustrate the new dynamic programming recursions through the well-known inventory control and machine replacement examples. Finally, Section 4.5 concludes by discussing the most important results obtained in this chapter.

In Chapter 5, we introduce the dynamic programming subject to total variation distance ambiguity on the conditional distribution of the controlled process for infinite horizon Markov Control Models (MCM), with optimality criterion, the expected discounted reward and the average pay-off per unit time. In Section 5.2, we consider the infinite horizon Markov control model with the expected discounted reward as an optimality criterion, and we show that the operator associated with the dynamic programming equation is contractive, and we introduce a new policy iteration algorithm. In Section 5.3, we study the infinite horizon Markov Control Model with the average pay-off per unit time as an optimality criterion. We derive the new dynamic programming equations under total variation distance ambiguity with and without imposing the irreducibility condition. In addition, we introduce the corresponding policy iteration algorithms for average cost dynamic programming. In Section 5.4, we illustrate an application of the infinite horizon minimax problem for both discounted and average pay-off. Moreover, we present an additional example for average optimality criterion, without imposing the irreducibility assumption. Finally, Section 5.5 concludes by discussing the most important results obtained in this chapter.

In Chapter 6, we approximate a finite state Markov process with a large number of states by a lower-dimensional process, called the approximating process. In Section 6.1, we introduce the approximation problem using two different methods. The first method, utilizes the total variation distance to discriminate the transition probabilities of a high-dimensional Markov process and a reduced order Markov process. The second method, utilizes total variation distance as a measure of discriminating the invariant probability of a Markov process by the approximating process, and the approximation is obtained via two alternative formulations: (a) maximizing a linear functional of the occupancy distribution of the Markov process, and (b) maximizing the entropy of the approximating process invariant probability. For both formulations, once the reduced invariant probability is obtained, a further approximation by a Markov process is proposed which minimizes the Kullback-Leibler divergence. In Section 6.2, a direct method for Markov by Markov approximation based on Method 1 is derived. In Section 6.3, the solution of the approximation problem based on Method 2 is given for both formulations, and the corresponding recursive algorithms and the optimal partition functions are presented. In Section 6.4, several examples are presented to illustrate the approximation methods. Finally, Section 6.5 concludes by discussing the most important results obtained in this chapter.

Chapter 7, is the concluding chapter, where in Section 7.1, we give a brief summary of our work and indicate its main contributions, and in Section 7.2 we mention some topics for further research.

2

Background Material

In this chapter, we briefly summarize background material. In particular, in Section 2.1, basic mathematical concepts of probability theory and certain results from signed measures are reviewed. In Section 2.2, total variation distance metric is introduced and some of its properties are discussed. Moreover, in Section 2.3 the relations of total variation distance to other distance metrics are briefly stated and some of its applications are discussed.

2.1. Mathematical Preliminaries

In this section, we review the basic concepts of probability theory following a measure-theoretic approach. We also summarize certain results from signed measures which are particularly relevant to the characterization of the extremum measures.

2.1.1. Elements of Functional Analysis

The following definitions can be found in [5, 37, 53].

A set $\mathcal{X}$ which is furnished with a measure of distance between any two elements of the set is called a metric space. A function $d: \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is called a metric if it satisfies the following properties.

(i) $d(x,y) \geq 0, \;\; \forall x, y \in \mathcal{X}$, and $d(x,y) = 0$ if and only if $x = y$;

(ii) $d(x, y) = d(y, x), \ \ \forall x, y \in \mathcal{X}$;

(iii) $d(x, z) \leq d(x, y) + d(y, z), \ \ \forall x, y, z \in \mathcal{X}$.

Definition 2.1. *(Continuous Function) Let $(\mathcal{X}, d_{\mathcal{X}})$ and $(\mathcal{Y}, d_{\mathcal{Y}})$ be two metric spaces. A function $f : \mathcal{X} \mapsto \mathcal{Y}$ is continuous at $x_0 \in \mathcal{X}$ if for every $\epsilon > 0$ there is a $\delta > 0$ such that $d_{\mathcal{X}}(x, x_0) < \delta$ implies $d_{\mathcal{Y}}(f(x), f(x_0)) < \epsilon$. The function f is continuous on $\mathcal{X}$ if it is continuous at every point in $\mathcal{X}$.*

Let $(\mathcal{X}, d)$ be a metric space. The open ball $B_r(\alpha)$, with radius $r > 0$ and center $\alpha \in \mathcal{X}$ is the set

$$B_r(\alpha) = \{x \in \mathcal{X} | d(x, \alpha) < r\}.$$

The closed ball, $\overline{B}_r(\alpha)$, is the set

$$\overline{B}_r(\alpha) = \{x \in \mathcal{X} | d(x, \alpha) \leq r\}.$$

Definition 2.2. *(Open and Closed Sets) A subset G of a metric space $\mathcal{X}$ is open if for every $x \in G$ there is an $r > 0$ such that $B_r(x)$ is contained in G. A subset F of $\mathcal{X}$ is closed if its complement $F^c = \mathcal{X} \setminus F$ is open.*

For example, an open ball is an open set, and a closed ball is a closed set.

Definition 2.3. *(Vector Space) Let $\mathcal{X}$ denote an arbitrary set and let $F = \mathbb{R}$ or $\mathbb{C}$ denote the field of real or complex numbers. A vector space $\mathcal{X}$ over a field F is a structure $(\mathcal{X}, F)$ consisting of two operations $+ : \mathcal{X} \times \mathcal{X} \mapsto \mathcal{X}$, $\times : F \times \mathcal{X} \mapsto \mathcal{X}$, called respectively vector addition and scalar multiplication such that the following conditions hold:*

(i) Additive commutativity. $x + y = y + x$, $\forall x, y \in \mathcal{X}$.

(ii) Additive associativity. $(x + y) + z = x + (y + z)$, $\forall x, y, z \in \mathcal{X}$.

(iii) Additive identity. There exists an element in $\mathcal{X}$, denoted by 0, such that $x + 0 = x$, $\forall x \in \mathcal{X}$.

(iv) Additive inverse. For all $x \in \mathcal{X}$ there exists an unique element in $\mathcal{X}$, denoted $-x$, such that $x + (-x) = 0$.

(v) Multiplicative associativity. $(a \times b) \times x = a \times (b \times x)$, $\forall a, b \in F$, $x \in \mathcal{X}$.

(vi) Distributivity.

$$a \times (x + y) = a \times x + a \times y, \quad \forall x, y \in \mathcal{X}, a \in F;$$
$$(a + b) \times x = a \times x + b \times x, \quad \forall x \in \mathcal{X}, a, b \in F.$$

(vii) *If* $1 \in F$ *is the multiplicative identity of* F *then* $1 \times x = x$, $\forall x \in \mathcal{X}$.

Definition 2.4. *(Normed Space) A Normed Space is a vector space* $(\mathcal{X}, F)$ *furnished with a norm* $\|\cdot\|_{\mathcal{X}}$ *and denoted by* $(\mathcal{X}, F, \|\cdot\|_{\mathcal{X}})$. *The norm* $\|\cdot\|_{\mathcal{X}} : \mathcal{X} \to F$ *must satisfy the following properties.*

(i) $\|x\| \geq 0, \;\; \forall x \in \mathcal{X}$;

(ii) $\|x\| = 0$, *if and only if* $x = 0$;

(iii) $\|ax\| = |\alpha| \|x\|, \; \forall x \in \mathcal{X}, \;\; \alpha \in F$;

(iv) $\|x + y\| \leq \|x\| + \|y\|, \; \forall x, y \in \mathcal{X}$.

Definition 2.5. *(Cauchy sequence) Let* $(\mathcal{X}, F, \|\cdot\|_{\mathcal{X}})$ *be a normed space. A sequence* $\{x_n : n \in \mathbb{N}\} \in \mathcal{X}$ *is said to be a Cauchy sequence if*

$$\lim_{n\to\infty} \|x_{n+p} - x_n\| = 0 \;\; \textit{for every} \;\; p \geq 1.$$

Definition 2.6. *(Banach Space) A normed space* $(\mathcal{X}, F, \|\cdot\|_{\mathcal{X}})$ *is said to be complete if every Cauchy sequence of* $\mathcal{X}$ *has a limit in* $\mathcal{X}$. *A complete normed space is called a Banach space.*

Definition 2.7. *(Topological Space) Let* $\mathcal{Z}$ *be a set and* $\mathcal{B}_{\mathcal{Z}}$ *a collection of subsets of* $\mathcal{Z}$. *Then* $\mathcal{B}_{\mathcal{Z}}$ *is called a topology in* $\mathcal{Z}$ *if the following properties hold [36].*

(i) $\emptyset \in \mathcal{B}_{\mathcal{Z}}$ *and* $\mathcal{Z} \in \mathcal{B}_{\mathcal{Z}}$;

(ii) *If* $\mathcal{Z}_i \in \mathcal{B}_{\mathcal{Z}}, i = 1, 2, ..., n$, *then* $\bigcap_{i=1}^{n} \mathcal{Z}_i \in \mathcal{B}_{\mathcal{Z}}$;

(iii) *If* $\{\mathcal{Z}_i\}$ *is an arbitrary collection of elements of* $\mathcal{Z}$ *(finite, countable, or uncountable), then* $\bigcup_i \mathcal{Z}_i \in \mathcal{B}_{\mathcal{Z}}$.

The pair $(\mathcal{Z}, \mathcal{B}_{\mathcal{Z}})$ is called a topological space and the members of $\mathcal{B}_{\mathcal{Z}}$ are called open sets in $\mathcal{Z}$. If $f : (\mathcal{Z}, \mathcal{B}_{\mathcal{Z}}) \to (\mathcal{Y}, \mathcal{B}_{\mathcal{Y}})$, then f is continuous provided $f^{-1}(\mathcal{Y}_i) \subset \mathcal{Z}$ is an open set for every open set $\mathcal{Y}_i \subset \mathcal{Y}$. Moreover, f is continuous at the point $x_0 \in \mathcal{Z}$ if for every neighborhood A of $f(x_0)$ there exists a neighborhood B of x_0 such that $f(B) \subset A$.

Definition 2.8. *(Convex Set) A set* S *in a vector space* $(\mathcal{X}, F)$ *is called a convex set if the line segment joining any pair of points of* S *lies entirely in* S. *The former statement is equivalent to saying that for any pair of vectors* $u \in S$, $v \in S$, *the vector* $(1-t)u + tv \in S$, $\forall t \in [0, 1]$.

Definition 2.9. *(Convex Function) Let* $(\mathcal{X}, \mathbb{R}, \|\cdot\|_{\mathcal{X}})$ *be any Banach space and* f *a (possibly extended) real valued function defined on* $\mathcal{X}$*,* $f : \mathcal{X} \mapsto \mathbb{R} \cup \{+\infty, -\infty\}$*. The function* f *is said to be convex if* $\mathcal{X}$ *is a convex set and if for any* $x, y \in \mathcal{X}$*,* $x \neq y$*, and* $\alpha \in [0, 1]$

$$f((1-\alpha)x + ay) \leq (1-\alpha)f(x) + \alpha f(y).$$

Similarly, if $\mathcal{X}$ *is replaced by a closed convex subset* $\Gamma \subseteq \mathcal{X}$*, and* f *satisfies the above inequality for all* $x, y \in \Gamma$*, we say that* f *is convex on* Γ*. The function* f *is said to be concave if* $-f$ *is convex.*

Definition 2.10. *Let* $(\mathcal{X}, \mathbb{R}, \|\cdot\|_{\mathcal{X}})$ *be any Banach space and* f *a real-valued function* $f : \mathcal{X} \mapsto \mathbb{R}$*. The function* f *is said to be lower semi-continuous at* $x \in \mathcal{X}$ *if for every sequence* $\{x_n : n \in \mathbb{Z}_+\}$ *converging to* x

$$f(x) \leq \lim_{n\to\infty} \inf f(x_n) = \sup_{n\geq 1} \inf_{k\geq n} f(x_k)$$

and it is said to be upper semi-continuous at x *if*

$$f(x) \geq \lim_{n\to\infty} \sup f(x_n) = \inf_{n\geq 1} \sup_{k\geq n} f(x_k)$$

and it is said to be lower or upper semi-continuous on set $\Gamma \subset \mathcal{X}$*, if the corresponding statements hold for all* $x \in \Gamma$*.*

2.1.2. Measurable Space and Probability Space

Probability theory deals with random experiments associated with elementary outcomes Ω and the set of all events of interest $\mathcal{F}$.

Definition 2.11. *(Algebra) Let* Ω *be a set of elementary outcomes and* $\mathcal{F}$ *be a non-empty collection of subsets of* Ω*. Then* $\mathcal{F}$ *is called an Algebra on* Ω *if the following properties hold.*

(i) $\Omega \in \mathcal{F}$ *(The Sample Space is an element of* $\mathcal{F}$*);*

(ii) If $A \in \mathcal{F}$ *then* $A^c = \Omega - A \in \mathcal{F}$*, where* A^c *is the complementation of A relative to* Ω *(if a subset of* Ω *belongs to* $\mathcal{F}$*, then so is its complement);*

(iii) If $A_i \in \mathcal{F}, i = 1, 2, ..., n$*, then* $\bigcup_{i=1}^{n} A_i \in \mathcal{F}$ *(if a finite number of subsets belong to* $\mathcal{F}$*, then so is their union).*

Clearly, an algebra is a collection of subsets of a set Ω, which a) contains Ω and b) is closed under complementation and finite unions. The members of $\mathcal{F}$ are called $\mathcal{F}$-measurable sets or measurable sets.

Definition 2.12. *(σ-Algebra) An algebra $\mathcal{F}$ on Ω is called a σ-Algebra on Ω if it is closed under countable unions, that is, if the following properties hold.*

(i) $\Omega \in \mathcal{F}$ *;*

(ii) If $A \in \mathcal{F}$ *then* $A^c \in \mathcal{F}$*;*

(iii) If $A_i \in \mathcal{F}$, $i = 1, 2, \ldots$ *then* $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

If $\mathcal{F}$ is a field the pair $(\Omega, \mathcal{F})$ is called a measurable space and the elements of $\mathcal{F}$ are called events and are said to be measurable sets in Ω. Fields and σ-fields are convenient mathematical objects which express how much we know about the outcome of the experiment.

Probability Space. In order to grade the possibility of occurrences of events associated with a random experiment we need to define a function (set function) which attaches a numerical value to events $A \in \mathcal{F}$. A function

$$\mu : \mathcal{F} \mapsto [0, \infty], \quad \mu(A) \in [0, \infty], \quad \forall A \in \mathcal{F}$$

is called a finite-additive set function, if μ satisfies the following two conditions.

$$\begin{aligned} &\text{(i)} \;\; \mu(\emptyset) = 0; \\ &\text{(ii)} \;\; \mu(A \cup B) = \mu(A) + \mu(B), \text{ if } A, B \in \mathcal{F} \text{ and } A \bigcap B = \emptyset. \end{aligned}$$

A finite-additive set function μ on a σ-algebra is called a measure, if it is countably-additive, and a probability measure if it is countably-additive and $\mu(\Omega) = 1$, hence the following definition.

Definition 2.13. *(Probability Measure) Let $(\Omega, \mathcal{F})$ be a measurable space. The function*

$$\mathbb{P} : \mathcal{F} \to [0, 1], \; \mathbb{P}(A) \in [0, 1], \; \forall A \in \mathcal{F}$$

is called a probability measure on $(\Omega, \mathcal{F})$ if it satisfies the following properties.

(i) $\mathbb{P}(\emptyset) = 0$*;*

(ii) $\mathbb{P}(\Omega) = 1$*;*

(iii) $\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$*, if* $A_i \in \mathcal{F}, \forall i$ *and* $\{A_j\}_{j=1}^{\infty}$ *are disjoint, e.g.,* $A_i \bigcap A_j = \emptyset, \forall i \neq j$.

The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a probability space.

A probability Space $(\Omega, \mathcal{F}, \mathbb{P})$ is called complete if whenever $B \in \mathcal{F}$ and $\mathbb{P}(B) = 0$ then $A \in \mathcal{F}$ for all $A \subset B$. The subsets A of an event B of zero probability is called a null set, therefore $(\Omega, \mathcal{F}, \mathbb{P})$ is complete if $\mathcal{F}$ includes all events of zero probability. If $\mathcal{F}_1$, $\mathcal{F}_2$ are two σ-algebras on Ω then the σ-algebra generated by $\mathcal{F}_1$, $\mathcal{F}_2$ is denoted by $\mathcal{F}_1 \bigvee \mathcal{F}_2$. Any probability space $(\Omega, \mathcal{F}, \mathbb{P})$ which is not complete can be uniquely extended to the σ-algebra $\overline{\mathcal{F}} = \mathcal{F} \bigvee \{\text{null sets}\}$.

Since the intersection of arbitrary σ-algebras of a subset of Ω is a σ-algebra of subsets of Ω, then for an arbitrary family $\mathcal{A}$ of subsets of Ω there is a smallest σ-algebra $\mathcal{F}$ in Ω such that $\mathcal{A} \subset \mathcal{F}$.

Theorem 2.1. *(Smallest σ-algebra) Let Ω be a sample space and $\mathcal{A}$ be a collection of subsets of Ω. There exists a smallest σ-algebra $\mathcal{F}(\mathcal{A})$ on Ω containing $\mathcal{A}$, which is constructed by*

$$\mathcal{F}(\mathcal{A}) = \bigcap_i \Big\{\mathcal{N}_i : \mathcal{N}_i \textit{ is a } \sigma\textit{-algebra on } \Omega, \mathcal{A} \subset \mathcal{N}_i\Big\}.$$

This is called the σ-algebra generated by $\mathcal{A}$, and it is often denoted by $\mathcal{F}(\mathcal{A}) = \sigma(\mathcal{A})$.

Borel Set. Let $\mathcal{X}$ be a topological space (e.g., $\mathcal{X} = \mathbb{R}^n$: the collection of all n-tuples $\{x = (x_1, x_2, ..., x_n) : \ x_i \in \mathbb{R}, \ 1 \leq i \leq n\}$). Then there exists a smallest σ-algebra $\mathcal{F}$ on $\mathcal{X}$ such that every open set $\mathcal{A} \subset \mathcal{X}$ belongs to $\mathcal{F}$. The elements $A \in \mathcal{F}$ are called Borel sets and the σ-algebra $\mathcal{F} = \mathcal{F}(\mathcal{X})$ is called a Borel σ-algebra. For example, if $\mathcal{X} = \mathbb{R}^n$, and $\mathcal{A}$ is the collection of all open sets of $\mathbb{R}^n$, the Borel σ-algebra denoted by $\mathcal{B}(\mathbb{R}^n)$, contains all open sets, their complements (closed sets), all the countable unions of open sets, and all the countable unions of closed sets. In fact, $\mathcal{B}(\mathbb{R}^n)$ is the smallest σ-algebra of subsets of $\mathbb{R}^n$ containing all sets of the form $\{x = (x_1, x_2, ..., x_n) : \quad x_1 \in A_1, x_2 \in A_2, ..., x_n \in A_n\}$, where $A_j, j = 1, ..., n$ are intervals in $\mathbb{R}$, which are closed, open, semi-open, points, etc. Clearly, $\mathcal{A} = \{\text{Collection of all open intervals of } \mathbb{R}^n\}$ is not a σ-algebra, but there exists many σ-algebra containing $\mathcal{A}$ as a subset. The smallest σ-algebra containing $\mathcal{A}$ is the σ-algebra generated by $\mathcal{A}$. The pair $\Big(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)\Big)$ is a measurable space, called, the Borel measurable space, and probability measures defined on it are called Borel probability measures.

2.1.3. Measurable Functions and Random Variables

Let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be two measurable spaces, and let $f : (\Omega_1, \mathcal{F}_1) \to (\Omega_2, \mathcal{F}_2)$. Then, the function f is called $\mathcal{F}_1/\mathcal{F}_2$ or $\mathcal{F}_1$ measurable if

$$f^{-1}(A) \stackrel{\triangle}{=} \{\omega : f(\omega) \in A\} \in \mathcal{F}_1, \ \forall A \in \mathcal{F}_2.$$

The set $f^{-1}(A)$ is called the inverse image of $A \in \mathcal{F}_2$. If $f : \Omega \to \mathcal{Y}$ where $(\Omega, \mathcal{F})$ is a measurable space, $\mathcal{Y}$ is a topological space (e.g., $\mathbb{R}^n$), then f is $\mathcal{F}/\mathcal{B}(\mathbb{R}^n)$-measurable provided $f^{-1}(V) \in \mathcal{F}$ for every open set $V \subset \mathcal{B}(\mathbb{R}^n)$.

The σ-algebra $\mathcal{F}(f)$ generated by f is the smallest σ-algebra on Ω containing all the sets $\{f^{-1}(V) : V \subset \mathcal{Y} \text{ is open}\}$ and f will be $\mathcal{F}(f)/\mathcal{Y}$. Moreover, if $\mathcal{Y} = \mathbb{R}^n$ then

$$\mathcal{F}(f) = \{f^{-1}(V) : V \in \mathcal{B}(\mathbb{R}^n)\}.$$

Clearly, if $(\Omega, \mathcal{B})$ is a Borel measurable space and $f : \Omega \to \mathcal{Y}$, where $\mathcal{Y}$ is a topological space and f is a continuous function, then from the definition of continuous function

$$f^{-1}(V) \in \mathcal{B}, \ \forall \text{ open set } V \subset \mathcal{Y}. \tag{2.1}$$

Hence, every continuous function is Borel measurable, called Borel function, e.g., $f : (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)) \to (\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$ is a Borel function.

If a probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ is defined, where $X : (\Omega, \mathcal{F}) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, is a measurable function then X is called a Random Variable (RV) defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 2.14. *(Random Variable) Let $X : (\Omega, \mathcal{F}) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ be a function defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then X is called an* n-dimensional random variable *RV (measurable function)*

$$X : (\Omega, \mathcal{F}) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$$

if for every $A \in \mathcal{B}(\mathbb{R}^n)$ the set

$$X^{-1}(A) \stackrel{\triangle}{=} \{\omega : X(\omega) \in A\} \in \mathcal{F}.$$

Clearly, the $\sigma - algebra$ $\mathcal{F}(X)$ generated by X is the smallest $\sigma - algebra$ on Ω containing all the sets

$$X^{-1}(A) \ : \ A \subset \mathbb{R}^n \text{ is open}$$

under which X is measurable. Equivalently,

$$\mathcal{F}(X) = X^{-1}\Big(\mathcal{B}(\mathbb{R}^n)\Big) = \{X^{-1}(B) : B \in \mathcal{B}(\mathbb{R}^n)\}$$

is the smallest σ-algebra on Ω under which X is measurable. However, $\mathcal{B}(\mathbb{R}^n)$ is generated by products of open sets of the form

$$\{x{=}(x_1, x_2, ..., x_n){:} -\infty < x_1 \le \alpha_1, \dots, -\infty < x_n \le \alpha_n\}, \ \alpha_j \in \mathbb{R}, \ 1 \le j \le n. \tag{2.2}$$

Therefore, for X to be a RV it is sufficient for every set of the form

$$\{\omega : X_1(\omega) \le \alpha_1, \dots, X_n(\omega) \le \alpha_n\}, \ \alpha_j \in \mathbb{R}, \ 1 \le j \le n$$

to be an event. This is because $\mathcal{B}(\mathbb{R}^n)$ is the family of subsets obtained by starting with (2.2) and taking repeatedly all complements, countable unions, intersections. Also, if $\{X_t : 0 \leq t \leq T\}$ is a family of random variables $\mathcal{F}_{0,T}^X \triangleq \sigma(X_t : 0 \leq t \leq T) = \bigvee_{t\in T} \mathcal{F}(X_t) = \sigma(\bigcup_{t\in T} \mathcal{F}(X_t))$ is the smallest σ-algebra on Ω under which $\{X_t : 0 \leq t \leq T\}$ are measurable.

2.1.4. Probability Distribution Function

Let $\left(\Omega, \mathcal{F}, \mathbb{P}\right)$ be a probability space and let $X : (\Omega, \mathcal{F}) \to (\Omega_1, \mathcal{F}_1)$ be an $\mathcal{F}/\mathcal{F}_1$- measurable RV. From the point of view of computations, it is often convenient to work with an induced measure on $\mathcal{F}_1$. This amounts to defining the probability measure induced by the RV on its range space rather than treat points with respect to the measure $\mathbb{P}$ and work with a probability measure on $\mathcal{F}_1$ with $\omega \in \Omega_1$ as its sample values.

The RV, $X : (\Omega, \mathcal{F}) \to (\Omega_1, \mathcal{F}_1)$ induces a probability measure P_X on $(\Omega_1, \mathcal{F}_1)$ by

$$P_X(A_1) \triangleq \mathbb{P} \circ X^{-1}(A_1) = \mathbb{P}\Big(\{\omega : X(\omega) \in A_1\}\Big) = P(X \in A_1), \qquad A_1 \in \mathcal{F}_1.$$

If $(\Omega_1, \mathcal{F}_1) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ then we can work with a probability measure on $\mathcal{B}(\mathbb{R})$ with $x \in \mathbb{R}$ as its sample points.

Definition 2.15. *(Probability Distribution) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a Probability Space and $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ a RV. The function $F_X(\cdot)$ defined as*

$$F_X(x) \triangleq \mathbb{P}(\{\omega : X(\omega) \leq x\}) = P_X(X \leq x)$$

is called the (cumulative) probability distribution of the RV X.

Thus, the relationship

$$P_X(A) = \mathbb{P}(\{\omega : X(\omega) \in A\})$$

defines a probability measure P_X on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Note that $F_X(x)$ is a probability distribution defined on $\mathbb{R}$, e.g., it corresponds to the probability measure corresponding to $\mathbb{P}$ induced by $X(\cdot)$ on $\mathbb{R}$.

Suppose $X_1, X_2, ..., X_n$ are n real-valued RV's and $X \triangleq (X_1, X_2, ..., X_n)$, then

$$X : (\Omega, \mathcal{F}) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$$

is a measurable function. The function

$$F_X(x) = F_X(x_1, x_2, ...x_n) = \mathbb{P}(\{\omega : \bigcap_{i=1}^{n} \{X_i(\omega) \leq x_i\}\}), \qquad x \in \mathbb{R}^n$$

is called the joint probability distribution function of X. Similarly as above, the relationship

$$P_X(A) = \mathbb{P}(\{\omega : X(\omega) \in A\}), \;\; A \in \mathcal{B}(\mathbb{R}^n)$$

defines a Borel probability measure.

A real-valued RV X is said to be discrete if there exists a countable set $\Sigma = \{x_i : i \in \mathbb{Z}_+\}$ such that

$$\sum_{x_i \in \Sigma} \mathbb{P}(\{\omega : X(\omega) = x_i\}) = 1.$$

If X is discrete, then the distribution function F_X is a function which is constant except for jumps at $x_i, \;\; i = 1, 2, ...$, the size of the jump at x_i being $\mathbb{P}(\{\omega : X(\omega) = x_i\})$. For an arbitrary Borel set A, we have

$$P_X(A) = \sum_{x_i \in A \bigcap \Sigma} \mathbb{P}(\{\omega : X(\omega) = x_i\}).$$

Let P be a probability measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. It is said to be singular (with respect to the Lebesgue measure) if there exist a set $S \in \mathcal{B}(\mathbb{R}^n)$ such that $P(S) = 1$ and the Lebesgue measure of S is zero. On the other hand, P is said to be absolutely continuous (w.r.t. the Lebesgue measure) if for every measurable set A the Lebesgue measure of A equals zero implies $P(A) = 0$. Clearly, if $X_1, X_2, ..., X_n$ are discrete RV's, then P_X is singular. If $X_1, X_2, ..., X_n$ are such that P_X is absolutely continuous, then there exists a non-negative Borel function $p_X(x), \; x \in \mathbb{R}^n$ such that

$$P_X(A) = \int_A p_X(x)dx, \qquad A \in \mathcal{B}(\mathbb{R}^n).$$

The function p_X is called the probability density function for X. In terms of the distribution

$$F_X(x_1, x_2, ...x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} p_X(x_1, ..., x_n)dx_1...dx_n$$

which implies

$$p_X(x_1, ...x_n) \triangleq \frac{\partial^n}{\partial x_1 ... \partial x_n} F_X(x_1, x_2, ..., x_n).$$

Let $(\Omega, \mathcal{F})$ be a measurable space, and $\mathbb{Q}$ and $\mathbb{P}$ are two probability measures defined on $(\Omega, \mathcal{F})$. Then $\mathbb{Q}$ is called absolutely continuous with respect to $\mathbb{P}$ (denoted by $\mathbb{Q} \ll \mathbb{P}$) if $\mathbb{Q}(A) = 0$ whenever $A \in \mathcal{F}$ and $\mathbb{P} = 0$. If $\mathbb{P} \ll \mathbb{Q}$ and $\mathbb{Q} \ll \mathbb{P}$ then $\mathbb{P}$ and $\mathbb{Q}$ are called equivalent probability measures and this is denoted by $\mathbb{P} \sim \mathbb{Q}$.

Theorem 2.2. *(Radon-Nikodym) Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space, and let* Q *be another measure defined also on* $\mathcal{F}$ *such that* $Q \ll P$. *Then there exists an* $\mathcal{F}$*-measurable function* $\phi : \Omega \to [0, \infty]$, *such that* $\phi \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ *and,*

$$Q(B) = \int_B \phi(\omega)dP(\omega), \forall B \in \mathcal{F}.$$

The function ϕ is unique except on a subset of P-measure zero. This function ϕ is often written as $\phi = \frac{dQ}{dP}\Big|_{\mathcal{F}}$ and is called Radon-Nikodym derivative (RND) since it satisfies

$$Q(B) = \int_B dQ = \int_B \phi dP,\ \forall B \in \mathcal{F}.$$

Definition 2.16. *(Regular Conditional Probability Measure) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. A regular conditional probability measure $P(\cdot|\mathcal{G})(\cdot)$ on $(\Omega, \mathcal{F})$ is a function $P(A|\mathcal{G})(\omega),\ A \in \mathcal{F},\ \omega \in \Omega$ having the following properties [12].*

(a) For each $A \in \mathcal{F}$, the function mapping $\omega \in \Omega \longmapsto P(A|\mathcal{G})(\omega)$ is measurable with respect to $\mathcal{G}$.

(b) For each $\omega \in \Omega$, $P(\cdot|\mathcal{G})(\omega)$ is a probability measure on $\mathcal{F}$.

(c) For each $A \in \mathcal{F}$, $P(A|\mathcal{G})(\omega)$ is a version of the conditional probability of A given $\mathcal{G}$. Moreover,

$$P(A \cap B) = \int_B P(A|\mathcal{G})(\omega) P_{\mathcal{G}}(d\omega),\ \forall A \in \mathcal{G}$$

where $P_{\mathcal{G}}$ is the restriction of P to $\mathcal{G}$.

Statements (a) and (c) state that $P(A|\mathcal{G})(\omega)$ is a version of the conditional probability of A given $\mathcal{G}$ (and it is a function of ω). If such a version $P(\cdot|\mathcal{G})(\cdot)$ exists then it is unique in the sense that, if $\bar{P}(\cdot|\mathcal{G})(\cdot)$ is another function with these properties, then there exists a $P_{\mathcal{G}}-$null set N such that $P(A|\mathcal{G})(\omega) = \bar{P}(A|\mathcal{G})(\omega),\ \forall A \in \mathcal{F}$ and $\omega \in N^c$ (e.g., $P(\cdot|\mathcal{G})(\omega) = \bar{P}(\cdot|\mathcal{G})(\omega),\ P_{\mathcal{G}} - a.s.$). Thus, a regular conditional probability measure exists if it can be shown that a version of the conditional probability measure can be chosen to be a probability measure on $\mathcal{F}$ for each $\omega \in \Omega$. Although in general, a regular conditional probability measure may not exist, for the case when $\mathcal{G}$ is generated by a countable partition of Ω, a regular conditional probability measure given $\mathcal{G}$ always exists. Moreover, if (Ω, d) is a metric space which is complete and separable (Polish space), and $\mathcal{F}$ is a Borel $\sigma-$algebra, then for any probability measure P on $(\Omega, \mathcal{F})$ and any sub-σ-algebra $\mathcal{G} \subseteq \mathcal{F}$, a regular conditional probability measure of P given $\mathcal{G}$ always exists.

Lemma 2.1. *(Absolute Continuity of Probability Measures)*

a) Suppose $Q_{\mathcal{G}} \ll P_{\mathcal{G}}$. If $Q(\cdot|\mathcal{G})(\omega) \ll P(\cdot|\mathcal{G})(\omega),\ Q_{\mathcal{G}} - a.s.$, then $Q \ll P$.

b) Conversely, if $Q \ll P$, then $Q(\cdot|\mathcal{G})(\omega) \ll P(\cdot|\mathcal{G})(\omega),\ P(\cdot|\mathcal{G})(\omega) - a.s.$

Note that if $Y : (\Omega, \mathcal{F}) \longmapsto (\mathcal{Y}, \mathcal{A})$ is a RV on $(\Omega, \mathcal{F})$ into a measurable space $(\mathcal{Y}, \mathcal{A})$ and $\mathcal{Y}$ is a Polish space, then a regular conditional distribution for Y given the sub-σ-algebra $\mathcal{G}$ of $\mathcal{F}$ denoted by $P(dy|\mathcal{G})(\omega)$ is defined according to Definition 2.16, and this always exists. Additionally, if $X : (\Omega, \mathcal{F}) \longmapsto (\mathcal{X}, \mathcal{B})$ is a RV on $(\Omega, \mathcal{F})$ into a measurable space $(\mathcal{X}, \mathcal{B})$, and $\mathcal{G}$ is the sub-σ-algebra of $\mathcal{F}$ generated by X, then $P(dy|X)(\omega)$ is called the *regular conditional distribution* of Y given X. One can go one step further to define a regular conditional distribution for Y given $X = x$ as a quantity $P(dy|X = x)$, and introduce an equivalent definition called stochastic kernel.

Definition 2.17. *(Stochastic Kernel) Given a measurable space $(\Omega, \mathcal{F})$ on which the RVs, X and Y are defined, via $X : (\Omega, \mathcal{F}) \to (\mathcal{X}, \Sigma_{\mathcal{X}})$, and $Y : (\Omega, \mathcal{F}) \to (\mathcal{Y}, \Sigma_{\mathcal{Y}})$, respectively, then the relation between the RV X and the RV Y is defined via a probabilistic mapping. The mapping $\mu : \Sigma_{\mathcal{Y}} \times \mathcal{X} \to [0, 1]$ satisfies the following two conditions.*

(i) For every $x \in \mathcal{X}$, the set function $\mu(\cdot|x)$ is a probability measure on $\Sigma_{\mathcal{Y}}$ (possibly finite additive);

(ii) For every $F \in \Sigma_{\mathcal{Y}}$, the function $\mu(F|\cdot)$ is $\mathcal{X}$-measurable.

The mapping $\mu(\cdot;\cdot)$ is called a stochastic kernel or transition probability. The set of all such stochastic kernels is denoted by $\mathcal{Q}(\mathcal{Y};\mathcal{X})$.

Definition 2.18. *Let $(\Omega, \mathcal{F})$ denote a measurable space and μ a positive measure on Ω. Let $f : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be a measurable function. Define*

$$\|f\|_p \triangleq \left\{ \int_\Omega |f|^p(\omega) d\mu(\omega) \right\}^p, \quad 1 \leq p < \infty.$$

$L_p(\Omega, \mathcal{F}, \mu)$ is the set of all measurable functions f on $(\Omega, \mathcal{F})$ for which $\|f\|_p < \infty$ (often denoted by $L_p(\Omega, \mathcal{F}, \mu) \equiv L_p(\mu)$), and $\|f\|_p$ denotes the L_p-norm of f.

2.1.5. Signed Measures

This section summarizes results for signed measures that are particularly relevant to the characterization of the extremum measures. Additional details and the proofs of the following theorems can be found in [31].

Definition 2.19. *Let $(\Sigma, \mathcal{B}(\Sigma))$ be a measurable space. A set function $\xi : \mathcal{B}(\Sigma) \mapsto \overline{\mathbb{R}} \triangleq \mathbb{R} \cup \{-\infty, \infty\}$ is called a signed measure on $(\Sigma, \mathcal{B}(\Sigma))$ if*

(i) $\xi(\emptyset) = 0$.

(ii) $\xi(A)$ *assumes at most one of the values* $\pm\infty$ *for all* $A \in \mathcal{B}(\Sigma)$.

(iii) $\xi(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \xi(A_n)$ *for all pairwise disjoint sets* $A_1, A_2, \ldots \in \mathcal{B}(\Sigma)$.

A signed measure ξ on $(\Sigma, \mathcal{B}(\Sigma))$ is called a non-negative signed measure,[1] if and only if ξ is a measure. If μ_1, μ_2 are two measures on $(\Sigma, \mathcal{B}(\Sigma))$, at least one of which is finite, then $\xi = \mu_1 - \mu_2$ is well defined and it is a signed measure on $(\Sigma, \mathcal{B}(\Sigma))$. Note that a signed measure is not, in general, monotone[2].

The next theorem describes sequentially continuity for signed measures.

Theorem 2.3. *Let* ξ *be a signed measure on* $(\Sigma, \mathcal{B}(\Sigma))$.

(i) Let $A, B \in \mathcal{B}(\Sigma)$ *and* $A \subseteq B$. *If* $|\xi(B)| < +\infty$, *then*

$$|\xi(A)| < +\infty.$$

(ii) If $A_1, A_2, \ldots \in \mathcal{B}(\Sigma)$ *and* $A_n \subseteq A_{n+1}$ *for all* n, *then*

$$\xi(\cup_{n=1}^{\infty} A_n) = \lim_{n \to +\infty} \xi(A_n).$$

(iii) If $A_1, A_2, \ldots \in \mathcal{B}(\Sigma)$ *and* $A_n \supseteq A_{n+1}$ *for all* n, *then*

$$\xi(\cap_{n=1}^{\infty} A_n) = \lim_{n \to +\infty} \xi(A_n).$$

The concepts of positive and negative sets are introduced next, since these are important in representing signed measures via its Jordan decomposition.

Definition 2.20. *Let* ξ *be a signed measure on* $(\Sigma, \mathcal{B}(\Sigma))$. *A set* $P \in \mathcal{B}(\Sigma)$ *is called positive for* ξ *if* $\xi(A) \geq 0$ *for every* $A \in \mathcal{B}(\Sigma)$ *and* $A \subseteq P$. *A set* $N \in \mathcal{B}(\Sigma)$ *is called negative for* ξ *if* $\xi(A) \leq 0$ *for every* $A \in \mathcal{B}(\Sigma)$ *and* $A \subseteq N$. *A set which is both positive and negative for* ξ *is a null set for* ξ.

The next theorem known as Hahn decomposition establishes existence of positive and negative sets partitioning the space Σ.

Theorem 2.4. *(Hahn Decomposition Theorem) If* ξ *is a signed measure on* $(\Sigma, \mathcal{B}(\Sigma))$ *then there exist a positive set* $P \in \mathcal{B}(\Sigma)$ *and a negative set* $N \in \mathcal{B}(\Sigma)$ *for* ξ *so that* $P \cup N = \Sigma$ *and* $P \cap N = \emptyset$.

Note that, the sets P and N are not unique.

[1] If $\xi(A) \geq 0$ for every $A \in \mathcal{B}(\Sigma)$

[2] If $A, B \in \mathcal{B}(\Sigma)$ and $A \subseteq B$, then $\xi(B) = \xi(A) + \xi(B \setminus A) \leq \xi(A)$ whenever $\xi(B \setminus A) \leq 0$.

Definition 2.21. *Let ξ_1, ξ_2 be two signed measures on $(\Sigma, \mathcal{B}(\Sigma))$. We say that they are mutually singular or that ξ_1 is singular with respect to ξ_2 or vice versa if there exists $A_1, A_2 \in \mathcal{B}(\Sigma)$ so that $A_1 \cup A_2 = \Sigma$, $A_1 \cap A_2 = \emptyset$ such that A_1 is null for ξ_2 and A_2 is null for ξ_1. We use the symbol $\xi_1 \perp \xi_2$ to denote that ξ_1, ξ_2 are mutually singular.*

The next theorem is the Jordan decomposition of a signed measure into its positive and negative parts which are mutually singular.

Theorem 2.5. *(Jordan Decomposition Theorem) Let ξ be a signed measure on $(\Sigma, \mathcal{B}(\Sigma))$. There exist two non-negative measures ξ^+, ξ^-, at least one of which is finite, such that*

$$\xi = \xi^+ - \xi^-, \qquad \xi^+ \perp \xi^-.$$

Consider any Hahn decomposition of Σ for ξ. For every $A \in \mathcal{B}(\Sigma)$, define the set functions $\xi^+, \xi^- : \mathcal{B}(\Sigma) \longmapsto [0, +\infty]$ by

$$\xi^+(A) = \xi(A \cap P), \qquad \xi^-(A) = -\xi(A \cap N).$$

We say that the non-negative signed measures ξ^+, ξ^- constitute the Jordan decomposition of ξ. ξ^+ is called the positive variation of ξ and ξ^- the negative variation of ξ. The measure $|\xi| = \xi^+ + \xi^-$ is called the absolute variation of ξ, while $|\xi|(\Sigma)$ is called the total variation of ξ and is equal to

$$|\xi|(\Sigma) = \xi^+(\Sigma) + \xi^-(\Sigma) = \xi(P) - \xi(N)$$

where the sets P, N constitute a Hahn decomposition of Σ for ξ. Hence, the total variation of ξ is equal to the difference between the largest and the smallest values of ξ.

Theorem 2.6. *Let ξ be a signed measure on $(\Sigma, \mathcal{B}(\Sigma))$ and $A \in \mathcal{B}(\Sigma)$. If $A_1, \ldots, A_n \in \mathcal{B}(\Sigma)$ are pairwise disjoint and $A = \bigcup_{k=1}^n A_k$, then for every $A \in \mathcal{B}(\Sigma)$,*

$$\xi^+(A) = \sup\{\sum_{k=1}^n \xi^+(A_k) : n \in \mathbb{N}\}$$
$$\xi^-(A) = \sup\{\sum_{k=1}^n \xi^-(A_k) : n \in \mathbb{N}\}$$
$$|\xi|(A) = \sup\{\sum_{k=1}^n |\xi(A_k)| : n \in \mathbb{N}\}.$$

where $\sup$ *is over all measurable partitions of the set Σ.*

Definition 2.22. *(Total Variation of a Signed Measure) Let ξ be a signed measure of the measurable space $(\Sigma, \mathcal{B}(\Sigma))$. The total variation norm of ξ is*

$$\|\xi\|_{TV} = \xi^+(\Sigma) + \xi^-(\Sigma)$$

where (ξ^+, ξ^-) is the Hahn-Jordan decomposition of ξ.

If Σ is finite or countable and ξ is a signed measure, then

$$\|\xi\|_{TV} = \sum_{x \in \Sigma} |\xi(x)|.$$

If ξ has a density f with respect to λ (a measure on $(\Sigma, \mathcal{B}(\Sigma))$), then

$$\|\xi\|_{TV} = \int |f(x)|\lambda(dx), \quad f(x) \triangleq \frac{d\xi(\cdot)}{d\lambda(\cdot)}(x)$$

2.2. Total Variation Distance Metric

In this section, we introduce the total variation distance metric and we review some of its properties.

Definition 2.23. *(Total Variation Distance) Let ξ_1, ξ_2 be two measures on the measurable space $(\Sigma, \mathcal{B}(\Sigma))$. The total variation distance between ξ_1 and ξ_2 is the total variation norm of a signed measure $\xi_1 - \xi_2$.*

Let

$$\mathcal{M}_{sm}(\Sigma) = \text{set of finite signed measures on } \mathcal{B}(\Sigma)$$
$$\mathcal{M}_1(\Sigma) = \text{set of probability measures on } \mathcal{B}(\Sigma)$$
$$\mathcal{M}_0(\Sigma) = \text{set of finite signed measures on } (\Sigma, \mathcal{B}(\Sigma)) \text{ satisfying } \xi(\Sigma) = 0.$$

Let $BM(\Sigma)$ denote the space of measurable real valued functions, and $\|f\|_\infty = \sup\{|f|(x) : x \in \Sigma\}$, and $BM^+(\Sigma) \triangleq \{f \in BM(\Sigma) : f \geq 0\}$. Note that, $BM(\Sigma)$ endowed with the sup norm is a Banach space. For any $\xi \in \mathcal{M}(\Sigma)$ and $f \in BM(\Sigma)$ define

$$\xi(f) = \int f d\xi$$

Any signed measure $\xi \in \mathcal{M}_{sm}(\Sigma)$ defines a linear function on the Banach space $(BM(\Sigma), \|\cdot\|_\infty)$.

Lemma 2.2.

(i) For any $\xi \in \mathcal{M}_{sm}(\Sigma)$ and $f \in BM(\Sigma)$

$$\left| \int f d\xi \right| \leq \|\xi\|_{TV} \|f\|_\infty.$$

(ii) For any $\xi \in \mathcal{M}_{sm}(\Sigma)$

$$\|\xi\|_{TV} = \sup\{\xi(f) : f \in BM(\Sigma), \|f\|_\infty = 1\}.$$

(iii) For any $f \in BM(\Sigma)$

$$\|f\|_\infty = \sup\{\xi(f) : \xi \in \mathcal{M}(\Sigma), \|\xi\|_{TV} = 1\}.$$

Proof. See Appendix A.1. ■

Let $\xi \in \mathcal{M}_0(\Sigma)$, $f \in BM(\Sigma)$. Since $\xi(\Sigma) = 0$ for any $c \in \mathbb{R}$ then

$$\xi(f) = \int_\Sigma f(x)d\xi(x) = \int_\Sigma (f(x) - c)d\xi(x)$$

and, hence

$$|\xi(f)| \leq \|\xi\|_{TV}\|f - c\|_\infty$$
$$|\xi(f)| \leq \|\xi\|_{TV} \inf_{c\in\mathbb{R}} \|f - c\|_\infty.$$

For $f \in BM(\Sigma)$, $\inf_{c\in\mathbb{R}} \|f - c\|_\infty$ is related to the oscillation semi-norm of f, called the global modulus of continuity, by

$$\text{osc}(f) \triangleq \sup_{(x_1,x_2)\in\Sigma\times\Sigma} |f(x_1) - f(x_2)| = 2\inf_{c\in\mathbb{R}} \|f - c\|_\infty. \tag{2.3}$$

For $f \in BM^+(\Sigma)$

$$\text{osc}(f) = \sup_{x\in\Sigma} |f(x)| - \inf_{x\in\Sigma} |f(x)|.$$

Lemma 2.3. *For any $\xi \in \mathcal{M}(\Sigma)$ and $f \in BM(\Sigma)$*

$$|\xi(f)| \leq \sup_{(x_1,x_2)\in\Sigma\times\Sigma} |\xi^+(\Sigma)f(x_1) - \xi^-(\Sigma)f(x_2)|$$

where (ξ^+, ξ^-) is the Hahn-Jordan decomposition of ξ. In particular, for any $\xi \in \mathcal{M}_0(\Sigma)$ and $f \in BM(\Sigma)$

$$|\xi(f)| \leq \frac{1}{2}\|\xi\|_{TV} 2\inf_{c\in\mathbb{R}} \|f - c\|_\infty = \frac{1}{2}\|\xi\|_{TV} osc(f).$$

Proof. See Appendix A.2. ■

Hence, for $\xi \in \mathcal{M}_0(\Sigma)$, $\|\xi\|_{TV}$ is the operator norm of ξ considered as an operator over the space $BM(\Sigma)$ equipped with oscillation semi-norm. As an application take $\xi_1, \xi_2 \in \mathcal{M}_1(\Sigma)$ and define $\xi_1 - \xi_2 \in \mathcal{M}_0(\Sigma)$. Then, for any $f \in BM(\Sigma)$

$$|\xi_1(f) - \xi_2(f)| \leq \frac{1}{2}\|\xi_1 - \xi_2\|_{TV}\text{osc}(f)$$

This is tighter than $|\xi_1(f) - \xi_2(f)| \leq \|\xi_1 - \xi_2\|_{TV}\|f\|_\infty$, since $\text{osc}(f) \leq 2\|f\|_\infty$.

Lemma 2.4. *For any $\xi_1, \xi_2 \in \mathcal{M}_1(\Sigma)$*

$$\frac{1}{2}\|\xi_1 - \xi_2\|_{TV} = \sup_A |\xi_1(A) - \xi_2(A)| \tag{2.4}$$

where the supremum is taken over all measurable subsets of Σ.

Proof. See Appendix A.3. ■

2.3. Relation of Total Variation Distance to Other Metrics

In this section, we briefly state relations of the total variation distance to other distance metrics, and we discuss some of its applications.

L_1 Distance.

Given $(\Sigma, \mathcal{B}(\Sigma))$, let $\sigma \in \mathcal{M}_1(\Sigma)$ be a fixed measure (as well as $\mu \in \mathcal{M}_1(\Sigma)$). Define the Radon-Nykodym derivatives $\psi \triangleq \frac{d\mu}{d\sigma}$, $\varphi \triangleq \frac{d\nu}{d\sigma}$, i.e., densities with respect to a fixed $\sigma \in \mathcal{M}_1(\Sigma)$. Then,

$$\|\nu - \mu\|_{TV} = \int |\varphi(x) - \psi(x)|\sigma(dx).$$

This can be used to model uncertainty as follows. Define

$$\mathbf{B}_R(\mu) \triangleq \{\nu \in \mathcal{M}_1(\Sigma) : \|\nu - \mu\|_{TV} \leq R\}$$

and, consider a subset of $\mathbf{B}_R(\mu)$ defined by

$$\mathbf{B}_{R,\sigma}(\mu) \triangleq \{\nu \in \mathbf{B}_R(\mu) : \nu << \sigma, \mu << \sigma\} \subseteq \mathbf{B}_R(\mu).$$

Then,

$$\mathbf{B}_{R,\sigma}(\mu) = \left\{\varphi \in L_1(\sigma), \varphi \geq 0, \sigma - a.s. : \int_\Sigma |\varphi(x) - \psi(x)|\sigma(dx) \leq R\right\}.$$

Thus, under the absolute continuity of measures the total variation distance reduces to L_1 distance. Robustness via L_1 distance uncertainty on the space of spectral densities is investigated in the context of Wiener-Kolmogorov theory in an estimation and decision framework in [47, 57].

Relative Entropy.

The relative entropy of $\nu \in \mathcal{M}_1(\Sigma)$ with respect to $\mu \in \mathcal{M}_1(\Sigma)$ is a mapping $H(\cdot|\cdot) : \mathcal{M}_1(\Sigma) \times \mathcal{M}_1(\Sigma) \longmapsto [0, \infty]$ defined by [23]

$$H(\nu|\mu) \triangleq \begin{cases} \int_\Sigma \log(\frac{d\nu}{d\mu})d\nu, & \text{if } \nu << \mu \text{ and } \log(\frac{d\nu}{d\mu}) \in L_1(\nu) \\ +\infty, & \text{otherwise.} \end{cases}$$

It is well known that $H(\nu|\mu) \geq 0, \forall \nu, \mu \in \mathcal{M}_1(\Sigma)$, while $H(\nu|\mu) = 0 \Leftrightarrow \nu = \mu$. Total variation distance is bounded above by relative entropy via Pinsker's inequality [45], giving

$$\|\nu - \mu\|_{TV} \leq \sqrt{2H(\nu|\mu)}, \quad \nu, \mu \in \mathcal{M}_1(\Sigma). \tag{2.5}$$

This can be used to model uncertainty as follows. Given a known or nominal probability measure $\mu \in \mathcal{M}_1(\Sigma)$ the uncertainty set based on relative entropy is defined by

$$A_{\tilde{R}}(\mu) \triangleq \left\{\nu \in \mathcal{M}_1(\Sigma) : H(\nu|\mu) \leq \tilde{R}\right\}, \quad \tilde{R} \in [0, \infty).$$

Clearly, the uncertainty set determined by the total variation distance d_{TV}, is larger than that determined by the relative entropy. In other words, in view of Pinsker's inequality (2.5), for any $r \in [0, \infty)$

$$\left\{\nu \in \mathcal{M}_1(\Sigma), \nu << \mu : H(\nu|\mu) \leq \frac{r^2}{2}\right\} \subseteq \mathbf{B}_R(\mu) \equiv \left\{\nu \in \mathcal{M}_1(\Sigma) : \|\nu - \mu\|_{TV} \leq r\right\}.$$

Hence, even for those measures which satisfy $\nu << \mu$, the uncertainty set described by relative entropy is a subset of the much larger total variation distance uncertainty set. Moreover, by Pinsker's inequality, distance in total variation of probability measures is a lower bound on their relative entropy or Kullback-Leibler distance.

Over the last few years, relative entropy uncertainty model has received particular attention due to various properties (convexity, compact level sets), its simplicity and its connection to risk sensitive pay-off, minimax games, and large deviations [2, 15, 44, 48, 56]. Recently, an uncertainty model along the spirit of Radon-Nikodym derivative is employed in [42] for portfolio optimization under uncertainty. Unfortunately, relative entropy uncertainty modeling has two disadvantages: 1) it does not define a true metric on the space of measures; 2) relative entropy between two measures is not defined if the measures are not absolutely continuous. The latter rules out the possibility of measures $\nu \in \mathcal{M}_1(\Sigma)$ and $\mu \in \mathcal{M}_1(\tilde{\Sigma})$, $\tilde{\Sigma} \subset \Sigma$ to be defined on different spaces[3]. It is one of the main disadvantages in employing relative entropy in the context of uncertainty modelling for stochastic controlled diffusions (or SDE's) [46]. Specifically, by invoking a change of measure it can be shown that relative entropy modelling allows uncertainty in the drift coefficient of stochastic controlled diffusions, but not in the diffusion coefficient, because the latter kind of uncertainty leads to measures which are not absolutely continuous with respect to the nominal measure [48].

Kakutani-Hellinger Distance.

Another measure of distance of two probability measures which relates to their distance in variation is the Kakutani-Hellinger distance [30]. Consider $\nu \in \mathcal{M}_1(\Sigma)$, $\mu \in \mathcal{M}_1(\Sigma)$ and a fixed measure $\sigma \in \mathcal{M}_1(\Sigma)$ such that $\nu << \sigma$, $\mu << \sigma$ and define $\varphi \triangleq \frac{d\nu}{d\sigma}$, $\psi \triangleq \frac{d\mu}{d\sigma}$. The

[3]This corresponds to the case in which the nominal system is a simplified version of the true system and is defined on a lower dimension space.

Kakutani-Hellinger distance is a mapping $d_{KH} : L_1(\sigma) \times L_1(\sigma) \mapsto [0, \infty)$ defined by

$$d^2_{KH}(\nu, \mu) \triangleq \frac{1}{2} \int \left(\sqrt{\varphi(x)} - \sqrt{\psi(x)} \right)^2 d\sigma(x). \tag{2.6}$$

Indeed, the function d_{KH} given by (2.6) is a metric on the set of probability measures. A related quantity is the Hellinger integral of measures $\nu \in \mathcal{M}_1(\Sigma)$ and $\mu \in \mathcal{M}_1(\Sigma)$ defined by

$$H(\nu, \mu) \triangleq \int \sqrt{\varphi(x)\psi(x)} d\sigma(x), \tag{2.7}$$

which is related to the Kakutani-Hellinger distance via $d^2_{KH}(\nu, \mu) = 1 - H(\nu, \mu)$. The relations between distance in variation and Kakutani-Hellinger distance (and Hellinger integral) are given by the following inequalities:

$$2\Big(1 - H(\nu, \mu)\Big) \leq \|\nu - \mu\|_{TV} \leq \sqrt{8\Big(1 - H(\nu, \mu)\Big)}, \tag{2.8}$$

$$\|\nu - \mu\|_{TV} \leq 2\sqrt{1 - H^2(\nu, \mu)}, \tag{2.9}$$

$$2d^2_{KH}(\nu, \mu) \leq \|\nu - \mu\|_{TV} \leq \sqrt{8} d_{KH}(\nu, \mu). \tag{2.10}$$

The above inequalities imply that these distances define the same topology on the space of probability measure on $(\Sigma, \mathcal{B}(\Sigma))$. Specifically, convergence in total variation of probability measures defined on a metric space $(\Sigma, \mathcal{B}(\Sigma), d)$, implies their weak convergence with respect to the Kakutani-Hellinger distance metric, [30]. In [27], the Hellinger distance on the space of spectral densities is used to define a pay-off subject to constraints in the context of approximation theory.

Levy-Prohorov Distance.

Given a metric space $(\Sigma, \mathcal{B}(\Sigma), d)$, and a family of probability measures $\mathcal{M}_1(\Sigma)$ on $(\Sigma, \mathcal{M}_1(\Sigma))$ it is possible to "metrize" weak convergence of probability measure, denoted by $P_n \overset{w}{\to} P$, where $\{P_n : n \in \mathbb{N}\} \subset \mathcal{M}_1(\Sigma)$, $P \in \mathcal{M}_1(\Sigma)$ via the so called Levy-Prohorov metric denoted by $d_{LP}(\nu, \mu)$ [23]. Thus, this metric is also a candidate for a measure of proximity between two probability measures. The Levi-Prohorov metric is related to distance in variation via the upper bound [30],

$$d_{LP}(\nu, \mu) \leq \min\left\{||\nu - \mu||_{TV}, 1\right\}, \quad \forall \ \ \nu \in \mathcal{M}_1(\Sigma), \mu \in \mathcal{M}_1(\Sigma).$$

The function defined by $L(\nu, \mu) = \max\left\{d_{LP}(\nu, \mu), d_{LP}(\mu, \nu)\right\}$, is actually a distance metric (it satisfies the properties of distance).

In view of the relations between different metrics, such as relative entropy, Levy-Prohorov metric, Kakutani-Hellinger metric, etc, it is clear that the extremum problems under investigation will give sub-optimal solution to the same extremum problems with distance in variation replaced by these metrics. An anthology of other distances and distance metrics related to total variation distance can be found in [23].

3

Extremum Problems

In this chapter, we investigate extremum problems with pay-off being the total variation distance metric defined on the space of probability measures, subject to linear functional constraints on the space of probability measures, and vice-versa; that is, with the roles of total variation metric and linear functional interchanged. Utilizing concepts from signed measures, the extremum probability measures of such problems are obtained in closed form, by identifying the partition of the support set and the mass of these extremum measures on the partition. Throughout the derivations we make extensive use of lower and upper bounds of pay-offs which are achievable. The results are derived for abstract spaces; specifically, complete separable metric spaces known as Polish spaces, while the high level ideas are also discussed for denumerable spaces endowed with the discrete topology. The results of this part include:

- characterization of the properties of the extremum problems under investigation;
- characterization of extremum measures on abstract spaces, and closed form solutions of the extremum measures for finite alphabet spaces;
- convexity and concavity properties of the extremum solutions;
- simulations which illustrate the different scenarios of the extremum solution of the various problems, and an application to the area of information theory.

3.1. Problem Formulation

In this section, we will introduce the extremum problems we investigate in this chapter. Let (Σ, d_Σ) denote a complete, separable metric space and $(\Sigma, \mathcal{B}(\Sigma))$ the corresponding measurable space, where $\mathcal{B}(\Sigma)$ is the σ-algebra generated by open sets in Σ. Let $\mathcal{M}_1(\Sigma)$ denote the set of probability measures on $\mathcal{B}(\Sigma)$. The total variation distance[1] is a metric [22] $d_{TV} : \mathcal{M}_1(\Sigma) \times \mathcal{M}_1(\Sigma) \to [0, \infty)$ defined by

$$d_{TV}(\alpha, \beta) \equiv \|\alpha - \beta\|_{TV} \triangleq \sup_{P \in \mathcal{P}(\Sigma)} \sum_{F_i \in P} |\alpha(F_i) - \beta(F_i)| \tag{3.1}$$

where $\alpha, \beta \in \mathcal{M}_1(\Sigma)$ and $\mathcal{P}(\Sigma)$ denotes the collection of all finite partitions of Σ. With respect to this metric, $(\mathcal{M}_1(\Sigma), d_{TV})$ is a complete metric space. The total variation distance is a true metric, hence it is a measure of difference between two distributions, $\nu, \mu \in \mathcal{M}_1(\Sigma)$. By the properties of the distance metric then $\|\nu - \mu\|_{TV} \le \|\nu\|_{TV} + \|\mu\|_{TV} = 2$, hence $\|\cdot\|_{TV}$ is further restricted to the interval $[0, 2]$. The two extreme cases are $\|\cdot\|_{TV} = 0$ implying $\nu = \mu$, and $\|\cdot\|_{TV} = 2$ implying that the support sets of ν and μ denoted by supp(ν) and supp(μ), respectively, are non-overlapping, that is, supp$(\nu) \cap$ supp$(\mu) = \emptyset$. In minimax problems one can introduce an uncertainty set based on distance in variation as follows. Suppose the probability measure $\nu \in \mathcal{M}_1(\Sigma)$ is unknown, while modeling techniques give access to a nominal probability measure $\mu \in \mathcal{M}_1(\Sigma)$. Having constructed the nominal probability measure, one may construct from empirical data, the distance of the two measures with respect to the total variation distance $\|\nu - \mu\|_{TV}$. This will provide an estimate of the radius R, such that $\|\nu - \mu\|_{TV} \le R$, and hence characterize the set of all possible true measures $\nu \in \mathcal{M}_1(\Sigma)$, centered at the nominal distribution $\mu \in \mathcal{M}_1(\Sigma)$, and lying within the ball of radius R, with respect to the total variation distance $\|\cdot\|_{TV}$. Such a procedure is used in information theory to define strong typicality of sequences. Unlike other distances used in the past such as relative entropy [2,15,44,48,56], quantifying uncertainty via the metric $\|\cdot\|_{TV}$ does not require absolute continuity of measures[2], i.e., singular measures are admissible, and hence ν and μ need not be defined on the same space. Thus, the support set of μ may be $\tilde{\Sigma} \subset \Sigma$, hence $\mu(\Sigma \setminus \tilde{\Sigma}) = 0$ but $\nu(\Sigma \setminus \tilde{\Sigma}) \neq 0$ is allowed. For measures induced by stochastic differential equations (SDE's), variation distance uncertainty set models situations in which both the drift and diffusion coefficient of SDE's are unknown.

[1]The definition of total variation distance can be extended to signed measures (see Chapter 2).

[2]$\nu \in \mathcal{M}_1(\Sigma)$ is absolutely continuous with respect to $\mu \in \mathcal{M}_1(\Sigma)$, denoted by $\nu << \mu$, if $\mu(A) = 0$ for some $A \in \mathcal{B}(\Sigma)$ then $\nu(A) = 0$.

Define the spaces

$$\begin{aligned}
BC(\Sigma) &\triangleq \Big\{\text{Bounded continuous functions } \ell:\Sigma\longrightarrow\mathbb{R}: \quad ||\ell|| \triangleq \sup_{x\in\Sigma}|\ell(x)|<\infty\Big\}\\
BC^+(\Sigma) &\triangleq \{\ell\in BC(\Sigma):\ell\geq 0\}\\
BM(\Sigma) &\triangleq \Big\{\text{Bounded measurable functions } \ell:\Sigma\longrightarrow\mathbb{R}: \quad ||\ell||<\infty\Big\}\\
BM^+(\Sigma) &\triangleq \{\ell\in BM(\Sigma):\ell\geq 0\}\\
C(\Sigma) &\triangleq \Big\{\text{Continuous functions } \ell:\Sigma\longrightarrow\mathbb{R}: \quad ||\ell||<\infty\Big\}\\
C^+(\Sigma) &\triangleq \{\ell\in C(\Sigma):\ell\geq 0\}.
\end{aligned}$$

Note that, $BC(\Sigma)$, $BM(\Sigma)$ and $C(\Sigma)$ endowed with the sup norm $\|\ell\| \triangleq \sup_{x\in\Sigma}|\ell(x)|$, are Banach spaces [22]. We derive the maximizing measure for $\ell\in BC^+(\Sigma)$ or $BM^+(\Sigma)$. However, the results can be generalized to real-valued functions $\ell\in L^{\infty,+}(\Sigma,\mathcal{B}(\Sigma),\nu)$, the set of all $\mathcal{B}(\Sigma)$-measurable, non-negative essentially bounded functions defined $\nu-a.e.$ endowed with the essential supremum norm $||\ell||_{\infty,\nu} = \nu\text{-ess}\sup_{x\in\Sigma}\ell(x) \triangleq \inf_{\Delta\in\mathcal{N}_\eta}\sup_{x\in\Delta^c}\|\ell(x)\|$, where $\mathcal{N}_\nu=\{A\in\mathcal{B}(\Sigma):\nu(A)=0\}$.

Before we proceed with the formulation of the extremum problems, we introduce first two main definitions.

Definition 3.1. *Given a fixed nominal distribution $\mu\in\mathcal{M}_1(\Sigma)$ and a parameter $R\in[0,2]$, define the class of true distributions by*

$$\mathbf{B}_R(\mu) \triangleq \Big\{\nu\in\mathcal{M}_1(\Sigma):\|\nu-\mu\|_{TV}\leq R\Big\} \tag{3.2}$$

and the average pay-off with respect to any $\nu\in\mathbf{B}_R(\mu)$ by

$$\mathbb{L}_1(\nu) \triangleq \int_\Sigma \ell(x)\nu(dx), \qquad \ell\in BC^+(\Sigma) \;\; or \;\; BM^+(\Sigma). \tag{3.3}$$

Definition 3.2. *Given a fixed nominal distribution $\mu\in\mathcal{M}_1(\Sigma)$ and a parameter $D\in[0,\infty)$, define the class of true distributions by*

$$\boldsymbol{Q}(D) \triangleq \Big\{\nu\in\mathcal{M}_1(\Sigma):\int_\Sigma \ell(x)\nu(dx)\leq D\Big\}, \qquad \ell\in BC^+(\Sigma) \;\; or \;\; BM^+(\Sigma) \tag{3.4}$$

and the total variation pay-off with respect to the true probability measure $\nu\in\boldsymbol{Q}(D)$ by

$$\mathbb{L}_2(\nu) \triangleq \|\nu-\mu\|_{TV}. \tag{3.5}$$

3.1.1. Maximization Problems

In this section we introduce the maximization problems under investigation[3].

[3] In all these optimization problems we assume that the solution exists.

Problem 3.1. *Find the solution of the extremum problem*

$$D^+(R) \triangleq \sup_{\nu \in \mathbf{B}_R(\mu)} \mathbb{L}_1(\nu) = \sup_{\nu \in \mathcal{M}_1(\Sigma): \|\nu - \mu\|_{TV} \leq R} \int_\Sigma \ell(x)\nu(dx), \quad \forall R \in [0, 2]. \tag{3.6}$$

Problem 3.1 is a convex optimization problem on the space of probability measures. In the context of minimax theory, Problem 3.1 is important in minimax stochastic control, estimation, and decision. Such formulations are found in [2, 15, 44, 48, 56] utilizing relative entropy to describe a class of models, and in [47, 57] utilizing L_1 distance to describe a class of power spectral densities.

Problem 3.2. *Find the solution of the extremum problem*

$$R^+(D) \triangleq \sup_{\nu \in \boldsymbol{Q}(D)} \mathbb{L}_2(\nu) = \sup_{\nu \in \mathcal{M}_1(\Sigma): \int_\Sigma \ell(x)\nu(dx) \leq D} \|\nu - \mu\|_{TV}, \quad \forall D \in [0, \infty). \tag{3.7}$$

Problem 3.2, i.e., $R^+(D)$, is the inverse mapping of Problem 3.1, i.e., $D^+(R)$, and hence the solution of $D^+(R)$ gives the solution of $R^+(D)$. $D^+(R)$ is investigated in [50] in the context of minimax stochastic control, following an alternative approach which utilizes large deviation theory to express the extremum measure by a convex combination of a tilted and the nominal probability measures. The two disadvantages of the method pursued in [2, 15, 44, 56] are the following. 1) No explicit closed form expression for the extremum measure is given, and as a consequence, 2) its application to dynamic programming is restricted to a class of uncertain probability measures which are absolutely continuous with respect to the nominal measure $\mu(\Sigma) \in \mathcal{M}_1(\Sigma)$.

3.1.2. Minimization Problems

In this section we introduce the minimization problems under investigation.

Problem 3.3. *Find the solution of the extremum problem*

$$D^-(R) \triangleq \inf_{\nu \in \mathbf{B}_R(\mu)} \mathbb{L}_1(\nu) = \inf_{\nu \in \mathcal{M}_1(\Sigma): \|\nu - \mu\|_{TV} \leq R} \int_\Sigma \ell(x)\nu(dx), \quad \forall R \in [0, \infty). \tag{3.8}$$

Problem 3.3 is important in approximating a class of probability distributions or spectral measures by reduced ones. In fact, the solution of (3.8) is obtained precisely as that of Problem 3.1, with a reverse computation of the partition of the space Σ and the mass of the extremum measure on the partition moving in the opposite direction.

Problem 3.4. *Find the solution of the extremum problem*

$$R^-(D) \triangleq \inf_{\nu \in \boldsymbol{Q}(D)} \mathbb{L}_2(\nu) = \inf_{\nu \in \mathcal{M}_1(\Sigma): \int_\Sigma \ell(x)\nu(dx) \leq D} \|\nu - \mu\|_{TV}, \quad \forall D \in [0, \infty) \tag{3.9}$$

whenever $\int_\Sigma \ell(x)\mu(dx) > D$. *If* $\int_\Sigma \ell(x)\mu(dx) \leq D$ *then* $\nu^* = \mu$ *is the trivial extremum measure of (3.9).*

Problem 3.4 is important in the context of approximation theory, since distance in variation is a measure of proximity of two probability distributions subject to constraints. It is also important in spectral measure or density approximation as follows. Recall that a function $\{R(\tau)\colon -\infty \leq \tau \leq \infty\}$ is the covariance function of a quadratic mean continuous and wide-sense stationary process if and only if it is of the form [60]

$$R(\tau) = \int_{-\infty}^{\infty} e^{2\pi\nu\tau} F(d\nu),$$

where $F(\cdot)$ is a finite Borel measure on $\mathbb{R}$, called spectral measure. Thus, by proper normalization of $F(\cdot)$ via $F_N(d\nu) \triangleq \frac{1}{R(0)}F(d\nu)$, then $F_N(d\nu)$ is a probability measure on $\mathcal{B}(\mathbb{R})$, and hence Problem 3.2 can be used to approximate the class of spectral measures with moment estimates belonging to the class described by inequality constraints. Spectral estimation problems are discussed extensively in [26–29, 43], utilizing relative entropy and Hellinger distances, under moment estimates involving equality constraints. However, in these references, the approximated spectral density is absolutely continuous with respect to the nominal spectral density; hence, it can not deal with reduced order approximation. In this respect, distance in total variation between spectral measures is very attractive.

3.1.3. Related Extremum Problems

Problems 3.1-3.4 are related to additional extremum problems which are introduced below.

1. Let ν and μ be absolutely continuous with respect to the Lebesgue measure so that $\varphi(x) \triangleq \frac{d\nu}{dx}(x)$, $\psi(x) \triangleq \frac{d\mu}{dx}(x)$ (e.g., $\varphi(\cdot)$, $\psi(\cdot)$ are the probability density functions of $\nu(\cdot)$ and $\mu(\cdot)$, respectively). Then, $||\nu-\mu||_{TV} = \int_\Sigma |\varphi(x)-\psi(x)|dx$ and hence, (3.6) and (3.7) reduce to

$$D^+(R) = \sup_{\varphi\in L_1:\int_\Sigma |\varphi(x)-\psi(x)|dx\leq R} \int_\Sigma \ell(x)\varphi(x)dx$$
$$R^+(D) = \sup_{\varphi\in L_1:\int_\Sigma \ell(x)\varphi(x)dx\leq D} \int_\Sigma |\varphi(x)-\psi(x)|dx.$$

2. Let Σ be a non-empty denumerable set endowed with the discrete topology including finite cardinality $|\Sigma|$, with $\mathcal{M}_1(\Sigma)$ identified with the standard probability simplex in $\mathbb{R}^{|\Sigma|}$, that is, the set of all $|\Sigma|$-dimensional vectors which are probability vectors, and $\ell(x) \triangleq -\log \nu(x), x\in\Sigma$, where $\{\nu(x) : x\in\Sigma\} \in \mathcal{M}_1(\Sigma)$, $\{\mu(x) : x\in\Sigma\} \in \mathcal{M}_1(\Sigma)$. Then, (3.6) is equivalent to maximizing the entropy of $\{\nu(x) : x\in\Sigma\}$ subject to total variational distance metric constraint defined by

$$D^+(R) = \sup_{\nu\in\mathcal{M}_1(\Sigma):\sum_{x\in\Sigma}|\nu(x)-\mu(x)|\leq R} H(\nu), \quad H(\nu) = -\sum_{x\in\Sigma}\log(\nu(x))\nu(x). \tag{3.10}$$

Problem (3.10) is of interest when the concept of insufficient reasoning (e.g., Jayne's maximum entropy principle [34, 35]) is applied to construct a model for $\nu \in \mathcal{M}_1(\Sigma)$, subject to information quantified via the total variational distance metric between ν and an empirical distribution μ. In the context of stochastic uncertain control systems and its relation to robustness, Problem (3.10) with the total variational distance constraint replaced by relative entropy distance constraint is investigated in [3, 51].

3.2. Characterization of Measures on Abstract Spaces

This section utilizes signed measures and some of their properties to convert Problems 3.1-3.4 into equivalent extremum problems, and to characterize the extremum measures on abstract spaces. We describe the results using abstract spaces to avoid excluding measures defined on Borel spaces.

Let $\mathcal{M}_{sm}(\Sigma)$ denote the set of finite signed measures. Then, any $\eta \in \mathcal{M}_{sm}(\Sigma)$ has a Jordan decomposition $\left\{\eta^+, \eta^-\right\}$ such that $\eta = \eta^+ - \eta^-$, and the total variation of η is defined by $\|\eta\|_{TV} \triangleq \eta^+(\Sigma) + \eta^-(\Sigma)$. Define the following subset

$$\mathcal{M}_{sm}^0(\Sigma) \triangleq \left\{\eta \in \mathcal{M}_{sm}(\Sigma) : \eta(\Sigma) = 0\right\}.$$

For $\xi \in \mathcal{M}_{sm}^0(\Sigma)$, then $\xi(\Sigma) = 0$, which implies that $\xi^+(\Sigma) = \xi^-(\Sigma)$, and hence $\xi^+(\Sigma) = \xi^-(\Sigma) = \frac{\|\xi\|_{TV}}{2}$. For any $\nu, \mu \in \mathcal{M}_1(\Sigma)$ then $\xi \triangleq \nu - \mu \in \mathcal{M}_{sm}^0(\Sigma)$ and hence

$$\xi = (\nu - \mu)^+ - (\nu - \mu)^- \equiv \xi^+ - \xi^-.$$

3.2.1. Equivalent Formulation of Maximization Problems

In this section we investigate maximization Problems 3.1 and 3.2.

Equivalent Formulation of $D^+(R)$

Before we proceed with the equivalent formulation of Problem 3.1, we discuss first some of its properties.

Lemma 3.1. *Consider Problem 3.1. Then*

1) *$D^+(R)$ is a non-decreasing concave function of R.*

2) *If $R \leq R_{\max}$,*

$$D^+(R) = \sup_{\|\nu-\mu\|_{TV}=R} \int_\Sigma \ell(x)\nu(dx) \tag{3.11}$$

where $R_{\max}$ is the smallest non-negative number belonging to $[0,2]$ such that $D^+(R)$ is constant in $[R_{\max}, 2]$.

Proof. 1) Suppose $0 \leq R_1 \leq R_2$, then for every $\nu \in \mathbf{B}_{R_1}(\mu)$ we have $\|\nu - \mu\|_{TV} \leq R_1 \leq R_2$, and therefore $\nu \in \mathbf{B}_{R_2}(\mu)$, hence

$$\sup_{\nu \in \mathbf{B}_{R_1}(\mu)} \int_\Sigma \ell(x)\nu(dx) \leq \sup_{\nu \in \mathbf{B}_{R_2}(\mu)} \int_\Sigma \ell(x)\nu(dx)$$

which is equivalent to $D^+(R_1) \leq D^+(R_2)$. So $D^+(R)$ is a non-decreasing function of R. Now consider two points $(R_1, D^+(R_1))$ and $(R_2, D^+(R_2))$ on the linear functional curve, such that $\nu_1 \in \mathbf{B}_{R_1}(\mu)$ achieves the supremum of (3.6) for R_1, and $\nu_2 \in \mathbf{B}_{R_2}(\mu)$ achieves the supremum of (3.6) for R_2. Then, $\|\nu_1 - \mu\|_{TV} \leq R_1$ and $\|\nu_2 - \mu\|_{TV} \leq R_2$. For any $\lambda \in (0,1)$, we have

$$\|\lambda\nu_1 + (1-\lambda)\nu_2 - \mu\|_{TV} \leq \lambda\|\nu_1 - \mu\|_{TV} + (1-\lambda)\|\nu_2 - \mu\|_{TV} \leq \lambda R_1 + (1-\lambda)R_2 = R.$$

Define $\nu^* \triangleq \lambda\nu_1 + (1-\lambda)\nu_2$, $R \triangleq \lambda R_1 + (1-\lambda)R_2$. The previous equation implies that $\nu^* \in \mathbf{B}_R(\mu)$, hence $D^+(\lambda R_1 + (1-\lambda)R_2) \geq \int_\Sigma \ell(x)\nu^*(dx)$. Therefore,

$$\begin{aligned} D^+(R) &= \sup_{\nu \in \mathbf{B}_R(\mu)} \int_\Sigma \ell(x)\nu(dx) \geq \int_\Sigma \ell(x)\nu^*(dx) = \int_\Sigma \ell(x)\left(\lambda\nu_1(dx) + (1-\lambda)\nu_2(dx)\right) \\ &= \lambda \int_\Sigma \ell(x)\nu_1(dx) + (1-\lambda)\int_\Sigma \ell(x)\nu_2(dx) = \lambda D^+(R_1) + (1-\lambda)D^+(R_2). \end{aligned}$$

So, $D^+(R)$ is a concave function of R.

2) The right side of (3.11), say $\bar{D}^+(R)$, is a concave function of R. But $D^+(R) = \sup_{R' \leq R} \bar{D}^+(R')$ which completes the derivation of (3.11). ∎

Consider the pay-off of Problem 3.1, for $\ell \in BC^+(\Sigma)$. The solution is based on finding an upper bound which is achievable. The following inequalities hold.

$$\begin{aligned} \mathbb{L}_1(\nu) \triangleq \int_\Sigma \ell(x)\nu(dx) &= \int_\Sigma \ell(x)\Big(\nu(dx) - \mu(dx)\Big) + \int_\Sigma \ell(x)\mu(dx) \\ &\overset{(a)}{=} \int_\Sigma \ell(x)\left(\xi^+(dx) - \xi^-(dx)\right) + \int_\Sigma \ell(x)\mu(dx) \\ &= \int_\Sigma \ell(x)\xi^+(dx) - \int_\Sigma \ell(x)\xi^-(dx) + \int_\Sigma \ell(x)\mu(dx) \\ &\overset{(b)}{\leq} \sup_{x\in\Sigma} \ell(x)\xi^+(\Sigma) - \inf_{x\in\Sigma} \ell(x)\xi^-(\Sigma) + \int_\Sigma \ell(x)\mu(dx) \\ &\overset{(c)}{=} \sup_{x\in\Sigma} \ell(x)\frac{\|\xi\|_{TV}}{2} - \inf_{x\in\Sigma} \ell(x)\frac{\|\xi\|_{TV}}{2} + \int_\Sigma \ell(x)\mu(dx) \\ &= \left\{\sup_{x\in\Sigma} \ell(x) - \inf_{x\in\Sigma} \ell(x)\right\}\frac{\|\xi\|_{TV}}{2} + \int_\Sigma \ell(x)\mu(dx) \qquad (3.12) \end{aligned}$$

where (a) follows from the Jordan decomposition of $(\nu-\mu)$, (b) follows due to $\ell \in BC^+(\Sigma)$, (c) follows because any $\xi \in \mathcal{M}_{sm}^0(\Sigma)$ satisfies $\xi^+(\Sigma) = \xi^-(\Sigma) = \frac{1}{2}\|\xi\|_{TV}$. For a given $\mu \in \mathcal{M}_1(\Sigma)$ and $\nu \in \mathbf{B}_R(\mu)$ define the set

$$\widetilde{\mathbf{B}}_R(\mu) \triangleq \Big\{\xi \in \mathcal{M}_{sm}^0(\Sigma) : \xi = \nu - \mu, \nu \in \mathcal{M}_1(\Sigma), \|\xi\|_{TV} \leq R\Big\}.$$

The upper bound in the right hand side of (3.12) is achieved by $\xi^* \in \widetilde{\mathbf{B}}_R(\mu)$ as follows. Let

$$x^0 \in \Sigma^0 \triangleq \Big\{x \in \overline{\Sigma} : \ell(x) = \sup\{\ell(y) : y \in \Sigma\} \equiv \ell_{\max}\Big\}$$
$$x_0 \in \Sigma_0 \triangleq \Big\{x \in \overline{\Sigma} : \ell(x) = \inf\{\ell(y) : y \in \Sigma\} \equiv \ell_{\min}\Big\}.$$

where $\overline{\Sigma}$ denotes the closure[4] of Σ. Take

$$\xi^*(dx) = \nu^*(dx) - \mu(dx) = \frac{R}{2}\left(\delta_{x^0}(dx) - \delta_{x_0}(dx)\right) \tag{3.13}$$

where $\delta_y(dx)$ denotes the Dirac measure concentrated at $y \in \Sigma$. This is indeed a signed measure with total variation $\|\xi^*\|_{TV} = \|\nu^* - \mu\|_{TV} = R$, and $\int_\Sigma \ell(x)(\nu^* - \mu)(dx) = \frac{R}{2}\left(\ell_{\max} - \ell_{\min}\right)$. Hence, by using (3.13) as a candidate of the maximizing distribution then the extremum Problem 3.1 is equivalent to

$$\int_\Sigma \ell(x)\nu^*(dx) = \frac{R}{2}\left\{\sup_{x\in\Sigma} \ell(x) - \inf_{x\in\Sigma} \ell(x)\right\} + \int_\Sigma \ell(x)\mu(dx) \tag{3.14}$$

where ν^* satisfies the constraint $\|\xi^*\|_{TV} = \|\nu^* - \mu\|_{TV} = R$, it is normalized $\nu^*(\Sigma) = 1$, and $0 \leq \nu^*(A) \leq 1$ on any $A \in \mathcal{B}(\Sigma)$. Alternatively, the pay-off $\int_\Sigma \ell(x)\nu^*(dx)$ can be written as

$$\begin{aligned} D^+(R) &= \int_\Sigma \ell(x)\nu^*(dx) \\ &= \int_{\Sigma^0} \ell_{\max}\nu^*(dx) + \int_{\Sigma_0} \ell_{\min}\nu^*(dx) + \int_{\Sigma\setminus\Sigma^0\cup\Sigma_0} \ell(x)\mu(dx). \end{aligned} \tag{3.15}$$

Hence, the optimal distribution $\nu^* \in \mathbf{B}_R(\mu)$ satisfies

$$\int_{\Sigma^0} \nu^*(dx) = \mu(\Sigma^0) + \frac{R}{2} \in [0,1] \tag{3.16a}$$

$$\int_{\Sigma_0} \nu^*(dx) = \mu(\Sigma_0) - \frac{R}{2} \in [0,1] \tag{3.16b}$$

$$\nu^*(A) = \mu(A), \quad \forall A \subseteq \Sigma \setminus \Sigma^0 \cup \Sigma_0. \tag{3.16c}$$

For any $R \in [0,2]$ such that $\nu^*(\Sigma^0) < 1$ and $\nu^*(\Sigma_0) > 0$, then (3.16) is the maximizing distribution while the resulting pay-off is (3.15). When these conditions are violated the measure ν^* on the sets Σ^0, Σ_0 and $\Sigma \setminus \Sigma^0 \cup \Sigma_0$ remains to be identified so the maximizing measure ν^* is characterized for all $R \in [0,2]$. The complete characterization of the extremum measure ν^* will be given in the Section 3.3.1 building on the discussion of this section.

[4]Closure of a set Σ consists of all points in Σ plus the limit points of Σ.

Remark 3.1.

(i) *For* $\mu \in \mathcal{M}_1(\Sigma)$ *which do not include point mass, and for* $f \in BC^+(\Sigma)$*, if* Σ^0 *and* Σ_0 *are countable, then (3.16) is* $\mu(\Sigma^0) = \mu(\Sigma_0) = 0$, $\nu^*(\Sigma_0) = 0$, $\nu^*(\Sigma^0) = \frac{R}{2}$, $\nu^*(\Sigma \setminus \Sigma^0 \cup \Sigma_0) = \mu(\Sigma \setminus \Sigma^0 \cup \Sigma_0) - \frac{R}{2}$.

(ii) *The first right side term in (3.14) is related to the oscillator semi-norm of* $f \in BM(\Sigma)$ *called global modulus of continuity defined by (2.3). Clearly, for* $f \in BM^+(\Sigma)$ *then*

$$\mathrm{osc}(f) = \sup_{x\in\Sigma} |f(x)| - \inf_{x\in\Sigma} |f(x)| = \sup_{x\in\Sigma} f(x) - \inf_{x\in\Sigma} f(x).$$

Corollary 3.1. *The value of* $R_{\max}$ *described in Lemma 3.1 is given by*

$$R_{\max} = 2\left(1 - \mu\left(\Sigma^0\right)\right). \tag{3.17}$$

Proof. We know that $D^+(R) \leq \sup_{x\in\Sigma} \ell(x)$, $\forall R \geq 0$, hence $D^+(R_{\max})$ can be at most $\sup_{x\in\Sigma} \ell(x)$. Since $D^+(R)$ is non-decreasing then

$$D^+(R_{\max}) \leq D^+(R) \leq \sup_{x\in\Sigma} \ell(x), \quad \text{for any } R \geq R_{\max}. \tag{3.18}$$

Consider a ν that achieves the supremum in (3.18). Let $\mu(\Sigma^0)$ and $\nu(\Sigma^0)$ to denote the nominal and true probability measures on Σ^0, respectively. If $\nu(\Sigma^0) = 1$ then $\nu(\Sigma \setminus \Sigma^0) = 0$. Therefore,

$$\begin{aligned}
\|\nu - \mu\|_{TV} &= \sum_{x\in\Sigma^0} |\nu(x) - \mu(x)| + \sum_{x\in\Sigma\setminus\Sigma^0} |\nu(x) - \mu(x)| \\
&\overset{(a)}{=} \sum_{x\in\Sigma^0} |\nu(x) - \mu(x)| + \sum_{x\in\Sigma\setminus\Sigma^0} |-\mu(x)| \\
&\overset{(b)}{=} \sum_{x\in\Sigma^0} \nu(x) - \sum_{x\in\Sigma^0} \mu(x) + \sum_{x\in\Sigma\setminus\Sigma^0} \mu(x) \\
&= 1 - \sum_{x\in\Sigma^0} \mu(x) + \sum_{x\in\Sigma\setminus\Sigma^0} \mu(x) \\
&= 2\left(1 - \sum_{x\in\Sigma^0} \mu(x)\right) = 2\left(1 - \mu(\Sigma^0)\right)
\end{aligned}$$

where (a) follows due to $\nu(\Sigma \setminus \Sigma^0) = 0$ which implies $\nu(x) = 0$ for any $x \in \Sigma \setminus \Sigma^0$, and (b) follows because $\nu(x) \geq \mu(x)$ for all $x \in \Sigma^0$. Therefore, $R_{\max} = 2(1 - \mu(\Sigma^0))$ implies that $D^+(R_{\max}) = \sup_{x\in\Sigma} \ell(x)$. Hence, $D^+(R) = \sup_{x\in\Sigma} \ell(x)$, for any $R \geq R_{\max}$. ∎

Equivalent Formulation of $R^{+}(D)$

Before we proceed with the equivalent formulation of Problem 3.2, we discuss first some of its properties.

Lemma 3.2. *Consider Problem 3.2. Then*

1) $R^{+}(D)$ is a non-decreasing convex function of D.

2) If $D \leq D_{\max}$,

$$R^{+}(D) = \sup_{\int_{\Sigma} \ell(x)\nu(dx)=D} \|\nu - \mu\|_{TV} \tag{3.19}$$

where $D_{\max}$ is the smallest non-negative number belonging to $[0, \infty)$ such that $R^{-}(D)$ is constant in $[D_{\max,\infty})$.

Proof. 1) Suppose $0 \leq D_1 \leq D_2$, then $\mathbf{Q}(D_1) \subset \mathbf{Q}(D_2)$, and $\sup_{\nu\in\mathbf{Q}(D_1)} \|\nu - \mu\|_{TV} \leq \sup_{\nu\in\mathbf{Q}(D_2)} \|\nu - \mu\|_{TV}$ which is equivalent to $R^{+}(D_1) \leq R^{+}(D_2)$. Hence, $R^{+}(D)$ is a non-decreasing function of D. Convexity is obtained by using the fact that $R^{+}(D)$ is the inverse mapping of $D^{+}(R)$. So, $R^{+}(D)$ is a convex function of D.

2) The right side of (3.19), say $\bar{R}^{+}(D)$, is convex function of D. But, $R^{+}(D) = \sup_{D'\leq D} \bar{R}^{+}(D')$ which completes the derivation of (3.19). ■

Consider the constraint of Problem 3.2, for $\ell \in BC^{+}(\Sigma)$. By following the same procedure as in Section 3.2.1, we obtain (3.14), that is

$$\int_{\Sigma} \ell(x)\nu^{*}(dx) = \frac{R}{2}\Big\{\sup_{x\in\Sigma} \ell(x) - \inf_{x\in\Sigma} \ell(x)\Big\} + \int_{\Sigma} \ell(x)\mu(dx). \tag{3.20}$$

Solving the above equation with respect to $R \equiv R^{+}$, the extremum Problem 3.2 is equivalent to

$$R^{+}(D) = \frac{2\Big(D - \int_{\Sigma} \ell(x)\mu(dx)\Big)}{\Big\{\sup_{x\in\Sigma} \ell(x) - \inf_{x\in\Sigma} \ell(x)\Big\}} \tag{3.21}$$

where ν^{*} in (3.20) satisfies the constraint $\int_{\Sigma} \ell(x)\nu^{*}(dx) = D$, it is normalized $\nu^{*}(\Sigma) = 1$, and $0 \leq \nu(A) \leq 1$ on any $A \in \mathcal{B}(\Sigma)$. We can now identify $D_{\max}$ described in Lemma 3.2.

Corollary 3.2. *The value of $D_{\max}$ described in Lemma 3.2 is given by*

$$D_{\max} = \ell_{\max}. \tag{3.22}$$

Proof. We know that $R^+(D) \leq 2$ for all $D \geq 0$, hence $R^+(D_{\max})$ can be at most equal to 2. For the extreme case $R^+(D) = 2$, we have that ν and μ are disjoint in the sense that Σ can be partitioned into two disjoint subsets Σ^0 and $\Sigma \setminus \Sigma^0$ such that ν puts all of its probability mass in Σ^0, that is, $\nu(\Sigma^0) = 1$ and hence $\nu(\Sigma \setminus \Sigma^0) = 0$, and μ puts all of its in $\Sigma \setminus \Sigma^0$, that is, $\mu(\Sigma \setminus \Sigma^0) = 1$ and hence $\mu(\Sigma^0) = 0$.
Without loss of generality, assume that μ puts all of its probability mass in $\Sigma \setminus \Sigma^0$. Let $D_{\max} = \ell_{\max}$, then it is obvious that $\nu(\Sigma^0) = 1$, and hence $R^+(D_{\max}) = 2$. Since $R^+(D)$ is non-decreasing then $R^+(D_{\max}) \leq R^+(D) \leq 2$ for any $D \geq D_{\max}$. Hence, $R^+(D) = 2$ for any $D \geq D_{\max}$. ■

3.2.2. Equivalent Formulation of Minimization Problems

In this section we investigate minimization Problems 3.3 and 3.4.

Equivalent Formulation of $D^-(R)$

Before we proceed with the equivalent formulation of Problem 3.3, we discuss first some of its properties.

Lemma 3.3. *Consider Problem 3.3. Then*

1) *$D^-(R)$ is a non-increasing convex function of R.*

2) *If $R \leq R_{\max}$,*

$$D^-(R) = \inf_{\|\nu-\mu\|_{TV}=R} \int_\Sigma \ell(x)\nu(dx) \tag{3.23}$$

where $R_{\max}$ is the smallest non-negative number belonging to $[0, 2]$ such that $D^-(R)$ is constant in $[R_{\max}, 2]$.

Proof. 1) Suppose $0 \leq R_1 \leq R_2$, then for every $\nu \in \mathbf{B}_{R_1}(\mu)$ we have $\|\nu-\mu\|_{TV} \leq R_1 \leq R_2$, and therefore $\nu \in \mathbf{B}_{R_2}(\mu)$, hence

$$\inf_{\nu \in \mathbf{B}_{R_1}(\mu)} \int_\Sigma \ell(x)\nu(dx) \geq \inf_{\nu \in \mathbf{B}_{R_2}(\mu)} \int_\Sigma \ell(x)\nu(dx)$$

which is equivalent to $D^-(R_1) \geq D^-(R_2)$. So $D^-(R)$ is a non-increasing function of R. Now consider two points $(R_1, D^+(R_1))$ and $(R_2, D^+(R_2))$ on the linear functional curve, such that $\nu_1 \in \mathbf{B}_{R_1}(\mu)$ achieves the infimum of (3.8) for R_1, and $\nu_2 \in \mathbf{B}_{R_2}(\mu)$ achieves the infimum of (3.8) for R_2. Then, $\|\nu_1 - \mu\|_{TV} \leq R_1$ and $\|\nu_2 - \mu\|_{TV} \leq R_2$. For any $\lambda \in (0, 1)$, we have

$$\|\lambda\nu_1 + (1-\lambda)\nu_2 - \mu\|_{TV} \leq \lambda\|\nu_1 - \mu\|_{TV} + (1-\lambda)\|\nu_2 - \mu\|_{TV} \leq \lambda R_1 + (1-\lambda)R_2 = R.$$

Define $\nu^* \triangleq \lambda\nu_1 + (1-\lambda)\nu_2$, $R \triangleq \lambda R_1 + (1-\lambda)R_2$. The previous equation implies that $\nu^* \in \mathbf{B}_R(\mu)$, hence $D^-(\lambda R_1 + (1-\lambda)R_2) \leq \int_\Sigma \ell(x)\nu^*(dx)$. Therefore,

$$\begin{aligned} D^-(R) &= \inf_{\nu\in\mathbf{B}_R(\mu)} \int_\Sigma \ell(x)\nu(dx) \leq \int_\Sigma \ell(x)\nu^*(dx) = \int_\Sigma \ell(x)\left(\lambda\nu_1(dx) + (1-\lambda)\nu_2(dx)\right) \\ &= \lambda \int_\Sigma \ell(x)\nu_1(dx) + (1-\lambda)\int_\Sigma \ell(x)\nu_2(dx) = \lambda D^-(R_1) + (1-\lambda)D^-(R_2). \end{aligned}$$

So, $D^-(R)$ is a convex function of R.

2) The right side of (3.23), say $\bar{D}^-(R)$, is convex function of R. But $D^-(R) = \inf_{R'\leq R} \bar{D}^+(R')$ which completes the derivation of (3.23). ■

Consider the pay-off of Problem 3.3, for $\ell \in BC^+(\Sigma)$. The solution is based on finding a lower bound which is achievable. The following inequalities hold.

$$\begin{aligned} \mathbb{L}_1(\nu) \triangleq \int_\Sigma \ell(x)\nu(dx) &= \int_\Sigma \ell(x)\Big(\nu(dx) - \mu(dx)\Big) + \int_\Sigma \ell(x)\mu(dx) \\ &\overset{(a)}{=} \int_\Sigma \ell(x)\left(\xi^+(dx) - \xi^-(dx)\right) + \int_\Sigma \ell(x)\mu(dx) \\ &= \int_\Sigma \ell(x)\xi^+(dx) - \int_\Sigma \ell(x)\xi^-(dx) + \int_\Sigma \ell(x)\mu(dx) \\ &\overset{(b)}{\geq} \inf_{x\in\Sigma} \ell(x)\xi^+(\Sigma) - \sup_{x\in\Sigma} \ell(x)\xi^-(\Sigma) + \int_\Sigma \ell(x)\mu(dx) \\ &\overset{(c)}{=} \inf_{x\in\Sigma} \ell(x)\frac{\|\xi\|_{TV}}{2} - \sup_{x\in\Sigma} \ell(x)\frac{\|\xi\|_{TV}}{2} + \int_\Sigma \ell(x)\mu(dx) \\ &= \left\{\inf_{x\in\Sigma} \ell(x) - \sup_{x\in\Sigma} \ell(x)\right\}\frac{\|\xi\|_{TV}}{2} + \int_\Sigma \ell(x)\mu(dx) \end{aligned} \tag{3.24}$$

where (a) follows from the Jordan decomposition of $(\nu-\mu)$, (b) follows due to $\ell \in BC^+(\Sigma)$, (c) follows because any $\xi \in \mathcal{M}_{sm}^0(\Sigma)$ satisfies $\xi^+(\Sigma) = \xi^-(\Sigma) = \frac{1}{2}\|\xi\|_{TV}$. For a given $\mu \in \mathcal{M}_1(\Sigma)$ and $\nu \in \mathbf{B}_R(\mu)$ define the set

$$\widetilde{\mathbf{B}}_R(\mu) \triangleq \left\{\xi \in \mathcal{M}_{sm}^0(\Sigma) : \xi = \nu - \mu, \nu \in \mathcal{M}_1(\Sigma), \|\xi\|_{TV} \leq R\right\}.$$

The lower bound in the right hand side of (3.24) is achieved by $\xi^* \in \widetilde{\mathbf{B}}_R(\mu)$ as follows. Let

$$\begin{aligned} x^0 \in \Sigma^0 &\triangleq \left\{x \in \overline{\Sigma} : \ell(x) = \sup\{\ell(y) : y \in \Sigma\} \equiv \ell_{\max}\right\} \\ x_0 \in \Sigma_0 &\triangleq \left\{x \in \overline{\Sigma} : \ell(x) = \inf\{\ell(y) : y \in \Sigma\} \equiv \ell_{\min}\right\}. \end{aligned}$$

Take

$$\xi^*(dx) = \nu^*(dx) - \mu(dx) = \frac{R}{2}\left(\delta_{x_0}(dx) - \delta_{x^0}(dx)\right) \tag{3.25}$$

where $\delta_y(dx)$ denotes the Dirac measure concentrated at $y \in \Sigma$. This is indeed a signed measure with total variation $\|\xi^*\|_{TV} = \|\nu^* - \mu\|_{TV} = R$, and $\int_\Sigma \ell(x)(\nu^* - \mu)(dx) =$

$\frac{R}{2}(\ell_{\min} - \ell_{\max})$. Hence, by using (3.25) as a candidate of the minimizing distribution then the extremum Problem 3.2 is equivalent to

$$\int_{\Sigma} \ell(x)\nu^*(dx) = \frac{R}{2}\left\{\inf_{x\in\Sigma} \ell(x) - \sup_{x\in\Sigma} \ell(x)\right\} + \int_{\Sigma} \ell(x)\mu(dx) \tag{3.26}$$

where ν^* satisfies the constraint $\|\xi^*\|_{TV} = \|\nu^* - \mu\|_{TV} = R$, it is normalized $\nu^*(\Sigma) = 1$, and $0 \leq \nu^*(A) \leq 1$ on any $A \in \mathcal{B}(\Sigma)$. Alternatively, the pay-off $\int_{\Sigma} \ell(x)\nu^*(dx)$ can be written as

$$\begin{aligned} D^-(R) &= \int_{\Sigma} \ell(x)\nu^*(dx) \\ &= \int_{\Sigma^0} \ell_{\max}\nu^*(dx) + \int_{\Sigma_0} \ell_{\min}\nu^*(dx) + \int_{\Sigma\setminus\Sigma^0\cup\Sigma_0} \ell(x)\mu(dx). \end{aligned} \tag{3.27}$$

Hence, the optimal distribution $\nu^* \in \mathbf{B}_R(\mu)$ satisfies

$$\int_{\Sigma_0} \nu^*(dx) = \mu(\Sigma^0) + \frac{R}{2} \in [0,1] \tag{3.28a}$$

$$\int_{\Sigma^0} \nu^*(dx) = \mu(\Sigma_0) - \frac{R}{2} \in [0,1] \tag{3.28b}$$

$$\nu^*(A) = \mu(A), \quad \forall A \subseteq \Sigma \setminus \Sigma^0 \cup \Sigma_0. \tag{3.28c}$$

For any $R \in [0,2]$ such that $\nu^*(\Sigma_0) < 1$ and $\nu^*(\Sigma^0) > 0$, then (3.28) is the minimizing distribution while the resulting pay-off is (3.27). When these conditions are violated the measure ν^* on the sets Σ^0, Σ_0 and $\Sigma \setminus \Sigma^0 \cup \Sigma_0$ remains to be identified so the minimizing measure ν^* is characterized for all $R \in [0,2]$. The complete characterization of the extremum measure ν^* will be given in the Section 3.3.4 building on the discussion of this section. Next, we identify $R_{\max}$ described in Lemma 3.3.

Corollary 3.3. *The value of $R_{\max}$ described in Lemma 3.3 is given by*

$$R_{\max} = 2(1 - \mu(\Sigma_0)). \tag{3.29}$$

Proof. We know that $D^-(R) \geq \inf_{x\in\Sigma} \ell(x)$, $\forall R \geq 0$, hence $D^-(R_{\max})$ can be at least $\inf_{x\in\Sigma} \ell(x)$. Since $D^-(R)$ is non-increasing then $\inf_{x\in\Sigma} \ell(x) \leq D^-(R) \leq D^-(R_{\max})$, for any $R \geq R_{\max}$. Consider a ν that achieves this infimum. Let $\mu(\Sigma_0)$ and $\nu(\Sigma_0)$ to denote the nominal and true probability measures on Σ_0, respectively. If $\nu(\Sigma_0) = 1$ then $\nu(\Sigma \setminus \Sigma_0) = 0$.

Therefore,

$$\begin{aligned}
\|\nu-\mu\|_{TV} &= \sum_{x\in\Sigma_0}|\nu(x)-\mu(x)| + \sum_{x\in\Sigma\setminus\Sigma_0}|\nu(x)-\mu(x)| \\
&\overset{(a)}{=} \sum_{x\in\Sigma_0}|\nu(x)-\mu(x)| + \sum_{x\in\Sigma\setminus\Sigma_0}|-\mu(x)| \\
&\overset{(b)}{=} \sum_{x\in\Sigma_0}\nu(x) - \sum_{x\in\Sigma_0}\mu(x) + \sum_{x\in\Sigma\setminus\Sigma_0}\mu(x) \\
&= 1 - \sum_{x\in\Sigma_0}\mu(x) + \sum_{x\in\Sigma\setminus\Sigma_0}\mu(x) \\
&= 2\left(1-\sum_{x\in\Sigma_0}\mu(x)\right) = 2\left(1-\mu(\Sigma_0)\right)
\end{aligned}$$

where (a) follows due to $\nu(\Sigma\setminus\Sigma_0)=0$ which implies $\nu(x)=0$ for any $x\in\Sigma\setminus\Sigma_0$, and (b) follows because $\nu(x)\geq\mu(x)$ for all $x\in\Sigma_0$. Therefore, $R_{\max}=2(1-\mu(\Sigma_0))$ implies that $D^-(R_{\max})=\inf_{x\in\Sigma}\ell(x)$. Hence, $D^-(R)=\inf_{x\in\Sigma}\ell(x)$, for any $R\geq R_{\max}$. ∎

Equivalent Formulation of $R^-(D)$

Before we proceed with the equivalent formulation of Problem 3.4, we discuss first some of its properties.

Lemma 3.4. *Consider Problem 3.4. Then*

1) $R^-(D)$ is a non-increasing convex function of D.

2) If $D\leq D_{\max}$,

$$R^-(D) = \inf_{\int_\Sigma \ell(x)\nu(dx)=D}\|\nu-\mu\|_{TV} \tag{3.30}$$

where $D_{\max}$ is the smallest non-negative number belonging to $[0,\infty)$ such that $R^-(D)=0$ for any $D\in[D_{\max},\infty)$.

Proof. 1) Suppose $0\leq D_1\leq D_2$, then $\mathbf{Q}(D_1)\subset\mathbf{Q}(D_2)$, and $\inf_{\nu\in\mathbf{Q}(D_1)}\|\nu-\mu\|_{TV}\geq\inf_{\nu\in\mathbf{Q}(D_2)}\|\nu-\mu\|_{TV}$ which is equivalent to $R^-(D_1)\geq R^-(D_2)$. Hence, $R^-(D)$ is a non-increasing function of D. Now consider two points $(D_1,R^-(D_1))$ and $(D_2,R^-(D_2))$ on the total variation curve. Let $D\triangleq\lambda D_1+(1-\lambda)D_2$, $\nu^*\triangleq\lambda\nu_1+(1-\lambda)\nu_2$ and $\nu_1\in\mathbf{Q}(D_1)$, $\nu_2\in\mathbf{Q}(D_2)$ such that $\|\nu_1-\mu\|_{TV}=R^-(D_1)$ and $\|\nu_2-\mu\|_{TV}=R^-(D_2)$. Then, $\int_\Sigma\ell(x)\nu_1(dx)\leq D_1$ and $\int_\Sigma\ell(x)\nu_2(dx)\leq D_2$. Taking convex combination leads to

$$\lambda\int_\Sigma\ell(x)\nu_1(dx)+(1-\lambda)\int_\Sigma\ell(x)\nu_2(dx)\leq\lambda D_1+(1-\lambda)D_2=D$$

and hence $\nu^* \in \mathbf{Q}(D)$. So,

$$R^-(D) = \inf_{\nu \in \mathbf{Q}(D)} \|\nu - \mu\|_{TV} \leq \|\nu^* - \mu\|_{TV} = \|\lambda\nu_1 + (1-\lambda)\nu_2 - \mu\|_{TV}$$
$$\leq \lambda\|\nu_1 - \mu\|_{TV} + (1-\lambda)\|\nu_2 - \mu\|_{TV} = \lambda R^-(D_1) + (1-\lambda)R^-(D_2).$$

This shows that $R^-(D)$ is convex function of D.

2) The right side of (3.30), say $\bar{R}^-(D)$, is convex function of D. But, $R^-(D) = \inf_{D' \leq D} \bar{R}^-(D')$ which completes the derivation of (3.30). ■

Consider the constraint of Problem 3.4, for $\ell \in BC^+(\Sigma)$. By following the same procedure as in Section 3.2.2, we obtain (3.26), that is

$$\int_\Sigma \ell(x)\nu^*(dx) = \frac{R}{2}\left\{\inf_{x\in\Sigma} \ell(x) - \sup_{x\in\Sigma} \ell(x)\right\} + \int_\Sigma \ell(x)\mu(dx). \tag{3.31}$$

Solving the above equation with respect to R the extremum Problem 3.4 (for $D < \int_\Sigma \ell(x)\mu(dx)$) is equivalent to

$$R^-(D) = \frac{2(D - \int_\Sigma \ell(x)\mu(dx))}{\left\{\inf_{x\in\Sigma} \ell(x) - \sup_{x\in\Sigma} \ell(x)\right\}} \tag{3.32}$$

where ν^* satisfies the constraint $\int_\Sigma \ell(x)\nu^*(dx) = D$, it is normalized $\nu^*(\Sigma) = 1$, and $0 \leq \nu(A) \leq 1$ on any $A \in \mathcal{B}(\Sigma)$. We can now identify $D_{\max}$ described in Lemma 3.4.

Corollary 3.4. *The value of $D_{\max}$ described in Lemma 3.4 is given by*

$$D_{\max} = \int_\Sigma \ell(x)\mu(dx).$$

Proof. We know that $R^-(D) \geq 0$ for all $D \geq 0$ hence $R^-(D_{\max})$ can be at least zero. Let $D_{\max} = \int_\Sigma \ell(x)\mu(dx)$, then it is obvious that $R^-(D_{\max}) = 0$. Since $R^-(D)$ in non-increasing, then $0 \leq R^-(D) \leq R^-(D_{\max})$, for any $D \geq D_{\max}$. Hence, $R^-(D) = 0$, for any $D \geq D_{\max}$. ■

3.3. Characterization of Measures for Finite Alphabets

This section uses the results of Section 3.2 to compute closed form expressions for the extremum measures ν^* for any $R \in [0,2]$, when Σ is a finite alphabet space to give the intuition into the solution procedure. This is done by identifying the sets Σ^0, Σ_0, $\Sigma \setminus \Sigma^0 \cup \Sigma_0$, and the measure ν^* on these sets for any $R \in [0,2]$. Although this can be done for probability measures on complete separable metric spaces (Polish spaces) (Σ, d_Σ), and for $\ell \in BM^+(\Sigma)$,

$\ell \in BC^+(\Sigma)$, $L^{\infty,+}(\Sigma, \mathcal{B}(\Sigma), \nu)$, we prefer to discuss the finite alphabet case to gain additional insight into these problems. At section 3.3.7 we shall use the finite alphabet case to discuss the extensions to countable alphabet and to $\ell \in L^{\infty,+}(\Sigma, \mathcal{B}(\Sigma), \nu)$.

Consider the finite alphabet case $(\Sigma, \mathcal{M})$, where $\text{card}(\Sigma) = |\Sigma|$ is finite, $\mathcal{M} = 2^{|\Sigma|}$. Thus, ν and μ are point mass distributions on Σ. Define the set of probability vectors on Σ by

$$\mathbb{P}(\Sigma) \triangleq \Big\{ p = (p_1, \ldots, p_{|\Sigma|}) : p_i \geq 0, i = 0, \ldots, |\Sigma|, \sum_{i \in \Sigma} p_i = 1 \Big\}. \tag{3.33}$$

Thus, $p \in \mathbb{P}(\Sigma)$ is a probability vector in $\mathbb{R}_+^{|\Sigma|}$. Also let $\ell \triangleq \{\ell_1, \ldots, \ell_{|\Sigma|}\}$ so that $\ell \in \mathbb{R}_+^{|\Sigma|}$ (e.g., set of non-negative vectors of dimension $|\Sigma|$). Next, we introduce some basic definitions which will be used for the solution of the extremum problems.

Define the maximum and minimum values of the sequence $\{\ell_1, \ldots, \ell_{|\Sigma|}\}$ by

$$\ell_{\max} \triangleq \max_{i \in \Sigma} \ell_i, \qquad \ell_{\min} \triangleq \min_{i \in \Sigma} \ell_i$$

and its corresponding support sets by

$$\Sigma^0 \triangleq \{i \in \Sigma : \ell_i = \ell_{\max}\}, \qquad \Sigma_0 \triangleq \{i \in \Sigma : \ell_i = \ell_{\min}\}.$$

For any remaining sequence, $\{\ell_i : i \in \Sigma \setminus \Sigma^0 \cup \Sigma_0\}$, and for $1 \leq r \leq |\Sigma \setminus \Sigma^0 \cup \Sigma_0|$ define recursively the set of indices for which ℓ achieves its $(k+1)^{th}$ smallest value by

$$\Sigma_k \triangleq \left\{ i \in \Sigma : \ell_i = \min \left\{ \ell_\alpha : \alpha \in \Sigma \setminus \Sigma^0 \cup \Big(\bigcup_{j=1}^{k} \Sigma_{j-1} \Big) \right\} \right\}, \ k \in \{1, 2, \ldots, r\} \tag{3.34}$$

and the set of indices for which ℓ achieves its $(k+1)^{th}$ largest value by

$$\Sigma^k \triangleq \left\{ i \in \Sigma : \ell_i = \max \left\{ \ell_\alpha : \alpha \in \Sigma \setminus \Sigma_0 \cup \Big(\bigcup_{j=1}^{k} \Sigma^{j-1} \Big) \right\} \right\}, \ k \in \{1, 2, \ldots, r\} \tag{3.35}$$

till all the elements of Σ are exhausted (i.e., k is at most $|\Sigma \setminus \Sigma^0 \cup \Sigma_0|$). Define the corresponding values of the sequence of sets in (3.34) by

$$\ell\Big(\Sigma_k\Big) \triangleq \min_{i \in \Sigma \setminus \Sigma^0 \cup \left(\bigcup_{j=1}^{k} \Sigma_{j-1} \right)} \ell_i, \quad k \in \{1, 2, \ldots, r\} \tag{3.36}$$

and the corresponding values of the sequence of sets in (3.35) by

$$\ell\left(\Sigma^k\right) \triangleq \max_{i \in \Sigma \setminus \Sigma_0 \cup \left(\bigcup_{j=1}^{k} \Sigma^{j-1} \right)} \ell_i, \quad k \in \{1, 2, \ldots, r\} \tag{3.37}$$

where r is the number of the support sets which is at most $|\Sigma \setminus \Sigma^0 \cup \Sigma_0|$; for example, when $k = 1$, $\ell\left(\Sigma_1\right) = \min_{i \in \Sigma \setminus \Sigma^0 \cup \Sigma_0} \ell_i$ and $\ell\left(\Sigma^1\right) = \max_{i \in \Sigma \setminus \Sigma^0 \cup \Sigma_0} \ell_i$, when $k = 2$, $\ell\left(\Sigma_2\right) =$

$\min_{i\in\Sigma\setminus\Sigma^0\cup\Sigma_0\cup\Sigma_1} \ell_i$ and $\ell(\Sigma^2) = \max_{i\in\Sigma\setminus\Sigma^0\cup\Sigma_0\cup\Sigma^1} \ell_i$, etc. Furthermore, if $\ell_1 < \ell_2 < \cdots < \ell_{|\Sigma|}$ then $\Sigma^0 = \{|\Sigma|\}$, $\Sigma_0 = \{1\}$ and $\Sigma_k = \{k+1\}$, $\Sigma^k = \{|\Sigma| - k\}$ for $k = 1, 2, \ldots, |\Sigma| - 2$. Note that, (3.34) and (3.36) will be used exclusively for the solution of maximization Problems 3.1 and 3.2, while (3.35) and (3.37) will be used exclusively for the solution of minimization Problems 3.3 and 3.4.

3.3.1. $D^+(R)$: The Finite Alphabet Case

Suppose $\nu \in \mathbb{P}(\Sigma)$ is the true probability vector and $\mu \in \mathbb{P}(\Sigma)$ is the nominal fixed probability vector. The extremum problem reduces to

$$D^+(R) \triangleq \max_{\nu\in\mathbf{B}_R(\mu)} \sum_{i\in\Sigma} \ell_i\nu_i = \max_{\nu\in\mathbb{P}(\Sigma):\sum_{i\in\Sigma}|\nu_i-\mu_i|\leq R} \sum_{i\in\Sigma} \ell_i\nu_i. \tag{3.38}$$

Next, we apply the results of Section 3.2 to characterize the optimal ν^* for any $R \in [0, 2]$. By defining, $\xi_i \triangleq \nu_i - \mu_i, i = 1, \ldots, |\Sigma|$ and $\xi \in \mathcal{M}^0_{sm}(\Sigma)$, Problem 3.1 can be reformulated as follows.

$$\max_{\nu\in\mathbf{B}_R(\mu)} \sum_{i\in\Sigma} \ell_i\nu_i \longrightarrow \sum_{i\in\Sigma} \ell_i\mu_i + \max_{\xi\in\widetilde{\mathbf{B}}_R(\mu)} \sum_{i\in\Sigma} \ell_i\xi_i. \tag{3.39}$$

Note that $\xi \in \widetilde{\mathbf{B}}_R(\mu)$ is described by the constraints

$$\alpha \triangleq \sum_{i\in\Sigma} |\xi_i| \leq R, \quad \sum_{i\in\Sigma} \xi_i = 0, \quad 0 \leq \xi_i + \mu_i \leq 1, \quad \forall i \in \Sigma. \tag{3.40}$$

The positive and negative variation of the signed measure $\xi = \nu - \mu \in \mathcal{M}^0_{sm}(\Sigma)$ are defined by $\xi^+ = \max\{\xi, 0\}$ and $\xi^- = \max\{-\xi, 0\}$. Therefore,

$$\sum_{i\in\Sigma} \xi_i = \sum_{i\in\Sigma} \xi_i^+ - \sum_{i\in\Sigma} \xi_i^-, \qquad \sum_{i\in\Sigma} |\xi_i| = \sum_{i\in\Sigma} \xi_i^+ + \sum_{i\in\Sigma} \xi_i^- \tag{3.41}$$

and hence

$$\sum_{i\in\Sigma} \xi_i^+ = \frac{\sum_{i\in\Sigma} \xi_i + \sum_{i\in\Sigma} |\xi_i|}{2} \equiv \frac{\alpha}{2} \tag{3.42}$$

$$\sum_{i\in\Sigma} \xi_i^- = \frac{-\sum_{i\in\Sigma} \xi_i + \sum_{i\in\Sigma} |\xi_i|}{2} \equiv \frac{\alpha}{2}. \tag{3.43}$$

In addition,

$$\sum_{i\in\Sigma} \ell_i\xi_i = \sum_{i\in\Sigma} \ell_i\xi_i^+ - \sum_{i\in\Sigma} \ell_i\xi_i^-. \tag{3.44}$$

The following theorem characterizes the solution of Problem 3.1.

Theorem 3.1. *The solution of the finite alphabet version of Problem 3.1 is given by*

$$D^{+}(R) = \ell_{\max}\nu^{*}(\Sigma^{0}) + \ell_{\min}\nu^{*}(\Sigma_{0}) + \sum_{k=1}^{r}\ell(\Sigma_{k})\nu^{*}(\Sigma_{k}) \tag{3.45}$$

where the optimal probabilities are given by

$$\nu^{*}(\Sigma^{0}) \triangleq \sum_{i\in\Sigma^{0}}\nu_i^{*} = \sum_{i\in\Sigma^{0}}\mu_i + \frac{\alpha}{2} \tag{3.46a}$$

$$\nu^{*}(\Sigma_{0}) \triangleq \sum_{i\in\Sigma_{0}}\nu_i^{*} = \Big(\sum_{i\in\Sigma_{0}}\mu_i - \frac{\alpha}{2}\Big)^{+} \tag{3.46b}$$

$$\nu^{*}(\Sigma_{k}) \triangleq \sum_{i\in\Sigma_{k}}\nu_i^{*} = \Bigg(\sum_{i\in\Sigma_{k}}\mu_i - \Big(\frac{\alpha}{2} - \sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\mu_i\Big)^{+}\Bigg)^{+} \tag{3.46c}$$

$$\alpha = \min\left(R, R_{\max}\right), \qquad R_{\max} \triangleq 2(1 - \sum_{i\in\Sigma^{0}}\mu_i) \tag{3.46d}$$

with $k = 1, 2, \ldots, r$ and r is the number of Σ_k sets which is at most $|\Sigma \setminus \Sigma^{0} \cup \Sigma_{0}|$.

The solution of Problem 3.1 is obtained by identifying the partition of Σ into disjoint sets $\{\Sigma^{0}, \Sigma_{0}, \Sigma_{1}, \ldots, \Sigma_{k}\}$ and the measures on this partition. The main idea is to express the total variation distance constraint as a summation of the positive and negative variation of a signed measure, and then to find upper and lower bounds on the probabilities of Σ^{0} and $\Sigma\setminus\Sigma^{0}$, which are achievable. Utilizing the fact that the positive and negative variation parts of the total variation distance have equal mass concentrated on them, closed form expressions of the probability measures, on these sets, which achieve the upper and lower bounds are derived (i.e., using (3.41), then (3.42) holds).

In the following Lemma upper and lower bounds which are achievable are obtained. These they will be used for the derivation of Theorem 3.1.

Lemma 3.5. *Consider Problem 3.1.*

(a) Upper Bound.

$$\sum_{i\in\Sigma}\ell_i\xi_i^{+} \leq \ell_{\max}\left(\frac{\alpha}{2}\right). \tag{3.47}$$

The bound holds with equality if

$$\sum_{i\in\Sigma^{0}}\mu_i + \frac{\alpha}{2} \leq 1 \tag{3.48a}$$

$$\sum_{i\in\Sigma^{0}}\xi_i^{+} = \frac{\alpha}{2} \tag{3.48b}$$

$$\xi_i^{+} = 0 \ \ \textit{for} \ \ i \in \Sigma\setminus\Sigma^{0}. \tag{3.48c}$$

(b) Lower Bound.

Case 1. If $\sum_{i\in\Sigma_0}\mu_i - \frac{\alpha}{2} \geq 0$ *then*

$$\sum_{i\in\Sigma}\ell_i\xi_i^- \geq \ell_{\min}\left(\frac{\alpha}{2}\right). \tag{3.49}$$

The bound holds with equality if

$$\sum_{i\in\Sigma_0}\mu_i - \frac{\alpha}{2} \geq 0 \tag{3.50a}$$

$$\sum_{i\in\Sigma_0}\xi_i^- = \frac{\alpha}{2} \tag{3.50b}$$

$$\xi_i^- = 0 \ \textit{ for } \ i\in\Sigma\setminus\Sigma_0. \tag{3.50c}$$

Case 2. If $\sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\mu_i - \frac{\alpha}{2} \leq 0$ *for any* $k\in\{1,2,\ldots,r\}$ *then*

$$\sum_{i\in\Sigma}\ell_i\xi_i^- \geq \ell(\Sigma_k)\Big(\frac{\alpha}{2} - \sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\mu_i\Big) + \sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\ell_i\mu_i. \tag{3.51}$$

Moreover, equality holds if

$$\sum_{i\in\Sigma_{j-1}}\xi_i^- = \sum_{i\in\Sigma_{j-1}}\mu_i, \ \textit{ for all } \ j=1,2,\ldots,k \tag{3.52a}$$

$$\sum_{i\in\Sigma_k}\xi_i^- = \Big(\frac{\alpha}{2} - \sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\mu_i\Big) \tag{3.52b}$$

$$\sum_{j=0}^{k}\sum_{i\in\Sigma_j}\mu_i - \frac{\alpha}{2} \geq 0 \tag{3.52c}$$

$$\xi_i^- = 0 \ \textit{ for all } \ i\in\Sigma\setminus\Sigma_0\cup\Sigma_1\cup\ldots\cup\Sigma_k. \tag{3.52d}$$

Proof. Part (a) and Part (b), case 1, follows from Section 3.2.1 (equivalent formulation of $D^+(R)$). For Part (b), case 2, we proceed as follows. Consider any $k\in\{1,2,\ldots,r\}$. First, we show that inequality holds. From Part (b), case 1, we have that

$$\sum_{i\in\Sigma\setminus\bigcup_{j=1}^{k}\Sigma_{j-1}}\ell_i\xi_i^- \geq \min_{j\in\Sigma\setminus\bigcup_{j=1}^{k}\Sigma_{j-1}}\ell_j\sum_{i\in\Sigma\setminus\bigcup_{j=1}^{k}\Sigma_{j-1}}\xi_i^-$$
$$= \ell(\Sigma_k)\sum_{i\in\Sigma\setminus\bigcup_{j=1}^{k}\Sigma_{j-1}}\xi_i^- = \ell(\Sigma_k)\Big(\sum_{i\in\Sigma}\xi_i^- - \sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\xi_i^-\Big).$$

Hence

$$\sum_{i\in\Sigma}\ell_i\xi_i^- - \sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\ell_i\xi_i^- \geq \ell(\Sigma_k)\Big(\frac{\alpha}{2} - \sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\mu_i\Big)$$

which implies

$$\sum_{i\in\Sigma}\ell_i\xi_i^- \geq \ell(\Sigma_k)\Big(\frac{\alpha}{2}-\sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\mu_i\Big)+\sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\ell_i\mu_i.$$

Next, we show under the stated conditions that equality holds.

$$\begin{aligned}\sum_{i\in\Sigma}\ell_i\xi_i^- &= \sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\ell_i\xi_i^- + \sum_{i\in\Sigma_k}\ell_i\xi_i^- + \sum_{i\in\Sigma\setminus\cup_{j=0}^{k}\Sigma_j}\ell_i\xi_i^-\\ &= \sum_{j=1}^{k}\ell(\Sigma_{j-1})\sum_{i\in\Sigma_{j-1}}\xi_i^- + \ell(\Sigma_k)\sum_{i\in\Sigma_k}\xi_i^-\\ &= \sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\ell_i\mu_i + \ell(\Sigma_k)\Big(\frac{\alpha}{2}-\sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\mu_i\Big).\end{aligned}$$

■

Proof of Theorem 3.1. From Lemma 3.1, and Corollary 3.1, we know that for $R \leq R_{\max}$, where $R_{\max} = 2(1-\mu(\Sigma^0))$, the total variation constraint holds with equality, that is, $||\xi||_{TV} = R$. Let $\alpha = ||\xi||_{TV}$. From (3.39) and (3.40), Problem 3.1 is given by

$$D^+(R) = \sum_{i\in\Sigma}\ell_i\mu_i + \max_{\xi\in\widetilde{\mathbf{B}}_R(\mu)}\sum_{i\in\Sigma}\ell_i\xi_i. \tag{3.53}$$

where $\xi \in \widetilde{\mathbf{B}}_R(\mu)$ is described by the constraints

$$\alpha \triangleq \sum_{i\in\Sigma}|\xi_i| = R, \quad \sum_{i\in\Sigma}\xi_i = 0, \quad 0 \leq \xi_i+\mu_i \leq 1, \quad \forall i\in\Sigma. \tag{3.54}$$

To maximize (3.53) we employ (3.44). It is obvious that an upper and a lower bound must be obtained for $\sum_{i\in\Sigma}\ell_i\xi_i^+$ and $\sum_{i\in\Sigma}\ell_i\xi_i^-$, respectively.

From Lemma 3.5, Part (a), the upper bound (3.47), holds with equality if conditions given by (3.48) are satisfied. Note that, $\sum_{i\in\Sigma^0}\mu_i + \frac{\alpha}{2} \leq 1$ is always satisfied and from (3.48b) we have that $\sum_{i\in\Sigma^0}\nu_i = \sum_{i\in\Sigma^0}\mu_i + \frac{\alpha}{2}$ and hence the optimal probability on Σ^0 is given by

$$\nu^*(\Sigma^0) \triangleq \sum_{i\in\Sigma^0}\nu_i^* = \sum_{i\in\Sigma^0}\mu_i + \frac{\alpha}{2}. \tag{3.55}$$

From Lemma 3.5, Part (b), case 1, the lower bound (3.49), holds with equality if conditions given by (3.50) are satisfied. Furthermore, from (3.50b) we have that $\sum_{i\in\Sigma_0}\nu_i = \sum_{i\in\Sigma_0}\mu_i - \frac{\alpha}{2}$ and condition (3.50a) must be satisfied, hence the optimal probability on Σ_0 is given by

$$\nu^*(\Sigma_0) \triangleq \sum_{i\in\Sigma_0}\nu_i^* = \Big(\sum_{i\in\Sigma_0}\mu_i - \frac{\alpha}{2}\Big)^+. \tag{3.56}$$

The extremum solution for any $R \leq R_{\max}$, under equality conditions (3.48) and (3.50) is given by

$$D^+(R) = \Big\{\ell_{\max} - \ell_{\min}\Big\}\frac{\alpha}{2} + \sum_{i\in\Sigma}\ell_i\mu_i. \tag{3.57}$$

Lemma 3.5, Part (b), case 1, characterize the extremum solution for $\sum_{i\in\Sigma_0}\mu_i - \frac{\alpha}{2} \geq 0$. Next, the characterization of extremum solution when this condition is violated, that is, when $\sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\mu_i - \frac{\alpha}{2} \leq 0$ for any $k \in \{1, 2, \dots, r\}$, is discussed.
From Lemma 3.5, Part (b), case 2, the lower bound (3.51), holds with equality if conditions given by (3.52) are satisfied. Furthermore, from (3.52b) we have that

$$\sum_{i\in\Sigma_k}\nu_i = \sum_{i\in\Sigma_k}\mu_i - \Big(\frac{\alpha}{2} - \sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\mu_i\Big), \tag{3.58}$$

and conditions $\frac{\alpha}{2} - \sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\mu_i \geq 0$ and (3.52c) must be satisfied, hence the optimal probability on Σ_k is given by

$$\nu^*(\Sigma_k) \triangleq \sum_{i\in\Sigma_k}\nu_i^* = \Big(\sum_{i\in\Sigma_k}\mu_i - \Big(\frac{\alpha}{2} - \sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\mu_i\Big)^+\Big)^+. \tag{3.59}$$

The extremum solution for any $R \leq R_{\max}$, under equality conditions (3.48) and (3.52) is given by

$$\begin{aligned} D^+(R) &= \sum_{i\in\Sigma}\ell_i\xi_i^+ - \sum_{i\in\Sigma}\ell_i\xi_i^- + \sum_{i\in\Sigma}\ell_i\mu_i \\ &= \ell_{\max}\Big(\frac{\alpha}{2}\Big) - \ell(\Sigma_k)\Big(\frac{\alpha}{2} - \sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\mu_i\Big) + \sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\ell_i\mu_i + \sum_{i\in\Sigma}\ell_i\mu_i. \end{aligned}$$

For $R \in [R_{\max}, 2]$, Lemma 3.1, states that $D^+(R)$ is constant. Indeed for $\alpha = ||\xi||_{TV} = R_{\max} = 2(1 - \mu(\Sigma^0))$ equality conditions of Lemma 3.5, Part (a), become

$$\sum_{i\in\Sigma^0}\mu_i + \frac{\alpha}{2} = 1, \quad \sum_{i\in\Sigma^0}\xi_i^+ = \frac{\alpha}{2}, \quad \xi_i^+ = 0 \text{ for } i \in \Sigma\setminus\Sigma^0. \tag{3.60}$$

and hence

$$\sum_{i\in\Sigma\setminus\Sigma^0}\mu_i - \frac{\alpha}{2} = 0, \quad \sum_{i\in\Sigma\setminus\Sigma^0}\xi_i^- = \frac{\alpha}{2}, \quad \xi_i^- = 0 \text{ for } i \in \Sigma^0. \tag{3.61}$$

Therefore, $\sum_{i\in\Sigma\setminus\Sigma^0}\xi_i^- = \sum_{i\in\Sigma\setminus\Sigma^0}\mu_i$ and hence $\xi_i^- = \mu_i$ for all $i \in \Sigma\setminus\Sigma^0$. The extremum solution for any $R \in [R_{\max}, 2]$ is given by

$$\begin{aligned} D^+(R) &= \sum_{i\in\Sigma}\ell_i\xi_i^+ - \sum_{i\in\Sigma}\ell_i\xi_i^- + \sum_{i\in\Sigma}\ell_i\mu_i \overset{(a)}{=} \sum_{i\in\Sigma^0}\ell_i\xi_i^+ - \sum_{i\in\Sigma\setminus\Sigma^0}\ell_i\xi_i^- + \sum_{i\in\Sigma}\ell_i\mu_i \\ &= \ell_{\max}\Big(\frac{\alpha}{2}\Big) - \sum_{i\in\Sigma\setminus\Sigma^0}\ell_i\mu_i + \sum_{i\in\Sigma}\ell_i\mu_i = \ell_{\max}\Big(1 - \sum_{i\in\Sigma^0}\mu_i\Big) + \sum_{i\in\Sigma^0}\ell_i\mu_i = \ell_{\max} \end{aligned}$$

where (a) follows from (3.60) and (3.61). ■

3.3.2. $R^+(D)$: The Finite Alphabet Case

Consider the finite alphabet version of Problem 3.2, that is

$$R^+(D) \triangleq \max_{\nu\in\mathbf{Q}(D)} ||\nu-\mu||_{TV} = \max_{\nu\in\mathbb{P}(\Sigma):\sum_{i\in\Sigma}\ell_i\nu_i\leq D} \sum_{i\in\Sigma}|\nu_i-\mu_i|. \tag{3.62}$$

The solution of (3.62) is obtained from the solution of Problem 3.1, by finding the inverse mapping or by following a similar procedure to the one utilized to derive Theorem 3.1. Below, we give the main theorem.

Theorem 3.2. *The solution of the finite alphabet version of (3.62) is given by*

$$R^+(D) = \sum_{i\in\Sigma}|\nu_i^*-\mu_i|, \tag{3.63}$$

where the value of $R^+(D)$ is calculated as follows.

(1) If

$$D \geq \ell_{\max}\left(\sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\mu_i + \sum_{i\in\Sigma^0}\mu_i\right) + \sum_{j=k}^{r}\sum_{i\in\Sigma_j}\ell_i\mu_i \tag{3.64a}$$

and

$$D \leq \ell_{\max}\left(\sum_{j=0}^{k}\sum_{i\in\Sigma_j}\mu_i + \sum_{i\in\Sigma^0}\mu_i\right) + \sum_{j=k+1}^{r}\sum_{i\in\Sigma_j}\ell_i\mu_i \tag{3.64b}$$

then

$$R^+(D) = \frac{2\left(D - \ell_{\max}\sum_{i\in\Sigma^0}\mu_i - \ell\left(\Sigma_k\right)\sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\mu_i - \sum_{j=k}^{r}\sum_{i\in\Sigma_j}\ell_i\mu_i\right)}{\ell_{\max}-\ell\left(\Sigma_k\right)}. \tag{3.65}$$

(2) If

$$D \leq \left(\ell_{\max}-\ell_{\min}\right)\sum_{i\in\Sigma_0}\mu_i + \sum_{i\in\Sigma}\ell_i\mu_i \tag{3.66}$$

then

$$R^+(D) = \frac{2\left(D - \sum_{i\in\Sigma}\ell_i\mu_i\right)}{\ell_{\max}-\ell_{\min}} \tag{3.67}$$

The optimal probabilities are given by

$$\nu^*(\Sigma^0) \triangleq \sum_{i\in\Sigma^0} \nu_i^* = \sum_{i\in\Sigma^0} \mu_i + \alpha \tag{3.68a}$$

$$\nu^*(\Sigma_0) \triangleq \sum_{i\in\Sigma_0} \nu_i^* = \left(\sum_{i\in\Sigma_0} \mu_i - \alpha\right)^+ \tag{3.68b}$$

$$\nu^*(\Sigma_k) \triangleq \sum_{i\in\Sigma_k} \nu_i^* = \left(\sum_{i\in\Sigma_k} \mu_i - \left(\alpha - \sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}} \mu_i\right)^+\right)^+ \tag{3.68c}$$

$$\alpha = \min\left(R^+(D), 2(1 - \sum_{i\in\Sigma^0} \mu_i)\right). \tag{3.68d}$$

where $k = 1, 2, \dots, r$ *and* r *is the number of* Σ_k *sets which is at most* $|\Sigma \setminus \Sigma^0 \cup \Sigma_0|$.

Proof. For the derivation of the Theorem see Appendix B.1. ■

3.3.3. Example

This example demonstrates the inverse mapping relation of maximization Problems 3.1 and 3.2. The optimal solution is found by implementing Theorem 3.1 and 3.2 for a fixed value of R and D, respectively.

Let $\Sigma = \{1, 2, 3\}$ and for simplicity consider an ascending sequence of non-negative lengths $\ell : \Sigma \longmapsto [0, \infty)$, with corresponding nominal probability vector $\mu \in \mathcal{M}_1(\Sigma)$. Specifically, let $\ell = \{\ell(1), \ell(2), \ell(3) : \ell(1) = 4, \ell(2) = 6, \ell(3) = 8\}$ and $\mu = \{\mu(1), \mu(2), \mu(3) : \mu(1) = \frac{2}{6}, \mu(2) = \frac{2}{6}, \mu(3) = \frac{2}{6}\}$. The sets which correspond to the maximum, minimum and the remaining length are equal to $\Sigma^0 = \{3\}$, $\Sigma_0 = \{1\}$ and $\Sigma_1 = \{2\}$, respectively. We proceed first with the solution of Problem 3.1.

Let $R = \frac{1}{3}$. By implementing (3.46) we have that, $\alpha = \min\left(\frac{R}{2}, 1 - \mu(3)\right) = \min\left(\frac{1}{6}, 1 - \frac{2}{6}\right) = \frac{1}{6}$, and the optimal probabilities are given by

$$\nu^*\left(\Sigma^0\right) = \mu(3) + \alpha = \frac{2}{6} + \frac{1}{6} = \frac{3}{6}, \qquad \nu^*\left(\Sigma_0\right) = (\mu(1) - \alpha)^+ = \left(\frac{2}{6} - \frac{1}{6}\right)^+ = \frac{1}{6}$$

$$\nu^*\left(\Sigma_1\right) = \left(\mu(2) - (\alpha - \mu(1))^+\right)^+ = \left(\frac{2}{6} - \left(\frac{1}{6} - \frac{2}{6}\right)^+\right)^+ = \frac{2}{6}.$$

Hence, for $R = \frac{1}{3}$, the maximum pay-off (3.45), is given by

$$D^+(R) = \ell(3)\nu^*(\Sigma^0) + \ell(1)\nu^*(\Sigma_0) + \ell(2)\nu^*(\Sigma_1) = 8\left(\frac{3}{6}\right) + 4\left(\frac{1}{6}\right) + 6\left(\frac{2}{6}\right) = \frac{40}{6}.$$

Next, we proceed with the solution of Problem 3.2. Let $D = \frac{40}{6}$. By implementing (3.66) we get that,

$$(\ell_{\max} - \ell_{\min}) \sum_{i\in\Sigma_0} \mu_i + \sum_{i\in\Sigma} \ell_i\mu_i = (\ell(3) - \ell(1))\mu(1) + \ell(1)\mu(1) + \ell(2)\mu(2) + \ell(3)\mu(3) = \frac{44}{6} > D$$

and hence, from (3.67)

$$R^{+}(D)=\frac{2(D-\sum_{i\in\Sigma}\ell_i\mu_i)}{\ell_{\max}-\ell_{\min}}=\frac{2(\frac{40}{6}-\frac{36}{6})}{8-4}=\frac{1}{3}.$$

It is clear that the optimal probabilities given by (3.68) are equal to the ones already calculated for Problem 3.1.

3.3.4. $D^{-}(R)$: The Finite Alphabet Case

Consider the finite alphabet version of Problem 3.3, that is

$$D^{-}(R)\triangleq\min_{\nu\in\mathbf{B}_R(\mu)}\sum_{i\in\Sigma}\ell_i\nu_i=\min_{\nu\in\mathbb{P}(\Sigma):\sum_{i\in\Sigma}|\nu_i-\mu_i|\leq R}\sum_{i\in\Sigma}\ell_i\nu_i. \tag{3.69}$$

The solution of (3.69) is obtained from that of Problem 3.1, but with a reverse computation on the partition of Σ and the mass of the extremum measure on the partition moving in the opposite direction. Problem 3.3 can be reformulated as follows.

$$\min_{\nu\in\mathbf{B}_R(\mu)}\sum_{i\in\Sigma}\ell_i\nu_i\longrightarrow\sum_{i\in\Sigma}\ell_i\mu_i+\min_{\xi\in\widetilde{\mathbf{B}}_R(\mu)}\sum_{i\in\Sigma}\ell_i\xi_i. \tag{3.70}$$

Note that $\xi\in\widetilde{\mathbf{B}}_R(\mu)$ is described by the constraints

$$\alpha\triangleq\sum_{i\in\Sigma}|\xi_i|\leq R,\quad\sum_{i\in\Sigma}\xi_i=0,\quad 0\leq\xi_i+\mu_i\leq 1,\quad\forall i\in\Sigma. \tag{3.71}$$

Below, we give the main theorem.

Theorem 3.3. *The solution of the finite alphabet version of (3.69) is given by*

$$D^{-}(R)=\ell_{\max}\nu^*(\Sigma^0)+\ell_{\min}\nu^*(\Sigma_0)+\sum_{k=1}^{r}\ell(\Sigma^k)\nu^*(\Sigma^k) \tag{3.72}$$

where the optimal probabilities are given by

$$\nu^*(\Sigma_0)\triangleq\sum_{i\in\Sigma_0}\nu_i^*=\sum_{i\in\Sigma_0}\mu_i+\alpha \tag{3.73a}$$

$$\nu^*(\Sigma^0)\triangleq\sum_{i\in\Sigma^0}\nu_i^*=\Big(\sum_{i\in\Sigma^0}\mu_i-\alpha\Big)^{+} \tag{3.73b}$$

$$\nu^*(\Sigma^k)\triangleq\sum_{i\in\Sigma^k}\nu_i^*=\Big(\sum_{i\in\Sigma^k}\mu_i-\Big(\alpha-\sum_{j=1}^{k}\sum_{i\in\Sigma^{j-1}}\mu_i\Big)^{+}\Big)^{+} \tag{3.73c}$$

$$\alpha=\min(R,R_{\max}),\qquad R_{\max}\triangleq 2(1-\sum_{i\in\Sigma_0}\mu_i) \tag{3.73d}$$

with $k=1,2,\ldots,r$ and r is the number of Σ^k sets which is at most $|\Sigma\setminus\Sigma^0\cup\Sigma_0|$.

Proof. For the derivation of the Theorem see Appendix B.2. ■

3.3.5. $R^-(D)$: The Finite Alphabet Case

In this subsection we provide the solution of Problem 3.4, by following the procedure utilized to derive the solution of Problem 3.2. The extremum problem is defined by

$$R^-(D) \triangleq \min_{\nu\in\mathbf{Q}(D)} ||\nu-\mu||_{TV} = \min_{\nu\in\mathbb{P}(\Sigma):\sum_{i\in\Sigma}\ell_i\nu_i\leq D} \sum_{i\in\Sigma}|\nu_i-\mu_i|. \tag{3.74}$$

The main theorem which characterizes the extremum solution of Problem 3.4 is given below.

Theorem 3.4. *The solution of the finite alphabet version of Problem 3.4 is given by*

$$R^-(D) = \sum_{i\in\Sigma}|\nu_i^*-\mu_i| \tag{3.75}$$

where the value of $R^-(D)$ is calculated as follows.
(1) If

$$D \geq \ell_{\min}\Big(\sum_{j=0}^{k}\sum_{i\in\Sigma^j}\mu_i + \sum_{i\in\Sigma_0}\mu_i\Big) + \sum_{j=k+1}^{r}\sum_{i\in\Sigma^j}\ell_i\mu_i \tag{3.76a}$$

and

$$D \leq \ell_{\min}\Big(\sum_{j=1}^{k}\sum_{i\in\Sigma^{j-1}}\mu_i + \sum_{i\in\Sigma_0}\mu_i\Big) + \sum_{j=k}^{r}\sum_{i\in\Sigma^j}\ell_i\mu_i \tag{3.76b}$$

then

$$R^-(D) = \frac{2\Big(D-\ell_{\min}\sum_{i\in\Sigma_0}\mu_i - \ell(\Sigma^k)\sum_{j=1}^{k}\sum_{i\in\Sigma^{j-1}}\mu_i - \sum_{j=k}^{r}\sum_{i\in\Sigma^j}\ell_i\mu_i\Big)}{\ell_{\min}-\ell(\Sigma^k)}. \tag{3.77}$$

(2) If

$$D \geq (\ell_{\min}-\ell_{\max})\sum_{i\in\Sigma^0}\mu_i + \sum_{i\in\Sigma}\ell_i\mu_i \tag{3.78}$$

then

$$R^-(D) = \frac{2\Big(D-\sum_{i\in\Sigma}\ell_i\mu_i\Big)}{\ell_{\min}-\ell_{\max}}. \tag{3.79}$$

The optimal probabilities are given by

$$\nu^*(\Sigma_0) \triangleq \sum_{i\in\Sigma_0}\nu_i^* = \sum_{i\in\Sigma_0}\mu_i + \frac{\alpha}{2} \tag{3.80a}$$

$$\nu^*(\Sigma^0) \triangleq \sum_{i\in\Sigma^0}\nu_i^* = \Big(\sum_{i\in\Sigma^0}\mu_i - \frac{\alpha}{2}\Big)^+ \tag{3.80b}$$

$$\nu^*(\Sigma^k) \triangleq \sum_{i\in\Sigma^k}\nu_i^* = \Big(\sum_{i\in\Sigma^k}\mu_i - \Big(\frac{\alpha}{2}-\sum_{j=1}^{k}\sum_{i\in\Sigma^{j-1}}\mu_i\Big)^+\Big)^+ \tag{3.80c}$$

$$\alpha = \min\Big(R^-(D), 2(1-\sum_{i\in\Sigma_0}\mu_i)\Big) \tag{3.80d}$$

where $k=1,2,\ldots,r$ and r is the number of Σ^k sets which is at most $|\Sigma\setminus\Sigma^0\cup\Sigma_0|$.

Before we proceed with the proof of Theorem 3.4, we give the following Lemma in which lower and upper bounds, which are achievable, are obtained.

Lemma 3.6. *Consider Problem 3.4.*

(a) Lower Bound.

$$\sum_{i\in\Sigma}\ell_i\xi_i^+ \geq \ell_{\min}\left(\frac{\alpha}{2}\right). \tag{3.81}$$

The bound holds with equality if

$$\sum_{i\in\Sigma_0}\mu_i + \frac{\alpha}{2} \leq 1 \tag{3.82a}$$

$$\sum_{i\in\Sigma_0}\xi_i^+ = \frac{\alpha}{2} \tag{3.82b}$$

$$\xi_i^+ = 0 \;\; \textit{for} \;\; i \in \Sigma\setminus\Sigma_0. \tag{3.82c}$$

(b) Upper Bound.

Case 1. If $\sum_{i\in\Sigma^0}\mu_i - \frac{\alpha}{2} \geq 0$ *then*

$$\sum_{i\in\Sigma}\ell_i\xi_i^- \leq \ell_{\max}\left(\frac{\alpha}{2}\right). \tag{3.83}$$

The bound holds with equality if

$$\sum_{i\in\Sigma^0}\mu_i - \frac{\alpha}{2} \geq 0 \tag{3.84a}$$

$$\sum_{i\in\Sigma^0}\xi_i^- = \frac{\alpha}{2} \tag{3.84b}$$

$$\xi_i^- = 0 \;\; \textit{for} \;\; i \in \Sigma\setminus\Sigma^0. \tag{3.84c}$$

Case 2. If $\sum_{j=1}^{k}\sum_{i\in\Sigma^{j-1}}\mu_i - \frac{\alpha}{2} \leq 0$ *for any* $k \in \{1,2,\ldots,r\}$ *then:*

$$\sum_{i\in\Sigma}\ell_i\xi_i^- \leq \ell(\Sigma^k)\left(\frac{\alpha}{2} - \sum_{j=1}^{k}\sum_{i\in\Sigma^{j-1}}\mu_i\right) + \sum_{j=1}^{k}\sum_{i\in\Sigma^{j-1}}\ell_i\mu_i. \tag{3.85}$$

Moreover, equality holds if

$$\sum_{i\in\Sigma^{j-1}}\xi_i^- = \sum_{i\in\Sigma^{j-1}}\mu_i \;\; \textit{for all} \;\; j = 1,2,\ldots,k, \tag{3.86a}$$

$$\sum_{i\in\Sigma^k}\xi_i^- = \left(\frac{\alpha}{2} - \sum_{j=1}^{k}\sum_{i\in\Sigma^{j-1}}\mu_i\right) \tag{3.86b}$$

$$\sum_{j=0}^{k}\sum_{i\in\Sigma^j}\mu_i - \frac{\alpha}{2} \geq 0 \tag{3.86c}$$

$$\xi_i^- = 0 \;\; \textit{for all} \;\; i \in \Sigma\setminus\Sigma^0\cup\Sigma_1\cup\ldots\cup\Sigma^k. \tag{3.86d}$$

Proof. Part (a) and Part (b), case 1, follows from Section 3.2.2 (equivalent formulation of $R^{-}(D)$). The proof of Part (b), case 2, is similar to the proof given for Lemma 3.5, Part (b), case 2, with appropriate changes on Σ^k sets. ■

Proof of Theorem 3.4. From Lemma 3.4, and Corollary 3.4, we know that for $D \leq D_{\max}$, where $D_{\max} = \sum_{i\in\Sigma} \ell_i \mu_i$, the average constraint holds with equality, that is

$$\sum_{i\in\Sigma} \ell_i \nu_i = \sum_{i\in\Sigma} \ell_i \xi_i^{+} - \sum_{i\in\Sigma} \ell_i \xi_i^{-} + \sum_{i\in\Sigma} \ell_i \mu_i = D.$$

From Lemma 3.6, Part (a) and from Part (b), case 1, when equality conditions (3.82) and (3.84) are satisfied we have that

$$\ell_{\min}\left(\frac{\alpha}{2}\right) - \ell_{\max}\left(\frac{\alpha}{2}\right) + \sum_{i\in\Sigma} \ell_i \mu_i = D.$$

Solving the above equation with respect to α we get that

$$\alpha = \frac{2\left(D - \sum_{i\in\Sigma} \ell_i \mu_i\right)}{\ell_{\min} - \ell_{\max}}. \tag{3.87}$$

Since (3.82a) is always satisfied, it remains to ensure that (3.84a) is also satisfied. By substituting (3.87) into (3.84a) and solving with respect to D we get that if $D \geq (\ell_{\min} - \ell_{\max}) \sum_{i\in\Sigma^0} \mu_i + \sum_{i\in\Sigma} \ell_i \mu_i$ then $R^{-}(D)$ is given by (3.79). Moreover, the optimal probabilities given by (3.80a) and (3.80b) are obtained from (3.82b) and (3.84b), respectively.

Lemma 3.6, Part (b), case 1, characterize the extremum solution for $\sum_{i\in\Sigma^0} \mu_i - \frac{\alpha}{2} \geq 0$. Next, the characterization of extremum solution when this condition is violated, that is, when $\sum_{j=1}^{k} \sum_{i\in\Sigma^{j-1}} \mu_i - \frac{\alpha}{2} \leq 0$ for any $k \in \{1, 2, \ldots, r\}$, is discussed.

From Lemma 3.6, Part (b), case 2, the upper bound (3.85), holds with equality if conditions given by (3.86) are satisfied. Hence,

$$\ell_{\min}\left(\frac{\alpha}{2}\right) - \ell(\Sigma^k)\left(\frac{\alpha}{2} - \sum_{j=1}^{k} \sum_{i\in\Sigma^{j-1}} \mu_i\right) + \sum_{j=1}^{k} \sum_{i\in\Sigma^{j-1}} \ell_i \mu_i = D.$$

Solving the above equation with respect to α we get that

$$\alpha = \frac{2\Big(D - \ell_{\min} \sum_{i\in\Sigma_0} \mu_i - \ell(\Sigma^k) \sum_{j=1}^{k} \sum_{i\in\Sigma^{j-1}} \mu_i - \sum_{j=k}^{r} \sum_{i\in\Sigma^j} \ell_i \mu_i\Big)}{\ell_{\min} - \ell(\Sigma^k)}. \tag{3.88}$$

Substituting (3.88) into $\sum_{j=1}^{k} \sum_{i\in\Sigma^{j-1}} \mu_i - \frac{\alpha}{2} \leq 0$ and into (3.86c) and solving with respect to D we get that if

$$D \geq \ell_{\min}\left(\sum_{j=0}^{k} \sum_{i\in\Sigma^j} \mu_i + \sum_{i\in\Sigma_0} \mu_i\right) + \sum_{j=k+1}^{r} \sum_{i\in\Sigma^j} \ell_i \mu_i$$

$$D \leq \ell_{\min}\left(\sum_{j=1}^{k} \sum_{i\in\Sigma^{j-1}} \mu_i + \sum_{i\in\Sigma_0} \mu_i\right) + \sum_{j=k}^{r} \sum_{i\in\Sigma^j} \ell_i \mu_i$$

then $R^-(D)$ is given by (3.77). Moreover, the optimal probability on Σ^k given by (3.80c) is obtained from (3.86b).

For $D \in [D_{\max}, \infty)$, is straightforward that, the extremum measure is given by $\nu^* = \mu$ and hence $R^-(D) = 0$. ■

3.3.6. Example

This example demonstrates the inverse mapping relation of minimization Problems 3.3 and 3.4. The optimal solution is found by implementing Theorem 3.3 and 3.4 for a fixed value of R and D, respectively.

Let $\Sigma = \{1, 2, 3\}$ and for simplicity consider an ascending sequence of non-negative lengths $\ell : \Sigma \longmapsto [0, \infty)$, with corresponding nominal probability vector $\mu \in \mathcal{M}_1(\Sigma)$. Specifically, let $\ell = \{\ell(1), \ell(2), \ell(3) : \ell(1) = 4, \ell(2) = 6, \ell(3) = 8\}$ and $\mu = \{\mu(1), \mu(2), \mu(3) : \mu(1) = \frac{2}{6}, \mu(2) = \frac{2}{6}, \mu(3) = \frac{2}{6}\}$. The sets which correspond to the maximum, minimum and the remaining length are equal to $\Sigma^0 = \{3\}$, $\Sigma_0 = \{1\}$ and $\Sigma^1 = \{2\}$, respectively. We proceed first with the solution of Problem 3.3.

Let $R = \frac{1}{3}$. By implementing (3.73) we have that, $\alpha = \min\left(\frac{R}{2}, 1 - \mu(1)\right) = \min\left(\frac{1}{6}, 1 - \frac{2}{6}\right) = \frac{1}{6}$, and the optimal probabilities are given by

$$\nu^*(\Sigma_0) = \mu(1) + \alpha = \frac{2}{6} + \frac{1}{6} = \frac{3}{6}, \qquad \nu^*\left(\Sigma^0\right) = (\mu(3) - \alpha)^+ = \left(\frac{2}{6} - \frac{1}{6}\right)^+ = \frac{1}{6}$$

$$\nu^*\left(\Sigma^1\right) = \left(\mu(2) - (\alpha - \mu(3))^+\right)^+ = \left(\frac{2}{6} - \left(\frac{1}{6} - \frac{2}{6}\right)^+\right)^+ = \frac{2}{6}.$$

Hence, for $R = \frac{1}{3}$, the maximum pay-off (3.72), is given by

$$D^-(R) = \ell(3)\nu^*(\Sigma^0) + \ell(1)\nu^*(\Sigma_0) + \ell(2)\nu^*(\Sigma^1) = 8\left(\frac{1}{6}\right) + 4\left(\frac{3}{6}\right) + 6\left(\frac{2}{6}\right) = \frac{32}{6}.$$

Next, we proceed with the solution of Problem 3.3. Let $D = \frac{32}{6}$. By implementing (3.78) we get that,

$$(\ell_{\min} - \ell_{\max}) \sum_{i \in \Sigma^0} \mu_i + \sum_{i \in \Sigma} \ell_i \mu_i = (\ell(3) - \ell(1))\mu(3) + \ell(1)\mu(1) + \ell(2)\mu(2) + \ell(3)\mu(3) = \frac{28}{6} < D$$

and hence, from (3.79)

$$R^+(D) = \frac{2(D - \sum_{i \in \Sigma} \ell_i \mu_i)}{\ell_{\min} - \ell_{\max}} = \frac{2(\frac{32}{6} - \frac{36}{6})}{4 - 8} = \frac{1}{3}.$$

It is clear that the optimal probabilities given by (3.80) are equal to the ones already calculated for Problem 3.3.

3.3.7. Extension to Countable Alphabets

The statements of Theorems 3.1, 3.2, 3.3 and 3.4 are also valid for the countable alphabet case, because their derivations are not restricted to Σ being finite alphabet.
It also holds for any $\ell \in BC^+(\Sigma)$ as seen in Section 3.2.

The extensions of Theorems 3.1-3.4 to $\ell \in L^{\infty,+}(\Sigma, \mathcal{B}(\Sigma), \nu)$ can be shown as well; for example, $D^+(R)$ is given by

$$D^+(R) = \ell_{\max}\nu^*(\Sigma^0) + \ell_{\min}\nu^*(\Sigma_0) + \sum_{k=1}^{r} \ell(\Sigma_k)\nu^*(\Sigma_k) \tag{3.89}$$

where the optimal probabilities are given by

$$\begin{aligned}
&\nu^*(\Sigma^0) = \mu(\Sigma^0) + \frac{\alpha}{2}, \qquad \nu^*(\Sigma_0) = \Big(\mu(\Sigma_0) - \frac{\alpha}{2}\Big)^+ \\
&\nu^*(\Sigma_k) = \bigg(\mu(\Sigma_k) - \Big(\frac{\alpha}{2} - \sum_{j=1}^{k}\mu(\Sigma_{j-1})\Big)^+\bigg)^+ \\
&\alpha = \min\left(R, R_{\max}\right), \qquad R_{\max} \triangleq 2(1-\mu(\Sigma^0))
\end{aligned}$$

k is at most countable. We outline the main steps of the derivation. For any $n \in \mathbb{N}$, $\ell \in BC^+(\Sigma)$ define $\ell_n \triangleq \ell \wedge n$ (i.e., the minimum between ℓ and n), then $\ell_n \in BC^+(\Sigma)$, and for any $\nu \in \mathbb{B}_R(\mu)$ we have

$$\sup_{\nu \in \mathbf{B}_R(\mu)} \int_\Sigma \ell_n(x)\nu(dx) = \frac{R}{2}\left\{\sup_{x\in\Sigma}\ell_n(x) - \inf_{x\in\Sigma}\ell_n(x)\right\} + \int_\Sigma \ell_n(x)\mu(dx).$$

For any $\nu \in \mathbf{B}_R(\mu)$, we obtain the following

$$\begin{aligned}
\int_\Sigma \ell(x)\nu(dx) &= \sup_{n\in\mathbb{N}} \int_\Sigma \ell_n(x)\nu(dx) \\
&\le \sup_{n\in\mathbb{N}} \sup_{\nu\in\mathbf{B}_R(\mu)} \int_\Sigma \ell_n(x)\nu(dx) \\
&= \sup_{n\in\mathbb{N}} \left\{\frac{R}{2}\left(\sup_{x\in\Sigma}\ell_n(x) - \inf_{x\in\Sigma}\ell_n(x)\right) + \int_\Sigma \ell_n(x)\mu(dx)\right\} \\
&\le \sup_{n\in\mathbb{N}} \left\{\frac{R}{2}\left(\sup_{x\in\Sigma}\ell_n(x) - \inf_{x\in\Sigma}\ell_n(x)\right)\right\} + \sup_{n\in\mathbb{N}}\int_\Sigma \ell_n(x)\mu(dx) \\
&= \sup_{n\in\mathbb{N}} \left\{\frac{R}{2}\left(\sup_{x\in\Sigma}\ell_n(x) - \inf_{x\in\Sigma}\ell_n(x)\right)\right\} + \int_\Sigma \ell(x)\mu(dx).
\end{aligned}$$

Hence,

$$\sup_{\nu\in\mathbf{B}_R(\mu)} \int_\Sigma \ell(x)\nu(dx) \le \frac{R}{2}\sup_{n\in\mathbb{N}}\left\{\sup_{x\in\Sigma}\ell_n(x) - \inf_{x\in\Sigma}\ell_n(x)\right\} + \int_\Sigma \ell(x)\mu(dx).$$

Next, we show the reverse inequality. For any $\nu \in \mathbf{B}_R(\mu)$, we have that

$$\int_\Sigma \ell(x)\nu(dx) \geq \int_\Sigma \ell_n(x)\nu(dx)$$

and therefore

$$\begin{aligned}\sup_{\nu\in\mathbf{B}_R(\mu)} \int_\Sigma \ell(x)\nu(dx) &\geq \sup_{\nu\in\mathbf{B}_R(\mu)} \int_\Sigma \ell_n(x)\nu(dx)\\ &= \frac{R}{2}\left\{\sup_{x\in\Sigma}\ell_n(x) - \inf_{x\in\Sigma}\ell_n(x)\right\} + \int_\Sigma \ell_n(x)\mu(dx).\end{aligned}$$

Since the above inequality holds for all $n \in \mathbb{N}$, then

$$\begin{aligned}\sup_{\nu\in\mathbf{B}_R(\mu)} \int_\Sigma \ell(x)\nu(dx) &\geq \frac{R}{2}\sup_{n\in\mathbb{N}}\left\{\sup_{x\in\Sigma}\ell_n(x) - \inf_{x\in\Sigma}\ell_n(x)\right\} + \sup_{n\in\mathbb{N}}\int_\Sigma \ell_n(x)\mu(dx)\\ &= \frac{R}{2}\sup_{n\in\mathbb{N}}\left\{\sup_{x\in\Sigma}\ell_n(x) - \inf_{x\in\Sigma}\ell_n(x)\right\} + \int_\Sigma \ell(x)\mu(dx).\end{aligned}$$

Hence,

$$\sup_{\nu\in\mathbf{B}_R(\mu)} \int_\Sigma \ell(x)\nu(dx) = \frac{R}{2}\sup_{n\in\mathbb{N}}\left\{\sup_{x\in\Sigma}\ell_n(x) - \inf_{x\in\Sigma}\ell_n(x)\right\} + \int_\Sigma \ell(x)\mu(dx).$$

Utilizing the fact that $\sup_{n\in\mathbb{N}}\sup_{x\in\Sigma}\ell_n = \sup_n ||\ell_n||_{\infty,\nu}$ ($||\ell||_{\infty,\nu} = \inf_{\Delta\in\mathcal{N}}\sup_{x\in\Delta^c}\ell(x)$, $\mathcal{N} \triangleq \{A \in \mathcal{B}(\Sigma) : \nu(A) = 0\}$, and similarly for the infimum) we obtain the results.

Remark 3.2. *Consider the maximization* $\sup_{\nu\in\mathbf{B}_R(\mu)}\int_\Sigma \ell(x)\nu(dx)$, $\ell \in BC^+(\Sigma)$. *Let* $\mathcal{M}_1^{FS}(\Sigma)$ *denote the set of probability measures on* Σ *with finite support. Since the set of probability measures* $\mathcal{M}_1^{FS}(\Sigma)$ *are dense in* $\mathcal{M}_1(\Sigma)$ *([11, Theorem 4] , see also Appendix B.3), there exists a probability measure* μ^{FS} *with finite support* $\{x_1, \ldots, x_k\}$ *which approximates* $\mu \in \mathcal{M}_1(\Sigma)$ *and since* $\ell \in BC^+(\Sigma)$ *then*

$$\int_\Sigma \ell(x)\nu^*(dx) = \frac{R}{2}\left\{\sup_{x\in\{A_i:i=1,\ldots,k\}}\ell(x) - \inf_{x\in\{A_i:i=1,\ldots,k\}}\ell(x)\right\} + \sum_{j=1}^{k}\ell(x_j)\mu^{FS}(A_j).$$

3.4. Examples

In this section, we illustrate through examples how the optimal solution of the different extremum problems behaves, and in addition, we present an application to the area of information theory. In particular, we present calculations through Example 3.4.1 for $D^+(R)$ and $R^+(D)$, and calculations through Example 3.4.2 for $R^-(D)$ and $D^-(R)$ when the sequence $\ell = \{\ell_1 \;\ \ell_2 \;\ \ldots \;\ \ell_n\} \in \mathbb{R}_+^n$ consists of a number of ℓ_i's which are equal. We further present calculations through Example 3.4.3 for $D^+(R)$, $R^+(D)$ and $D^-(R)$, $R^-(D)$ using a large number of ℓ_i's which are not equal. In Example 3.4.4, the results are applied to universal lossless coding for a class of source distributions maximizing the entropy.

3.4.1. Maximization Problems with a Number of Equal Lengths

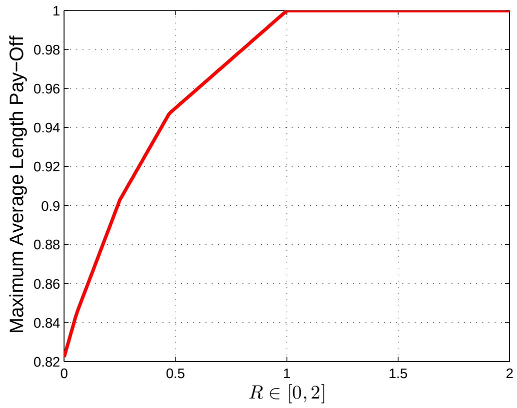
Let $\Sigma = \{i : i = 1, 2, \ldots, 8\}$ and for simplicity consider a descending sequence of lengths $\ell = \{\ell \in \mathbb{R}^8_+ : \ell_1 = \ell_2 > \ell_3 = \ell_4 > \ell_5 > \ell_6 = \ell_7 > \ell_8\}$ with corresponding nominal probability vector $\mu \in \mathbb{P}_1(\Sigma)$. Specifically, let $\ell = [1, 1, 0.8, 0.8, 0.6, 0.4, 0.4, 0.2]$, and $\mu = \left[\frac{23}{72}, \frac{13}{72}, \frac{10}{72}, \frac{9}{72}, \frac{8}{72}, \frac{4}{72}, \frac{3}{72}, \frac{2}{72}\right]$. Note that, the sets which correspond to the maximum, minimum and all the remaining lengths are equal to $\Sigma^0 = \{1, 2\}, \Sigma_0 = \{8\}, \Sigma_1 = \{7, 6\}, \Sigma_2 = \{5\}, \Sigma_3 = \{4, 3\}$.

Fig.3.1a depicts the maximum linear functional pay-off subject to total variational constraint, $D^+(R)$, given by Theorem 3.1. Fig.3.1b depicts the maximum total variational pay-off subject to linear functional constraint, $R^+(D)$, given by Theorem 3.2. Recall Lemma 3.1 and Corollary 3.1. Fig.3.1a shows that, $D^+(R)$ is a non-decreasing concave function of R and also that is constant in $[R_{\max}, 2]$, where $R_{\max} = 2\,(1 - \mu(\Sigma^0)) = 1$. Also, from Lemma 3.2 and Corollary 3.2, Fig.3.1b shows that, $R^+(D)$ is a non-decreasing convex function of D and is constant in $[D_{\max}, \infty)$ where $D_{\max} = \ell_{\max} = 1$. Fig.3.1c depicts the optimal probabilities as a function of the total variation parameter R. Note that, the optimal probabilities are the same for both problems.

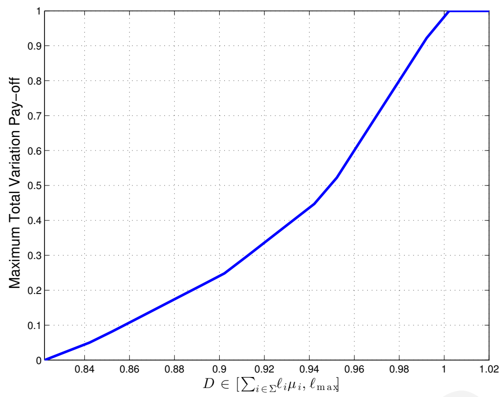
3.4.2. Minimization Problems with a Number of Equal Lengths

Let $\Sigma = \{i : i = 1, 2, \ldots, 8\}$ and for simplicity consider a descending sequence of lengths $\ell = \{\ell \in \mathbb{R}^8_+ : \ell_1 = \ell_2 > \ell_3 = \ell_4 > \ell_5 > \ell_6 = \ell_7 > \ell_8\}$ with corresponding nominal probability vector $\mu \in \mathbb{P}_1(\Sigma)$. Specifically, let $\ell = [1, 1, 0.8, 0.8, 0.6, 0.4, 0.4, 0.2]$ and $\mu = \left[\frac{2}{72}, \frac{3}{72}, \frac{4}{72}, \frac{8}{72}, \frac{9}{72}, \frac{10}{72}, \frac{12}{72}, \frac{24}{72}\right]$. Note that, the sets which correspond to the maximum, minimum and all the remaining lengths are equal to $\Sigma^0 = \{1, 2\}, \Sigma_0 = \{8\}, \Sigma^1 = \{3, 4\}, \Sigma^2 = \{5\}, \Sigma^3 = \{6, 7\}$.

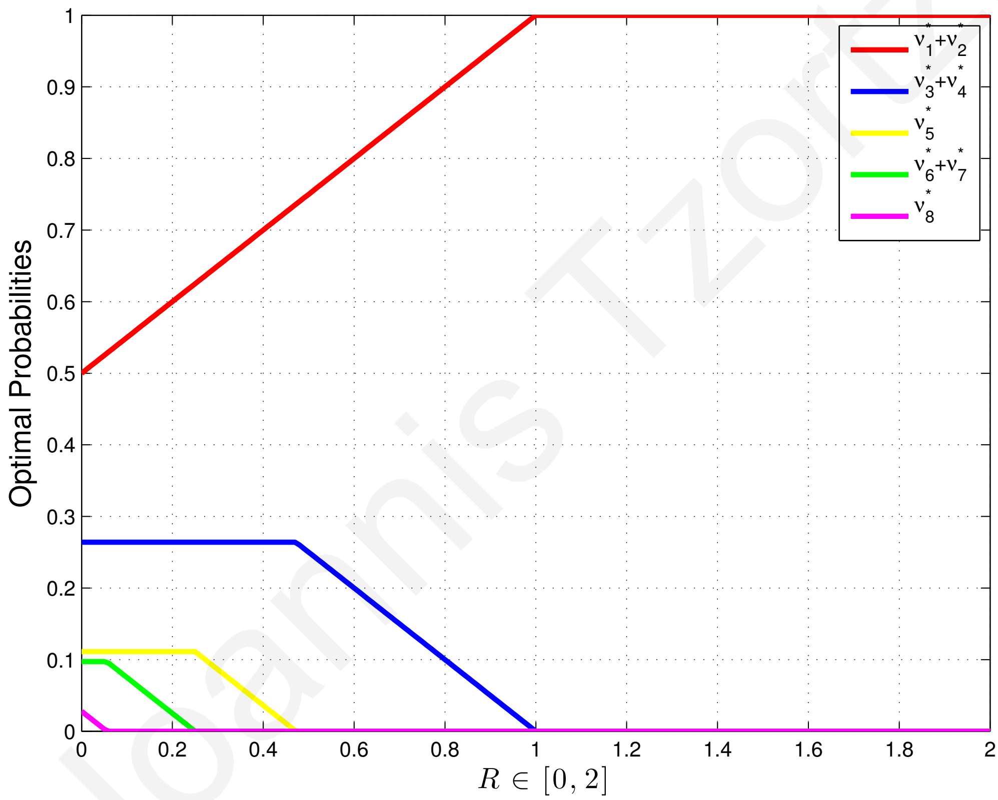
Fig.3.2a depicts the minimum linear functional pay-off subject to total variational constraint, $D^-(R)$, given by Theorem 3.3. Fig.3.2b depicts the minimum total variational pay-off subject to linear functional constraint, $R^-(D)$, given by Theorem 3.4. Recall Lemma 3.3 and Corollary 3.3. Fig.3.2b shows that $D^-(R)$ is non-increasing convex function of R and also that is constant in $[R_{\max}, 2]$, where $R_{\max} = 2\,(1 - \mu(\Sigma_0)) = 1.33$. Also, from Lemma 3.4 and Corollary 3.4, Fig.3.2b shows that, $R^-(D)$ is a non-increasing convex function of D, $D \in [\ell_{\min}, \sum_{i \in \Sigma} \ell_i \mu_i)$. Note that for $D < \ell_{\min} = 0.2$ no solution exists and $R^-(D)$ is zero in $[D_{\max}, \infty)$ where $D_{\max} = \sum_{i=1}^{8} \ell_i \mu_i = 0.73$. Fig.3.2c depicts the optimal probabilities as a function of the total variation parameter R. Note that, the optimal probabilities are the same for both problems.



(a)

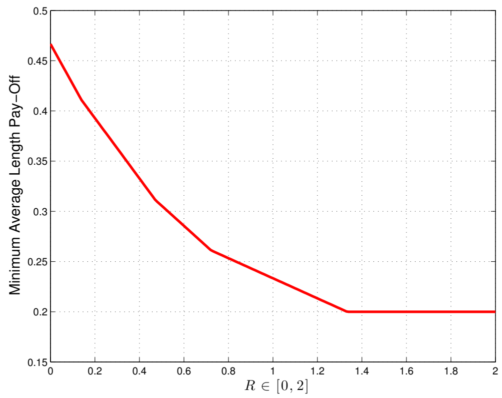


(b)

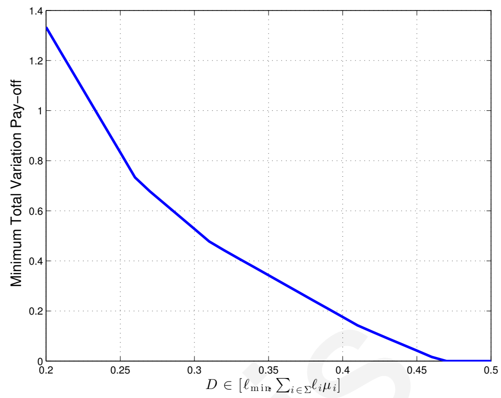


(c)

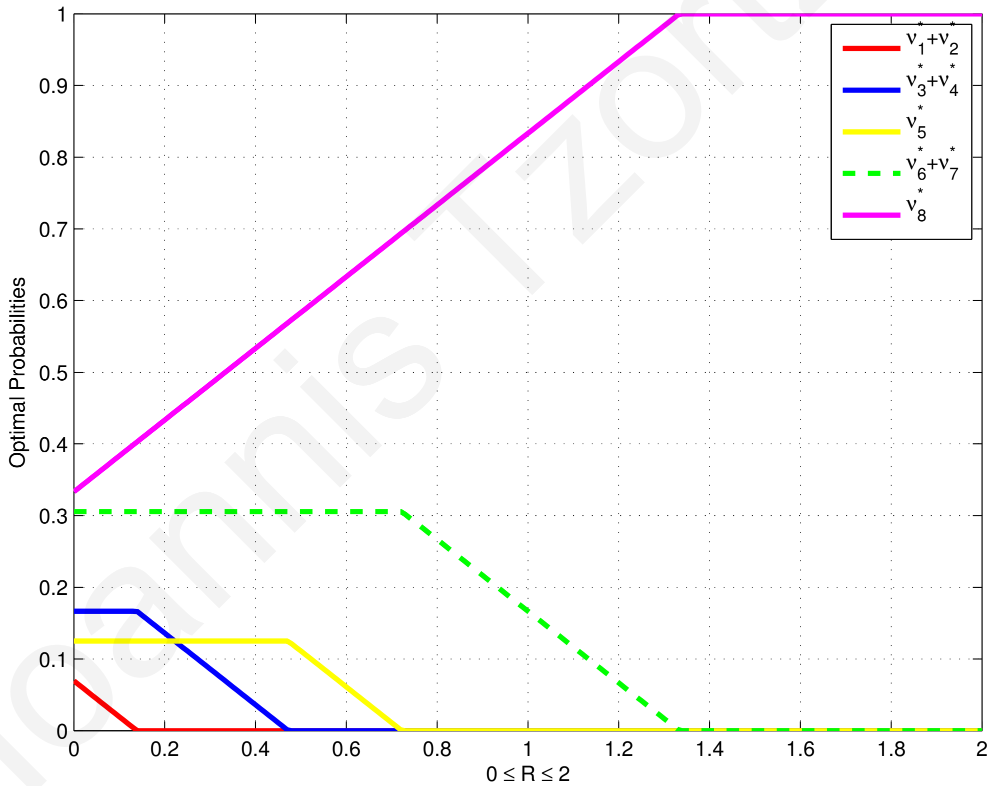
Figure 3.1.: Optimal Solution of Maximization Problems with a Number of Equal Lengths: (a) Linear functional pay-off subject to total variational constraint, $D^+(R)$; (b) Total variational pay-off subject to linear functional constraint, $R^+(D)$; and, (c) Optimal probabilities.



(a)



(b)



(c)

Figure 3.2.: Optimal Solution of Minimization Problems with a Number of Equal Lengths: (a) Linear functional pay-off subject to total variational constraint, $D^{-}(R)$; (b) Total variational pay-off subject to linear functional constraint, $R^{-}(D)$; and, (c) Optimal probabilities.

3.4.3. Extremum Problems with a Large Number of Not Equal Lengths

Let $\Sigma = \{i : i = 1, 2, \dots, 50\}$ and consider a descending sequence of lengths $\ell = \{\ell \in \mathbb{R}_+^{50}\}$ with corresponding nominal probability vector $\mu \in \mathbb{P}_1(\Sigma)$. For display purposes the support sets are denoted by Σ_x^y where $x, y = \{1, 2, \dots, 16\}$, though of course the subscript symbol x corresponds to the support sets of Problem $D^+(R)$, $R^+(D)$ and the superscript symbol y corresponds to the support sets of Problem $D^-(R)$ and $R^-(D)$. Let

$$\ell = \Big[20\ \ 20\ \ 20\ \ 20\ \ 19\ \ 19\ \ 19\ \ 18\ \ 17\ \ 17\ \ 16\ \ 14\ \ 14\ \ 13\ \ 13\ \ 13\ \ 13\ \ 12\ \ 10\ \ 10\ \ 10\ \ 10$$
$$10\ \ 9\ \ 9\ \ 9\ \ 8\ \ 8\ \ 8\ \ 8\ \ 8\ \ 8\ \ 8\ \ 7\ \ 7\ \ 6\ \ 5\ \ 4\ \ 3\ \ 3\ \ 3\ \ 3\ \ 3\ \ 3\ \ 2\ \ 2\ \ 2\ \ 2\ \ 1\ \ 1\Big],$$

and

$$\mu = \Big[26\ \ 1\ \ 5\ \ 3\ \ 2\ \ 19\ \ 16\ \ 14\ \ 13\ \ 4\ \ 6\ \ 5\ \ 4\ \ 13\ \ 25\ \ 22\ \ 15\ \ 16\ \ 12\ \ 5\ \ 10\ \ 15\ \ 7\ \ 12\ \ 2\ \ 3\ \ 12$$
$$5\ \ 11\ \ 6\ \ 8\ \ 21\ \ 7\ \ 8\ \ 5\ \ 12\ \ 10\ \ 4\ \ 7\ \ 16\ \ 9\ \ 6\ \ 5\ \ 20\ \ 18\ \ 9\ \ 1\ \ 11\ \ 6\ \ 8\Big]/500.$$

Note that, the sets which correspond to the maximum, minimum and all the remaining lengths are equal to $\Sigma^0 = \{1-4\}, \Sigma_0 = \{50, 49\}, \Sigma_1^{16} = \{48-45\}, \Sigma_2^{15} = \{44-39\}, \Sigma_3^{14} = \{38\}, , \Sigma_4^{13} = \{37\}, \Sigma_5^{12} = \{36\}, \Sigma_6^{11} = \{35, 34\}, \Sigma_7^{10} = \{33-27\}, \Sigma_8^9 = \{26-24\}, \Sigma_9^8 = \{23-19\}, \Sigma_{10}^7 = \{18\}, \Sigma_{11}^6 = \{17-14\}, \Sigma_{12}^5 = \{13, 12\}, \Sigma_{13}^4 = \{11\}, \Sigma_{14}^3 = \{10-9\}, \Sigma_{15}^2 = \{8\}, \Sigma_{16}^1 = \{7-5\}$.

Fig.3.3a-b depicts the maximum linear functional pay-off subject to total variational constraint, $D^+(R)$, and the maximum total variational pay-off subject to linear functional constraint, $R^+(D)$, given by Theorem 3.1, 3.2, respectively. Fig.3.3c-d depicts the minimum linear functional pay-off subject to total variational constraint, $D^-(R)$, and the minimum total variational pay-off subject to linear functional constraint, $R^-(D)$, given by Theorem 3.3, 3.4 respectively.

3.4.4. Variable Length Lossless Coding for a Variation Distance Class

In this section we illustrate an application of Problem 3.1 to the well-known lossless compression problem of finding uniquely decodable codes, which minimize the average codeword length, knowing as Shannon codes [17]. However, instead of designing codes for a simple probability distribution, we design codes for a class of probability distributions, also known as universal codes [4].

Given a fixed nominal distribution $\mu \in \mathbb{P}(\Sigma)$ and distance parameter $R \in [0, 2]$, define the average codeword length pay-off with respect to the true source probability distribution

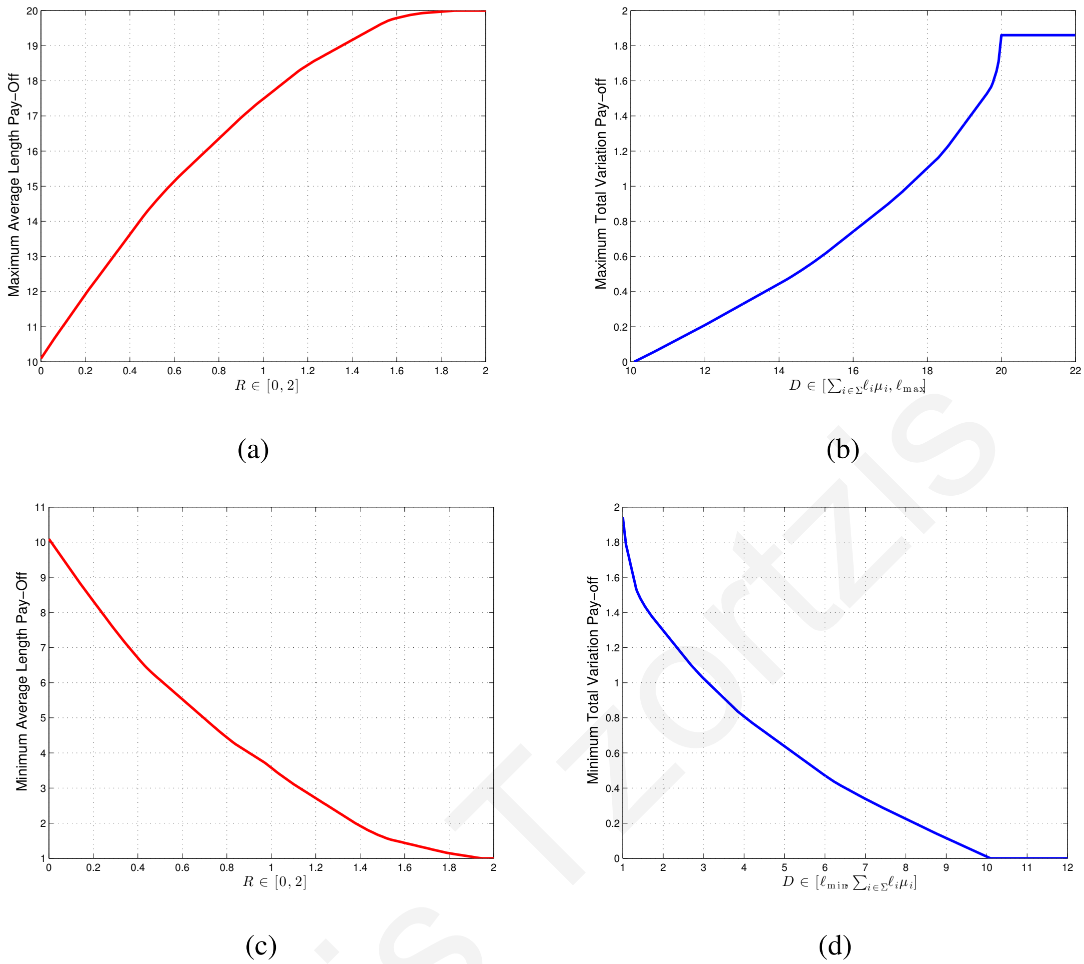


Figure 3.3.: Optimal Solution of Maximization and Minimization Problems with a Number of Not Equal Lengths: (a) Linear functional pay-off subject to total variational constraint, $D^+(R)$; (b) Total variational pay-off subject to linear functional constraint, $R^+(D)$; (c) Linear functional pay-off subject to total variational constraint, $D^-(R)$; and, (d) Total variational pay-off subject to linear functional constraint, $R^-(D)$.

$\nu \in \mathbb{B}_R(\mu) \subset \mathbb{P}(\Sigma)$ by

$$\mathbb{L}_1(\ell, \nu) \triangleq \sum_{i\in\Sigma} \ell_i \nu_i. \tag{3.90}$$

The objective is to find a prefix code length vector $\ell^* \in \mathbb{R}_+^{|\Sigma|}$, satisfying Kraft inequality, $\sum_{i\in\Sigma} D^{-\ell_i}$, where D corresponds to a D-ary code alphabet [17], which minimizes the maximum average codeword length pay-off defined by

$$\mathbb{L}_1(\ell, \nu^*) \triangleq \max_{\nu \in \mathbb{B}_R(\mu)} \sum_{i\in\Sigma} \ell_i \nu_i \tag{3.91}$$

for all $R \in [0,2]$. By Theorem 3.1, (3.91) is equivalent to

$$\mathbb{L}_1(\ell,\nu^*) = \frac{R}{2}\{\max_{i\in\Sigma}\ell_i - \min_{i\in\Sigma}\ell_i\} + \sum_{i\in\Sigma}\ell_i\mu_i.$$

Moreover, by introducing Lagrange multipliers, the problem can be expressed as follows.

$$\max_s \min_t \min_l \Big\{\alpha(t-s) + \sum_{i\in\Sigma}\ell_i\mu_i\Big\}, \qquad \alpha \triangleq \frac{R}{2} \tag{3.92}$$

subject to the Kraft inequality and the constraints $\ell_i \leq t\ \forall i \in \Sigma$ and $\ell_i \geq s$, $\forall i \in \Sigma$. By introducing the real-valued Lagrange multipliers λ_i associated with the constraint $\ell_i \leq t$, $\forall i \in \Sigma$, σ_i associated with the constraint $\ell_i \geq s$, $\forall i \in \Sigma$, and a real-valued Lagrange multiplier τ associate with the Kraft inequality, the augmented pay-off is defined by

$$\mathbb{L}^\alpha(\ell,\mu,\lambda,\sigma,\tau) \triangleq \alpha(t-s) + \sum_{i\in\Sigma}\ell_i\mu_i + \tau\Big(\sum_{i\in\Sigma}D^{-\ell_i} - 1\Big) + \sum_{i\in\Sigma}\lambda_i(\ell_i - t) + \sum_{i\in\Sigma}\sigma_i(s-\ell_i)\ .$$

The augmented pay-off is a convex and differentiable function with respect to ℓ, t and s. Denote the real-valued minimization over $\ell, t, s, \lambda, \sigma, \tau$ by $\ell^*, t^*, s^*, \lambda^*$, σ^* and τ^*. By the Karush-Kuhn-Tucker theorem, the following conditions are necessary and sufficient for optimality.

$$\begin{aligned}
\frac{\partial}{\partial \ell_i}\mathbb{L}^\alpha(\ell,\mu,t,s,\lambda,\sigma,\tau)|_{\ell=\ell^*,\lambda=\lambda^*,t=t^*,s=s^*,\sigma=\sigma^*,\tau=\tau^*} &= 0, \quad \forall i \in \Sigma,\\
\frac{\partial}{\partial t}\mathbb{L}^\alpha(\ell,\mu,t,s,\lambda,\sigma,\tau)|_{\ell=\ell^*,\lambda=\lambda^*,t=t^*,s=s^*,\sigma=\sigma^*,\tau=\tau^*} &= 0,\\
\frac{\partial}{\partial s}\mathbb{L}^\alpha(\ell,\mu,t,s,\lambda,\sigma,\tau)|_{\ell=\ell^*,\lambda=\lambda^*,t=t^*,s=s^*,\sigma=\sigma^*,\tau=\tau^*} &= 0,\\
\sum_{i\in\Sigma}D^{-\ell_i^*} - 1 &\leq 0,\\
\tau^*\cdot\Big(\sum_{i\in\Sigma}D^{-\ell_i^*} - 1\Big) &= 0, \quad \tau^* \geq 0,\\
\ell_i^* - t^* &\leq 0,\\
\lambda_i^*\cdot(\ell_i^* - t^*) &= 0, \quad \lambda_i^* \geq 0, \quad \forall i \in \Sigma,\\
s^* - \ell_i^* &\leq 0,\\
\sigma_i^*\cdot(s^* - \ell_i^*) &= 0, \quad \sigma_i^* \geq 0, \quad \forall i \in \Sigma.
\end{aligned}$$

Differentiating with respect to ℓ, the following equation is obtained:

$$\frac{\partial}{\partial \ell_i}\mathbb{L}^\alpha(\ell,\mu,\lambda,\tau)|_{\ell=\ell^*,\lambda=\lambda^*,t=t^*,\tau=\tau^*} = \mu_i - \tau^* D^{-\ell_i^*}\log_e D + \lambda_i^* - \sigma_i^* = 0,\ \forall i \in \Sigma \tag{3.93}$$

which after manipulation, it becomes

$$D^{-\ell_i^*} = \frac{\mu_i + \lambda_i^* - \sigma_i^*}{\tau^*\log_e D}, \qquad i \in \Sigma. \tag{3.94}$$

Differentiating with respect to t and s, the following equations are obtained:

$$\frac{\partial}{\partial t}\mathbb{L}^{\alpha}(\ell,\mu,\lambda,\tau)|_{\ell=\ell^*,\lambda=\lambda^*,t=t^*,\tau=\tau^*} = \alpha - \sum_{i\in\Sigma}\lambda_i^* = 0 \Longrightarrow \sum_{i\in\Sigma}\lambda_i^* = \alpha. \tag{3.95}$$

$$\frac{\partial}{\partial s}\mathbb{L}^{\alpha}(\ell,\mu,\lambda,\tau)|_{\ell=\ell^*,\lambda=\lambda^*,t=t^*,\tau=\tau^*} = -\alpha + \sum_{i\in\Sigma}\sigma_i^* = 0 \Longrightarrow \sum_{i\in\Sigma}\sigma_i^* = \alpha. \tag{3.96}$$

When $\tau^* = 0$, (3.93) gives $\mu_i = \sigma_i^* - \lambda_i^*, \forall i \in \Sigma$. Since $\sigma_i^* = \lambda_i^* = 0, \forall i \in \Sigma \setminus \Sigma^0 \cup \Sigma_0$, then it is concluded that $\mu_i = 0$. However, $\mu_i > 0$, $\forall i \in \Sigma \setminus \Sigma^0 \cup \Sigma_0$, and therefore, necessarily $\tau^* > 0$. Next, τ^* is found by substituting (3.94) and (3.95) into the Kraft equality to deduce

$$\sum_{i\in\Sigma} D^{-\ell_i^*} = \sum_{i\in\Sigma}\frac{\mu_i + \lambda_i^* - \sigma_i^*}{\tau^*\log_e D} = \frac{\sum_{i\in\Sigma}\mu_i}{\tau^*\log_e D} + \frac{\sum_{i\in\Sigma}\lambda_i^*}{\tau^*\log_e D} - \frac{\sum_{i\in\Sigma}\sigma_i^*}{\tau^*\log_e D} = \frac{1}{\tau^*\log_e D} = 1.$$

Therefore, $\tau^* = \frac{1}{\log_e D}$. Substituting τ^* into (3.94) yields

$$D^{-\ell_i^*} = \mu_i + \lambda_i^* - \sigma_i^*, \quad i \in \Sigma. \tag{3.97}$$

Let $w_i^* \triangleq D^{-\ell_i^*}$, i.e., the probabilities that correspond to the codeword lengths ℓ_i^*; also, let $\underline{w} \triangleq D^{-t^*}$ and $\overline{w} \triangleq D^{-s^*}$. From the Karush-Kuhn-Tucker conditions $\lambda_i^* \cdot (\ell_i^* - t^*) = 0$ and $\lambda_i^* \geq 0$, $\forall i \in \Sigma$ we deduce the following. For all $i \in \Sigma \setminus \Sigma^0 \cup \Sigma_0$, $\ell_i < t$ and $\ell_i > s$; hence $\lambda_i^* = 0$ and $\sigma_i^* = 0$. For all $i \in \Sigma_0$, $\ell_i < t$ and $\ell_i = s$; hence $\lambda_i^* = 0$ and $\sigma_i^* > 0$. For all $i \in \Sigma^0$, $\ell_i = t$ and $\ell_i > s$; hence $\lambda_i^* > 0$ and $\sigma_i^* = 0$. Therefore, we can distinguish (3.97) in the following cases:

$$D^{-\ell_i^*} = \mu_i, \qquad i \in \Sigma \setminus \Sigma^0 \cup \Sigma_0, \tag{3.98}$$

$$D^{-\ell_i^*} = \mu_i - \sigma_i^*, \quad i \in \Sigma_0, \tag{3.99}$$

$$D^{-\ell_i^*} = \mu_i + \lambda_i^*, \quad i \in \Sigma^0. \tag{3.100}$$

Substituting λ_i^* into (3.95) we have $\sum_{i\in\Sigma}\left(D^{-\ell_i^*} - \mu_i\right) = \alpha$, and substituting $w_i^* \triangleq D^{-\ell_i^*}$ we get

$$\sum_{i\in\Sigma}\left(w_i^* - \mu_i\right) = \alpha. \tag{3.101}$$

We know that $\lambda_i^* \neq 0$ only when $\ell_i^* = t^*$; otherwise, $w_i^* = \mu_i$. Hence, we can see that $w_i^* - \mu_i = (\underline{w} - \mu_i)^+$ and it is positive only when $\ell_i^* = t^*$. Hence, equation (3.101) becomes

$$\sum_{i\in\Sigma}\left(\underline{w} - \mu_i\right)^+ = \alpha, \tag{3.102}$$

where $(f)^+ = \max(0, f)$. This is the classical waterfilling equation [17, Section 9.4] and $\underline{w}$ is the water-level chosen, as shown in Figure 3.4.

If we also substitute the previously obtained expression of σ_i^* into (3.96) we have $\sum_{i\in\Sigma}\Big(\mu_i - D^{-\ell_i^*}\Big) = \alpha$, and substituting $w_i^\ddagger \triangleq D^{-\ell_i^*}$ we get

$$\sum_{i\in\Sigma}\Big(\mu_i - w_i^\ddagger\Big) = \alpha. \tag{3.103}$$

Hence, substituting $\overline{w} \triangleq D^{-s}$, equation (3.103) becomes

$$\sum_{i\in\Sigma}\Big(\mu_i - \overline{w}\Big)^+ = \alpha. \tag{3.104}$$

In this example, the solution to a minimax average codeword length lossless coding prob-

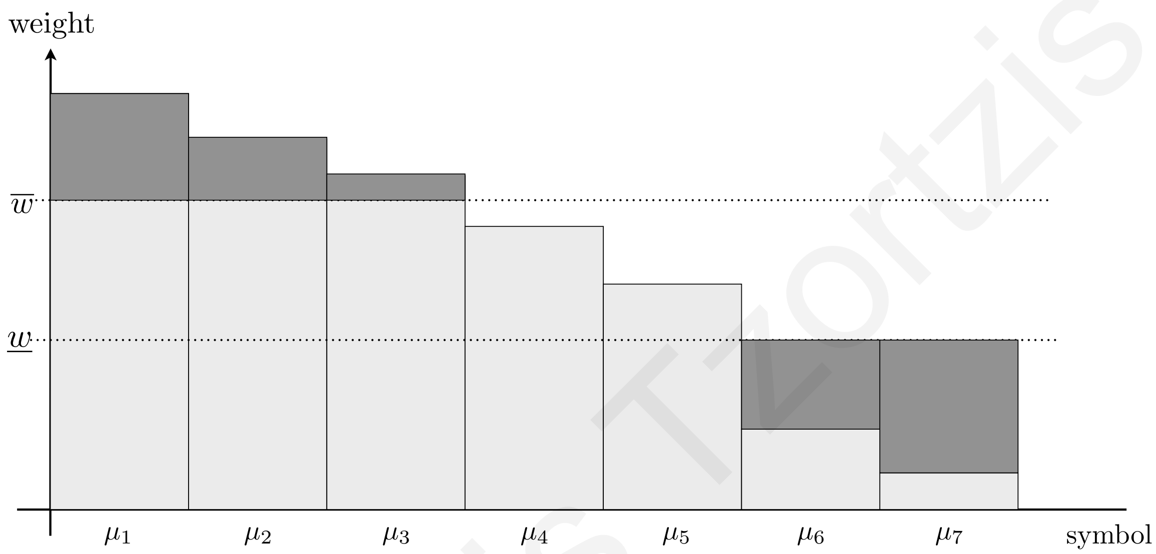


Figure 3.4.: Example demonstrating the solution of the coding problem using a watefilling-like fashion, where $\nu^* = \{\overline{w}, \overline{w}, \overline{w}, \mu_4, \mu_5, \underline{w}, \underline{w}\}$.

lem for the class of sources described by the total variational ball is transformed into an optimization one by finding the expression of the maximization over the total variational ball. Subsequently, a solution is given in terms of a waterfilling with two distinct levels.

Remark 3.3. *Note that the above solution can be used to approximate a high-dimensional alphabet probability distribution $\mu \in \mathbb{P}(\Sigma)$, by another lower-dimensional alphabet probability distribution $\nu \in \mathbb{P}(\overline{\Sigma})$, $\overline{\Sigma} \subseteq \Sigma$ by invoking Jayne's maximum entropy principle [34,35], subject to information quantified via the total variation distance between $\nu \in \mathbb{P}(\Sigma)$ and $\mu \in \mathbb{P}(\Sigma)$, because of the following fact. Since all conditions of the Von-Neumann minimax theorem [25] hold, then we have*

$$\min_{\ell\in\mathbb{R}_+^{|\Sigma|}:\sum_{i\in\Sigma} D^{-\ell_i}\le 1}\ \max_{\nu\in\mathbb{B}_R(\mu)} \sum_{i\in\Sigma}\ell_i\nu_i = \max_{\nu\in\mathbb{B}_R(\mu)}\ \min_{\ell\in\mathbb{R}_+^{|\Sigma|}:\sum_{i\in\Sigma} D^{-\ell_i}\le 1} \sum_{i\in\Sigma}\ell_i\nu_i$$
$$\overset{(a)}{=} \max_{\nu\in\mathbb{B}_R(\mu)} \sum_{i\in\Sigma}\ell_i^*\nu_i\big|_{\ell_i^*=-\log\nu_i} = \max_{\nu\in\mathbb{B}_R(\mu)} H(\nu). \tag{3.105}$$

where equality in (a) *follows from the classical solution of Shannon codes [17]. Hence, the solution of lossless coding for a variation class is equivalent to Jayne's maximum entropy formulation subject to total variation distance constraint. Thus, the approximated distribution is obtained by a waterfilling-like solution, as shown in Figure 3.4.*

3.5. Summary

This chapter is concerned with extremum problems involving total variation distance metric as a pay-off subject to linear functional constraints, and vice-versa; that is, with the roles of total variation metric and linear functional interchanged. These problems are formulated using concepts from signed measures while the theory is developed on abstract spaces. Certain properties and applications of the extremum problems are discussed, while closed form expressions of the extremum measures are derived for finite alphabet spaces. The fundamental water-filling property and the partitioning of the extremum problems are also elaborated. Finally, it is shown through examples how the extremum solution of the various problems behaves, and an application to the well-known lossless compression problem of finding uniquely decodable codes, which minimize the average codeword length is presented.

4

Dynamic Programming with TV Distance Ambiguity on a Finite Horizon

In this chapter we address optimality of stochastic control strategies on a finite horizon, via dynamic programming subject to total variation distance ambiguity on the conditional distribution of the controlled process. We formulate the stochastic control problem using minimax theory, in which the control minimizes the pay-off while the controlled process, maximizes it. By employing certain results of Chapter 3, the maximization of a linear functional on the space of probability measures, among those probability measures which are within a total variation distance from a nominal probability measure, we solve the minimax stochastic control problem with deterministic control strategies, under a Markovian and a non-Markovian assumption, on the conditional distributions of the controlled process. The results of this part include:

- minimax optimization subject to total variation distance ambiguity constraint;
- new dynamic programming recursions, which involve the oscillator seminorm of the value function, in addition to the standard terms;
- examples which illustrate the applications of our results.

4.1. Problem Formulation

In this section, we describe the abstract formulation of the minimax problem under total variation distance ambiguity.

4.1.1. Dynamic Programming of Finite Horizon Discounted-Markov Control Model

A finite horizon Discounted-Markov Control Model (D-MCM) with deterministic strategies is a septuple

$$\text{D-MCM}: \Big(\{\mathcal{X}_i\}_{i=0}^{n}, \{\mathcal{U}_i\}_{i=0}^{n-1}, \{\mathcal{U}_i(x_i) : x_i \in \mathcal{X}_i\}_{i=0}^{n-1}, \{Q_i(dx_i|x_{i-1}, u_{i-1}) : (x_{i-1}, u_{i-1}) \in \mathcal{X}_{i-1} \times \mathcal{U}_{i-1}\}_{i=0}^{n}, \{f_i\}_{i=0}^{n-1}, h_n, \alpha\Big) \tag{4.1}$$

consisting of the following.

(a) State Space. A sequence of Polish spaces (complete separable metric spaces) $\{\mathcal{X}_i : i = 0, \ldots, n\}$, which model the state space of the controlled random process $\{x_j \in \mathcal{X}_j : j = 0, \ldots, n\}$.

(b) Control or Action Space. A sequence of Polish spaces $\{\mathcal{U}_i : i = 0, \ldots, n-1\}$, which model the control or action set of the control random process $\{u_j \in \mathcal{U}_j : j = 0, \ldots, n-1\}$.

(c) Feasible Controls or Actions. A family $\{\mathcal{U}_i(x_i) : x_i \in \mathcal{X}_i\}$ of non-empty measurable subsets $\mathcal{U}_i(x_i) \subseteq \mathcal{U}_i$, where $\mathcal{U}_i(x_i)$ denotes the set of feasible controls or actions, when the controlled process is in state $x_i \in \mathcal{X}_i$, and the feasible state-actions pairs defined by $\mathbb{K}_i \triangleq \Big\{(x_i, u_i) : x_i \in \mathcal{X}_i, u_i \in \mathcal{U}_i(x_i)\Big\}$ are measurable subsets of $\mathcal{X}_i \times \mathcal{U}_i, i = 0, \ldots, n-1$.

(d) Controlled Process Distribution. A collection of conditional distributions or stochastic kernels $Q_i(dx_i|x_{i-1}, u_{i-1})$ on $\mathcal{X}_i$ given $(x_{i-1}, u_{i-1}) \in \mathbb{K}_{i-1} \subseteq \mathcal{X}_{i-1} \times \mathcal{U}_{i-1}, i = 0, \ldots, n$. The controlled process distribution is described by the sequence of transition probability distributions $\{Q_i(dx_i|x_{i-1}, u_{i-1}) : (x_{i-1}, u_{i-1}) \in \mathbb{K}_{i-1}, i = 0, \ldots, n\}$.

(e) Cost-Per-Stage. A collection of non-negative measurable functions $f_j : \mathbb{K}_j \to [0, \infty]$, called the cost-per-stage, such that $f_j(x, \cdot)$ does not take the value $+\infty$ for each $x \in \mathcal{X}_j, j = 0, \ldots, n-1$. The running pay-off functional is defined in terms of $\{f_j : j = 0, \ldots, n-1\}$.

(f) Terminal Cost. A bounded measurable non-negative function $h_n : \mathcal{X}_n \rightarrow [0, \infty)$ called the terminal cost. The pay-off functional at the last stage is defined in terms of h_n.

(g) Discounting Factor. A real number $\alpha \in (0, 1)$ called the discounting factor.

The definition of D-MCM envisions applications of systems described by discrete-time dynamical state space models, which include random external inputs, since such models give rise to a collection of controlled processes distributions $\{Q_i(dx_i|x_{i-1}, u_{i-1}){:}(x_{i-1}, u_{i-1}) \in \mathbb{K}_{i-1}, i = 0, \ldots, n\}$. For any integer $j \geq 0$, define the product spaces by $\mathcal{X}_{0,j} \triangleq \times_{i=0}^{j} \mathcal{X}_i$ and $\mathcal{U}_{0,j-1} \triangleq \times_{i=0}^{j-1} \mathcal{U}_i$, and the discounted sample pay-off by

$$F_{0,n}^{\alpha}(x_0, u_0, x_1, u_1, \ldots, x_{n-1}, u_{n-1}, x_n) \triangleq \sum_{j=0}^{n-1} \alpha^j f_j(x_j, u_j) + \alpha^n h_n(x_n). \tag{4.2}$$

The goal in feedback controlled optimization with deterministic strategies is to choose a control strategy or policy $g \triangleq \{g_j : j = 0, 1, \ldots, n-1\}$, $g_j : \mathcal{X}_{0,j} \times \mathcal{U}_{0,j-1} \longrightarrow \mathcal{U}_j(x_j)$, $u_j^g = g_j(x_0^g, x_1^g, \ldots, x_j^g, u_0^g, u_1^g, \ldots, u_{j-1}^g)$, $j = 0, 1, \ldots, n-1$ so as to minimize the pay-off functional

$$\begin{aligned} \mathbb{E}\Big\{\sum_{j=0}^{n-1} \alpha^j f_j(x_j^g, u_j^g) + \alpha^n h_n(x_n^g)\Big\} = \int_{\mathcal{X}_0 \times \mathcal{X}_1 \times \ldots \times \mathcal{X}_n} \\ F_{0,n}^{\alpha}(x_0, u_0^g(x_0), x_1, u_1^g(x_0, x_1), \ldots, x_{n-1}, u_{n-1}^g(x_0, x_1, \ldots, x_{n-1}), x_n) \\ Q_0(dx_0) Q_1(dx_1|x_0, u_0^g(x_0)) \ldots Q_n(dx_n|x_{n-1}, u_{n-1}^g(x_0, x_1, \ldots, x_{n-1})). \end{aligned} \tag{4.3}$$

Clearly, pay-off (4.3) is a functional of the collection of conditional distributions $\{Q_i(\cdot|\cdot) : i = 0, 1, \ldots, n\}$. Moreover, if this collection of distribution has countable support for each (x_{i-1}, u_{i-1}), $i = 0, \ldots, n$, then each integral in (4.3) is reduced to a countable summation.

A Markov property on the controlled process distributions, i.e.,

$$Q_i(dx_i|x^{i-1}, u^{i-1}) {=} Q_i(dx_i|x_{i-1}, u_{i-1}), \quad \forall (x^{i-1}, u^{i-1}) \in \times_{j=0}^{i-1} \mathbb{K}_j, \quad i = 0, 1, \ldots, n$$

under admissible non-Markov strategies, implies that Markov control strategies are optimal [39]. Therefore, $g_j : \mathcal{X}_j \longrightarrow \mathcal{U}_j(x_j)$. For $(i, x) \in \{0, 1 \ldots, n\} \times \mathcal{X}_i$, let $V_i^0(x) \in \mathbb{R}$ represent the minimal cost-to-go or value function on the time horizon $\{i, i{+}1, \ldots, n\}$ if the controlled process starts at state $x_i = x$ at time i, defined by

$$V_i^0(x) \triangleq \inf_{\substack{g_k \in \mathcal{U}_k(x_k) \\ k=i,\ldots,n-1}} \mathbb{E}_{i,x}^g \Big\{\sum_{j=i}^{n-1} \alpha^j f_j(x_j^g, u_j^g) + \alpha^n h_n(x_n^g)\Big\} \tag{4.4}$$

where $\mathbb{E}_{i,x}^g\{\cdot\}$ denotes expectation conditioned on $x_i^g = x$. Consequently, it can be shown that the value function (4.4) satisfies the following dynamic programming recursion relating the value functions $V_i^0(\cdot)$ and $V_{i+1}^0(\cdot)$ [39],

$$V_n^0(x) = \alpha^n h_n(x), \quad x \in \mathcal{X}_n \tag{4.5}$$

$$V_i^0(x) = \inf_{u \in \mathcal{U}_i(x)} \Big\{ \alpha^i f_i(x,u) + \int_{\mathcal{X}_{i+1}} V_{i+1}^0(z) Q_{i+1}(dz|x,u) \Big\}, \quad x \in \mathcal{X}_i. \tag{4.6}$$

Since the value function $V_i^0(x)$ defined by (4.4) and the dynamic programming recursion (4.5), (4.6) depend on the complete knowledge of the collection of conditional distributions $\{Q_i(\cdot|\cdot) : i = 0, \ldots, n\}$, any mismatch of the collection $\{Q_i(\cdot|\cdot) : i = 0, \ldots, n\}$ from the true collection of conditional distributions, will affect the optimality of the control strategies. Our objective is to address the impact of any ambiguity measured by the total variation distance between the true conditional distribution and a given nominal distribution on the cost-to-go (4.4), and dynamic programming recursion (4.5), (4.6).

4.1.2. Dynamic Programming with Total Variation Distance Ambiguity

The objective of this chapter is to investigate dynamic programming under ambiguity of the conditional distributions of the controlled processes

$$\Big\{ Q_i(dx_i|x_{i-1}, u_{i-1}) : (x_{i-1}, u_{i-1}) \in \mathbb{K}_{i-1} \Big\}, \quad i = 0, \ldots, n.$$

The ambiguity of the conditional distributions of the controlled process is modeled by the total variation distance. Specifically, given a collection of nominal controlled process distributions $\{Q_i^o(dx_i|x_{i-1}, u_{i-1}) : (x_{i-1}, u_{i-1}) \in \mathbb{K}_{i-1}\}$, $i = 0, \ldots, n$, the corresponding collection of true controlled process distributions $\{Q_i(dx_i|x_{i-1}, u_{i-1}) : (x_{i-1}, u_{i-1}) \in \mathbb{K}_{i-1}\}$, $i = 0, \ldots, n$, is modeled by a set described by the total variation distance centered at the nominal conditional distribution having radius $R_i \in [0, 2]$, $i = 0, \ldots, n$, defined by

$$\mathbf{B}_{R_i}(Q_i^o)(x_{i-1}, u_{i-1}) \triangleq \Big\{ Q_i(\cdot|x_{i-1}, u_{i-1}) : ||Q_i(\cdot|x_{i-1}, u_{i-1}) - Q_i^o(\cdot|x_{i-1}, u_{i-1})||_{TV} \leq R_i \Big\}.$$

The total variation distance model of ambiguity is quite general, and it includes linear, nonlinear, finite and/or countable state space models, etc, since no assumptions are impossed on the structure of the stochastic control dynamical system model, which induces the collection of conditional distributions $\{Q_i(\cdot|\cdot) : i = 0, \ldots, n\}$, $\{Q_i^o(\cdot|\cdot) : i = 0, \ldots, n\}$. Given the above description of ambiguity in distribution, we re-formulate the value function and dynamic programming recursion via minimax theory as follows.

For $(i, x) \in \{0, 1 \ldots, n\} \times \mathcal{X}_i$, let $V_i(x) \in \mathbb{R}$ represent the minimal cost-to-go on the time horizon $\{i, i+1, \ldots, n\}$ if the state of the controlled process starts at state $x_i = x$ at time i, defined by[1]

$$V_i(x) \triangleq \inf_{\substack{g_k \in \mathcal{U}_k(x_k)\\ k=i,\ldots,n-1}} \sup_{\substack{Q_{k+1}(\cdot|x_k,u_k) \in \mathbf{B}_{R_{k+1}}(Q^o_{k+1})(x_k,u_k)\\ k=i,\ldots,n-1}} \mathbb{E}^g_{i,x}\Big\{ \sum_{j=i}^{n-1} \alpha^j f_j(x^g_j, u^g_j) + \alpha^n h_n(x^g_n) \Big\}$$

where $\mathbb{E}^g_{i,x}$ denotes conditional expectation with respect to the true collection of conditional distribution $\{Q_k(\cdot|\cdot) : k = i, \ldots, n\}$. Even in the above minimax setting the Markov property of the controlled process distribution under an admissible non-Markov (i.e., feedback) strategy implies that Markov control strategies are optimal. Moreover, the value function satisfies the following dynamic programming recursion relating the value function $V_i(\cdot)$ and $V_{i+1}(\cdot)$, for all $i = 0, 1, \ldots, n-1$.

$$\begin{aligned} V_n(x) &= \alpha^n h_n(x), \quad x \in \mathcal{X}_n \\ V_i(x) &= \inf_{u \in \mathcal{U}_i(x)} \sup_{Q_{i+1}(\cdot|x,u) \in \mathbf{B}_{R_{i+1}}(Q^o_{i+1})(x,u)} \Big\{ \alpha^i f_i(x,u) + \int_{\mathcal{X}_{i+1}} V_{i+1}(z) Q_{i+1}(dz|x,u) \Big\}, \ x \in \mathcal{X}_i. \end{aligned}$$

Based on this formulation, if $V_{i+1}(\cdot)$ is bounded continuous non-negative, we show that the new dynamic programming equation involves the oscillator seminorm of the value function, in addition to the standard terms.

In addition to the D-MCM, we will also discuss the general discounted feedback control model (i.e., we relax the Markovian assumption). In summary, the issues discussed and results obtained in this chapter are the following:

1. formulation of finite horizon discounted stochastic optimal control subject to conditional distribution ambiguity described by total variation distance via minimax theory;
2. dynamic programming recursions for
   a) discounted-feedback control model
   b) nominal discounted-Markov control model

   under total variation distance ambiguity on the conditional distribution of the controlled process;
3. characterization of the maximizing conditional distribution belonging to the total variation distance set, and the corresponding new dynamic programming recursions;
4. applications of the finite horizon minimax problem to the well-known inventory control and machine replacement examples. Comparisons are included for the case for which the total variation constraint is replaced by the relative entropy constraint.

[1]Assuming the inf sup solution exists. However, for finite alphabet spaces such solution exists.

4.2. Minimax Stochastic Control

In this section, we first introduce the general definition of finite horizon Discounted-Feedback Control Model (D-FCM) with randomized and deterministic control policies, under total variation distance uncertainty (which includes the D-MCM introduced in Section 4.1), and then we apply the characterization of the maximizing distribution of Section 3.2.1 and 3.3.1 to the dynamic programming recursion.

Define $\mathbb{N}^n \triangleq \{0,1,2,\ldots,n\}, n \in \mathbb{N}$. The state space and the control space are sequences of Polish spaces $\{\mathcal{X}_j : j = 0,1,\ldots,n\}$ and $\{\mathcal{U}_j : j = 0,1,\ldots,n-1\}$, respectively. These spaces are associated with their corresponding measurable spaces $(\mathcal{X}_j, \mathcal{B}(\mathcal{X}_j)), \forall j \in \mathbb{N}^n$, $(\mathcal{U}_j, \mathcal{B}(\mathcal{U}_j))$, $\forall j \in \mathbb{N}^{n-1}$. Define the product spaces by $\mathcal{X}_{0,n} \triangleq \times_{i=0}^{n} \mathcal{X}_i$, $\mathcal{U}_{0,n-1} \triangleq \times_{i=0}^{n-1} \mathcal{U}_i$, and introduce their product measurable spaces, $(\mathcal{X}_{0,n}, \mathcal{B}(\mathcal{X}_{0,n}))$, $(\mathcal{U}_{0,n-1}, \mathcal{B}(\mathcal{U}_{0,n-1}))$, respectively, for $n \in \mathbb{N}^n$. The state process is denoted by $x^n \triangleq \{x_j : j = 0,1,\ldots,n\}$, and the control process is denoted by $u^{n-1} \triangleq \{u_j : j = 0,1,\ldots,n-1\}$.

Given $(\mathcal{X}_{0,n}, \mathcal{B}(\mathcal{X}_{0,n}))$, $(\mathcal{U}_{0,n-1}, \mathcal{B}(\mathcal{U}_{0,n-1}))$ the Borel state and control or action spaces, respectively, and the initial state distribution $\nu_0(dx_0)$, we introduce the space $H_{0,n}$ of admissible observable histories by

$$H_{0,n} \triangleq \mathbb{K}_0 \times \mathbb{K}_1 \times \ldots \times \mathbb{K}_{n-1} \times \mathcal{X}_n \equiv \times_{i=0}^{n-1} \mathbb{K}_i \times \mathcal{X}_n, \quad n \in \mathbb{N}, \quad H_{0,0} = \mathcal{X}_0$$

where $\mathbb{K}_i \triangleq \{(x_i, u_i) : x_i \in \mathcal{X}_i, u_i \in \mathcal{U}_i(x_i)\}$, denote the feasible state-action pairs, for $i = 0,1,\ldots,n-1$. A typical element $h_{0,n} \in H_{0,n}$ is a sequence of the form

$$h_{0,n} = (x_0, u_0, \ldots, x_{n-1}, u_{n-1}, x_n), \quad (x_i, u_i) \in \mathbb{K}_i, \quad i = 0, \ldots, n-1, \quad x_n \in \mathcal{X}_n.$$

Similarly, introduce

$$G_{0,n} = \mathcal{X}_0 \times \mathcal{U}_0 \times \ldots \times \mathcal{X}_{n-1} \times \mathcal{U}_{n-1} \times \mathcal{X}_n \equiv \times_{i=0}^{n-1} (\mathcal{X}_i \times \mathcal{U}_i) \times \mathcal{X}_n, \quad n \in \mathbb{N}$$
$$G_{0,0} = H_{0,0} = \mathcal{X}_0.$$

Thus, $H_{0,n}$ is a sequence of $G_{0,n}$ for each $n = 0,1,\ldots$. The spaces $G_{0,n}$ and $H_{0,n}$ are equipped with the natural σ-algebra $\mathcal{B}(G_{0,n})$ and $\mathcal{B}(H_{0,n})$, respectively (and by Kolmogorov's extension theorem they can be extended to $\mathcal{B}(G_{0,\infty})$ and $\mathcal{B}(H_{0,\infty})$). We shall use the Borel space $(H_\infty, \mathcal{B}(H_\infty))$ as the main measurable space $(\Omega, \mathcal{F})$. Next, we formulate the definition of discounted feedback control model.

Definition 4.1. *A finite horizon D-FCM is a septuple*

$$\textit{D-FCM}: \Big(\mathcal{X}_{0,n}, \mathcal{U}_{0,n-1}, \{\mathcal{U}_i(x_i) : x_i \in \mathcal{X}_i\}_{i=0}^{n-1}, \{Q_i(dx_i|x^{i-1}, u^{i-1}) : (x^{i-1}, u^{i-1}) \in \mathcal{X}_{0,i-1} \times \mathcal{U}_{0,i-1}\}_{i=0}^{n}, \{f_i\}_{i=0}^{n-1}, h_n, \alpha\Big) \quad (4.7)$$

*consisting of the items **(a)**-**(c)**, **(e)**-**(g)** of finite horizon D-MCM (4.1), while the controlled process distribution in **(d)** is replaced by the non-Markov collection* $\{Q_i(dx_i|x^{i-1},u^{i-1}) : (x^{i-1},u^{i-1}) \in \times_{j=0}^{i-1}\mathbb{K}_j\}_{i=0}^n$.

Next, we give the definitions of randomized, deterministic, and stationary control strategies or policies.

Definition 4.2. *A randomized control strategy is a sequence* $\pi \triangleq \{\pi_0,\ldots,\pi_{n-1}\}$ *of stochastic kernels* $\pi_i(\cdot|\cdot)$ *on* $(\mathcal{U}_i,\mathcal{B}(\mathcal{U}_i))$ *conditioned on* $(H_{0,i},\mathcal{B}(H_{0,i}))$ *(e.g.,* $\pi_i(du_i|x^i,u^{i-1})$ *) satisfying*

$$\pi_i(\mathcal{U}_i(x_i)|x^i,u^{i-1}) = 1 \quad \textit{for every} \quad (x^i,u^{i-1}) \in H_{0,i}, \quad i=0,1,\ldots,n-1.$$

The set of all such policies is denoted by $\mathbf{\Pi}_{0,n-1}$*. A strategy* $\pi \triangleq \{\pi_i : i=0,\ldots,n-1\} \in \mathbf{\Pi}_{0,n-1}$ *is called*

(a) randomized Markov strategy if there exists a sequence $\left\{\pi_i^M(\cdot|\cdot) : i=0,\ldots,n-1\right\}$ *of stochastic kernels* $\pi_i^M(\cdot|\cdot)$ *on* $(\mathcal{U}_i,\mathcal{B}(\mathcal{U}_i))$ *conditioned on* $(\mathcal{X}_i,\mathcal{B}(\mathcal{X}_i))$ *such that*

$$\pi_i\left(\cdot|x^i,u^{i-1}\right) = \pi_i^M\left(\cdot|x_i\right), \quad \forall \left(x^i,u^{i-1}\right) \in H_{0,i}, \quad i=0,1,\ldots,n-1.$$

The set of randomized Markov strategies is denoted by $\mathbf{\Pi}_{0,n-1}^{RM}$*;*

(b) randomized stationary Markov strategy if there exists a stochastic kernel $\pi^S(\cdot|\cdot)$ *on* $(\mathcal{U},\mathcal{B}(\mathcal{U}))$ *conditioned on* $(\mathcal{X},\mathcal{B}(\mathcal{X}))$ *such that*

$$\pi_i(\cdot|x^i,u^{i-1}) = \pi^S(\cdot|x_i), \quad \forall (x^i,u^{i-1}) \in H_{0,i}, \quad i=0,1,\ldots,n-1.$$

The set of randomized stationary Markov strategies is denoted by $\mathbf{\Pi}_{0,n-1}^{RS}$*;*

(c) deterministic feedback strategy if there exists a sequence $g \triangleq \{g_j : j=0,1,\ldots,n-1\}$ *of measurable functions* $g_j : \times_{i=0}^{j-1}\mathbb{K}_i \times \mathcal{X}_j \longmapsto \mathcal{U}_j$*, such that for all* $(x^j,u^{j-1}) \in H_{0,j}$*,* $j \in \mathbb{N}^{n-1}$*,* $g_j(x_0,u_0,x_1,u_1,\ldots,x_{j-1},u_{j-1},x_j) \in \mathcal{U}_j(x_j)$*, and* $\pi_j\left(\cdot|x^j,u^{j-1}\right)$ *assigns mass 1 to some point in* $\mathcal{U}_j$*, that is,*

$$\pi_i\left(A_i|x^i,u^{i-1}\right) = I_{A_i}\left(g_i\left(x^i,u^{i-1}\right)\right), \quad \forall A_i \in \mathcal{B}(\mathcal{U}_i), \quad i=0,1,\ldots,n-1,$$

where $I_{A_i}(\cdot)$ *is the indicator function of* $A_i \in \mathcal{B}(\mathcal{U}_i)$*.*
The set of deterministic feedback strategies is denoted by $\mathbf{\Pi}_{0,n-1}^{DF}$*;*

(d) deterministic Markov strategy if there exists a sequence $g \triangleq \{g_j : j=0,1,\ldots,n-1\}$ *of measurable functions* $g_j : \mathcal{X}_j \longmapsto \mathcal{U}_j$ *satisfying* $g_j(x_j) \in \mathcal{U}_j(x_j)$ *for all* $x_j \in \mathcal{X}_j$*,*

$j \in \mathbb{N}^{n-1}$, and $\pi_j(\cdot|x^j, u^{j-1})$ is concentrated at $g_j(x_j) \in \mathcal{U}_j(x_j)$ for all $(x^j, u^{j-1}) \in H_{0,j}$, $j \in \mathbb{N}^{n-1}$, that is,

$$\pi_i\left(A_i|x^i, u^{i-1}\right) = I_{A_i}\left(g_i\left(x_i\right)\right), \quad \forall A_i \in \mathcal{B}(\mathcal{U}_i), \quad i = 0, 1, \ldots, n-1,$$

The set of deterministic Markov strategies is denoted by $\mathbf{\Pi}^{DM}_{0,n-1}$;

(e) deterministic stationary Markov strategy if there exists a measurable function $g : \mathcal{X} \longmapsto \mathcal{U}$ such that $g(x_t) \in \mathcal{U}(x_t)$, $\forall x_t \in \mathcal{X}$, and $\pi_j(\cdot|x^j, u^{j-1})$ assigns mass to some point u_j, $\forall (x^j, u^{j-1}) \in H_{0,j}$, e.g.,

$$\pi_i(A_i|x^i, u^{i-1}) = I_{A_i}(g(x_i)), \quad \forall A_i \in \mathcal{B}(\mathcal{U}_i), \quad i = 0, \ldots, n-1.$$

The set of deterministic stationary Markov strategies is denoted by $\mathbf{\Pi}^{DS}_{0,n-1}$.

Let $\Pi^{D}_{0,n-1}$ denote the set of all deterministic policies, so that $\Pi^{D}_{0,n-1} \subset \Pi_{0,n-1}$.

The relationship between the classes of control strategies or policies is as follows: $\Pi^{DS}_{0,n-1} \subset \Pi^{RS}_{0,n-1} \subset \Pi^{RM}_{0,n-1} \subset \Pi_{0,n-1}$, $\Pi^{DS}_{0,n-1} \subset \Pi^{DM}_{0,n-1} \subset \Pi^{RM}_{0,n-1} \subset \Pi_{0,n-1}$, and $\Pi^{DS}_{0,n-1} \subset \Pi^{DM}_{0,n-1} \subset \Pi^{DF}_{0,n-1} \subset \Pi_{0,n-1}$. Thus, randomized feedback strategies or policies $\Pi_{0,n-1}$ contain all other classes of policies, and hence, are most general. On the other hand, stationary deterministic strategies or policies are contained in all other classes. According to Definition 4.2, the set of control policies is non-empty, since we have assumed existence of measurable functions $g_j : \mathbb{K}_{0,j-1} \times \mathcal{X}_j \longrightarrow \mathcal{U}_j$ such that $\forall x^j, u^{j-1} \in \mathbb{K}_{0,j-1} \times \mathcal{X}_j$, $g_j(x^j, u^{j-1}) \in \mathcal{U}_j(\mathcal{X}_j), \forall j \in \mathbb{N}^{n-1}$. Sufficient conditions for this to hold are in general obtained via measurable selection theorems [32]. For denumerable set (countable alphabet) $\mathcal{X}_j$ endowed with the discrete topology any function is measurable.

Given a controlled process $\{Q_i(\cdot|x^{i-1}, u^{i-1}) : (x^{i-1}, u^{i-1}) \in \mathbb{K}_{0,i-1}\}_{i=0}^n$ and a randomized control process $\{\pi_i(\cdot|x^i, u^{i-1}) : (x^i, u^{i-1}) \in \mathbb{K}_{0,i-1} \times \mathcal{X}_i\}_{i=0}^n \in \mathbf{\Pi}_{0,n-1}$ and the initial probability $\nu_0(\cdot) \in \mathcal{M}_1(\mathcal{X}_0)$, then by Ionescu-Tulceu theorem [10] there exists a unique probability measure $\mathbf{Q}^\pi_\nu$ on $(\Omega, \mathcal{F})$ defined by

$$\begin{aligned}
&\mathbf{Q}^\pi_\nu(x_0 \in A_0, u_0 \in B_0, \ldots, x_{n-1} \in A_{n-1}, u_{n-1} \in B_{n-1}, x_n \in A_n) = \\
&\quad Q_0(dx_0)\pi_0(du_0|x_0) \otimes Q_1(dx_1|x_0, u_0)\pi_1(du_1|x^1, u_0) \otimes \ldots \\
&\qquad \otimes Q_{n-1}(dx_{n-1}|x^{n-2}, u^{n-2})\pi_{n-1}(du_{n-1}|x^{n-1}, u^{n-2}) \otimes Q_n(dx_n|x^{n-1}, u^{n-1})
\end{aligned} \tag{4.8}$$

such that

$$\begin{aligned}
&\mathbf{Q}^\pi_\nu(x_0 \in A_0) = \nu(A_0), \quad A_0 \in \mathcal{B}(\mathcal{X}_0) \\
&\mathbf{Q}^\pi_\nu(u_j \in B_j|h_{0,j}) = \pi_j(B_j|h_{0,j}), \quad B_j \in \mathcal{B}(\mathcal{U}_j) \\
&\mathbf{Q}^\pi_\nu(x_{j+1} \in C_{j+1}|h_{0,j}, u_j) = Q_{j+1}(C_{j+1}|h_{0,j}, u_j), \quad C_{j+1} \in \mathcal{B}(\mathcal{X}_{j+1}).
\end{aligned}$$

Given the sample pay-off

$$F_{0,n}^{\alpha}(x_0,u_0,x_1,u_1,\ldots,x_{n-1},u_{n-1},x_n) \triangleq \sum_{j=0}^{n-1} \alpha^j f_j(x_j,u_j) + \alpha^n h_n(x_n) \tag{4.9}$$

its expectation is

$$\begin{aligned}\mathbb{E}_{\mathbf{Q}_\nu^\pi}\Big\{F_{0,n}^{\alpha}(x_0,u_0,\ldots,x_{n-1},u_{n-1},x_n)\Big\} &= \int F_{0,n}^{\alpha}(x_0,u_0,\ldots,x_{n-1},u_{n-1},x_n)\\ Q_0(dx_0)\pi_0(du_0|x_0)\otimes Q_1(dx_1|x_0,u_0)\pi_1(du_1|x^1,u_0)\otimes\ldots\\ \otimes Q_{n-1}(dx_{n-1}|x^{n-2},u^{n-2})\pi_{n-1}(du_{n-1}|x^{n-1},u^{n-2})\otimes Q_n(dx_n|x^{n-1},u^{n-1}).\end{aligned} \tag{4.10}$$

Note that the class of randomized strategies $\mathbf{\Pi}_{0,n-1}$ embeds deterministic feedback and Markov strategies.

4.2.1. Variation Distance Ambiguity

Next, we introduce the definitions of nominal controlled process distributions (for finite horizon D-FCM and D-MCM), and their corresponding ambiguous controlled process distributions.

For each $\pi \in \mathbf{\Pi}_{0,n-1}^{DF}$, $\pi \in \mathbf{\Pi}_{0,n-1}^{DM}$ and $\pi \in \mathbf{\Pi}_{0,n-1}^{DS}$ the nominal controlled process is described by a sequence of conditional distributions as follows.

Definition 4.3. *(Nominal Controlled Process Distributions). A nominal controlled state processes* $\{x^g = x_0^g, x_1^g, \ldots, x_n^g : \pi \in \mathbf{\Pi}_{0,n-1}^{DF}, \ \ \pi \in \mathbf{\Pi}_{0,n-1}^{DM}, \ \ or \ \ \pi \in \mathbf{\Pi}_{0,n-1}^{DS}\}$ *corresponds to a sequence of stochastic kernels as follows:*

(a) Feedback Controlled Process. For every $A \in \mathcal{B}(\mathcal{X}_j)$,

$$Prob(x_j \in A|x^{j-1},u^{j-1}) = Q_j^o(A|x^{j-1},u^{j-1})$$

where $Q_j^o(A|x^{j-1},u^{j-1}) \in \mathcal{Q}(\mathcal{X}_j|\mathbb{K}_{0,j-1}), \forall j \in \mathbb{N}_+^n$.

(b) Markov Controlled Process. For every $A \in \mathcal{B}(\mathcal{X}_j)$,

$$Prob(x_j \in A|x^{j-1},u^{j-1}) = Q_j^o(A|x_{j-1},u_{j-1})$$

where $Q_j^o(A|x_{j-1},u_{j-1}) \in \mathcal{Q}(\mathcal{X}_j|\mathbb{K}_{j-1}), \forall j \in \mathbb{N}_+^n$.

(c) Stationary Markov Controlled Process. For every $A \in \mathcal{B}(\mathcal{X})$

$$Prob(x_j \in A|x^{j-1},u^{j-1}) = Q^o(A|x_{j-1},u_{j-1})$$

where $Q^o(A|x_{j-1},u_{j-1}) \in \mathcal{Q}(\mathcal{X}|\mathbb{K})$.

The class of controlled processes is described by the sequence of stochastic kernels,

$$\{Q_j(dx_j|x^{j-1}, u^{j-1}) \in \mathcal{Q}(\mathcal{X}_j|\mathbb{K}_{0,j-1} : j = 0, \ldots, n\}$$

belonging to a total variation distance set as follows.

Definition 4.4. *(Class of Controlled Process Distribution) Given a nominal controlled process stochastic kernel of Definition 4.3, and $R_i \in [0,2], 0 \leq i \leq n$ the class of controlled process stochastic kernels is defined as follows:*

(a) Class with respect to Feedback Nominal Controlled Process. Given a fixed $Q_j^o(\cdot|\cdot) \in \mathcal{Q}(\mathcal{X}_j|\mathbb{K}_{0,j-1})$, $j = 0, 1, \ldots, n$ the class of stochastic kernels is defined by

$$\mathbf{B}_{R_i}(Q_i^o)(x^{i-1}, u^{i-1}) \triangleq \Big\{Q_i(dx_i|x^{i-1}, u^{i-1}) : \\ ||Q_i(\cdot|x^{i-1}, u^{i-1}) - Q_i^o(\cdot|x^{i-1}, u^{i-1})||_{TV} \leq R_i\Big\}, \quad i = 0, 1, \ldots, n.$$

(b) Class with respect to Markov Nominal Controlled Process. Given a fixed $Q_j^o(\cdot|\cdot) \in \mathcal{Q}(\mathcal{X}_j|\mathbb{K}_{j-1})$, $j = 0, 1, \ldots, n$ the class of stochastic kernels is defined by

$$\mathbf{B}_{R_i}(Q_i^o)(x^{i-1}, u^{i-1}) \triangleq \Big\{Q_i(dx_i|x^{i-1}, u^{i-1}) : \\ ||Q_i(\cdot|x^{i-1}, u^{i-1}) - Q_i^o(\cdot|x_{i-1}, u_{i-1})||_{TV} \leq R_i\Big\}, \quad i = 0, 1, \ldots, n.$$

(c) Class with respect to Stationary Markov Nominal Controlled Process. Given a fixed $Q^o(\cdot|\cdot) \in \mathcal{Q}(\mathcal{X}|\mathbb{K})$ the class of stochastic kernels is defined by

$$\mathbf{B}_R(Q^o)(x, u) \triangleq \Big\{Q(dz|x, u) : ||Q(\cdot|x, u) - Q^o(\cdot|x, u)||_{TV} \leq R\Big\}.$$

Note that in Definition 4.4 (a), (b), although we use the same notation $\mathbf{B}_{R_i}(Q_i^o)(x^{i-1}, u^{i-1})$ these sets are different because the nominal distribution $Q_i^o(\cdot|\cdot)$ can be of Feedback or Markov form. The above model is motivated by the fact that dynamic programming involves conditional expectation with respect to the collection of conditional distributions $\{Q_i(\cdot|\cdot) \in \mathcal{Q}(\mathcal{X}_i|\mathbb{K}_{0,i-1}) : i = 0, \ldots, n\}$. Therefore, any ambiguity in these distributions will affect the optimality of the strategies.

4.2.2. Pay-Off Functional

For each $\pi \in \mathbf{\Pi}_{0,n-1}^{DF}$ or $\pi \in \mathbf{\Pi}_{0,n-1}^{DM}$ the discounted average pay-off is defined by

$$J_{0,n}(\pi, Q_i : i = 0, \ldots, n) \triangleq \mathbb{E}_{\mathbf{Q}_\nu^\pi}\Big\{\sum_{j=0}^{n-1} \alpha^j f_j(x_j, u_j) + \alpha^n h_n(x_n)\Big\} \tag{4.11}$$

where $\mathbb{E}_{\mathbf{Q}_\nu^\pi}\{\cdot\}$ denotes expectation with respect to the true joint measure $\mathbf{Q}_\nu^\pi(dx^n, du^{n-1})$ defined by (4.8) such that $Q_i(\cdot|x^{i-1}, u^{i-1}) \in \mathbf{B}_{R_i}(Q_i^o)$, $i = 0, 1, \dots, n$ (e.g., it belongs to the total variation distance ball of Definition 4.4).

Next, we introduce assumptions so that the maximization over the class of ambiguous measures is well-defined.

Assumption 4.1. *The nominal system family satisfies the following assumption: The maps* $\{f_j : \mathcal{X}_j \times \mathcal{U}_j \longmapsto \mathbb{R} : j = 0, 1, \dots, n-1\}$, $h_n : \mathcal{X}_n \longmapsto \mathbb{R}$ *are bounded, continuous and non-negative.*

Note that it is possible to relax Assumption 4.1 to lower semi-continuous non-negative functions bounded from below. Next, for illustration purposes we introduce an example based on discrete-time recursion dynamics and deterministic strategies.

Example 4.1. *(Nominal Model) The nominal controlled processes is* $\{x^g = x_0^g, x_1^g, \dots, x_n^g : u \in \mathbf{\Pi}_{0,n-1}^{DF}\}$, *and corresponds to a sequence of stochastic kernels* $\{Q^o_{w_j|x^j,u^j}(dw|x^j, u^j) : j = 0, 1, \dots, n-1\}$, *functions* $\{b_j : \mathcal{X}_j \times \mathcal{U}_j \times \mathcal{W}_j \longmapsto \mathcal{X}_{j+1} : j = 0, 1, \dots, n-1\}$, *and noise processes* $\{w_j : j = 0, 1, \dots, n-1\}$ *adapted to a filtration* $\{\mathcal{F}_{0,i} : i = 0, \dots, n-1\}$ *such that the following hold.*

1. *For each* $j \in \mathbb{N}^{n-1}$, w_j *is* $\mathcal{F}_{0,j}-$ *measurable and* $\{x_0^g, x_1^g, \dots, x_n^g\}$ *are generated by the recursion*

$$x_{j+1}^g = b_j(x_j^g, u_j^g, w_j), \qquad x_0^g = x_0 \tag{4.12}$$

which implies that if x_0 *is* $\mathcal{F}_{0,0}-$*measurable then* x_j^g *is* $\mathcal{F}_{0,j-1}-$*measurable.*

2. *For every* $A \in \mathcal{B}(\mathcal{W}_j), j \in \mathbb{N}^{n-1}$

$$Prob(w_j \in A|x^j, u^j) = Q^o_{w_j}(A|x_j^g, u_j^g), \qquad a.s. \tag{4.13}$$

3. $Prob(x_0^g = x_0) = 1, \qquad \forall u \in \mathbf{\Pi}_{0,n-1}^{DF}.$

Notice that (4.13) assumes that the noise $\{w_j : j \in \mathbb{N}^{n-1}\}$ *is correlated with the state and control processes. It can be further simplified to an independent and identically distributed sequence* $\{w_j : j \in \mathbb{N}^{n-1}\}$, *independent of* x_j^g, *which then implies* $Q^o_{w_j}(dw_j|x_j^g, u_j^g) = Q^o_{w_j}(dw_j)$ *for almost all* $(x_j^g, u_j^g) \in \mathcal{X}_j \times \mathcal{U}_j$, $j = 0, \dots, n-1$.
The ambiguous model is constructed as follows.

(Uncertainty Stochastic Model) Suppose $\{\mathcal{G}_{0,j} : j = 0, 1, \dots, n-1\}$ *is the true filtration which is generated by some processes and that* $\mathcal{F}_{0,i} \subset \mathcal{G}_{0,i}$, $\forall i \in \mathbb{N}^{n-1}$. *We can model the*

class of true conditional distributions by, $Q_{w_j|\mathcal{G}_{0,j}}(dw_j|\mathcal{G}_{0,j}) \in \mathcal{M}_1(\mathcal{W}_j)$, $0 \leq j \leq n-1$, *such that they belong to variation distance class*

$$\mathbf{B}_{R_i}(Q^o_{w_i})(\mathcal{G}_{0,i}) \triangleq \Big\{Q_{w_i}(dw_i|\mathcal{G}_{0,i}) : ||Q_{w_i}(\cdot|\mathcal{G}_{0,i}) - Q^o_{w_i}(\cdot|x^g_i, u^g_i)||_{TV} \leq R_i\Big\},$$
$$R_i \in [0,2], \quad i = 0,1,\dots,n-1.$$

The above model is motivated by the fact that the value function often involves conditional expectation with respect to $Q_{w_i}(dw_i|\mathcal{G}_{0,i})$, *and that the true noise in (4.12) can be correlated with past information, such as, the information defined by* $\mathcal{G}_{0,i} \triangleq \sigma\{x_k, u_k, x_{k+1}, u_{k+1}, \dots, x_i, u_i\}$, $0 \leq k \leq i$.

4.3. Minimax Dynamic Programming

In this section we shall apply the results of Chapter 3 to formulate and solve the minimax stochastic control under finite horizon D-FCM and D-MCM ambiguities.

4.3.1. Discounted Feedback Control Model

Utilizing the above formulation, we define the minimax stochastic control problem, where the maximization is over a total variation distance ball, centered at the nominal conditional distribution $Q^o_i(dx_i|x^{i-1}, u^{i-1}) \in \mathcal{Q}(\mathcal{X}_i|\mathbb{K}_{0,i-1})$ having radius $R_i \in [0,2]$, for $i = 0,1,\dots,n$.

Problem 4.1. *Given a nominal feedback controlled process of Definition 4.3 (a), an admissible policy set* $\mathbf{\Pi}^{DF}_{0,n-1}$ *and an ambiguity class* $\mathbf{B}_{R_k}(Q^o_k)(x^{k-1}, u^{k-1})$, k=0, ..., n *of Definition 4.4 (a), find a* $\pi^* \in \mathbf{\Pi}^{DF}_{0,n-1}$ *and a sequence of stochastic kernels* $Q^*_k(dx_k|x^{k-1}, u^{k-1}) \in \mathbf{B}_{R_k}(Q^o_k)(x^{k-1}, u^{k-1})$, $k = 0, 1, ..., n$ *which solve the following minimax optimization problem.*

$$J_{0,n}(\pi^*, Q^*_k : k = 0,\dots,n) = \inf_{\pi \in \mathbf{\Pi}^{DF}_{0,n-1}} \Big\{ \sup_{\substack{Q_k(\cdot|x^{k-1},u^{k-1}) \in \mathbf{B}_{R_k}(Q^o_k)(x^{k-1},u^{k-1}) \\ k=0,1,\dots,n}} \mathbb{E}_{\mathbf{Q}^{\pi}_{\nu}} \Big\{ \sum_{k=0}^{n-1} \alpha^k f_k(x^g_k, u^g_k) + \alpha^n h_n(x^g_n) \Big\} \Big\}. \quad (4.14)$$

Next, we apply dynamic programming to characterize the solution of (4.14), by first addressing the maximization. Define the pay-off associated with the maximization problem

$$J_{0,n}(\pi, Q^*_k : k = 0,\dots,n) \triangleq \sup_{\substack{Q_k(\cdot|x^{k-1},u^{k-1}) \in \mathbf{B}_{R_k}(Q^o_k)(x^{k-1},u^{k-1}) \\ k=0,1,\dots,n}} J_{0,n}(\pi, Q_k : k = 0,\dots,n).$$

For a given $\pi \in \mathbf{\Pi}^{DF}_{0,n-1}$, which defines $\{g_j : j = 0, \dots, n-1\}$, and $\pi_{[k,m]} \equiv u^g_{[k,m]}$, denoting the restriction of policies in $[k,m]$, $0 \leq k \leq m \leq n-1$, define the conditional expectation taken over the events $\mathcal{G}_{0,j} \triangleq \sigma\{x^g_0, \dots, x^g_j, u^g_0, \dots, u^g_j\}$ maximized over the class $\mathbf{B}_{R_k}(Q^o_k)(x^{k-1}, u^{k-1})$, $k = j+1, \dots, n$, as follows [13, 39]:

$$V_j(u^g_{[j,n-1]}, \mathcal{G}_{0,j}) \triangleq \sup_{\substack{Q_k(\cdot|x^{k-1},u^{k-1}) \in \mathbf{B}_{R_k}(Q^o_k)(x^{k-1},u^{k-1}) \\ k=j+1,\dots,n}} \mathbb{E}_{\mathbf{Q}^\pi_\nu}\Big\{ \sum_{k=j}^{n-1} \alpha^k f_k(x^g_k, u^g_k) + \alpha^n h_n(x^g_n) | \mathcal{G}_{0,j} \Big\} \tag{4.15}$$

where $\mathbb{E}_{\mathbf{Q}^\pi_\nu}\{\cdot|\mathcal{G}_{0,j}\}$ denotes conditional expectation with respect to $\mathcal{G}_{0,j}$ calculated on the probability measure $\mathbf{Q}^\pi_\nu$. Then, $V_j(u^g_{[j,n-1]}, \mathcal{G}_{0,j})$ satisfies the following dynamic programming equation [39],

$$V_n(\mathcal{G}_{0,n}) = \alpha^n h_n(x^g_n) \tag{4.16}$$

$$V_j(u^g_{[j,n-1]}, \mathcal{G}_{0,j}) = \sup_{Q_{j+1}(\cdot|x^j,u^j) \in \mathbf{B}_{R_{j+1}}(Q^o_{j+1})(x^j,u^j)} \Big\{ \mathbb{E}_{Q_{j+1}(\cdot|x^j,u^j)}\Big\{ \alpha^j f_j(x^g_j, u^g_j) + V_{j+1}(u^g_{[j+1,n-1]}, \mathcal{G}_{0,j+1}) \Big\}\Big\} \tag{4.17}$$

where $\mathbb{E}_{Q_{j+1}(\cdot|x^j,u^j)}\{\cdot\}$ denotes expectation with respect to $Q_{j+1}(dx_{j+1}|\mathbb{K}_{0,j})$.

Next, we present the dynamic programming recursion for the minimax problem. Let $V_j(\mathcal{G}_{0,j})$ represent the minimax pay-off on the future time horizon $\{j, j+1, \dots, n\}$ at time $j \in \mathbb{N}^n_+$ defined by

$$V_j(\mathcal{G}_{0,j}) \triangleq \inf_{\pi \in \mathbf{\Pi}^{DF}_{j,n-1}} \sup_{\substack{Q_k(\cdot|x^{k-1},u^{k-1}) \in \mathbf{B}_{R_k}(Q^o_k)(x^{k-1},u^{k-1}) \\ k=j+1,\dots,n}} \Big\{ \mathbb{E}_{\mathbf{Q}^\pi_\nu}\Big\{ \sum_{k=j}^{n-1} \alpha^k f_k(x^g_k, u^g_k) + \alpha^n h_n(x^g_n) | \mathcal{G}_{0,j} \Big\}\Big\} = \inf_{\pi \in \mathbf{\Pi}^{DF}_{j,n-1}} V_j(u^g_{[j,n-1]}, \mathcal{G}_{0,j}). \tag{4.18}$$

Then by reconditioning we obtain

$$V_j(\mathcal{G}_{0,j}) \triangleq \inf_{u \in \mathcal{U}_{ad}[j,n-1]} \sup_{\substack{Q_k(\cdot|x^{k-1},u^{k-1}) \in \mathbf{B}_{R_k}(Q^o_k)(x^{k-1},u^{k-1}) \\ k=j+1,\dots,n}} \Big\{ \mathbb{E}_{\mathbf{Q}^\pi_\nu}\Big\{ \alpha^j f_j(x^g_j, u^g_j) + \mathbb{E}_{\mathbf{Q}^\pi_\nu}\Big\{ \sum_{k=j+1}^{n-1} \alpha^k f_k(x^g_k, u^g_k) + \alpha^n h_n(x^g_n) | \mathcal{G}_{0,j+1} \Big\} | \mathcal{G}_{0,j} \Big\}\Big\}. \tag{4.19}$$

Hence, we deduce the following dynamic programming recursion

$$V_n(\mathcal{G}_{0,n}) = \alpha^n h_n(x^g_n) \tag{4.20}$$

$$V_j(\mathcal{G}_{0,j}) \triangleq \inf_{u_j \in \mathcal{U}_j(x)} \sup_{Q_{j+1}(\cdot|x^j,u^j) \in \mathbf{B}_{R_{j+1}}(Q^o_{j+1})(x^j,u^j)} \Big\{ \mathbb{E}_{Q_{j+1}(\cdot|x^j,u^j)}\Big\{ \alpha^j f_j(x^g_j, u^g_j) + V_{j+1}(\mathcal{G}_{0,j+1}) \Big\}\Big\}. \tag{4.21}$$

By applying the results of Section 3.2.1 and 3.3.1 to (4.20), (4.21) we obtain the following theorem.

Theorem 4.1. *Suppose there exist an optimal policy for Problem 4.1, and assume* $V_{j+1}(\cdot):\mathcal{X}_{0,j+1}\times\mathcal{U}_{0,j}\longmapsto[0,\infty)$ *in (4.18) is bounded continuous in* $x\in\mathcal{X}_{j+1}$, $j=0,\ldots,n-1$.

1) The dynamic programming recursion is given by

$$V_n(\mathcal{G}_{0,n}) = \alpha^n h_n(x_n^g) \tag{4.22}$$

$$\begin{aligned} V_j(\mathcal{G}_{0,j}) = \inf_{u_j\in\mathcal{U}_j(x)} \Big\{ & \mathbb{E}_{Q_{j+1}^o}\Big(\alpha^j f_j(x_j^g,u_j^g) + V_{j+1}(\mathcal{G}_{0,j+1})|\mathcal{G}_{0,j}\Big) \\ & + \frac{R_j}{2}\Big(\sup_{x_{j+1}\in\mathcal{X}_{j+1}} V_{j+1}(\mathcal{G}_{0,j},x_{j+1}) - \inf_{x_{j+1}\in\mathcal{X}_{j+1}} V_{j+1}(\mathcal{G}_{0,j},x_{j+1})\Big)\Big\}. \end{aligned} \tag{4.23}$$

Moreover,

$$V_j(\mathcal{G}_{0,j}) = \inf_{u_j\in\mathcal{U}_j(x)} \mathbb{E}_{Q_{j+1}^*}\Big\{\alpha^j f_j(x_j^g,u_j^g) + V_{j+1}(\mathcal{G}_{0,j},x_{j+1})|\mathcal{G}_{0,j}\Big\} \tag{4.24}$$

where, the optimal conditional distributions $\{Q_j^* : j=0,1,\ldots,n-1\}$ *are given by*

$$Q_{j+1}^*\left(\mathcal{X}_{j+1}^+|x^j,u^j\right) = Q_{j+1}^o(\mathcal{X}_{j+1}^+|x^j,u^j) + \frac{R_{j+1}}{2} \in [0,1], \ \ (x^j,u^j)\in\mathbb{K}_{0,j} \tag{4.25}$$

$$Q_{j+1}^*\left(\mathcal{X}_{j+1}^-|x^j,u^j\right) = Q_{j+1}^o(\mathcal{X}_{j+1}^-|x^j,u^j) - \frac{R_{j+1}}{2} \in [0,1], \ \ (x^j,u^j)\in\mathbb{K}_{0,j} \tag{4.26}$$

$$Q_{j+1}^*\left(A|x^j,u^j\right) = Q_{j+1}^o(A|x^j,u^j), \ \forall A\subseteq\mathcal{X}_{j+1}\setminus\mathcal{X}_{j+1}^+\cup\mathcal{X}_{j+1}^-, \ (x^j,u^j)\in\mathbb{K}_{0,j} \tag{4.27}$$

and [2]

$$\mathcal{X}_{j+1}^+ \triangleq \Big\{x_{j+1}\in\overline{\mathcal{X}}_{j+1}: V_{j+1}(\mathcal{G}_{0,j},x_{j+1}) = \sup\Big\{V_{j+1}(\mathcal{G}_{0,j},y_{j+1}): y_{j+1}\in\mathcal{X}_{j+1}\Big\}\Big\} \tag{4.28}$$

$$\mathcal{X}_{j+1}^- \triangleq \Big\{x_{j+1}\in\overline{\mathcal{X}}_{j+1}: V_{j+1}(\mathcal{G}_{0,j},x_{j+1}) = \inf\Big\{V_{j+1}(\mathcal{G}_{0,j},y_{j+1}): y_{j+1}\in\mathcal{X}_{j+1}\Big\}\Big\}. \tag{4.29}$$

2) The total pay-off is given by

$$J_{0,n}(\pi^*, Q_i^* : i=0,\ldots,n-1) = \sup_{Q_0(\cdot)\in\mathbf{B}_{R_0}(Q^o)} \mathbb{E}_{Q_0}\Big\{V_0(\mathcal{G}_{0,0})\Big\}. \tag{4.30}$$

Proof. 1) Consider (4.21) expressed in integral form

$$\begin{aligned} V_j(\mathcal{G}_{0,j}) = \inf_{u_j\in\mathcal{U}_j(x)} \Big\{ & \alpha^j f_j(x_j,u_j) \\ & + \sup_{Q_{j+1}(\cdot|x^j,u^j)\in\mathbf{B}_{R_{j+1}}(Q_{j+1}^o)(x^j,u^j)} \int V_{j+1}(\mathcal{G}_{0,j},z)Q_{j+1}(dz|x^j,u^j)\Big\}. \end{aligned} \tag{4.31}$$

By applying (3.46) we obtain (4.22), (4.23), while (4.25)-(4.29) follow as well.

2) By evaluating (4.18) at $j=0$ we obtain (4.30). This completes the derivation. ■

[2]Note the notation Σ^0 and Σ_0 in Chapter 3 is identical to the notation $\mathcal{X}_{j+1}^+$ and $\mathcal{X}_{j+1}^-$, respectively.

By Theorem 4.1, the maximizing measure is given by (4.25)-(4.27), and it is a functional of the nominal measure. At this stage we cannot claim that the maximizing measure is Markovian, and hence the optimal strategy is not necessarily Markov. Therefore, the computation of optimal strategies using non-Markov nominal controlled processes is computationally intensive. Next, we restrict the minimax formulation to Markov controlled nominal processes.

4.3.2. Discounted Markov Control Model

In this section we shall apply the results of Chapter 3 to formulate and solve minimax stochastic control under finite horizon D-MCM ambiguity. The derivations of the results are based on the classical results (without ambiguity on the controlled process) found in [54]. We define the minimax stochastic control problem as follows.

Problem 4.2. *Given a nominal Markov controlled process of Definition 4.3 (b), an admissible policy set* $\mathbf{\Pi}^{DF}_{0,n-1}$ *and an ambiguity class* $\mathbf{B}_{R_k}(Q^o_k)(x^{k-1},u^{k-1}), k{=}0,\ldots,n$ *of Definition 4.4 (b), find a* $\pi^*{\in}\mathbf{\Pi}^{DF}_{0,n-1}$ *and a sequence of stochastic kernels* $Q^*_k(dx_k|x^{k-1},u^{k-1}) \in \mathbf{B}_{R_k}(Q^o_k)(x^{k-1},u^{k-1})$*,* $k=0,1,\ldots,n$ *which solve the following minimax optimization problem.*

$$J_{0,n}(\pi^*,\{Q^*_k\}^n_{k=0}) = \inf_{\pi\in\mathbf{\Pi}^{DF}_{0,n-1}} \Big\{ \sup_{\substack{Q_k(\cdot|x^{k-1},u^{k-1})\in\mathbf{B}_{R_k}(Q^o_k)(x^{k-1},u^{k-1})\\ k=0,1,\ldots,n}} \mathbb{E}_{\mathbf{Q}^\pi_\nu}\Big\{\sum_{k=0}^{n-1}\alpha^k f_k(x^g_k,u^g_k)+\alpha^n h_n(x^g_n)\Big\}\Big\}. \tag{4.32}$$

In view of Section 3.2.1, specifically, the relation between the maximizing distribution and the nominal distribution (3.14)-(3.16), which also apply to conditional distributions, we deduce that the maximization conditional distribution $Q^*_i(dx_i|x^{i-1},u^{i-1})$ is Markovian, hence $Q^*_i(dx_i|x^{i-1},u^{i-1}) = Q^*_i(dx_i|x_{i-1},u_{i-1})$, $\forall(x^{i-1},u^{i-1})\in\mathbb{K}_{0,i-1}$. Hence, the minimax optimization problem (4.32) is reformulated as follows.

$$J_{0,n}(\pi^*,\{Q^*_k\}^n_{k=0}) = \inf_{\pi\in\mathbf{\Pi}^{DF}_{0,n-1}} \Big\{ \sup_{\substack{Q_k(\cdot|x_{k-1},u_{k-1})\in\mathbf{B}_{R_k}(Q^o_k)(x_{k-1},u_{k-1})\\ k=0,1,\ldots,n}} \mathbb{E}_{\mathbf{Q}^\pi_\nu}\Big\{\sum_{k=0}^{n-1}\alpha^k f_k(x^g_k,u^g_k)+\alpha^n h_n(x^g_n)\Big\}\Big\}. \tag{4.33}$$

Next, we apply dynamic programming to characterize the solution of (4.33), by first addressing the maximization. Define the pay-off associated with the maximization problem

$$J_{0,n}(\pi,\{Q^*_k\}^n_{k=0}) = \sup_{\substack{Q_k(\cdot|x_{k-1},u_{k-1})\in\mathbf{B}_{R_k}(Q^o_k)(x_{k-1},u_{k-1})\\ k=0,1,\ldots,n}} J_{0,n}(\pi,\{Q_k\}^n_{k=0}). \tag{4.34}$$

For a given $\pi \in \Pi^{DF}_{0,n-1}$, which defines $\{g_j : j = 0, 1, \ldots, n-1\}$, and $\pi_{[k,m]} \equiv u^g_{[k,m]}$ denoting the restriction of policies in $[k, m]$, $0 \leq k \leq m \leq n-1$, define the conditional cost-to-go or value function taken over the events $\mathcal{G}_{0,j} \triangleq \sigma\{x^g_0, \ldots, x^g_j, u^g_0, \ldots, u^g_j\}$ maximized over the class $\mathbf{B}_{R_k}(Q^o_k)(x_{k-1}, u_{k-1})$, $k = j+1, \ldots, n$, as follows.

$$V_j(u^g_{[j,n-1]}, \mathcal{G}_{0,j}) \triangleq \sup_{\substack{Q_k(\cdot|x_{k-1},u_{k-1}) \in \mathbf{B}_{R_k}(Q^o_k)(x_{k-1},u_{k-1}) \\ k=j+1,\ldots,n}} \Bigg\{ \mathbb{E}_{\mathbf{Q}^\pi_\nu}\Big\{ \sum_{k=j}^{n-1} \alpha^k f_k(x^g_k, u^g_k) + \alpha^n h_n(x^g_n) | \mathcal{G}_{0,j} \Big\} \Bigg\} \tag{4.35}$$

where $\mathbb{E}_{\mathbf{Q}^\pi_\nu}\{\cdot|\mathcal{G}_{0,j}\}$ denotes conditional expectation with respect to $\mathcal{G}_{0,j}$ calculated on the probability measure $\mathbf{Q}^\pi_\nu$. Here $V_j(u^g_{[j,n-1]}, \mathcal{G}_{0,j})$ depends on g only through $g_j, g_{j+1}, \ldots, g_{n-1}$. Hence, for $j = n$, $V_j(u^g_{[j,n-1]}, \mathcal{G}_{0,j}) \equiv V_n(\mathcal{G}_{0,n})$. Note further that

$$J_{0,n}(\pi, \{Q^*_k\}^n_{k=0}) = \mathbb{E}_{\mathbf{Q}^\pi_\nu}\Big[V_0(u^g_{[0,n-1]}, \mathcal{G}_{0,0}) \Big]. \tag{4.36}$$

Lemma 4.1. *Let $\pi \in \Pi^{DM}_{0,n-1}$, and define recursively the functions*

$$V^g_n(x) = \alpha^n h_n(x) \tag{4.37}$$

$$V^g_j(x) = \sup_{Q_{j+1}(\cdot|x,u) \in \mathbf{B}_{R_{j+1}}(Q^o_{j+1})(x,u)} \mathbb{E}_{Q_{j+1}(\cdot|x,u)}\Big\{ \alpha^j f_j(x, g_j(x)) + V^g_{j+1}(x^g_{j+1}) \Big\}. \tag{4.38}$$

Then the random variable $V^g_j(x^g_j)$ satisfies

$$V^g_j(x^g_j) = V_j(u^g_{[j,n-1]}, \mathcal{G}_{0,j}). \tag{4.39}$$

Proof. For the derivation of Lemma 4.1 see Appendix C.1. ∎

Lemma 4.2. *(The Comparison Principle) Let $V_j(x)$, $0 \leq j \leq n$, be any functions such that*

$$V_n(x) \leq \alpha^n h_n(x) \tag{4.40}$$

$$V_j(x) \leq \sup_{Q_{j+1}(\cdot|x,u) \in \mathbf{B}_{R_{j+1}}(Q^o_{j+1})(x,u)} \mathbb{E}_{Q_{j+1}(\cdot|x,u)}\Big\{ \alpha^j f_j(x, u) + V_{j+1}(x_{j+1}) \Big\} \tag{4.41}$$

for all $x \in \mathcal{X}_j$ and for all $u \in \mathcal{U}_j(x)$. Let $\pi \in \Pi^D_{0,n-1}$ be an arbitrary policy. Then w.p.1

$$V_j(x^g_j) \leq V_j(u^g_{[j,n-1]}, \mathcal{G}_{0,j}). \tag{4.42}$$

The interpretation of (4.42) is that, under an arbitrary deterministic policy, $V_j(x^g_j)$ is a lower bound of the conditional value function.

Proof. For the derivation of Lemma 4.2 see Appendix C.2. ∎

Corollary 4.1. *Let $V_j(x)$ be functions satisfying (4.40) and (4.41). Then $J_{0,n}(\pi^*, \{Q_k^*\}_{k=0}^n) \geq \mathbb{E}\{V_0(x_0)\}$. Hence, if $\pi \in \Pi^D$ is arbitrary and is such that $V_0(u^g_{[0,n-1]}, G_{0,0}) = V_0(x_0)$, then π is optimal.*

Proof. For any arbitrary $\pi \in \Pi^D_{0,n-1}$, we have that $V_0(u^g_{[0,n-1]}, G_{0,0}) \geq V_0(x_0)$ by (4.42). Taking expectations and using (4.36), $J_{0,n}(\pi, \{Q_k^*\}_{k=0}^n) \geq \mathbb{E}\{V_0(x_0)\}$ and since π was arbitrary this yields

$$J_{0,n}(\pi^*, \{Q_k^*\}_{k=0}^n) \geq \mathbb{E}\{V_0(x_0)\}.$$

Finally, if $V_0(u^g_{[0,n-1]}, G_{0,0}) = V_0(x_0)$ then

$$J_{0,n}(\pi, \{Q_k^*\}_{k=0}^n) = \mathbb{E}\{V_0(x_0)\} \leq J_{0,n}(\pi^*, \{Q_k^*\}_{k=0}^n)$$

so that π must be optimal and $J_{0,n}(\pi, \{Q_k^*\}_{k=0}^n) = J_{0,n}(\pi^*, \{Q_k^*\}_{k=0}^n)$. ■

Theorem 4.2. *Define recursively the functions*

$$V_n(x) = \alpha^n h_n(x) \tag{4.43}$$

$$V_j(x) = \inf_{u \in \mathcal{U}_j(x)} \sup_{Q_{j+1}(\cdot|x,u) \in \mathbf{B}_{R_{j+1}}(Q^o_{j+1})(x,u)} \mathbb{E}_{Q_{j+1}(\cdot|x,u)}\Big\{\alpha^j f_j(x,u) + V_{j+1}(x_{j+1})\Big\}. \tag{4.44}$$

1) *Let $\pi \in \Pi^D_{0,n-1}$ be arbitrary. Then $V_j(x^g_j) \leq V_j(u^g_{[j,n-1]}, \mathcal{G}_{0,j})$ w.p.1, in particular $J_{0,n}(\pi, \{Q_k^*\}_{k=0}^n) \geq \mathbb{E}\{V_0(x_0)\}$.*

2) *A Markov policy $\pi \in \Pi^{DM}_{0,n-1}$ which defines $\{g_j : j = 0, 1, \ldots, n-1\}$ is optimal if the infimum in (4.44) is achieved at $g_j(x)$, and then $V_j(x^g_j) = V_j(u^g_{[j,n-1]}, \mathcal{G}_{0,j})$ w.p.1 and $J_{0,n}(\pi^*, \{Q_k^*\}_{k=0}^n) = \mathbb{E}\{V_0(x_0)\}$.*

3) *A Markov policy $\pi \in \Pi^{DM}_{0,n-1}$ which defines $\{g_j : j = 0, 1, \ldots, n-1\}$ is optimal only if for each j, the infimum at x^g_j in (4.44) is achieved by $g_j(x^g_j)$, i.e.,*

$$V^g_j(x^g_j) = \sup_{Q_{j+1}(\cdot|x,u) \in \mathbf{B}_{R_{j+1}}(Q^o_{j+1})(x,u)} \mathbb{E}_{Q_{j+1}(\cdot|x,u)}\Big\{\alpha^j f_j(x^g_j, g_j(x^g_j)) + V^g_{j+1}(x^g_{j+1})\Big\}$$

w.p.1.

4) *Assume $V_{j+1}(\cdot) : \mathcal{X}_{j+1} \to [0,\infty)$ is bounded continuous in $x \in \mathcal{X}_{j+1}$, $j = 0, \ldots, n-1$, then the dynamic programming recursion is given by*

$$V_n(x) = \alpha^n h_n(x), \qquad x \in \mathcal{X}_n \tag{4.45}$$

$$V_j(x) = \inf_{u \in \mathcal{U}_j(x)} \Big\{\alpha^j f_j(x,u) + \int_{\mathcal{X}_{j+1}} V_{j+1}(z) Q^o_{j+1}(dz|x,u) + \frac{R_j}{2}\Big(\sup_{z \in \mathcal{X}_{j+1}} V_{j+1}(z) - \inf_{z \in \mathcal{X}_{j+1}} V_{j+1}(z)\Big)\Big\}, \qquad x \in \mathcal{X}_j. \tag{4.46}$$

Moreover,

$$V_j(x) = \inf_{u \in \mathcal{U}_j(x)} \mathbb{E}_{Q^*_{j+1}} \Big\{ \alpha^j f_j(x^g_j, u^g_j) + V_{j+1}(x_{j+1}) | x_j = x \Big\} \tag{4.47}$$

where the optimal conditional distribution $\{Q^*_j(\cdot|\cdot,\cdot) : j = 0, 1, \ldots, n-1\}$ *is given by*

$$Q^*_{j+1}\Big(\mathcal{X}^+_{j+1}|x_j,u_j\Big) = Q^o_{j+1}(\mathcal{X}^+_{j+1}|x_j,u_j) + \frac{R_{j+1}}{2} \in [0,1], \ \ (x_j,u_j) \in \mathbb{K}_j \tag{4.48}$$

$$Q^*_{j+1}\Big(\mathcal{X}^-_{j+1}|x_j,u_j\Big) = Q^o_{j+1}(\mathcal{X}^-_{j+1}|x_j,u_j) - \frac{R_{j+1}}{2} \in [0,1], \ \ (x_j,u_j) \in \mathbb{K}_j \tag{4.49}$$

$$Q^*_{j+1}(A|x_j,u_j) = Q^o_{j+1}(A|x_j,u_j), \ \forall A \subseteq \mathcal{X}_{j+1} \setminus \mathcal{X}^+_{j+1} \cup \mathcal{X}^-_{j+1}, \ \ (x_j,u_j) \in \mathbb{K}_j \tag{4.50}$$

and

$$\mathcal{X}^+_{j+1} \triangleq \Big\{ x_{j+1} \in \mathcal{X}_{j+1} : V_{j+1}(x_{j+1}) = \sup\{V_{j+1}(x_{j+1}) : x_{j+1} \in \mathcal{X}_{j+1}\} \Big\} \tag{4.51}$$

$$\mathcal{X}^-_{j+1} \triangleq \Big\{ x_{j+1} \in \mathcal{X}_{j+1} : V_{j+1}(x_{j+1}) = \inf\{V_{j+1}(x_{j+1}) : x_{j+1} \in \mathcal{X}_{j+1}\} \Big\}. \tag{4.52}$$

Proof. 1) The functions $V_j(x)$ defined by (4.43) and (4.44) clearly satisfy (4.40) and (4.41) and hence part 1) follows from Lemma 4.2.

2) To prove the sufficiency in part 2), let $g = \{g_j\}$ be a Markov policy that achieves the infimum in (4.44), so

$$V_j(x) = \sup_{Q_{j+1}(\cdot|x,u) \in \mathbf{B}_{R_{j+1}}(Q^o_{j+1})(x,u)} \mathbb{E}_{Q_{j+1}(\cdot|x,u)} \Big\{ \alpha^j f_j(x, g_j(x)) + V_{j+1}(x_{j+1}) \Big\}. \tag{4.53}$$

By Lemma 4.1 it follows that $V^g_j(x^g_j) = V_j(u^g_{[j,n-1]}, \mathcal{G}_{0,j})$ for all j and in particular $V_0(u^g_{[0,n-1]}, \mathcal{G}_{0,0}) = V^g_0(x^g_0)$. By Corollary 4.1, g is optimal and $J_{0,n}(\pi^*, \{Q^*_k\}^n_{k=0}) = \mathbb{E}\{V^g_0(x^g_0)\}$.

3) To prove the necessity in part 3) suppose the Markovian policy $\pi \in \Pi^{DM}_{0,n-1}$ is optimal. We prove by induction that $g_j(x^g_j)$ achieves the infimum in (4.44) at x^g_j with probability 1. Consider $j = n-1$. Suppose the assertion is false. Then there exists another function g'_{n-1} such that

$$\begin{aligned} &\sup_{Q_n(\cdot|x_{n-1},u_{n-1})} \mathbb{E}_{Q_n(\cdot|x_{n-1},u_{n-1})} \{\alpha^{n-1} f_{n-1}(x^g_{n-1}, g_{n-1}(x^g_{n-1})) + V_n(x^g_n)\} \\ &\quad \geq \sup_{Q_n(\cdot|x_{n-1},u_{n-1})} \mathbb{E}_{Q_n(\cdot|x_{n-1},u_{n-1})} \{\alpha^{n-1} f_{n-1}(x^g_{n-1}, g'_{n-1}(x^g_{n-1})) + V_n(x^{g'}_n)\}, \quad \text{w.p.1} \end{aligned}$$

and note the abuse of notation. Moreover the inequality is strict with positive probability. Using (4.43), we get

$$\begin{aligned} &\sup_{Q_n(\cdot|x_{n-1},u_{n-1})} \mathbb{E}_{Q_n(\cdot|x_{n-1},u_{n-1})} \{\alpha^{n-1} f_{n-1}(x^g_{n-1}, g_{n-1}(x^g_{n-1})) + \alpha^n h_n(x^g_n)\} \\ &\quad > \sup_{Q_n(\cdot|x_{n-1},u_{n-1})} \mathbb{E}_{Q_n(\cdot|x_{n-1},u_{n-1})} \{\alpha^{n-1} f_{n-1}(x^g_{n-1}, g'_{n-1}(x^g_{n-1})) + \alpha^n h_n(x^{g'}_n)\} \end{aligned} \tag{4.54}$$

Consider the Markov policy $g' = \{g_0, \dots, g_{n-2}, g'_{n-1}\}$. Evidently, $x_j^g = x_j^{g'}$ for all $0 \leq j \leq n-1$ and so $u_j^g = u_j^{g'}$, $0 \leq j \leq n-1$ and $u_{n-1}^{g'} = g'_{n-1}(x_{n-1}^g)$. Hence

$$\sup_{Q_j(\cdot|x_{j-1},u_{j-1})} \mathbb{E}_{Q_j(\cdot|x_{j-1},u_{j-1})}\{\alpha^j f_j(x_j^g, u_j^g)\}$$
$$= \sup_{Q_j(\cdot|x_{j-1},u_{j-1})} \mathbb{E}_{Q_j(\cdot|x_{j-1},u_{j-1})}\{\alpha^j f_j(x_j^g, u_j^{g'})\}, \quad 0 \leq j \leq n-2 \quad (4.55)$$

Adding (4.54) and (4.55) gives

$$\sup_{\substack{Q_k(\cdot|x_{k-1},u_{k-1})\\ k=0,1,\dots,n}} \mathbb{E}_{\mathbf{Q}_\nu^\pi}\{\sum_{k=0}^{n-2} \alpha^k f_k(x_k^g, u_k^g) + \alpha^{n-1} f_{n-1}(x_{n-1}^g, g_{n-1}(x_{n-1}^g)) + \alpha^n h_n(x_n^g)\}$$
$$> \sup_{\substack{Q_k(\cdot|x_{k-1},u_{k-1})\\ k=0,1,\dots,n}} \mathbb{E}_{\mathbf{Q}_\nu^\pi}\{\sum_{k=0}^{n-2} \alpha^k f_k(x_k^g, u_k^g) + \alpha^{n-1} f_{n-1}(x_{n-1}^g, g'_{n-1}(x_{n-1}^g)) + \alpha^n h_n(x_n^{g'})\}$$

and so g cannot be optimal contrary to the hypothesis. Thus $g_{n-1}(x_{n-1}^g)$ does achieve the infimum in (4.44) for $n-1$, and so $V_{n-1}(u_{[n-1,n-1]}^g, \mathcal{G}_{0,n-1}) = V_{n-1}(x_{n-1}^g)$. Now suppose by induction that $g_{j+1}(x_{j+1}^g)$ achieves the infimum and that $V_{j+1}(u_{[j+1,n-1]}^g, \mathcal{G}_{0,j+1}) = V_{j+1}(x_{j+1}^g)$. We prove this for j. Indeed, otherwise there is a function g'_j such that

$$\sup_{\substack{Q_k(\cdot|x_{k-1},u_{k-1})\\ k=j,j+1,\dots,n-1}} \mathbb{E}_{Q_j(\cdot|x_{j-1},u_{j-1})}\{\alpha^j f_j(x_j^g, g_j(x_j^g)) + V_{j+1}(x_{j+1}^g)\}$$
$$\geq \sup_{\substack{Q_k(\cdot|x_{k-1},u_{k-1})\\ k=j,j+1,\dots,n-1}} \mathbb{E}_{Q_j(\cdot|x_{j-1},u_{j-1})}\{\alpha^j f_j(x_j^g, g'_j(x_j^g)) + V_{j+1}(x_{j+1}^{g'})\}, \quad \text{w.p.1.} \quad (4.56)$$

This inequality is strict with positive probability. Consider the policy $g' = \{g_0, \dots, g_{j-1}, g'_j, g_{j+1}, \dots, g_{n-1}\}$. Then certainly

$$\sup_{Q_k(\cdot|x_{k-1},u_{k-1})} \mathbb{E}\{\alpha^k f_k(x_k^g, u_k^g)\} = \sup_{Q_k(\cdot|x_{k-1},u_{k-1})} \mathbb{E}\{\alpha^k f_k(x_k^{g'}, u_k^{g'})\}, \quad 0 \leq k \leq j-1 \quad (4.57)$$

Also, by the induction hypothesis $g_{j+1}, \dots, g_{n-1}$ achieve the infimum in (4.56) and so by Lemma 4.1

$$\sup_{Q_{j+1}(\cdot|x_j,u_j)} \mathbb{E}\{V_{j+1}^g(u_{[j+1,n-1]}^g, \mathcal{G}_{0,j+1})\} = \sup_{Q_{j+1}(\cdot|x_j,u_j)} \mathbb{E}\{V_{j+1}^g(x_{j+1}^g)\} \quad (4.58)$$
$$\sup_{Q_{j+1}(\cdot|x_j,u_j)} \mathbb{E}\{V_{j+1}^{g'}(u_{[j+1,n-1]}^{g'}, \mathcal{G}_{0,j+1})\} = \sup_{Q_{j+1}(\cdot|x_j,u_j)} \mathbb{E}\{V_{j+1}^{g'}(x_{j+1}^{g'})\} \quad (4.59)$$

From (4.56), (4.57), (4.58) and (4.59) it follows that

$$\sup_{\substack{Q_k(\cdot|x_{k-1},u_{k-1})\\ k=j,j+1,\dots,n-1}} \mathbb{E}_{Q_j(\cdot|x_{j-1},u_{j-1})}\{\sum_{k=0}^{j-1} \alpha^k f_k(x_k^g, u_k^g) + \alpha^j f_j(x_j^g, u_k^g) + V_{j+1}(x_{j+1}^g)\}$$
$$> \sup_{\substack{Q_k(\cdot|x_{k-1},u_{k-1})\\ k=j,j+1,\dots,n-1}} \mathbb{E}_{Q_j(\cdot|x_{j-1},u_{j-1})}\{\sum_{k=0}^{j-1} \alpha^k f_k(x_k^g, u_k^g) + \alpha^j f_j(x_j^g, u_k^{g'}) + V_{j+1}(x_{j+1}^{g'})\}, \text{ w.p.1.}$$

and so g cannot be optimal contrary to hypothesis. Thus, $g_j(x_j^g)$ must achieve the infimum and the result follows by induction.

4) By definition, (4.44) is also equivalent to

$$V_j(x) = \inf_{u \in \mathcal{U}(x)} \Big\{ \alpha^j f_j(x,u) + \sup_{Q_{j+1}(\cdot|x,u) \in \mathbf{B}_{R_{j+1}}(Q^o_{j+1})(x,u)} \int_{\mathcal{X}_{j+1}} V_{j+1}(z) Q_{j+1}(dz|x,u) \Big\}.$$

Hence, by applying the results of Section 3.2.1 we obtain (4.45)-(4.50). ■

Remark 4.1. *In many applications the nominal controlled process is described by*

$$x_{j+1}^g = b_j(x_j^g, u_j^g, \xi_j), \qquad x_0^g = x_0, \qquad j \in \mathbb{N}^{n-1}$$

where $\{\xi_j : j = 0, 1, \ldots, n-1\}$ *taking values in some metric space,* $\xi_j \in (\Xi_j, d)$*, is a deterministic exogenous input which belongs, for example to the space*

$$l^2(\Xi_{0,n-1}) \triangleq \left\{ \xi_j \in \Xi_j, j = 0, \ldots, n-1 : \sum_{j=0}^{n-1} |\xi_j|^2_{\Xi_j} < \infty \right\}.$$

In this case the nominal conditional distribution becomes

$$Q^o_{j+1}(A|x,u) = I_A(b_j(x,u,\xi)), \qquad A \in \mathcal{X}_j, \qquad \xi \in \Xi_j, \qquad j = 0, \ldots, n-1.$$

Remark 4.2. *We make the following observations regarding Theorem 4.2.*

(a) The dynamic programming equation (4.45), (4.46) has the interpretation of minimizing the future ambiguity. It involves in its right hand side the oscillator seminorm of $V_{j+1}(\cdot)$*, called the global modulus of continuity of* $V_{j+1}(\cdot)$*, which measures the difference between the maximum and the minimum values of* $V_{j+1}(\cdot)$.

(b) For finite and countable alphabet spaces $\mathcal{X}_j$*,* $\mathcal{X}$*, the integrals in the right hand side of (4.46) are replaced by summations.*

(c) The dynamic programming recursion (4.45), (4.46) can be applied to a controlled process with continuous alphabets and to a controlled process with finite or countable alphabets, such as Markov Decision models.

Next, we show that for any $j \in \mathbb{N}^{n-1}$, the minimax pay-off

$$V_j(x) = \inf_{\pi \in \Pi^{DM}_{j,n-1}} \sup_{\substack{Q_k(\cdot|x_{k-1},u_{k-1}) \in \mathbf{B}_{R_k}(x_{k-1},u_{k-1}) \\ k=j+1,\ldots,n}} \left\{ \sum_{k=j}^{n-1} \alpha^k f_k(x_k^g, u_k^g) + \alpha^n h_n(x_n^g) | x \right\} \tag{4.60}$$

as a function of R_j is non-decreasing and concave.

Lemma 4.3. *Suppose the conditions of Theorem 4.2 hold and in addition $R_j = R$, $j = 1,\ldots,n$. The minimax pay-off $V_j^R(x) \equiv V_j(x)$ defined by (4.60) is a non-decreasing concave function of R.*

Proof. Consider two values for $R^1, R^2 \in \mathbb{R}^+$ such that $0 \leq R^1 \leq R^2$. Since

$$\mathbf{B}_{R^1}(Q_k^o)(x_{k-1},u_{k-1}) \subseteq \mathbf{B}_{R^2}(Q_k^o)(x_{k-1},u_{k-1})$$

then for every $Q_k(\cdot,x_{k-1},u_{k-1}) \in \mathbf{B}_{R^1}(Q_k^o)(x_{k-1},u_{k-1})$ we have $Q_k(\cdot,x_{k-1},u_{k-1}) \in B_{R^2}(Q_k^o)(x_{k-1},u_{k-1})$, $k=j+1,\ldots,n-1$. Hence, $V_j^{R^1}(x) \leq V_j^{R^2}(x)$ and thus, $V_j^R(x)$ is a non-decreasing function of $R \in \mathbb{R}^+$.

Next, for a fixed $\pi \in \Pi_{j,n-1}^{DM}$ consider two points (R^1, V_j^{π,R^1}), (R^2, V_j^{π,R^2}) such that $\{Q_k^1(\cdot|x_{k-1},u_{k-1}) : k=j+1,\ldots,n\}$ achieves the supremum in (4.35) for R^1, and $\{Q_k^2(\cdot|x_{k-1},u_{k-1}) : k=j+1,\ldots,n\}$ achieves the supremum in (4.35) for R^2. Then

$$\begin{aligned}
||Q_k^1(\cdot|x_{k-1},u_{k-1}) - Q_k^o(\cdot|x_{k-1},u_{k-1})||_{TV} &\leq R^1, \quad k=j+1,\ldots,n-1\\
||Q_k^2(\cdot|x_{k-1},u_{k-1}) - Q_k^o(\cdot|x_{k-1},u_{k-1})||_{TV} &\leq R^2, \quad k=j+1,\ldots,n-1.
\end{aligned}$$

For any $\lambda \in (0,1)$ we have

$$\begin{aligned}
&||\lambda Q_k^1(\cdot|x_{k-1},u_{k-1}) + (1-\lambda)Q_k^2(\cdot|x_{k-1},u_{k-1}) - Q_k^o(\cdot|x_{k-1},u_{k-1})||_{TV}\\
&\quad \leq \lambda||Q_k^1(\cdot|x_{k-1},u_{k-1}) - Q_k^o(\cdot|x_{k-1},u_{k-1})||_{TV} + (1-\lambda)||Q_k^2(\cdot|x_{k-1},u_{k-1})\\
&\qquad - Q_k^o(\cdot|x_{k-1},u_{k-1})||_{TV} \leq \lambda R^1 + (1-\lambda)R^2, \quad k=j+1,\ldots,n. \qquad (4.61)
\end{aligned}$$

Define $Q_k^*(\cdot|x_{k-1},u_{k-1}) \triangleq \lambda Q_k^1(\cdot|x_{k-1},u_{k-1}) + (1-\lambda)Q_k^2(\cdot|x_{k-1},u_{k-1})$, $R = \lambda R^1 + (1-\lambda)R^2$. By (4.61), $Q_k^* \in \mathbf{B}_R(Q_k^o)(x_{k-1},u_{k-1})$, $k=j+1,\ldots,n$. Define the unique probability measure

$$Q_{j+1,n}^*(dx^n||u^n) \triangleq \lambda \otimes_{k=j+1}^n Q_k^1(dx_k|x_{k-1},u_{k-1}) + (1-\lambda) \otimes_{k=j+1}^n Q_k^2(dx_k|x_{k-1},u_{k-1}).$$

Then,

$$V_j^{\pi,R}(x) \geq \int \Big(\sum_{k=j}^{n-1} f_k(x_k,u_k) + h_n(x_n)\Big) Q_{j+1,n}^*(dx^n||u^n)).$$

Hence,

$$\begin{aligned}
V_j^{\pi,R}(x_j) &= \text{RHS of (4.35)}\\
&\geq \lambda \int \Big(\sum_{k=j}^{n-1} f_k(x_k,u_k) + h_n(x_n)\Big) \otimes_{k=j+1}^n Q_k^1(dx_k|x_{k-1},u_{k-1})\\
&\quad +(1-\lambda) \int \Big(\sum_{k=j}^{n-1} f_k(x_k,u_k) + h_n(x_n)\Big) \otimes_{k=j+1}^n Q_k^2(dx_k|x_{k-1},u_{k-1})\\
&= \lambda V_j^{\pi,R^1}(x_j) + (1-\lambda) V_j^{\pi,R^2}(x_j), \quad j=0,\ldots,n-1.
\end{aligned}$$

Hence, for any $\pi \in \mathbf{\Pi}^{DM}_{j,n-1}$, $V_j^{\pi,R}(x_j)$ is a concave function of R, and thus it is also concave for the $\pi \in \mathbf{\Pi}^{DM}_{j,n-1}$, which achieve the infimum in (4.60). ∎

This concavity property of the pay-off is also verified in the examples presented in Section 4.4.

Remark 4.3. *The previous results apply to randomized strategies as well.*

Relative Entropy and Exponential Functions

Related work on modeling uncertainty in probability distribution utilizes relative entropy [3, 15, 33, 44, 56] defined by (see also Chapter 2.3)

$$H(\alpha||\beta) \triangleq \begin{cases} \int_\Sigma \log(\frac{\alpha(dx)}{\beta(dx)})\alpha(dx), & \text{if } \alpha(\cdot) << \beta(\cdot), \text{ and } \log\frac{\alpha}{\beta} \in L^1(\alpha) \\ +\infty, & \text{otherwise.} \end{cases} \tag{4.62}$$

However, by Pinsker's inequality (2.5), distance in total variation of probability measures is a lower bound on relative entropy or Kullback-Leibler distance. Hence, for any fixed $\beta \in \mathcal{M}_1(\Sigma)$ then

$$\Big\{\alpha \in \mathcal{M}_1(\Sigma) : H(\alpha||\beta) \leq \frac{r^2}{2}\Big\} \subseteq \mathbb{B}_R(\beta) \equiv \Big\{\alpha \in \mathcal{M}_1(\Sigma) : ||\alpha-\beta||_{TV} \leq r\Big\}.$$

Moreover, by the definition of relative entropy (4.62), for any finite $r \in [0,\infty]$, and fixed $\beta \in \mathcal{M}_1(\Sigma)$, any ambiguity set described by relative entropy consists of only those measures $\alpha \in \mathcal{M}_1(\Sigma)$ which are absolutely continuous with $\beta \in \mathcal{M}_1(\Sigma)$. The relative entropy constraint set is defined by

$$\mathbf{A}_r(Q_i^o)(x_{i-1},u_{i-1}) \triangleq \Big\{Q_i(\cdot|x_{i-1},u_{i-1}) : H(Q_i||Q_i^o)(x_{i-1},u_{i-1}) \leq r(x_{i-1})\Big\},\ i=0,1,\ldots,n$$

where $r : \mathcal{X} \mapsto [0,\infty)$. The minimax optimization problem subject to relative entropy constraint on the conditional distribution of the controlled process is formulated as follows.

$$J_{0,n}(\pi^*, Q_k^* : k=0,\ldots,n) = \inf_{\substack{u_k\in\mathcal{U}_k(x_k)\\ k=0,1,\ldots,n-1}} \sup_{\substack{Q_k(\cdot|x_{k-1},u_{k-1})\in\mathbf{A}_r(Q_k^o)(x_{k-1},u_{k-1})\\ k=0,1,\ldots,n}} \mathbb{E}_{\mathbf{Q}_\nu^\pi}\Big\{\sum_{k=0}^{n-1}\alpha^k f_k(x_k^g,u_k^g) + \alpha^n h_n(x_n^g)\Big\}. \tag{4.63}$$

We formulate the stochastic control problem in alignment with the dynamic programming equation of total variation distance constraint. For $(i,x) \in \{0,1,\ldots,n\}\times\mathcal{X}_i$, let $V_i(x) \in \mathbb{R}$ represent the minimal cost-to-go defined by

$$V_i(x) = \inf_{\substack{u_k\in\mathcal{U}_k(x_k)\\ k=i,\ldots,n-1}} \sup_{\substack{Q_{k+1}(\cdot|x_k,u_k)\in\mathbf{A}_r(Q_{k+1}^o)(x_k,u_k)\\ k=i,\ldots,n-1}} \mathbb{E}_{i,x}^g\Big\{\sum_{j=i}^{n-1}\alpha^j f_j(x_j^g,u_j^g) + \alpha^n h_n(x_n^g)\Big\}.$$

The dynamic programming equations are given by

$$\begin{aligned} V_n(x) &= \alpha^n h_n(x), \quad x \in \mathcal{X}_n \\ V_i(x) &= \inf_{u \in \mathcal{U}_i(x)} \sup_{Q_{i+1}(\cdot|x,u) \in \mathbf{A}_r(Q^o_{i+1})(x,u)} \Big\{ \alpha^i f_i(x,u) + \int_{\mathcal{X}_{i+1}} V_{i+1}(z) Q_{i+1}(dz|x,u) \Big\}. \end{aligned}$$

By Lagrange duality theorem [40], then

$$V_i(x) = \inf_{u \in \mathcal{U}_i(x)} \inf_{s(x) \geq 0} \sup_{Q_{i+1}(\cdot|x,u): H(Q_{i+1}||Q^o_{i+1})(x,u) < \infty} \Big\{ \alpha^i f_i(x,u) + \int_{\mathcal{X}_{i+1}} V_{i+1}(z) Q_{i+1}(dz|x,u) - s(x)\Big(H(Q_{i+1}||Q^o_{i+1})(x,u) - r(x)\Big) \Big\} \tag{4.64}$$

where $s(x)$ is the Lagrange multiplier. By [48] (Proposition 2.3), the supremum over $Q_{i+1}(\cdot|x,u)$ with $H(Q_{i+1}||Q^o_{i+1})(x,u) < \infty$ is attained at

$$Q^*_{i+1}(dz|x,u) = \frac{\exp\Big(\frac{1}{s(x)} V_{i+1}(z)\Big) Q^o_{i+1}(dz|x,u)}{\int_{\mathcal{X}_{i+1}} \exp\Big(\frac{1}{s(x)} V_{i+1}(z)\Big) Q^o_{i+1}(dz|x,u)}. \tag{4.65}$$

Substituting (4.65) into (4.64) yields

$$V_n(x) = \alpha^n h_n(x), \quad x \in \mathcal{X}_n \tag{4.66}$$

$$V_i(x) = \inf_{u \in \mathcal{U}_i(x)} \inf_{s(x) \geq 0} \Big\{ \alpha^i f_i(x,u) + s(x) \log \int_{\mathcal{X}_{i+1}} \exp\Big(\frac{1}{s(x)} V_{i+1}(z)\Big) Q^o_{i+1}(dz|x,u) \Big\} + s(x) r(x). \tag{4.67}$$

The Lagrange multipliers $\inf_{s(x) \geq 0}\{\cdot\}$ can be found by the relative entropy constraint which holds with equality, i.e.,

$$H\Big(Q^*_{i+1}||Q^o_{i+1}\Big)(x,u)\Big|_{s(x)=s^*(x)} = r(x), \quad \text{for } i = 0, 1, \ldots, N-1.$$

A further elaboration on the connections between stochastic optimal control with risk-sensitive pay-off and minimax stochastic control in which the maximization is with respect to relative entropy ambiguity is found in [3, 15, 33, 44, 48, 56] (where all duality relations require that relative entropy is finite). A specific example which illustrates the differences between relative entropy and the total variation distance ambiguity, is presented analytically in Section 4.4.1.

In the next section, we illustrate through examples how the theoretical results obtained in preceding sections are applied.

4.4. Examples

In Section 4.4.1 we illustrate an application of the finite horizon minimax problem to the well-known inventory control example. Comparisons are included for the case for which the total variation constraint is replaced by the relative entropy between the nominal and true probability distributions. In Section 4.4.2 we illustrate an application of the finite horizon minimax problem to the well-known machine replacement example.

4.4.1. Inventory Control Example

Consider an inventory control example inspired by [9]. Specifically, an optimal inventory ordering policy of a quantity of a certain item at each of the N periods must be found so as to meet a stochastic demand. Let us denote

- x_k, stock available at the beginning of the kth period;
- u_k, stock ordered at the beginning of the kth period;
- w_k, demand during kth period with given probability distribution;
- h, holding cost per unit item remaining unsold at the end of the kth period;
- c, cost per unit stock ordered;
- p, shortage cost per unit demand unfilled.

The random disturbance at time k, w_k may depend on values of x_k and u_k but not on values of prior disturbances $w_0, \ldots, w_{k-1}$. Excess demand is backlogged and filled as soon as additional inventory becomes available. Inventory and demand are non-negative integers variables. Thus, we assume a nominal system given by

$$x_{k+1} = \max(0, x_k + u_k - w_k). \tag{4.68}$$

and a total sample pay-off over N periods given by

$$\sum_{k=0}^{N-1} \left(cu_k + h\max(0, x_k + u_k - w_k) + p\max(0, w_k - x_k - u_k)\right).$$

We further assume that w_k is independent and identically distributed according to $\mu_{w_k}(\cdot) = \mu_w(\cdot)$. We formulate the problem as a minimax optimization of the expected cost as follows.

$$\begin{aligned}\min_{u_k \in U_k(x_k)} \max_{\substack{\nu_{w_k}(\cdot):||\nu_{w_k}(\cdot)-\mu_w(\cdot)||_{TV}\leq R \\ k=0,\ldots,N-1}} \mathbb{E}\Bigg\{ \sum_{k=0}^{N-1} \bigg(cu_k \\ + h\max(0, x_k + u_k - w_k) + p\max(0, w_k - x_k - u_k)\bigg)\Bigg\}.\end{aligned} \tag{4.69}$$

Assume the following:

- the nominal and the true distribution of $\{w_k : k = 0, 1, \ldots, N-1\}$ is $\mu_{w_k}(\cdot) = \mu_w(\cdot)$, and $\nu_{w_k}(\cdot)$, respectively, $k = 0, 1, \ldots, N-1$;
- the maximum capacity $(x_k + u_k)$ for stock is 2 units;
- the planning horizon $N = 3$ periods;
- the holding cost h and the ordering cost c are both 1 unit;
- the shortage cost p is 3 units;
- the demand w_k has a nominal probability distribution given by, $\mu_w(w_k = 0) = 0.2$, $\mu_w(w_k = 1) = 0.7$, and $\mu_w(w_k = 2) = 0.1$, $k = 0, 1, \ldots, N-1$.

Dynamic Programming Subject to Total Variation Distance Constraint

The dynamic programming algorithm for the minimax problem subject to total variation distance uncertainty is given by

$$V_N(x_N) = 0, \tag{4.70a}$$

$$\begin{aligned} V_k(x_k) &= \min_{0 \le u_k \le 2 - x_k} \max_{\nu_{w_k}(\cdot):||\nu_{w_k}(\cdot) - \mu_w(\cdot)||_{TV} \le R} \\ &\mathbb{E}\Big\{u_k + \max(0, x_k + u_k - w_k) + 3\max(0, w_k - x_k - u_k) + V_{k+1}(\max(0, x_k + u_k - w_k))\Big\} \\ &= \min_{0 \le u_k \le 2 - x_k} \max_{\nu_{w_k}(\cdot):||\nu_{w_k}(\cdot) - \mu_w(\cdot)||_{TV} \le R} \mathbb{E}\Big\{\ell_k(x_k, u_k, w_k)\Big\} \\ &\equiv \min_{0 \le u_k \le 2 - x_k} D^+(x_k, u_k, R), \quad k = 0, 1, \ldots, N-1, \end{aligned} \tag{4.70b}$$

where

$$\begin{aligned} \ell_k(x_k, u_k, w_k) = u_k &+ \max(0, x_k + u_k - w_k) \\ &+ 3\max(0, w_k - x_k - u_k) + V_{k+1}(\max(0, x_k + u_k - w_k)). \end{aligned}$$

To address the maximization problem in (4.70b), for each $k = 0, 1, \ldots, N-1$, $x_k \in \{0, 1, 2\}$ and $0 \le u_k \le 2 - x_k$, define the maximum and minimum values of $\ell(x_k, u_k, w_k)$ by

$$\ell_{\max}(x_k, u_k) \triangleq \max_{w_k \in \{0,1,2\}} \ell(x_k, u_k, w_k), \qquad \ell_{\min}(x_k, u_k) \triangleq \min_{w_k \in \{0,1,2\}} \ell(x_k, u_k, w_k)$$

and its corresponding support sets by

$$\begin{aligned} \Sigma^0 &= \Big\{w_k \in \{0, 1, 2\} : \ell(x_k, u_k, w_k) = \ell_{\max}(x_k, u_k)\Big\}, \\ \Sigma_0 &= \Big\{w_k \in \{0, 1, 2\} : \ell(x_k, u_k, w_k) = \ell_{\min}(x_k, u_k)\Big\}. \end{aligned}$$

For all remaining sequence $\{\ell(x_k,u_k,w_k) \ : \ w_k \in \{0,1,2\} \setminus \Sigma^0 \cup \Sigma_0\}$ and for $1 \leq r \leq |\{0,1,2\} \setminus \Sigma^0 \cup \Sigma_0|$ define recursively the set of indices for which $\ell(x_k,u_k,u_k)$ achieves its $(j+1)^{th}$ smallest value by

$$\Sigma_j \triangleq \Big\{w_k \in \{0,1,2\} : \ell(x_k,u_k,w_k) = \min\Big\{\ell(x_k,u_k,\alpha_k) : \alpha_k \in \{0,1,2\} \setminus \Sigma^0 \cup \Big(\bigcup_{i=1}^{j} \Sigma_{i-1}\Big)\Big\}\Big\}, \quad j \in \{1,2,\ldots,r\},$$

till all the elements of $\{0,1,2\}$ are exhausted. Further, define

$$\ell_{\Sigma_j}(x_k,u_k) \triangleq \min_{w_k \in \{0,1,2\} \setminus \Sigma^0 \cup (\bigcup_{i=1}^{j} \Sigma_{i-1})} \ell(x_k,u_k,w_k).$$

where $j \in \{1,2,\ldots,r\}$. Once we identify the support sets and the corresponding values of the sequence $\ell(x_k,u_k,w_k)$ on these sets, we employ (3.45), (3.46) to calculate the maximizing distribution $\nu^*_{w_k}(\cdot)$ and the extremum solution of $D^+(x_k,u_k,R)$. Finally, by employing (4.70) the optimal cost-to-go and hence the optimal ordering policy are obtained. Alternatively, from the definition of the oscillator seminorm (Remark 3.1, second part), (4.70) can be expressed as follows.

$$V_N(x_N) = 0, \tag{4.71a}$$

$$\begin{aligned} V_k(x_k) = \min_{0 \leq u_k \leq 2-x_k} \Big\{ & \mathbb{E}_{\mu_w}\Big\{u_k + \max(0, x_k+u_k-w_k) + 3\max(0, w_k-x_k-u_k) \\ & + V_{k+1}(\max(0, x_k+u_k-w_k))\Big\} + \frac{R_k}{2}\Big(\max_{w_k}\Big\{u_k + \max(0, x_k+u_k-w_k) \\ & + 3\max(0, w_k-x_k-u_k) + V_{k+1}(\max(0, x_k+u_k-w_k))\Big\} - \min_{w_k}\Big\{u_k + \max(0, x_k+u_k-w_k) \\ & + 3\max(0, w_k-x_k-u_k) + V_{k+1}(\max(0, x_k+u_k-w_k))\Big\}\Big)\Big\}, \end{aligned} \tag{4.71b}$$

where $R_k = R \in [0,2]$. The problem is solved for two possible values of R for each period resulting in optimal ordering policies as shown in Table 4.1.

By setting $R = 0$, we choose to calculate the optimal control policy, when the true probability distribution $\nu_{w_k}(\cdot) = \mu_w(\cdot), \ k = 0,1,2$. This corresponds to the classical dynamic programming algorithm. From Table 4.1, the resulting optimal ordering policy for each period is to order one unit if the current stock is zero and order nothing otherwise.

By setting the total variation distance $R = 1$, we choose to calculate the optimal control policy, when the true probability distribution is $\nu_{w_k}(\cdot) \neq \mu_w(\cdot), \ k = 0,1,2$. The maximizing distribution $\nu^*_{w_k}(\cdot)$ and its corresponding support sets for each stock available, and the resulting optimal ordering policies at each stage are given in Table 4.2. Taking into consideration the maximization (that is, by setting $R > 0$) the dynamic programming algorithm results in optimal control policies which are more robust with respect to uncertainty, but with the sacrifice of low present and future costs. In cases where the planner needs to balance the desire for low costs with the undesirability of scenarios with high uncertainty, he must choose values of R between 0 and 1. From Table 4.1, the resulting optimal ordering policy for the first two periods is to order two, one and zero units if the current stock is zero, one and two, respectively. For the last period the optimal ordering policy is to order one unit if the current stock is zero and order nothing otherwise. The optimal cost-to-go and the optimal control policy, for each period and for each possible state, as a function of $R \in [0, 2]$, are illustrated in Fig 4.1. Clearly, Fig.4.1a depicts that the optimal cost-to-go is a non-decreasing concave function of R as shown in Lemma 3.1.

Dynamic Programming Subject to Relative Entropy Constraint

The dynamic programming algorithm for the minimax problem subject to relative entropy constraint is given by

$$
\begin{aligned}
V_N(x_N) &= 0\\
V_k(x_k) &= \min_{0\leq u_k\leq 2-x_k}\ \min_{s(x_k)\geq 0}\Big\{s(x_k)\log \mathbb{E}_{\mu_w}\Big\{\exp\Big(\frac{1}{s(x_k)}(u_k+\max(0,x_k+u_k-w_k)\\
&\quad +3\max(0,w_k-x_k-u_k)+V_{k+1}(\max(0,x_k+u_k-w_k)))\Big)\Big\}+s(x_k)r(x_k)\Big\}.
\end{aligned}
$$

The above dynamic programming equations are obtained by slightly modifying dynamic programming equations (4.66)-(4.67), since the cost of the inventory control example is also a function of the demand w_k. The problem with relative entropy is a convex optimization problem, and the maximization of the cost over the relative entropy is a concave non-decreasing function of $r(x) \in [0, r_{\max})$ where $r_{\max}$ can be computed. In addition, since the ambiguity set described by relative entropy is a subset of the much larger total variation ambiguity set[3], lower values of the optimal cost-to-go are obtained compared to the ones obtained under total variation ambiguity.

Fig 4.2, depicts the optimal cost-to-go and the optimal control policy, for each period and for each possible state, as a function of the relative entropy constraint. Fig 4.3, depicts a realization of the inventory control example, under the resulting optimal control policy for three possible scenarios, (i) without ambiguity[4], (ii) with ambiguity based on total variation distance and, (iii) with ambiguity based on relative entropy. In particular, the comparison

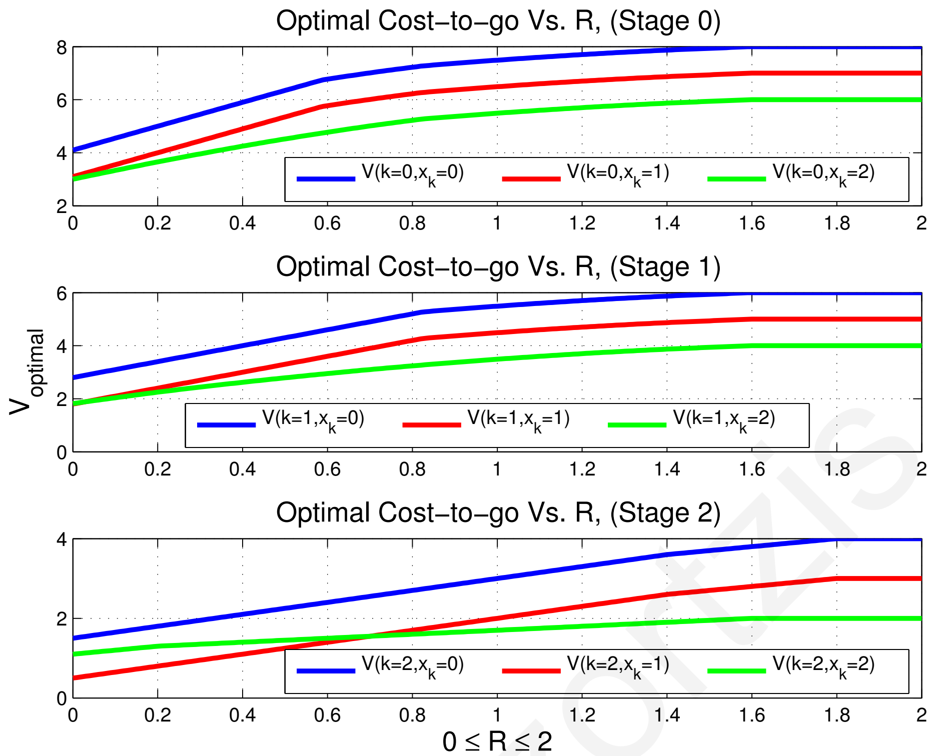
[3] See Pinsker's inequality (2.5).

[4] This scenario corresponds to the classical dynamic programming, see Fig 4.1 and/or Fig 4.2, for $R = 0$ and $r = 0$, respectively

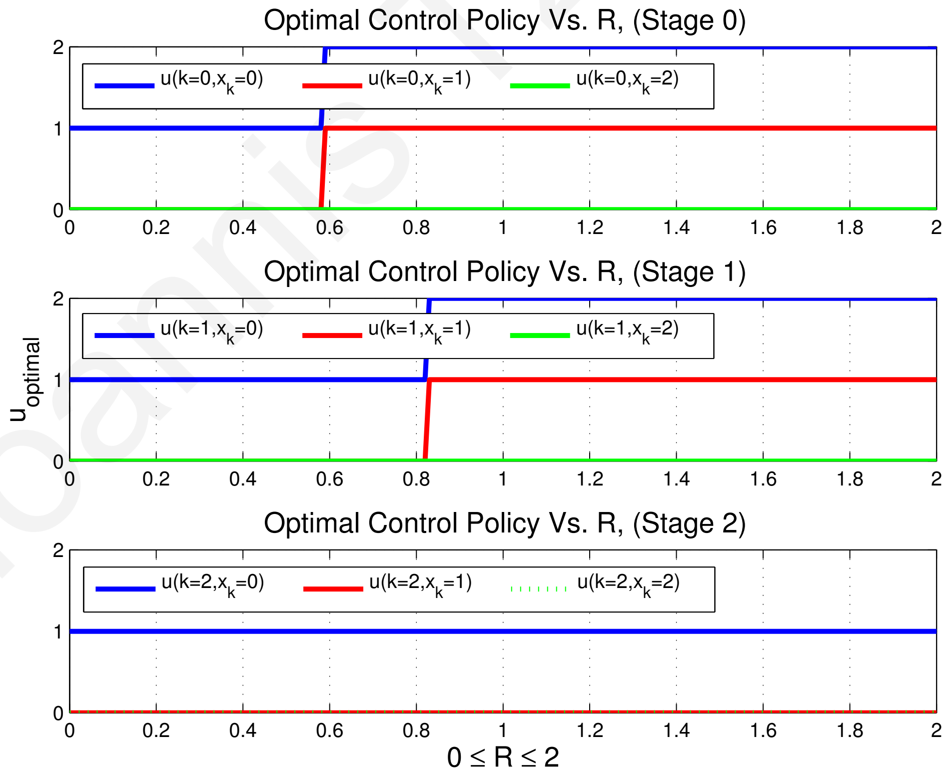
is performed by first choosing the maximizing distribution $\nu^* = [0.1 \ \ 0.3 \ \ 0.6]$ and by calculating the total variation and the relative entropy parameters. The resulting total variation parameter R is equal to one, while the resulting relative entropy parameter r is equal to 0.75. Then by extracting the optimal control policies from Fig 4.1b and 4.2b (for the corresponding value of total variation and relative entropy parameter), and by selecting the stock available x_k, and the demand w_k, for each period as shown in Fig 4.3, it is clear that, optimal control policy under total variation distance ambiguity is more robust with respect to optimal control policies with no ambiguity and with relative entropy ambiguity in which excess demand is lost. In conclusion, the dynamic programming based on relative entropy is not as general as the dynamic programming based on total variation, and in addition it has the disadvantage that it does not admit distributions which are singular with respect to the nominal distribution, and this rules out the cases in which the nominal systems are simplified versions of the true systems. This is in contrast to the dynamic programming based on total variation distance.

$R = 1$			$R = 0$		
Stock	Stage 0 Cost-to-go	Stage 0 Optimal Stock to Purchase	Stock	Stage 0 Cost-to-go	Stage 0 Optimal Stock to Purchase
0	7.49	2	0	4.1	1
1	6.49	1	1	3.1	0
2	5.49	0	2	3	0
Stock	Stage 1 Cost-to-go	Stage 1 Optimal Stock to Purchase	Stock	Stage 1 Cost-to-go	Stage 1 Optimal Stock to Purchase
0	5.49	2	0	2.8	1
1	4.49	1	1	1.8	0
2	3.49	0	2	1.82	0
Stock	Stage 2 Cost-to-go	Stage 2 Optimal Stock to Purchase	Stock	Stage 2 Cost-to-go	Stage 2 Optimal Stock to Purchase
0	3	1	0	1.5	1
1	2	0	1	0.5	0
2	1.7	0	2	1.1	0

Table 4.1.: Dynamic Programming Algorithm Results for Inventory Control Example

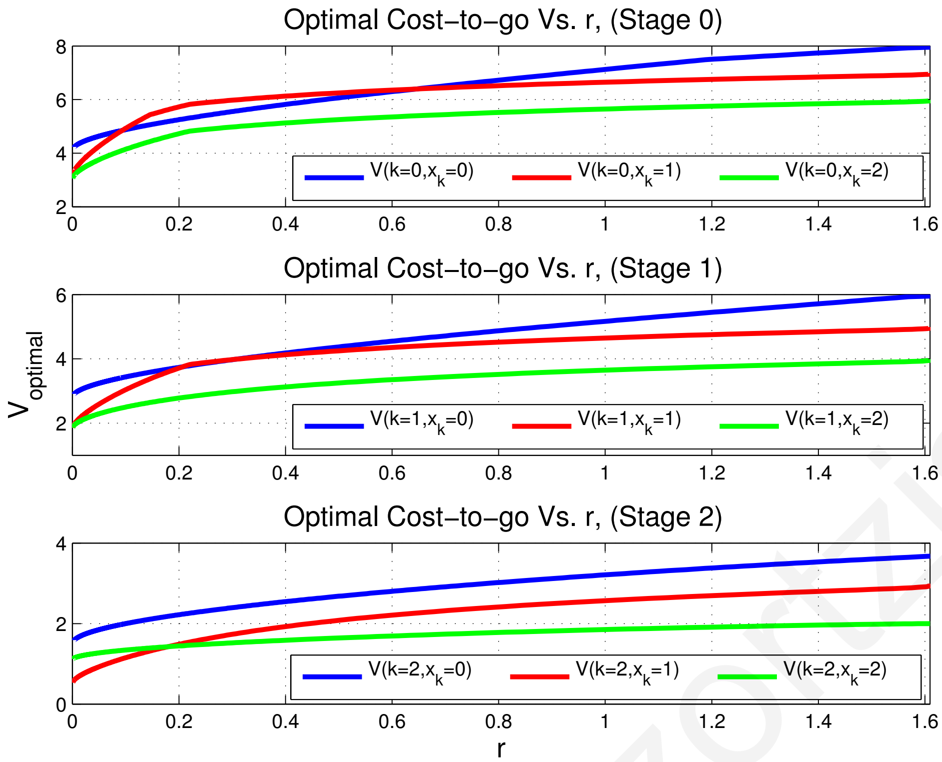


(a)

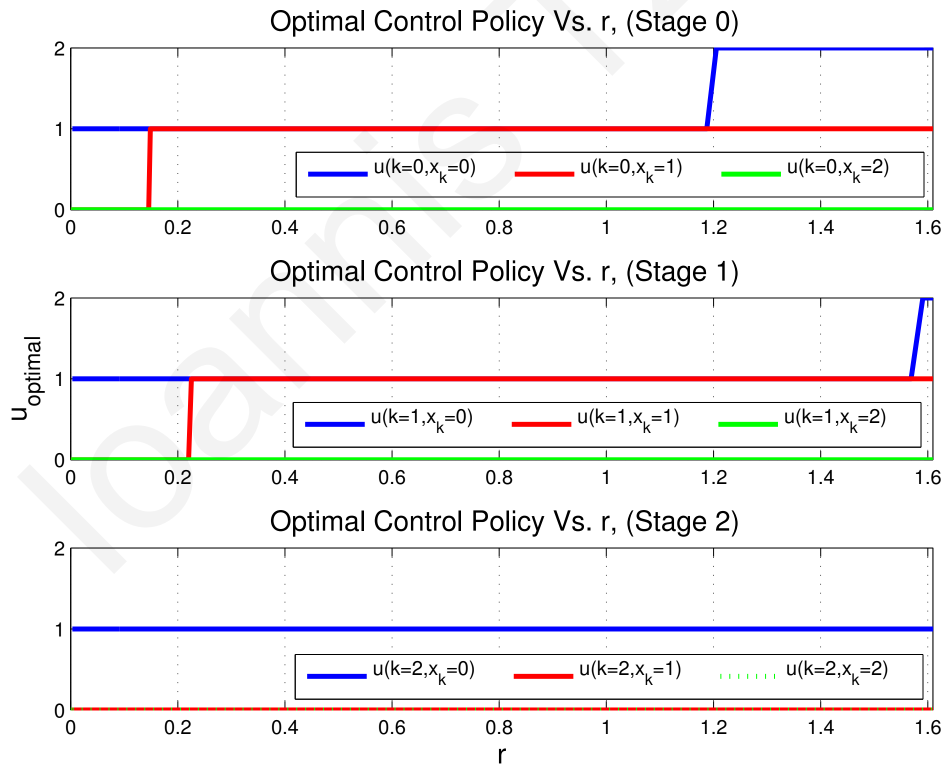


(b)

Figure 4.1.: Inventory Control Cxample with Total Variation as a Constraint: Plot (a) depicts the optimal cost-to-go. Plot (b) depicts the optimal control policy.



(a)



(b)

Figure 4.2.: Inventory Control Example with Relative Entropy as a Constraint: Plot (a) depicts the optimal cost-to-go. Plot (b) depicts the optimal control policy.

	Stock	Optimal Ordering	Support Sets			Maximizing Distribution		
			Σ^0	Σ_0	Σ_1	$\nu^*_{w_k}(\Sigma^0)$	$\nu^*_{w_k}(\Sigma_0)$	$\nu^*_{w_k}(\Sigma_1)$
Stage 0	0	2	{0,1,2}	-	-	1	-	-
	1	1	{0,1,2}	-	-	1	-	-
	2	0	{0,1,2}	-	-	1	-	-
Stage 1	0	2	{0}	{1,2}	-	0.7	0.3	-
	1	1	{0}	{1,2}	-	0.7	0.3	-
	2	0	{0}	{1,2}	-	0.7	0.3	-
Stage 2	0	1	{2}	{1}	{0}	0.6	0.2	0.2
	1	0	{2}	{1}	{0}	0.6	0.2	0.2
	2	0	{0}	{2}	{1}	0.7	0	0.3

Table 4.2.: Maximizing Distribution and Support Sets, when R=1, for Inventory Control Example

4.4.2. Machine Replacement Example

Consider a machine replacement example inspired by [9]. Specifically, we have a machine that is either running or is broken down. If it runs throughout one week, it makes a profit of € 100 for that week. If it fails during the week, the profit is zero for that week. If it is running at the beginning of the week and we perform preventive maintenance, the probability that it will fail during the week is 0.4. If we do not perform such maintenance, the probability of failure is 0.7. The maintenance cost is set to € 20. When the machine is broken down at the start of the week, it may either be repaired at a cost of € 40, in which case it will fail during the week with a probability of 0.4, or it may be replaced at a cost of € 150 by a new machine that is guaranteed to run through its first week of operation. Assume that after $N>1$ weeks the machine, irrespective of its state, is scrapped with no cost.

The system dynamics is of the form

$$x_{k+1} = f_k(x_k, u_k, w_k), \quad k = 0, 1, \ldots, N-1$$

where the state x_k is an element of a space $S_k = \{\mathbf{R}, \mathbf{B}\}$, $\mathbf{R}$ = machine running, $\mathbf{B}$ = machine broken down, the control u_k is an element of a space $\mathcal{U}_k(x_k)$, $\mathcal{U}_k(\mathbf{R}) = \{m, nm\}$, m = maintenance, nm = no maintenance, $\mathcal{U}_k(\mathbf{B}) = \{r, s\}$, r = repair, s = replace. The random disturbance has a nominal conditional distribution $w_k \sim \mu(\cdot|x_k, u_k)$.

Such a system can be described in terms of the discrete-time system equation

$$x_{k+1} = w_k, \quad k = 0, 1, \ldots, N-1$$

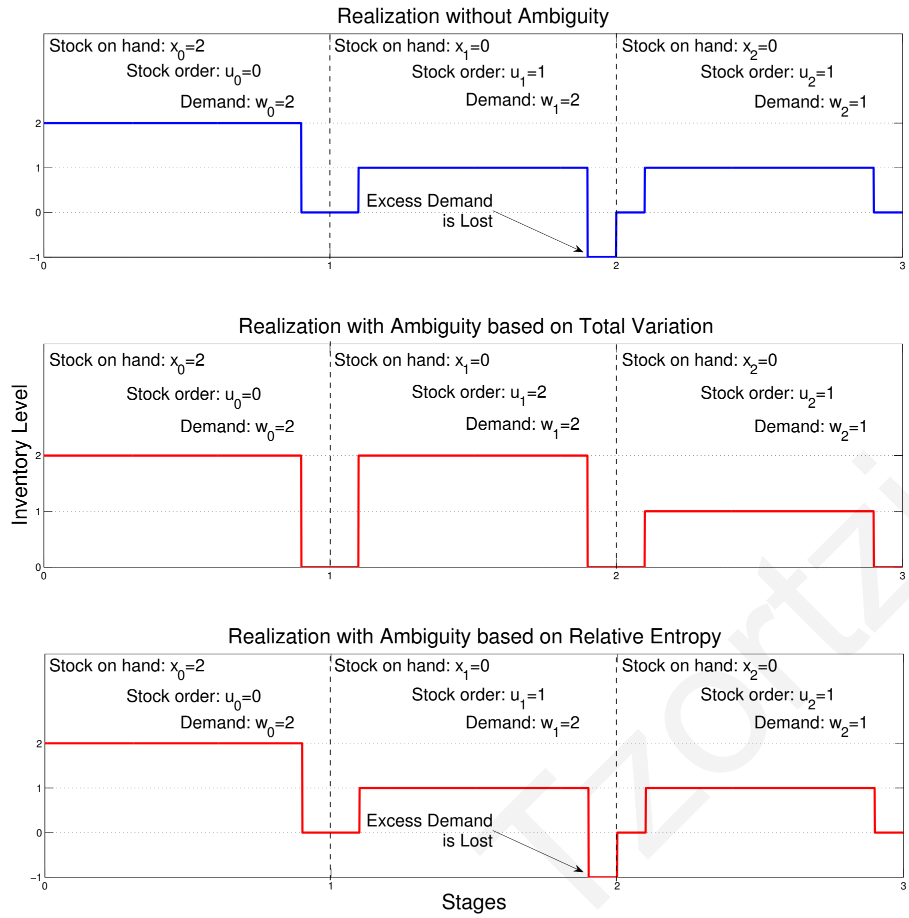


Figure 4.3.: Realization of the Inventory Control Example Under the Resulting Optimal Policy.

where the nominal probability distribution of w_k is given by

$$\begin{aligned}
&\mu(w_k = \mathbf{R}|x_k = \mathbf{R}, u_k = m) = 0.6, && \mu(w_k = \mathbf{B}|x_k = \mathbf{R}, u_k = m) = 0.4,\\
&\mu(w_k = \mathbf{R}|x_k = \mathbf{R}, u_k = nm) = 0.3, && \mu(w_k = \mathbf{B}|x_k = \mathbf{R}, u_k = nm) = 0.7,\\
&\mu(w_k = \mathbf{R}|x_k = \mathbf{B}, u_k = r) = 0.6, && \mu(w_k = \mathbf{B}|x_k = \mathbf{B}, u_k = r) = 0.4,\\
&\mu(w_k = \mathbf{R}|x_k = \mathbf{B}, u_k = s) = 1, && \mu(w_k = \mathbf{B}|x_k = \mathbf{B}, u_k = s) = 0
\end{aligned}$$

and the input costs C_u are given by: if $u = m$ then $C_m = €20$, if $u = nm$ then $C_{nm} = €0$, if $u = r$ then $C_r = €40$, and if $u = s$ then $C_s = €150$. The cost per stage is $g_k(x_k, u_k, w_k) = C_{u_k}$ if $w_k = \mathbf{R}$, and $g_k(x_k, u_k, w_k) = C_{u_k} + 100$ if $w_k = \mathbf{B}$. Since it is assumed that after N weeks the machine, irrespective of its state, is scrapped without incurring any cost the terminal cost is $g_N(\mathbf{R}) = g_N(\mathbf{B}) = 0$. The dynamic programming algorithm for the minimax problem subject to total variation distance uncertainty is given by

$$V_N(x_N) = 0 \tag{4.72}$$

$$\begin{aligned} V_k(x_k) &= \min_{u_k \in \mathcal{U}_k(x_k)} \max_{\nu(dw_k|x_k,u_k):||\nu(\cdot|x_k,u_k)-\mu(\cdot|x_k,u_k)||_{TV}\leq R} \Big\{ \\ &\quad \mathbb{E}\Big\{ g_k(x_k,u_k,w_k) + V_{k+1}(f(x_k,u_k,w_k)) \Big\} \Big\} \\ &= \min_{u_k \in \mathcal{U}_k(x_k)} \max_{\nu(dw_k|x_k,u_k):||\nu(\cdot|x_k,u_k)-\mu(\cdot|x_k,u_k)||_{TV}\leq R} \mathbb{E}\Big\{ \ell_k(x_k,u_k,w_k) \Big\} \end{aligned} \tag{4.73}$$

where $\ell_k(x_k,u_k,w_k) = g_k(x_k,u_k,w_k) + V_{k+1}(w_k)$, $k = 0,1,\ldots,N-1$. To address the maximization problem in (4.73), for each $k = 0,1,\ldots,N-1$, $x_k \in \{\mathbf{R},\mathbf{B}\}$ and $u_k \in \{m,nm,r,s\}$, define the maximum and minimum values of $\ell(x_k,u_k,w_k)$ by

$$\ell_{\max}(x_k,u_k) \triangleq \max_{w_k\in\{\mathbf{R},\mathbf{B}\}} \ell(x_k,u_k,w_k), \qquad \ell_{\min}(x_k,u_k) \triangleq \min_{w_k\in\{\mathbf{R},\mathbf{B}\}} \ell(x_k,u_k,w_k)$$

and its corresponding support sets by $\Sigma^0{=}\{w_k{\in}\{\mathbf{R},\mathbf{B}\}{:}\ell(x_k,u_k,w_k){=}\ell_{\max}(x_k,u_k)\}$, and $\Sigma_0{=}\{w_k{\in}\{\mathbf{R},\mathbf{B}\}{:}\ell(x_k,u_k,w_k){=}\ell_{\min}(x_k,u_k)\}$. By employing (3.46), the maximizing conditional probability distribution of the random parameter w_k is given by

$$\alpha = \min\left(\frac{R}{2}, 1-\mu(\Sigma^0|x_k,u_k)\right) \tag{4.74a}$$

$$\nu^*(\Sigma^0|x_k,u_k) = \mu(\Sigma^0|x_k,u_k)+\alpha, \quad \nu^*(\Sigma_0|x_k,u_k) = \Big(\mu(\Sigma_0|x_k,u_k)-\alpha\Big)^+. \tag{4.74b}$$

Based on this formulation, the dynamic programming equation is given by

$$V_N(x_N) = 0 \tag{4.75}$$

$$V_k(x_k) = \min_{u_k\in\mathcal{U}_k(x_k)} \mathbb{E}_{\nu^*(\cdot|\cdot,\cdot)}\Big\{ g_k(x_k,u_k,w_k) + V_{k+1}(f(x_k,u_k,w_k)) \Big\}. \tag{4.76}$$

We assume that the planning horizon is $N = 3$. The optimal cost-to-go and the optimal control policy, for each week and each possible state, as a function of $R \in [0,2]$ are illustrated in Fig.4.4. Clearly, Fig.4.4a depicts that the optimal cost-to-go is a non-decreasing concave function of R as stated in Lemma 4.3.

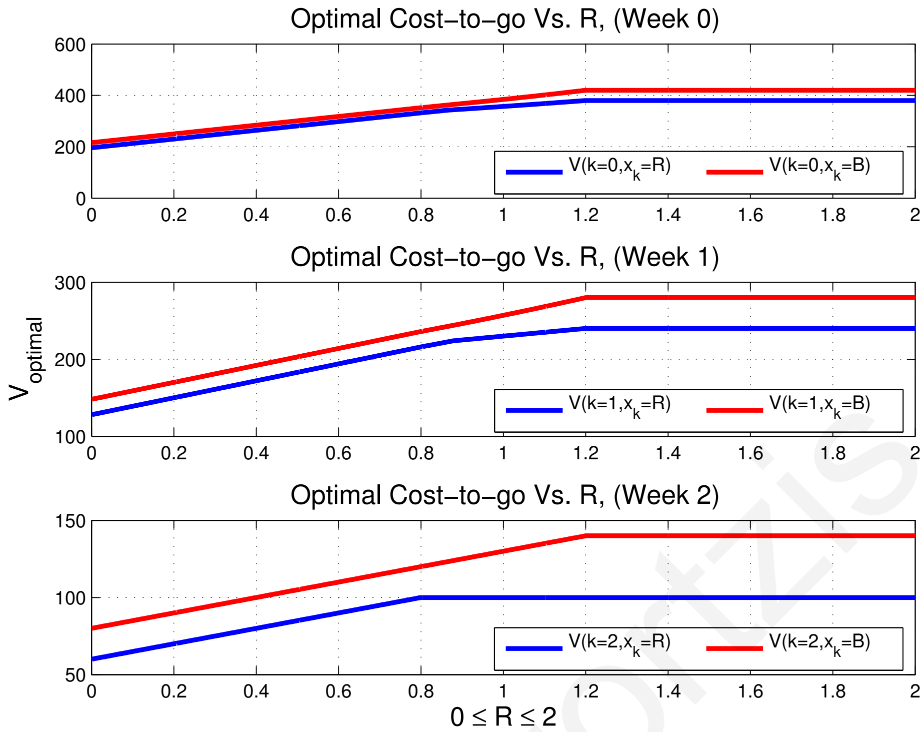
In addition, the optimum solution for two possible values of R and for each week results in optimal control policies as depicted in Table 4.3. By setting $R{=}0$, we choose to calculate the optimal control policy when the true conditional probability $\nu(\cdot|x_k,u_k) = \mu(\cdot|x_k,u_k)$, $k{=}0,1,2$. This corresponds to the classical dynamic programming algorithm. By setting $R{=}0.85$, we choose to calculate the optimal control policy when the true conditional distribution $\nu(\cdot|x_k,u_k) \neq \mu(\cdot|x_k,u_k)$, $k{=}0,1,2$. Taking into consideration the maximization (that is, by setting $R{>}0$) the dynamic programming algorithm results in optimal control policies which are more robust with respect to uncertainty, but with the sacrifice of low present and future costs. In cases in which we need to balance the desire for low costs with the undesirability of scenarios with high uncertainty, we must choose the appropriate value of R by using Fig.4.4b.

		Week 0		Week 1		Week 2	
	State	Cost-to-go	Optimal Policy	Cost-to-go	Optimal Policy	Cost-to-go	Optimal Policy
$R = 0$	**R**	196	m	128	m	60	m
	B	216	r	148	r	80	r
$R = 0.85$	**R**	340	m	221	m	100	nm
	B	360	r	241	r	122	r

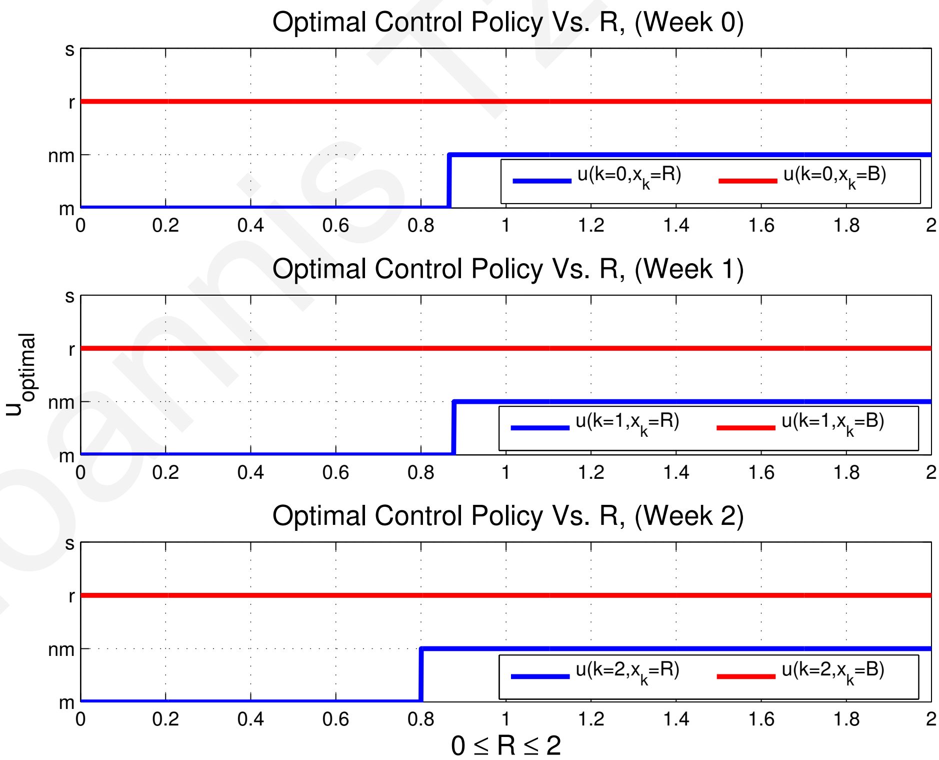
Table 4.3.: Dynamic Programming Algorithm Results for Machine Replacement Example

4.5. Summary

In this chapter, we examined the optimality of stochastic control strategies via dynamic programming, when the ambiguity class is described by the total variation distance between the conditional distribution of the controlled process and the nominal conditional distribution. The problem is formulated using minimax strategies in which the control process seeks to minimize the pay-off while the controlled process seeks to maximize it over the total variation ambiguity class. By employing certain results of Section 3, in particular, the maximization of a linear functional on the space of probability measures, among those probability measures which are within a total variation distance from a nominal probability measure, we solve the minimax stochastic control problem with deterministic control strategies, under a Markovian and a non-Markovian assumption, on the conditional distributions of the controlled process. The new dynamic programming recursion, in addition to the standard terms, it also includes the oscillator seminorm of the value function which codify the level of ambiguity with respect to total variation distance ball. Hence, the new dynamic programming recursions result in optimal control policies which are more robust with respect to uncertainty, but with the sacrifice of low present and future costs. Finally, we apply our results to the inventory control example and to the machine replacement example.



(a)



(b)

Figure 4.4.: Optimal Solution of Machine Replacement Example: Plot (a) depicts the optimal cost-to-go. Plot (b) depicts the optimal control policy ("m"= maintenance, "nm= no maintenace", "r=repair", "s=replace")

5

Dynamic Programming with TV Distance Ambiguity on an Infinite Horizon

In this chapter we address optimality of stochastic control strategies on an infinite horizon via dynamic programming subject to total variation distance ambiguity on the conditional distribution of the controlled process. For optimality criterion, we consider both the expected discounted reward and the average pay-off per unit time. Throughout this chapter, we pay particular attention to policy iteration algorithms for computing the optimal policies, which in contrast to the classical case, the policy improvement and policy evaluation steps are performed using the maximizing conditional distribution obtained under total variation distance ambiguity constraint. The new policy iterations algorithms are expected to converge to a stationary policy in a finite number of iterations, and that at each iteration a better stationary policy will be obtained. The results of this part include:

- minimax optimization subject to total variation distance ambiguity constraint;
- new infinite horizon discounted and average dynamic programming equations;
- new policy iteration algorithms;
- examples which illustrate the use of our recommended policy iteration algorithms.

5.1. Problem Formulation

In this section, we describe the abstract formulation of the minimax problem under total variation distance ambiguity, with optimality criterions being the expected discounted reward and the average pay-off per unit time.

5.1.1. Dynamic Programming of Infinite Horizon MCM

The infinite horizon D-MCM with deterministic strategies is a special case of the finite horizon D-MCM (see Definition 4.1), specified by a sextuple

$$\Big(\mathcal{X},\mathcal{U},\{\mathcal{U}(x):x\in\mathcal{X}\},\{Q(dz|x,u):(x,u)\in\mathcal{X}\times\mathcal{U}\},f,\alpha\Big) \tag{5.1}$$

consisting of the following.

(a) State Space. A Polish space $\mathcal{X}$, which model the state space of the controlled random process $\{x_k\in\mathcal{X}:k\in\mathbb{N}\}$.

(b) Control or Action Space. A Polish space $\mathcal{U}$, which model the control or action set of the control random process $\{u_k\in\mathcal{U}:k\in\mathbb{N}\}$.

(c) Feasible Controls or Actions. A family $\{\mathcal{U}(x):x\in\mathcal{X}\}$ of non-empty measurable subsets $\mathcal{U}(x)$ of $\mathcal{U}$, where $\mathcal{U}(x)$ denotes the set of feasible controls or actions, when the controlled process is in state $x\in\mathcal{X}$, and the feasible state-actions pairs defined by $\mathbb{K}\triangleq\Big\{(x,u):x\in\mathcal{X},u\in\mathcal{U}(x)\Big\}$ are measurable subsets of $\mathcal{X}\times\mathcal{U}$.

(d) Controlled Process Distribution. A conditional distribution or stochastic kernel $Q(dx|x,u)$ on $\mathcal{X}$ given $(x,u)\in\mathbb{K}\subseteq\mathcal{X}\times\mathcal{U}$. The controlled process distribution is described by the transition probability distribution $\{Q(dx|x,u):(x,u)\in\mathbb{K}\}$.

(e) One-Stage-Cost. A non-negative measurable function $f:\mathbb{K}\rightarrow[0,\infty]$, called the one-stage-cost, such that $f(x,\cdot)$ does not take the value $+\infty$ for each $x\in\mathcal{X}$.

(f) Discounting Factor. A real number $\alpha\in(0,1)$ called the discounting factor.

The dynamic programming equation of the infinite horizon D-MCM as given by [54] is a function $v^0_\infty:\mathcal{X}\longrightarrow\mathbb{R}$ satisfying

$$v^0_\infty(x)=\inf_{u\in\mathcal{U}(x)}\Big\{f(x,u)+\alpha\int_{\mathcal{X}}v^0_\infty(z)Q(dz|x,u)\Big\},\quad x\in\mathcal{X}. \tag{5.2}$$

This equation is also obtained from (4.4), and (4.5), (4.6), by assuming $V_n^0(x) = 0$, $f_i = f$, $\mathcal{X}_i = \mathcal{X}, \mathcal{U}_i = \mathcal{U}$, $Q_i(\cdot|\cdot) = Q(\cdot|\cdot)$, $\forall i$, as follows. Define

$$v_i^0(x) \triangleq \alpha^{i-n} V_{n-i}^0(x)$$

and subsequently find the equation for $v_i^0(x)$ from (4.5), (4.6), and then take the limit as $n \longrightarrow \infty$ to obtain $v_\infty^0(x)$ satisfying (5.2), which implies

$$v_n^0(x) = \inf_{\substack{g_k \in \mathcal{U}(x_k)\\ k=0,1,\ldots,n-1}} \mathbb{E}_{0,x}^g \Big\{ \sum_{j=0}^{n-1} \alpha^j f(x_j^g, u_j^g) \Big\} \tag{5.3}$$

$$v_\infty^0(x) = \inf_{\substack{g_k \in \mathcal{U}(x_k)\\ k=0,1,\ldots}} \mathbb{E}_{0,x}^g \Big\{ \sum_{j=0}^{\infty} \alpha^j f(x_j^g, u_j^g) \Big\}. \tag{5.4}$$

Similarly to the finite horizon D-MCM, the dynamic programming equation (5.2) depends on the conditional distribution $Q(dz|x,u)$. Hence, any ambiguity or mismatch of $Q(dz|x,u)$ from the true distribution affects optimality of the strategies.

In this chapter, we will also study the dynamic programming of infinite horizon MCM with an average pay-off per unit time. First, recall the definition of infinite horizon D-MCM specified by (5.1), with discounting factor $\alpha = 1$, and consider the problem of minimizing the average pay-off per unit time

$$J(\pi) = \limsup_{j \to \infty} \Big\{ \frac{1}{j} \mathbb{E}_{\mathbf{Q}_\nu^\pi} \Big\{ \sum_{k=0}^{j-1} f(x_k, u_k) \Big\} \Big\}. \tag{5.5}$$

In [54], it is shown that under the assumption that for every stationary Markov control law the transition probability matrix $Q(g)$ is irreducible, then there exists a solution $V : \mathcal{X} \longmapsto \mathbb{R}$ and $J^* \in \mathbb{R}$ to the dynamic programming equation

$$J^* + V(x) = \inf_{u \in \mathcal{U}(x)} \Big\{ f(x,u) + \int_{\mathcal{X}} Q(dz|x,u) V(z) \Big\}. \tag{5.6}$$

By the same reasoning as before, the above dynamic programming equation depends on the conditional distribution $Q(dz|x,u)$, hence any ambiguity or mismatch from the true distribution affects the optimality of the strategies.

5.1.2. Dynamic Programming with Total Variation Distance Ambiguity

The objective of this chapter is to investigate dynamic programming under ambiguity of the conditional distributions of the controlled process

$$\{Q(dz|x,u) : (x,u) \in \mathbb{K}\}. \tag{5.7}$$

Specifically, given a collection of nominal controlled process distributions $\{Q^o(dz|x,u) : (x,u) \in \mathbb{K}\}$ the corresponding collection of true controlled process distributions $\{Q(dz|x,u) : (x,u) \in \mathbb{K}\}$ is modeled by a set described by the total variation distance centered at the nominal conditional distribution having radius $R \in [0,2]$ defined by

$$\mathbf{B}_R(Q^o)(x,u) \triangleq \Big\{Q(\cdot|x,u) : ||Q(\cdot|x,u) - Q^o(\cdot|x,u)||_{TV} \leq R\Big\}. \tag{5.8}$$

Given the above description the dynamic programming equation of the infinite horizon D-MCM is given by

$$v_\infty(x) = \inf_{u\in\mathcal{U}(x)} \sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \Big\{f(x,u) + \alpha\int_{\mathcal{X}} v_\infty(z)Q(dz|x,u)\Big\}, \quad x \in \mathcal{X} \tag{5.9}$$

and, the dynamic programming equation of the infinite horizon MCM with an average pay-off is given by

$$J^* + V(x) = \inf_{u\in\mathcal{U}(x)} \sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \Big\{f(x,u) + \int_{\mathcal{X}} V(z)Q(dz|x,u)\Big\}. \tag{5.10}$$

In summary, the contributions of this chapter are the following:

(a) formulation of the infinite horizon MCM and dynamic programming equation under conditional distribution ambiguity described by total variation distance via minimax theory;

(b) characterization of the maximizing conditional distribution belonging to the total variation distance set, and the corresponding new dynamic programming recursions;

(c) contraction property of the infinite horizon D-MCM dynamic programming;

(d) discussion of the limitations of the infinite horizon MCM with an average pay-off, under irreducibility assumption, and new general dynamic programming recursions;

(e) new policy iteration algorithms;

(f) examples for the infinite horizon case.

5.2. Minimax Stochastic Control - Discounted Cost

In this section, we consider the infinite horizon version of the finite horizon D-MCM, and we derive similar results. In addition, we show that the operator associated with the dynamic programming equation is contractive, and we introduce a new policy iteration algorithm.

Consider the problem of minimizing the finite horizon cost

$$\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^\pi}\Big\{\sum_{j=0}^{n-1}\alpha^j f(x_j^g,u_j^g)\Big\} \tag{5.11}$$

with $0<\alpha<1$. By Theorem 4.2 the value function of (5.11), denoted by $V_j(x)$, $j=0,\ldots,n$, $x\in\mathcal{X}_j$ satisfies the dynamic programming equations (4.45), (4.46) with $h_n=0$, $f_j=f$, $R_j=R$, $\mathcal{X}_j=\mathcal{X}$, $\mathcal{U}_j=\mathcal{U}$, $\mathcal{U}_j(x)=\mathcal{U}(x)$ and $Q_j^o(\cdot|\cdot)=Q^o(\cdot|\cdot)$. Define $v_i(x)=\alpha^{i-n}V_{n-i}(x)$, where $0\le i\le n$ is the time to go, (see [54]). Then,

$$v_0(x)=0 \tag{5.12}$$

$$v_i(x)=\inf_{u\in\mathcal{U}(x)}\Big\{f(x,u)+\alpha\int_{\mathcal{X}}v_{i-1}(z)Q^o(dz|x,u)+\alpha\frac{R}{2}\Big(\sup_{z\in\mathcal{X}}v_{i-1}(z)-\inf_{z\in\mathcal{X}}v_{i-1}(z)\Big)\Big\} \tag{5.13}$$

which is obtained as follows:

$$v_0(x)=\alpha^{-n}V_n(x)=0,\quad (\text{since } h_n=0)$$

and

$$\begin{aligned}v_i(x)&=\alpha^{i-n}V_{n-i}(x)\\&=\inf_{u\in\mathcal{U}(x)}\Big\{\alpha^{i-n}\alpha^{n-i}f(x,u)+\alpha^{i-n}\int_{\mathcal{X}}V_{n-i+1}(z)Q^o(dz|x,u)\\&\qquad\qquad+\alpha^{i-n}\frac{R}{2}\Big(\sup_{z\in\mathcal{X}}V_{n-i+1}(z)-\inf_{z\in\mathcal{X}}V_{n-i+1}(z)\Big)\Big\}\\&=\inf_{u\in\mathcal{U}(x)}\Big\{f(x,u)+\alpha\int_{\mathcal{X}}v_{i-1}(z)Q^o(dz|x,u)+\alpha\frac{R}{2}\Big(\sup_{z\in\mathcal{X}}v_{i-1}(z)-\inf_{z\in\mathcal{X}}v_{i-1}(z)\Big)\Big\}.\end{aligned}$$

In contrast with finite horizon case, the one given by (5.12)-(5.13) proceeds from lower to higher values of indices i. The dynamic programming for the discounted cost

$$\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^\pi}\Big\{\sum_{j=0}^{\infty}\alpha^j f(x_j^g,u_j^g)\Big\} \tag{5.14}$$

is given by

$$v_\infty(x)=\inf_{u\in\mathcal{U}(x)}\Big\{f(x,u)+\alpha\int_{\mathcal{X}}v_\infty(z)Q^o(dz|x,u)+\alpha\frac{R}{2}\Big(\sup_{z\in\mathcal{X}}v_\infty(z)-\inf_{z\in\mathcal{X}}v_\infty(z)\Big)\Big\}. \tag{5.15}$$

The maximizing conditional distribution is

$$Q^*\left(\mathcal{X}^+|x,u\right)=Q^o(\mathcal{X}^+|x,u)+\frac{R}{2}\in[0,1],\quad (x,u)\in\mathbb{K} \tag{5.16}$$

$$Q^*\left(\mathcal{X}^-|x,u\right)=Q^o(\mathcal{X}^-|x,u)-\frac{R}{2}\in[0,1],\quad (x,u)\in\mathbb{K} \tag{5.17}$$

$$Q^*\left(A|x,u\right)=Q^o(A|x,u),\quad \forall A\subseteq\mathcal{X}\setminus\mathcal{X}^+\cup\mathcal{X}^-,\quad (x,u)\in\mathbb{K} \tag{5.18}$$

where

$$\mathcal{X}^{+} \triangleq \Big\{x{\in}\mathcal{X} : V(x) = \sup\{V(x) : x{\in}\mathcal{X}\}\Big\} \tag{5.19}$$

$$\mathcal{X}^{-} \triangleq \Big\{x{\in}\mathcal{X} : V(x) = \inf\{V(x) : x{\in}\mathcal{X}\}\Big\}. \tag{5.20}$$

Next, we recall the following theorem from [54], which we invoke to show that the operator in the right hand side of (5.15) is contractive.

Theorem 5.1. *Let $(L, ||\cdot||)$ be a complete normed space and let $T : L \longrightarrow L$ satisfy the following inequality for some $0 < \alpha < 1$,*

$$||TV_1 - TV_2|| \leq \alpha ||V_1 - V_2||, \quad \textit{for all} \quad V_1, V_2 \in L. \tag{5.21}$$

A mapping T satisfying (5.21) is called a contraction mapping.
The following hold.

1) There exists a unique $w \in L$ satisfying $Tw = w$, called the fixed point of T.

2) For $V \in L$, define $\{T^n V : n \in \mathbb{Z}_+\}$ by $TV = V$, $T^{n+1}V = T^n(TV)$ then

$$\lim_{n\longrightarrow\infty} ||T^n V - w|| = 0, \quad \textit{for all} \quad V \in L, \tag{5.22}$$

where w is the fixpoint defined in 1).

Lemma 5.1. *Let L be the class of all measurable functions $V : \mathcal{X} \longrightarrow \mathbb{R}$, with finite norm $||V|| \triangleq \max_{x\in\mathcal{X}} |V(x)|$, and $T : L \longmapsto L$ defined by*

$$(TV)(x) = \inf_{u\in\mathcal{U}(x)} \Big\{ f(x,u) + \alpha \int_{\mathcal{X}} V(z) Q^o(dz|x,u) + \alpha\frac{R}{2}\Big(\sup_{z\in\mathcal{X}} V(z) - \inf_{z\in\mathcal{X}} V(z)\Big)\Big\}. \tag{5.23}$$

If $V \in BC^+(\mathcal{X})$ and $\sup_{z\in\mathcal{X}} V(z)$, $\inf_{z\in\mathcal{X}} V(z)$ are finite, then T is a contraction.

Proof. For $V_1, V_2 \in L$,

$$\begin{aligned}(TV_1)(x) - (TV_2)(x) = &\\ \inf_{u\in\mathcal{U}(x)} &\Big\{ f(x,u) + \alpha \int_{\mathcal{X}} V_1(z) Q^o(dz|x,u) + \alpha\frac{R}{2}\Big(\sup_{z\in\mathcal{X}} V_1(z) - \inf_{z\in\mathcal{X}} V_1(z)\Big)\Big\} \\ &- \inf_{u\in\mathcal{U}(x)} \Big\{ f(x,u) + \alpha \int_{\mathcal{X}} V_2(z) Q^o(dz|x,u) + \alpha\frac{R}{2}\Big(\sup_{z\in\mathcal{X}} V_2(z) - \inf_{z\in\mathcal{X}} V_2(z)\Big)\Big\}.\end{aligned}$$

Let

$$v \triangleq \arg \inf_{u\in\mathcal{U}(x)} \Big\{ f(x,u) + \alpha \int_{\mathcal{X}} V_2(z) Q^o(dz|x,u) + \alpha\frac{R}{2}\Big(\sup_{z\in\mathcal{X}} V_2(z) - \inf_{z\in\mathcal{X}} V_2(z)\Big)\Big\}.$$

Then,

$$
\begin{aligned}
&(TV_1)(x) - (TV_2)(x)\\
&= \inf_{u \in \mathcal{U}(x)} \Big\{ f(x,u) + \alpha \int_{\mathcal{X}} V_1(z) Q^o(dz|x,u) + \alpha \frac{R}{2} \Big(\sup_{z\in\mathcal{X}} V_1(z) - \inf_{z\in\mathcal{X}} V_1(z) \Big) \Big\}\\
&\qquad\qquad - \Big\{ f(x,v) + \alpha \int_{\mathcal{X}} V_2(z) Q^o(dz|x,v) + \alpha \frac{R}{2} \Big(\sup_{z\in\mathcal{X}} V_2(z) - \inf_{z\in\mathcal{X}} V_2(z) \Big) \Big\}\\
&\leq \Big\{ f(x,v) + \alpha \int_{\mathcal{X}} V_1(z) Q^o(dz|x,v) + \alpha \frac{R}{2} \Big(\sup_{z\in\mathcal{X}} V_1(z) - \inf_{z\in\mathcal{X}} V_1(z) \Big) \Big\}\\
&\qquad\qquad - \Big\{ f(x,v) + \alpha \int_{\mathcal{X}} V_2(z) Q^o(dz|x,v) + \alpha \frac{R}{2} \Big(\sup_{z\in\mathcal{X}} V_2(z) - \inf_{z\in\mathcal{X}} V_2(z) \Big) \Big\}\\
&\overset{(a)}{=} \Big\{ \alpha \int_{\mathcal{X}} V_1(z) Q^{V_1}(dz|x,v) \Big\} - \Big\{ \alpha \int_{\mathcal{X}} V_2(z) Q^o(dz|x,v) + \alpha \frac{R}{2} \Big(\sup_{z\in\mathcal{X}} V_2(z) - \inf_{z\in\mathcal{X}} V_2(z) \Big) \Big\}\\
&\overset{(b)}{\leq} \Big\{ \alpha \int_{\mathcal{X}} V_1(z) Q^{V_1}(dz|x,v) \Big\} - \Big\{ \alpha \int_{\mathcal{X}} V_2(z) Q^{V_1}(dz|x,v) \Big\}\\
&= \alpha \int_{\mathcal{X}} \big(V_1(z) - V_2(z)\big) Q^{V_1}(dz|x|v) \leq \alpha \sup_{z\in\mathcal{X}} |V_1(z) - V_2(z)| = \alpha ||V_1 - V_2||
\end{aligned}
$$

where (a) is obtained by applying (3.14), with $\ell \equiv \alpha V_1$, $\nu^*(\cdot) \equiv Q^{V_1}(\cdot|\cdot)$, $\mu(\cdot) \equiv Q^o(\cdot|\cdot)$, and (b) is obtained by first applying (3.14) as in (a) with $Q^{V_2}(\cdot|\cdot)$ and then replace $Q^{V_2}(\cdot|\cdot)$ by $Q^{V_1}(\cdot|\cdot)$ which is suboptimal hence, the upper bound. By reversing the roles of V_1 and V_2 we get $(TV_2)(x) - (TV_1)(x) \leq \alpha ||V_2 - V_1||$. Hence, $|(TV_1)(x) - (TV_2)(x)| \leq \alpha ||V_1 - V_2||$ for all $x \in \mathcal{X}$, and

$$
||TV_1 - TV_2|| \triangleq \max_{x\in\mathcal{X}} |(TV_1)(x) - (TV_2)(x)| \leq \alpha ||V_1 - V_2||
$$

which implies that the operator $T : L \longmapsto L$ is a contraction. ■

Utilizing Lemma 5.1 we obtain the following theorem which is analogous to the classical result given in [54].

Theorem 5.2. *Assume $v_\infty \in BC^+(\mathcal{X})$ and $\sup_{z\in\mathcal{X}} v_\infty(z)$, $\inf_{z\in\mathcal{X}} v_\infty(z)$ are finite.*

(1) The dynamic programming equation

$$
v_\infty(x) = \inf_{u\in\mathcal{U}(x)} \Big\{ f(x,u) + \alpha \int_{\mathcal{X}} v_\infty(z) Q^o(dz|x,u) + \alpha \frac{R}{2} \Big(\sup_{z\in\mathcal{X}} v_\infty(z) - \inf_{z\in\mathcal{X}} v_\infty(z) \Big) \Big\}
$$

has a unique solution.

(2) Moreover,

$$
v_\infty(x) = \inf_{g\in\mathcal{U}(x)} \mathbb{E}_{Q^*} \Big\{ \sum_{j=0}^{\infty} \alpha^j f(x_j, u_j) | x_0 = x \Big\}.
$$

(3) The mapping T defined by

$$(TV)(x)= \inf_{u\in\mathcal{U}(x)} \Big\{f(x,u)+\alpha\int_{\mathcal{X}} V(z)Q^o(dz|x,u)+\alpha\frac{R}{2}\Big(\sup_{z\in\mathcal{X}} V(z)- \inf_{z\in\mathcal{X}} V(z)\Big)\Big\}$$

is a contraction mapping with respect to the norm $||V|| = \max_{x\in\mathcal{X}} |V(x)|$.

(4) For any V, $\lim_{n\to\infty} ||T^n V - v_\infty|| = 0$ and so

$$\lim_{n\longrightarrow\infty} (T^n V)(x) = v_\infty(x), \quad \textit{for all } \ x \in \mathcal{X}.$$

Proof. (1) Follows from [54] (Theorem 6.3.6, part (a)).

(2) We need to show that $v_\infty(x)$ is the minimum value of $\mathbb{E}_{Q^*}\left\{\sum_{j=0}^{\infty} \alpha^j f(x_j,u_j)\right\}$ starting in state $x_0 = x$. Recall that $0 \leq f(x,u) \leq M$ for all $x \in \mathcal{X}$, $u \in \mathcal{U}(x)$. Clearly, with $x_0 = x$ and for all n,

$$\inf_{g\in\mathcal{U}(x)} \mathbb{E}_{Q^*}\Big\{\sum_{j=0}^{\infty} \alpha^j f(x_j,u_j)\Big\} \geq \inf_{g\in\mathcal{U}(x)} \mathbb{E}_{Q^*}\Big\{\sum_{j=0}^{n-1} \alpha^j f(x_j,u_j)\Big\} = v_n(x).$$

Hence, $\inf_{g\in\mathcal{U}(x)} \mathbb{E}_{Q^*}\Big\{ \sum_{j=0}^{\infty} \alpha^j f(x_j,u_j)\Big\} \geq \lim_{n\to\infty} v_n(x) = v_\infty(x)$. Conversely, for all n

$$\inf_{g\in\mathcal{U}(x)} \mathbb{E}_{Q^*}\Big\{\sum_{j=0}^{\infty} \alpha^j f(x_j,u_j)\Big\}\leq \inf_{g\in\mathcal{U}(x)} \mathbb{E}_{Q^*}\Big\{\sum_{j=0}^{n-1} \alpha^j f(x_j,u_j)\Big\}+\sum_{j=n}^{\infty} \alpha^j M=v_n(x)+\frac{\alpha^n M}{1-\alpha}$$

and so

$$\inf_{g\in\mathcal{U}(x)} \mathbb{E}_{Q^*}\Big\{\sum_{j=0}^{\infty} \alpha^j f(x_j,u_j)\Big\} \leq \lim_{n\to\infty} \Big[v_n(x) + \frac{\alpha^n M}{1-\alpha}\Big] = v_\infty(x).$$

Hence, $\inf_{g\in\mathcal{U}(x)} \mathbb{E}_{Q^*}\Big\{ \sum_{j=0}^{\infty} \alpha^j f(x_j,u_j)\Big\} = v_\infty(x)$.

(3) This follows from Lemma 5.1.

(4) Follows from [54] (Theorem 6.3.6, part (b)). ■

5.2.1. Policy Iteration Algorithm

Next, we present a modified version of the classical policy iteration algorithm [39,54]. From part 4 of Theorem 4.2, the policy improvement and policy evaluation steps of a policy iteration algorithm must be performed using the maximizing conditional distribution obtained under total variation distance ambiguity constraint. Hence, in addition to the classical case, in which the policy improvement and evaluation steps are performed using the nominal conditional distribution, here, under the assumption that $f(\cdot)$ is bounded and non-negative, by invoking the results developed in earlier sections we propose a modified algorithm which is expected to converge to a stationary policy in a finite number of iterations, since both state space $\mathcal{X}$ and control space $\mathcal{U}$ are finite sets, and that at each iteration a better stationary policy will be obtained.

First, we introduce some notation. Since the state space $\mathcal{X}$ is a finite set, with say, n elements, any function $V : \mathcal{X} \longrightarrow \mathbb{R}^n$ may be represented by vector in $\mathbb{R}^n$ defined by

$$V(x) \triangleq \begin{pmatrix} V(x_1) & \cdots & V(x_n) \end{pmatrix}^T \in \mathbb{R}^n.$$

Write $z \leq y$, if $z(i) \leq y(i)$, for $\forall i \in \mathbb{Z}^n \triangleq \{1, 2, \ldots, n\}$; and $z < y$ if $z \leq y$ and $z \neq y$. For a stationary control law g, let

$$f(g) = \begin{pmatrix} f(x_1, g(x_1)) & \cdots & f(x_n, g(x_n)) \end{pmatrix}^T$$

and define each entry of the transition matrix $Q^o(g) \in \mathbb{R}^{n \times n}$ by $Q^o_{ij}(g) = Q^o(x_j|x_i, g(x_i)) \equiv Q^{g,o}(x_i|x_j)$. Rewrite (5.23) (with $\sup_{z \in \mathcal{X}} V(z)$ denoting componentwise supremum, and similarly for the infimum) as

$$TV = \min_{g \in \mathbb{R}^n} \left\{ f(g) + \alpha Q^o(g)V + \alpha \frac{R}{2} \left\{ \sup_{z \in \mathcal{X}} V(z) - \inf_{z \in \mathcal{X}} V(z) \right\} \right\}$$

which by Theorem 4.2 is equivalent to

$$TV = \min_{g \in \mathbb{R}^n} \left\{ f(g) + \alpha Q^*(g)V \right\}$$

where $Q^*(g) \in \mathbb{R}^{n \times n}$ and is given by (5.16)-(5.18). Note that, the minimization is taken componentwise, i.e., $g(x_1)$ is the minimum of the first component of $f(g) + \alpha Q^*(g)V$ and so on. For each stationary policy g, define $T(g) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ by

$$T(g)V = f(g) + \alpha Q^*(g)V.$$

Then, $T(g)$ is a contraction mapping on the space of bounded continuous functions to itself, and from Theorem 5.2 it follows that

$$V(g) = T(g)V = f(g) + \alpha Q^*(g)V$$

has a unique solution $V(g) \in \mathbb{R}^n$. Next, we give the policy iteration algorithm.

Algorithm 5.3 (Policy Iteration)**.** *Consider the notation above.*

Initialization. Let $m = 0$ and select $g_0 : \mathcal{X} \longmapsto \mathcal{U}$ be an arbitrary stationary control law. Solve the equation

$$f(g_0) + \alpha Q^o(g_0)V_{Q^o}(g_0) = V_{Q^o}(g_0) \;\; for \;\; V_{Q^o}(g_0) \in \mathbb{R}^n. \tag{5.24}$$

Identify the support sets using (5.19)-(5.20) and the analogue of Σ_k of Section 3.3.1, and construct the matrix $Q^*(g_0)$ using (5.16)-(5.18). Solve the equation

$$f(g_0) + \alpha Q^*(g_0) V_{Q^*}(g_0) = V_{Q^*}(g_0) \;\; \textit{for} \;\; V_{Q^*}(g_0) \in \mathbb{R}^n. \tag{5.25}$$

1. *For $m = m + 1$ while*

$$\min_{g \in \mathbb{R}^n} \Big\{ f(g) + \alpha Q^*(g) V_{Q^*}(g_{m-1}) \Big\} < V_{Q^*}(g_{m-1}) \tag{5.26}$$

do:

   (a) *(Policy Improvement) Let $g_m \in \mathbb{R}^n$ be such that*

$$f(g_m) + \alpha Q^*(g_m) V_{Q^*}(g_{m-1}) = \min_{g \in \mathbb{R}^n} \Big\{ f(g) + \alpha Q^*(g) V_{Q^*}(g_{m-1}) \Big\}. \tag{5.27}$$

   (b) *(Policy Evaluation) Solve the following equation for $V_{Q^o}(g_m) \in \mathbb{R}^n$*

$$f(g_m) + \alpha Q^o(g_m) V_{Q^o}(g_m) = V_{Q^o}(g_m). \tag{5.28}$$

   Identify the support sets using (5.19)-(5.20), and construct the matrix $Q^*(g_m)$ using (5.16)-(5.18). Solve the equation

$$f(g_m) + \alpha Q^*(g_m) V_{Q^*}(g_m) = V_{Q^*}(g_m) \;\; \textit{for} \;\; V_{Q^*}(g_m) \in \mathbb{R}^n. \tag{5.29}$$

2. *Set $g^* = g_m$.*

5.3. Minimax Stochastic Control - Average Cost

In this section, we study the infinite horizon Markov Control Model with the average pay-off per unit time as an optimality criterion. We derive the new dynamic programming equations under total variation distance ambiguity with and without imposing the irreducibility condition. In addition, we introduce the corresponding policy iteration algorithms for average cost dynamic programming. The derivations of our results are based on the classical results found in [32, 54].

Recall the definition of the infinite horizon D-MCM specified by (5.1), with discounting factor $\alpha = 1$, and consider the problem of minimizing the average pay-off per unit time

$$J(\pi) = \limsup_{j \to \infty} \Big\{ \sup_{Q(\cdot|x,u) \in \mathbf{B}_R(Q^o)(x,u)} \frac{1}{j} \mathbb{E}_{\mathbf{Q}^\pi_\nu} \Big\{ \sum_{k=0}^{j-1} f(x_k, u_k) \Big\} \Big\}. \tag{5.30}$$

For the finite-horizon optimal stochastic control problem with pay-off

$$\sup_{Q(\cdot|x,u) \in \mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}^\pi_\nu} \Big\{ \sum_{k=0}^{n-1} f(x_k, u_k) \Big\} \tag{5.31}$$

the value function satisfies the dynamic programming equation (see Chapter 4, Theorem 4.2, with discounted factor $\alpha = 1$, finite state-space $\mathcal{X}$ and finite input set $\mathcal{U}$)

$$V_j(x) = \inf_{u \in \mathcal{U}(x)} \sup_{Q(\cdot|x,u) \in \mathbf{B}_R(Q^o)(x,u)} \left\{ f(x,u) + \int_{\mathcal{X}} V_{j+1}(z) Q(dz|x,u) \right\} \tag{5.32}$$

which is equivalent to (since Q^* exists and is given by (3.46))

$$\begin{aligned} V_j(x) &= \inf_{u \in \mathcal{U}(x)} \left\{ f(x,u) + \int_{\mathcal{X}} V_{j+1}(z) Q^*(dz|x,u) \right\} \\ &= \inf_{u \in \mathcal{U}(x)} \left\{ f(x,u) + \int_{\mathcal{X}} V_{j+1}(z) Q^o(dz|x,u) + \frac{R}{2} \Big(\sup_{z \in \mathcal{X}} V_{j+1}(z) - \inf_{z \in \mathcal{X}} V_{j+1}(z) \Big) \right\}. \end{aligned} \tag{5.33}$$

Define $\overline{V}_j(x) = V_{n-j}(x)$. Then $\overline{V}$ satisfies the equation

$$\overline{V}_j(x) = \inf_{u \in \mathcal{U}(x)} \sup_{Q(\cdot|x,u) \in \mathbf{B}_R(Q^o)(x,u)} \left\{ f(x,u) + \int_{\mathcal{X}} \overline{V}_{j-1}(z) Q(dz|x,u) \right\}. \tag{5.34}$$

Rewrite this as

$$\begin{aligned} \overline{V}_j(x) + \frac{1}{j}\overline{V}_j(x) = \inf_{u \in \mathcal{U}(x)} \sup_{Q(\cdot|x,u) \in \mathbf{B}_R(Q^o)(x,u)} \Big\{ f(x,u) \\ + \int_{\mathcal{X}} Q(dz|x,u) \Big(\overline{V}_{j-1}(z) + \frac{1}{j}\overline{V}_j(x) \Big) \Big\}. \end{aligned} \tag{5.35}$$

Assume that there exists a V and a $J^* \in \mathbb{R}$ such that

$$\lim_{j \to \infty} \Big(\overline{V}_j(x) - jJ^* \Big) = V(x), \quad \forall x \in \mathcal{X}. \tag{5.36}$$

Then

$$\lim_{j \to \infty} \frac{1}{j} \overline{V}_j(x) = J^*, \quad \forall x \in \mathcal{X} \tag{5.37}$$

which limit does not depend on $x \in \mathcal{X}$. In addition, taking supremum with respect to $x \in \mathcal{X}$ on both sides of (5.36), and by the finite cardinality of $\mathcal{X}$

$$\lim_{j \to \infty} \sup_{x \in \mathcal{X}} \Big(\overline{V}_j(x) - jJ^* \Big) = \sup_{x \in \mathcal{X}} V(x). \tag{5.38}$$

Then, by (5.36) and (5.37)

$$\begin{aligned} J^* + V(x) &= \lim_{j \to \infty} \Big(\frac{1}{j} \overline{V}_j(x) + (\overline{V}_j(x) - jJ^*) \Big) \\ &\overset{(a)}{=} \lim_{j \to \infty} \inf_{u \in \mathcal{U}(x)} \sup_{Q(\cdot|x,u) \in \mathbf{B}_R(Q^o)(x,u)} \Big\{ f(x,u) + \int_{\mathcal{X}} Q(dz|x,u) \Big(\overline{V}_{j-1}(z) + \frac{1}{j} \overline{V}_j(x) \Big) - jJ^* \Big\} \\ &\overset{(b)}{=} \lim_{j \to \infty} \inf_{u \in \mathcal{U}(x)} \Big\{ f(x,u) + \int_{\mathcal{X}} Q^o(dz|x,u) \Big(\overline{V}_{j-1}(z) + \frac{1}{j} \overline{V}_j(x) \Big) \\ &\qquad + \frac{R}{2} \Big(\sup_{z \in \mathcal{X}} \Big(\overline{V}_{j-1}(z) + \frac{1}{j} \overline{V}_j(x) \Big) - \inf_{z \in \mathcal{X}} \Big(\overline{V}_{j-1}(z) + \frac{1}{j} \overline{V}_j(x) \Big) \Big) - jJ^* \Big\} \\ &\overset{(c)}{=} \lim_{j \to \infty} \inf_{u \in \mathcal{U}(x)} \Big\{ f(x,u) + \int_{\mathcal{X}} Q^o(dz|x,u) \Big(\overline{V}_{j-1}(z) - (j-1)J^* + \frac{1}{j} \overline{V}_j(x) - J^* \Big) \\ &\qquad + \frac{R}{2} \Big(\sup_{z \in \mathcal{X}} \Big(\overline{V}_{j-1}(z) - jJ^* \Big) - \inf_{z \in \mathcal{X}} \Big(\overline{V}_{j-1}(z) - jJ^* \Big) \Big) \Big\} \end{aligned}$$

where (a) is obtained by (5.35), (b) is obtained by equivalent formulation (5.33), and (c) by adding and subtracting $J^*(1+j\frac{R}{2})$. By assuming that $\mathcal{U}$ and $\mathcal{X}$ are finite and by definition of V, interchange of the limit and the minimization and maximization operations is allowed, and hence, the dynamic programming equation for the average pay-off is given by

$$J^* + V(x) = \min_{u\in\mathcal{U}(x)} \Big\{ f(x,u) + \int_{\mathcal{X}} Q^o(dz|x,u)V(z) + \frac{R}{2}\Big(\sup_{z\in\mathcal{X}} V(z) - \inf_{z\in\mathcal{X}} V(z)\Big)\Big\} \tag{5.39}$$

which is equivalent to

$$J^* + V(x) = \min_{u\in\mathcal{U}(x)} \max_{Q(\cdot|x,u)\in \mathbf{B}_R(Q^o)(x,u)} \Big\{ f(x,u) + \int_{\mathcal{X}} Q(dz|x,u)V(z)\Big\}. \tag{5.40}$$

Next, we show that J^* satisfying (5.40) is the minimal average pay-off, and further that if π^* is a stationary policy such that $g^*(x)$ achieves the minimum on the right-hand side of (5.40) for every $x \in \mathcal{X}$, then π^* is optimal.

Theorem 5.4. *Suppose there exists a solution (V, J^*) of the dynamic programming equation (5.39). Let π^* be a stationary policy such that $g^*(x)$ attains the minimum in the right-hand side of (5.39) for every x. Then π^* is an optimal policy and J^* is the minimum average cost.*

Proof. Let $\pi \in \Pi^D$ be any policy, possibly non-stationary, and $u \in \mathcal{U}(x)$. Then

$$\begin{aligned} &\max_{Q(\cdot|x,u)\in \mathbf{B}_R(Q^o)(x,u)} \Big\{ f(x,u) + \int_{\mathcal{X}} Q(dz|x,u)V(z)\Big\} \\ &\geq \max_{Q(\cdot|x,u)\in \mathbf{B}_R(Q^o)(x,u)} \Big\{ f(x,g^*(x)) + \int_{\mathcal{X}} Q(dz|x,g^*(x))V(z)\Big\} = J^* + V(x) \end{aligned} \tag{5.41}$$

since (V, J^*) satisfy the dynamic programming equation and by definition of π^*. Hence, by (5.41)

$$\begin{aligned} \mathbb{E}^g_{Q^*}\Big(f(x_j,u_j)\Big) &\geq J^* + \mathbb{E}^g_{Q^*}\Big(V(x_j)\Big) - \mathbb{E}^g_{Q^*}\Big(\int_{\mathcal{X}} Q^*(dz|x_j,u_j)V(z)\Big) \\ &= J^* + \mathbb{E}^g_{Q^*}\Big(V(x_j)\Big) - \mathbb{E}^g_{Q^*}\Big(V(x_{j+1})\Big). \end{aligned} \tag{5.42}$$

Then

$$\begin{aligned} J(\pi) &\geq \liminf_{j\to\infty} \Big(\frac{1}{j}\sum_{k=0}^{j-1} \mathbb{E}^g_{Q^*}\Big(f(x_k,u_k)\Big)\Big) \\ &\overset{(a)}{\geq} \liminf_{j\to\infty} \Big(J^* + \frac{1}{j}\Big(\mathbb{E}^g_{Q^*}\Big(V(x_0)\Big) - \mathbb{E}^g_{Q^*}\Big(V(x_j)\Big)\Big)\Big) \\ &\overset{(b)}{=} J^* \end{aligned}$$

where (a) is obtained by (5.42), and (b) because the last term vanishes as $j \to \infty$. Thus, $J^* \leq \inf_{\pi\in\Pi^D} J(\pi)$. However, when π^* replaces π equality holds throughout and as a result π^* is optimal, that is, $J^* = J(\pi^*) = \inf_{\pi\in\Pi^D} J(\pi)$, $\pi^* \in \Pi^D$ is an optimal control law and J^* is the value. ∎

5.3.1. Existence

Dynamic programming equation (5.39) is valid under the assumption that (5.36) and (5.38) are satisfied. Here, we derive more general conditions under which (5.39) is valid. First, we introduce some notation similar to Section 5.2.1.

Let the state space $\mathcal{X}$ be a finite set, with say, n elements. Then, any function $V : \mathcal{X} \longrightarrow \mathbb{R}^n$ may be represented by a vector in $\mathbb{R}^n$. Any stationary control law $\pi \in \Pi^{DS}$ defines $g : \mathcal{X} \longmapsto \mathbb{R}$ which may also be identified with a $g \in \mathbb{R}^n$. For any g let $Q(j) \in \mathbb{R}^{n\times n}$ and

$$f(g) = \begin{pmatrix} f(x_1, g(x_1)) & \cdots & f(x_n, g(x_n)) \end{pmatrix}^T \in \mathbb{R}^n.$$

Let $q_0 \in \mathbb{R}^n$ be defined by

$$q_0(i) \triangleq Q(\{x_0 = i\})$$

and

$$e \triangleq (1, \cdots, 1)^T \in \mathbb{R}^n.$$

Maximizing the expected cost on a finite horizon is then

$$\begin{aligned} \max_{Q(\cdot|x,u)\in \mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}^\pi_\nu}\left\{\sum_{k=0}^{j-1} f(x_k,u_k)\right\} &= \max_{Q(\cdot|x,u)\in \mathbf{B}_R(Q^o)(x,u)} \left\{\sum_{k=0}^{j-1} q_0^T Q(g)^k f(g)\right\} \\ &= \max_{Q(\cdot|x,u)\in \mathbf{B}_R(Q^o)(x,u)} q_0^T\left\{\sum_{k=0}^{j-1} Q(g)^k\right\} f(g). \end{aligned} \tag{5.43}$$

Let $Q^*(\cdot|x,u)$ denote the maximizing conditional distribution of (5.43). Then (5.43) is equivalent

$$\max_{Q(\cdot|x,u)\in \mathbf{B}_R(Q^o)(x,u)} q_0^T\left\{\sum_{k=0}^{j-1} Q(g)^k\right\} f(g) = q_0^T\left\{\sum_{k=0}^{j-1} Q^*(g)^k\right\} f(g). \tag{5.44}$$

Maximizing the average cost per unit time is then

$$\begin{aligned} J(\pi) &= \limsup_{j\to\infty} \frac{1}{j} \max_{Q(\cdot|x,u)\in \mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}^\pi_\nu}\left\{\sum_{k=0}^{j-1} f(x_k,u_k)\right\} \\ &= \limsup_{j\to\infty} \frac{1}{j} q_0^T\left\{\sum_{k=0}^{j-1} Q^*(g)^k\right\} f(g). \end{aligned}$$

Since $q_0 \in \mathbb{R}^n$ and $f(g) \in \mathbb{R}^n$ are independent of j, we only need to investigate the conditions under which the limit of

$$\lim_{j\to\infty} \frac{1}{j} \sum_{k=0}^{j-1} Q^*(g)^k$$

exists. The following results are stated without proof.

Lemma 5.2. *[54] If $Q^* \in \mathbb{R}_+^{n\times n}$ is a stochastic matrix, then the Cesaro limit*

$$\lim_{j\to\infty} \frac{1}{j}\sum_{k=0}^{j-1}(Q^*)^k = Q_1^* \tag{5.45}$$

always exist. The matrix $Q_1^* \in \mathbb{R}_+^{n\times n}$ is a stochastic matrix and is the solution of the equation

$$Q_1^* Q^* = Q_1^*. \tag{5.46}$$

Thus, the maximization of the average cost of a stationary Markov control law is given by

$$J(\pi) = q_0^T Q_1^*(g) f(g) \tag{5.47}$$

where $Q_1^*(g)$ and $Q^*(g)$ are related by (5.46).

Definition 5.1. *[54] A stochastic matrix $P \in \mathbb{R}_+^{n\times n}$ is said to be reducible if by row and column permutations it can be placed into block upper-triangular form*

$$P = \begin{pmatrix} P_1 & P_2 \\ 0 & P_3 \end{pmatrix}, \quad \text{where } P_1,\ P_2 \text{ are square matrices.}$$

A stochastic matrix that is not reducible is said to be irreducible.

Lemma 5.3. *[54] Let $Q^* \in \mathbb{R}_+^{n\times n}$ be an irreducible stochastic matrix. Then, there exists a unique vector q such that*

$$Q^* q = q, \quad e^T q = 1, \quad q(i) > 0 \text{ for all } i \in \mathbb{Z}_n.$$

Moreover, the matrix Q_1^* associated with Q^* in (5.46) has all rows equal to q.

Note that, (5.47) depends on the probability distribution q_0 of the initial state. However, if Q_1^* is assumed to be an irreducible stochastic matrix, then by Lemma 5.3

$$J(\pi) = q_0^T Q_1^*(g) f(g) = q(g)^T f(g) \tag{5.48}$$

where $q(g)$ is the unique invariant probability distribution, that is, $Q^* q(g) = q(g)$. From (5.48), the average cost $J(g)$ is independent of the initial distribution. Hence, for the remainder of this section, we will assume that for every stationary Markov control law the stochastic matrix Q^* is irreducible. The next proposition summarizes the above results.

Proposition 5.1. *[54] Let π be a stationary Markov control law which defines $g : \mathcal{X} \longmapsto \mathcal{U}$. Assume that $Q^*(g) \in \mathbb{R}_+^{n\times n}$ is irreducible.*

(a) There exists a unique $q(g) \in \mathbb{R}_+^n$ such that

$$Q^*(g) q(g) = q(g), \quad e^T g = 1, \quad e = (1, \ldots, 1)^T. \tag{5.49}$$

(b) *The average cost associated with the control law* $\pi \in \Pi^{DS}$ *is*

$$J(\pi) = q(g)^T f(g). \tag{5.50}$$

(c) *There exists a* $V(g) \in \mathbb{R}^n$ *such that*

$$J(\pi)e + V(g) = f(g) + Q^*(g)V(g). \tag{5.51}$$

Proof. Part (a) and (b) follows from Lemma (5.3) and the discussion above. For part (c) see [54]. ∎

Lemma 5.4. *Assume that:*

1. *For any stationary control law* $\pi \in \Pi^{DS}$, $Q^*(g) \in \mathbb{R}_+^{n\times n}$ *is irreducible.*

2. *There exists a stationary Markov control law* $\pi \in \Pi^{DS}$ *such that*

$$J(\pi^*) = \inf_{\pi\in\Pi^{DS}} J(\pi)$$

Then there exists $(V(g^*,\cdot), J(\pi^*))$, $V(g^*,\cdot) : \mathcal{X} \longmapsto \mathbb{R}$ *and* $J(\pi) \in \mathbb{R}$ *that is a solution to the dynamic programming equation*

$$J(\pi^*) + V(g^*, x) = \min_{u\in\mathcal{U}} \Big\{ f(x,u) + \sum_{z\in\mathcal{X}} Q^*(z|x,u)V(g^*,z) \Big\}. \tag{5.52}$$

Proof. See Appendix D.1. ∎

Theorem 5.5. *Assume that for all stationary Markov control laws* $\pi \in \Pi^{DS}$, *and for a given total variation parameter* R, *the maximizing transition matrix* $Q^*(g)$ *is irreducible.*

(a) *There exists a solution* $V : \mathcal{X} \longmapsto \mathbb{R}$ *and* $J^* \in \mathbb{R}$ *to the dynamic programming equation*

$$J^* + V(x) = \min_{u\in\mathcal{U}} \Big\{ f(x,u) + \sum_{z\in\mathcal{X}} Q^*(z|x,u)V(z) \Big\}. \tag{5.53}$$

The maximizing conditional distribution is

$$Q^*(\mathcal{X}^+|x,u) = Q^o(\mathcal{X}^+|x,u) + \frac{R}{2} \in [0,1] \tag{5.54}$$

$$Q^*(\mathcal{X}^-|x,u) = Q^o(\mathcal{X}^-|x,u) + \frac{R}{2} \in [0,1] \tag{5.55}$$

$$Q^*(A|x,u) = Q^o(A|x,u), \quad \forall A \subseteq \mathcal{X} \setminus \mathcal{X}^+ \cup \mathcal{X}^- \tag{5.56}$$

where

$$\mathcal{X}^+ \triangleq \Big\{ x \in \mathcal{X} : V(x) = \sup\{V(x) : x \in \mathcal{X}\} \Big\} \tag{5.57}$$

$$\mathcal{X}^- \triangleq \Big\{ x \in \mathcal{X} : V(x) = \inf\{V(x) : x \in \mathcal{X}\} \Big\}. \tag{5.58}$$

Moreover,

$$J^* + V(x) = \min_{u \in \mathcal{U}} \Big\{ f(x,u) + \sum_{z \in \mathcal{X}} Q^o(z|x,u)V(z) + \frac{R}{2}\Big(\sup_{z \in \mathcal{X}} V(z) - \inf_{z \in \mathcal{X}} V(z) \Big) \Big\}. \tag{5.59}$$

(b) *If $g^*(x)$ attains the minimum in (5.53) for every x, then g^* is an optimal policy.*

(c) *The minimum cost is J^*.*

Proof. Theorem 5.5 is obtained by combining Lemmas 5.4 and 5.4 and by applying the results of Section 4.3.2. ■

5.3.2. Policy Iteration Algorithm

In this section, we provide a modified version of the classical policy iteration algorithm for average cost dynamic programming [39,54]. From part (a) of Theorem 5.5, policy evaluation and policy improvement steps of a policy iteration algorithm must be performed using the maximizing conditional distribution obtained under total variation distance ambiguity constraint. Moreover, one needs to guarantee that for the given total variation parameter R the resulting matrix Q^* is irreducible, otherwise, Algorithm 5.6 may not be sufficient to give the optimal policy and the minimum cost.

Algorithm 5.6 (Policy iteration for average-cost dynamic programming)**.**

1. *Let $m = 0$ and select an arbitrary stationary Markov control law $g_0 : \mathcal{X} \longmapsto \mathcal{U}$.*

2. *(Policy Evaluation) Solve the equation*

$$J_{Q^o}(g_m)e + V_{Q^o}(g_m) = f(g_m) + Q^o(g_m)V_{Q^o}(g_m), \quad e = (1, \ldots, 1)^T \tag{5.60}$$

for $J_{Q^o}(g_m) \in \mathbb{R}$ and $V_{Q^o}(g_m) \in \mathbb{R}^n$. Identify the support sets of (5.60) *using* (5.57)-(5.58)*, and construct the matrix $Q^*(g_m)$ using* (5.54)-(5.56)*. Solve the equation*

$$J_{Q^*}(g_m)e + V_{Q^*}(g_m) = f(g_m) + Q^*(g_m)V_{Q^*}(g_m), \quad e = (1, \ldots, 1)^T \tag{5.61}$$

for $J_{Q^*}(g_m) \in \mathbb{R}$ and $V_{Q^*}(g_m) \in \mathbb{R}^n$.

3. *(Policy Improvement) Let*

$$g_{m+1} = \operatorname*{argmin}_{g \in \mathbb{R}^n} \Big\{ f(g) + Q^*(g)V_{Q^*}(g_m) \Big\}. \tag{5.62}$$

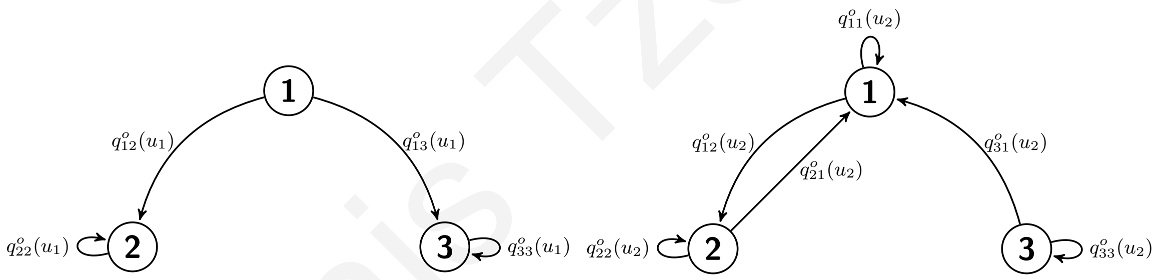
4. *If $g_{m+1} = g_m$, let $g^* = g_m$; else let $m = m+1$ and return to step 2.*

In Section 5.4.2, we illustrate how policy iteration algorithm for infinite horizon average cost dynamic programming is implemented through an inventory control example.

5.3.3. Limitations

Part (a) of Theorem 5.5, showed that for a stationary Markov control policy $\pi \in \Pi^{DS}$ and for an irreducible stochastic matrix Q^* there exists a solution to the dynamic programming equation (5.53). Moreover, the maximizing stochastic matrix Q^* which is given by (5.54)-(5.56), is calculated based on the support sets (5.57)-(5.58), the nominal stochastic matrix Q^o, and the value of the total variation parameter R. Hence, in order to apply policy iteration algorithm for average-cost dynamic programming one needs to know in advance that, for a given total variation parameter $R \in [0, 2]$ and an irreducible nominal stochastic matrix Q^o, the maximizing stochastic matrix Q^* is also irreducible. Otherwise, policy iteration algorithm may not be sufficient to give the optimal policy and the minimum cost. In particular, as we show in the next example, if irreducibility condition is not satisfied then policy iteration algorithm need not have a solution.

Example 5.1. *Consider the stochastic control system shown in Fig.5.1, with state-space* $\mathcal{X} = \{1, 2, 3\}$ *and control set* $\mathcal{U} = \{u_1, u_2\}$*. Let the nominal transition probability under*



(a) (b)

Figure 5.1.: Transition Probability Graph of Q^o under controls u_1 and u_2. Plot (a) depicts matrix Q^o under control u_1. Plot (b) depicts matrix Q^o under control u_2.

controls u_1 *and* u_2 *to be given by*

$$Q^o(u_1) = \frac{1}{9}\begin{pmatrix} 0 & 5 & 4 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix}, \quad Q^o(u_2) = \frac{1}{9}\begin{pmatrix} 2 & 7 & 0 \\ 3 & 6 & 0 \\ 8 & 0 & 1 \end{pmatrix}$$

and the cost function under each state and action is

$$f(1, u_1) = 2,\ f(2, u_1) = 1,\ f(3, u_1) = 3,\ f(1, u_2) = 0.5,\ f(2, u_2) = 3,\ f(3, u_2) = 0.$$

Clearly, for this control system the nominal transition probability matrix, under both controls, is reducible, since the system under controls u_1 *and* u_2 *contains more than one recurrent class. Using policy iteration Algorithm 5.6 with initial policies* $g_0(1) = g_0(2) =$

$g_0(3) = u_1$, *the optimality equation* (5.60) *for this system may be written as*

$$J_{Q^o}\begin{pmatrix}1\\1\\1\end{pmatrix} + V_{Q^o}(g_0) = \begin{pmatrix}2\\1\\3\end{pmatrix} + \frac{1}{9}\begin{pmatrix}0 & 5 & 4\\0 & 9 & 0\\0 & 0 & 9\end{pmatrix} V_{Q^o}(g_0)$$

and hence

$$\begin{aligned}
J_{Q^o} + V_{Q^o}(g_0,1) &= 2 + \frac{5}{9}V_{Q^o}(g_0,2) + \frac{4}{9}V_{Q^o}(g_0,3)\\
J_{Q^o} + V_{Q^o}(g_0,2) &= 1 + V_{Q^o}(g_0,2)\\
J_{Q^o} + V_{Q^o}(g_0,3) &= 3 + V_{Q^o}(g_0,3).
\end{aligned}$$

The second and third equations show that the system is inconsistent, and hence, the policy iteration algorithm fails to give the optimal policy and the minimum cost.

Moreover, even if Q^o is an irreducible stochastic matrix, it turns out that, as the value of total variation parameter increases, the maximizing stochastic matrix Q^*, eventually, will be transformed into a reducible stochastic matrix. Hence, our proposed method for solving minimax stochastic control problem with average cost is valid only for a specific range of values of total variation parameter. In particular, if Q^o is an irreducible stochastic matrix then, for any given partition of the state-space, there exists an $R_{\min} \in [0, 2)$ for which we distinguish the following two cases:

(a) for $0 \leq R < R_{\min}$, Q^* is an irreducible stochastic matrix. Theorem 5.5 is valid and policy iteration algorithm gives the optimal policy and the minimum cost.

(b) for $R \geq R_{\min}$, Q^* is a reducible stochastic matrix. Theorem 5.5 is not valid and policy iteration algorithm need not have a solution.

Remark 5.1. *An extended solution through a reduced dimensional state-space may be obtained as follows. Consider the case for which* $R \geq R_{\min}$. *Due to the water-filling behavior of maximizing conditional distribution* (5.54)-(5.56)*, columns of* Q^* *which correspond to states belonging to* $\mathcal{X} \setminus \mathcal{X}^0$*, become columns with all zero's as total variation parameter* R *increases. Whenever an all zero column appears, one can augment the corresponding state of that column, and hence* Q^* *will be transformed back into an irreducible stochastic matrix of reduced order.*

5.3.4. General Dynamic Programming

In this section, we propose an alternative method, the so-called, general dynamic programming for average cost which overcomes the limitations discussed in section 5.3.3. Despite the fact that it is more complex, it completely solves the minimax stochastic control problem with average cost. In particular, we introduce a general dynamic programming for average cost, without imposing the assumption that for all stationary Markov control laws, and for a given total variation parameter R, the maximizing stochastic matrix $Q^* \in \mathbb{R}_+^{n\times n}$ is irreducible.

As we have already discussed, when considering control systems with more than one recurrent class, policy iteration algorithm may not be sufficient to give the optimal policy and the minimum cost. In addition, due to the water-filling behavior of the maximizing conditional distribution, eventually, Q^* will be transformed into a reducible stochastic matrix. Hence, the proposed methodology of previous sections and policy iteration Algorithm 5.6, solve the minimax stochastic control with average cost only for a specific range of values of total variation parameter. The general dynamic programming equations which completely solves minimax stochastic control problem with average cost are introduced next.

General Dynamic Programming Equations

Example 5.1 in Section 5.3.3, showed that a unique dynamic programming equation may not be sufficient to give the optimal policy and the minimum cost when there is more than one recurrent class. A general solution is obtained by introducing an additional dynamic programming equation. We refer to the pair of dynamic programming equations as general dynamic programming equations, since they completely solve the minimax stochastic control problem with average cost, and they are given by

$$
\begin{aligned}
J^*(x) &= \inf_{u\in\mathcal{U}(x)} \sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \int_{\mathcal{X}} Q(dz|x,u)J^*(z) && \text{(5.63a)}\\
&= \inf_{u\in\mathcal{U}(x)} \int_{\mathcal{X}} Q^*(dz|x,u)J^*(z) && \text{(5.63b)}\\
&= \inf_{u\in\mathcal{U}(x)} \Big\{ \int_{\mathcal{X}} Q^o(dz|x,u)J^*(z) + \frac{R}{2}\Big(\sup_{z\in\mathcal{X}} J^*(z) - \inf_{z\in\mathcal{X}} J^*(z)\Big)\Big\} && \text{(5.63c)}
\end{aligned}
$$

and

$$
\begin{aligned}
J^*(x)+V(x) &= \inf_{u\in\mathcal{U}(x)} \sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \Big\{ f(x,u)+\int_{\mathcal{X}} Q(dz|x,u)V(z)\Big\} && \text{(5.64a)}\\
&= \inf_{u\in\mathcal{U}(x)} \Big\{ f(x,u) + \int_{\mathcal{X}} Q^*(dz|x,u)V(z)\Big\} && \text{(5.64b)}\\
&= \inf_{u\in\mathcal{U}(x)} \Big\{ f(x,u)+\int_{\mathcal{X}} Q^o(dz|x,u)V(z)+\frac{R}{2}\Big(\sup_{z\in\mathcal{X}} V(z)- \inf_{z\in\mathcal{X}} V(z)\Big)\Big\}. && \text{(5.64c)}
\end{aligned}
$$

We refer to (5.63) as the first general dynamic programming equation and to (5.64) as the second general dynamic programming equation. The maximizing conditional distribution of

(5.63b) is given by (5.54)-(5.56), where

$$\mathcal{X}^+ \triangleq \Big\{x \in \mathcal{X} : J^*(x) = \sup\{J^*(x) : x \in \mathcal{X}\}\Big\} \tag{5.65}$$

$$\mathcal{X}^- \triangleq \Big\{x \in \mathcal{X} : J^*(x) = \inf\{J^*(x) : x \in \mathcal{X}\}\Big\}. \tag{5.66}$$

Similarly, the maximizing conditional distribution of (5.64b) is given by (5.54)-(5.56), where

$$\mathcal{X}^+ \triangleq \Big\{x \in \mathcal{X} : V(x) = \sup\{V(x) : x \in \mathcal{X}\}\Big\} \tag{5.67}$$

$$\mathcal{X}^- \triangleq \Big\{x \in \mathcal{X} : V(x) = \inf\{V(x) : x \in \mathcal{X}\}\Big\}. \tag{5.68}$$

Definition 5.2. *[32] Let ρ and h be real-valued measurable functions on $\mathcal{X}$ and φ^* a given stationary selector. Then (ρ, h, φ^*) is said to be a canonical triplet if, for every $x \in \mathcal{X}$ and $j = 0, 1, \ldots$*

$$\begin{aligned}&\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}}\Big\{\sum_{k=0}^{j-1} f(x_k,u_k) + h(x_j)\Big\}\\ &= \inf_{\pi\in\Pi^{DS}} \sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\pi}}\Big\{\sum_{k=0}^{j-1} f(x_k,u_k) + h(x_j)\Big\} = j\rho(x) + h(x).\end{aligned} \tag{5.69}$$

If a stationary selector φ^* is an element of a canonical triplet, then it is called canonical.

Theorem 5.7. *[32] (ρ, h, φ^*) is a canonical triplet if and only if, for every $x \in \mathcal{X}$,*

$$\begin{aligned}\rho(x) &= \inf_{u\in\mathcal{U}(x)} \sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \int_{\mathcal{X}} \rho(z)Q(dz|x,u)\\ &= \sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \int_{\mathcal{X}} \rho(z)Q(dz|x,\varphi^*(x))\end{aligned} \tag{5.70}$$

and

$$\begin{aligned}\rho(x) + h(x) &= \inf_{u\in\mathcal{U}(x)} \sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \Big\{f(x,u) + \int_{\mathcal{X}} h(z)Q(dz|x,u)\Big\}\\ &= \sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \Big\{f(x,\varphi^*(x)) + \int_{\mathcal{X}} h(z)Q(dz|x,\varphi^*(x))\Big\}.\end{aligned} \tag{5.71}$$

Note that, if (ρ, h, φ^*) solve (5.70) and (5.71), then so does $(\rho, h + N, \varphi^*)$ for any constant N.

Proof. $(\Longrightarrow)$ Suppose that (ρ, h, φ^*) is a canonical triplet, i.e., (5.69) holds $\forall x \in \mathcal{X}$ and $j \geq 0$. From dynamic programming equation (5.32) we have that

$$V_j(x) = \min_{u\in\mathcal{U}(x)} \max_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \Big\{f(x,u) + \int_{\mathcal{X}} V_{j+1}(z)Q(dz|x,u)\Big\}. \tag{5.72}$$

Define $\overline{V}_j(x) = V_{n-j}(x)$, $(j = 0, \ldots, n)$. Then (5.72) may be written in the "forward" form

$$\overline{V}_{j+1}(x) = \min_{u \in \mathcal{U}(x)} \max_{Q(\cdot|x,u) \in \mathbf{B}_R(Q^o)(x,u)} \Big\{ f(x,u) + \int_{\mathcal{X}} \overline{V}_j(z) Q(dz|x,u) \Big\}. \tag{5.73}$$

Thus, from (5.69)

$$(j+1)\rho(x) + h(x) = \min_{u \in \mathcal{U}(x)} \max_{Q(\cdot|x,u) \in \mathbf{B}_R(Q^o)(x,u)} \Big\{ f(x,u) + \int_{\mathcal{X}} \Big(j\rho(z) + h(z) \Big) Q(dz|x,u) \Big\}$$

which yields the first equality in (5.70) when $j = 0$ and multiplying by $1/j$ and letting $j \rightarrow \infty$ it also gives the first equality in (5.71). Finally, for any deterministic stationary policy $\varphi^* \in \Pi^{DS}$ which satisfies (5.69), we have that

$$(j+1)\rho(x) + h(x) = \max_{Q(\cdot|x,u) \in \mathbf{B}_R(Q^o)(x,u)} \Big\{ f(x,\varphi^*) + \int_{\mathcal{X}} \Big(j\rho(z) + h(z) \Big) Q(dz|x,\varphi^*) \Big\}$$

which, as before, gives the second equality in (5.70) and (5.71).

($\Longleftarrow$) Conversely, suppose now that (ρ, h, φ^*) satisfy (5.70) and (5.71). Then iteration of (5.71) using (5.70) implies, $\forall x \in \mathcal{X}$ and $j \geq 0$

$$\begin{aligned} & j\rho(x) + h(x) \\ &= \sup_{Q(\cdot|x,u) \in \mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}} \Big\{ \sum_{k=0}^{j-1} f(x_k, u_k) \Big\} + \sup_{Q(\cdot|x,u) \in \mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}} \Big\{ h(x_j) \Big\} \\ &= \sup_{Q(\cdot|x,u) \in \mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}} \Big\{ \sum_{k=0}^{j-1} f(x_k, u_k) + h(x_j) \Big\}. \end{aligned} \tag{5.74}$$

Thus to complete the proof of (5.69) it only remains to show that

$$\begin{aligned} & \sup_{Q(\cdot|x,u) \in \mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}} \Big\{ \sum_{k=0}^{j-1} f(x_k, u_k) + h(x_j) \Big\} \\ & \qquad = \inf_{\pi \in \Pi^{DS}} \sup_{Q(\cdot|x,u) \in \mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\pi}} \Big\{ \sum_{k=0}^{j-1} f(x_k, u_k) + h(x_j) \Big\}. \end{aligned} \tag{5.75}$$

Proceeding by induction, we note first that (5.75) is obvious when $n = 0$. Suppose now that (5.75) holds for some $j \geq 0$. Then, from (5.72) and (5.74), together with the induction hypothesis

$$\begin{aligned}
\overline{V}_j(x) &= \min_{u\in\mathcal{U}(x)} \max_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \Big\{ f(x,u) + \int_{\mathcal{X}} \overline{V}_j(z)Q(dz|x,u) \Big\} \\
&= \min_{u\in\mathcal{U}(x)} \max_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \Big\{ f(x,u) + \int_{\mathcal{X}} \big(j\rho(z)+h(z)\big)Q(dz|x,u) \Big\} \\
&\geq \min_{u\in\mathcal{U}(x)} \max_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \Big\{ f(x,u) + \int_{\mathcal{X}} h(z)Q(dz|x,u) \Big\} \\
&\qquad\qquad + j \min_{u\in\mathcal{U}(x)} \max_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \Big\{ \int_{\mathcal{X}} \rho(z)Q(dz|x,u) \Big\} \\
&= (j+1)\rho(x) + h(x) \\
&= \sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}} \Big\{ \sum_{k=0}^{j-1} f(x_k,u_k) + h(x_j) \Big\}.
\end{aligned}$$

Hence, since the reverse inequality trivially holds, then (5.75) holds for $n+1$, and the proof is complete. ■

Theorem 5.8. *Suppose the cost function f is bounded below, and (ρ, h, φ^*) be a canonical triplet.*

(a) If for any $\pi \in \Pi^D$ and any $x \in \mathcal{X}$

$$\lim_{j\to\infty} \Big[\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\pi}} \Big\{ \frac{h(x_j)}{j} \Big\} \Big] = 0 \tag{5.76}$$

then φ^* is an optimal strategy and ρ is the average cost value function

$$\begin{aligned}
V(x) = \rho(x) &= \limsup_{j\to\infty} \frac{1}{j} \Big[\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}} \Big\{ \sum_{k=0}^{j-1} f(x_k,u_k) \Big\} \Big] \\
&= \lim_{j\to\infty} \frac{1}{j} \Big[\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}} \Big\{ \sum_{k=0}^{j-1} f(x_k,u_k) \Big\} \Big].
\end{aligned} \tag{5.77}$$

(b) If for any $x \in \mathcal{X}$

$$\lim_{j\to\infty} \sup_{\pi\in\Pi^D} \Big[\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\pi}} \Big\{ \frac{h(x_j)}{j} \Big\} \Big] = 0 \tag{5.78}$$

then for all π and $x \in \mathcal{X}$

$$\begin{aligned}
&\limsup_{j\to\infty} \frac{1}{j} \Big[\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}} \Big\{ \sum_{k=0}^{j-1} f(x_k,u_k) \Big\} \Big] \\
&\qquad \leq \liminf_{j\to\infty} \frac{1}{j} \Big[\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\pi}} \Big\{ \sum_{k=0}^{j-1} f(x_k,u_k) \Big\} \Big]
\end{aligned} \tag{5.79}$$

and

$$\lim_{j\to\infty}\Big[\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)}\mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}}\Big\{\sum_{k=0}^{j-1}f(x_k,u_k)\Big\} - \inf_{\pi\in\Pi^D}\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)}\mathbb{E}_{\mathbf{Q}_\nu^{\pi}}\Big\{\sum_{k=0}^{j-1}f(x_k,u_k)\Big\}\Big]/j=0. \quad (5.80)$$

Proof. (a) From (5.69)

$$\begin{aligned} j\rho(x)+h(x) &= \inf_{\pi\in\Pi^D}\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)}\mathbb{E}_{\mathbf{Q}_\nu^{\pi}}\Big\{\sum_{k=0}^{j-1}f(x_k,u_k)+h(x_j)\Big\}\\ &\le \sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)}\mathbb{E}_{\mathbf{Q}_\nu^{\pi}}\Big\{\sum_{k=0}^{j-1}f(x_k,u_k)\Big\}+\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)}\mathbb{E}_{\mathbf{Q}_\nu^{\pi}}\Big\{h(x_j)\Big\}\end{aligned}$$

$\forall\pi\in\Pi^D$ and $x\in\mathcal{X}$. Hence, multiplying by $1/j$ and taking the $\limsup$ as $j\to\infty$

$$\rho(x)\le\limsup_{j\to\infty}\frac{1}{j}\Big[\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)}\mathbb{E}_{\mathbf{Q}_\nu^{\pi}}\Big\{\sum_{k=0}^{j-1}f(x_k,u_k)\Big\}\Big],\qquad\forall\pi,x$$

which implies

$$\rho(x)\le\inf_{\pi\in\Pi^D}\limsup_{j\to\infty}\frac{1}{j}\Big[\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)}\mathbb{E}_{\mathbf{Q}_\nu^{\pi}}\Big\{\sum_{k=0}^{j-1}f(x_k,u_k)\Big\}\Big],\qquad\forall x. \quad (5.81)$$

Furthermore, from (5.69) again

$$\begin{aligned}&\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)}\mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}}\Big\{\sum_{k=0}^{j-1}f(x_k,u_k)+h(x_j)\Big\}\\ &=\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)}\mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}}\Big\{\sum_{k=0}^{j-1}f(x_k,u_k)\Big\}+\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)}\mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}}\Big\{h(x_j)\Big\}\\ &=j\rho(x)+h(x).\end{aligned}$$

Finally, multiplying by $1/n$ and then taking both $\limsup$ and $\liminf$ as $j\to\infty$ we obtain the last two equalities in (5.77), which in turn, together with (5.81), yield the first one since

$$\begin{aligned}\rho(x)&\le\inf_{\pi\in\Pi^D}\limsup_{j\to\infty}\frac{1}{j}\Big[\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)}\mathbb{E}_{\mathbf{Q}_\nu^{\pi}}\Big\{\sum_{k=0}^{j-1}f(x_k,u_k)\Big\}\Big]\\ &\le\limsup_{j\to\infty}\frac{1}{j}\Big[\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)}\mathbb{E}_{\mathbf{Q}_\nu^{\pi}}\Big\{\sum_{k=0}^{j-1}f(x_k,u_k)\Big\}\Big]=\rho(x).\end{aligned}$$

(b) From (5.69)

$$\begin{aligned}&\inf_{\pi\in\Pi^D}\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)}\mathbb{E}_{\mathbf{Q}_\nu^{\pi}}\Big\{\sum_{k=0}^{j-1}f(x_k,u_k)+h(x_j)\Big\} \quad (5.82)\\ &=\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)}\mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}}\Big\{\sum_{k=0}^{j-1}f(x_k,u_k)\Big\}+\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)}\mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}}\Big\{h(x_j)\Big\}\end{aligned}$$

and, on the other hand,

$$
\begin{aligned}
&\inf_{\pi\in\Pi^D} \sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^\pi}\Big\{\sum_{k=0}^{j-1} f(x_k,u_k)+h(x_j)\Big\}\\
&= \inf_{\pi\in\Pi^D}\Big[\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}}\Big\{\sum_{k=0}^{j-1} f(x_k,u_k)\Big\}+\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}}\Big\{h(x_j)\Big\}\Big]\\
&\le \inf_{\pi\in\Pi^D}\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}}\Big\{\sum_{k=0}^{j-1} f(x_k,u_k)\Big\}+\sup_{\pi\in\Pi^D}\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}}\Big\{h(x_j)\Big\}
\end{aligned}
$$

Hence, if h satisfies (5.78) then (5.80) is obtained. To prove (5.79), we use (5.82) to obtain

$$
\begin{aligned}
&\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}}\Big\{\sum_{k=0}^{j-1} f(x_k,u_k)\Big\}+\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}}\Big\{h(x_j)\Big\}\\
&\le \sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\pi}}\Big\{\sum_{k=0}^{j-1} f(x_k,u_k)\Big\}+\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}}\Big\{h(x_j)\Big\}, \qquad \forall \pi, x, j
\end{aligned}
$$

so that, from (5.78)

$$
\begin{aligned}
&\liminf_{j\to\infty}\frac{1}{j}\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}}\Big\{\sum_{k=0}^{j-1} f(x_k,u_k)\Big\}\\
&\qquad\qquad\le \liminf_{j\to\infty}\frac{1}{j}\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\pi}}\Big\{\sum_{k=0}^{j-1} f(x_k,u_k)\Big\}.
\end{aligned}
$$

From part (a) of Theorem 5.8, the left-hand side equals

$$
\limsup_{j\to\infty}\frac{1}{j}\sup_{Q(\cdot|x,u)\in\mathbf{B}_R(Q^o)(x,u)} \mathbb{E}_{\mathbf{Q}_\nu^{\varphi^*}}\Big\{\sum_{k=0}^{j-1} f(x_k,u_k)\Big\}
$$

and hence (5.79) is obtained and the proof is complete. ∎

5.3.5. General Policy Iteration Algorithm

In this section, we provide a policy iteration algorithm for average cost dynamic programming, which solves general dynamic programming equations (5.63) and (5.64). Because policy evaluation and policy improvement steps must be performed using the maximizing conditional distribution obtained under total variation distance ambiguity constraint for a pair of equations, the proposed algorithm is considerably more complex compared to Algorithm 5.6. However, as we mentioned earlier, it solves minimax stochastic control problem with average cost for all range of values of total variation parameter $R \in [0, 2]$, and without imposing the irreducibility condition.

Algorithm 5.9 (General policy iteration for average-cost dynamic programming).

1. *Let* $m = 0$ *and select an arbitrary stationary Markov control law* $g_0 : \mathcal{X} \longmapsto \mathcal{U}$.

2. *(Policy Evaluation) Solve the equations*

$$J_{Q^o}(g_m) = Q^o(g_m) J_{Q^o}(g_m) \tag{5.83}$$

$$J_{Q^o}(g_m) + V_{Q^o}(g_m) = f(g_m) + Q^o(g_m) V_{Q^o}(g_m) \tag{5.84}$$

for $J_{Q^o}(g_m)$ *and* $V_{Q^o}(g_m)$. *Identify the support sets based on the values of* V_{Q^o} *using* (5.67)-(5.68), *and construct the matrix* $Q^*(g_m)$ *using* (5.54)-(5.56). *Solve the equations*

$$J_{Q^*}(g_m) = Q^*(g_m) J_{Q^*}(g_m) \tag{5.85}$$

$$J_{Q^*}(g_m) + V_{Q^*}(g_m) = f(g_m) + Q^*(g_m) V_{Q^*}(g_m) \tag{5.86}$$

for $J_{Q^*}(g_m)$ *and* $V_{Q^*}(g_m)$.

3. *(Policy Improvement) Let*

$$g_{m+1} = \operatorname*{argmin}_{g \in \mathbb{R}^n} \Big\{ f(g) + Q^*(g) V_{Q^*}(g_m) \Big\}. \tag{5.87}$$

4. *If* $g_{m+1} = g_m$ *let* $g^* = g_m$; *else let* $m = m + 1$ *and return to step 2.*

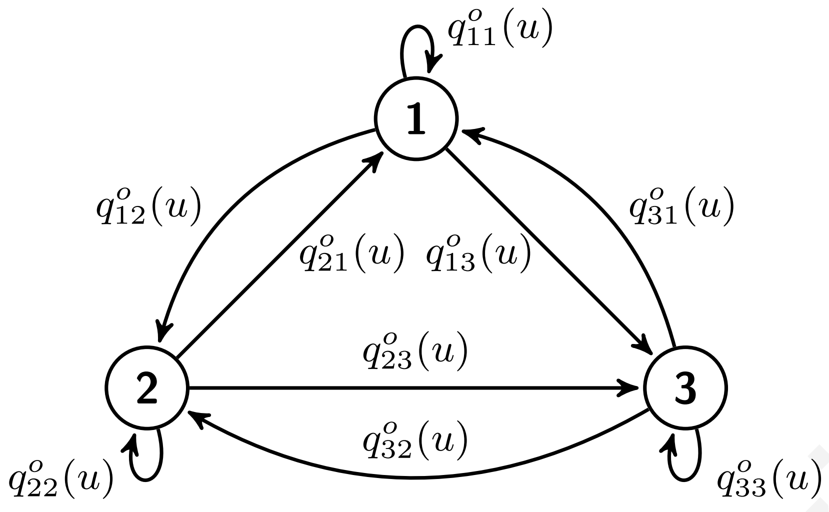
In Section 5.4.3, we illustrate through an example how Algorithm 5.9 is applied.

5.4. Examples

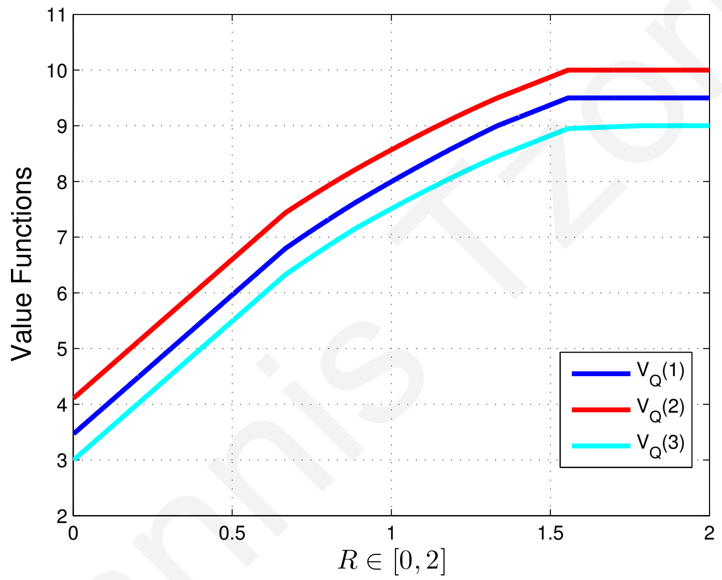
In this Section we illustrate the new dynamic programming equations and the corresponding policy iteration algorithms through examples. In particular, in Section 5.4.1 we present an application of the infinite horizon minimax problem for discounted cost by employing the policy iteration algorithm 5.3. In Section 5.4.2, the example under consideration is identical to the previous one, except that an average cost is considered and policy iteration algorithm 5.6 is employed. In Section 5.4.3, we illustrate an application of the infinite horizon minimax problem for average cost, however, now the stochastic control system under consideration is described by a transition probability graph which is reducible, and hence general policy iteration algorithm 5.9 is applied.

5.4.1. Infinite Horizon D-MCM

Here, we illustrate an application of the infinite horizon minimax problem for discounted cost, by considering the stochastic control system shown in Fig.5.2a, with state space $\mathcal{X} = \{1, 2, 3\}$ and control set $\mathcal{U} = \{u_1, u_2\}$.



(a)



(b)

Figure 5.2.: Optimal Solution of Infinite Horizon D-MCM Example: Plot (a) depicts the transition probability graph. Plot (b) depicts the optimal value as a function of total variation parameter.

Assume the nominal transition probabilities are given under controls u_1 and u_2 by

$$Q^o(u_1) = \frac{1}{9}\begin{pmatrix} 3 & 1 & 5 \\ 4 & 2 & 3 \\ 1 & 6 & 2 \end{pmatrix}, \quad Q^o(u_2) = \frac{1}{9}\begin{pmatrix} 1 & 2 & 6 \\ 4 & 2 & 3 \\ 4 & 1 & 4 \end{pmatrix} \tag{5.88}$$

the discount factor is $\alpha = 0.9$, the total variation distance radius is $R = \frac{6}{9}$, and the cost function under each state and action is

$$f(1,u_1)=2,\ f(2,u_1)=1,\ f(3,u_1)=3,\ f(1,u_2)=0.5,\ f(2,u_2)=3,\ f(3,u_2)=0.$$

Using policy iteration algorithm 5.3, with initial policies $g_0(1)=u_1$, $g_0(2)=u_2$, $g_0(3)=u_2$, the algorithm converge to the following optimal policy and value after two iterations.

$$g^*=g_2\triangleq\begin{pmatrix}g_2(1)\\g_2(2)\\g_2(3)\end{pmatrix}=\begin{pmatrix}u_2\\u_1\\u_2\end{pmatrix},\quad V_{Q^*}(g^*)=V_{Q^*}(g_2)\triangleq\begin{pmatrix}V_{Q^*}(1)\\V_{Q^*}(2)\\V_{Q^*}(3)\end{pmatrix}=\begin{pmatrix}6.79\\7.43\\6.32\end{pmatrix}.$$

Fig.5.2b depicts the optimal value functions for all possible values of R, and shows that, the value functions are non-decreasing and concave functions of R. For the analytic solution of this example, refer to Appendix D.2.

5.4.2. Infinite Horizon Average MCM - Policy Iteration Algorithm 5.6

This example is identical to the previous one Example 5.4.1, except that an average cost function is considered. The stochastic control system is as shown in Fig.5.2a, with state space $\mathcal{X}=\{1,2,3\}$ and $\mathcal{U}=\{u_1,u_2\}$. Assume that the nominal transition probabilities are given by (5.88). The average cost function under each state and action is

$$f(1,u_1)=2,\ f(2,u_1)=1,\ f(3,u_1)=3,\ f(1,u_2)=0.5,\ f(2,u_2)=3,\ f(3,u_2)=0,$$

and the total variation distance radius is $R=\frac{6}{9}$. To obtain an optimal stationary policy of the infinite horizon minimax problem for average cost, policy iteration algorithm 5.6 is applied.

A. Let $m=0$.

1. Select the initial policies as follows

$$g_0(1)=u_1,\quad g_0(2)=u_2,\quad g_0(3)=u_2. \tag{5.89}$$

2. Solve the equation $J_{Q^o}(g_0)e+V_{Q^o}(g_0)=f(g_0)+Q^o(g_0)V_{Q^o}(g_0)$, for $J_{Q^o}(g_0)\in\mathbb{R}$ and $V_{Q^o}(g_0)\in\mathbb{R}^3$, or,

$$J_{Q^o}(g_0)\begin{pmatrix}1\\1\\1\end{pmatrix}+\begin{pmatrix}V_{Q^o}(g_0,1)\\V_{Q^o}(g_0,2)\\V_{Q^o}(g_0,3)\end{pmatrix}=$$
$$\begin{pmatrix}f(1,g_0(1))\\f(2,g_0(2))\\f(3,g_0(3))\end{pmatrix}+\begin{pmatrix}q^o_{11}(g_0(1)) & q^o_{12}(g_0(1)) & q^o_{13}(g_0(1))\\q^o_{21}(g_0(2)) & q^o_{22}(g_0(2)) & q^o_{23}(g_0(2))\\q^o_{31}(g_0(3)) & q^o_{32}(g_0(3)) & q^o_{33}(g_0(3))\end{pmatrix}\begin{pmatrix}V_{Q^o}(g_0,1)\\V_{Q^o}(g_0,2)\\V_{Q^o}(g_0,3)\end{pmatrix}$$

which is given by

$$J_{Q^o}(g_0)\begin{pmatrix}1\\1\\1\end{pmatrix}+\begin{pmatrix}V_{Q^o}(g_0,1)\\V_{Q^o}(g_0,2)\\V_{Q^o}(g_0,3)\end{pmatrix}=\begin{pmatrix}2\\3\\0\end{pmatrix}+\frac{1}{9}\begin{pmatrix}3&1&5\\4&2&3\\4&1&4\end{pmatrix}\begin{pmatrix}V_{Q^o}(g_0,1)\\V_{Q^o}(g_0,2)\\V_{Q^o}(g_0,3)\end{pmatrix}.$$

Since $V_{Q^o}(g_0)$ is uniquely determined up to an additive constant, let $V_{Q^o}(g_0,3)=0$. The solution is

$$\begin{pmatrix}V_{Q^o}(g_0,1)\\V_{Q^o}(g_0,2)\\V_{Q^o}(g_0,3)\end{pmatrix}=\begin{pmatrix}1.8\\3.375\\0\end{pmatrix},\qquad J_{Q^o}(g_o)=1.175. \tag{5.90}$$

Note that, $V_{Q^o}\triangleq\{V_{Q^o}(1),V_{Q^o}(2),V_{Q^o}(3)\}$, $|\mathcal{X}|=3$, and hence $\mathcal{X}^+=\{2\}$, $\mathcal{X}^-=\{3\}$ and $\mathcal{X}_1=\{1\}$. Since the nominal transition probabilities are given by (5.88), the total variation distance is equal to $R=6/9$ and the resulting partition is the same as in the initialization step of Example 5.4.1 (see Appendix D.2) then $Q^*(u_1)$ and $Q^*(u_2)$ are given by (D.3) and (D.4), respectively. The transition probability graph of Q^*, under controls u_1 and u_2, is depicted in Fig.5.3. Note that, under both controls, matrix $Q^*(u)$ remains irreducible.

Next, we proceed to solve the equation $J_{Q^*}(g_0)e+V_{Q^*}(g_0)=f(g_0)+Q^*(g_0)V_{Q^*}(g_0)$, for $J_{Q^*}(g_0)\in\mathbb{R}$ and $V_{Q^*}(g_0)\in\mathbb{R}^3$, or,

$$J_{Q^*}(g_0)\begin{pmatrix}1\\1\\1\end{pmatrix}+\begin{pmatrix}V_{Q^*}(g_0,1)\\V_{Q^*}(g_0,2)\\V_{Q^*}(g_0,3)\end{pmatrix}=$$
$$\begin{pmatrix}f(1,g_0(1))\\f(2,g_0(2))\\f(3,g_0(3))\end{pmatrix}+\begin{pmatrix}q^*_{11}(g_0(1))&q^*_{12}(g_0(1))&q^*_{13}(g_0(1))\\q^*_{21}(g_0(2))&q^*_{22}(g_0(2))&q^*_{23}(g_0(2))\\q^*_{31}(g_0(3))&q^*_{32}(g_0(3))&q^*_{33}(g_0(3))\end{pmatrix}\begin{pmatrix}V_{Q^*}(g_0,1)\\V_{Q^*}(g_0,2)\\V_{Q^*}(g_0,3)\end{pmatrix}$$

which is given by

$$J_{Q^*}(g_0)\begin{pmatrix}1\\1\\1\end{pmatrix}+\begin{pmatrix}V_{Q^*}(g_0,1)\\V_{Q^*}(g_0,2)\\V_{Q^*}(g_0,3)\end{pmatrix}=\begin{pmatrix}2\\3\\0\end{pmatrix}+\frac{1}{9}\begin{pmatrix}3&4&2\\4&5&0\\4&4&1\end{pmatrix}\begin{pmatrix}V_{Q^*}(g_0,1)\\V_{Q^*}(g_0,2)\\V_{Q^*}(g_0,3)\end{pmatrix}.$$

Since $V_{Q^*}(g_0)$ is uniquely determined up to an additive constant, let $V_{Q^*}(g_0,3)=0$. The solution is

$$\begin{pmatrix}V_{Q^*}(g_0,1)\\V_{Q^*}(g_0,2)\\V_{Q^*}(g_0,3)\end{pmatrix}=\begin{pmatrix}1.8\\3.375\\0\end{pmatrix},\qquad J_{Q^*}(g_0)=2.3. \tag{5.91}$$

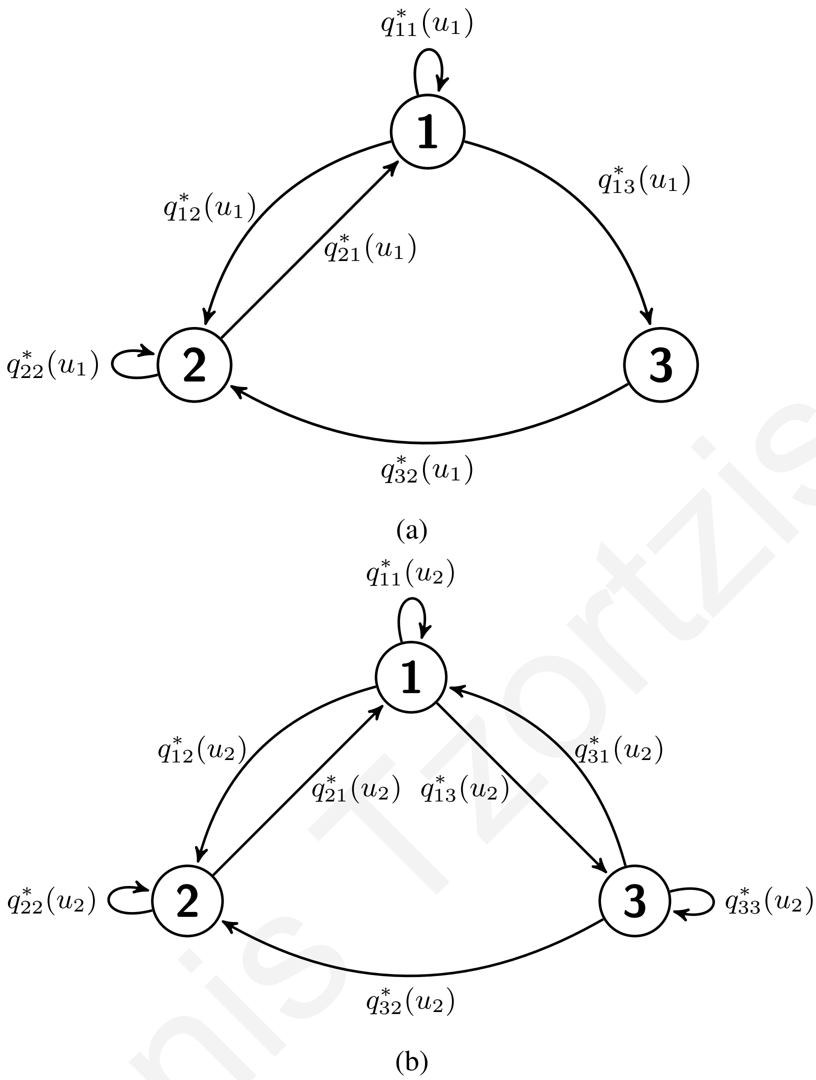


Figure 5.3.: Transition Probability Graph of Q^* under controls u_1 and u_2. Plot (a) depicts matrix Q^* under control u_1. Plot (b) depicts matrix Q^* under control u_2.

3. Let $g_1 = \operatorname{argmin}_{g \in \mathbb{R}^3}\{f(g) + Q^*(g)V_{Q^*}(g_0)\}$, that is,

$$g_1(1) = \operatorname{argmin}\Big\{f(1,u_1) + g^*_{11}(u_1)V_{Q^*}(g_0,1) + g^*_{12}(u_1)V_{Q^*}(g_0,2) + g^*_{13}(u_1)V_{Q^*}(g_0,3)$$
$$f(1,u_2) + g^*_{11}(u_2)V_{Q^*}(g_0,1) + g^*_{12}(u_2)V_{Q^*}(g_0,2) + g^*_{13}(u_2)V_{Q^*}(g_0,3)\Big\}$$
$$= \operatorname{argmin}\Big\{4.099, 2.573\Big\} = \{2\}. \qquad \text{Hence, } g_1(1) = u_2.$$

$$g_1(2) = \operatorname{argmin}\Big\{f(2,u_1) + g^*_{21}(u_1)V_{Q^*}(g_0,1) + g^*_{22}(u_1)V_{Q^*}(g_0,2) + g^*_{23}(u_1)V_{Q^*}(g_0,3)$$
$$f(2,u_2) + g^*_{21}(u_2)V_{Q^*}(g_0,1) + g^*_{22}(u_2)V_{Q^*}(g_0,2) + g^*_{23}(u_2)V_{Q^*}(g_0,3)\Big\}$$
$$= \operatorname{argmin}\Big\{3.673, 5.673\Big\} = \{1\}. \qquad \text{Hence, } g_1(2) = u_1.$$

$$g_1(3) = \operatorname{argmin}\Big\{ f(3,u_1) + g^*_{31}(u_1)V_{Q^*}(g_0,1) + g^*_{32}(u_1)V_{Q^*}(g_0,2) + g^*_{33}(u_1)V_{Q^*}(g_0,3)$$
$$f(3,u_2) + g^*_{31}(u_2)V_{Q^*}(g_0,1) + g^*_{32}(u_2)V_{Q^*}(g_0,2) + g^*_{33}(u_2)V_{Q^*}(g_0,3)\Big\}$$
$$= \operatorname{argmin}\Big\{6.375, 2.3\Big\} = \{2\}. \qquad \text{Hence, } g_1(3) = u_2.$$

Since, $g_1 \neq g_0$, let $m = 1$ and return to step 2.

B. **Let $m = 1$.**

2. Solve the equation $J_{Q^o}(g_1)e + V_{Q^o}(g_1) = f(g_1) + Q^o(g_1)V_{Q^o}(g_1)$, for $J_{Q^o}(g_1) \in \mathbb{R}$ and $V_{Q^o}(g_1) \in \mathbb{R}^3$, or,

$$J_{Q^o}(g_1)\begin{pmatrix}1\\1\\1\end{pmatrix} + \begin{pmatrix}V_{Q^o}(g_1,1)\\V_{Q^o}(g_1,2)\\V_{Q^o}(g_1,3)\end{pmatrix} =$$
$$\begin{pmatrix}f(1,g_1(1))\\f(2,g_1(2))\\f(3,g_1(3))\end{pmatrix} + \begin{pmatrix}q^o_{11}(g_1(1)) & q^o_{12}(g_1(1)) & q^o_{13}(g_1(1))\\ q^o_{21}(g_1(2)) & q^o_{22}(g_1(2)) & q^o_{23}(g_1(2))\\ q^o_{31}(g_1(3)) & q^o_{32}(g_1(3)) & q^o_{33}(g_1(3))\end{pmatrix}\begin{pmatrix}V_{Q^o}(g_0,1)\\V_{Q^o}(g_0,2)\\V_{Q^o}(g_0,3)\end{pmatrix}$$

which is given by

$$J_{Q^o}(g_1)\begin{pmatrix}1\\1\\1\end{pmatrix} + \begin{pmatrix}V_{Q^o}(g_1,1)\\V_{Q^o}(g_1,2)\\V_{Q^o}(g_1,3)\end{pmatrix} = \begin{pmatrix}0.5\\1\\0\end{pmatrix} + \frac{1}{9}\begin{pmatrix}1 & 2 & 6\\4 & 2 & 3\\4 & 1 & 4\end{pmatrix}\begin{pmatrix}V_{Q^o}(g_1,1)\\V_{Q^o}(g_1,2)\\V_{Q^o}(g_1,3)\end{pmatrix}.$$

Let $V_{Q^o}(g_1,3) = 0$. The solution is

$$\begin{pmatrix}V_{Q^o}(g_1,1)\\V_{Q^o}(g_1,2)\\V_{Q^o}(g_1,3)\end{pmatrix} = \begin{pmatrix}0.468\\1.125\\0\end{pmatrix}, \qquad J_{Q^o}(g_o) = 0.333. \tag{5.92}$$

Therefore, $\mathcal{X}^+ = \{2\}$, $\mathcal{X}^- = \{3\}$ and $\mathcal{X}_1 = \{1\}$. Since the partition is the same as in $m = 0$ then $Q^*(u_1)$ and $Q^*(u_2)$ are given by (D.3) and (D.4), respectively.

Solve the equation $J_{Q^*}(g_1)e + V_{Q^*}(g_1) = f(g_1) + Q^*(g_1)V_{Q^*}(g_1)$, for $J_{Q^*}(g_1) \in \mathbb{R}$ and $V_{Q^*}(g_1) \in \mathbb{R}^3$, or,

$$J_{Q^*}(g_1)\begin{pmatrix}1\\1\\1\end{pmatrix} + \begin{pmatrix}V_{Q^*}(g_1,1)\\V_{Q^*}(g_1,2)\\V_{Q^*}(g_1,3)\end{pmatrix} =$$
$$\begin{pmatrix}f(1,g_1(1))\\f(2,g_1(2))\\f(3,g_1(3))\end{pmatrix} + \begin{pmatrix}q^*_{11}(g_1(1)) & q^*_{12}(g_1(1)) & q^*_{13}(g_1(1))\\ q^*_{21}(g_1(2)) & q^*_{22}(g_1(2)) & q^*_{23}(g_1(2))\\ q^*_{31}(g_1(3)) & q^*_{32}(g_1(3)) & q^*_{33}(g_1(3))\end{pmatrix}\begin{pmatrix}V_{Q^*}(g_1,1)\\V_{Q^*}(g_1,2)\\V_{Q^*}(g_1,3)\end{pmatrix}$$

which is given by

$$J_{Q^*}(g_1)\begin{pmatrix}1\\1\\1\end{pmatrix}+\begin{pmatrix}V_{Q^*}(g_1,1)\\V_{Q^*}(g_1,2)\\V_{Q^*}(g_1,3)\end{pmatrix}=\begin{pmatrix}0.5\\1\\0\end{pmatrix}+\frac{1}{9}\begin{pmatrix}1&5&3\\4&5&0\\4&4&1\end{pmatrix}\begin{pmatrix}V_{Q^*}(g_1,1)\\V_{Q^*}(g_1,2)\\V_{Q^*}(g_1,3)\end{pmatrix}.$$

Let $V_{Q^*}(g_1,3)=0$. The solution is

$$\begin{pmatrix}V_{Q^*}(g_1,1)\\V_{Q^*}(g_1,2)\\V_{Q^*}(g_1,3)\end{pmatrix}=\begin{pmatrix}0.468\\1.125\\0\end{pmatrix},\qquad J_{Q^*}(g_1)=0.708. \tag{5.93}$$

3. Let $g_2=\operatorname{argmin}_{g\in\mathbb{R}^3}\{f(g)+Q^*(g)V_{Q^*}(g_1)\}$, that is,

$$\begin{aligned}g_2(1)=\operatorname{argmin}\Big\{&f(1,u_1)+g^*_{11}(u_1)V_{Q^*}(g_1,1)+g^*_{12}(u_1)V_{Q^*}(g_1,2)+g^*_{13}(u_1)V_{Q^*}(g_1,3)\\&f(1,u_2)+g^*_{11}(u_2)V_{Q^*}(g_1,1)+g^*_{12}(u_2)V_{Q^*}(g_1,2)+g^*_{13}(u_2)V_{Q^*}(g_1,3)\Big\}\\=\operatorname{argmin}&\Big\{2.656,1.177\Big\}=\{2\}.\qquad\text{Hence, }\ g_2(1)=u_2.\end{aligned}$$

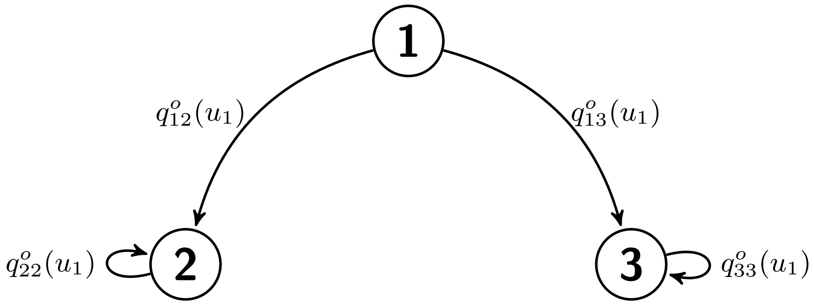
$$\begin{aligned}g_2(2)=\operatorname{argmin}\Big\{&f(2,u_1)+g^*_{21}(u_1)V_{Q^*}(g_1,1)+g^*_{22}(u_1)V_{Q^*}(g_1,2)+g^*_{23}(u_1)V_{Q^*}(g_1,3)\\&f(2,u_2)+g^*_{21}(u_2)V_{Q^*}(g_1,1)+g^*_{22}(u_2)V_{Q^*}(g_1,2)+g^*_{23}(u_2)V_{Q^*}(g_1,3)\Big\}\\=\operatorname{argmin}&\Big\{1.831,3.831\Big\}=\{1\}.\qquad\text{Hence, }\ g_2(2)=u_1.\end{aligned}$$

$$\begin{aligned}g_2(3)=\operatorname{argmin}\Big\{&f(3,u_1)+g^*_{31}(u_1)V_{Q^*}(g_1,1)+g^*_{32}(u_1)V_{Q^*}(g_1,2)+g^*_{33}(u_1)V_{Q^*}(g_1,3)\\&f(3,u_2)+g^*_{31}(u_2)V_{Q^*}(g_1,1)+g^*_{32}(u_2)V_{Q^*}(g_1,2)+g^*_{33}(u_2)V_{Q^*}(g_1,3)\Big\}\\=\operatorname{argmin}&\Big\{4.125,0.708\Big\}=\{2\}.\qquad\text{Hence, }\ g_2(3)=u_2.\end{aligned}$$

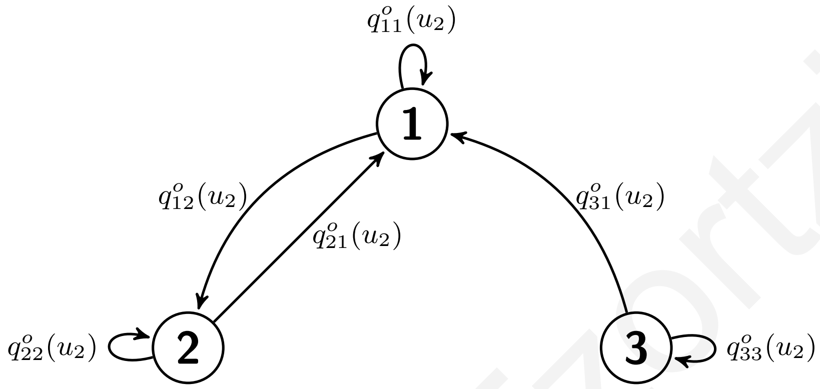
4. Because, $g_2=g_1$, then $g^*=g_1$ is an optimal control law with $J_{Q^*}=0.708$, $V_{Q^*}(1)=0.468$, $V_{Q^*}(2)=1.125$ and $V_{Q^*}(3)=0$.

5.4.3. Infinite Horizon Average MCM - General Policy Iteration Algorithm 5.9

In this example, we illustrate an application of the infinite horizon minimax problem for average cost, by considering the stochastic control system shown in Fig.5.4, with $\mathcal{X}=\{1,2,3\}$ and control set $\mathcal{U}=\{u_1,u_2\}$. The essential difference between this example and the previous one, is that here, the stochastic control system under consideration is described by a



(a) Matrix Q^o under control u_1



(b) Matrix Q^o under control u_2

Figure 5.4.: Transition Probability Graph of Q^o under controls u_1 and u_2

transition probability graph which is reducible, and hence general policy iteration algorithm 5.9 is applied.

Assume the nominal transition probabilities are given under controls u_1 and u_2 by

$$Q^o(u_1) = \frac{1}{9}\begin{pmatrix} 0 & 5 & 4 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix}, \quad Q^o(u_2) = \frac{1}{9}\begin{pmatrix} 2 & 7 & 0 \\ 3 & 6 & 0 \\ 8 & 0 & 1 \end{pmatrix} \tag{5.94}$$

the total variation distance radius is $R = 14/9$, and the cost function under each state and action is

$$f(1,u_1) = 2,\ f(2,u_1) = 1,\ f(3,u_1) = 3,\ f(1,u_2) = 0.5,\ f(2,u_2) = 3,\ f(3,u_2) = 0.$$

A. Let $m = 0$.

1. Select the initial policies as follows

$$g_0(1) = u_1, \quad g_0(2) = u_1, \quad g_0(3) = u_1. \tag{5.95}$$

2. Solve the equation $J_{Q^o}(g_0) = Q^o(g_0) J_{Q^o}(g_0)$, or,

$$\begin{pmatrix} J_{Q^o}(g_0,1) \\ J_{Q^o}(g_0,2) \\ J_{Q^o}(g_0,3) \end{pmatrix} = \begin{pmatrix} q^o_{11}(g_0(1)) & q^o_{12}(g_0(1)) & q^o_{13}(g_0(1)) \\ q^o_{21}(g_0(2)) & q^o_{22}(g_0(2)) & q^o_{23}(g_0(2)) \\ q^o_{31}(g_0(3)) & q^o_{32}(g_0(3)) & q^o_{33}(g_0(3)) \end{pmatrix} \begin{pmatrix} J_{Q^o}(g_0,1) \\ J_{Q^o}(g_0,2) \\ J_{Q^o}(g_0,3) \end{pmatrix}$$
$$= \frac{1}{9} \begin{pmatrix} 0 & 5 & 4 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix} \begin{pmatrix} J_{Q^o}(g_0,1) \\ J_{Q^o}(g_0,2) \\ J_{Q^o}(g_0,3) \end{pmatrix}.$$

The optimality equations (5.83) are

$$J_{Q^o}(g_0,1) = \frac{5}{9} J_{Q^o}(g_0,2) + \frac{4}{9} J_{Q^o}(g_0,3), \tag{5.96a}$$
$$J_{Q^o}(g_0,2) = J_{Q^o}(g_0,2), \tag{5.96b}$$
$$J_{Q^o}(g_0,3) = J_{Q^o}(g_0,3). \tag{5.96c}$$

Next, solve the equation $J_{Q^o}(g_0) + V_{Q^o}(g_0) = f(g_0) + Q^o(g_0) V_{Q^o}(g_0)$, for $J_{Q^o}(g_0) \in \mathbb{R}^3$ and $V_{Q^o}(g_0) \in \mathbb{R}^3$, or,

$$\begin{pmatrix} J_{Q^o}(g_0,1) \\ J_{Q^o}(g_0,2) \\ J_{Q^o}(g_0,3) \end{pmatrix} + \begin{pmatrix} V_{Q^o}(g_0,1) \\ V_{Q^o}(g_0,2) \\ V_{Q^o}(g_0,3) \end{pmatrix} =$$
$$\begin{pmatrix} f(1,g_0(1)) \\ f(2,g_0(2)) \\ f(3,g_0(3)) \end{pmatrix} + \begin{pmatrix} q^o_{11}(g_0(1)) & q^o_{12}(g_0(1)) & q^o_{13}(g_0(1)) \\ q^o_{21}(g_0(2)) & q^o_{22}(g_0(2)) & q^o_{23}(g_0(2)) \\ q^o_{31}(g_0(3)) & q^o_{32}(g_0(3)) & q^o_{33}(g_0(3)) \end{pmatrix} \begin{pmatrix} V_{Q^o}(g_0,1) \\ V_{Q^o}(g_0,2) \\ V_{Q^o}(g_0,3) \end{pmatrix}$$

which is given by

$$\begin{pmatrix} J_{Q^o}(g_0,1) \\ J_{Q^o}(g_0,2) \\ J_{Q^o}(g_0,3) \end{pmatrix} + \begin{pmatrix} V_{Q^o}(g_0,1) \\ V_{Q^o}(g_0,2) \\ V_{Q^o}(g_0,3) \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} + \frac{1}{9} \begin{pmatrix} 0 & 5 & 4 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix} \begin{pmatrix} V_{Q^o}(g_0,1) \\ V_{Q^o}(g_0,2) \\ V_{Q^o}(g_0,3) \end{pmatrix}$$

The optimality equations (5.84) are given by

$$J_{Q^o}(g_0,1) + V_{Q^o}(g_0,1) = 2 + \frac{5}{9} V_{Q^o}(g_0,2) + \frac{4}{9} V_{Q^o}(g_0,3) \tag{5.97a}$$
$$J_{Q^o}(g_0,2) + V_{Q^o}(g_0,2) = 1 + V_{Q^o}(g_0,2) \tag{5.97b}$$
$$J_{Q^o}(g_0,3) + V_{Q^o}(g_0,3) = 3 + V_{Q^o}(g_0,3) \tag{5.97c}$$

The solution of (5.96) and (5.97) has

$$V_{Q^o}(g_0,1) = \frac{1}{9} + \frac{5}{9}\alpha + \frac{4}{9}\beta, \qquad V_{Q^o}(g_0,2) = \alpha, \qquad V_{Q^o}(g_0,3) = \beta,$$
$$J_{Q^o}(g_0,1) = 1.888, \qquad J_{Q^o}(g_0,2) = 1, \qquad J_{Q^o}(g_0,3) = 3.$$

Setting $\alpha = 1$ and $\beta = 0$ yields

$$V_{Q^o}(g_0,1) = 0.666, \qquad V_{Q^o}(g_0,2) = 1, \qquad V_{Q^o}(g_0,3) = 0.$$

Note that, $V_{Q^o} = \{V_{Q^o}(1), V_{Q^o}(2), V_{Q^o}(3)\}$, and hence the support sets based on the values of V_{Q^o} are $\mathcal{X}^+ = \{2\}$, $\mathcal{X}^- = \{3\}$ and $\mathcal{X}_1 = \{1\}$. Once the partition is been identified, (5.54)-(5.56) is applied to obtain

$$Q^*(u_1) = \begin{pmatrix} \left(q^o_{11}(u_1) - \left(\frac{R}{2} - q^o_{13}(u_1)\right)^+\right)^+ & \min\left(1, q^o_{12}(u_1) + \frac{R}{2}\right) & \left(q^o_{13}(u_1) - \frac{R}{2}\right)^+ \\ \left(q^o_{21}(u_1) - \left(\frac{R}{2} - q^o_{23}(u_1)\right)^+\right)^+ & \min\left(1, q^o_{22}(u_1) + \frac{R}{2}\right) & \left(q^o_{23}(u_1) - \frac{R}{2}\right)^+ \\ \left(q^o_{31}(u_1) - \left(\frac{R}{2} - q^o_{33}(u_1)\right)^+\right)^+ & \min\left(1, q^o_{32}(u_1) + \frac{R}{2}\right) & \left(q^o_{33}(u_1) - \frac{R}{2}\right)^+ \end{pmatrix}$$
$$= \frac{1}{9}\begin{pmatrix} 0 & 9 & 0 \\ 0 & 9 & 0 \\ 0 & 7 & 2 \end{pmatrix} \tag{5.98}$$

and

$$Q^*(u_2) = \begin{pmatrix} \left(q^o_{11}(u_2) - \left(\frac{R}{2} - q^o_{13}(u_2)\right)^+\right)^+ & \min\left(1, q^o_{12}(u_2) + \frac{R}{2}\right) & \left(q^o_{13}(u_2) - \frac{R}{2}\right)^+ \\ \left(q^o_{21}(u_2) - \left(\frac{R}{2} - q^o_{23}(u_2)\right)^+\right)^+ & \min\left(1, q^o_{22}(u_2) + \frac{R}{2}\right) & \left(q^o_{23}(u_2) - \frac{R}{2}\right)^+ \\ \left(q^o_{31}(u_2) - \left(\frac{R}{2} - q^o_{33}(u_2)\right)^+\right)^+ & \min\left(1, q^o_{32}(u_2) + \frac{R}{2}\right) & \left(q^o_{33}(u_2) - \frac{R}{2}\right)^+ \end{pmatrix}$$
$$= \frac{1}{9}\begin{pmatrix} 0 & 9 & 0 \\ 0 & 9 & 0 \\ 2 & 7 & 0 \end{pmatrix}. \tag{5.99}$$

Next, solve the equation $J_{Q^*}(g_0) = Q^*(g_0)J_{Q^*}(g_0)$, or,

$$\begin{pmatrix} J_{Q^*}(g_0,1) \\ J_{Q^*}(g_0,2) \\ J_{Q^*}(g_0,3) \end{pmatrix} = \begin{pmatrix} q^*_{11}(g_0(1)) & q^*_{12}(g_0(1)) & q^*_{13}(g_0(1)) \\ q^*_{21}(g_0(2)) & q^*_{22}(g_0(2)) & q^*_{23}(g_0(2)) \\ q^*_{31}(g_0(3)) & q^*_{32}(g_0(3)) & q^*_{33}(g_0(3)) \end{pmatrix} \begin{pmatrix} J_{Q^*}(g_0,1) \\ J_{Q^*}(g_0,2) \\ J_{Q^*}(g_0,3) \end{pmatrix}$$
$$= \frac{1}{9}\begin{pmatrix} 0 & 9 & 0 \\ 0 & 9 & 0 \\ 0 & 7 & 2 \end{pmatrix} \begin{pmatrix} J_{Q^*}(g_0,1) \\ J_{Q^*}(g_0,2) \\ J_{Q^*}(g_0,3) \end{pmatrix}.$$

The optimality equations (5.85) are

$$J_{Q^*}(g_0,1) = J_{Q^*}(g_0,2), \tag{5.100a}$$
$$J_{Q^*}(g_0,2) = J_{Q^*}(g_0,2), \tag{5.100b}$$
$$J_{Q^*}(g_0,3) = \frac{7}{9}J_{Q^*}(g_0,2) + \frac{2}{9}J_{Q^*}(g_0,3) \tag{5.100c}$$

and hence, $J_{Q^*}(g_0, 1) = J_{Q^*}(g_0, 2) = J_{Q^*}(g_0, 3)$.

Next, solve the equation $J_{Q^*}(g_0) + V_{Q^*}(g_0) = f(g_0) + Q^*(g_0)V_{Q^*}(g_0)$, for $J_{Q^*}(g_0) \in \mathbb{R}^3$ and $V_{Q^*}(g_0) \in \mathbb{R}^3$, or,

$$\begin{pmatrix} J_{Q^*}(g_0,1) \\ J_{Q^*}(g_0,2) \\ J_{Q^*}(g_0,3) \end{pmatrix} + \begin{pmatrix} V_{Q^*}(g_0,1) \\ V_{Q^*}(g_0,2) \\ V_{Q^*}(g_0,3) \end{pmatrix} = \begin{pmatrix} f(1,g_0(1)) \\ f(2,g_0(2)) \\ f(3,g_0(3)) \end{pmatrix} + \begin{pmatrix} q^*_{11}(g_0(1)) & q^*_{12}(g_0(1)) & q^*_{13}(g_0(1)) \\ q^*_{21}(g_0(2)) & q^*_{22}(g_0(2)) & q^*_{23}(g_0(2)) \\ q^*_{31}(g_0(3)) & q^*_{32}(g_0(3)) & q^*_{33}(g_0(3)) \end{pmatrix} \begin{pmatrix} V_{Q^*}(g_0,1) \\ V_{Q^*}(g_0,2) \\ V_{Q^*}(g_0,3) \end{pmatrix}$$

which is given by

$$\begin{pmatrix} J_{Q^*}(g_0,1) \\ J_{Q^*}(g_0,2) \\ J_{Q^*}(g_0,3) \end{pmatrix} + \begin{pmatrix} V_{Q^*}(g_0,1) \\ V_{Q^*}(g_0,2) \\ V_{Q^*}(g_0,3) \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} + \frac{1}{9}\begin{pmatrix} 0 & 9 & 0 \\ 0 & 9 & 0 \\ 0 & 7 & 2 \end{pmatrix} \begin{pmatrix} V_{Q^*}(g_0,1) \\ V_{Q^*}(g_0,2) \\ V_{Q^*}(g_0,3) \end{pmatrix}$$

The optimality equations (5.86) are given by

$$J_{Q^*}(g_0,1) + V_{Q^*}(g_0,1) = 2 + V_{Q^*}(g_0,2) \tag{5.101a}$$

$$J_{Q^*}(g_0,2) + V_{Q^*}(g_0,2) = 1 + V_{Q^*}(g_0,2) \tag{5.101b}$$

$$J_{Q^*}(g_0,3) + V_{Q^*}(g_0,3) = 3 + \frac{7}{9}V_{Q^*}(g_0,2) + \frac{2}{9}V_{Q^*}(g_0,3) \tag{5.101c}$$

The solution of (5.100) and (5.101) has

$$V_{Q^*}(g_0,1) = 1 + \alpha, \qquad V_{Q^*}(g_0,2) = \alpha, \qquad V_{Q^*}(g_0,3) = \frac{18}{7} + \alpha,$$
$$J_{Q^*}(g_0,1) = 1, \qquad J_{Q^*}(g_0,2) = 1, \qquad J_{Q^*}(g_0,3) = 1.$$

Setting $\alpha = 1$ yields

$$V_{Q^*}(g_0,1) = 2, \qquad V_{Q^*}(g_0,2) = 1, \qquad V_{Q^*}(g_0,3) = 3.57.$$

3. Let $g_1 = \text{argmin}_{g \in \mathbb{R}^3}\{f(g) + Q^*(g)V_{Q^*}(g_0)\}$, that is,

$$\begin{aligned} g_1(1) = \text{argmin}\Big\{ & f(1,u_1) + g^*_{11}(u_1)V_{Q^*}(g_0,1) + g^*_{12}(u_1)V_{Q^*}(g_0,2) + g^*_{13}(u_1)V_{Q^*}(g_0,3), \\ & f(1,u_2) + g^*_{11}(u_2)V_{Q^*}(g_0,1) + g^*_{12}(u_2)V_{Q^*}(g_0,2) + g^*_{13}(u_2)V_{Q^*}(g_0,3)\Big\} \\ = \text{argmin}\Big\{3, 1.5\Big\} = \{2\}. \qquad \text{Hence, } g_1(1) = u_2. \end{aligned}$$

$$\begin{aligned}g_1(2) = \operatorname{argmin} \Big\{ &\\ &f(2,u_1) + g^*_{21}(u_1)V_{Q^*}(g_0,1) + g^*_{22}(u_1)V_{Q^*}(g_0,2) + g^*_{23}(u_1)V_{Q^*}(g_0,3), \\ &\quad f(2,u_2) + g^*_{21}(u_2)V_{Q^*}(g_0,1) + g^*_{22}(u_2)V_{Q^*}(g_0,2) + g^*_{23}(u_2)V_{Q^*}(g_0,3)\Big\} \\ &= \operatorname{argmin}\Big\{2,4\Big\} = \{1\}. \qquad \text{Hence, } g_1(2) = u_1.\end{aligned}$$

$$\begin{aligned}g_1(3) = \operatorname{argmin} \Big\{ &\\ &f(3,u_1) + g^*_{31}(u_1)V_{Q^*}(g_0,1) + g^*_{32}(u_1)V_{Q^*}(g_0,2) + g^*_{33}(u_1)V_{Q^*}(g_0,3), \\ &\quad f(3,u_2) + g^*_{31}(u_2)V_{Q^*}(g_0,1) + g^*_{32}(u_2)V_{Q^*}(g_0,2) + g^*_{33}(u_2)V_{Q^*}(g_0,3)\Big\} \\ &= \operatorname{argmin}\Big\{4.57, 1.222\Big\} = \{2\}. \qquad \text{Hence, } g_1(3) = u_2.\end{aligned}$$

Since $g_1 \neq g_0$, let $m = 1$ and return to step 2.

B. Let $m = 1$.

2. Solve the equation $J_{Q^o}(g_1) = Q^o(g_1)J_{Q^o}(g_1)$, or,

$$\begin{pmatrix} J_{Q^o}(g_1,1) \\ J_{Q^o}(g_1,2) \\ J_{Q^o}(g_1,3) \end{pmatrix} = \begin{pmatrix} q^o_{11}(g_1(1)) & q^o_{12}(g_1(1)) & q^o_{13}(g_1(1)) \\ q^o_{21}(g_1(2)) & q^o_{22}(g_1(2)) & q^o_{23}(g_1(2)) \\ q^o_{31}(g_1(3)) & q^o_{32}(g_1(3)) & q^o_{33}(g_1(3)) \end{pmatrix} \begin{pmatrix} J_{Q^o}(g_1,1) \\ J_{Q^o}(g_1,2) \\ J_{Q^o}(g_1,3) \end{pmatrix}$$
$$= \frac{1}{9}\begin{pmatrix} 2 & 7 & 0 \\ 0 & 9 & 0 \\ 8 & 0 & 1 \end{pmatrix} \begin{pmatrix} J_{Q^o}(g_1,1) \\ J_{Q^o}(g_1,2) \\ J_{Q^o}(g_1,3) \end{pmatrix}.$$

The optimality equations (5.83) are

$$J_{Q^o}(g_1,1) = \frac{2}{9}J_{Q^o}(g_1,1) + \frac{7}{9}J_{Q^o}(g_1,2), \tag{5.102a}$$
$$J_{Q^o}(g_1,2) = J_{Q^o}(g_1,2), \tag{5.102b}$$
$$J_{Q^o}(g_1,3) = \frac{8}{9}J_{Q^o}(g_1,1) + \frac{1}{9}J_{Q^o}(g_1,3). \tag{5.102c}$$

and hence, $J_{Q^o}(g_1,1) = J_{Q^o}(g_1,2) = J_{Q^o}(g_1,3)$.

Next, solve the equation $J_{Q^o}(g_1) + V_{Q^o}(g_1) = f(g_1) + Q^o(g_1)V_{Q^o}(g_1)$, for $J_{Q^o}(g_1) \in \mathbb{R}^3$

and $V_{Q^o}(g_1) \in \mathbb{R}^3$, or,

$$\begin{pmatrix} J_{Q^o}(g_1,1) \\ J_{Q^o}(g_1,2) \\ J_{Q^o}(g_1,3) \end{pmatrix} + \begin{pmatrix} V_{Q^o}(g_1,1) \\ V_{Q^o}(g_1,2) \\ V_{Q^o}(g_1,3) \end{pmatrix} =$$
$$\begin{pmatrix} f(1,g_1(1)) \\ f(2,g_1(2)) \\ f(3,g_1(3)) \end{pmatrix} + \begin{pmatrix} q^o_{11}(g_1(1)) & q^o_{12}(g_1(1)) & q^o_{13}(g_1(1)) \\ q^o_{21}(g_1(2)) & q^o_{22}(g_1(2)) & q^o_{23}(g_1(2)) \\ q^o_{31}(g_1(3)) & q^o_{32}(g_1(3)) & q^o_{33}(g_1(3)) \end{pmatrix} \begin{pmatrix} V_{Q^o}(g_1,1) \\ V_{Q^o}(g_1,2) \\ V_{Q^o}(g_1,3) \end{pmatrix}$$

which is given by

$$\begin{pmatrix} J_{Q^o}(g_1,1) \\ J_{Q^o}(g_1,2) \\ J_{Q^o}(g_1,3) \end{pmatrix} + \begin{pmatrix} V_{Q^o}(g_1,1) \\ V_{Q^o}(g_1,2) \\ V_{Q^o}(g_1,3) \end{pmatrix} = \begin{pmatrix} 0.5 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{9}\begin{pmatrix} 2 & 7 & 0 \\ 0 & 9 & 0 \\ 8 & 0 & 1 \end{pmatrix} \begin{pmatrix} V_{Q^o}(g_1,1) \\ V_{Q^o}(g_1,2) \\ V_{Q^o}(g_1,3) \end{pmatrix}$$

The optimality equations (5.84) are given by

$$J_{Q^o}(g_1,1) + V_{Q^o}(g_1,1) = 0.5 + \frac{2}{9}V_{Q^o}(g_1,1) + \frac{7}{9}V_{Q^o}(g_1,2) \tag{5.103a}$$
$$J_{Q^o}(g_1,2) + V_{Q^o}(g_1,2) = 1 + V_{Q^o}(g_1,2) \tag{5.103b}$$
$$J_{Q^o}(g_1,3) + V_{Q^o}(g_1,3) = \frac{8}{9}V_{Q^o}(g_1,1) + \frac{1}{9}V_{Q^o}(g_1,3) \tag{5.103c}$$

The solution of (5.102) and (5.103) has

$$V_{Q^o}(g_1,1) = \alpha + \frac{9}{8}, \qquad V_{Q^o}(g_1,2) = \alpha + \frac{99}{56}, \qquad V_{Q^o}(g_1,3) = \alpha,$$
$$J_{Q^o}(g_1,1) = 1, \qquad J_{Q^o}(g_1,2) = 1, \qquad J_{Q^o}(g_1,3) = 1.$$

Setting $\alpha = 1$ yields

$$V_{Q^o}(g_1,1) = 2.125, \qquad V_{Q^o}(g_1,2) = 2.76, \qquad V_{Q^o}(g_1,3) = 1.$$

Hence, we proceed with the identification of the support sets, which are $\mathcal{X}^+ = \{2\}$, $\mathcal{X}^- = \{3\}$ and $\mathcal{X}_1 = \{1\}$. Since the partition is the same as in $m = 0$ then $Q^*(u_1)$ and $Q^*(u_2)$ are equal to (5.98) and (5.99), respectively.

Next, solve the equation $J_{Q^*}(g_1) = Q^*(g_1)J_{Q^*}(g_1)$, or,

$$\begin{pmatrix} J_{Q^*}(g_1,1) \\ J_{Q^*}(g_1,2) \\ J_{Q^*}(g_1,3) \end{pmatrix} = \begin{pmatrix} q^*_{11}(g_1(1)) & q^*_{12}(g_1(1)) & q^*_{13}(g_1(1)) \\ q^*_{21}(g_1(2)) & q^*_{22}(g_1(2)) & q^*_{23}(g_1(2)) \\ q^*_{31}(g_1(3)) & q^*_{32}(g_1(3)) & q^*_{33}(g_1(3)) \end{pmatrix} \begin{pmatrix} J_{Q^*}(g_1,1) \\ J_{Q^*}(g_1,2) \\ J_{Q^*}(g_1,3) \end{pmatrix}$$
$$= \frac{1}{9}\begin{pmatrix} 0 & 9 & 0 \\ 0 & 9 & 0 \\ 2 & 7 & 0 \end{pmatrix} \begin{pmatrix} J_{Q^*}(g_0,1) \\ J_{Q^*}(g_0,2) \\ J_{Q^*}(g_0,3) \end{pmatrix}.$$

The optimality equations (5.85) are

$$J_{Q^*}(g_1,1) = J_{Q^*}(g_1,2), \tag{5.104a}$$

$$J_{Q^*}(g_1,2) = J_{Q^*}(g_1,2), \tag{5.104b}$$

$$J_{Q^*}(g_1,3) = \frac{2}{9}J_{Q^*}(g_1,1) + \frac{7}{9}J_{Q^*}(g_1,2) \tag{5.104c}$$

and hence, $J_{Q^*}(g_1,1) = J_{Q^*}(g_1,2) = J_{Q^*}(g_1,3)$.

Next, solve the equation $J_{Q^*}(g_1) + V_{Q^*}(g_1) = f(g_1) + Q^*(g_1)V_{Q^*}(g_1)$, for $J_{Q^*}(g_1) \in \mathbb{R}^3$ and $V_{Q^*}(g_1) \in \mathbb{R}^3$, or,

$$\begin{pmatrix} J_{Q^*}(g_1,1) \\ J_{Q^*}(g_1,2) \\ J_{Q^*}(g_1,3) \end{pmatrix} + \begin{pmatrix} V_{Q^*}(g_1,1) \\ V_{Q^*}(g_1,2) \\ V_{Q^*}(g_1,3) \end{pmatrix} = \begin{pmatrix} f(1,g_1(1)) \\ f(2,g_1(2)) \\ f(3,g_1(3)) \end{pmatrix} + \begin{pmatrix} q^*_{11}(g_1(1)) & q^*_{12}(g_1(1)) & q^*_{13}(g_1(1)) \\ q^*_{21}(g_1(2)) & q^*_{22}(g_1(2)) & q^*_{23}(g_1(2)) \\ q^*_{31}(g_1(3)) & q^*_{32}(g_1(3)) & q^*_{33}(g_1(3)) \end{pmatrix} \begin{pmatrix} V_{Q^*}(g_1,1) \\ V_{Q^*}(g_1,2) \\ V_{Q^*}(g_1,3) \end{pmatrix}$$

which is given by

$$\begin{pmatrix} J_{Q^*}(g_1,1) \\ J_{Q^*}(g_1,2) \\ J_{Q^*}(g_1,3) \end{pmatrix} + \begin{pmatrix} V_{Q^*}(g_1,1) \\ V_{Q^*}(g_1,2) \\ V_{Q^*}(g_1,3) \end{pmatrix} = \begin{pmatrix} 0.5 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{9}\begin{pmatrix} 0 & 9 & 0 \\ 0 & 9 & 0 \\ 2 & 7 & 0 \end{pmatrix} \begin{pmatrix} V_{Q^*}(g_1,1) \\ V_{Q^*}(g_1,2) \\ V_{Q^*}(g_1,3) \end{pmatrix}$$

The optimality equations (5.86) are given by

$$J_{Q^*}(g_1,1) + V_{Q^*}(g_1,1) = 0.5 + V_{Q^*}(g_1,2) \tag{5.105a}$$

$$J_{Q^*}(g_1,2) + V_{Q^*}(g_1,2) = 1 + V_{Q^*}(g_1,2) \tag{5.105b}$$

$$J_{Q^*}(g_1,3) + V_{Q^*}(g_1,3) = \frac{2}{9}V_{Q^*}(g_1,1) + \frac{7}{9}V_{Q^*}(g_1,2) \tag{5.105c}$$

The solution of (5.104) and (5.105) has

$$V_{Q^*}(g_1,1) = \alpha + \frac{11}{18}, \qquad V_{Q^*}(g_1,2) = \alpha + \frac{10}{9}, \qquad V_{Q^*}(g_1,3) = \alpha,$$

$$J_{Q^*}(g_1,1) = 1, \qquad J_{Q^*}(g_1,2) = 1, \qquad J_{Q^*}(g_1,3) = 1.$$

Setting $\alpha = 1$ yields

$$V_{Q^*}(g_1,1) = 1.611, \qquad V_{Q^*}(g_1,2) = 2.111, \qquad V_{Q^*}(g_1,3) = 1.$$

3. Let $g_2 = \text{argmin}_{g\in\mathbb{R}^3}\{f(g) + Q^*(g)V_{Q^*}(g_1)\}$, that is,

$$g_2(1) = \text{argmin}\Big\{ f(1,u_1) + g^*_{11}(u_1)V_{Q^*}(g_1,1) + g^*_{12}(u_1)V_{Q^*}(g_1,2) + g^*_{13}(u_1)V_{Q^*}(g_1,3), f(1,u_2) + g^*_{11}(u_2)V_{Q^*}(g_1,1) + g^*_{12}(u_2)V_{Q^*}(g_1,2) + g^*_{13}(u_2)V_{Q^*}(g_1,3)\Big\}$$
$$= \text{argmin}\Big\{4.111, 2.611\Big\} = \{2\}. \qquad \text{Hence, } g_2(1) = u_2.$$

$$g_2(2) = \text{argmin}\Big\{ f(2,u_1) + g^*_{21}(u_1)V_{Q^*}(g_1,1) + g^*_{22}(u_1)V_{Q^*}(g_1,2) + g^*_{23}(u_1)V_{Q^*}(g_1,3), f(2,u_2) + g^*_{21}(u_2)V_{Q^*}(g_1,1) + g^*_{22}(u_2)V_{Q^*}(g_1,2) + g^*_{23}(u_2)V_{Q^*}(g_1,3)\Big\}$$
$$= \text{argmin}\Big\{3.111, 5.111\Big\} = \{1\}. \qquad \text{Hence, } g_2(2) = u_1.$$

$$g_2(3) = \text{argmin}\Big\{ f(3,u_1) + g^*_{31}(u_1)V_{Q^*}(g_1,1) + g^*_{32}(u_1)V_{Q^*}(g_1,2) + g^*_{33}(u_1)V_{Q^*}(g_1,3), f(3,u_2) + g^*_{31}(u_2)V_{Q^*}(g_1,1) + g^*_{32}(u_2)V_{Q^*}(g_1,2) + g^*_{33}(u_2)V_{Q^*}(g_1,3)\Big\}$$
$$= \text{argmin}\Big\{4.864, 1.999\Big\} = \{2\}. \qquad \text{Hence, } g_2(3) = u_2.$$

4. Because, $g_2 = g_1$, then $g^* = g_1$ is an optimal control law with $J_{Q^*}(1) = J_{Q^*}(2) = J_{Q^*}(3) = 1$, $V_{Q^*}(1) = 1.611$, $V_{Q^*}(2) = 2.111$ and $V_{Q^*}(3) = 1$.

5.5. Summary

In this chapter, we examined the optimality of stochastic control strategies via dynamic programming on an infinite horizon, when the ambiguity class is described by the total variation distance between the conditional distribution of the controlled process and the nominal conditional distribution. For optimality criteria we considered both the expected discounted reward and the average pay-off per unit time. For the infinite horizon case with a discounted pay-off we showed that the operator associated with the resulting dynamic programming equation under total variation distance ambiguity is contractive, and we introduced a new policy iteration algorithm. For the infinite horizon case with an average pay-off, under the assumption that for every stationary Markov control law the maximizing stochastic matrix is irreducible, we derived a new dynamic programming equation and a new policy iteration algorithm. However, due to the water-filling behavior of the maximizing conditional distribution, it turns out that our proposed method of solution is limited only to a specific range of values of total variation distance, and hence, we derived a general dynamic programming equation by introducing a pair of dynamic programming equations, and, consequently a new policy iteration algorithm, which despite the fact that it is more complex it completely solves the minimax stochastic control problem. Finally, the application of our recommended policy iteration algorithms is shown via illustrative examples.

6

Approximation of Markov Processes by Lower Dimensional Processes

In this chapter, we approximate a finite-state Markov process by another process with fewer states, called herein the approximating process. The approximation problem is formulated using two different methods. The first method, utilizes the total variation distance to discriminate the transition probabilities of a high-dimensional Markov process and a reduced order Markov process. The second method, utilizes total variation distance as a measure of discriminating the invariant probability of a Markov process by the approximating process. Once the reduced invariant probability is obtained, which does not correspond to a Markov process, a further approximation by a Markov process is proposed which minimizes the Kullback-Leibler divergence. The results of this part include:

- based on the first method, a direct procedure for Markov by Markov approximation;
- and, based on the second method, extremum measures which exhibit a water-filling behavior and solve the approximation problems;
- optimal partition functions and new iterative approximation algorithms which compute the invariant distribution of the approximating process;
- examples which illustrate the methodology and the behavior of the approximations.

6.1. Problem Formulation

In this section, the approximation problems under investigation are introduced.

6.1.1. Preliminaries and discrepancy measures

We consider a discrete-time homogeneous Markov process $\{X_t : t = 0, 1, \ldots\}$, with state-space $\mathcal{X}$ of finite cardinality $card(\mathcal{X}) = |\mathcal{X}|$, and transition probability matrix P with elements $\{p_{ij} : i, j = 1, \ldots, |\mathcal{X}|\}$ defined by

$$p_{ij} \triangleq \mathbb{P}(X_{t+1} = j | X_t = i), \quad i, j \in \mathcal{X}, \quad t = 0, 1, \ldots.$$

The Markov process is assumed to be irreducible, aperiodic having a unique invariant distribution $\mu = [\mu_1 \; \mu_2 \ldots \mu_{|\mathcal{X}|}]$ satisfying

$$\mu = \mu P.$$

For the rest of the chapter we adopt the notation $(\mu, P, \mathcal{X})$ to denote a stationary FSM process P with stationary distribution μ and state-space $\mathcal{X}$.

The distance metrics we will use to define the discrepancy between two probability distributions (and conditional probability distributions) are the Total Variation distance, and the Kullback-Leibler divergence. The latter is introduced below (see also Section 2.3).

Relative Entropy distance

The relative entropy of $\nu \in \mathbb{P}(\mathcal{X})$ with respect to $\mu \in \mathbb{P}(\mathcal{X})$ is a mapping $\mathbb{D}(\cdot||\cdot) : \mathbb{P}(\mathcal{X}) \times \mathbb{P}(\mathcal{X}) \longrightarrow [0, \infty]$ defined by

$$\mathbb{D}(\nu||\mu) \triangleq \sum_{i \in \mathcal{X}} \nu_i \log \frac{\nu_i}{\mu_i}$$

It is well known that $\mathbb{D}(\nu||\mu) \geq 0, \forall \nu, \mu \in \mathbb{P}_1(\mathcal{X})$, while $\mathbb{D}(\nu||\mu) = 0 \Leftrightarrow \nu = \mu$.

Let $(\mu, P, \mathcal{X})$ and $(\nu, \Phi, \mathcal{X})$ be two stationary FSM processes. A version of the KL divergence used in [17, 49], is defined by

$$\mathbb{D}_\mu(P||\Phi) \triangleq \sum_{i,j \in \mathcal{X}} \mu_i P_{ij} \log \left(\frac{P_{ij}}{\Phi_{ij}} \right), \tag{6.1}$$

where $P_{i\bullet}$ is assumed to be absolutely continuous with respect to $\Phi_{i\bullet}$, that is, for any $i \in \mathcal{X}$, $\Phi_{ij} = 0$ for some $j \in \mathcal{X}$ then $P_{ij} = 0$. Note that (6.1) is used to compare stationary Markov processes which are defined on the same state-space. For Markov processes which are defined on different state-spaces, (6.1) is modified by introducing a lifted version of the lower-dimensional Markov process (see [19]), defined by

$$\widehat{\Phi}_{ij} = \frac{\mu_j}{\sum_{k\in\psi(j)}\mu_k}\Phi_{\varphi(i)\varphi(j)}, \qquad i,j \in \mathcal{X}, \tag{6.2}$$

where $\psi(j)$ denotes the set of states belonging to the same group as the jth state, and φ denotes a partition function from $\mathcal{X}$ onto $\mathcal{Y}$. For the rest of the chapter we will use the notation $\mathbb{D}^{(\varphi)}(P||\Phi) = \mathbb{D}_\mu(P||\widehat{\Phi})$ to denote the KL divergence distance between two Markov processes defined on different state-spaces.

6.1.2. Approximation problems

The two different methods proposed to approximate FSM processes by lower-dimensional processes, are the following.

Method 1

This method is based on comparing two FSM processes $(\mu, P, \mathcal{X})$ and $(\nu, \Phi, \mathcal{Y})$, $\mathcal{Y} \subset \mathcal{X}$, by working directly on their transition probability matrices P and Φ. The approximation problem is formulated as a maximization of a linear functional, defined on the transition probabilities of the reduced order FSM process $(\nu, \Phi, \mathcal{Y})$, subject to a TV constraint distance between the transition probabilities of the high and low-dimensional FSM processes. The precise problem formulation is given below.

Problem 6.1. *Given a FSM process $(\mu, P, \mathcal{X})$, find a transition probability matrix Φ which solves the maximization problem defined by*

$$\max_{\Phi_{i\bullet}\in\mathbb{P}(\mathcal{Y}),\forall i\in\mathcal{Y}} \quad \sum_{i\in\mathcal{X}}\sum_{j\in\mathcal{X}} \ell_j \Phi_{ij}\mu_i, \quad \ell \in \mathbb{R}_+^{|\mathcal{X}|}$$

$$\textit{subject to} \quad \sum_{i\in\mathcal{X}}\sum_{j\in\mathcal{X}} |\Phi_{ij} - P_{ij}|\mu_i \le R, \quad \forall R \in [0,2].$$

The optimal transition probability matrix Φ which solves optimization Problem 6.1 is obtained for all values of TV parameter $R \in [0,2]$, and exhibits a *water-filling* solution. In addition, as the TV parameter increases, it turns out that the dimension of the transition matrix Φ is reduced, and hence, a reduced order FSM process is obtained.

Method 2

Given a FSM process $(\mu, P, \mathcal{X})$, an invariant distribution $\mu \in \mathbb{P}(\mathcal{X})$, and a parameter $R \in [0,2]$, define the average pay-off with respect to the stationary distribution $\nu \in \mathbf{B}_R(\mu) \subset$

$\mathbb{P}(\mathcal{X})$ by

$$\mathbb{L}(\nu) = \sum_{i \in \mathcal{X}} \ell_i \nu_i, \quad \ell \in \mathbb{R}_+^{|\mathcal{X}|} \tag{6.3}$$

The objective is to approximate $\mu \in \mathbb{P}(\mathcal{X})$ by $\nu \in \mathbf{B}_R(\mu)$, by solving the maximization problem defined by

$$\mathbb{L}(\nu^*) = \max_{\substack{\nu \in \mathbf{B}_R(\mu) \\ \mu = \mu P}} \mathbb{L}(\nu), \quad \forall R \in [0, 2], \tag{6.4}$$

for two alternative choices of the parameters $\ell \in \mathbb{R}_+^{|\mathcal{X}|}$, as follows.

Formulation (a) (*Approximation Based on Occupancy Distribution*)
Let $\ell_i \triangleq \mu_i, \forall i \in \mathcal{X}$, which implies (6.4) is equivalent to maximizing the stationary distribution $\{\nu_i : i \in \mathcal{X}\} \in \mathbb{P}(\mathcal{X})$ subject to the approximation constraint. This formulation leads to an approximation algorithm described via reduction of the states (i.e., by deleting certain states of the original Markov process) to obtain the approximating reduced state process. Intuitively, the optimal solution has the property of maintaining and strengthening the states with the highest invariant probability, while removing the states with the smallest invariant probability, as shown in Fig. 6.1.

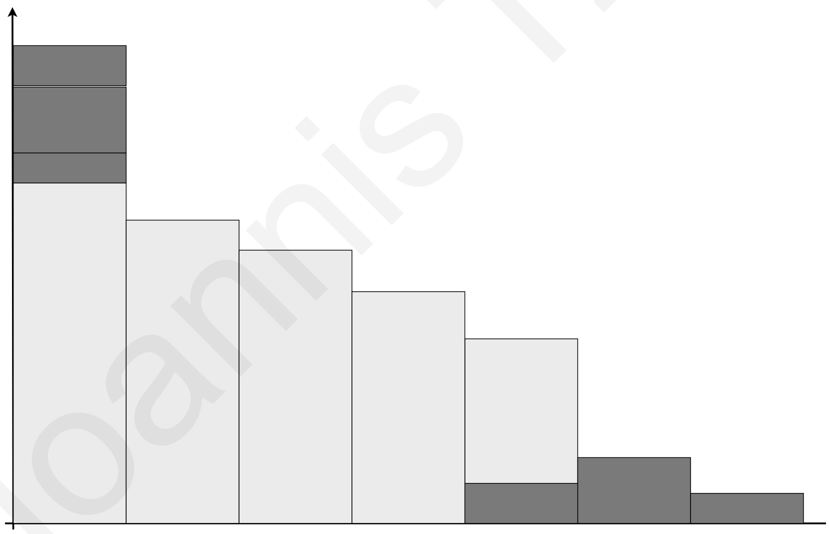


Figure 6.1.: *Water-filling* behavior of invariant distribution based on occupancy distribution

Formulation (b) (*Approximation Based on Maximum Entropy Principle*)
Let $\ell_i \triangleq -\log \nu_i, \forall i \in \mathcal{X}$, which implies that (6.4) equivalent to the problem of finding the approximating distribution corresponding to the minimum description codeword length [4]. This formulation leads to an optimal approximation algorithm described via aggregation of the states (i.e., by grouping certain states of the original Markov process) to obtain the approximated reduced state process, as shown in Fig. 6.2, which is a Hidden Markov process.

This formulation is equivalent to finding the minimum description length of the approximating process, and it is related to minimizing the average codeword length of the approximated Markov process, subject to a fidelity set.

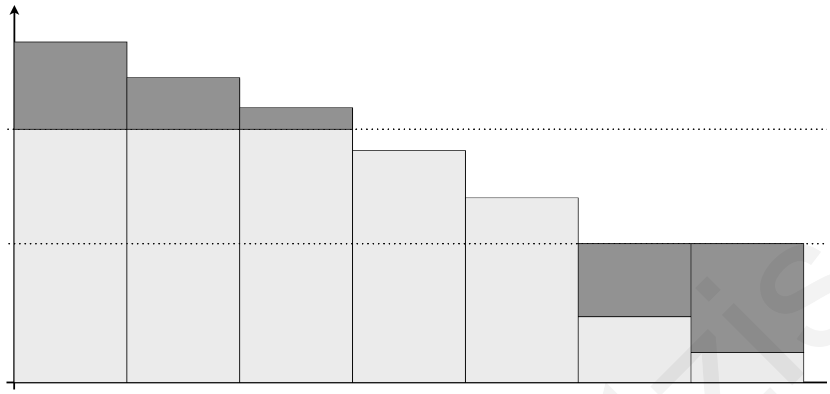


Figure 6.2.: *Water-filling like* behavior of invariant distribution based on maximum entropy principle

The approximated probability vector is based on the following concept. Given a FSM process $(\mu, P, \mathcal{X})$, the optimal probabilities of the reduced process are defined on $\mathcal{X}$, which is partitioned into disjoint sets $\mathcal{X} = \cup_{i=1}^{K} \mathcal{X}_i$, $K \leq |\mathcal{X}|$. The solution of the optimization problems based on Method 2(a) and 2(b), give the maximizing probability $\nu^*(\mathcal{X}_i)$, $i = 1, \ldots, K$, on this partition.

For Method $2(a)$, as R increases the maximizing probability vector, ν^*, is given by a *water-filling* solution, having the property that states of the initial probability vector $\mu \in \mathbb{P}(\mathcal{X})$ are deleted to form a new partition of $\mathcal{X}$, denoted by $\mathcal{X} = \cup_{i=1}^{M} \mathcal{Y}_i$, $M \leq K \leq |\mathcal{X}|$. The approximated probability vector is defined as follows.

Definition 6.1. *(Approximated Probability Vector based on Occupancy Distribution) Define the restriction of ν^* on only those elements of the partition $\{\mathcal{Y}_1, \ldots, \mathcal{Y}_M\}$ which have non-zero probability by*

$$\nu^*|_{\text{supp}(\nu^*)\neq 0} : \{\mathcal{Y}_{i_1}, \mathcal{Y}_{i_2}, \ldots, \mathcal{Y}_{i_k}\} \longmapsto [0, 1], \tag{6.5}$$

where $\{\mathcal{Y}_{i_1}, \mathcal{Y}_{i_2}, \ldots, \mathcal{Y}_{i_k}\} \subseteq \{\mathcal{Y}_1, \ldots, \mathcal{Y}_M\}$, and $i_1, i_2, \ldots, i_k \in \{1, 2, \ldots, M\}$. The approximated probability vector based on occupancy distribution is defined by

$$\bar{\nu} = \nu^*|_{\text{supp}(\nu^*)\neq 0}, \tag{6.6}$$

having states which are in one-to-one correspondence with $\{1, 2, \ldots, k\}$, via the mapping $\mathcal{Y}_{i_1} \longmapsto 1$, $\mathcal{Y}_{i_2} \longmapsto 2, \ldots, \mathcal{Y}_{i_k} \longmapsto k$, with corresponding process $\{Y_t : t = 0, 1, \ldots\}$ having state-space $\mathcal{Y} = \{1, 2, \ldots, k\}$.

For Method $2(b)$, as R increases the maximizing probability vector ν^*, exhibits a *water-filling like* solution, with the property that states of $\mu \in \mathbb{P}(\mathcal{X})$ are aggregated together to form a new partition of $\mathcal{X}$. The approximated probability vector is defined as follows.

Definition 6.2. *(Approximated Probability Vector based on Maximum Entropy Principle) Define $\bar{\nu} = \nu^*$ if all elements of $\nu^*(\mathcal{X}_k)$ are not equal and the state-space of $\bar{\nu}$ is $\mathcal{Y} = \{1, \ldots, K\}$. If any of the $\nu^*(\mathcal{X}_k)$, $k \in \{1, \ldots, K\}$ become equal then a new probability vector $\bar{\nu}$ is defined by adding together those $\nu^* \in \mathbb{P}(\mathcal{X})$ which are equal, and $\bar{\nu} \triangleq \nu^*(\mathcal{X}_k)$ for the $\nu^*(\mathcal{X}_k)$ whose elements are not equal. The resulting approximated probability vector based on maximum entropy principle $\bar{\nu} \in \mathbb{P}(\mathcal{Y})$, with corresponding process $\{Y_t : t = 0, 1, \ldots\}$, is defined on a state-space $\mathcal{Y}$, whose cardinality is less or equal to $|\mathcal{X}|$.*

Remark 6.1. *The reduction based on Method $2(a), (b)$, in general does not produce a Markov chain though in specific cases it may be a Markov chain.*

Once the approximating process[1] $\{Y_t : t = 0, 1, \ldots\}$ is obtained, we move one step further to investigate the problem of approximating a FSM process by another FSM process $(\bar{\nu}, \Phi, \mathcal{Y})$, $\mathcal{Y} \subset \mathcal{X}$. Here, the objective is to find an optimal partition function φ and a transition matrix Φ which minimizes the KL divergence rate [19] defined by

$$\mathbb{D}^{(\varphi)}(P||\Phi) = \sum_{i,j \in \mathcal{X}} \mu_i P_{ij} \log\left(\frac{P_{ij}}{\widehat{\Phi}_{ij}}\right), \tag{6.7}$$

where $\widehat{\Phi}$ is given by (6.2), and denotes the lifted version of the lower-dimensional Markov chain Φ by using an optimal partition function φ. By employing certain results from [19], the transition matrix Φ which solves (6.7) is obtained. What remains, is to find an optimal partition function φ, for the approximation problems of Method $2(a)$ and $2(b)$. This Markov by Markov approximation is found by working only with values of TV parameter for which a reduction of the states occurs, that is, $|\mathcal{Y}| < |\mathcal{X}|$.

Given a FSM process $(\mu, P, \mathcal{X})$, an algorithm is presented, which describes how to construct the transition probability matrix $Q^\dagger$, from the maximizing distribution ν^* of problem (6.4) for Method $2(a)$ and $2(b)$. Then, using Definitions 6.1 and 6.2, a lower probability distribution $\bar{\nu} \in \mathbb{P}(\mathcal{Y})$ is obtained. Under the restriction that the lower-dimensional process is also a FSM process $(\bar{\nu}, \Phi, \mathcal{Y})$, $\mathcal{Y} \subset \mathcal{X}$, an optimal partition function φ and a transition probability matrix Φ, are found which minimize the KL divergence rate between P and $\widehat{\Phi}$. The approximation procedure for Method $2(a)$ and $2(b)$, is shown in Fig.6.3.

The precise problem definition of approximation Method 2 based on occupancy distribution is given below.

[1] The reduced approximating process is obtained without a priori imposing the assumption that it is also a Markov process.

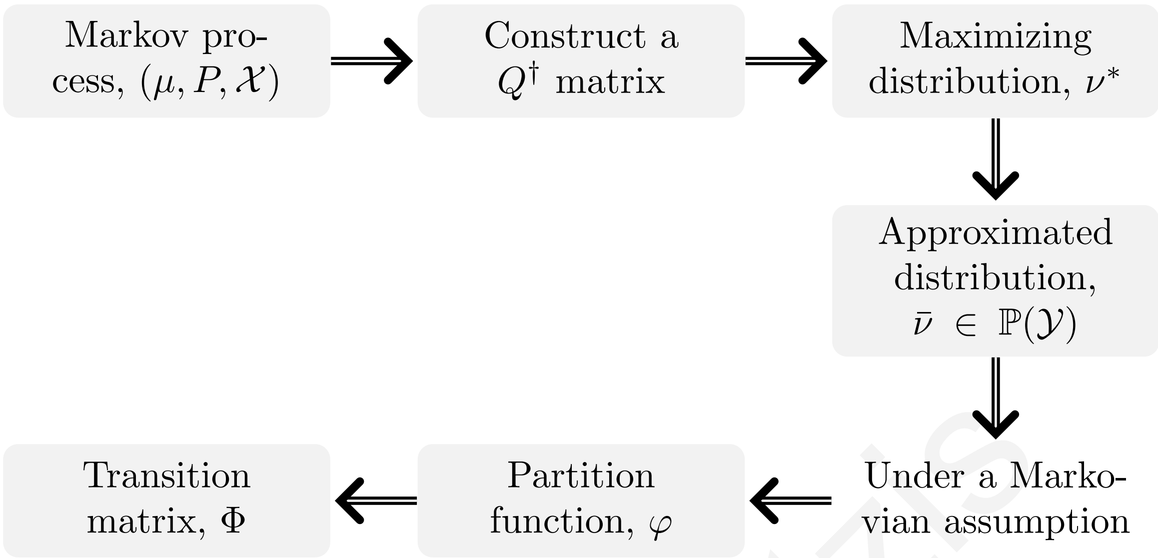


Figure 6.3.: Procedure of *Method 2*.

Problem 6.2. *(Approximation Based on Occupancy Distribution)*
Let $\{\ell_i : i \in \mathcal{X}\} \in \mathbb{R}_+^{|\mathcal{X}|}$ *denote the occupancy distribution of a FSM process* $(\mu, P, \mathcal{X})$ *defined by* $\ell_i \triangleq \mu_i, \forall i \in \mathcal{X}$. *Find* $\{\nu_i : i \in \mathcal{X}\} \in \mathbb{P}(\mathcal{X})$ *which solves*

$$\max_{\substack{\nu \in \mathbb{B}_R(\mu) \\ \mu = \mu P}} \sum_{i \in \mathcal{X}} \mu_i \nu_i. \tag{6.8}$$

Given the optimal solution of (6.8), *let* $\bar{\nu}$ *of Definition 6.1 denote the invariant distribution of a lower-dimensional FSM process* $(\bar{\nu}, \Phi, \mathcal{Y})$, $\mathcal{Y} \subset \mathcal{X}$.
Find an optimal partition function φ, *and calculate the transition probability matrix* Φ, *which satisfies* $\bar{\nu} = \bar{\nu}\Phi$, *and minimizes the KL divergence rate defined by*

$$\min_{\substack{\varphi, \Phi \\ \bar{\nu} = \bar{\nu}\Phi}} \mathbb{D}^{(\varphi)}(P||\Phi). \tag{6.9}$$

Other reasonable choices, are possible by letting $\ell \in \mathbb{R}_+^{|\mathcal{X}|}$ correspond to a reward or a profit, a cost or a loss, etc., whenever a node is visited.

Next, the precise problem definition of approximation Method 2 based on maximum entropy principle is given.

Problem 6.3. *(Approximation Based on Maximum Entropy Principle)*
Maximize the entropy of $\{\nu_i : i \in \mathcal{X}\} \in \mathbb{P}(\mathcal{X})$ *subject to total variation fidelity set, defined by*

$$\max_{\substack{\nu \in \mathbb{B}_R(\mu) \\ \mu = \mu P}} H(\nu), \quad H(\nu) \triangleq -\sum_{i \in \mathcal{X}} \log(\nu_i)\nu_i. \tag{6.10}$$

Given the optimal solution of (6.10)*, let* $\bar{\nu}$ *of Definition* 6.2 *denote the invariant distribution of a lower-dimensional Markov process* $(\bar{\nu}, \Phi, \mathcal{Y})$*,* $\mathcal{Y} \subset \mathcal{X}$*.*
Find an optimal partition function φ*, and calculate the transition probability matrix* Φ*, which satisfies* $\bar{\nu} = \bar{\nu}\Phi$*, and minimizes the KL divergence rate defined by*

$$\min_{\substack{\varphi,\Phi \\ \bar{\nu}=\bar{\nu}\Phi}} \mathbb{D}^{(\varphi)}(P||\Phi). \tag{6.11}$$

Problem (6.10) is of interest when the concept of insufficient reasoning (e.g., Jayne's maximum entropy principle[2] [34]) is applied to construct a model for $\nu \in \mathbb{P}(\mathcal{X})$, subject to information quantified via the fidelity set defined by the variation distance between ν and μ.

It is not difficult to show that the maximum entropy approximation defined by (6.10) is precisely equivalent to the problem of finding the approximating distribution corresponding to the minimum description codeword length, also known as the universal coding problem [4, 52], as follows. Let $\{\ell_i : i \in \mathcal{X}\} \in \mathbb{R}_+^{|\mathcal{X}|}$ denote the positive codeword lengths corresponding to each symbol of the approximated distribution, which satisfy the Kraft inequality of lossless Shannon codes $\sum_{i\in\mathcal{X}} D^{-\ell_i} \leq 1$, where the codeword alphabet is D-ary (unless specified otherwise $\log(\cdot) \triangleq \log_D(\cdot)$). Then, by the Von-Neumann's theorem, which holds due to compactness and convexity of the constraints, it follows that

$$\min_{\ell\in\mathbb{R}_+^{|\mathcal{X}|}:\sum_{i\in\mathcal{X}} D^{-\ell_i}\leq 1} \max_{\substack{\nu\in\mathbb{B}_R(\mu) \\ \mu=\mu P}} \sum_{i\in\mathcal{X}} \ell_i\nu_i = \max_{\substack{\nu\in\mathbb{B}_R(\mu) \\ \mu=\mu P}} \min_{\ell\in\mathbb{R}_+^{|\mathcal{X}|}:\sum_{i\in\mathcal{X}} D^{-\ell_i}\leq 1} \sum_{i\in\mathcal{X}} \ell_i\nu_i = \max_{\substack{\nu\in\mathbb{B}_R(\mu) \\ \mu=\mu P}} H(\nu).$$

Hence, for $\ell_i \triangleq -\log \nu_i$, $\forall i \in \mathcal{X}$, the optimization (6.4) is equivalent to optimization (6.10).

6.2. Method 1: Solution of Approximation Problem

In this section, we give the main theorem which characterizes the solution of Problem 6.1.

Similar to Chapter 3, define the maximum and minimum values of the sequence $\{\ell_1, \ldots, \ell_{|\mathcal{X}|}\} \in \mathbb{R}_+^{|\mathcal{X}|}$ by

$$\ell_{\max} \triangleq \max_{i\in\mathcal{X}} \ell_i, \qquad \ell_{\min} \triangleq \min_{i\in\mathcal{X}} \ell_i$$

and its corresponding support sets by

$$\mathcal{X}^0 \triangleq \{i \in \mathcal{X} : \ell_i = \ell_{\max}\}, \qquad \mathcal{X}_0 \triangleq \{i \in \mathcal{X} : \ell_i = \ell_{\min}\}.$$

[2] The maximum entropy principle states that, subject to precisely stated prior data, the probability distribution which best represents the current state of knowledge is the one with largest entropy.

For all remaining elements of the sequence, $\{\ell_i : i \in \mathcal{X} \setminus \mathcal{X}^0 \cup \mathcal{X}_0\}$, define recursively the set of indices for which ℓ achieves its $(k+1)$th smallest value by $\mathcal{X}_k$, where $k \in \{1, 2, \dots, |\mathcal{X} \setminus \mathcal{X}^0 \cup \mathcal{X}_0|\}$, till all the elements of $\mathcal{X}$ are exhausted (i.e., k is at most $|\mathcal{X} \setminus \mathcal{X}^0 \cup \mathcal{X}_0|$), and the corresponding values of the sequence on the $\mathcal{X}_k$ sets by $\ell(\mathcal{X}_k)$.

For a fixed $i \in \mathcal{X}$, define the total variation of a signed measure

$$\Xi_{ij} \triangleq \Phi_{ij} - P_{ij}, \quad \forall j \in \mathcal{X}$$

to be equal to the summation of its positive and its negative part, that is,

$$||\Xi_{i\bullet}||_{TV} \triangleq \sum_{j\in\mathcal{X}} \Xi_{ij}^+ + \sum_{j\in\mathcal{X}} \Xi_{ij}^-, \qquad \forall i \in \mathcal{X}. \tag{6.12}$$

By utilizing the fact that $\sum_{j\in\mathcal{X}} \Xi_{ij} = 0$, $\forall i \in \mathcal{X}$ then

$$\sum_{j\in\mathcal{X}} \Xi_{ij}^+ = \sum_{j\in\mathcal{X}} \Xi_{ij}^- = \frac{||\Xi_{i\bullet}||_{TV}}{2}, \qquad \forall i \in \mathcal{X}. \tag{6.13}$$

Let $\alpha_i \triangleq ||\Xi_{i\bullet}||_{TV}$, $\forall i \in \mathcal{X}$, then the constraint of Problem 6.1 is equivalent to

$$\sum_{i\in\mathcal{X}} \alpha_i \mu_i \leq R. \tag{6.14}$$

and the pay-off can be reformulated as follows.

$$\max_{\Phi_{i\bullet}\in\mathbb{P}(\cdot)} \sum_{i\in\mathcal{X}} \sum_{j\in\mathcal{X}} \ell_j \Phi_{ij} \mu_i \equiv \sum_{i\in\mathcal{X}} \sum_{j\in\mathcal{X}} \ell_j P_{ij} \mu_i + \max_{\Phi_{i\bullet}\in\mathbb{P}(\cdot)} \sum_{i\in\mathcal{X}} \sum_{j\in\mathcal{X}} \ell_j \Xi_{ij} \mu_i. \tag{6.15}$$

In addition,

$$\sum_{i\in\mathcal{X}} \sum_{j\in\mathcal{X}} \ell_j \Xi_{ij} \mu_i = \sum_{i\in\mathcal{X}} \sum_{j\in\mathcal{X}} \ell_j \Xi_{ij}^+ \mu_i - \sum_{i\in\mathcal{X}} \sum_{j\in\mathcal{X}} \ell_j \Xi_{ij}^- \mu_i. \tag{6.16}$$

The solution of Problem 6.1 is obtained by identifying the partition of $\mathcal{X}$ into disjoint sets $\{\mathcal{X}^0, \mathcal{X}_0, \mathcal{X}_1, \dots, \mathcal{X}_k\}$ and the transitions on this partition. The main idea is to express $\Xi_{i\bullet}$ as the difference of its positive and negative part and then find upper and lower bounds on the transition probabilities of $\mathcal{X}^0$ and $\mathcal{X} \setminus \mathcal{X}^0$ which are achievable. Closed form expressions of the transition probability measures, on these sets, which achieve the bounds are derived.

The following Theorem characterizes the solution of Problem 6.1.

Theorem 6.1. *The solution of Problem 6.1 is given by*

$$\sum_{i\in\mathcal{X}} \sum_{j\in\mathcal{X}} \ell_j \Phi_{ij}^\dagger \mu_i = \ell_{\max} \sum_{i\in\mathcal{X}^0} \sum_{j\in\mathcal{X}} \mu_j \Phi_{ji}^\dagger + \ell_{\min} \sum_{i\in\mathcal{X}_0} \sum_{j\in\mathcal{X}} \mu_j \Phi_{ji}^\dagger + \sum_{k=1}^{r} \ell(\mathcal{X}_k) \sum_{i\in\mathcal{X}_k} \sum_{j\in\mathcal{X}} \mu_j \Phi_{ji}^\dagger, \tag{6.17}$$

where for any $i \in \mathcal{X}$,

$$\Phi^{\dagger}_{ij} = P_{ij} + \frac{\alpha_i}{2|\mathcal{X}^0|}, \quad \forall j \in \mathcal{X}^0, \tag{6.18a}$$

$$\Phi^{\dagger}_{ij} = \left(P_{ij} - \frac{\alpha_i}{2|\mathcal{X}_0|}\right)^+, \quad \forall j \in \mathcal{X}_0, \tag{6.18b}$$

$$\Phi^{\dagger}_{ij} = \left(P_{ij} - \left(\frac{\alpha_i}{2|\mathcal{X}_k|} - \sum_{j=1}^{k}\sum_{z \in \mathcal{X}_{j-1}} P_{iz}\right)^+\right)^+, \forall j \in \mathcal{X}_k \tag{6.18c}$$

$$\alpha_i = \min(R, R_{\max,i}), \quad R_{\max,i} = 2(1 - \sum_{j \in \mathcal{X}^0} P_{ij}), \tag{6.18d}$$

$k = 1, 2, \ldots, r$ and r is the number of $\mathcal{X}_k$ sets which is at most $|\mathcal{X} \setminus \mathcal{X}^0 \cup \mathcal{X}_0|$. Once the $\Phi^{\dagger}$ matrix is constructed as a function of TV parameter R, then the transition matrix Φ which solves Problem 6.1 is given by removing all zero columns and the respective rows of $\Phi^{\dagger}$ matrix.

Proof. See Appendix E. ■

Clearly, the optimal transition matrix Φ is obtained via a *water-filling* solution.

Remark 6.2. *Note that, if we replace the maximization of Problem 6.1 with minimization, then the solution of the new problem is obtained precisely as that of Problem 6.1, but with a reverse computation of the partition of the space $\mathcal{X}$ and the mass of the transition probability on the partition moving in the opposite direction.*

6.3. Method 2: Solution of Approximation Problems

In this section, we recall some results from Chapter 3, which are vital in providing the solution of Problem (6.4), and consequently the solution of approximation Problems (6.2) and (6.3).

First recall, from Section 6.2, the definitions of the support sets $\mathcal{X}^0$, $\mathcal{X}_0$, $\mathcal{X}_k$ and the definitions of the corresponding values of the sequence on these sets given by $\ell_{\max}$, $\ell_{\min}$ and $\ell(\mathcal{X}_k)$. Given $\ell \in \mathbb{R}_+^{|\mathcal{X}|}$, $\mu \in \mathbb{P}(\mathcal{X})$, it is shown in Chapter 3, that the solution of optimization (6.4) is given by[3]

$$\mathbb{L}(\nu^*) = \ell_{\max}\nu^*(\mathcal{X}^0) + \ell_{\min}\nu^*(\mathcal{X}_0) + \sum_{k=1}^{r} \ell(\mathcal{X}_k)\nu^*(\mathcal{X}_k), \tag{6.19}$$

[3]Note the notation Σ^0, Σ_0 and Σ_k in Chapter 3, is identical to the notation $\mathcal{X}^0$, $\mathcal{X}_0$ and $\mathcal{X}_k$, respectively.

and the optimal probabilities are obtained via *water-filling*, as follows

$$\nu^*(\mathcal{X}^0)\triangleq \sum_{i\in\mathcal{X}^0} \nu_i^* = \sum_{i\in\mathcal{X}^0} \mu_i + \frac{\alpha}{2}, \tag{6.20a}$$

$$\nu^*(\mathcal{X}_0)\triangleq \sum_{i\in\mathcal{X}_0} \nu_i^* = \Big(\sum_{i\in\mathcal{X}_0} \mu_i - \frac{\alpha}{2}\Big)^+, \tag{6.20b}$$

$$\nu^*(\mathcal{X}_k)\triangleq \sum_{i\in\mathcal{X}_k} \nu_i^* = \Big(\sum_{i\in\mathcal{X}_k} \mu_i - \Big(\frac{\alpha}{2} - \sum_{j=1}^{k} \sum_{i\in\mathcal{X}_{j-1}} \mu_i\Big)^+\Big)^+, \tag{6.20c}$$

$$\alpha = \min\left(R, R_{\max}\right), \qquad R_{\max} \triangleq 2(1 - \sum_{i\in\mathcal{X}^0} \mu_i), \tag{6.20d}$$

where, $k = 1, 2, \ldots, r$ and r is the number of $\mathcal{X}_k$ sets which is at most $|\mathcal{X} \setminus \mathcal{X}^0 \cup \mathcal{X}_0|$. The optimal probabilities given by (6.20a)-(6.20c), can be expressed in matrix form as follows

$$\nu^* = \mu Q^\dagger = \mu P Q^\dagger. \tag{6.21}$$

In Sections 6.3.1 and 6.3.2, we provide algorithms for constructing the desired $Q^\dagger$ matrix for the optimizations (6.8) and (6.10), respectively.

Remark 6.3. *The identification of the support sets $\mathcal{X}^0$, $\mathcal{X}_0$ and $\mathcal{X}_k$, $k = 1, 2, \ldots, r$, is based on the values of ℓ_i's, $\forall i \in \mathcal{X}$. If the cardinality of any of the support sets is greater than one, i.e., $|\mathcal{X}^0| > 1$, and $\ell_i = \ell_{i+1} = \ldots, \forall i, i+1, \dots \in \mathcal{X}^0$ then by* (6.20a)

$$\nu_i^* = \frac{\nu^*(\mathcal{X}^0)}{|\mathcal{X}^0|}, \quad \forall i \in \mathcal{X}^0, \tag{6.22a}$$

and similarly for the rest, that is, if $|\mathcal{X}_0| > 1$ then

$$\nu_i^* = \frac{\nu^*(\mathcal{X}_0)}{|\mathcal{X}_0|}, \quad \forall i \in \mathcal{X}_0, \tag{6.22b}$$

and if $|\mathcal{X}_k| > 1$, for $k = 1, \ldots, r$, then

$$\nu_i^* = \frac{\nu^*(\mathcal{X}_k)}{|\mathcal{X}_k|}, \quad \forall i \in \mathcal{X}_k. \tag{6.22c}$$

The resulting optimal probability ν^* is a $(2 + r)$ row vector and hence, by (6.21) *$Q^\dagger$ is an $|\mathcal{X}| \times (2 + r)$ matrix. Then by employing* (6.22) *we extract the optimal probabilities ν_i^* for all $i \in \mathcal{X}$, which are then used in definition of the optimal partition functions (see Definition 6.3 and 6.4).*

For the approximation based on occupancy distribution, we let the matrix $Q^\dagger$ to be an $|\mathcal{X}| \times |\mathcal{X}|$ matrix, instead of an $|\mathcal{X}| \times (2 + r)$ matrix. The reason for doing so, is that we want to take into account the cases for which ℓ_i's, $\forall i \in \mathcal{X}$, might be defined to represent a cost or profit etc., whenever a node is visited. In such cases, (6.22) *is not valid anymore,* *since $\ell_i = \ell_j$ does not necessarily imply $\mu_i = \mu_j$, $\forall i, j \in \mathcal{X}$. As we will show in Section 6.3.1, Algorithm 6.2 constructs a $Q^\dagger$ matrix which in addition to occupancy distribution, considers those alternative cases as well.*

By Definition 6.1 and 6.2, the approximated probability vector $\bar{\nu} \in \mathbb{P}(\mathcal{Y})$ is readily available and satisfies

$$\bar{\nu} = \mu Q = \mu P Q, \tag{6.23}$$

where Q matrix is modified accordingly.

Once the reduced state process is obtained, we utilize its solution to solve the optimizations (6.9) and (6.11). The relation between $\mu(t), \mu(t+1) \in \mathbb{P}(\mathcal{X})$ and $\bar{\nu}(t), \bar{\nu}(t+1) \in \mathbb{P}(\mathcal{Y})$ is shown in Fig.6.4.

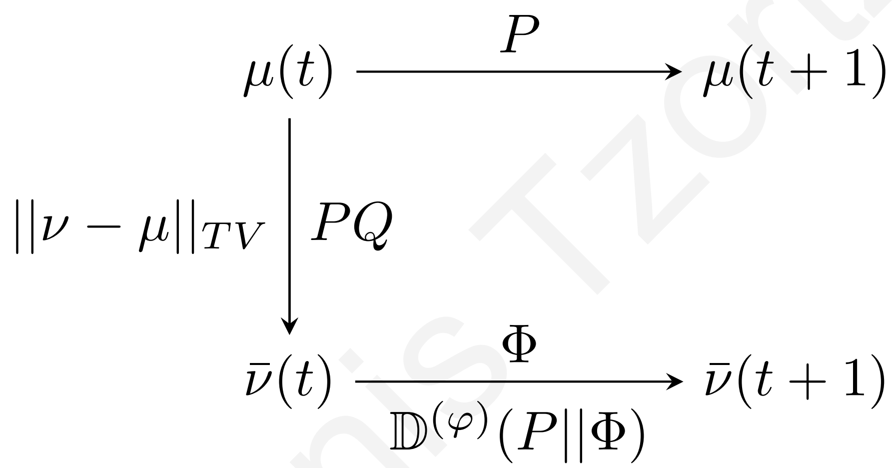


Figure 6.4.: *Method 2*. Diagram that shows the relationship of the initial and the lower probability distributions.

6.3.1. Solution of Approximation Problem based on Occupancy Distribution

In this section, we first give an algorithm to construct the $Q^\dagger$ matrix which solves (6.8). Then, under an additional assumption that the reduced process is also Markov, we give the solution of (6.9).

Let $k = 0, 1, \ldots, r-1$, where r denotes the number of $\mathcal{X}_k$ sets, that is, $1 \le r \le |\mathcal{X} \setminus \mathcal{X}^0|$ (note that, $\mathcal{X}_0$ set is included). For all $j = 1, 2, \ldots, |\mathcal{X}_k|$, $\mathcal{X}_{k,j} \triangleq \{jth$ element of $\mathcal{X}_k$ set$\}$, (note that, if $|\mathcal{X}_k| = 1$ then $\mathcal{X}_{k,j} = \mathcal{X}_k$). Similarly, $\mathcal{X}^{0,j} \triangleq \{jth$ element of $\mathcal{X}^0$ set$\}$, (note that, if $|\mathcal{X}^0| = 1$ then $\mathcal{X}^{0,j} = \mathcal{X}^0$).

Algorithm 6.2.

1. *Initialization step:*
   a) *Arrange* ℓ_i, $i \in \mathcal{X}$, *in a descending order.*
   b) *Identify the support sets* $\mathcal{X}^0$, $\mathcal{X}_0$ *and* $\mathcal{X}_k$ *for all* $k \in \{1, 2, \ldots, |\mathcal{X} \setminus \mathcal{X}^0 \cup \mathcal{X}_0|\}$.
   c) *Calculate the value of* r.

For any $R \in [0, 2]$*:*

2) *Step.1 (Indicator functions):*
   a) *Let*

$$\mu^R(\mathcal{X}^0) \triangleq \sum_{i \in \mathcal{X}^0} \mu_i + \frac{R}{2}.$$

Define

$$I^{\mathcal{X}^0} \triangleq \begin{cases} 1, & \textit{if } \mu^R(\mathcal{X}^0) \geq 1, \\ 0, & \textit{otherwise}. \end{cases} \tag{6.24}$$

   b) *For* $k = 0, 1, \ldots, r-1$ *let*

$$\mu^R(\mathcal{X}_k) \triangleq \sum_{j=0}^{k} \sum_{i \in \mathcal{X}_j} \mu_i - \frac{R}{2}.$$

Define

$$I^{\mathcal{X}_k} \triangleq \begin{cases} 1, & \textit{if } \mu^R(\mathcal{X}_k) \geq 0, \\ 0, & \textit{otherwise}. \end{cases}$$

$$I^{\mathcal{X}_{[0,k-1]}} \triangleq \begin{cases} 1, & \textit{if } \mu^R(\mathcal{X}_i) < 0, \ \forall i = 0, 1, \ldots, k-1 \\ 0, & \textit{otherwise}, \end{cases}$$

and

$$I^{\mathcal{X}_k, \mathcal{X}_{[0,k-1]}} = I^{\mathcal{X}_k} I^{\mathcal{X}_{[0,k-1]}}. \tag{6.25}$$

   c) *For* $k = 0, 1, \ldots, r-1$, *if* $|\mathcal{X}_k| > 1$, *then for all* j=$1 \ldots, |\mathcal{X}_k|$, *let*

$$\mu^R(\mathcal{X}_{k,j}) \triangleq \mu_{\mathcal{X}_{k,j}} - \frac{(R/2 - \sum_{i \in \cup_{j=0}^{k-1} \mathcal{X}_j} \mu_i)}{|\mathcal{X}_k|}.$$

Define

$$I^{\mathcal{X}_{k,j}} \triangleq \begin{cases} 1, & \textit{if } \mu^R(\mathcal{X}_{k,j}) \geq 0, \\ 0, & \textit{otherwise}, \end{cases} \tag{6.26}$$

3) *Step.2 (The $Q^\dagger$ matrix):*

Let $Q^\dagger$ be an $|\mathcal{X}| \times |\mathcal{X}|$ matrix and $i = 1, 2, \dots, |\mathcal{X}|$ to denote the ith column of $Q^\dagger$ matrix.

a) *For all $i \in \mathcal{X}^0$, the elements of the ith column are given as follows.*

i) *Let the $(Q^\dagger)_{i,i}$ element be equal to*

$$\sum_{k=0}^{r-1} I^{\mathcal{X}_k,\mathcal{X}_{[0,k-1]}}\Big(1+\frac{R/2}{|\mathcal{X}^0|}\Big) + I^{\mathcal{X}^0}\frac{\big(\mu_{\mathcal{X}^{0,i}} + \frac{\sum_{j\in\mathcal{X}\setminus\mathcal{X}^0}\mu_j}{|\mathcal{X}^0|}\big)}{\mu_{\mathcal{X}^{0,i}}}. \tag{6.27}$$

ii) *Let all the remaining elements of the ith column be equal to*

$$\sum_{k=0}^{r-1} I^{\mathcal{X}_k,\mathcal{X}_{[0,k-1]}}\frac{R/2}{|\mathcal{X}^0|}. \tag{6.28}$$

b) *For all $i \in \mathcal{X}_k$, $k = 0, 1, \dots, r-1$, and $j \in \Big\{\psi \in \{1, 2, \dots, |\mathcal{X}_k|\} : i \in \mathcal{X}_k$ is in the ψth position on $\mathcal{X}_k$ set$\Big\}$, the elements of the ith column are as follows.*

i) *Let the $(Q^\dagger)_{i,i}$ element be equal to*

$$\sum_{j=0}^{k-1} I^{\mathcal{X}_j,\mathcal{X}_{[0,j-1]}} + I^{\mathcal{X}_{k,j}} I^{\mathcal{X}_k,\mathcal{X}_{[0,k-1]}}\Big(1-\frac{R/2}{\sum_{j=1}^{|\mathcal{X}_k|} I^{\mathcal{X}_{k,j}}}\Big). \tag{6.29}$$

ii) *If $|\mathcal{X}_k| > 1$, then for all $z \in \mathcal{X}_k \setminus \mathcal{X}_{k,j}$, let the $(Q^\dagger)_{z,i}$ element be equal to*

$$I^{\mathcal{X}_{k,j}} I^{\mathcal{X}_k,\mathcal{X}_{[0,k-1]}}\Bigg\{\prod_{j=1}^{|\mathcal{X}_k|} I^{\mathcal{X}_{k,j}}\Big(\frac{-R/2}{\sum_{j=1}^{|\mathcal{X}_k|} I^{\mathcal{X}_{k,j}}}\Big) + \Big(1-\frac{R/2}{\sum_{j=1}^{|\mathcal{X}_k|} I^{\mathcal{X}_{k,j}}}\Big)\Big(1-\prod_{j=1}^{|\mathcal{X}_k|} I^{\mathcal{X}_{k,j}}\Big)\Bigg\}. \tag{6.30}$$

iii) *For all $z \in \mathcal{X} \setminus \mathcal{X}^0 \cup \mathcal{X}_k$ and only if $z > i$ let the $(Q^\dagger)_{z,i}$ element be equal to*

$$I^{\mathcal{X}_{k,j}} I^{\mathcal{X}_k,\mathcal{X}_{[0,k-1]}}\Bigg\{\prod_{j=1}^{|\mathcal{X}_k|} I^{\mathcal{X}_{k,j}}\Big(\frac{1}{|\mathcal{X}_k|}-\frac{R/2}{\sum_{j=1}^{|\mathcal{X}_k|} I^{\mathcal{X}_{k,j}}}\Big)+\Big(1-\frac{R/2}{\sum_{j=1}^{|\mathcal{X}_k|} I^{\mathcal{X}_{k,j}}}\Big)\Big(1-\prod_{j=1}^{|\mathcal{X}_k|} I^{\mathcal{X}_{k,j}}\Big)\Bigg\}. \tag{6.31}$$

iv) *Let all the remaining elements of the ith column be equal to*

$$I^{\mathcal{X}_{k,j}} I^{\mathcal{X}_k,\mathcal{X}_{[0,k-1]}}\Big(\frac{-R/2}{\sum_{j=1}^{|\mathcal{X}_k|} I^{\mathcal{X}_{k,j}}}\Big). \tag{6.32}$$

Once the $Q^\dagger$ matrix is constructed, as a function of the TV parameter R, then by (6.21) the resulting optimal probability, ν^*, is an $1 \times |\mathcal{X}|$ row vector. However, recall from Remark 6.3 that by definition ν^* is just an $1 \times (2 + r)$ row vector. By using all the information that the support sets provide to us we can easily transform the $1\times|\mathcal{X}|$ row vector to an $1\times(2+r)$ row vector, by simply adding together the optimal probabilities, ν_i^*, $\forall i \in \mathcal{X}$, which belong to the same support sets. Given the optimal solution of optimization (6.8), then by Definition 6.1 the lower-dimensional process $\{Y_t : t = 0, 1 \dots\}$ with invariant distribution $\bar{\nu}$ is obtained, either by removing all zero elements of $\nu^* \in \mathbb{P}(\mathcal{X})$, or by defining a Q matrix to be equal to $Q^\dagger$ after the deletion of all zero columns, and hence

$$\bar{\nu} = \mu Q = \mu P Q, \tag{6.33}$$

where the dimensions of Q matrix are based on the value of the TV parameter $R \in [0, 2]$.

Before we proceed with the solution of (6.9), we provide a simple, yet useful example in order to explain each step of Algorithm 6.2.

Example 6.1. *Let $\ell = [\ell_1\ \ell_2\ \ell_3\ \ell_4]$, where $\ell_1 > \ell_2 > \ell_3 > \ell_4$, and $|\mathcal{X}| = 4$. For simplicity it is assumed that the optimum probabilities ν_i^*, $i \in \mathcal{X}$, as a function of R are known, as presented in Fig.6.5.*

Initialization step. The support sets are equal to $\mathcal{X}^0 = \{1\}$, $\mathcal{X}_0 = \{4\}$, $\mathcal{X}_1 = \{3\}$ and $\mathcal{X}_2 = \{2\}$. The number of $\mathcal{X}_k$ sets is equal to $r = 3$.

Step.1 From (6.24), the indicator function $I^{\mathcal{X}^0}$ is given by

$$I^{\mathcal{X}^0} \triangleq \begin{cases} 1, & \text{if } \mu_1 + \frac{R}{2} \geq 1, \\ 0, & \text{otherwise}. \end{cases}$$

From (6.25), the indicator functions $I^{\mathcal{X}_0}$, $I^{\mathcal{X}_1, \mathcal{X}_{[0,0]}}$ and $I^{\mathcal{X}_2, \mathcal{X}_{[0,1]}}$ are given by

$$I^{\mathcal{X}_0} \triangleq \begin{cases} 1, & \text{if } \mu_4 - \frac{R}{2} \geq 0, \\ 0, & \text{otherwise}. \end{cases}$$

$$I^{\mathcal{X}_1, \mathcal{X}_{[0,0]}} \triangleq \begin{cases} 1, & \text{if } \mu_3 + \mu_4 - \frac{R}{2} \geq 0 \text{ and } \mu_4 - \frac{R}{2} \leq 0, \\ 0, & \text{otherwise}. \end{cases}$$

$$I^{\mathcal{X}_2, \mathcal{X}_{[0,1]}} \triangleq \begin{cases} 1, & \text{if } \mu_2 + \mu_3 + \mu_4 - \frac{R}{2} \geq 0 \\ & \text{and } \mu_3 + \mu_4 - \frac{R}{2} \leq 0 \text{ and } \mu_4 - \frac{R}{2} \leq 0, \\ 0, & \text{otherwise}. \end{cases}$$

The values of the indicator functions for $R \in [0, 2]$ are given below.

$0 \leq R < R_1$	$R_1 \leq R < R_2$	$R_2 \leq R < R_3$	$R_3 \leq R \leq 2$
$I^{\mathcal{X}^0} = 0$	$I^{\mathcal{X}^0} = 0$	$I^{\mathcal{X}^0} = 0$	$I^{\mathcal{X}^0} = 1$
$I^{\mathcal{X}_0} = 1$	$I^{\mathcal{X}_0} = 0$	$I^{\mathcal{X}_0} = 0$	$I^{\mathcal{X}_0} = 0$
$I^{\mathcal{X}_1, \mathcal{X}_{[0,0]}} = 0$	$I^{\mathcal{X}_1, \mathcal{X}_{[0,0]}} = 1$	$I^{\mathcal{X}_1, \mathcal{X}_{[0,0]}} = 0$	$I^{\mathcal{X}_1, \mathcal{X}_{[0,0]}} = 0$
$I^{\mathcal{X}_2, \mathcal{X}_{[0,1]}} = 0$	$I^{\mathcal{X}_2, \mathcal{X}_{[0,1]}} = 0$	$I^{\mathcal{X}_2, \mathcal{X}_{[0,1]}} = 1$	$I^{\mathcal{X}_2, \mathcal{X}_{[0,1]}} = 0$

For $0 \leq R < R_1$*, all indicator functions are equal to one, except the one which corresponds to* $\mathcal{X}_0$ *set, that is,* $I^{\mathcal{X}_0} = 1$*. As soon as* $\mu^R(\mathcal{X}_0) = 0$*, then* $I^{\mathcal{X}_0}$ *becomes equal to zero and* $I^{\mathcal{X}_1, \mathcal{X}_{[0,0]}}$ *equal to one. This procedure is repeated until the value of* $R = R_{\max} = R_3$*, see Fig.6.5, in which* $I^{\mathcal{X}^0}$ *becomes equal to one, and all other indicator functions equal to zero, and* $I^{\mathcal{X}^0}$ *remains active for all* $R \geq R_{\max} = R_3$*.*

Step.2 Let $Q^\dagger$ *be an* 4×4 *matrix. For* $0 \leq R < R_1$*,*

$$Q^\dagger = \begin{pmatrix} 1+R/2 & 0 & 0 & -R/2 \\ R/2 & 1 & 0 & -R/2 \\ R/2 & 0 & 1 & -R/2 \\ R/2 & 0 & 0 & 1-R/2 \end{pmatrix},$$

and since no zero column exist then $Q^\dagger = Q$*. For* $R_1 \leq R < R_2$*,*

$$Q^\dagger = \begin{pmatrix} 1+R/2 & 0 & -R/2 & 0 \\ R/2 & 1 & -R/2 & 0 \\ R/2 & 0 & 1-R/2 & 0 \\ R/2 & 0 & 1-R/2 & 0 \end{pmatrix},$$

and hence

$$Q = \begin{pmatrix} 1+R/2 & 0 & -R/2 \\ R/2 & 1 & -R/2 \\ R/2 & 0 & 1-R/2 \\ R/2 & 0 & 1-R/2 \end{pmatrix}.$$

For $R_2 \leq R < R_3$*,*

$$Q^\dagger = \begin{pmatrix} 1+R/2 & -R/2 & 0 & 0 \\ R/2 & 1-R/2 & 0 & 0 \\ R/2 & 1-R/2 & 0 & 0 \\ R/2 & 1-R/2 & 0 & 0 \end{pmatrix},$$

and hence

$$Q = \begin{pmatrix} 1+R/2 & -R/2 \\ R/2 & 1-R/2 \\ R/2 & 1-R/2 \\ R/2 & 1-R/2 \end{pmatrix}.$$

For $R \geq R_3$*,*

$$Q^\dagger = \begin{pmatrix} \frac{1}{\mu_1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \Longrightarrow Q = \begin{pmatrix} \frac{1}{\mu_1} \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Note that, the number of columns of Q matrix is based on the value of total variation

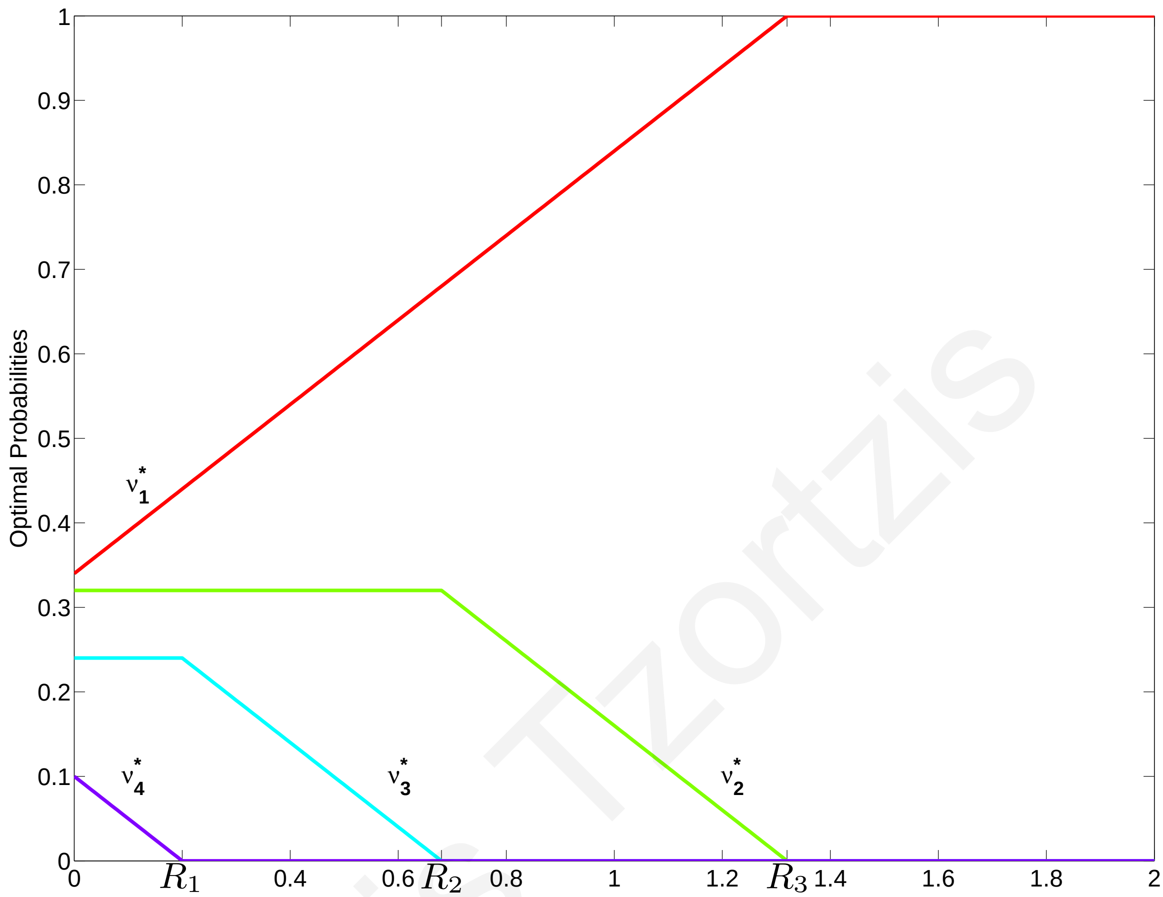


Figure 6.5.: Optimal probabilities as a function of R.

parameter R. For $0 \leq R < R_1$, its dimension is equal to $(|\mathcal{X}|) \times (1 + r)$. Whenever an indicator function becomes equal to zero, all elements of the respective column become equal to zero, and hence the column is deleted, until $R \geq R_3$, where the Q matrix will be transformed into a column vector of dimension $(|\mathcal{X}|) \times (1)$.

Next, we proceed with the solution of (6.9), by letting $\bar{\nu} \in \mathbb{P}(\mathcal{Y})$ to denote the invariant distribution of a lower-dimensional Markov process $(\bar{\nu}, \Phi)$. As mentioned in [19], the main difficulty in solving (6.9) is in finding an optimal partition function φ. However, once an optimal partition is given then the solution of Φ can be easily obtained. Toward this end, next we define an optimal partition function for the approximation problem based on occupancy distribution at values of TV parameter R for which a reduction of the states occurs (i.e., see Example 6.1, Fig.6.5, for values of $R = R_1, R_2$ and R_3).

Definition 6.3. *(Partition function) Let $\mathcal{X}$ and $\mathcal{Y}$ be two finite dimensional state-spaces with*

$|\mathcal{Y}| < |\mathcal{X}|$. *Define a surjective (partition) function* $\varphi : \mathcal{X} \longmapsto \mathcal{Y}$ *as follows.*

$$\forall i \in \mathcal{X}^0, \quad \varphi(i) = 1 \in \mathcal{Y},$$
$$\forall i \in \mathcal{X} \setminus \mathcal{X}^0, \quad \varphi(i) = \begin{cases} 1, & \text{if } \nu_i^* = 0, \\ k \in \mathcal{Y}, & \text{if } \nu_i^* > 0. \end{cases}$$

Note that, once the optimal probabilities ν_i^*, $\forall i \in \mathcal{X}$ are obtained, we can easily identify the values of R for which a reduction of the states occurs. In addition, since the solution behavior of (6.8) is to remove probability mass from states with the smallest invariant probability and strengthening the states with the highest invariant probability, this property of the partition function φ is intuitive and expected.

Next, we reproduce the main theorem of [21], which gives the solution of Φ that solves (6.9).

Theorem 6.3. *Let* $(\mu, P, \mathcal{X})$ *be a given FSM process and* φ *be the partition function of Definition 6.3. For optimization (6.9), the solution of* Φ *is given by*

$$\Phi_{kl} = \frac{u^{(k)} \Pi P u^{(\ell)'}}{\bar{\nu}_k}, \quad k, \ell \in \mathcal{Y}, \tag{6.34}$$

where $\Pi = diag(\mu)$, $u^{(k)'}$ *is the transpose of* $u^{(k)}$, *and* $u^{(k)}$ *is a* $1 \times |\mathcal{X}|$ *row vector defined by*

$$u_i^{(k)} = \begin{cases} 1, & \text{if } \varphi(i) = k, \\ 0, & \text{otherwise}. \end{cases} \tag{6.35}$$

Proof. See [19]. ■

6.3.2. Solution of Approximation Problem based on Maximum Entropy Principle

In this subsection, we first give an algorithm to construct the Q matrix which solves (6.10). Then, under the assumption that the reduced process is also Markov, we give the solution of (6.11). Before giving the algorithm, we introduce some notation.

Let r denote the number of $\mathcal{X}_k$ sets, that is, $1 \leq r \leq |\mathcal{X} \setminus \mathcal{X}^0 \cup \mathcal{X}_0|$ (note that, $\mathcal{X}_0$ set is excluded, in contrast with the definition of r in Section 6.3.1). Furthermore, let r^+ and r^- denote the number of μ_i, $i \in \mathcal{X}$, such that $\mu_i \geq \frac{1}{|\mathcal{X}|}$ and $\mu_i < \frac{1}{|\mathcal{X}|}$, respectively. In addition, $\mu_i \neq \mu_j$ should also be satisfied for all $i \neq j$, $i, j \in \mathcal{X}$.

Remark 6.4. *The initialization step of the following algorithm is performed by letting* $R = 0$. *In this case,* $\nu_i = \mu_i$, $\forall i \in \mathcal{X}$, *and hence,* $\ell_i \triangleq -\log \nu_i = -\log \mu_i$.

Algorithm 6.4.

1. Initialization step:

a) Arrange μ_i, $i \in \mathcal{X}$, in a descending order and let $R = 0$.

b) Identify the support sets $\mathcal{X}^0$, $\mathcal{X}_0$ and $\mathcal{X}_k$ for all $k \in \{1, 2, \ldots, |\mathcal{X} \setminus \mathcal{X}^0 \cup \mathcal{X}_0|\}$.

c) Calculate the value of r, r^- and r^+.

For any $R \in [0, 2]$:

2) Step.1 (Indicator functions):

a) For $k = 1, 2 \ldots, r^- - 1$ let

$$\mu_-^R(\mathcal{X}_k) \triangleq \frac{\sum_{i \in \cup_{j=0}^{k-1} \mathcal{X}_j} \mu_i - R/2}{\sum_{j=0}^{k-1} |\mathcal{X}_j|}.$$

Define

$$I_-^{\mathcal{X}_k} \triangleq \begin{cases} 1, & \text{if } \mu_-^R(\mathcal{X}_k) \leq \frac{\sum_{i \in \mathcal{X}_k} \mu_i}{|\mathcal{X}_k|}, \\ 0, & \text{otherwise}. \end{cases} \tag{6.36}$$

For $k = r^-$ let

$$\mu_-^R(\mathcal{X}_{r^-}) \triangleq \frac{\sum_{i \in \cup_{j=0}^{r^- -1} \mathcal{X}_j} \mu_i - R/2}{\sum_{j=0}^{r^- -1} |\mathcal{X}_j|}.$$

Define

$$I_-^{\mathcal{X}_{r^-}} \triangleq \begin{cases} 1, & \text{if } \mu_-^R(\mathcal{X}_{r^-}) \leq \frac{1}{|\mathcal{X}|}, \\ 0, & \text{otherwise}. \end{cases} \tag{6.37}$$

b) For $k = 1, 2 \ldots, r^+ - 1$ let

$$\mu_+^R(\mathcal{X}_k) \triangleq \frac{\sum_{i \in \mathcal{X} \setminus \cup_{j=r}^{k-1} \mathcal{X}_{r-j}} \mu_i + R/2}{|\mathcal{X} \setminus \cup_{j=r}^{k-1} \mathcal{X}_{r-j}|}.$$

Define

$$I_+^{\mathcal{X}_k} \triangleq \begin{cases} 1, & \text{if } \mu_+^R(\mathcal{X}_k) \geq \frac{\sum_{i \in \mathcal{X}_{r-k+1}} \mu_i}{|\mathcal{X}_{r-k+1}|}, \\ 0, & \text{otherwise}. \end{cases} \tag{6.38}$$

For $k = r^+$ let

$$\mu_+^R(\mathcal{X}_{r^+}) \triangleq \frac{\sum_{i \in \mathcal{X} \setminus \cup_{j=r}^{r^+ -1} \mathcal{X}_{r-j}} \mu_i + \frac{R}{2}}{|\mathcal{X} \setminus \cup_{j=r}^{r^+ -1} \mathcal{X}_{r-j}|}.$$

Define

$$I_+^{\mathcal{X}_{r^+}} \triangleq \begin{cases} 1, & \textit{if } \mu_+^R(\mathcal{X}_{r^+}) \geq \frac{1}{|\mathcal{X}|}, \\ 0, & \textit{otherwise}. \end{cases} \tag{6.39}$$

3) *Step.2 (The $Q^\dagger$ matrix):*

Let $Q^\dagger$ be an $(|\mathcal{X}|) \times (2+r)$ matrix.

a) *The elements of the first column are given as follows.*

i) *For all $i \in \mathcal{X}_0$, let the $(Q^\dagger)_{i,1}$ be equal to*

$$\frac{1-R/2}{|\mathcal{X}_0| + \sum_{j=1}^{r^{\downarrow}-1} I_-^{\mathcal{X}_j}|\mathcal{X}_j|}\left(I_-^{\mathcal{X}_{r^-}}\right)^{\mathsf{c}} + \frac{I_-^{\mathcal{X}_{r^-}}}{|\mathcal{X}|}. \tag{6.40}$$

ii) *For all $i \in \mathcal{X}_k$, $k = 1, 2, \dots, r^- {-} 1$, let the $(Q^\dagger)_{i,1}$ be equal to*

$$\frac{I_-^{\mathcal{X}_k} - R/2}{|\mathcal{X}_0| + \sum_{j=1}^{r^{\downarrow}-1} I_-^{\mathcal{X}_j}|\mathcal{X}_j|}\left(I_-^{\mathcal{X}_{r^-}}\right)^{\mathsf{c}} + \frac{I_-^{\mathcal{X}_{r^-}}}{|\mathcal{X}|}. \tag{6.41}$$

iii) *Let all the remaining elements be equal to*

$$\frac{-R/2}{|\mathcal{X}_0| + \sum_{j=1}^{r^{\downarrow}-1} I_-^{\mathcal{X}_j}|\mathcal{X}_j|}\left(I_-^{\mathcal{X}_{r^-}}\right)^{\mathsf{c}} + \frac{I_-^{\mathcal{X}_{r^-}}}{|\mathcal{X}|}. \tag{6.42}$$

b) *The elements of the last column are given by*

i) *For all $i \in \mathcal{X}^0$, let the $(Q^\dagger)_{i,r+2}$ be equal to*

$$\frac{1+R/2}{|\mathcal{X}^0| + \sum_{j=1}^{r^{\uparrow}-1} I_+^{\mathcal{X}_j}|\mathcal{X}_{r-j+1}|}(I_+^{\mathcal{X}_{r^+}})^{\mathsf{c}}. \tag{6.43}$$

ii) *For all $i \in \mathcal{X}_{r-k+1}$, $k = 1, 2, \dots, r^{\uparrow} - 1$ let the $(Q^\dagger)_{i,r+2}$ be equal to*

$$\frac{I_+^{\mathcal{X}_k} + R/2}{|\mathcal{X}^0| + \sum_{j=1}^{r^{\uparrow}-1} I_+^{\mathcal{X}_j}|\mathcal{X}_{r-j+1}|}(I_+^{\mathcal{X}_{r^+}})^{\mathsf{c}}. \tag{6.44}$$

iii) *Let all the remaining elements be equal to*

$$\frac{R/2}{|\mathcal{X}^0| + \sum_{j=1}^{r^{\uparrow}-1} I_+^{\mathcal{X}_j}|\mathcal{X}_{r-j+1}|}(I_+^{\mathcal{X}_{r^+}})^{\mathsf{c}}. \tag{6.45}$$

c) *The elements of all remaining columns are given by*

i) *For all $i \in \mathcal{X}_k$, $k = 1, 2, \dots, r^- {-} 1$ let*

$$(Q^\dagger)_{i,z} = \frac{(I_-^{\mathcal{X}_k})^{\mathsf{c}}}{|\mathcal{X}_k|}, \tag{6.46}$$

where $z = 1 + k$ denotes the zth column. Let all the remaining elements of the zth column be equal to zero. However, if $I_{-}^{\mathcal{X}_k} = 1$, then let all the elements of the zth column be equal with the corresponding elements of the first column, that is,

$$(Q^{\dagger})_{1,z} = (Q^{\dagger})_{1,1}, \quad (Q^{\dagger})_{2,z} = (Q^{\dagger})_{2,1}, \quad \dots \quad (Q^{\dagger})_{|\mathcal{X}|,z} = (Q^{\dagger})_{|\mathcal{X}|,1}. \tag{6.47}$$

ii) For all $i \in \mathcal{X}_{r-k+1}$, $k = 1, 2, \dots, r^{+}-1$ let

$$(Q^{\dagger})_{i,z} = \frac{(I_{+}^{\mathcal{X}_k})^{\mathsf{c}}}{|\mathcal{X}_k|}, \tag{6.48}$$

where $z = r + 2 - k$ denotes the zth column. Let all the remaining elements of the zth column be equal to zero. However, if $I_{+}^{\mathcal{X}_k} = 1$, then let all the elements of the zth column be equal with the corresponding elements of the last column, that is,

$$(Q^{\dagger})_{1,z} = (Q^{\dagger})_{1,|\mathcal{X}|}, \quad (Q^{\dagger})_{2,z} = (Q^{\dagger})_{2,|\mathcal{X}|}, \quad \dots \quad (Q^{\dagger})_{|\mathcal{X}|,z} = (Q^{\dagger})_{|\mathcal{X}|,|\mathcal{X}|}. \tag{6.49}$$

Once the $Q^{\dagger}$ matrix is constructed, as a function of the TV parameter R, then by (6.21) the solution of optimization (6.10) is readily available, and hence, by Definition 6.2, the lower-dimensional process $\{Y_t : t = 0, 1 \dots\}$ with invariant distribution $\bar{\nu}$ is obtained, either by adding all equal elements of $\nu^* \in \mathbb{P}(\mathcal{X})$, or by defining a Q matrix to be equal to $Q^{\dagger}$, after the merging of all equal columns (by adding them). Hence

$$\bar{\nu} = \mu Q = \mu P Q, \tag{6.50}$$

where the dimensions of Q matrix are based on the value of the TV parameter $R \in [0, 2]$.

Before we proceed with the solution of (6.11), we provide a simple example in order to explain each step of Algorithm 6.4.

Example 6.2. *Let $\mu = [\mu_1\ \mu_2\ \mu_3\ \mu_4]$, where $\mu_1 > \mu_2 > \mu_3 > \mu_4$, and also assume that $\mu_1 > \mu_2 > \frac{1}{|\mathcal{X}|}$ and $\mu_4 < \mu_3 < \frac{1}{|\mathcal{X}|}$, where $|\mathcal{X}| = 4$. For simplicity of presentation it is assumed that the optimum probabilities ν_i^*, $i \in \mathcal{X}$, as a function of R are as shown in Fig.6.6.*

Initialization step. For $R = 0$, and from Remark 6.4, we conclude that $\ell_1 < \ell_2 < \ell_3 < \ell_4$, and therefore the support sets are equal to $\mathcal{X}^0 = \{4\}$, $\mathcal{X}_0 = \{1\}$, $\mathcal{X}_1 = \{2\}$ and $\mathcal{X}_2 = \{3\}$. The number of the $\mathcal{X}_k$ sets is equal to $r = 2$. The number of μ_i, $i \in \mathcal{X}$, which are greater (or equal) than $\frac{1}{|\mathcal{X}|} = 0.25$ (and also $\mu_i \neq \mu_j$, $i, j \in \mathcal{X}$) is $r^{-} = 2$. Similarly, the number of μ_i which are strictly smaller than $\frac{1}{|\mathcal{X}|} = 0.25$ (and also not equal to each other) is also $r^{+} = 2$.

Step.1 From (6.36)-(6.37)*, the indicator functions* $I_-^{\mathcal{X}_1}$ *and* $I_-^{\mathcal{X}_2}$ *are given by*

$$I_-^{\mathcal{X}_1} \triangleq \begin{cases} 1, & \textit{if } \mu_1 - \frac{R}{2} \leq \mu_2, \\ 0, & \textit{otherwise}, \end{cases} \qquad I_-^{\mathcal{X}_2} \triangleq \begin{cases} 1, & \textit{if } \frac{\mu_1+\mu_2-R/2}{2} \leq 0.25, \\ 0, & \textit{otherwise}, \end{cases}$$

and from (6.38)-(6.39)*, the indicator functions* $I_+^{\mathcal{X}_1}$ *and* $I_+^{\mathcal{X}_2}$ *are given by*

$$I_+^{\mathcal{X}_1} \triangleq \begin{cases} 1, & \textit{if } \mu_4 + \frac{R}{2} \geq \mu_3, \\ 0, & \textit{otherwise}, \end{cases} \qquad I_+^{\mathcal{X}_2} \triangleq \begin{cases} 1, & \textit{if } \frac{\mu_3+\mu_4+R/2}{2} \geq 0.25, \\ 0, & \textit{otherwise}. \end{cases}$$

The values of the indicator functions for $R \in [0, 2]$ *are shown in Fig.*6.6*. For* $0 \leq R < R_1$*, that is, before a merge occurs, all indicator functions are equal to zero. If a merge occurs the respective indicator functions become equal to one, until for some* $R \geq R_3$*, where all indicator functions are equal to one.*

Step.2 Let $Q^\dagger$ *be an* 4×4 *matrix. For* $0 \leq R < R_1$*,*

$$Q^\dagger = \begin{pmatrix} 1-R/2 & 0 & 0 & R/2 \\ -R/2 & 1 & 0 & R/2 \\ -R/2 & 0 & 1 & R/2 \\ -R/2 & 0 & 0 & 1+R/2 \end{pmatrix},$$

and since no equal columns exist then $Q^\dagger = Q$*. For* $R_1 \leq R < R_2$*,*

$$Q^\dagger = \begin{pmatrix} \frac{1-R/2}{2} & \frac{1-R/2}{2} & 0 & R/2 \\ \frac{1-R/2}{2} & \frac{1-R/2}{2} & 0 & R/2 \\ -R/4 & -R/4 & 1 & R/2 \\ -R/4 & -R/4 & 0 & 1+R/2 \end{pmatrix},$$

and hence

$$Q = \begin{pmatrix} 1-R/2 & 0 & R/2 \\ 1-R/2 & 0 & R/2 \\ -R/2 & 1 & R/2 \\ -R/2 & 0 & 1+R/2 \end{pmatrix}.$$

For $R_2 \leq R < R_3$*,*

$$Q^\dagger = \begin{pmatrix} \frac{1-R/2}{2} & \frac{1-R/2}{2} & R/4 & R/4 \\ \frac{1-R/2}{2} & \frac{1-R/2}{2} & R/4 & R/4 \\ -R/4 & -R/4 & \frac{1+R/2}{2} & \frac{1+R/2}{2} \\ -R/4 & -R/4 & \frac{1+R/2}{2} & \frac{1+R/2}{2} \end{pmatrix},$$

and hence

$$Q = \begin{pmatrix} 1 - R/2 & R/2 \\ 1 - R/2 & R/2 \\ -R/2 & 1 + R/2 \\ -R/2 & 1 + R/2 \end{pmatrix}.$$

For $R \geq R_3$,

$$Q^{\dagger} = \begin{pmatrix} 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \end{pmatrix} \Longrightarrow Q = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

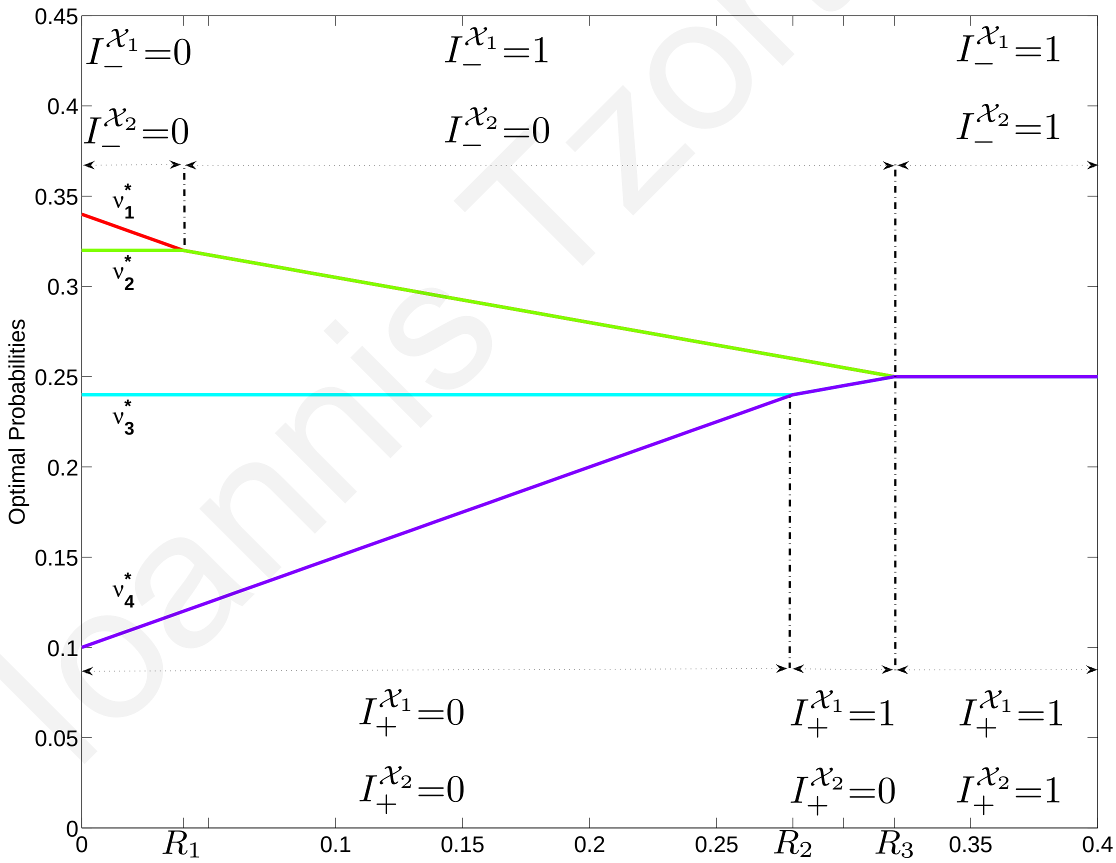


Figure 6.6.: Optimal Probabilities as a function of R.

Note that, the dimension of matrix Q is based on the value of total variation distance parameter R. For $0 < R \leq R_1$ its dimension is equal to $(|\mathcal{X}|) \times (2 + r)$. Whenever two columns become equal (that is, an indicator function is activated) they are merged, until for some $R \geq R_2$, where matrix Q is transformed into column vector of dimension $(|\mathcal{X}|) \times (1)$.

Next, we proceed with the solution of (6.11), by letting $\bar{\nu}$ to denote the invariant distribution of a lower-dimensional Markov process $(\bar{\nu}, \Phi)$. To this end, we next define an optimal partition function for the approximation problem, based on maximum entropy principle at values of TV parameter R, for which an aggregation of the states occurs (i.e., see Example 6.2, Fig.6.6, for values of $R = R_1$, R_2 and R_3.).

Definition 6.4. *(partition function) Let $\mathcal{X}$ and $\mathcal{Y}$ be two finite dimensional state-spaces with $|\mathcal{Y}| < |\mathcal{X}|$. Define a surjective (partition) function $\varphi : \mathcal{X} \longmapsto \mathcal{Y}$ as follows*

$$\forall i, j \in \mathcal{X}, \quad \varphi(i) = \varphi(j) = k \in \mathcal{Y} \quad \text{if} \quad \nu_i^* = \nu_j^*. \tag{6.51}$$

Note that, once the optimal probabilities ν_i^*, $\forall i \in \mathcal{X}$ are obtained, we can easily identify the values of R for which an aggregation of the states occurs. Next, we reproduce the main theorem of [21], which gives the solution of Φ that solves (6.11).

Theorem 6.5. *Let $(\mu, P, \mathcal{X})$ be a FSM process and φ be the partition function of Definition 6.4. For optimization (6.11), the solution of Φ is given by*

$$\Phi_{kl} = \frac{u^{(k)} \Pi P u^{(\ell)'}}{\bar{\nu}_k}, \quad k, \ell \in \mathcal{Y} \tag{6.52}$$

where $\Pi = diag(\nu^*)$, $u^{(k)'}$ is the transpose of $u^{(k)}$, and $u^{(k)}$ is a $1 \times |\mathcal{X}|$ row vector defined by

$$u_i^{(k)} = \begin{cases} 1 & \text{if } \varphi(i) = k \\ 0 & \text{otherwise} \end{cases} \tag{6.53}$$

Proof. See [19]. ∎

6.4. Examples

In this section, the theoretical results of both methods are applied to specific examples to illustrate the methodology, and the water-filling behavior of the approximations.

6.4.1. Markov Chain Approximation with a Small Number of States

In this example we employ the theoretical results obtained in preceding sections to approximate a 4-state FSM process $(\mu, P, \mathcal{X})$ with transition probability matrix given by

$$P = \begin{bmatrix} 0.4 & 0.2 & 0.3 & 0.1 \\ 0.3 & 0.5 & 0.1 & 0.1 \\ 0.2 & 0.3 & 0.4 & 0.1 \\ 0.6 & 0.2 & 0.1 & 0.1 \end{bmatrix}, \tag{6.54}$$

and steady state nominal probability vector equal to

$$\mu = [0.34 \;\; 0.32 \;\; 0.24 \;\; 0.1]. \tag{6.55}$$

In particular, in Section 6.4.1, we solve approximation problem based on *Method 1*. In Section 6.4.1 we solve the approximation problem based on occupancy distribution, and in Section 6.4.1 based on entropy principle of *Method 2*.

Solution of Problem 6.1

Let $\ell = \{\ell \in \mathbb{R}_+^4 : \ell_1 > \ell_2 > \ell_3 > \ell_4\}$, then the support sets are given by $\mathcal{X}^0 = \{1\}$, $\mathcal{X}_0 = \{4\}$, $\mathcal{X}_1 = \{3\}$ and $\mathcal{X}_2 = \{2\}$, and by (6.18d), $R_{\max,1} = 1.2$, $R_{\max,2} = 1.4$, $R_{\max,3} = 1.6$ and $R_{\max,4} = 0.8$. By employing Theorem 6.1, the optimal $\Phi^\dagger$ and Φ matrices are obtained as a function of TV parameter R, as shown in Table 6.1. Note that, in contrast with Problems 6.2-6.3, where the approximation is performed only for values of R for which a reduction of the states occurs, the solution of Problem 6.1 is obtained for all values of total variation parameter.

R	$\Phi^\dagger$	Φ
0	$\begin{bmatrix} .4 & .2 & .3 & .1 \\ .3 & .5 & .1 & .1 \\ .2 & .3 & .4 & .1 \\ .6 & .2 & .1 & .1 \end{bmatrix}$	$\begin{bmatrix} .4 & .2 & .3 & .1 \\ .3 & .5 & .1 & .1 \\ .2 & .3 & .4 & .1 \\ .6 & .2 & .1 & .1 \end{bmatrix}$
0.2	$\begin{bmatrix} .5 & .2 & .3 & 0 \\ .4 & .5 & .1 & 0 \\ .3 & .3 & .4 & 0 \\ .7 & .2 & .1 & 0 \end{bmatrix}$	$\begin{bmatrix} .5 & .2 & .3 \\ .4 & .5 & .1 \\ .3 & .3 & .4 \end{bmatrix}$
1	$\begin{bmatrix} .9 & .1 & 0 & 0 \\ .8 & .2 & 0 & 0 \\ .7 & .3 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} .9 & .1 \\ .8 & .2 \end{bmatrix}$
1.4	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 \end{bmatrix}$

Table 6.1.: Optimal results obtained by the Approximation based on *Method 1*.

Solution of Problem 6.2

By employing Algorithm 6.2, with $\ell_i \triangleq \mu_i$, $i = 1, \ldots, 4$, and support sets given by $\mathcal{X}^0 = \{1\}$, $\mathcal{X}_0 = \{4\}$, $\mathcal{X}_1 = \{3\}$ and $\mathcal{X}_2 = \{2\}$ the maximizing distribution of (6.8) exhibits a water-filling behavior as depicted in Fig.6.5. For values of TV parameter $0 \leq R \leq R_1 = 0.2$, all maximizing probabilities ν_i^*, $i = 1, \ldots, 4$, are greater than zero and hence $|\mathcal{Y}| = 4 = |\mathcal{X}|$ and $\bar{\nu}_i = \nu_i^*$, $i = 1, \ldots, 4$. However, for $R_1 \leq R < R_2 = 0.68$, $|\mathcal{Y}| = 3 < |\mathcal{X}| = 4$ since ν_4^* becomes equal to zero and hence $\bar{\nu}_i = \nu_i^*$, $i = 1, 2, 3$. The procedure follows until for some $R \geq R_3 = 1.32$ in which $|\mathcal{Y}| = 1$ and $\bar{\nu}_1 = \nu_1^* = 1$.

From the above discussion, it is clear that, the solution of approximation problem based on occupancy distribution is described via a water-filling deletion of states with the smallest invariant probability and maintaining and strengthening the states with the highest invariant probability, and hence a lower-dimensional distribution $\bar{\nu}$ is obtained which is then applied to the problem of Markov by Markov approximation. For the solution of (6.9), first we find an optimal partition function φ and then we calculate a transition probability matrix Φ which best approximates transition matrix P only for values of R for which a reduction of states occurs, that is, for $R = 0,\ 0.2,\ 0.68$ and 1.32. The optimal results are depicted in Table 6.2.

Solution of Problem 6.3

By employing Algorithm 6.4, with $\ell_i \triangleq -\log \nu_i$, $i = 1, \ldots, 4$, the support sets are calculated for $R = 0$, where $\nu_i^* = \mu_i$ and hence $\ell_i = -\log \mu_i$, and are equal to $\mathcal{X}^0 = \{4\}$, $\mathcal{X}_0 = \{1\}$, $\mathcal{X}_1 = \{2\}$ and $\mathcal{X}_2 = \{3\}$. The maximizing distribution of (6.10) exhibits a water-filling like behavior as depicted in Fig.6.6. For values of $0 \leq R < R_1 = 0.04$, $|\mathcal{Y}| = 4 = |\mathcal{X}|$ since $\nu_i^* \neq \nu_j^*$ for $i \neq j$, $i, j = 1, \ldots, 4$ and hence $\bar{\nu}_i = \nu_i^*$, $i = 1, \ldots, 4$. For $R_1 \leq R < R_2 = 0.28$, $|\mathcal{Y}| = 3 < |\mathcal{X}| = 4$ since ν_1^* becomes equal to ν_2^* and hence $\bar{\nu}_1 = \nu_1^* + \nu_2^*$ and $\bar{\nu}_i = \nu_i^*$, $i = 3, 4$. The procedure follows until for some $R \geq R_3 = 0.32$ in which $|\mathcal{Y}| = 1$ and $\bar{\nu}_1 = \sum_{i=1}^{4} \nu_i^* = \frac{1}{4}$.

In summary, the solution of approximation problem based on entropy principle is described via aggregation of states, that is, by grouping certain states of the original Markov chain to obtain the approximating reduced state process.Then the lower-dimensional distribution $\bar{\nu}$ is applied to problem (6.11). The optimal partition function φ and the transition probability matrix Φ which minimizes the KL divergence rate for values of $R = 0,\ 0.04,\ 0.28$ and 0.32 are as shown in Table 6.3.

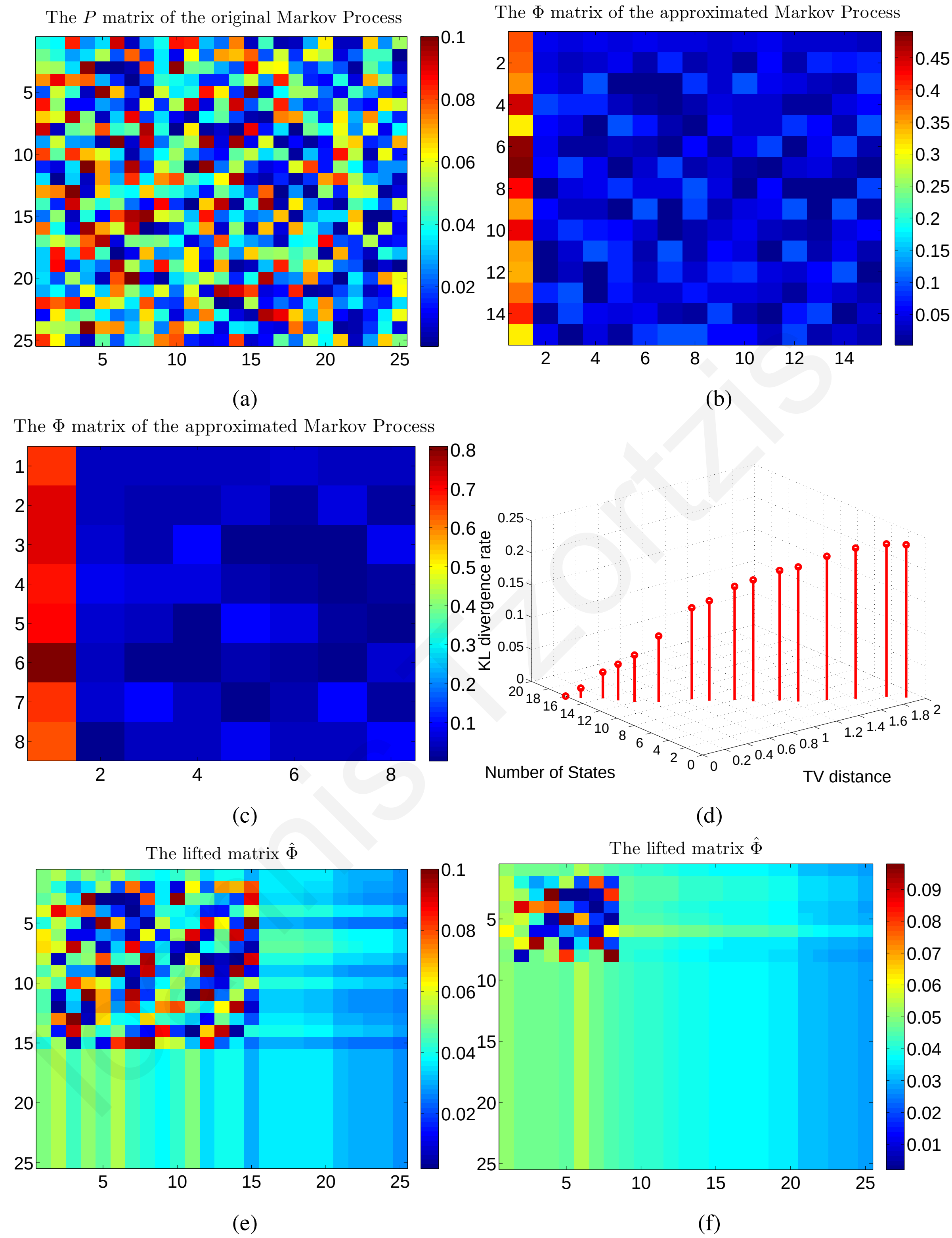


Figure 6.7.: Approximation results based on occupancy distribution: Plot (a) depicts the P matrix of the original Markov process. Plot (b) depicts a 15-state approximation. Plot (c) depicts an 8-state approximation. Plot (d) depicts the KL divergence rate. Plot (e) depicts the lifted $\hat{\Phi}$ matrix for the 15-state approximation. Plot (f) depicts the lifted $\hat{\Phi}$ matrix for the 8-state approximation.

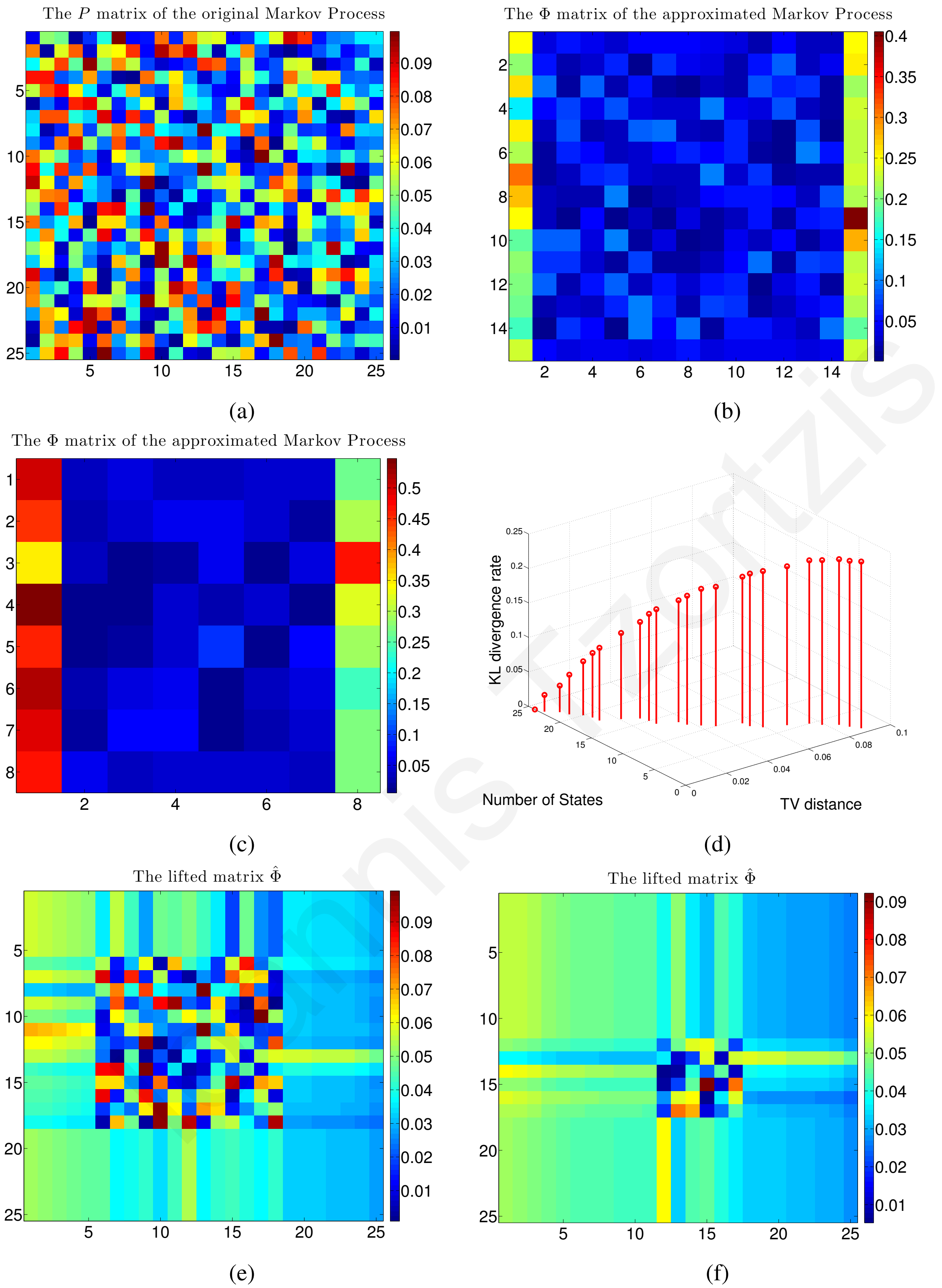


Figure 6.8.: Approximation results based on maximum entropy: Plot (a) depicts the P matrix of the original Markov process. Plot (b) depicts a 15-state approximation. Plot (c) depicts an 8-state approximation. Plot (d) depicts the KL divergence rate. Plot (e) depicts the lifted $\hat{\Phi}$ matrix for the 15-state approximation. Plot (f) depicts the lifted $\hat{\Phi}$ matrix for the 8-state approximation.

R	$\bar{\nu}$	Q	φ	Φ
0	[.34 .32 .24 .1]	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	$\varphi(1) = 1$ $\varphi(2) = 2$ $\varphi(3) = 3$ $\varphi(4) = 4$	$\begin{bmatrix} .4 & .2 & .3 & .1 \\ .3 & .5 & .1 & .1 \\ .2 & .3 & .4 & .1 \\ .6 & .2 & .1 & .1 \end{bmatrix}$
0.2	[.44 .32 .24]	$\begin{bmatrix} 1.1 & 0 & -.1 \\ .1 & 1 & -.1 \\ .1 & 0 & .9 \\ .1 & 0 & .9 \end{bmatrix}$	$\varphi(1) = 1$ $\varphi(2) = 2$ $\varphi(3) = 3$ $\varphi(4) = 1$	$\begin{bmatrix} .5455 & .2 & .2545 \\ .4 & .5 & .1 \\ .3 & .3 & .4 \end{bmatrix}$
0.68	[0.68 0.32]	$\begin{bmatrix} 1.34 & -.34 \\ .34 & .66 \\ .34 & .66 \\ .34 & .66 \end{bmatrix}$	$\varphi(1) = 1$ $\varphi(2) = 2$ $\varphi(3) = 1$ $\varphi(4) = 1$	$\begin{bmatrix} .7647 & .2353 \\ .5 & .5 \end{bmatrix}$
1.32	[1]	$\begin{bmatrix} 2.94 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\varphi(1) = 1$ $\varphi(2) = 1$ $\varphi(3) = 1$ $\varphi(4) = 1$	$\begin{bmatrix} 1 \end{bmatrix}$

Table 6.2.: Optimal results obtained by the Approximation based on occupancy distribution.

6.4.2. Markov Chain Approximation based on Occupancy Distribution with a Large Number of States

In this example we approximate a 25-state Markov process based on occupancy distribution. The transition matrix P of the original Markov process is as shown in Fig.6.7a, in which the color of the ith row and jth column represents the P_{ij} element as indicated by the color bar. Then, based on the resulting values of μ_i, $\forall i \in \mathcal{X}$, the state space $\mathcal{X}$ is partitioned into 16 disjoint sets, where

$$\mathcal{X}^0 = \{1\},\ \mathcal{X}_0 = \{25\},\ \mathcal{X}_1 = \{24, 23\}, \mathcal{X}_2 = \{22\},\ \mathcal{X}_3 = \{21\},\ \mathcal{X}_4 = \{20, 19\},$$
$$\mathcal{X}_5 = \{18, 16\},\ \mathcal{X}_6 = \{15\},\ \mathcal{X}_7 = \{14, 13\},\ \mathcal{X}_8 = \{12\},\ \mathcal{X}_9 = \{11, 10\},$$
$$\mathcal{X}_{10} = \{9\},\ \mathcal{X}_{11} = \{8, 7\},\ \mathcal{X}_{12} = \{6, 5\},\ \mathcal{X}_{13} = \{4, 3\},\ \mathcal{X}_{14} = \{2\}.$$

Fig.6.7d depicts the KL divergence rate as a function of the number of the states of the approximated Markov process and also as a function of the TV parameter R for values where a reduction of the states occurs, due to the *water-filling* behaviour of the solution. Fig.6.7b-e depict the Φ matrix and the corresponding lifted matrix $\hat{\Phi}$ of the approximated Markov

R	$\bar{\nu}$	Q	φ	Φ
0	[.34 .32 .24 .1]	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	$\varphi(1)=1$ $\varphi(2)=2$ $\varphi(3)=3$ $\varphi(4)=4$	$\begin{bmatrix} .4 & .2 & .3 & .1 \\ .3 & .5 & .1 & .1 \\ .2 & .3 & .4 & .1 \\ .6 & .2 & .1 & .1 \end{bmatrix}$
0.04	[.64 .24 .12]	$\begin{bmatrix} .98 & 0 & .02 \\ .98 & 0 & .02 \\ -.02 & 1 & .02 \\ -.02 & 0 & 1.02 \end{bmatrix}$	$\varphi(1)=1$ $\varphi(2)=1$ $\varphi(3)=2$ $\varphi(4)=3$	$\begin{bmatrix} .7 & .2 & .1 \\ .5 & .4 & .1 \\ .8 & .1 & .1 \end{bmatrix}$
0.28	[0.52 0.48]	$\begin{bmatrix} .86 & .14 \\ .86 & .14 \\ -.14 & 1.14 \\ -.14 & 1.14 \end{bmatrix}$	$\varphi(1)=1$ $\varphi(2)=1$ $\varphi(3)=2$ $\varphi(4)=2$	$\begin{bmatrix} .7 & .3 \\ .65 & .35 \end{bmatrix}$
0.32	[1]	$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$	$\varphi(1)=1$ $\varphi(2)=1$ $\varphi(3)=1$ $\varphi(4)=1$	$\begin{bmatrix} 1 \end{bmatrix}$

Table 6.3.: Optimal results obtained by the Approximation based on entropy principle.

process, when the 25-state Markov process is approximated by a 15-state Markov process. Similarly, Fig.6.7c-f depict Φ and $\hat{\Phi}$ when the 25-state Markov process is approximated by an 8-state Markov process.

6.4.3. Markov Chain Approximation based on Maximum Entropy with a Large Number of States

In this example we approximate a 25-state Markov process based on maximum entropy. The transition matrix P of the original Markov process is as shown in Fig.6.8a. By Remark 6.4, the state-space $\mathcal{X}$ is partitioned into 25 disjoint sets, where $\mathcal{X}^0 = \{25\}$, $\mathcal{X}_0 = \{1\}$ and $\mathcal{X}_k = \{k+1\}$ for $k = 1, \ldots, 23$. Similarly to example 6.4.2, Fig.6.8d depicts the KL divergence rate as a function of the number of the states of the approximated Markov process and as a function of TV parameter for values where an aggregation of the states occurs. It is worth noting, that the approximation based on maximum entropy principle is much faster, in terms of TV parameter, compared to the approximation based on occupancy and this is due to the *water-filling like* behavior of the solution. Fig.6.8b-e and 6.8c-f depict the Φ matrix and the corresponding lifted matrix $\hat{\Phi}$ when the original Markov process is approximated by a 15-state and an 8-state Markov process, respectively.

6.5. Summary

In this chapter, we present two methods of approximating a FSM process by another process with fewer states. The first method, utilizes the total variation distance to discriminate the transition probabilities of a high-dimensional FSM process by a reduced order Markov process, and hence, a direct method for a Markov by Markov approximation is obtained. The second method, utilizes total variation distance as a new discrepancy measure, and the problem is formulated using: (a) maximization of an average pay-off functional with respect to the approximated invariant probability, and, (b) maximization of the entropy of the approximated invariant probability, both subject to a constraint on the total variation distance metric between the invariant probability of the original Markov process and that of the approximated process. Then, by utilizing the obtained solution, we studied the problem of approximating a FSM process with another FSM process of reduced order with respect to the Kullback-Leibler divergence rate. Examples are included to demonstrate the approximation approach for each of the two methods.

7
Conclusion

In this chapter, a summary of the main findings of the thesis is presented, and suggestions for future research are discussed.

7.1. Summary and Main Contributions

Extremum problems with total variation distance metric on the space of probability measures are of fundamental importance; they have applications in stochastic optimal control, decision theory, information theory, mathematical finance, etc. In this thesis, such extremum problems are introduced and analyzed. Subsequently, the results are applied to minimax stochastic optimal control via dynamic programming on a finite and on an infinite horizon, and to approximate high-dimensional Markov processes by lower-dimensional processes. Below, we give a brief summary followed by the main contributions of the thesis.

Extremum Problems

In this chapter, we investigated extremum problems with pay-off being the total variation distance metric defined on the space of probability measures, subject to linear functional constraints on the space of probability measures and vice-versa; that is with the roles of total variation metric and linear functional interchanged. First, we introduced the precise definitions of the extremum problems under investigation, and then we studied some of their most important properties. Next, by utilizing concepts from signed measures we characterized the extremum measures, which exhibit a water-filling behavior, on abstract and on finite alphabet spaces. In particular, the construction of the extremum measures involves the identification of the partition of their support set, and their mass defined on these partitions. Due to the convexity of these extremum problems, the optimal solution of all problems is obtained explicitly, by finding upper and lower bounds which are achievable. The main contributions of this part are the following.

- Characterization of the properties of the extremum problems;
- characterization of extremum measures on abstract spaces;
- closed form expressions of the extremum measures for finite alphabets.

Dynamic Programming on a Finite Horizon

In this chapter, we addressed optimality of stochastic optimal control strategies on a finite horizon, via dynamic programming subject to total variation distance ambiguity on the conditional distribution of the controlled process. The solution of the minimax stochastic control with deterministic strategies, is obtained under a Markovian and a non-Markovian assumption, on the conditional distribution of the controlled process. Although, optimality of the control strategies subject to uncertainty have been addressed previously by several authors using relative entropy constraints, our approach is novel in the sense that the class of models is described by total variation distance between the nominal and the true distributions. The essence of our approach lies in the fact that total variation is more general than relative entropy, and in addition, it has the advantage of admitting distributions which are singular with respect to the nominal distribution. Hence, stochastic control under total variation results in optimal policies which are more robust with respect to uncertainty. The main contributions of this part are the following.

- Minimax optimization subject to total variation distance ambiguity constraint;
- new dynamic programming recursions, which includes in addition to the classical terms, the oscillator semi-norm of the future cost-to-go.

Dynamic Programming on an Infinite Horizon

In this chapter, we addressed optimality of stochastic control strategies on an infinite horizon, via dynamic programming subject to total variation distance ambiguity on the conditional distribution of the controlled process, by considering as optimality criterion both the expected discounted reward and the average pay-off per unit time. New policy iteration algorithms, with convergence properties, are developed for computing the optimal policies which in contrast to the classical ones, policy improvement and policy evaluation steps are performed using the maximizing conditional distribution obtained under total variation distance ambiguity constraints. The main contributions of this part are the following.

- New infinite horizon discounted dynamic programming equation, associated contractive property, and a new policy iteration algorithm;
- new infinite horizon average dynamic programming equations, and new policy iteration algorithms.

Approximation of Markov Processes by Lower Dimensional Processes
In this chapter, we investigated the problem of approximating a finite-state Markov process by another process with fewer states, called the approximating process. The approximation problem is formulated as an optimization problem with respect to a certain pay-off subject to a fidelity criterion defined by the total variation distance, using two different methods. In the first method, we approximated the transition probabilities of a Markov process by another Markov process with reduced transition probability matrix, while in the second method we approximated a Markov process by another process which is non-necessarily Markov, but with lower-dimensional state-space. For both methods, the resulting approximating processes are given by water-filling solutions, and new recursive algorithms are developed to compute the invariant distribution of the approximating processes. The main contributions of this part are the following.

- A direct method for Markov by Markov approximation based on the transition probabilities of the original FSM process and the reduced one;
- optimal partition functions which aggregate the original finite-state Markov process to form the reduced order finite-state Markov process;
- iterative algorithms to compute the invariant distribution of the approximating process.

7.2. Topics for Further Research

Extremum problems with total variation distance metric on the space of probability measures can be further generalized as follows.

1. In this thesis, extremum problems with total variation distance metric on the space of probability measures have been formulated and solved, and their solutions are applied to the areas of minimax stochastic control and Markov process approximation. It will

be interesting to consider applications of extremum problems on Information Theory such as capacity of channel for a class of channel distributions, and lossy compression with fidelity criterion for a class of sources via the Rate Distortion Function. Moreover, it will be interesting to extend the example on lossless coding of section 3.4.4 to universal coding and modeling for the purpose of introducing lossless codes for a class of source distributions described by total variation distance metric via minimax theory.

2. In chapters 4 and 5 we have been exclusively concerned with Markov controlled optimization using deterministic control strategies. It would be desirable to solve minimax stochastic optimal control problems with randomized control strategies, under a Markovian and a non-Markovian assumption. In addition, it would be very interesting to develop new dynamic programming algorithms by employing performance criteria of different types, i.e., of exponential type.

3. In the literature several authors investigate optimality of stochastic control on problems with complete observations and on problems with partial observations. The extension of our work and results to the partially observed case it would be a very interesting and challenging problem.

4. Recall the discussion of chapter 5.3.3, where policy iteration algorithm may not be sufficient to give the optimal policy and the minimum cost, if irreducibility condition is not satisfied. It would be challenging, to investigate the problem of approximating the reducible maximizing stochastic matrix, by employing certain results and techniques presented in chapter 6, by an irreducible stochastic matrix of lower dimension, and then to compare the resulting policies with the corresponding optimal policies obtained by solving the general dynamic programming equations.

5. In chapter 6 we have been exclusively concerned with the problem of approximating a finite-state Markov process by another process (non-necessarily Markov) with fewer states. A natural extension of the proposed Markov process approximations is to consider the problem of approximating a hidden Markov process, instead. Furthermore, it would be interesting to investigate the problem of approximating joint distributions by lower-dimensional joint distributions.

A
Total Variation Distance

A.1. Proof of Lemma 2.2

(i) Let H be a Hahn-Jordan set of ξ. Then $\xi^+(H) = \xi(H)$ and $\xi^-(H^c) = -\xi(H^c)$. For $f \in BM(\Sigma)$

$$\begin{aligned}\Big|\xi(f)\Big| &= \Big|\int f(x)d\xi^+(x) - \int f(x)d\xi^-(x)\Big|\\ &\le \Big|\xi^+(f)\Big| + \Big|\xi^-(f)\Big|\\ &= \Big|\int f(x)d\xi^+(x)\Big| + \Big|\int f(x)d\xi^-(x)\Big|\\ &\le \|f\|_\infty\Big(\xi^+(\Sigma) + \xi^-(\Sigma)\Big)\\ &= \|f\|_\infty\|\xi\|_{TV}.\end{aligned}$$

(ii) From part (i),

$$\sup\{\xi(f) : f \in BM(\Sigma), \|f\|_\infty = 1\} \le \|\xi\|_{TV}$$

and

$$\sup\{\xi(f) : \xi \in \mathcal{M}(\Sigma), \|\xi\|_{TV} = 1\} \le \|f\|_\infty.$$

Note that[1], $\|\mathbb{1}_H - \mathbb{1}_{H^c}\|_\infty = 1$ and $\xi(\mathbb{1}_H - \mathbb{1}_{H^c}) = \xi(H) - \xi(H^c) = \|\xi\|_{TV}$. Taking $f = \mathbb{1}_H - \mathbb{1}_{H^c}$, establishes equality in (ii).

[1] $\mathbb{1}_H$ denotes the indicator function of H.

(iii) Let $f \in BM(\Sigma)$ and let $\{x_n\}$ be a sequence in Σ such that $\lim_{n\to\infty} |f(x_n)| = \|f\|_\infty$. Then, $\|f\|_\infty = \lim_{n\to\infty} |\delta_{x_n}(f)|$ proving (iii), (e.g., $\xi = \lim_{n\to\infty} \delta_{x_n}$ such that $\lim_{n\to\infty} |f(x_n)| = \|f\|_\infty$).

A.2. Proof of Lemma 2.3

Let $\xi \in \mathcal{M}(\Sigma)$ and $f \in BM(\Sigma)$. Then

$$\begin{aligned}\xi(f) &= \int f(x)\xi^+(dx) - \int f(x)\xi^-(dx)\\ &= \frac{\int(\int f(x)\xi^+(dx))\xi^-(dx')}{\xi^-(\Sigma)} - \frac{\int(\int f(x')\xi^+(dx))\xi^-(dx')}{\xi^+(\Sigma)}.\end{aligned}$$

Hence,

$$\begin{aligned}|\xi(f)| &\leq \int\int \left|\frac{f(x)}{\xi^-(\Sigma)} - \frac{f(x')}{\xi^+(\Sigma)}\right| \xi^+(dx)\xi^-(dx')\\ &\leq \sup_{(x,x')\in\Sigma\times\Sigma} \left|\frac{f(x)}{\xi^-(\Sigma)} - \frac{f(x')}{\xi^+(\Sigma)}\right| \xi^+(\Sigma)\xi^-(\Sigma)\end{aligned}$$

which proves the first part. If $\xi(\Sigma) = 0$, then $\xi^+(\Sigma) = \xi^-(\Sigma) = \frac{1}{2}\|\xi\|_{TV}$, which proves the second part.

A.3. Proof of Lemma 2.4

For any $\xi_1, \xi_2 \in \mathcal{M}_1(\Sigma)$, write

$$\xi_1(A) - \xi_2(A) = (\xi_1 - \xi_2)\mathbb{1}_A$$

e.g., $f = \mathbb{1}_A$ and note that

$$\text{osc}(\mathbb{1}_A) = 1.$$

Thus, by $|\xi_1(f) - \xi_2(f)| \leq \frac{1}{2}\|\xi_1 - \xi_2\|_{TV}\text{osc}(f)$, letting $f = \mathbb{1}_A$ then

$$|\xi_1(A) - \xi_2(A)| \leq \frac{\|\xi_1 - \xi_2\|_{TV}}{2}.$$

Hence,

$$\sup_A |\xi_1(A) - \xi_2(A)| \leq \frac{\|\xi_1 - \xi_2\|_{TV}}{2}. \tag{A.1}$$

Next, consider reverse inequality. Let H be a Jordan set of the signed measure $\xi_1 - \xi_2$. Then

$$\begin{aligned}\xi_1(H) - \xi_2(H) &= (\xi_1 - \xi_2)^+(H) - (\xi_1 - \xi_2)^-(H)\\ &= (\xi_1 - \xi_2)^+(H) - (-(\xi_1 - \xi_2)(H^c))\\ &= (\xi_1 - \xi_2)^+(\Sigma) - (-(\xi_1 - \xi_2)(\emptyset))\\ &= \frac{1}{2}\|\xi_1 - \xi_2\|_{TV}.\end{aligned}$$

Hence

$$\sup_{A} |\xi_1(H) - \xi_2(H)| \geq \frac{1}{2} \|\xi_1 - \xi_2\|_{TV}$$

and the proof is complete.

B

Extremum Measures

B.1. Proof of Theorem 3.2

For the following proof we employ lemma 3.5, which is also valid for Problem 3.2.

From Lemma 3.2, and Corollary 3.2, we know that for $D \leq D_{\max}$, where $D_{\max} = \ell_{\max}$, the average constraint holds with equality, that is

$$\sum_{i \in \Sigma} \ell_i \nu_i = \sum_{i \in \Sigma} \ell_i \xi_i^+ - \sum_{i \in \Sigma} \ell_i \xi_i^- + \sum_{i \in \Sigma} \ell_i \mu_i = D.$$

From Lemma 3.5, Part (a) and from Part (b), case 1, when equality conditions (3.48) and (3.50) are satisfied we have that

$$\ell_{\max} \left(\frac{\alpha}{2} \right) - \ell_{\min} \left(\frac{\alpha}{2} \right) + \sum_{i \in \Sigma} \ell_i \mu_i = D.$$

Solving the above equation with respect to α we get that

$$\alpha = \frac{2 \left(D - \sum_{i \in \Sigma} \ell_i \mu_i \right)}{\ell_{\max} - \ell_{\min}}. \tag{B.1}$$

Since (3.48a) is always satisfied, it remains to ensure that (3.50a) is also satisfied. By substituting (B.1) into (3.50a) and solving with respect to D we get that if $D \geq (\ell_{\max} - \ell_{\min}) \sum_{i \in \Sigma_0} \mu_i + \sum_{i \in \Sigma} \ell_i \mu_i$ then $R^+(D)$ is given by (3.67). Moreover, the optimal probabilities given by (3.68a) and (3.68b) are obtained from (3.48b) and (3.50b), respectively.

Lemma 3.5, Part (b), case 1, characterize the extremum solution for $\sum_{i\in\Sigma_0}\mu_i - \frac{\alpha}{2} \geq 0$. Next, the characterization of extremum solution when this condition is violated, that is, when $\sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\mu_i - \frac{\alpha}{2} \leq 0$ for any $k \in \{1,2,\ldots,r\}$, is discussed.
From Lemma 3.5, Part (b), case 2, the lower bound (3.51), holds with equality if conditions given by (3.52) are satisfied. Hence,

$$\ell_{\max}\left(\frac{\alpha}{2}\right) - \ell(\Sigma_k)\left(\frac{\alpha}{2} - \sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\mu_i\right) + \sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\ell_i\mu_i = D.$$

Solving the above equation with respect to α we get that

$$\alpha = \frac{2\left(D - \ell_{\max}\sum_{i\in\Sigma^0}\mu_i - \ell(\Sigma_k)\sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\mu_i - \sum_{j=k}^{r}\sum_{i\in\Sigma_j}\ell_i\mu_i\right)}{\ell_{\max} - \ell(\Sigma_k)}. \tag{B.2}$$

Substituting (B.2) into $\sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\mu_i - \frac{\alpha}{2} \leq 0$ and into (3.52c) and solving with respect to D we get that if

$$D \geq \ell_{\max}\left(\sum_{j=1}^{k}\sum_{i\in\Sigma_{j-1}}\mu_i + \sum_{i\in\Sigma^0}\mu_i\right) + \sum_{j=k}^{r}\sum_{i\in\Sigma_j}\ell_i\mu_i$$
$$D \leq \ell_{\max}\left(\sum_{j=0}^{k}\sum_{i\in\Sigma_j}\mu_i + \sum_{i\in\Sigma^0}\mu_i\right) + \sum_{j=k+1}^{r}\sum_{i\in\Sigma_j}\ell_i\mu_i$$

then $R^+(D)$ is given by (3.65). Moreover, the optimal probability on Σ_k given by (3.68c) is obtained from (3.52b).
For $D \in [D_{\max},\infty)$, is straightforward that, the extremum measure $\nu^*(\Sigma^0) = 1$ and $\nu^*(\Sigma\setminus\Sigma^0) = 0$, and hence $R^+(D) = 2(1-\mu(\Sigma^0))$.

B.2. Proof of Theorem 3.3

From Lemma 3.3, and Corollary 3.3, we know that for $R \leq R_{\max}$, where $R_{\max} = 2(1-\mu(\Sigma_0))$, the total variation constraint holds with equality, that is, $||\xi||_{TV} = R$. Let $\alpha = ||\xi||_{TV}$. From (3.70) and (3.71), Problem 3.3 is given by

$$D^+(R) = \sum_{i\in\Sigma}\ell_i\mu_i + \min_{\xi\in\widetilde{\mathbb{B}}_R(\mu)}\sum_{i\in\Sigma}\ell_i\xi_i \tag{B.3}$$

where $\xi \in \widetilde{\mathbb{B}}_R(\mu)$ is described by the constraints

$$\alpha \triangleq \sum_{i\in\Sigma}|\xi_i| = R, \quad \sum_{i\in\Sigma}\xi_i = 0, \quad 0 \leq \xi_i + \mu_i \leq 1, \quad \forall i \in \Sigma. \tag{B.4}$$

To minimize (B.3) we employ (3.44). It is obvious that a lower and an upper bound must be obtained for $\sum_{i\in\Sigma}\ell_i\xi_i^+$ and $\sum_{i\in\Sigma}\ell_i\xi_i^-$, respectively.

From Lemma 3.6, Part (a), the lower bound (3.81), holds with equality if conditions given by (3.84) are satisfied. Note that, (3.82a) is always satisfied and from (3.82b) we have that $\sum_{i\in\Sigma_0}\nu_i = \sum_{i\in\Sigma_0}\mu_i + \frac{\alpha}{2}$ and hence the optimal probability on Σ_0 is given by

$$\nu^*(\Sigma_0) \triangleq \sum_{i\in\Sigma_0}\nu_i^* = \sum_{i\in\Sigma_0}\mu_i + \frac{\alpha}{2}. \tag{B.5}$$

From Lemma 3.6, Part (b), case 1, the upper bound (3.83), holds with equality if conditions given by (3.84) are satisfied. Furthermore, from (3.84b) we have that $\sum_{i\in\Sigma^0}\nu_i = \sum_{i\in\Sigma^0}\mu_i - \frac{\alpha}{2}$ and condition (3.84a) must be satisfied, hence the optimal probability on Σ^0 is given by

$$\nu^*(\Sigma^0) \triangleq \sum_{i\in\Sigma^0}\nu_i^* = \Big(\sum_{i\in\Sigma^0}\mu_i - \frac{\alpha}{2}\Big)^+. \tag{B.6}$$

The extremum solution for any $R \leq R_{\max}$, under equality conditions (3.82) and (3.84) is given by

$$D^-(R) = \Big\{\ell_{\min} - \ell_{\max}\Big\}\frac{\alpha}{2} + \sum_{i\in\Sigma}\ell_i\mu_i. \tag{B.7}$$

Lemma 3.6, Part (b), case 1, characterize the extremum solution for $\sum_{i\in\Sigma^0}\mu_i - \frac{\alpha}{2} \geq 0$. Next, the characterization of extremum solution when this condition is violated, that is, when $\sum_{j=1}^k\sum_{i\in\Sigma^{j-1}}\mu_i - \frac{\alpha}{2} \leq 0$ for any $k \in \{1,2,\ldots,r\}$, is discussed.
From Lemma 3.6, Part (b), case 2, the upper bound (3.85), holds with equality if conditions given by (3.86) are satisfied. Furthermore, from (3.86b) we have that

$$\sum_{i\in\Sigma^k}\nu_i = \sum_{i\in\Sigma^k}\mu_i - \Big(\frac{\alpha}{2} - \sum_{j=1}^k\sum_{i\in\Sigma^{j-1}}\mu_i\Big), \tag{B.8}$$

and conditions $\frac{\alpha}{2} - \sum_{j=1}^k\sum_{i\in\Sigma^{j-1}}\mu_i \geq 0$ and (3.86c) must be satisfied, hence the optimal probability on Σ^k is given by

$$\nu^*(\Sigma^k) \triangleq \sum_{i\in\Sigma^k}\nu_i^* = \Big(\sum_{i\in\Sigma^k}\mu_i - \Big(\frac{\alpha}{2} - \sum_{j=1}^k\sum_{i\in\Sigma^{j-1}}\mu_i\Big)^+\Big)^+. \tag{B.9}$$

The extremum solution for any $R \leq R_{\max}$, under equality conditions (3.82) and (3.86) is given by

$$\begin{aligned}D^-(R) &= \sum_{i\in\Sigma}\ell_i\xi_i^+ - \sum_{i\in\Sigma}\ell_i\xi_i^- + \sum_{i\in\Sigma}\ell_i\mu_i\\ &= \ell_{\min}\Big(\frac{\alpha}{2}\Big) - \ell(\Sigma^k)\Big(\frac{\alpha}{2} - \sum_{j=1}^k\sum_{i\in\Sigma^{j-1}}\mu_i\Big) + \sum_{j=1}^k\sum_{i\in\Sigma^{j-1}}\ell_i\mu_i + \sum_{i\in\Sigma}\ell_i\mu_i.\end{aligned}$$

For $R \in [R_{\max}, 2]$, Lemma 3.3, states that $D^-(R)$ is constant. Indeed for $\alpha = ||\xi||_{TV} = R_{\max} = 2(1-\mu(\Sigma_0))$ equality conditions of Lemma 3.6, Part (a), become

$$\sum_{i\in\Sigma_0}\mu_i + \frac{\alpha}{2} = 1, \quad \sum_{i\in\Sigma_0}\xi_i^+ = \frac{\alpha}{2}, \quad \xi_i^+ = 0 \text{ for } i \in \Sigma\setminus\Sigma_0 \tag{B.10}$$

and hence

$$\sum_{i\in\Sigma\setminus\Sigma_0}\mu_i-\frac{\alpha}{2}=0,\qquad \sum_{i\in\Sigma\setminus\Sigma_0}\xi_i^-=\frac{\alpha}{2},\qquad \xi_i^-=0 \text{ for } i\in\Sigma_0. \tag{B.11}$$

Therefore $\sum_{i\in\Sigma\setminus\Sigma_0}\xi_i^-=\sum_{i\in\Sigma\setminus\Sigma_0}\mu_i$ and hence $\xi_i^-=\mu_i$ for all $i\in\Sigma\setminus\Sigma_0$. The extremum solution for any $R\in[R_{\max},2]$ is given by

$$\begin{aligned}D^-(R)&=\sum_{i\in\Sigma}\ell_i\xi_i^+-\sum_{i\in\Sigma}\ell_i\xi_i^-+\sum_{i\in\Sigma}\ell_i\mu_i\overset{(a)}{=}\sum_{i\in\Sigma_0}\ell_i\xi_i^+-\sum_{i\in\Sigma\setminus\Sigma_0}\ell_i\xi_i^-+\sum_{i\in\Sigma}\ell_i\mu_i\\&=\ell_{\min}\left(\frac{\alpha}{2}\right)-\sum_{i\in\Sigma\setminus\Sigma_0}\ell_i\mu_i+\sum_{i\in\Sigma}\ell_i\mu_i=\ell_{\min}\Big(1-\sum_{i\in\Sigma_0}\mu_i\Big)+\sum_{i\in\Sigma_0}\ell_i\mu_i=\ell_{\min}.\end{aligned}$$

where (a) follows from (B.10) and (B.11).

B.3. Weak Convergence of Probability Measures

Below, we give a brief description of concepts of weak convergence of probability measures which are relevant to our work.

1) Let (Σ,d_Σ) be a metric space with $\mathbb{B}(\Sigma)$ its Borel sets. Recall that a sequence $\{P_n: n=1,2,\ldots\}$ of probability measures on Σ converges weakly to a probability P on Σ if

$$\lim_{n\longrightarrow\infty}\int f(x)P^n(dx)=\int f(x)P(dx),\qquad \forall f\in BC(\Sigma).$$

This convergence is denoted by $P^n\xrightarrow{w}P$. Weak convergence defines a topology τ on the set of probability measures $\mathcal{M}_1(\Sigma)$, that is, $P^n\xrightarrow{w}P$ if and only if for each neighborhood $N(P)\in\tau$ of P then $P^n\in N(P)$ for n sufficiently large. Moreover, if the metric space (Σ,d_Σ) is separable, the topology τ is metrizable, and hence there exist a metric $d:\mathcal{M}_1(\Sigma)\times\mathcal{M}_1(\Sigma)\longmapsto\mathbb{R}$ which generates this topology. The Prohorov metric is such a metric on $\mathcal{M}_1(\Sigma)$.

In order to introduce it, we define for $A\subset\Sigma$, $\epsilon>0$, $A^{(\epsilon)}\triangleq\{x\in\Sigma: d(x,A)<\epsilon\}$. Then $d_P:\mathcal{M}_1(\Sigma)\times\mathcal{M}_1(\Sigma)\longmapsto\mathbb{R}$ defined by

$$d_P(Q,P)\triangleq\inf\{\epsilon>0: Q(F)\leq P(F^{(\epsilon)})+\epsilon,\ \ \forall\ \text{closed subset}\ \ F\subset\Sigma.$$

Moreover, if $d_P(P^n,P)\longrightarrow 0$ then $P^n\xrightarrow{w}P$.

An important property of the topology τ is that for (Σ,d_Σ) a separable metric space the set of probability measures with finite support denoted by $\mathcal{M}_1^{FS}(\Sigma)$ is dense in $\mathcal{M}_1(\Sigma)$, that is, the closure of $\mathcal{M}_1^{FS}(\Sigma)=\mathcal{M}_1(\Sigma)$ [11, Appendix III, Theorem 4]. Therefore, for each $P\in\mathcal{M}_1(\Sigma)$ there exists a sequence $\{P^n: n=1,2,\ldots\}$ of probability measures with finite support from $\mathcal{M}_1^{FS}(\Sigma)$ converging (i.e., to $P\in\mathcal{M}_1(\Sigma)$).

2) Consider a neighborhood $N(P) \triangleq \{Q \in \mathcal{M}_1(\Sigma) : Q(F_i) < P(F_i) + \epsilon, i = 1, \ldots, n\}$, where $F_i, i = 1, \ldots, n$ are closed subset of Σ, $\epsilon > 0$. Consider, a partition $\{A_1, A_2, \ldots, A_k\}$ generated by $\{F_1, \ldots, F_n\}$. Then for each $i = 1, \ldots, k$ there exists an $x_i \in A_i$ such that

$$f(x_i)P(A_i) \leq \int_{A_i} f(x_i)P(dx) \;\; \text{holds,} \;\;\; \text{for a measurable function} \;\; f : \Sigma \longmapsto \mathbb{R}.$$

On the points $\{x_i, \ldots, x_k\}$ put mass $\{P(A_i) : i = 1, \ldots, k\}$ and denote this probability measure by $Q^{FS} \in \mathcal{M}_1^{FS}(\Sigma)$. Then

$$\int f(x)Q^{FS}(dx) = \sum_{j=1}^{k} f(x_j)P(A_j).$$

C

Finite Horizon Dynamic Programming

C.1. Proof of Lemma 4.1

Let $\pi \in \Pi^{DM}_{0,n-1}$, which defines $\{g_j : j = 0,1,\ldots,n-1\}$. It follows that $\{x^g_j : j = 0,1,\ldots,n\}$ is a Markov process. Then

$$\begin{aligned} V^g_n(x^g_n) &= \alpha^n h_n(x^g_n) \\ &= \mathbb{E}_{\mathbf{Q}^\pi_\nu}\{\alpha^n h_n(x^g_n)|\mathcal{G}_{0,n}\} \\ &= \sup_{Q_n(\cdot|x_{n-1},u_{n-1})\in\mathbf{B}_{R_n}(Q^o_n)(x_{n-1},u_{n-1})} \mathbb{E}_{\mathbf{Q}^\pi_\nu}\{\alpha^n h_n(x^g_n)|\mathcal{G}_{0,n}\} \\ &= V_n(\mathcal{G}_{0,n}) \end{aligned}$$

so that it is true for $j = n$.

Suppose that (4.39) holds for $j+1, j+2, \ldots, n$. Then

$$
\begin{aligned}
V_j^g(x_j^g) &= \sup_{Q_{j+1}(\cdot|x_j^g,u_j^g)\in \mathbf{B}_{R_{j+1}}(Q_{j+1}^o)(x_j^g,u_j^g)} \mathbb{E}_{Q_{j+1}(\cdot|x_j,u_j)}\Big\{\alpha^j f_j(x_j^g, g_j(x_j^g)) + V_{j+1}^g(x_{j+1}^g)\Big\} \\
&= \sup_{Q_{j+1}(\cdot|x_j^g,u_j^g)\in \mathbf{B}_{R_{j+1}}(Q_{j+1}^o)(x_j^g,u_j^g)} \mathbb{E}_{\mathbf{Q}_\nu^\pi}\Big\{\alpha^j f_j(x_j^g, g_j(x_j^g)) + V_{j+1}^g(x_{j+1}^g)|\mathcal{G}_{0,j}\Big\} \\
&= \sup_{Q_{j+1}(\cdot|x_j^g,u_j^g)\in \mathbf{B}_{R_{j+1}}(Q_{j+1}^o)(x_j^g,u_j^g)} \mathbb{E}_{\mathbf{Q}_\nu^\pi}\Big\{\alpha^j f_j(x_j^g, g_j(x_j^g)) \\
&\quad + \sup_{\substack{Q_k(\cdot|x_{k-1},u_{k-1})\in \mathbf{B}_{R_k}(Q_k^o)(x_{k-1},u_{k-1})\\ k=j+2,\ldots,n}} \mathbb{E}_{\mathbf{Q}_\nu^\pi}\Big\{\sum_{k=j+1}^{n-1}\alpha^k f_k(x_k^g, g_k(x_k^g)) + \alpha^n h_n(x_n^g)|\mathcal{G}_{0,j+1}\Big\}|\mathcal{G}_{0,j}\Big\} \\
&= \sup_{\substack{Q_k(\cdot|x_{k-1},u_{k-1})\in \mathbf{B}_{R_k}(Q_k^o)(x_{k-1},u_{k-1})\\ k=j+1,\ldots,n}} \Bigg\{\mathbb{E}_{\mathbf{Q}_\nu^\pi}\Big\{\alpha^j f_j(x_j^g, g_j(x_j^g))|\mathcal{G}_{0,j}\Big\} \\
&\quad + \mathbb{E}_{\mathbf{Q}_\nu^\pi}\Big\{\mathbb{E}_{\mathbf{Q}_\nu^\pi}\Big\{\sum_{k=j+1}^{n-1}\alpha^k f_k(x_k^g, g_k(x_k^g)) + \alpha^n h_n(x_n^g)|\mathcal{G}_{0,j+1}\Big\}|\mathcal{G}_{0,j}\Big\}\Bigg\} \\
&= \sup_{\substack{Q_k(\cdot|x_{k-1},u_{k-1})\in \mathbf{B}_{R_k}(Q_k^o)(x_{k-1},u_{k-1})\\ k=j+1,\ldots,n}} \mathbb{E}_{\mathbf{Q}_\nu^\pi}\Big\{\sum_{k=j}^{n-1}\alpha^k f_k(x_k^g, g_k(x_k^g)) + \alpha^n h_n(x_n^g)|\mathcal{G}_{0,j}\Big\} \\
&= V_j(u_{[j,n-1]}^g, \mathcal{G}_{0,j}) \\
&= \sup_{\substack{Q_k(\cdot|x_{k-1},u_{k-1})\in \mathbf{B}_{R_k}(Q_k^o)(x_{k-1},u_{k-1})\\ k=j+1,\ldots,n}} \mathbb{E}_{\mathbf{Q}_\nu^\pi}\Big\{\sum_{k=j}^{n-1}\alpha^k f_k(x_k^g, g_k(x_k^g)) + \alpha^n h_n(x_n^g)|x_j\Big\}
\end{aligned}
$$

because $\{x_j^g : j = 0, 1, \ldots, n\}$ is a Markov process.

C.2. Proof of Lemma 4.2

Let $\pi \in \Pi_{0,n-1}^D$ be arbitrary. Then

$$
\begin{aligned}
V_n^g(x_n^g) &\le \alpha^n h_n(x_n^g) \\
&= \mathbb{E}_{\mathbf{Q}_\nu^\pi}\{\alpha^n h_n(x_n^g)|\mathcal{G}_{0,n}\} \\
&= \sup_{Q_n(\cdot|x_{n-1},u_{n-1})\in \mathbf{B}_{R_n}(Q_n^o)(x_{n-1},u_{n-1})} \mathbb{E}_{\mathbf{Q}_\nu^\pi}\{\alpha^n h_n(x_n^g)|\mathcal{G}_{0,n}\} \\
&= V_n(\mathcal{G}_{0,n})
\end{aligned}
$$

so that it is true for $j=n$.

Suppose that (4.42) holds for $j+1, j+2, \ldots, n$. Then

$$\begin{aligned}
V_j^g(x_j^g) &\leq \sup_{Q_{j+1}(\cdot|x_j^g,u_j^g)\in \mathbf{B}_{R_{j+1}}(Q_{j+1}^o)(x_j^g,u_j^g)} \mathbb{E}_{Q_{j+1}(\cdot|x_j,u_j)}\Big\{\alpha^j f_j(x_j^g,u_j^g) + V_{j+1}^g(x_{j+1}^g)\Big\}\\
&= \sup_{Q_{j+1}(\cdot|x_j^g,u_j^g)\in \mathbf{B}_{R_{j+1}}(Q_{j+1}^o)(x_j^g,u_j^g)} \mathbb{E}_{\mathbf{Q}_\nu^\pi}\Big\{\alpha^j f_j(x_j^g,u_j^g) + V_{j+1}^g(x_{j+1}^g)|\mathcal{G}_{0,j}\Big\}\\
&\leq \sup_{Q_{j+1}(\cdot|x_j^g,u_j^g)\in \mathbf{B}_{R_{j+1}}(Q_{j+1}^o)(x_j^g,u_j^g)} \mathbb{E}_{\mathbf{Q}_\nu^\pi}\Big\{\alpha^j f_j(x_j^g,u_j^g)\\
&\quad + \sup_{\substack{Q_k(\cdot|x_{k-1},u_{k-1})\in \mathbf{B}_{R_k}(Q_k^o)(x_{k-1},u_{k-1})\\ k=j+2,\ldots,n}} \mathbb{E}_{\mathbf{Q}_\nu^\pi}\Big\{\sum_{k=j+1}^{n-1}\alpha^k f_k(x_k^g,u_k^g) + \alpha^n h_n(x_n^g)|\mathcal{G}_{0,j+1}\Big\}|\mathcal{G}_{0,j}\Big\}\\
&= \sup_{\substack{Q_k(\cdot|x_{k-1},u_{k-1})\in \mathbf{B}_{R_k}(Q_k^o)(x_{k-1},u_{k-1})\\ k=j+1,\ldots,n}} \Bigg\{\mathbb{E}_{\mathbf{Q}_\nu^\pi}\Big\{\alpha^j f_j(x_j^g,u_j^g)|\mathcal{G}_{0,j}\Big\}\\
&\quad + \mathbb{E}_{\mathbf{Q}_\nu^\pi}\Big\{\mathbb{E}_{\mathbf{Q}_\nu^\pi}\Big\{\sum_{k=j+1}^{n-1}\alpha^k f_k(x_k^g,u_k^g) + \alpha^n h_n(x_n^g)|\mathcal{G}_{0,j+1}\Big\}|\mathcal{G}_{0,j}\Big\}\Bigg\}\\
&= \sup_{\substack{Q_k(\cdot|x_{k-1},u_{k-1})\in \mathbf{B}_{R_k}(Q_k^o)(x_{k-1},u_{k-1})\\ k=j+1,\ldots,n}} \mathbb{E}_{\mathbf{Q}_\nu^\pi}\Big\{\sum_{k=j}^{n-1}\alpha^k f_k(x_k^g,u_k^g) + \alpha^n h_n(x_n^g)|\mathcal{G}_{0,j}\Big\}\\
&= V_j(u_{[j,n-1]}^g, \mathcal{G}_{0,j}).
\end{aligned}$$

D

Infinite Horizon Dynamic Programming

D.1. Proof of Lemma 5.4

By Proposition 5.1 (c), there exists a $V(g^*, \cdot) : \mathcal{X} \longmapsto \mathbb{R}$ and $J(\pi^*)$ such that for all $x \in \mathcal{X}$

$$J(\pi^*) + V(g^*, x) = f(x, g^*(x)) + \sum_{z \in \mathcal{X}} Q^*(z|x, g^*(x))V(g^*, z). \tag{D.1}$$

Then, for all $x \in \mathcal{X}$

$$J(\pi^*) + V(g^*, x) \geq \min_{u \in \mathcal{U}} \Big\{ f(x, u) + \sum_{z \in \mathcal{X}} Q^*(z|x, u)V(g^*, z) \Big\}.$$

Define $g_1 : \mathcal{X} \longmapsto \mathcal{U}$ as

$$g_1(x) = \operatorname*{argmin}_{u \in \mathcal{U}} \Big\{ f(x, u) + \sum_{z \in \mathcal{X}} Q^*(z|x, u)V(g^*, z) \Big\}.$$

Suppose that for some $x_2 \in \mathcal{X}$ strict inequality holds in (D.1), then

$$J(\pi^*) + V(g^*, x) > \min_{u \in \mathcal{U}} \Big\{ f(x_2, u) + \sum_{z \in \mathcal{X}} Q^*(z|x_2, u)V(g^*, z) \Big\}. \tag{D.2}$$

Then multiplying (D.2) by $q(g_1)(x_0) > 0$ and summing over $x_0 \in \mathcal{X}$ yields

$$\begin{aligned} J(\pi^*) + \sum_{x_0 \in \mathcal{X}} q(g_1) V(g^*, x_0) \\ > \min \Big\{ \sum_{x_0 \in \mathcal{X}} q(g_1)(x_0) f(x_0, u) + \sum_{x_0 \in \mathcal{X}} q(g_1)(x_0) \sum_{z \in \mathcal{X}} Q^*(z|x_0, u) V(g^*, z) \Big\} \\ = \sum_{x_0 \in \mathcal{X}} q(g_1)(x_0) f(x_0, g_1(x_0)) + \sum_{x_0 \in \mathcal{X}} q(g_1)(x_0) \sum_{z \in \mathcal{X}} Q^*(z|x_0, g_1(x_0)) V(g^*, z) \\ = J(g_1) + \sum_{z \in \mathcal{X}} q(g_1) V(g^*, z), \quad \text{by Proposition 5.1 (a)} \end{aligned}$$

which gives $J(\pi^*) > J(g_1)$, contradicting assumption 2. Hence, equality holds in (D.1), for every $x \in \mathcal{X}$.

D.2. Analytic Solution of Example 5.4.1

To obtain an optimal stationary policy of the infinite horizon minimax problem for discounted cost, policy iteration algorithm 5.3 is applied.

1. Initialization. Solve the equation $f(g_0)+\alpha Q^o(g_0)V_{Q^o}(g_0)=V_{Q^o}(g_0)$, for $V_{Q^o}(g_0) \in \mathbb{R}^3$, or,

$$\begin{pmatrix} f(1, g_0(1)) \\ f(2, g_0(2)) \\ f(3, g_0(3)) \end{pmatrix} + \alpha \begin{pmatrix} q^o_{11}(g_0(1)) & q^o_{12}(g_0(1)) & q^o_{13}(g_0(1)) \\ q^o_{21}(g_0(2)) & q^o_{22}(g_0(2)) & q^o_{23}(g_0(2)) \\ q^o_{31}(g_0(3)) & q^o_{32}(g_0(3)) & q^o_{33}(g_0(3)) \end{pmatrix} \begin{pmatrix} V_{Q^o}(1) \\ V_{Q^o}(2) \\ V_{Q^o}(3) \end{pmatrix} = \begin{pmatrix} V_{Q^o}(1) \\ V_{Q^o}(2) \\ V_{Q^o}(3) \end{pmatrix}$$

which is given by

$$\begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} + \frac{0.9}{9} \begin{pmatrix} 3 & 1 & 5 \\ 4 & 2 & 3 \\ 4 & 1 & 4 \end{pmatrix} V_{Q^o}(g_0) = V_{Q^o}(g_0) \Longrightarrow V_{Q^o}(g_0) \triangleq \begin{pmatrix} V_{Q^o}(1) \\ V_{Q^o}(2) \\ V_{Q^o}(3) \end{pmatrix} = \begin{pmatrix} 12.42 \\ 13.93 \\ 10.60 \end{pmatrix}.$$

Note that, $V_{Q^o} \triangleq \{V_{Q^o}(1), V_{Q^o}(2), V_{Q^o}(3)\}$, $|\mathcal{X}| = 3$, and hence

$$\begin{aligned} \mathcal{X}^+ &\triangleq \{x \in \mathcal{X} : V_{Q^o}(x) = \max\{V_{Q^o}(x) : x \in \mathcal{X}\}\} = \{x \in \mathcal{X} : V_{Q^o}(x) = V_{Q^o}(2)\} = \{2\}, \\ \mathcal{X}^- &\triangleq \{x \in \mathcal{X} : V_{Q^o}(x) = \min\{V_{Q^o}(x) : x \in \mathcal{X}\}\} = \{x \in \mathcal{X} : V_{Q^o}(x) = V_{Q^o}(3)\} = \{3\}, \\ \mathcal{X}_1 &\triangleq \Big\{x \in \mathcal{X} : V_{Q^o}(x) = \min\{V_{Q^o}(\alpha) : \alpha \in \mathcal{X} \setminus \mathcal{X}^+ \cup \mathcal{X}^-\}\Big\} = \{1\}. \end{aligned}$$

Once the partition is been identified, (5.16)-(5.18) is applied to obtain

$$Q^*(u_1) = \begin{pmatrix} \left(q_{11}^o(u_1) - \left(\frac{R}{2} - q_{13}^o(u_1)\right)^+\right)^+ & \min\left(1, q_{12}^o(u_1) + \frac{R}{2}\right) & \left(q_{13}^o(u_1) - \frac{R}{2}\right)^+ \\ \left(q_{21}^o(u_1) - \left(\frac{R}{2} - q_{23}^o(u_1)\right)^+\right)^+ & \min\left(1, q_{22}^o(u_1) + \frac{R}{2}\right) & \left(q_{23}^o(u_1) - \frac{R}{2}\right)^+ \\ \left(q_{31}^o(u_1) - \left(\frac{R}{2} - q_{33}^o(u_1)\right)^+\right)^+ & \min\left(1, q_{32}^o(u_1) + \frac{R}{2}\right) & \left(q_{33}^o(u_1) - \frac{R}{2}\right)^+ \end{pmatrix}$$

$$= \frac{1}{9}\begin{pmatrix} 3 & 4 & 2 \\ 4 & 5 & 0 \\ 0 & 9 & 0 \end{pmatrix} \tag{D.3}$$

and

$$Q^*(u_2) = \begin{pmatrix} \left(q_{11}^o(u_2) - \left(\frac{R}{2} - q_{13}^o(u_2)\right)^+\right)^+ & \min\left(1, q_{12}^o(u_2) + \frac{R}{2}\right) & \left(q_{13}^o(u_2) - \frac{R}{2}\right)^+ \\ \left(q_{21}^o(u_2) - \left(\frac{R}{2} - q_{23}^o(u_2)\right)^+\right)^+ & \min\left(1, q_{22}^o(u_2) + \frac{R}{2}\right) & \left(q_{23}^o(u_2) - \frac{R}{2}\right)^+ \\ \left(q_{31}^o(u_2) - \left(\frac{R}{2} - q_{33}^o(u_2)\right)^+\right)^+ & \min\left(1, q_{32}^o(u_2) + \frac{R}{2}\right) & \left(q_{33}^o(u_2) - \frac{R}{2}\right)^+ \end{pmatrix}$$

$$= \frac{1}{9}\begin{pmatrix} 1 & 5 & 3 \\ 4 & 5 & 0 \\ 4 & 4 & 1 \end{pmatrix}. \tag{D.4}$$

Solve the equation $f(g_0) + \alpha Q^*(g_0)V_{Q^*}(g_0) = V_{Q^*}(g_0)$, for $V_{Q^*}(g_0) \in \mathbb{R}^3$, or,

$$\begin{pmatrix} f(1, g_0(1)) \\ f(2, g_0(2)) \\ f(3, g_0(3)) \end{pmatrix} + \alpha \begin{pmatrix} q_{11}^*(g_0(1)) & q_{12}^*(g_0(1)) & q_{13}^*(g_0(1)) \\ q_{21}^*(g_0(2)) & q_{22}^*(g_0(2)) & q_{23}^*(g_0(2)) \\ q_{31}^*(g_0(3)) & q_{32}^*(g_0(3)) & q_{33}^*(g_0(3)) \end{pmatrix} \begin{pmatrix} V_{Q^*}(1) \\ V_{Q^*}(2) \\ V_{Q^*}(3) \end{pmatrix} = \begin{pmatrix} V_{Q^*}(1) \\ V_{Q^*}(2) \\ V_{Q^*}(3) \end{pmatrix}$$

which is given by

$$\begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} + \frac{0.9}{9}\begin{pmatrix} 3 & 4 & 2 \\ 4 & 5 & 0 \\ 4 & 4 & 1 \end{pmatrix} V_{Q^*}(g_0) = V_{Q^*}(g_0) \Longrightarrow V_{Q^*}(g_0) \triangleq \begin{pmatrix} V_{Q^*}(1) \\ V_{Q^*}(2) \\ V_{Q^*}(3) \end{pmatrix} = \begin{pmatrix} 22.42 \\ 23.93 \\ 20.60 \end{pmatrix}.$$

2. Let $m = 1$. (a) Determine $g_1 \in \mathbb{R}^3$ such that

$$f(g_1) + \alpha Q^*(g_1)V_{Q^*}(g_0) = \min_{g \in \mathbb{R}^3} \{f(g) + \alpha Q^*(g)V_{Q^*}(g_0)\}$$
$$= \begin{pmatrix} \min\{22.42, 20.88\} \\ \min\{21.93, 23.93\} \\ \min\{24.53, 20.60\} \end{pmatrix} = \begin{pmatrix} 20.88 \\ 21.93 \\ 20.60 \end{pmatrix}.$$

Hence, $g_1(1) = u_2$, $g_1(2) = u_1$ and $g_1(3) = u_2$.

(b) Solve the equation $f(g_1) + \alpha Q^o(g_1)V_{Q^o}(g_1) = V_{Q^o}(g_1)$, for $V_{Q^o}(g_1) \in \mathbb{R}^3$, or,

$$\begin{pmatrix} f(1, g_1(1)) \\ f(2, g_1(2)) \\ f(3, g_1(3)) \end{pmatrix} + \alpha \begin{pmatrix} q^o_{11}(g_1(1)) & q^o_{12}(g_1(1)) & q^o_{13}(g_1(1)) \\ q^o_{21}(g_1(2)) & q^o_{22}(g_1(2)) & q^o_{23}(g_1(2)) \\ q^o_{31}(g_1(3)) & q^o_{32}(g_1(3)) & q^o_{33}(g_1(3)) \end{pmatrix} \begin{pmatrix} V_{Q^o}(1) \\ V_{Q^o}(2) \\ V_{Q^o}(3) \end{pmatrix} = \begin{pmatrix} V_{Q^o}(1) \\ V_{Q^o}(2) \\ V_{Q^o}(3) \end{pmatrix}$$

which is given by

$$\begin{pmatrix} 0.5 \\ 1 \\ 0 \end{pmatrix} + \frac{0.9}{9} \begin{pmatrix} 1 & 2 & 6 \\ 4 & 2 & 3 \\ 4 & 1 & 4 \end{pmatrix} V_{Q^o}(g_1) = V_{Q^o}(g_1) \Longrightarrow V_{Q^o}(g_1) \triangleq \begin{pmatrix} V_{Q^o}(1) \\ V_{Q^o}(2) \\ V_{Q^o}(3) \end{pmatrix} = \begin{pmatrix} 3.46 \\ 4.10 \\ 2.99 \end{pmatrix}.$$

Therefore, $\mathcal{X}^+ = \{2\}$, $\mathcal{X}^- = \{3\}$ and $\mathcal{X}_1 = \{1\}$. Since the partition is the same as in $m = 0$ then $Q^*(u_1)$, $Q^*(u_2)$ are the same as (D.3) and (D.4), respectively.

Solve the equation $b(g_1) + \alpha Q^*(g_1)V_{Q^*}(g_1) = V_{Q^*}(g_1)$, for $V_{Q^*}(g_1) \in \mathbb{R}^3$, or,

$$\begin{pmatrix} f(1, g_1(1)) \\ f(2, g_1(2)) \\ f(3, g_1(3)) \end{pmatrix} + \alpha \begin{pmatrix} q^*_{11}(g_1(1)) & q^*_{12}(g_1(1)) & q^*_{13}(g_1(1)) \\ q^*_{21}(g_1(2)) & q^*_{22}(g_1(2)) & q^*_{23}(g_1(2)) \\ q^*_{31}(g_1(3)) & q^*_{32}(g_1(3)) & q^*_{33}(g_1(3)) \end{pmatrix} \begin{pmatrix} V_{Q^*}(1) \\ V_{Q^*}(2) \\ V_{Q^*}(3) \end{pmatrix} = \begin{pmatrix} V_{Q^*}(1) \\ V_{Q^*}(2) \\ V_{Q^*}(3) \end{pmatrix}$$

which is given by

$$\begin{pmatrix} 0.5 \\ 1 \\ 0 \end{pmatrix} + \frac{0.9}{9} \begin{pmatrix} 1 & 5 & 3 \\ 4 & 5 & 0 \\ 4 & 4 & 1 \end{pmatrix} V_{Q^*}(g_1) = V_{Q^*}(g_1) \Longrightarrow V_{Q^*}(g_1) \triangleq \begin{pmatrix} V_{Q^*}(1) \\ V_{Q^*}(2) \\ V_{Q^*}(3) \end{pmatrix} = \begin{pmatrix} 6.79 \\ 7.43 \\ 6.32 \end{pmatrix}.$$

Note that

$$\min_{g \in \mathbb{R}^3} \{f(g) + \alpha Q^*(g)V_{Q^*}(g_0)\} = \begin{pmatrix} 20.88 \\ 21.93 \\ 20.60 \end{pmatrix} < \begin{pmatrix} 22.42 \\ 23.93 \\ 20.60 \end{pmatrix} = V_{Q^*}(g_0). \tag{D.5}$$

3. Let $m = 2$. (a) Determine $g_2 \in \mathbb{R}^3$ such that

$$\begin{aligned} f(g_2) + \alpha Q^*(g_2)V_{Q^*}(g_1) &= \min_{g \in \mathbb{R}^3} \{f(g) + \alpha Q^*(g)V_{Q^*}(g_1)\} \\ &= \begin{pmatrix} \min\{8.27, 6.79\} \\ \min\{7.43, 9.43\} \\ \min\{9.69, 6.32\} \end{pmatrix} = \begin{pmatrix} 6.79 \\ 7.43 \\ 6.32 \end{pmatrix}. \end{aligned}$$

Hence, $g_2(1) = u_2$, $g_2(2) = u_1$ and $g_2(3) = u_2$. Note that, $g_2 = g_1$.

(b) Note that the solution of (5.28) is such that $V_{Q^o}(g_2) = V_{Q^o}(g_1)$, and hence the solution of (5.29) is such that $V_{Q^*}(g_2) = V_{Q^*}(g_1)$. Furthermore, condition (5.26) is satisfied since

$$\min_{g \in \mathbb{R}^3} \{f(g) + \alpha Q^*(g) V_{Q^*}(g_1)\} = \begin{pmatrix} 6.79 \\ 7.43 \\ 6.32 \end{pmatrix} = V_{Q^*}(g_1).$$

Hence, policy iteration algorithm has converged and the optimal policy and value are

$$g^* = g_2 \triangleq \begin{pmatrix} g_2(1) \\ g_2(2) \\ g_2(3) \end{pmatrix} = \begin{pmatrix} u_2 \\ u_1 \\ u_2 \end{pmatrix}, \quad V_{Q^*}(g^*) = V_{Q^*}(g_2) \triangleq \begin{pmatrix} V_{Q^*}(1) \\ V_{Q^*}(2) \\ V_{Q^*}(3) \end{pmatrix} = \begin{pmatrix} 6.79 \\ 7.43 \\ 6.32 \end{pmatrix}.$$

E

Markov Process Approximation

Before we proceed with the proof of Theorem 6.1, we give the following Lemma in which lower and upper bounds, which are achievable, are obtained.

E.1. Upper and Lower Bounds

Lemma E.1. *(a) Upper Bound.*

$$\sum_{j\in\mathcal{X}} \ell_j \Xi^{+}_{ij} \mu_i \leq \ell_{\max} \left(\frac{\alpha_i \mu_i}{2} \right). \tag{E.1}$$

The bound holds with equality if

$$\sum_{j\in\mathcal{X}^0} P_{ij} + \frac{\alpha_i}{2} \leq 1, \quad \sum_{j\in\mathcal{X}^0} \Xi^{+}_{ij} = \frac{\alpha_i}{2}, \quad \Xi^{+}_{ij} = 0,\ \forall j \in \mathcal{X} \backslash \mathcal{X}^0. \tag{E.2}$$

(b) Lower Bound. Case 1. If $\sum_{j\in\mathcal{X}_0} P_{ij} - \frac{\alpha_i}{2} \geq 0$ *then*

$$\sum_{j\in\mathcal{X}} \ell_j \Xi^{-}_{ij} \mu_i \geq \ell_{\min} \left(\frac{\alpha_i \mu_i}{2} \right). \tag{E.3}$$

The bound holds with equality if

$$\sum_{j\in\mathcal{X}_0} P_{ij} - \frac{\alpha_i}{2} \geq 0, \quad \sum_{j\in\mathcal{X}_0} \Xi^{-}_{ij} = \frac{\alpha_i}{2}, \quad \Xi^{-}_{ij} = 0,\ \forall j \in \mathcal{X} \backslash \mathcal{X}_0. \tag{E.4}$$

Case 2. If $\sum_{s=1}^{k}\sum_{j\in\mathcal{X}_{s-1}} P_{ij} - \frac{\alpha_i}{2} \leq 0$ *for any* $k \in \{1,2,\ldots,r\}$ *then*

$$\sum_{j\in\mathcal{X}} \ell_j \Xi^-_{ij}\mu_i \geq \ell(\mathcal{X}_k)\Big(\frac{\alpha_i\mu_i}{2} - \sum_{s=1}^{k}\sum_{j\in\mathcal{X}_{s-1}} P_{ij}\mu_i\Big) + \sum_{s=1}^{k}\sum_{j\in\mathcal{X}_{s-1}} \ell_j P_{ij}\mu_i. \tag{E.5}$$

Moreover, equality holds if

$$\sum_{j\in\mathcal{X}_{s-1}} \Xi^-_{ij} = \sum_{j\in\mathcal{X}_{s-1}} P_{ij}, \quad \textit{for all} \quad s = 1,2,\ldots,k, \tag{E.6a}$$

$$\sum_{j\in\mathcal{X}_k} \Xi^-_{ij} = \Big(\frac{\alpha_i}{2} - \sum_{s=1}^{k}\sum_{j\in\mathcal{X}_{s-1}} P_{ij}\Big), \tag{E.6b}$$

$$\sum_{s=0}^{k}\sum_{j\in\mathcal{X}_s} P_{ij} - \frac{\alpha_i}{2} \geq 0, \tag{E.6c}$$

$$\Xi^-_{ij} = 0 \;\; \textit{for all} \;\; j \in \mathcal{X} \setminus \mathcal{X}_0 \cup \mathcal{X}_1 \cup \ldots \cup \mathcal{X}_k. \tag{E.6d}$$

Proof. Part (a): First, we show that inequality (E.1) holds.

$$\sum_{j\in\mathcal{X}} \ell_j \Xi^+_{ij}\mu_i \leq \ell_{\max}\mu_i \sum_{j\in\mathcal{X}} \Xi^+_{ij} = \ell_{\max}\left(\frac{\alpha_i\mu_i}{2}\right).$$

Next, we show that under the stated conditions (E.2) equality holds.

$$\begin{aligned}\sum_{j\in\mathcal{X}} \ell_j \Xi^+_{ij}\mu_i &= \sum_{j\in\mathcal{X}^0} \ell_j \Xi^+_{ij}\mu_i + \sum_{j\in\mathcal{X}\setminus\mathcal{X}^0} \ell_j \Xi^+_{ij}\mu_i \\ &= \ell_{\max}\mu_i \sum_{j\in\mathcal{X}^0} \Xi^+_{ij} + \sum_{j\in\mathcal{X}\setminus\mathcal{X}^0} \ell_j \Xi^+_{ij}\mu_i = \ell_{\max}\left(\frac{\alpha_i\mu_i}{2}\right).\end{aligned}$$

Part (b), case 1: First, we show that inequality (E.3) holds.

$$\sum_{j\in\mathcal{X}} \ell_j \Xi^-_{ij}\mu_i \geq \ell_{\min}\mu_i \sum_{j\in\mathcal{X}} \Xi^-_{ij} = \ell_{\min}\left(\frac{\alpha_i\mu_i}{2}\right).$$

Next, we show that under the stated conditions (E.4) equality holds.

$$\begin{aligned}\sum_{j\in\mathcal{X}} \ell_j \Xi^-_{ij}\mu_i &= \sum_{j\in\mathcal{X}_0} \ell_j \Xi^-_{ij}\mu_i + \sum_{j\in\mathcal{X}\setminus\mathcal{X}_0} \ell_j \Xi^-_{ij}\mu_i \\ &= \ell_{\min}\mu_i \sum_{j\in\mathcal{X}_0} \Xi^-_{ij} + \sum_{j\in\mathcal{X}\setminus\mathcal{X}_0} \ell_j \Xi^+_{ij}\mu_i = \ell_{\min}\left(\frac{\alpha_i\mu_i}{2}\right).\end{aligned}$$

Part (b), case 2: First, we show that inequality (E.5) holds. Consider any $k \in \{1,2,\ldots,r\}$. From Part (b), case 1, we have that

$$\begin{aligned}\sum_{j\in\mathcal{X}\setminus\cup_{s=1}^{k}\mathcal{X}_{s-1}} \ell_j \Xi^-_{ij}\mu_i &\geq \min_{j\in\mathcal{X}\setminus\cup_{s=1}^{k}\mathcal{X}_{s-1}} \ell_j \sum_{j\in\mathcal{X}\setminus\cup_{s=1}^{k}\mathcal{X}_{s-1}} \Xi^-_{ij}\mu_i \\ &= \ell(\mathcal{X}_k)\sum_{j\in\mathcal{X}\setminus\cup_{s=1}^{k}\mathcal{X}_{s-1}} \Xi^-_{ij}\mu_i = \ell(\mathcal{X}_k)\Big(\sum_{j\in\mathcal{X}} \Xi^-_{ij}\mu_i - \sum_{s=1}^{k}\sum_{j\in\mathcal{X}_{s-1}} \Xi^-_{ij}\mu_i\Big).\end{aligned}$$

Hence,

$$\sum_{j\in\mathcal{X}} \ell_j \Xi_{ij}^- \mu_i - \sum_{s=1}^{k} \sum_{j\in\mathcal{X}_{s-1}} \ell_j \Xi_{ij}^- \mu_i \geq \ell(\mathcal{X}_k)\Big(\frac{\alpha_i\mu_i}{2} - \sum_{s=1}^{k} \sum_{j\in\mathcal{X}_{s-1}} P_{ij}\mu_i\Big),$$

which implies

$$\sum_{j\in\mathcal{X}} \ell_j \Xi_{ij}^- \mu_i \geq \ell(\mathcal{X}_k)\Big(\frac{\alpha_i\mu_i}{2} - \sum_{s=1}^{k} \sum_{j\in\mathcal{X}_{s-1}} P_{ij}\mu_i\Big) + \sum_{s=1}^{k} \sum_{j\in\mathcal{X}_{s-1}} \ell_j P_{ij}\mu_i.$$

Next, we show under the stated conditions (E.6) that equality holds.

$$\begin{aligned}
\sum_{j\in\mathcal{X}} \ell_j \Xi_{ij}^- \mu_i &= \sum_{s=1}^{k} \sum_{j\in\mathcal{X}_{s-1}} \ell_j \Xi_{ij}^- \mu_i + \sum_{j\in\mathcal{X}_k} \ell_j \Xi_{ij}^- \mu_i + \sum_{j\in\mathcal{X}\setminus\cup_{s=0}^{k}\mathcal{X}_s} \ell_j \Xi_{ij}^- \mu_i \\
&= \sum_{s=1}^{k} \ell(\mathcal{X}_{s-1}) \sum_{j\in\mathcal{X}_{s-1}} \Xi_{ij}^- \mu_i + \ell(\mathcal{X}_k) \sum_{j\in\mathcal{X}_k} \Xi_{ij}^- \mu_i \\
&= \sum_{s=1}^{k} \sum_{j\in\mathcal{X}_{s-1}} \ell_j P_{ij}\mu_i + \ell(\mathcal{X}_k)\Big(\frac{\alpha_i\mu_i}{2} - \sum_{s=1}^{k} \sum_{j\in\mathcal{X}_{s-1}} P_{ij}\mu_i\Big).
\end{aligned}$$

■

E.2. Proof of Theorem 6.1

We provide the main steps for the derivation of Theorem 6.1, since the methodology followed for solving Problem 6.1 is similar to the one followed in Chapter 3, Theorem 3.1. In particular, for a fixed $i \in \mathcal{X}$, the solution of Problem 6.1 is given by (6.19) and (6.20), with proper substitution of $\nu^* \rightarrow \Phi^\dagger$ and $\mu \rightarrow P$.

From (6.15), the pay-off of Problem 6.1 is given by

$$\sum_{i\in\mathcal{X}} \sum_{j\in\mathcal{X}} \ell_j P_{ij}\mu_i + \max_{\Xi_{ij}} \sum_{i\in\mathcal{X}} \sum_{j\in\mathcal{X}} \ell_j \Xi_{ij}\mu_i. \tag{E.7}$$

To maximize (E.7) we employ the fact that Ξ is a signed measure satisfying (6.16). It is obvious that for each $i \in \mathcal{X}$ an upper and a lower bound must be obtained for $\sum_{j\in\mathcal{X}} \ell_j \Xi_{ij}^+ \mu_i$ and $\sum_{j\in\mathcal{X}} \ell_j \Xi_{ij}^- \mu_i$, respectively. Before proceeding with the derivation of the optimal transition probabilities $\Phi^\dagger$ based on upper and lower bounds, we discuss first the solution behavior in terms of the TV constraint given by (6.14), that is $\sum_{i\in\mathcal{X}} \alpha_i\mu_i \leq R$.

Let α_i, $\forall i \in \mathcal{X}$, to be given by (6.18d) (see Chapter 3, Lemma 3.1 and Corollary 3.1); then, it can be verified that for $R \leq R_{\max,i}$, $\forall i \in \mathcal{X}$, the TV constraint holds with equality, and also that as R increases (i.e., $R_{\max,i} \leq R \leq R_{\max,i+1}$, $\forall i, i+1 \in \mathcal{X}$), the TV constraint holds with inequality. However, the solution of Problem 6.1 with respect to the specific $i \in \mathcal{X}$ for which $R \geq R_{\max,i}$ is constant and hence the overall solution of Problem 6.1 is not affected.

Finally, for values of $R \geq R_{\max,i}$, $\forall i \in \mathcal{X}$ the overall solution of Problem 6.1 is constant, in particular, is equal to $\ell_{\max}$. The relation of TV constraint $\sum_{i\in\mathcal{X}} \alpha_i \mu_i$ with the TV parameter R, is depicted in Fig.E.1. Next we proceed with the derivation of (6.18).

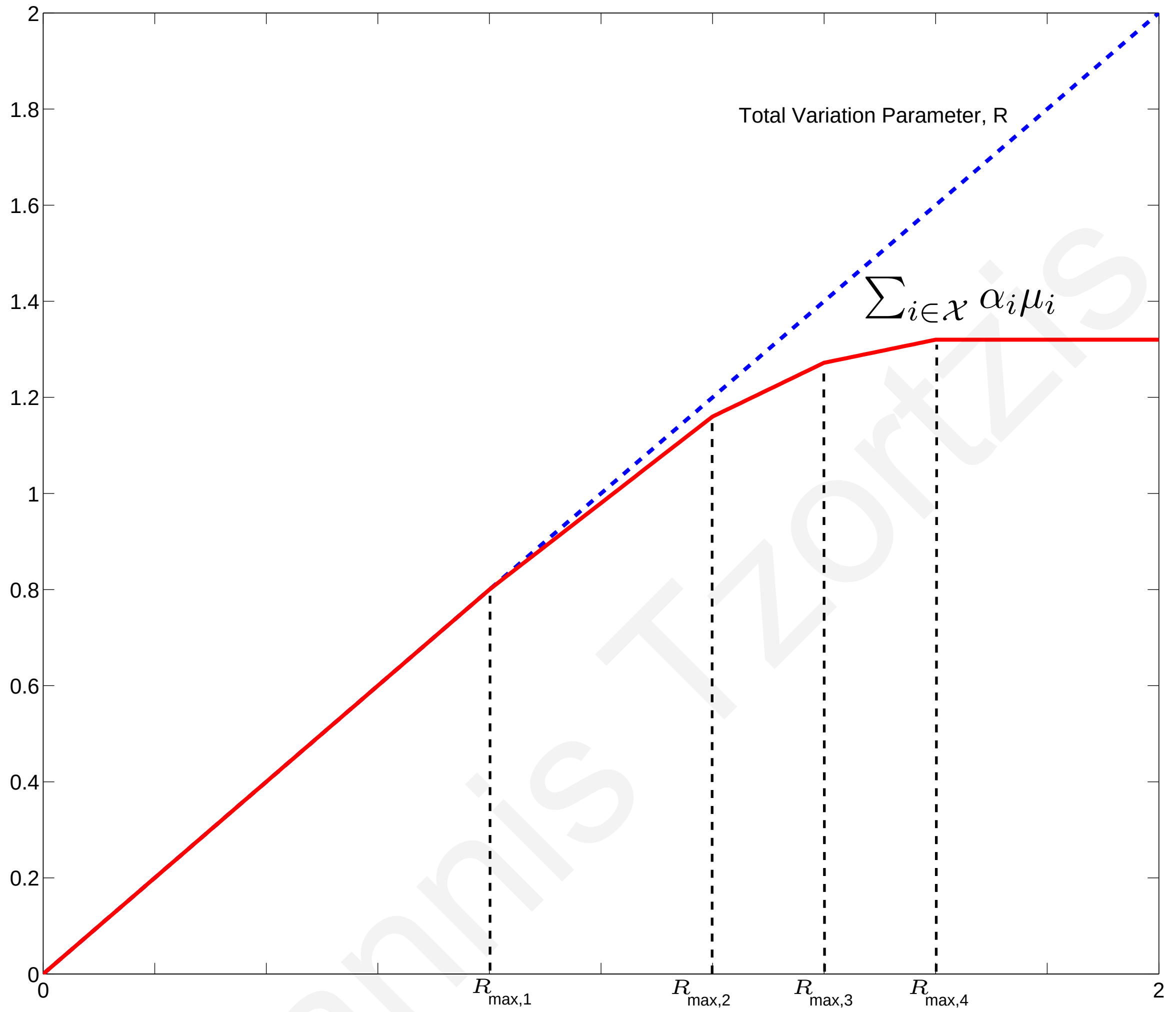


Figure E.1.: Total Variation Constraint vs. Total Variation Parameter

From Lemma E.1, Part (a), the upper bound (E.1), holds with equality if conditions given by (E.2) are satisfied. Note that, the first condition of (E.2) is always satisfied and from the second condition we have that $\sum_{j\in\mathcal{X}^0} \Phi_{ij} = \sum_{j\in\mathcal{X}^0} P_{ij} + \frac{\alpha_i}{2}$ and hence the optimal transition probability of each $j \in \mathcal{X}^0$ is given by

$$\Phi^{\dagger}_{ij} = P_{ij} + \frac{\alpha_i}{2|\mathcal{X}^0|}, \qquad \forall j \in \mathcal{X}^0.$$

From Lemma E.1, Part (b), case 1, the lower bound (E.3), holds with equality if conditions given by (E.4) are satisfied. Furthermore, from the second condition of (E.4) we have that $\sum_{j\in\mathcal{X}_0} \Phi_{ij} = \sum_{j\in\mathcal{X}_0} P_{ij} - \frac{\alpha_i}{2}$, and also the first condition must be satisfied, hence the optimal transition probability of each $j \in \mathcal{X}_0$ is given by

$$\Phi^{\dagger}_{ij} = \left(P_{ij} - \frac{\alpha_i}{2|\mathcal{X}_0|}\right)^+, \qquad \forall j \in \mathcal{X}_0.$$

Lemma E.1, Part (b), case 1, characterize the solution for $\sum_{j\in\mathcal{X}_0} P_{ij} + \frac{\alpha_i}{2} \geq 0$. Next, the characterization of solution when this condition is violated, that is, when $\sum_{s=1}^{k}\sum_{j\in\mathcal{X}_{s-1}} P_{ij} - \frac{\alpha_i}{2} \leq 0$ for any $k \in \{1, 2, \ldots, r\}$ is discussed.

From Lemma E.1, Part (b), case 2, the lower bound (E.5), holds with equality if conditions given by (E.6) are satisfied. Furthermore, from (E.6b) we have that

$$\sum_{j\in\mathcal{X}_k} \Phi_{ij} = \sum_{j\in\mathcal{X}_k} P_{ij} - \Big(\frac{\alpha_i}{2} - \sum_{s=1}^{k}\sum_{j\in\mathcal{X}_{s-1}} P_{ij}\Big),$$

and conditions $\frac{\alpha_i}{2} - \sum_{s=1}^{k}\sum_{j\in\mathcal{X}_{s-1}} P_{ij} \geq 0$ and (E.6c) must be satisfied, hence the optimal transition probability of each $j \in \mathcal{X}_k$ is given by

$$\Phi^{\dagger}_{ij} = \Big(P_{ij} - \Big(\frac{\alpha_i}{2|\mathcal{X}_k|} - \sum_{j=1}^{k}\sum_{z\in\mathcal{X}_{j-1}} P_{iz}\Big)^+\Big)^+.$$

For additional details concerning the steps for the solution of Problem 6.1, see the proof of Theorem 3.1.

Bibliography

[1] N.U. Ahmed. *Linear and nonlinear filtering for scientists and engineers*. World Scientific Publishers, Singapore, New Jersey, London, Hong Kong, 1999.

[2] N.U. Ahmed and C.D. Charalambous. Minimax games for stochastic systems subject to relative entropy uncertainty: Applications to sde's on hilbert spaces. *J. Math. Control Signals Systems*, 19(1):65–91, February 2007.

[3] J.S. Baras and M. Rabi. Maximum entropy models, dynamic games, and robust output feedback control for automata. In *Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference*, pages 1043–1049, December 2005.

[4] A. Barron, J. Rissanen, and B. Yu. The minimum description length principle in coding and modeling. *IEEE Trans. Inf. Theory*, 44(6):2743–2760, October 1998.

[5] R.G. Bartle. *The elements of integration and Lebesgue measure*. Wiley Classics Library. Wiley, 2011.

[6] T.S. Basar and P. Bernhard. *H-infinity optimal control and related minimax design problems: A dynamic game approach*. Collection Systèmes complexes. Birkhuser, 1995.

[7] C. Beck, S. Lall, T. Liang, and M. West. Model reduction, optimal prediction, and the Mori-Zwanzig representation of Markov chains. In *Proceedings of the 48th IEEE Conference on Decision and Control*, pages 3282–3287, December 2009.

[8] A. Bensoussan and R. Elliot. A finite dimensional risk-sensitive control problem. *SIAM J. Control Optim.*, 33(6):1834–1846, 1995.

[9] D.P. Bertsekas. *Dynamic programming and optimal control*. Athena Scientific, 2005.

[10] D.P. Bertsekas and S.E. Shreve. *Stochastic optimal control: The discrete-time case*. Athena Scientific, 2007.

[11] P. Billingsley. *Convergence of probability measures*. John Wiley and Sons, New York, 1968.

[12] L. Breiman. *Probability*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 1968.

[13] P.E. Caines. *Linear stochastic systems*. John Wiley & Sons, Inc., New York, 1988.

[14] C.D. Charalambous and J. Hibey. Minimum principle for partially observable nonlinear risk-sensitive control problems using measure-valued decompositions. *Stoch.Stoch.Rep.*, 57(3-4):247–288, 1996.

[15] C.D. Charalambous and F. Rezaei. Stochastic uncertain systems subject to relative entropy constraints: Induced norms and monotonicity properties of minimax games. *IEEE Trans. Autom. Control*, 52(4):647–663, April 2007.

[16] C.D. Charalambous, I. Tzortzis, and F. Rezaei. Stochastic optimal control of discrete-time systems subject to conditional distribution uncertainty. In *Proceedings of the 50th IEEE Conference on Decision and Control, and the European Control Conference*, pages 6407–6412, December 2011.

[17] T.M. Cover and J.A. Thomas. *Elements of information theory*. John Wiley and Sons, Inc., 1991.

[18] K. Deng, P.G. Mehta, and S.P. Meyn. Aggregation-based model reduction of a hidden Markov model. In *Proceedings of the 49th IEEE Conference on Decision and Control*, pages 6183–6188, 2010.

[19] K. Deng, P.G. Mehta, and S.P. Meyn. Optimal Kullback-Leibler aggregation via spectral theory of Markov chains. *IEEE Trans. Autom. Control*, 56(12):2793–2808, December 2011.

[20] K. Deng, P.G. Mehta, S.P. Meyn, and M. Vidyasagar. A recursive learning algorithm for model reduction of hidden Markov models. In *Proceedings of the 50th IEEE Conference on Decision and Control, and the European Control Conference*, pages 4674–4679, December 2011.

[21] K. Deng, Y. Sun, P.G. Mehta, and S.P. Meyn. An information-theoretic framework to aggregate a Markov chain. In *Proceedings of the American Control Conference*, pages 731–736, June 2009.

[22] N. Dunford and J.T. Schwartz. *Linear operators: General theory*. Interscience Publishers, Inc., New York, 1958.

[23] P. Dupuis and R.S. Ellis. *A weak convergence approach to the theory of large deviations*. John Wiley & Sons, Inc., New York, 1997.

[24] R.J. Elliott, L. Aggoun, and J.B. Moore. *Hidden Markov models: Estimation and control*. Springer, 1995.

[25] K. Fan. Minimax theorems. *Proc. Nat. Acad. Sci.*, 39(1):42–47, 1953.

[26] A. Ferrante, M. Pavon, and F. Ramponi. Constrained approximation in the Hellinger distance. In *Proceedings of European Control Conference*, pages 322–327, July 2007.

[27] A. Ferrante, M. Pavon, and F. Ramponi. Hellinger vs. Kullback-Leibler multivariable spectrum approximation. *IEEE Trans. Autom. Control*, 53(5):954–967, May 2008.

[28] T.T. Georgiou. Relative entropy and the multivariable multidimensional moment problem. *IEEE Trans. Inform. Theory*, 52(3):1052–1066, March 2006.

[29] T.T. Georgiou and A. Lindquist. Kullback-Leibler approximation of spectral density functions. *IEEE Trans. Inform. Theory*, 49(11):2910–2917, November 2003.

[30] A.L. Gibbs and F.E. Su. On choosing and bounding probability metrics. *Internat. Statist. Rev*, 70(3):419–435, December 2002.

[31] P.R. Halmos. *Measure theory*. Springer-Verlag New York Inc., 1974.

[32] O. Hernandez-Lerma and J.B. Lasserre. *Discrete-time Markov control processes: Basic optimality criteria*. Number v. 1 in Applications of Mathematics Stochastic Modelling and Applied Probability. Springer Verlag, 1996.

[33] M. James, J. Baras, and R. Elliot. Risk-sensitive control and dynamic games for partially observed discrete-time nonlinear systems. *IEEE Trans. Autom. Control*, 39(4):780–792, 1994.

[34] E.T. Jaynes. Information theory and statistical mechanics. *Physics Review*, 106:620–630, 1957.

[35] E.T. Jaynes. Information theory and statistical mechanics ii. *Physics Review*, 108:171–190, 1957.

[36] J.L. Kelley. *General Topology*. Graduate Texts in Mathematics. Springer, 1975.

[37] E. Kreyszig. *Introductory functional analysis with applications*. Wiley Classics Library. John Wiley & Sons, 1978.

[38] P.R. Kumar and J.H. Van Schuppen. On the optimal control of stochastic systems with an exponential-of-integral performance index. *Journal of Mathematical Analysis and Applications*, 80(2):312–332, April 1981.

[39] P.R. Kumar and P. Varaiya. *Stochastic systems: Estimation, identification, and adaptive control*. Prentice Hall, 1986.

[40] D.G. Luenberger. *Optimization by vector space methods*. Professional Series. Wiley, 1968.

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

[42] B. Oksendal and A. Sulem. Portfolio optimization under model uncertainty and bsde games. Rapport de Recherche 7554, Institut National de Recherche en Informatique et en Automatique, 2011.

[43] M. Pavon and A. Ferrante. On the Georgiou-Lindquist approach to constrained Kullback-Leibler approximation of spectral densities. *IEEE Trans. Autom. Control*, 51(4):639–644, April 2006.

[44] I.R. Petersen, M.R. James, and P. Dupuis. Minimax optimal control of stochastic uncertain systems with relative entropy constraints. *IEEE Trans. Autom. Control*, 45(3):398–412, March 2000.

[45] M.S. Pinsker. *Information and information stability of random variables and processes*. Holden-Day, 1964.

[46] M.S. Pinsker. *Information and information stability of random variables and processes*. Holden-Day, San Francisco, 1964.

[47] H.V. Poor. On robust Wiener filtering. *IEEE Trans. Autom. Control*, 25(3):531–536, June 1980.

[48] P. Dai Pra, L. Meneghini, and W.J. Runggaldier. Connections between stochastic control and dynamic games. *Math. Control Signals Systems*, 9(4):303–326, 1996.

[49] Z. Rached, F. Alalaji, and L.L. Campbell. The Kullback-Leibler divergence rate between Markov sources. *IEEE Trans. Inf. Theory*, 50(5):917–921, May 2004.

[50] F. Rezaei, C.D. Charalambous, and N.U. Ahmed. Optimal control of uncertain stochastic systems subject to total variation distance uncertainty. *SIAM J. Control Optim.*, 50(5):2683–2725, September 2012.

[51] F. Rezaei, C.D. Charalambous, and N.U. Ahmed. *Optimization of stochastic uncertain systems: Entropy rate functionals, minimax games and robustness. A festschrift in honor of Robert J Elliott*. Advances in Statistics, Probability and Actuarial Science. World Scientific Publishing Company Incorporated, 2012.

[52] J. Rissanen. Modeling by shortest data description. *Automatica*, 14(5):465–471, September 1978.

[53] W. Rudin. *Functional analysis*. International series in pure and applied mathematics. McGraw-Hill, 1991.

[54] J.H. Van Schuppen. *Mathematical control and system theory of discrete-time stochastic systems*. Preprint, 2014.

[55] P. Sharma, S.M. Salapaka, and C.L. Beck. Entropy-based framework for dynamic coverage and clustering problems. *IEEE Trans. Autom. Control*, 57(1):135–150, January 2012.

[56] V.A. Ugrinovskii and I.R. Petersen. Finite horizon minimax optimal control of stochastic partially observed time varying uncertain systems. *Math. Control Signals Systems*, 12(1):1–23, 1999.

[57] K.S. Vastola and H.V. Poor. On robust Wiener-Kolmogorov theory. *IEEE Trans. Inform. Theory*, 30(2):315–327, March 1984.

[58] M. Vidyasagar. Reduced-order modeling of Markov and hidden Markov processes via aggregation. In *Proceedings of the 49th IEEE Conference on Decision and Control*, pages 1810–1815, December 2010.

[59] P. Whittle. A risk-sensitive maximum principle: The case of imperfect state observations. *IEEE Trans. Autom. Control*, 36(7):793–801, July 1991.

[60] E. Wong and B. Hajek. *Stochastic processes in engineering systems*. Springer-Verlag, New York, 1985.

[61] Yunwen Xu, S.M. Salapaka, and C.L. Beck. On reduction of graphs and Markov chain models. In *Proceedins of the 50th IEEE Conference on Decision and Control, and the European Control Conference*, pages 2317–2322, December 2011.

Curriculum Vitae

Ioannis Tzortzis
E-mail: tzortzis.ioannis@ucy.ac.cy

Education

2009-2015	PhD in Electrical/Electronic Engineering, Adviser: Professor Charalambos D. Charalambous Electrical and Computer Engineering Department, University of Cyprus, Nicosia, Cyprus.
2007-2009	M.Sc in Electrical/Electronic Engineering, Adviser: Professor Charalambos D. Charalambous Electrical and Computer Engineering Department, University of Cyprus, Nicosia, Cyprus.
2005-2007	B.Sc in Electrical/Electronic Engineering, Adviser: Professor Andras Barta Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, Hungary.
2000-2003	Diploma in Electrical Engineering, Adviser: Professor Christos Marouchos Electrical Engineering Department, Higher Technical Institute, Nicosia, Cyprus.

Publications

Thesis

1. I. Tzortzis. Extremum problems with total variation distance metric on the space of probability measures and applications. *Ph.D. thesis*, Department of Electrical and Computer Engineering, University of Cyprus, January 2015.

2. I. Tzortzis. Mathematical models for demography and its applications to Cyprus population. *Master thesis*, Department of Electrical and Computer Engineering, University of Cyprus, May 2009.

3. I. Tzortzis. Real time control of the ball and beam system. *Bachelor thesis*, Department of Electrical Engineering and Informatics, Budapest University of Technology and Economics, June 2007.

4. I. Tzortzis. Analysis of switching converters. *Diploma thesis*, Electrical Engineering Department, Higher Technical Institute, June 2003.

Journals

1. I. Tzortzis, C.D. Charalambous, T. Charalambous, C. N. Hadjicostis and M. Johansson. Approximation of Markov processes by lower dimensional processes via total variation metrics. *Submitted to IEEE Transactions of Automatic Control (IEEE TAC)*, October 2014, pages 15. [Online]. Available: http://arxiv.org/abs/1410.3976.

2. I. Tzortzis, C.D. Charalambous, and T. Charalambous. Dynamic programming subject to total variation distance ambiguity. *Submitted to SIAM Journal on Control and Optimization (SICON)*, February 2014, pages 26. [Online]. Available: http://arxiv.org/abs/1402.1009.

3. C.D. Charalambous, I. Tzortzis, S. Loyka, and T. Charalambous. Extremum problems with total variation distance and their applications. *IEEE Transactions of Automatic Control (IEEE TAC)*, vol.59, no.9, pp.2353-2368, Sept.2014.

Conferences

1. I. Tzortzis, C.D. Charalambous, S. Loyka, T. Charalambous, C.N. Hadjicostis, and M. Johansson. Approximation of Markov processes by lower dimensional processes.

In: *53rd IEEE Conference on Decision and Control (CDC 14)*, Los Angeles, California, 15–17 December 2014.

2. C.D. Charalambous, I. Tzortzis, S. Loyka, and T. Charalambous. Extremum problems with total variation distance. In: *52nd IEEE Conference on Decision and Control (CDC 13)*, Florence, Italy, 10–13 December 2013, pp. 1204–1209.

3. C.D. Charalambous, I. Tzortzis, and T. Charalambous. Dynamic programming with total variational distance uncertainty. In: *51st IEEE Conference on Decision and Control (CDC 12)*, Maui, Hawaii, 10–13 December 2012, pp. 1909–1914.

4. C.D. Charalambous, I. Tzortzis, and F. Rezaei. Stochastic optimal control of discrete-time systems subject to conditional distribution uncertainty In: *50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC 11)*, Orlando, Florida, 12–15 December 2011, pp. 6407–6412.

5. C.D. Charalambous, I. Tzortzis, and F. Rezaei. Stochastic optimal control subject to ambiguity. In: *18th International Federation of Automatic Control World Congress (IFAC 11)*, Milano, Italy, August 28–September 2, 2011, pp. 4459–4464.

6. C.D. Charalambous, F. Rezaei, and I. Tzortzis. Optimal stochastic control of discrete-time systems subject to total variation distance uncertainty. In: *49th IEEE Conference on Decision and Control (CDC 10)*, Atlanta, Georgia, USA, 15–17 December 2010, pp. 1442–1447.

7. I. Tzortzis, and C.D. Charalambous. Optimum immigration policies based on linear quadratic theory. In: *4th International Symposium on Communications, Control and Signal Processing (ISCCSP 10)*, Limassol, Cyprus, 3–5 March 2010, pp. 1–5.

8. I. Tzortzis, and C.D. Charalambous. Dynamic demographic models and parameter identification: simulations based on statistical data. In: *10th European Control Conference (ECC 09)*, Budapest, Hungary, 23–26 August 2009.