Fundamental limits of remote estimation

Jhelum Chakravorty

Electrical and Computer Engineering
McGill University
Montreal, Canada

July 2017

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

To my parents

Abstract

In this thesis we focus on an application of the dynamic team theory, the *Remote Estimation* problem, that has been gaining interest in today's unified world of control and communication. Fit into the realm of decentralized control problems, which is almost ubiquitous nowadays due to its own merits, the remote estimation problem finds application in various networked systems. The remote estimation system consists of a sensor (transmitter) and an estimator (receiver). The sensor observes (*senses*) a dynamic source and sends the data (of its observation) over a lossy communication channel to a remotely placed receiver. In such scenarios, often the data communication is expensive, which gives the incentive for intermittent transmission. On the other hand the receiver has to estimate the source realization based on the received data. The accuracy of the estimation is measured by an estimation error. The transmission and the estimation cost together gives the total per-step cost of communication. In such applications, a fundamental question is to make an optimized trade-off between the transmission cost and the estimation error. In this thesis we investigate such an optimization problem and provide a unified framework to analyze the costly and constrained communication problems with finite and infinite horizon. We recognize that in the applications of remote estimation, often there lie certain symmetry and monotonicity properties in the state-dynamics. We exploit this observation to establish the optimality of nicely structured optimal strategies, namely the threshold-based strategies. Threshold-based strategies have their appeal in simplicity of implementation. In this work, we provide a complete characterization of the optimal thresholds and optimal performances pertaining to the remote estimation problem. We derive analytic expressions for the optimal thresholds and the optimal performances. Furthermore, we develop algorithms to find numerical solutions, which creates the scope for extension of the current research to higher dimensions.

Résumé

Dans cette thèse, l'accent est mis sur une application de la théorie de la dynamique d'équipe, le problème de l'estimation à distance, qui a progressivement gagné de l'intérêt dans le monde unifié d'aujourd'hui, un monde de contrôle et de communication. Inscrit dans la sphère des problèmes du contrôle décentralisé qui est presque omniprésente de nos jours en raison de ses caractéristiques propres, le problème de l'estimation à distance (ED) trouve son application dans divers systèmes en réseaux. Le système de l'estimation à distance consiste en un capteur (transmetteur), et en un estimateur (receveur). Le capteur observe (perceptions) une source dynamique et envoie les données (de son observation) sur un canal de communication avec perte, vers un receveur placé à distance. Souvent, dans de tels scénarios, la communication de données est chère, ce qui encourage les transmissions intermittentes. D'un autre côté, le receveur doit estimer la source de réalisation basée sur les données reçues. L'exactitude de l'estimation est mesurée par une erreur estimée. Ensemble, le coût de la transmission et de l'estimation donnent le coût total par étape de la communication. Dans de telles applications, la question fondamentale est de faire un compromis optimisé, entre le coût de la transmission et l'erreur de l'estimation. Dans cette thèse, nous étudions un tel problème d'optimisation et nous fournissons une structure unifiée pour analyser les problèmes à l'horizon fini et infini dans le cadre d'une communication coûteuse et contraignante. Nous remarquons que dans l'application de l'estimation à distance résident souvent certaines propriétés de symétrie et de monotonicité dans la dynamique stable. Nous exploitons cette information pour établir l'optimalité de stratégies optimales très bien structurées, appelées les stratégies de seuil. Les stratégies de seuil sont attrayantes grâce à la simplicité de leur mise en oeuvre. Dans ce travail, nous apportons une caractérisation complète du seuil optimal et de la performance optimale relative au problème de l'estimation à distance. Nous établissons une formule pour le seuil optimal et les performance optimales. Par ailleurs, nous développons des algorithmes pour trouver des solutions numériques qui permettent d'étendre d'extension de notre recherche actuelle vers les dimensions supérieures.

Acknowledgments

I started my journey around four years ago as a doctoral candidate with a dilemma regarding the choice of the topic of my research. Having an academic background in control systems, a curiosity in delving into an interdisciplinary research environment involving but not limited to control theory and no prior exposure to stochastic systems and communication theory, I suppose I had a rather rocky advent. Had there been no incessant support, warm encouragement and keen guidance from my supervisor, Professor Aditya Mahajan, I wonder if I could learn to cope with numerous overwhelming days, especially towards the beginning of the program. He introduced to me the world of communication, the flavor of control in it and the wonderful intertwining of History, Mathematics and Philosophy pertaining to this domain. I convey my heartfelt gratitude to him for bestowing upon me the pleasure of having worked in a coming-of-age area of research and enriching myself through vigorous interactions along the course of my study.

I would like to take this opportunity to thank the two other members of my advisory committee, Professors Peter E. Caines and Ioannis Psaromiligkos for sharing with me their insights and suggestions that instigated me to reflect on various important aspects of my work and address the practical impacts of my study. Their invaluable advice has helped me improve my presentation skill over past few years. I am grateful for their time and support even with matters outside the scope of my research. I thank all of my course instructors, Professors Benoit Boulet, Peter E. Caines, Aditya Mahajan, Hannah Michalska and Ioannis Psaromiligkos for enhancing the joy of learning. Special thanks to Stephanie Larocque, my superior at the Office of Students with Disabilities at McGill, for providing me with a wonderful work culture and exposing me to the terrific opportunity to experience professionalism woven with humanity and benevolence.

I am grateful to have been in the presence of my colleagues at McConnel Engineering and friends in and outside McGill. Special thanks to Jayakumar for his inputs in many productive discussions and his helpful suggestions on many occasions. His passion for coding has certainly inflicted upon me the urge to hone my skill in it, which previously I used to ignore to some extent. Many thanks to Lydie for helping me with the French in writing the abstract of my thesis. My sincere regards for Gogol, one of my closest friends in Montreal who hails from my hometown and my undergraduate school, to have offered me the warmth of a wonderful blend of camaraderie and nostalgia. I know no way to properly express the extent of my gratitude to Laure, Jack and Simon, my very close friends and past roommates, for showing me the magic of transparent communication, bringing out the best in me, helping me restore the faith in myself during low phases and dragging me into learning all those fun sports. I thank all of my friends whom I have met in Montreal for lending the warmth of meaningful connections which helped me survive the chilly winters. Never have I been a part of such a motley of enthusiastic, welcoming and driven personalities.

Last but not the least, I wish to convey my sincere gratitude to my family for their unconditional love and support. Thanks to my parents for their invaluable advices in every turn of events in my life, for being patient listener to all my rantings, for inculcating in me the values which molded my personality quite deeply. I thank my grandparents, who unfortunately have passed away, for showering me with indulgent love and offering me in my tender age the persona I could look up to. Finally, special thanks to my little sister–a friend, an advisor, a confidante and a partner–for it is her presence that has prohibited my life from ever becoming boring.

Contents

List of Figures

List of Tables

This page is intentionally left blank.

List of Acronyms

AdaM Adaptive Moments

FIE Fredholm Integral Equation of second kind

KW Kiefer-Wolfowitz

MDP Markov Decision Processes

POMDP Partially Observable Markov Decision Process

RE Remote Estimation

RM Robbins-Monro

RMC Renewal Monte Carlo

SA Stochastic Approximation

SF Smoothed Functional

SG Stochastic Gradient

SPSA Simultaneous Perturbation Stochastic Approximation

This page is intentionally left blank.

List of Notations

$\mathbb{X}$	State space of the source process
$\mathbb{Z}$	Space of integers
$\mathbb{R}$	Space of reals
X_t	State of Markov process at time t
x_t	Realization of the random variable X_t
$\hat{X}_t$	Estimate of the random variable X_t
$X_{0:t}$	Shorthand for $\{X_0, X_1, \cdots, X_t\}$
$\mathbb{E}[\cdot]$	Expectation of a random variable
$\mathbb{P}(\cdot)$	Probability of a random variable
$\mathcal{N}(\mu, \sigma^2)$	Gaussian distribution with mean μ and variance σ^2
β	Discount factor
U_t	Action taken at time t
λ	Transmission cost
$d(\cdot)$	Distortion function
v_i	i-th element of a vector v
$\nabla_x Y$	The gradient of a vector Y w.r.t. a variable x
$\odot$	Hadamard product operator
$\otimes$	Kronecker product operator
$L_\beta^{(k)}$	Expected distortion until first successful reception
$M_\beta^{(k)}$	Expected time until first successful reception
$K_\beta^{(k)}$	Expected number of transmissions until first successful reception
$D_\beta^{(k)}$	Expected distortion
$N_\beta^{(k)}$	Expected number of transmissions
$C_\beta^{(k)}$	Expected costly performance

This page is intentionally left blank.

Chapter 1

Introduction

"The single biggest problem in communication is the illusion that it has taken place." – George Bernard Shaw

1.1 Motivation

This thesis collates an in-depth study of the Remote Estimation (RE) problem, which is a coming-of-age theoretical problem finding its applications in the field of control and communication. The rich framework provided by sophisticated mathematical tools connects the two aforementioned domains of research with the possibilities of diverse applications. The remote estimation problem is undoubtedly one among such applications, with the potential of extensive theoretical exploration as well as practical applicability.

In today's world, applications such as network control systems, smart grids, cyber physical systems, environmental monitoring et cetera are growing in importance. One of the salient inherent features of such applications is the *decentralized control.* With the advantage of lower installation and control cost over the centralized counterpart, the decentralized control and communication systems offer promising solutions to complex optimization problems. In the field of systems engineering, from the aspect of the role played by control actions, multi-agent systems can be broadly classified in two groups, namely the *centralized* and the *decentralized* systems. The latter group of systems are the ones where control action (or *decision*) is taken by components of the system, based on the available *local information* to each of them, which put together reflects into the *global* (or holistic) outcome. The decentralized systems are different from their centralized counterpart by their accessibility to the information available to each agent. The decentralized systems essentially have the *non-classical* information structure (a set of data available to each agent) where no two decision makers have the same information available to them.

1.1.1 The theory of *Teams*

Based on the objective of the control action, the decentralized systems can be categorized as *Teams* or *Games*. The decentralized control problem that we consider in our work falls into the category of *Team Theory* [1–3]. This realm of research deals with multiple agents and considers the *organizational behavior* (a term introduced by Marschak), where the agents have access to different information about the state of the world and take their own decisions based on the information available to them in order to optimize a common objective function.

Teams have two solution concepts: the equilibrium solution and the optimal solution. An equilibrium solution is the set of strategies of all agents where the system performance cannot be improved by any unilateral deviation of the agents. In the context of team problems, this is called the *person by person* solution (an equivalent of such a solution in games being the Nash equilibrium). An equilibrium solution is the set of strategies of all agents in a team, for which no other strategy can yield a better performance. As opposed to the *globally optimal* solution in teams, the equilibrium solution can be viewed as *locally optimal*.

1.1.2 The RE problem

Over the years, several applications of the team theory have emerged, and remote estimation problem is one such instance. In remote estimation problems, one or more *agent(s)* observe(s) the realization of the state of a *source* process and send(s) the observation to remotely placed agents over communication channels between them. This essentially consists of two (or more) agents, one of which observes the source and then takes a decision to transmit its observation to a remotely placed receiver. The challenge of the receiver, who has access to an information different than the observer and does not see the source, is to generate in real time (i.e., with zero delay) an estimate of the source so as to minimize the estimation error. In most applications of remote estimation, the transmission of the data is costly and hence the observer sends its observation intermittently. A natural quest then arises to make an optimal trade-off between the transmission cost and the estimation error.

1.1.3 Role of information in RE problem

In a RE system, an observer transmits the observed source symbol over a communication channel to a remotely placed receiver, which tries to regenerate the source symbol. The channel could be *noiseless* or *noisy*. In a noiseless or ideal channel, the symbol observed by the transmitter (which is the input to the channel) is the same as that received by the estimator (when the input alphabet is the set of reals, an ideal channel has infinite capacity). In contrast, a noisy channel has the output different than its input. A special type of noisy channel that we consider in this thesis is an *erasure* channel, where loss of information contained in the symbol occurs due to packet drops in the channel. The issues like successful transmission and recovery of a message (or information) in a symbol were addressed by Shannon [4]. This was the pioneering work in the realm of *classical Information Theory* which dealt with the coding (or compressing) of a source and a channel (error correction).

Although the RE problem is a communication system in which the transmitter transmits information about a source to a receiver, the tools from classical information theory are not appropriate to analyze the optimal strategies and the optimal performance. This is because an RE problem is a *real-time communication system*, which is described in the next section. Recently, in [5], the authors have characterized bounds of the maximal achievable rate for a given finite block-length and error-probability over block-fading channels. These bounds are tight for moderate block-lengths. However, using a block or streaming code of block-length n introduces a delay of n. Such schemes are not applicable for the models described in this thesis where zero-delay reconstruction is required.

1.1.4 RE as real-time communication system

A real-time communication system [3,6–8] is a point-to-point communication system, consisting of a causal transmitter and a causal receiver. At each instant t, upon receiving a symbol, the receiver needs to generate an estimate of the source symbol at $t - \delta$, where $\delta > 0$ is the delay in communication, which can be zero or non-zero finite. Depending on the type of the communication channel connecting the agents, there are broadly four categories of real-time communication systems where:

- the transmitter and the estimator are connected by one-way noiseless channel
- the transmitter and the estimator are connected by one-way noisy channel
- the transmitter and the estimator are connected by two-way channel with noisy forward and noiseless feedback
- the transmitter and the estimator are connected by two-way channel with noisy forward and feedback.

Remote estimation problems may be viewed as a special case of real-time communication. The salient features of RE are as follows:

(F1) The decisions are made sequentially.

(F2) The reconstruction/estimation at the receiver must be done with zero-delay.

(F3) When a packet does get through, it is received without noise.

The main impediment in formulating the RE problem in the classical information theoretic framework is that the fundamental aspects of the latter, such as the *entropy* of a source, the *capacity* of the channel or the *rate-distortion* function are asymptotic in nature, which call for a large value of delay in communication. In RE problem, an estimate of the source is to be generated in real-time, which means zero or finite delay and hence the notions of the classical information theory do not apply here.

Furthermore, as in real-time communication, the key conceptual difficulty is that the data available at the transmitter and the receiver is increasing with time. Thus, the domain of the transmission and the estimation function (the *decision functions* of the transmitter and the estimator respectively) increases with time. This makes the optimization problem combinatorial and hence *intractable*.

To circumvent this difficulty one needs to identify sufficient statistics for the data at the transmitter and the data at the receiver. In the real-time communication literature, dynamic team theory (or decentralized stochastic control theory) is used to identify such sufficient statistics as well as to identify a dynamic program to determine the optimal transmission and estimation strategies. Similar ideas are also used in remote-estimation literature. In addition, feature (F3) allows one to further simplify the structure of optimal transmission and estimation strategies.

In particular, when the source is a first-order autoregressive process, majorization theory is used to show that the optimal transmission strategies is characterized by a threshold [9–13]. In particular, it is optimal to transmit when the instantaneous distortion due to not transmitting is greater than a threshold. The optimal thresholds can be computed either using dynamic programming [9, 10] or using renewal relationships [13, 14].

In the next section, we provide a brief overview of the results existing in the literature and discuss the ones close to our line of research in a little more details.

1.2 Previous works on RE

Two approaches have been used in the literature to investigate real-time zero-delay communication. The first approach considers coding of individual sequences [15–18]; the second approach considers coding of Markov sources [3, 6–8, 19, 20]. The model presented above fits with the latter approach. In particular, it may be viewed as real-time transmission, which is noiseless but expensive. In most of the results in the literature, the focus has been on identifying sufficient statistics (or information states) at the transmitter and the receiver; for some of the models, a dynamic programming decomposition has also been derived. However, very little is known about the solution of these dynamic programs.

The communication system described in this thesis is much simpler than the general real-time communication setup due to the following feature: whenever the transmitter transmits, it sends the current state to the receiver. These transmitted events *reset* the estimation error to zero. We exploit these special features to identify an analytic solution to the dynamic program corresponding to the above communication system.

The motivation of RE comes from networked control systems. The earliest instance in the form of a static (one shot) RE problem was first considered in [1] in the context of information gathering in organizations. The problem of optimal off line choice of measurement times was considered in [21], whereas the problem of optimal on-line choice of measurement times was considered in [22]. In the field of control, the early work on the separation of control and estimation through identification of *information states* was considered in [23] and in [24]. In recent years, several variations of remote estimation have been considered. The closely related problem of event-based sampling (also called Lebesgue sampling) was considered in [25]. In addition, several variations of the remote estimation problem have been considered in the literature. The most closely related models are [9–11, 26–28], which are summarized below. Other related work includes censoring sensors [29, 30] (where a sensor takes a measurement and decides whether to transmit it or not; in the context of sequential hypothesis testing), estimation with measurement cost [31–33] (where the receiver decides when the sensor should transmit), sensor sleep scheduling [34–37] (where the sensor is allowed to sleep for a pre-specified amount of time); and event-based communication [38–40] (where the sensor transmits when a certain event takes place). We contrast our model with [9–11] below.

In [26], optimal remote estimation of i.i.d. Gaussian processes is investigated under a constraint on the total number of transmissions. The optimal estimation strategy is derived when the transmitter is restricted to be of threshold-type.

In [27], the optimal remote estimation of a continuous-time autoregressive Markov process driven by Brownian motion is considered under a constraint on the number of transmissions. The optimal transmission strategy is derived under an assumption on the structure of the optimal estimation strategy. It is shown that the optimal transmission strategy is of a threshold-type, where the thresholds are determined by solving a sequence of nested optimal stopping problems.

In [28] optimal remote estimation of Gauss-Markov processes is investigated when there is a cost associated with each transmission. The optimal transmission strategy is derived when the estimation strategy is restricted to be Kalman-like.

In [9–11], optimal remote estimation of autoregressive Markov processes is investigated when there is a cost associated with each transmission. It is assumed that the autoregressive process is driven by a symmetric and unimodal noise process but no assumption is imposed on the structure of the transmitter or the receiver. Using different solution approaches ([9, 10] use majorization theory while [11] uses person-by-person optimality), it is shown that the optimal transmission strategy is threshold-based and the optimal estimation strategy is Kalman-like (the precise form of these strategies is stated in Theorem 2.6.1). Thus, the optimal transmission and estimation strategies are easy to implement.

1.3 Scope of this thesis: an overview

In this thesis, we concentrate on a two-agent problem consisting of a *transmitter* and a *estimator*. A transmitter observes (alternatively, senses) a Markov state process and decides whether or not to transmit the state realization to a remotely placed estimator. Hence, the transmitter is also called a *sensor*. The estimator receives a symbol which reaches it passing through a *communication channel*. For this reason an estimator is often called a *receiver*. We use the terms transmitter/sensor and estimator/receiver interchangeably. We consider two cases of the real-time communication: i) one-way with noiseless channel and ii) two-way with noisy forward channel and noiseless feedback. It is perhaps worth noting here that we consider no *source coding* and consequently, the transmitter is assumed to transmit the observed state of the source intact. The rationale behind this is that in many applications of remote estimation, the cost of transmitting a packet is much more significant than the size of the packet. which can be verified by looking at a 'standard' sensor and by comparing the energy consumed in sensing and in communication. An example of such an application would be network control systems with battery-powered transmitters sending data over a packet-switched network.

We recognize the remote estimation problem as a *Decentralized Markov Decision Process* (or, Dec-MDP). In this setup, the source is a Markov process. A transmitter observes the realization of the state and takes a decision to transmit the observation to a remote receiver. The transmitter gets an acknowledgment of the reception of the data (or the lack of so) and hence has the knowledge of the information available to the receiver. Upon receiving a symbol, which may or may not be the data, the receiver attempts to estimate the source. The estimator does not know the information available to the transmitter. Thus, from the estimator's point of view, the source process is a Partially Observable Markov Decision Process (POMDP). Nevertheless, the overall two-agent system is *jointly fully observable*, hence the name.

In the Dec-MDP, as described above, one could pose the team problem with non-classical information structure to find the globally optimal solution by formulating a POMDP-like dynamic program. The main conceptual difficulty in finding the global solution is that the information set at the agents grows with time, making it a combinatorial problem, which is inherently intractable.

It is shown by Witsenhausen in his famous counterexample [41] that for decentralized systems, linear control strategies are not optimal, even for LQG (Linear Quadratic Gaussian) systems. In the remote estimation problem, which is a decentralized team problem, we seek a globally optimal solution, which comprises the optimal *communication strategies*, i.e., the transmission and the estimation strategies. The optimal strategies are expected to be non-linear and in order to be able to find the global solution, we need convexity in the strategy space.

A key feature of the RE problem is the sequential communication of the control actions, that arises in a team with decentralized control. Although the strategies of the two agents–the transmitter and the estimator–are common knowledge, since the estimator does not have access to the information available to the transmitter, it has to make a *belief* on the transmitter's information in order to come up with it's own strategy. This eventually leads to a dynamic program, which is like a POMDP but the minimization is over a functional space, which is computationally heavy.

In this thesis, we address the fundamental trade-off between the transmission cost and the estimation error and formulate the *costly* and *constrained* communication problems. We choose a stylized model of the source process and the per-step cost function and seek to establish the structures of the optimal transmission and estimation strategies as well completely characterize the optimal performance for both the cases with an ideal and a erasure communication channel. We provide a unified approach to find the global optimal solutions and complete characterization and computation of the optimal performances for a large class of RE problems without the need to solve a dynamic program explicitly for every change of the transmission cost.

On a slightly different note, we investigate a somewhat related dynamic team problem that has been gaining attention recently. In the light of analyzing the optimality of the threshold-based strategies for such an *interactive communication* problem, we study a stylized model of a two-user network, where each user observes a noisy alphabet of interest and through sequential communication between them try to generate an estimate of that symbol. In our discussion we have considered the alphabet of interest to be a static random variable and mention the difficulty of formulating such a problem for a dynamic (Markov) symbol.

Lastly, we recognize the role of symmetry and monotonicity in the structural results throughout this thesis and attempt to analyze the sufficiency conditions that make the value function and the optimal strategies even and quasi-convex. In a fairly generalized framework, we establish the easily verifiable conditions which leads to the optimality of threshold-based strategies. For the compactness of the thesis, we have put together the content of the last two topics in two separate chapters in a standalone fashion, where we discuss the results along with the relevant literature overview to create the context.

1.4 Organization of this thesis

This thesis is structured as follows:

Chapter 2 discusses the remote estimation problem with an *ideal* communication channel, i.e., a channel with no packet drop. Hence, a transmitted symbol always reaches the estimator. We study the *discounted* and *long-term average* setups of an *infinite horizon* optimization problem. We formulate the costly and constrained communication problems and characterize the optimal communication strategies and optimal performances.

Chapter 3 introduces two different types of *erasure* communication channels in the remote estimation problem. We establish a theoretical base for the structural results on which we develop the characterization of the optimal performances.

Chapter 4 focuses on the numerical methods of solving a remote estimation problem with packet drops. Using the theoretical results provided in the previous chapters, we provide stochastic approximation algorithms to numerically solve the costly and constrained remote estimation problems. In addition, we build a framework with variants of *Stochastic Gradient Descent* algorithms, which creates the scope of solving other problems fitting closely into the structure of the remote estimation problems, such as the *inventory problem*.

Chapter 5 is the first addendum of the thesis, which introduces the problem of interactive communication, where two agents (alternatively, users) playing the roles of a transmitter and an estimator sequentially communicate between them. In such a scenario, we investigate the structures of the optimal strategies.

Chapter 6 is the second addendum, which provides the sufficient conditions for the value function and the optimal strategies of the remote estimation problem to be even and quasi-convex.

Chapter 7 concludes the work by summarizing the findings in the previous chapters, discussing the interpretations obtained from the results and presenting some of the future scopes of this research.

1.5 Claims of originality and publications

1.5.1 Claims of originality

The following original contributions are presented in this thesis:

- *Developing a unified framework for the remote estimation problem that takes into account the finite horizon and both costly and constrained communication in infinite horizon.* We consider the source state-space to be discrete as well as continuous.

- *Extending structural results for countable source with finite support [10] to infinite support.* Also, extension of the structural results to infinite horizon by arguing the effective compactness of the state-space and the boundedness of the cost function under threshold-based strategies.

- *Extending the assumption of the Gaussian noise* for the continuous state-dynamics to any arbitrary noise process possessing symmetry and unimodality.

- *Introducing in a unified framework the erasure channel with (Markovian) and without (i.i.d.) memory.* Complete characterization of the structural results and the performance of the optimal strategies.

- *Computing analytical expressions for optimal thresholds and optimal performances for lossless and lossy (i.i.d.) channels.* For the discrete state-space, we provide closed-form expressions for the optimal performaces as well as provide a look-up table which enables one to compute the optimal results on-line without explicitly solving a dynamic program. For continuous state-space, the results can be found by solving FIE.

- *Analyzing the symmetry and monotonicity of the value function and optimal strategies* by properly defining a *folding operator*. In a fairly generic framework, we analyze sufficient conditions for these to hold for the power allocation problem in RE (Chapter 6).

- *Integrating the Remote Estimation problem with the domain of learning.* To our knowledge, this is the first approach to integrate stochastic approximation based methods into an application such as remote estimation. We provide the validation of the optimality results by comparing the simulation-based results with those found by the analytical methods. This lays the ground for extending the work to the higher dimension, where the threshold-based strategies have relevance due to their simple implementation, despite the possibility of them being suboptimal.

- *Establishing the analytical results with the examples* of first-order scalar autoregressive processes (birth-death process as a representative of discrete state-space and Gaussian process as a representative of continuous state-space).

1.5.2 List of publications

Journal papers:

(J1) Chakravorty J. and Mahajan A., "Fundamental limits of remote estimation of autoregressive Markov processes under communication constraints," IEEE Transactions on Automatic Control, vol. 62, pp. 1109–1124, March 2017.

Submitted journal papers:

(S1) Chakravorty, J. and Mahajan, A., "Sufficient conditions for the value function and optimal strategy to be even and quasi-convex," arXiv: 1703.10746, submitted to IEEE Transactions on Automatic Control, 2017.

Journal papers under preparation:

(P1) Chakravorty J. and Mahajan A., "Remote estimation with packet drops," in preparation.

(P2) Chakravorty, J., Subramanian, J. and Mahajan, A., "Stochastic approximation based methods to compute optimal thresholds for remote estimation with packet drops," in preparation.

Peer-reviewed conference papers:

(C1) Chakravorty J. and Mahajan A., "Structure of optimal strategies for remote estimation over Gilbert-Elliott channel with feedback", Proceedings of IEEE International Symposium on Information Theory (ISIT), Aachen, Germany, Jun. 25-30, 2017.

(C2) Chakravorty, J., Subramanian, J. and Mahajan, A., "Stochastic approximation based methods for computing the optimal thresholds in remote-state estimation with packet drops," Proceedings of the American Control Conference (ACC), Seattle, WA, May 26-26, 2017.

(C3) Chakravorty J. and Mahajan A., "Remote-state estimation with packet drops", Proceedings of the IFAC Workshop on Distributed Estimation and Control in Networked Systems, Tokyo, Japan, Sep. 8-9, 2016. Recipient of **Best Student Paper Award**.

(C4) Chakravorty J. and Mahajan A., "Structural results for two-user interactive communication", Proceedings of IEEE International Symposium on Information Theory (ISIT), Barcelona, Spain, Jul. 10-15, 2016.

(C5) Chakravorty J. and Mahajan A., "Distortion-transmission trade-off in real-time transmission of Gauss-Markov sources", Proceedings of IEEE International Symposium on Information Theory (ISIT), Hong Kong, Hong Kong, Jun. 14-19, 2015.

(C6) Chakravorty J. and Mahajan A., "Distortion-transmission trade-off in real-time transmission of Markov sources", Proceedings of the Information Theory Workshop (ITW), pp. 1–5, Jerusalem, Israel, Apr. 26-May 1, 2015.

(C7) Chakravorty J. and Mahajan A., "On the optimal thresholds in remote state estimation with communication costs", Proceedings of the IEEE Conference on Decision and Control (CDC), pp. 1041–1046, Los Angeles, CA, Dec. 15-17, 2014.

(C8) Chakravorty J. and Mahajan A., "Average cost optimal threshold strategies for remote state estimation with communication costs", Proceedings of the Allerton conference on communication, control, and computing, Monticello, IL, Oct. 1-3, 2014.

Conference abstracts:

(A1) Chakravorty J., Subramanian J., and Mahajan A.,"Renewal theory based reinforcement learning for Markov processes with controlled restarts", GERAD Optimization Day, May 8-10, 2017.

(A2) Chakravorty J. and Mahajan A., "The distortion transmission function for remote estimation under communication constraints", International Symposium on Modeling and Optimization in Mobile, Ad-Hoc and Wireless Networks (WiOpt), Tempe, AZ, May 9-13, 2016. (Invited Talk)

(A3) Chakravorty J. and Mahajan A., "The distortion transmission function for transmitting autoregressive Markov processes under communication constraints", Information

Theory and Application (ITA) Workshop, San Diego, CA, Feb. 1-5, 2016. (Invited talk)

(A4) Chakravorty J. and Mahajan A., "When to communicate information in two-player teams?", Sixth Workshop on Dynamic Games in Management Science, Montreal, QC, Oct. 22-23, 2015.

(A5) Chakravorty J. and Mahajan A., "Remote estimation of Markov processes under communication constraints", Fourth Rutgers Applied Probability Conference, Piscataway, NJ, Oct. 2-3, 2015. (Invited talk)

Indexed technical reports:

(T1) Chakravorty J. and Mahajan A., "Structural results for two-user interactive communication", Les Cahiers du GERAD, no. G-2016-40, July 2016.

(T2) Chakravorty J. and Mahajan A., "Fundamental limits of remote estimation of Markov processes under communication constraints", Les Cahiers du GERAD, no. G-2015-53, May 2015.

(T3) Chakravorty J. and Mahajan A., "Distortion-transmission trade-off in real-time transmission of Markov sources", Les Cahiers du GERAD, no. G-2014-104, Dec. 2015.

1.5.3 Contributions of co-authors

In (J1), (C4) and (S1), J. Chakravorty and A. Mahajan contributed equally to the analysis, while J. Chakravorty was primarily responsible for the derivations and the writing of the results. In (P1), (C1) and (C3), J. Chakravorty and A. Mahajan contributed equally to develop the model and J. Chakravorty was primarily responsible for the analysis and the writing of the papers. In (P2), (C2), all three authors contributed towards developing of the idea, J. Chakavorty and J. Subramanian were primarily responsible for conducting the numerical experiments and J. Chakravorty was primarily responsible for the writing of the papers.

This page is intentionally left blank.

Chapter 2

Remote estimation over an ideal channel

2.1 Introduction

In many applications such as networked control systems, sensor and surveillance networks, and transportation networks, etc., data must be transmitted sequentially from one node to another under a strict delay deadline. In many of such *real-time* communication systems, the transmitter is a battery powered device that transmits over a wireless packet-switched network; the cost of switching on the radio and transmitting a packet is significantly more important than the size of the data packet. Therefore, the transmitter does not transmit all the time; but when it does transmit, the transmitted packet is as big as needed to communicate the current source realization. In this chapter, we characterize fundamental trade-offs between the estimation error (or distortion) and the cost or average number of transmissions in such systems.

In particular, we investigate a two-agent communication system consisting of a transmitter which observes the state process of a source and transmits its observation to a remotely placed estimator over a lossless communication channel, i.e., a transmitted packet always reaches the estimator. Throughout this thesis, we use the terms tranmitter/sensor and estimator/receiver interchangeably. We consider a stylized model, where a sensor observes a first-order autoregressive Markov process. At each time instant, based on the current state of the process and the history of its past decisions, the sensor determines whether or not to transmit the current state. If the sensor does not transmit, the receiver must estimate the state using the previously transmitted values. A per-step distortion function measures the estimation error. We investigate two fundamental trade-offs in this setup: (i) when there is a cost associated with each communication, what is the minimum expected estimation error plus communication cost; and (ii) when there is a constraint on the average number of transmissions, what is the minimum estimation error. For both these cases, we characterize the transmission and estimation strategies that achieve the optimal trade-off.

It is worth noting here that the aforementioned model fits into the realm of *Team theory* [1–3, 24], where multiple agents work toward optimizing a common objective. Team problems being a decentralized control problem where there are multiple decision makers (in our case there are two decision makers: a transmitter/sensor and and estimator/receiver), the goal of finding a global optimum is essentially challenging. We therefore seek to analyze the structure of the optimal strategies and characterize the optimal performance for a stylized model as introduced in the last paragraph.

2.2 Original contribution

Following the research of the predecessors, an immediate question is how to identify the optimal transmission and estimation strategies for a given communication cost. It is shown in [9–11] that the optimal estimation strategy does not depend on the communication cost while the optimal transmission strategy can be computed by solving an appropriate dynamic program. However, the dynamic programs presented in [9–11] do not exploit the threshold structure of the optimal strategy.

In this chapter, we provide an alternative approach to identify the optimal transmission strategies for the remote estimation problem with ideal communication channel (i.e., channel with no packet drop). We consider infinite horizon remote estimation problem and show that there is no loss of optimality in restricting attention to transmission strategies that use a time homogeneous threshold. To determine the optimal threshold, we first provide computable expressions for the performance of a generic threshold-based transmission strategy and then use these expressions to identify the best threshold-based strategy. Thus, we show that the structure of optimal strategies derived in [9–11] is also useful to compute the optimal strategy.

We investigate remote estimation for two models of Markov processes—discrete state autoregressive Markov processes (Model A) and continuous state autoregressive Markov processes (Model B); both driven by symmetric and unimodal innovations process—under two infinite horizon setups: the discounted setup with discount factor $\beta \in (0, 1)$ and the long term average setup, which we denote by $\beta = 1$ for uniformity of notation. For both models, we consider two fundamental trade-offs:

1. *Costly communication*: When each transmission costs λ units, what is the minimum achievable cost of communication plus estimation error, which we denote by $C^*_\beta(\lambda)$?

2. *Constrained communication*: When the average number of transmissions are constrained by $\alpha \in (0, 1)$, what is the minimum achievable estimation error, which we denote by $D^*_\beta(\alpha)$ and refer to as the *distortion-transmission* trade-off?

We completely characterize both trade-offs. In particular, we show that

- In Model A, $C^*_\beta(\lambda)$ is continuous, increasing, piecewise-linear, and concave in λ while $D^*_\beta(\alpha)$ is continuous, decreasing, piecewise-linear, and convex in α. We derive explicit expressions (in terms of simple matrix products) for the corner points of both these curves.

- In Model B, $C^*_\beta(\lambda)$ is continuous, increasing, and concave in λ while $D^*_\beta(\alpha)$ is continuous, decreasing, and convex in α. We derive an algorithmic procedure to compute these curves by using solutions of Fredholm integral equations of the second kind. When the innovations process is Gaussian, we characterize how these curves scale as a function of the variance σ^2.

We also explicitly identify transmission and estimation strategies that achieve any point on these trade-off curves. For all cases, we show that: (i) there is no loss of optimality in restricting attention to time-homogeneous strategies; (ii) the optimal estimation strategy is Kalman-like; (iii) the optimal transmission strategy is a randomized threshold-based strategy for Model A and is a deterministic threshold-based strategy for Model B.

In addition,

- In Model A, the optimal threshold as a function of λ or α can be computed using a look-up table.

- In Model B, the optimal threshold as function of λ or α can be computed using the solutions of Fredholm integral equations of the second kind.

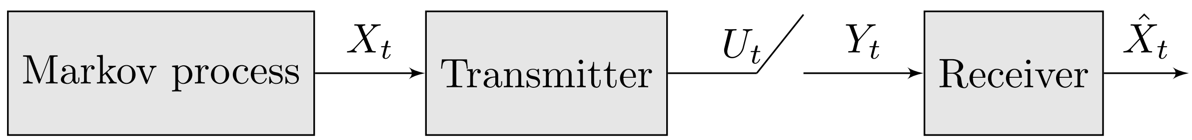


Fig. 2.1 Block diagram of a remote estimation system with ideal channel.

2.3 Model and problem formulation

2.3.1 Model

Consider the following two models of a discrete-time Markov process $\{X_t\}_{t=0}^{\infty}$ with the initial state $X_0 = 0$ and for $t \geq 0$,

$$X_{t+1} = aX_t + W_t, \tag{2.1}$$

where $\{W_t\}_{t=0}^{\infty}$ is an i.i.d. innovations process. We consider two specific models:

- **Model A:** $a, X_t, W_t \in \mathbb{Z}$ and W_t is distributed according to a unimodal and symmetric pmf (probability mass function) p, i.e., for all $e \in \mathbb{Z}_{\geq 0}$, $p_e = p_{-e}$ and $p_e \geq p_{e+1}$. To avoid trivial cases, we assume p_0 is strictly less than 1.

- **Model B:** $a, X_t, W_t \in \mathbb{R}$ and W_t is distributed according to a unimodal, differentiable and symmetric pdf (probability density function) ϕ, i.e., for all $e \in \mathbb{R}_{\geq 0}$, $\phi(e) = \phi(-e)$ and for any $\delta \in \mathbb{R}_{>0}$, $\phi(e) \geq \phi(e + \delta)$.

Remark 1 We consider in this work that the parameter a is known to both the agents. If not, the transmitter can estimate a using its observation and when it transmits, it transmits $[X_t, \hat{a}_t]^\intercal$, where $\hat{a}_t$ denotes its current estimate. The receiver on receiving the vector $[X_t, \hat{a}_t]^\intercal$ solves a filtering problem and generates its estimate of the state realization using $\hat{a}_t$ as the model parameter.

For uniformity of notation, define $\mathbb{X}$ to be equal to $\mathbb{Z}$ for Model A and equal to $\mathbb{R}$ for Model B. $\mathbb{X}_{\geq 0}$ and $\mathbb{X}_{>0}$ are defined similarly.

A sensor sequentially observes the process and at each time, chooses whether or not to transmit the current state. This decision is denoted by $U_t \in \{0, 1\}$, where $U_t = 0$ denotes no transmission and $U_t = 1$ denotes transmission. The decision to transmit is made using a *transmission strategy* $f = \{f_t\}_{t=0}^{\infty}$, where

$$U_t = f_t(X_{0:t}, U_{0:t-1}). \tag{2.2}$$

We use the short-hand notation $X_{0:t}$ to denote the sequence $(X_0, \dots, X_t)$. Similar interpretations hold for $U_{0:t-1}$.

The transmitted symbol, which is denoted by Y_t, is given by

$$Y_t = \begin{cases} X_t, & \text{if } U_t = 1; \\ \mathfrak{E}, & \text{if } U_t = 0, \end{cases}$$

where $Y_t = \mathfrak{E}$ denotes no transmission.

The receiver sequentially observes $\{Y_t\}_{t=0}^{\infty}$ and generates an estimate $\{\hat{X}_t\}_{t=0}^{\infty}$, $\hat{X} \in \mathbb{X}$, using an *estimation strategy* $g = \{g_t\}_{t=0}^{\infty}$, i.e.,

$$\hat{X}_t = g_t(Y_{0:t}). \tag{2.3}$$

The fidelity of the estimation is measured by a per-step distortion $d(X_t - \hat{X}_t)$. Also, it is assumed that the pmf (for Model A. Alternatively, the pdf for Model B) of W_t is known to both the sensor and the receiver.

For both models, we assume the following: for any $e \in \mathbb{X}$,

- $d(0) = 0$ and for $e \neq 0$, $d(e) > 0$;
- $d(\cdot)$ is even, i.e., for all e, $d(e) = d(-e)$;
- $d(\cdot)$ is increasing, i.e., for $e_1 > e_2 > 0$, $e_1, e_2 \in \mathbb{X}$, $d(e_1) \geq d(e_2)$;
- For Model B, we assume that $d(\cdot)$ is differentiable.

We also characterize our results to the following special case of Model B:

- **Gauss-Markov model**: the density ϕ is zero-mean Gaussian with variance σ^2 and the distortion is quadratic, i.e., for any $e \in \mathbb{X}$,

$$\phi(e) = \frac{1}{\sqrt{2\pi}\sigma} \exp\big(-e^2/(2\sigma^2)\big) \quad \text{and} \quad d(e) = e^2.$$

2.3.2 Performance measures

Given a transmission and estimation strategy (f, g) and a discount factor $\beta \in (0, 1]$, we define the expected distortion and the expected number of transmissions as follows. For $\beta \in (0, 1)$, the expected *discounted* distortion is given by

$$D_\beta(f, g) \coloneqq (1-\beta)\mathbb{E}^{(f,g)}\Big[\sum_{t=0}^{\infty} \beta^t d(X_t - \hat{X}_t) \;\Big|\; X_0 = 0\Big] \tag{2.4}$$

and for $\beta = 1$, the expected *long-term average* distortion is given by

$$D_1(f, g) \coloneqq \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}^{(f,g)}\Big[\sum_{t=0}^{T-1} d(X_t - \hat{X}_t) \;\Big|\; X_0 = 0\Big]. \tag{2.5}$$

Similarly, for $\beta \in (0, 1)$, the expected *discounted* number of transmissions is given by

$$N_\beta(f, g) \coloneqq (1-\beta)\mathbb{E}^{(f,g)}\Big[\sum_{t=0}^{\infty} \beta^t U_t \;\Big|\; X_0 = 0\Big] \tag{2.6}$$

and for $\beta = 1$, the expected *long-term average* number of transmissions is given by

$$N_1(f, g) \coloneqq \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}^{(f,g)}\Big[\sum_{t=0}^{T-1} U_t \;\Big|\; X_0 = 0\Big]. \tag{2.7}$$

Remark 2 We use a normalizing factor of $(1-\beta)$ to have a unified scaling for both discounted and long-term average setups. In particular, we will show that for any strategy (f, g)

$$C_1(f, g; \lambda) = \lim_{\beta\uparrow 1} C_\beta(f, g; \lambda), \quad \text{and} \quad D_1(f, g) = \lim_{\beta\uparrow 1} D_\beta(f, g),$$

where λ is the per-step transmission cost and $C_\beta(f, g; \lambda) = D_\beta(f, g) + \lambda N_\beta(f, g)$. Similar notation is used in [42].

2.3.3 Problem formulations

Let us call the tuple (f, g) the *communication strategy*. We are interested in the following two optimization problems.

Problem 2.3.1 (Costly communication) *In the model of Section 2.3.1, given a discount factor $\beta \in (0, 1]$ and a communication cost $\lambda \in \mathbb{R}_{>0}$, find a communication strategy (f^*, g^*) such that*

$$C^*_\beta(\lambda) \coloneqq C_\beta(f^*, g^*; \lambda) = \inf_{(f,g)} C_\beta(f, g; \lambda), \tag{2.8}$$

where

$$C_\beta(f, g; \lambda) \coloneqq D_\beta(f, g) + \lambda N_\beta(f, g)$$

is the total communication cost and the infimum in (2.8) is taken over all history-dependent strategies.

Problem 2.3.2 (Constrained communication) *In the model of Section 2.3.1, given a discount factor $\beta \in (0, 1]$ and a constraint $\alpha \in (0, 1)$, find a communication strategy (f^*, g^*) such that*

$$D^*_\beta(\alpha) \coloneqq D_\beta(f^*, g^*) = \inf_{(f,g):N_\beta(f,g)\le\alpha} D_\beta(f, g), \tag{2.9}$$

where the infimum is taken over all history-dependent strategies.

Remark 3 It can be shown for $|a| < 1$, $\lim_{\alpha\to 0} D^*_1(\alpha) < \infty$ and for $|a| \ge 1$ that $\lim_{\alpha\to 0} D^*_1(\alpha) = \infty$[1], and in both cases $\lim_{\alpha\to 1} D^*_\beta(\alpha) = 0$.

The function $D^*_\beta(\alpha)$, $\beta \in (0, 1]$, represents the minimum expected distortion that can be achieved when the expected number of transmissions are less than or equal to α. It is analogous to the distortion-rate function in Information Theory; for that reason, we call it the *distortion-transmission function.*

[1]For $|a| < 1$, a symmetric Markov chain as given by (2.1) has a stationary distribution whereas for $|a| \ge 1$, (2.1) does not. Therefore, in the limit of no transmission, the expected long-term average distortion is finite for $|a| < 1$ and diverges to ∞ for $|a| \ge 1$.

2.4 The main results

2.4.1 Structure of optimal strategies

To completely characterize the functions $C^*_\beta(\lambda)$ and $D^*_\beta(\alpha)$, we first establish the structure of optimal transmitter and receiver.

Theorem 2.4.1 (Structural results) *Consider Problem 2.3.1 for $\beta \in (0, 1]$. Then, for both Models A and B, we have the following.*

1. Structure of optimal estimation strategy*: The optimal estimation strategy $\hat{X}_0 = 0$ and for $t > 0$ is as follows:*

$$\hat{X}_t = \begin{cases} Y_t, & \text{if } Y_t \neq \mathfrak{E} \\ a\hat{X}_{t-1}, & \text{if } Y_t = \mathfrak{E}, \end{cases}$$

or equivalently,

$$\hat{X}_t = \begin{cases} X_t, & \text{if } U_t = 1 \\ a\hat{X}_{t-1}, & \text{if } U_t \neq 1. \end{cases}$$

We denote this strategy by g^*.

2. Structure of optimal transmission strategy*: Define $E_t := X_t - a\hat{X}_{t-1}$, which we call the* error process*. Then there exists a time-invariant threshold k such that the transmission strategy*

$$U_t = f^{(k)}(E_t) := \begin{cases} 1, & \text{if } |E_t| \geq k \\ 0, & \text{if } |E_t| < k \end{cases} \tag{2.10}$$

is optimal.

The proof of the theorem is given in Section 2.6.

Similar structural results were established for the finite horizon setup in [9–11], which we use to establish Theorem 2.4.1. See Section 2.6 for details. The transmission strategy of the form (2.10) are also called *event-driven transmission* or *delta sampling.*

Remark 4 Each transmission *resets* the state of the error process to $w \in \mathbb{X}$ with probability p_w in Model A and with probability density $\phi(w)$ in Model B. In between two consecutive transmissions, the error process evolves in a Markovian manner. Thus $\{E_t\}_{t=0}^{\infty}$ is a regenerative process.

2.4.2 Performance of generic threshold-based strategies

Let $\mathcal{F}$ denote the class of all time-homogeneous threshold-based strategies of the form (2.10). For $\beta \in (0,1]$ and $e \in \mathbb{X}$, define the following for a system that starts in state e and follows strategy $f^{(k)}$:

- $L_\beta^{(k)}(e)$: the expected distortion until the first transmission;
- $M_\beta^{(k)}(e)$: the expected time until the first transmission;
- $D_\beta^{(k)}(e)$: the expected distortion;
- $N_\beta^{(k)}(e)$: the expected number of transmissions;
- $C_\beta^{(k)}(e;\lambda)$: the expected total cost, i.e.,

$$C_\beta^{(k)}(e;\lambda) = D_\beta^{(k)}(e) + \lambda N_\beta^{(k)}(e), \quad \lambda \geq 0.$$

Note that $D_\beta^{(k)}(0) = D_\beta(f^{(k)}, g^*)$, $N_\beta^{(k)}(0) = N_\beta(f^{(k)}, g^*)$ and $C_\beta^{(k)}(0;\lambda) = C_\beta(f^{(k)}, g^*;\lambda)$. Define $S^{(k)}$ as follows:

$$S^{(k)} := \begin{cases} \{-(k-1), \cdots, k-1\}, & \text{for Model A;} \\ (-k, k), & \text{for Model B.} \end{cases}$$

Under strategy $f^{(k)}$, the transmitter does not transmit if $E_t \in S^{(k)}$. For that reason, we call $S^{(k)}$ the *silent set*. Define linear operator $\mathcal{B}^{(k)}$ as follows:

- **Model A**: For any $v^{(k)} : S^{(k)} \to \mathbb{R}$, define operator $\mathcal{B}^{(k)}$ as

$$[\mathcal{B}^{(k)} v](e) := \sum_{n \in S^{(k)}} p_{n-ae} v(n), \quad \forall e \in S^{(k)}.$$

- **Model B**: For any $v^{(k)} : S^{(k)} \to \mathbb{R}$, define operator $\mathcal{B}^{(k)}$ as

$$[\mathcal{B}^{(k)} v](e) := \int_{S^{(k)}} \phi(n - ae) v(n) dn, \quad \forall e \in S^{(k)}.$$

Recall from Remark 4 that the state E_t evolves in a Markovian manner until the first transmission. We may equivalently consider the Markov process until it is absorbed in $(-\infty, -k] \cup [k, \infty)$. Thus, from balance equation for Markov processes, we have for all $e \in S^{(k)}$,

$$L_\beta^{(k)}(e) = d(e) + \beta[\mathcal{B}^{(k)} L_\beta^{(k)}](e), \tag{2.11}$$

$$M_\beta^{(k)}(e) = 1 + \beta[\mathcal{B}^{(k)} M_\beta^{(k)}](e). \tag{2.12}$$

Lemma 2.4.1 *For any $\beta \in (0, 1]$, equations* (2.11) *and* (2.12) *have unique and bounded solutions $L_\beta^{(k)}$ and $M_\beta^{(k)}$ that are*

(a) strictly increasing in k,

(b) continuous and differentiable in k for Model B,

(c) $\lim_{\beta \uparrow 1} L_\beta^{(k)}(e) = L_1^{(k)}(e)$, $\lim_{\beta \uparrow 1} M_\beta^{(k)}(e) = M_1^{(k)}(e)$, for all e.

The proof of the lemma is given in Appendix A.2.

Theorem 2.4.2 (Renewal relationships) *For any $\beta \in (0, 1]$, the performance of strategy $f^{(k)}$ in both Models A and B is given as follows:*

1. *$D_\beta(f^{(0)}, g^*) = 0$, $N_\beta(f^{(0)}, g^*) = 1$, and $C_\beta(f^{(0)}, g^*; \lambda) = \lambda$.*

2. *For $k \in \mathbb{X}_{>0}$,*

$$D_\beta(f^{(k)}, g^*) = \frac{L_\beta^{(k)}(0)}{M_\beta^{(k)}(0)},$$

$$N_\beta(f^{(k)}, g^*) = \frac{1}{M_\beta^{(k)}(0)} - (1 - \beta),$$

and

$$C_\beta(f^{(k)}, g^*; \lambda) = \frac{L_\beta^{(k)}(0) + \lambda}{M_\beta^{(k)}(0)} - \lambda(1 - \beta).$$

The proof of the theorem is given in Section 2.7.

Remark 5 There is a $-1/(1-\beta)$ term in the expression of $N_\beta^{(k)}(0)$ because for $k > 0$, $U_0 = 0$. Had we defined $U_0 = 1$, then we would have obtained the usual renewal relationship of $N_\beta^{(k)}(0) = 1/M_\beta^{(k)}(0)$.

Thus, to compute $D_\beta(f^{(k)}, g^*)$ and $N_\beta(f^{(k)}, g^*)$, one needs to compute only $L_\beta^{(k)}(0)$ and $M_\beta^{(k)}(0)$. Computation of the latter expressions is given in the next section.

Proposition 2.4.1 *For both Models A and B,*

1. $C_\beta^{(k)}(0;\lambda)$ *is submodular in* (k,λ), *i.e., for* $l > k$, $C_\beta^{(l)}(0;\lambda) - C_\beta^{(k)}(0;\lambda)$ *is decreasing in* λ.

2. *Let* $k_\beta^*(\lambda) = \arg\inf_{k\geq 0} C_\beta^{(k)}(0;\lambda)$ *be the optimal* k *for a fixed* λ. *Then* $k_\beta^*(\lambda)$ *is increasing in* λ.

The proof of the proposition is in Appendix A.3.

2.4.3 Computation of $L_\beta^{(k)}$ and $M_\beta^{(k)}$

Model A

For Model A, the values of $L_\beta^{(k)}$ and $M_\beta^{(k)}$ can be computed by observing that the operator $\mathcal{B}^{(k)}$ is equivalent to a matrix multiplication. In particular, define the matrix $P^{(k)}$ as

$$P_{ij}^{(k)} := p_{i-j}, \quad \forall i,j \in S^{(k)}.$$

Then,

$$[\mathcal{B}^{(k)}v](e) = \sum_{n\in S^{(k)}} p_{n-ae}v(n) = \sum_{n\in S^{(k)}} P_{n,ae}^{(k)} v(n) = [P^{(k)}v]_{ae}. \tag{2.13}$$

With a slight abuse of notation, we are using v both as a function and a vector. Define the matrix $Q^{(k)}$ and the vector $d^{(k)}$ as follows:

$$Q_\beta^{(k)} := [I_{2k-1} - \beta P^{(k)}]^{-1}, \quad d^{(k)} := [d(-k+1), \ldots, d(k-1)]^\intercal.$$

Then, (2.11), (2.12) and (2.13) imply the following:

Proposition 2.4.2 *In Model A, for any $\beta \in (0,1]$,*

$$L_\beta^{(k)} = [I_{2k-1} - \beta P^{(k)}]^{-1} d^{(k)} \tag{2.14}$$
$$M_\beta^{(k)} = [I_{2k-1} - \beta P^{(k)}]^{-1} \mathbf{1}_{2k-1}. \tag{2.15}$$

See Section 2.4.6 for an example of these calculations.

Model B

For Model B, for any $\beta \in (0,1]$, (2.11) and (2.12) are Fredholm integral equations of second kind [43]. The solution can be computed by identifying the inverse operator

$$\mathcal{Q}_\beta^{(k)} = [I - \beta \mathcal{B}^{(k)}]^{-1},$$

which is given by

$$[\mathcal{Q}_\beta^{(k)} v](e) = \int_{-k}^{k} R_\beta^{(k)}(e, w; a) v(w) dw, \tag{2.16}$$

where for any given a, $R_\beta^{(k)}(\cdot, \cdot; a)$ is the resolvent of ϕ and can be computed using the Liouville-Neumann series. See [43] for details. Since ϕ is smooth, (2.11) and (2.12) can also be solved by discretizing the integral equation using quadrature methods. A Matlab implementation of this approach is available in [44].

2.4.4 Main results for Model A

Results for costly communication

Theorem 2.4.3 *For $\beta \in (0,1]$, let $\mathbb{K}$ denote $\{k \in \mathbb{Z}_{\geq 0} : D_\beta^{(k+1)}(0) > D_\beta^{(k)}(0)\}$. For $k_n \in \mathbb{K}$, define:*

$$\lambda_\beta^{(k_n)} := \frac{D_\beta^{(k_{n+1})}(0) - D_\beta^{(k_n)}(0)}{N_\beta^{(k_n)}(0) - N_\beta^{(k_{n+1})}(0)}. \tag{2.17}$$

Then, we have the following.

1. *For any $k_n \in \mathbb{K}$ and any $\lambda \in (\lambda_\beta^{(k_{n-1})}, \lambda_\beta^{(k_n)}]$, the strategy $f^{(k_n)}$ is optimal for Problem 2.3.1 with communication cost λ.*

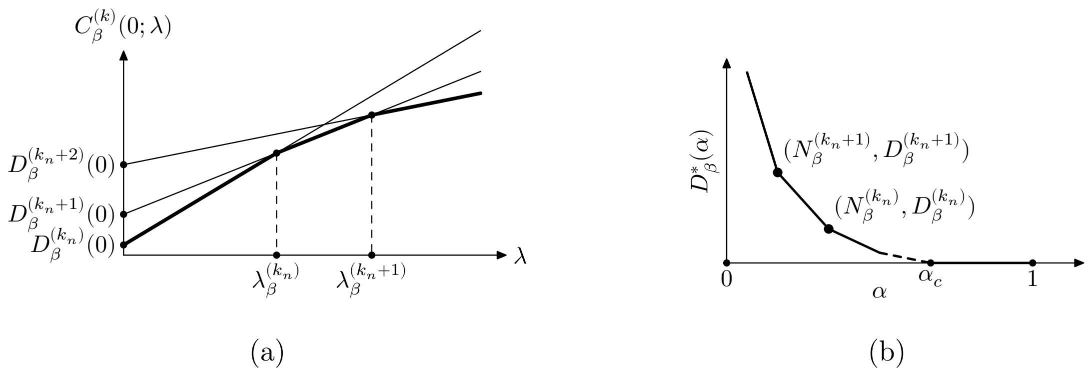


Fig. 2.2 In Model A, (a) the optimal costly communication cost $C^*_\beta(\lambda)$; (b) the distortion-transmission function $D^*_\beta(\alpha)$.

2. *The optimal performance $C^*_\beta(\lambda)$ is continuous, concave, increasing and piecewise linear in λ. The corner points of $C^*_\beta(\lambda)$ are given by $\{(\lambda_\beta^{(k_n)}, D_\beta^{(k_n)}(0)+\lambda_\beta^{(k_n)} N_\beta^{(k_n)}(0))\}_{k_n \in \mathbb{K}}$ (see Fig 2.2(a)).*

The proof of the theorem is given in Section 2.8.

Results for constrained communication

To describe the solution of Problem 2.3.2, we first define Bernoulli randomized strategy and Bernoulli randomized simple strategy [45].

Definition 2.4.1 *Suppose we are given two (non-randomized) time-homogeneous strategies f_1 and f_2 and a randomization parameter $\theta \in (0,1)$. The* Bernoulli randomized strategy *(f_1, f_2, θ) is a strategy that randomizes between f_1 and f_2 at each stage; choosing f_1 with probability θ and f_2 with probability $(1-\theta)$. Such a strategy is called a Bernoulli randomized* simple *strategy if f_1 and f_2 differ on exactly one state, i.e., there exists a state e_0 such that*

$$f_1(e) = f_2(e), \quad \forall e \neq e_0.$$

Theorem 2.4.4 *For any $\beta \in (0,1]$ and $\alpha \in (0,1)$, define*

$$\begin{aligned} k^*_\beta(\alpha) &= \sup\{k \in \mathbb{Z}_{\geq 0} : N_\beta(f^{(k)}, g^*) \geq \alpha\} \\ &= \sup\left\{k \in \mathbb{Z}_{\geq 0} : M_\beta^{(k)} \leq \frac{1}{1+\alpha-\beta}\right\} \end{aligned} \tag{2.18}$$

and

$$\begin{aligned}\theta^*_\beta(\alpha) &= \frac{\alpha - N_\beta(f^{(k^*_\beta(\alpha)+1)}, g^*)}{N_\beta(f^{(k^*_\beta(\alpha))}, g^*) - N_\beta(f^{(k^*_\beta(\alpha)+1)}, g^*)} \\ &= \frac{M^{(k^*+1)}_\beta - \frac{1}{1+\alpha-\beta}}{M^{(k^*+1)}_\beta - M^{(k^*)}_\beta}. \end{aligned} \tag{2.19}$$

For ease of notation, we use $k^* = k^*_\beta(\alpha)$ *and* $\theta^* = \theta^*_\beta(\alpha)$.

Let f^* *be the Bernoulli randomized simple strategy* $(f^{(k^*)}, f^{(k^*+1)}, \theta^*)$, *i.e.,*

$$f^*(e) = \begin{cases} 0, & \text{if } |e| < k^*; \\ 0, & \text{w.p. } 1-\theta^*, \text{ if } |e| = k^*; \\ 1, & \text{w.p. } \theta^*, \text{ if } |e| = k^*; \\ 1, & \text{if } |e| > k^*. \end{cases} \tag{2.20}$$

Then

1. (f^*, g^*) *is optimal for the constrained Problem 2.3.2 with constraint* α.

2. *Let* $\alpha^{(k)} = N_\beta(f^{(k)}, g^*)$. *Then, for* $\alpha \in (\alpha^{(k+1)}, \alpha^{(k)})$, $k^* = k$ *and* $\theta^* = (\alpha - \alpha^{(k+1)})/(\alpha^{(k)} - \alpha^{(k+1)})$, *and the distortion-transmission function is given by*

$$D^*_\beta(\alpha) = \theta^* D^{(k)}_\beta + (1-\theta^*) D^{(k+1)}_\beta. \tag{2.21}$$

Moreover, the distortion-transmission function is is continuous, convex, decreasing and piecewise linear in α. *Thus, the corner points of* $D^*_\beta(\alpha)$ *are given by* $\{(N^{(k)}_\beta(0), D^{(k)}_\beta(0))\}_{k=1}^\infty$ *(see Fig 2.2(b)).*

The proof of the theorem is given in Section 2.8.

Corollary 2.4.1 *In Model A, for any* $\beta \in (0,1]$,

$$D_\beta(f^{(1)}, g^*) = 0, \quad \text{and} \quad N_\beta(f^{(1)}, g^*) = \beta(1-p_0) \coloneqq \alpha_c.$$

Algorithm 1: Computation of $C^*_\beta(\lambda)$ for Model B

input : $\lambda \in \mathbb{R}_{>0}$, $\beta \in (0,1]$, $\varepsilon \in \mathbb{R}_{>0}$
output: $C_\beta^{(k^\circ)}(\lambda)$, where $|k^\circ - k^*_\beta(\lambda)| < \varepsilon$

Let $\lambda^*_\beta(k)$ denote the left-hand side of (2.22), which is computed by finite-difference method
Pick $\underline{k}$ and $\bar{k}$ such that $\lambda^*_\beta(\underline{k}) < \lambda < \lambda^*_\beta(\bar{k})$
$k^\circ \leftarrow (\underline{k} + \bar{k})/2$
while $|\lambda^*_\beta(k^\circ) - \lambda| > \varepsilon$ **do**
  if $\lambda^*(k^\circ) < \lambda$ **then**
    $\underline{k} \leftarrow k^\circ$
  else
    $\bar{k} \leftarrow k^\circ$
  $k^\circ \leftarrow (\underline{k} + \bar{k})/2$
return $D_\beta^{(k^\circ)}(0) + \lambda N_\beta^{(k^\circ)}(0)$

2.4.5 Main results for Model B

Results for costly communication

Let $\partial_k D_\beta^{(k)}$, $\partial_k N_\beta^{(k)}$ and $\partial_k C_\beta^{(k)}$ denote the derivative of $D_\beta^{(k)}$, $N_\beta^{(k)}$ and $C_\beta^{(k)}$ with respect to k (in Lemma 2.9.1 we show that $D_\beta^{(k)}$, $N_\beta^{(k)}$ and $C_\beta^{(k)}$ are differentiable in k).

Theorem 2.4.5 *For $\beta \in (0,1]$, we have the following.*

1. *If the pair (λ, k) satisfies the following*

$$\lambda = -\frac{\partial_k D_\beta^{(k)}(0)}{\partial_k N_\beta^{(k)}(0)}, \tag{2.22}$$

   then, the strategy $(f^{(k)}, g^*)$ is optimal for Problem 2.3.1 with communication cost λ. Furthermore, for any $k > 0$, there exists a $\lambda \geq 0$ that satisfies (3.73).

2. *The optimal performance $C^*_\beta(\lambda)$ is continuous, concave and increasing function of λ.*

The proof of the theorem is given in Section 2.9. Algorithm 1 shows how to compute $C^*_\beta(\lambda)$.

Algorithm 2: Computation of $D^*_\beta(\alpha)$ for Model B

input : $\alpha \in (0,1)$, $\beta \in (0,1]$, $\varepsilon \in \mathbb{R}_{>0}$
output: $D^{(k^\circ)}_\beta(\alpha)$, where $|N^{(k^\circ)}_\beta(0) - \alpha| < \varepsilon$

Pick $\underline{k}$ and $\bar{k}$ such that $N^{(\underline{k})}_\beta(0) < \alpha < N^{(\bar{k})}_\beta(0)$
$k^\circ \leftarrow (\underline{k} + \bar{k})/2$
while $|N^{(k^\circ)}_\beta(0) - \alpha| > \varepsilon$ **do**
    if $N^{(k^\circ)}_\beta(0) < \alpha$ **then**
        $\underline{k} \leftarrow k^\circ$
    else
        $\bar{k} \leftarrow k^\circ$
    $k^\circ \leftarrow (\underline{k} + \bar{k})/2$
return $D^{(k^\circ)}_\beta(\alpha)$

Results for constrained communication

Theorem 2.4.6 *For any $\beta \in (0,1]$ and $\alpha \in (0,1)$, let $k^*_\beta(\alpha) \in \mathbb{R}_{\ge 0}$ be such that*

$$N^{(k^*_\beta(\alpha))}_\beta(0) = \alpha. \tag{2.23}$$

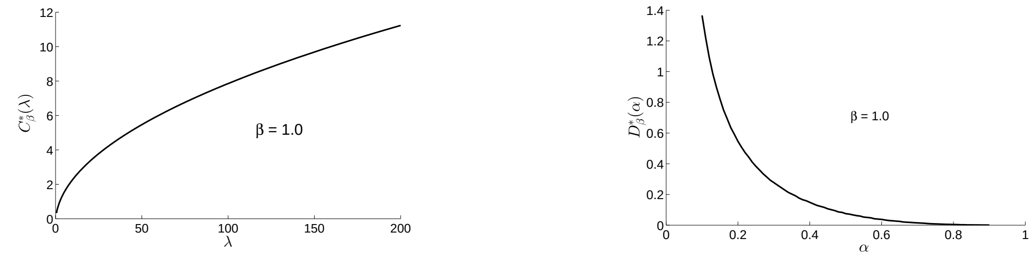
Such a $k^*_\beta(\alpha)$ always exists and we have the following:

1. *The strategy $(f^{(k^*_\beta(\alpha))}, g^*)$ is optimal for Problem 2.3.2 with constraint α.*
2. *The distortion-transmission function $D^*_\beta(\alpha)$ is continuous, convex and decreasing in α and is given by*
$$D^*_\beta(\alpha) = D^{(k^*_\beta(\alpha))}_\beta(0). \tag{2.24}$$

The proof of the theorem is given in Section 2.9. Algorithm 2 shows how to compute $D^*_\beta(\alpha)$.

Special case of Model B–Gauss-Markov model

In general, the optimal thresholds, and the functions $C^*_\beta(\lambda)$ and $D^*_\beta(\alpha)$ depend on the noise distribution $\phi(\cdot)$. For the Gauss-Markov model, the dependence on the variance σ^2 of the noise may be quantified exactly.



(a) (b)

Fig. 2.3 Gauss-Markov model ($\sigma^2 = 1$ and $a = 1$): (a) optimal costly communication cost $C_1^*(\alpha)$; (b) distortion-transmission function $D_1^*(\alpha)$.

For ease of notation, we drop the dependence on β from the notation, and instead, show the dependence on σ. Thus, $C_\sigma^*(\lambda)$ denotes the optimal value for the costly communication case when the noise variance is σ^2. Similar notation holds for other terms.

Theorem 2.4.7 *For the Gauss-Markov model for Problem 2.3.1, $k_\sigma^*(\lambda) = k_1^*(\lambda/a^2\sigma^2)$ and $C_\sigma^*(\lambda) = \sigma^2 C_1^*(\lambda/\sigma^2)$. For Problem 2.3.2, $k_\sigma^*(\alpha) = \sigma k_1^*(\alpha)$ and $D_\sigma^*(\alpha) = \sigma^2 D_1^*(\alpha)$.*

The proof of the theorem is given in Section 2.9.

An implication of the above theorem is that we only need to numerically compute $C_1^*(\lambda)$ and $D_1^*(\alpha)$, which are shown in Fig. 2.3. The optimal total communication cost and the distortion-transmission function for any other value σ^2 can be obtained by simply scaling $C_1^*(\lambda)$ and $D_1^*(\alpha)$ respectively.

2.4.6 An example for Model A: symmetric birth-death Markov chain

An example of a Markov process and a distortion function that satisfy Model A is the following:

Example 2.4.1 *Consider a Markov chain of the form* (2.1) *where the pmf (probability mass function) of W_t is given by*

$$p_n = \begin{cases} p, & \text{if } |n| = 1 \\ 1 - 2p, & \text{if } n = 0 \\ 0, & \text{otherwise,} \end{cases}$$

Table 2.1 Values of $D_\beta^{(k)}$, $N_\beta^{(k)}$ and $\lambda_\beta^{(k)}$ for different values of k and β for the Markov chain of Example 2.4.1 with $p = 0.3$. Note that $D_\beta^{(0)}(0) = D_\beta^{(1)}(0)$; therefore $\mathbb{K}$ defined in Theorem 2.4.3 equals $\mathbb{Z}_{>0}$.

(a) For $\beta = 0.9$ (b) For $\beta = 0.95$ (c) For $\beta = 1.0$

k	$D_\beta^{(k)}(0)$	$N_\beta^{(k)}(0)$	$\lambda_\beta^{(k)}$	k	$D_\beta^{(k)}(0)$	$N_\beta^{(k)}(0)$	$\lambda_\beta^{(k)}$	k	$D_\beta^{(k)}(0)$	$N_\beta^{(k)}(0)$	$\lambda_\beta^{(k)}$
0	0	1	–	0	0	1	–	0	0	1	–
1	0	0.5400	1.0989	1	0	0.5700	1.1050	1	0	0.6000	1.1111
2	0.4576	0.1236	4.1021	2	0.4790	0.1365	4.3657	2	0.5000	0.1500	4.6667
3	0.7695	0.0475	9.2839	3	0.8282	0.0565	10.6058	3	0.8889	0.0667	12.3810
4	1.0066	0.0220	16.2509	4	1.1218	0.0288	19.9550	4	1.2500	0.0375	25.9259
5	1.1844	0.0111	24.4478	5	1.3715	0.0163	32.0869	5	1.6000	0.0240	46.9697
6	1.3130	0.0058	33.4121	6	1.5811	0.0098	46.4727	6	1.9444	0.0167	77.1795
7	1.4029	0.0031	42.8289	7	1.7536	0.0061	62.5651	7	2.2857	0.0122	118.2222
8	1.4638	0.0017	52.5042	8	1.8927	0.0039	79.8921	8	2.6250	0.0094	171.7647
9	1.5040	0.0009	62.3245	9	2.0028	0.0025	98.0854	9	2.9630	0.0074	239.4737
10	1.5298	0.0005	72.2255	10	2.0884	0.0016	116.8739	10	3.0000	0.0060	323.0159

where $p \in (0, \frac{1}{3})$. *The distortion function is taken as* $d(e) = |e|$.

This Markov process corresponds to a symmetric, birth-death Markov chain defined over $\mathbb{Z}$ as shown in Fig. 2.4, with the transition probability matrix is given by

$$P_{ij} = \begin{cases} p, & \text{if } |i - j| = 1; \\ 1 - 2p, & \text{if } i = j; \\ 0, & \text{otherwise.} \end{cases}$$

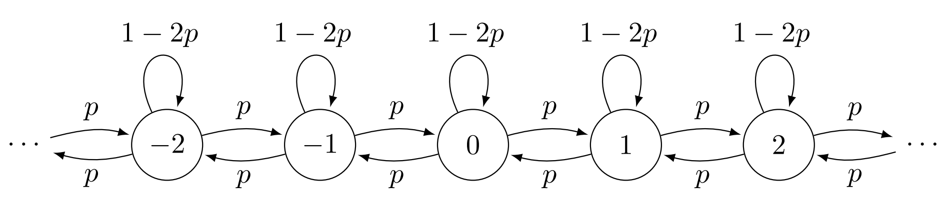
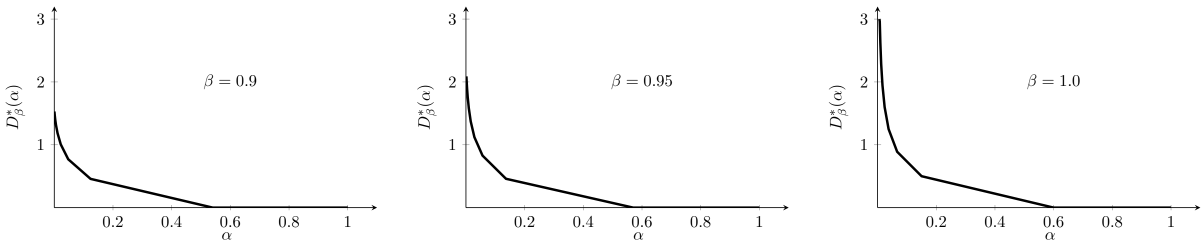


Fig. 2.4 A birth-death Markov chain



(a) $D^*_\beta(\alpha)$ vs α for $\beta = 0.9$ (b) $D^*_\beta(\alpha)$ vs α for $\beta = 0.95$ (c) $D^*_\beta(\alpha)$ vs α for $\beta = 1.0$

Fig. 2.5 Plots of $D^*_\beta(\alpha)$ vs α for different β for the birth-death Markov chain of Example 2.4.1 with $p = 0.3$.

Performance of a generic threshold-based strategy

Lemma 2.4.2 *Define for* $\beta \in (0, 1]$

$$K_\beta = -2 - \frac{(1-\beta)}{\beta p} \quad \textit{and} \quad m_\beta = \cosh^{-1}(-K_\beta/2).$$

Then,

1. *For* $\beta \in (0, 1)$,

$$D^{(k)}_\beta(0) = \frac{\sinh(k m_\beta) - k \sinh(m_\beta)}{2 \sinh^2(k m_\beta/2) \sinh(m_\beta)};$$
$$N^{(k)}_\beta(0) = \frac{2\beta p \sinh^2(m_\beta/2) \cosh(k m_\beta)}{\sinh^2(k m_\beta/2)} - (1-\beta).$$

2. *For* $\beta = 1$,

$$D^{(k)}_1 = \frac{k^2 - 1}{3k}; \quad N^{(k)}_1 = \frac{2p}{k^2};$$

and

$$\lambda^{(k)}_1 = \frac{k(k+1)(k^2+k+1)}{6p(2k+1)}.$$

The proof is given in Section 2.10.

Optimal strategy for costly communication

Using the above expressions for $D^{(k)}_\beta(0)$ and $N^{(k)}_\beta(0)$, we can identify $\mathbb{K}$ and for each $k_n \in \mathbb{K}$, compute $\lambda^{(k_n)}_\beta$ according to (3.68). These values are tabulated in Table 2.1 for differ-

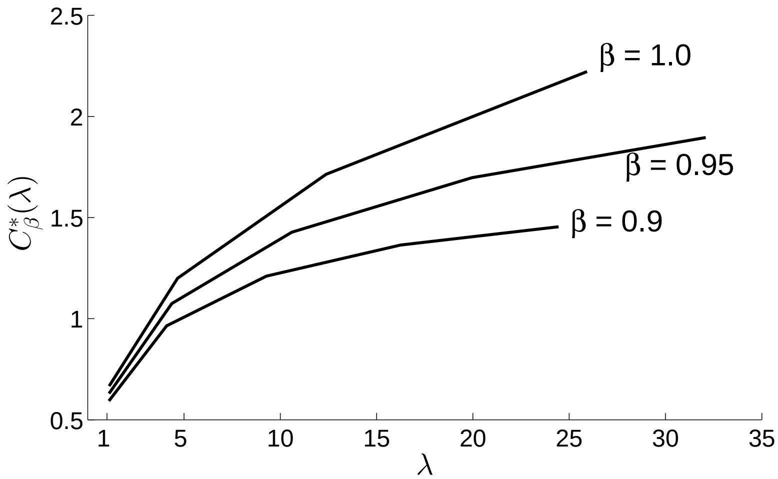


Fig. 2.6 Plot of $C^*_\beta(\lambda)$ vs λ for the Markov chain of Example 2.4.1 with $p = 0.3$.

ent values of β (all for $p = 0.3$). Using Table 2.1, we can compute the corner points $(\lambda^{(k_n)}_\beta, D^{(k_n)}_\beta(0) + \lambda^{(k_n)}_\beta N^{(k_n)}_\beta(0))$ of $C^*_\beta(\lambda)$. Joining these points by straight lines gives $C^*_\beta(\lambda)$, as shown in Fig. 2.6. The optimal strategy for a given λ can be computed from Table 2.1.

For example, for $\lambda = 20$, $\beta = 0.9$, we can find from Table 2.1a that $\lambda \in (\lambda^{(4)}_\beta, \lambda^{(5)}_\beta]$. Hence, $k^*_\beta = 5$ (i.e., the strategy $f^{(5)}$ is optimal) and the optimal total communication cost is

$$C^*_{0.9}(20) = D^{(5)}_{0.9}(0) + 20N^{(5)}_{0.9}(0) = 1.1844 + 20 \times 0.0111 = 1.4064.$$

Optimal strategy for constrained communication

Using the values in Table 2.1, we can also compute the corner points $(N^{(k)}_\beta(0), D^{(k)}_\beta(0))$ of $D^*_\beta(\alpha)$. Joining these points by straight lines gives $D^*_\beta(\alpha)$ (see Fig. 2.5). The optimal strategy for a given α can be computed from Table 2.1. For example, at $\alpha = 0.1$ and $\beta = 0.9$, $k^*_\beta(\alpha)$ is the largest value of k such that $N^{(k)}_\beta(0) \geq \alpha$. Thus, from Table 2.1a, we get that $k^* = 2$. Then, by (3.74),

$$\theta^* = \frac{\alpha - N^{(3)}_\beta}{N^{(2)}_\beta - N^{(3)}_\beta} = 0.6899.$$

Let $f^* = (f^{(2)}, f^{(3)}, \theta^*)$. Then the Bernoulli randomized simple strategy (f^*, g^*) is optimal for Problem 2.3.2 for $\beta \in (0,1)$. Furthermore, by (3.72), $D^*_\beta(\alpha) = 0.5543$.

2.5 Salient features and discussion

2.5.1 Comparison with periodic and randomized strategies

In our model, we assume that the transmission decision depends on the state of the Markov process. In some of the remote estimation literature, it is assumed that the transmission schedule does not depend on the state of the Markov process. Two such commonly used strategies are:

1. Periodic transmission strategy with period T:

$$U_t = f_p(t \bmod T),$$

   where $\sum_{t=0}^{T-1} f_p(t) = 1/\alpha$.

2. Random transmission strategy:

$$U_t = \begin{cases} 1, & \text{w.p. } \alpha \\ 0, & \text{w.p. } 1-\alpha. \end{cases}$$

Below, we compare the performance of the threshold-based strategy with these two strategies for the long-term average setup for Problem 2.3.2 for Model B with $a = 1$.

Performance of the periodic strategy

In general, the performance of a periodic transmission strategy depends on the choice of transmission function f_p. For ease of calculation we consider the values of (α, T) for which f_p is unique.

1. $\alpha = 1/T$, $T \in \mathbb{Z}_{>0}$, i.e., the transmitter remains silent for $(T-1)$ steps and then

transmits once. The expected distortion in this case is

$$\begin{aligned} D_{\text{per}}(\alpha) &= \frac{1}{T}\mathbb{E}\Big[\sum_{t=0}^{T-1} E_t^2\Big] \\ &\stackrel{(a)}{=} \frac{1}{T}\mathbb{E}\Big[\sum_{t=0}^{T-1} t\sigma^2\Big] = \frac{1}{T}\frac{(T-1)T}{2}\sigma^2 = \frac{\sigma^2}{2}\big(\frac{1}{\alpha}-1\big), \end{aligned}$$

where (a) uses $E_t = W_0 + W_1 + W_2 + \cdots + W_{t-1}$.

2. $\alpha = (T-1)/T$, $T \in \mathbb{Z}_{>0}$, i.e., the transmitter remains silent for 1 step and then transmits for $(T-1)$ steps. The expected distortion in this case is

$$D_{\text{per}}(\alpha) = \frac{1}{T}\mathbb{E}[E_1^2] = \frac{\sigma^2}{T} = \sigma^2(1-\alpha).$$

Performance generic stationary transmission strategy

Next, we derive an expression of $D_\beta(f, g^*)$ for arbitrary stationary transmission strategy f (that does not use the value of the state E_t to determine when to transmit; so the receiver is the same as in Theorem 2.4.1) for the long-term average setup for Model B when $a = 1$.

Proposition 2.5.1 *For $\beta = 1$ and $a = 1$ in Model B, let f be an arbitrary stationary transmission strategy. Let τ denote the stopping time of the first transmission under f. Then*

$$D_1(f, g^*) = \frac{\sigma^2}{2}\Big[\frac{\mathbb{E}(\tau^2)}{\mathbb{E}(\tau)} - 1\Big].$$

Proof For any $t < \tau$, $E_t = W_0^2 + \cdots + W_{t-1}^2$. Therefore, $\mathbb{E}[E_t^2] = t\sigma^2$ and define $\hat{L}(t) = \sum_{s=1}^{t-1}\mathbb{E}[E_s^2] = \frac{1}{2}t(t-1)\sigma^2$. Now, $L_1(0) = \mathbb{E}[\hat{L}(\tau)] = (\sigma^2/2)[\mathbb{E}(\tau^2) - \mathbb{E}(\tau)]$ and $M_1(0) = \mathbb{E}(\tau)$. By using the same argument as in the proof of Theorem 2.4.2, we get $D_1(f, g^*) = L_1(0)/M_1(0)$, which implies the result.

Performance of randomized transmission strategy

For the randomized strategy defined above, τ is a $\text{Geom}_1(\alpha)$ random variable. Therefore, $\mathbb{E}(\tau^2) = 2/\alpha^2 - 1/\alpha$ and $\mathbb{E}(\tau) = 1/\alpha$. Hence, following Proposition 2.5.1, we have

$$D_{\text{rand}}(\alpha) = \sigma^2\big[\frac{1}{\alpha} - 1\big].$$

Fig. 2.7 shows that threshold-based strategy performs considerably well compared to the periodic transmission strategy and the randomized transmission strategy.

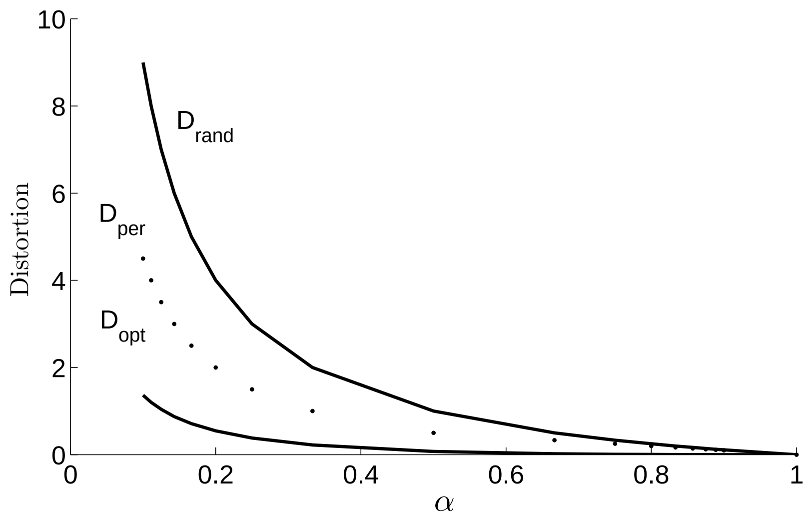


Fig. 2.7 Comparison of the performances of the threshold-based strategy (denoted by D_{opt}) with periodic and randomized transmission strategies (denoted by D_{per} and D_{rand}, respectively) for a Gauss-Markov process with $a = 1$ and $\sigma^2 = 1$.

2.5.2 Discussion on deterministic implementation

The optimal strategy shown in Theorem 2.4.4 chooses a randomized action in states $\{-k^*, k^*\}$. It is also possible to identify deterministic (non-randomized) but time-varying strategies that achieve the same performance. We describe two such strategies for the long-term average setup.

Steering strategies

Let a_t^0 (respectively, a_t^1) denote the number of times the action $u_t = 0$ (respectively, the action $u_t = 1$) has been chosen in states $\{-k^*, k^*\}$ in the past, i.e.

$$a_t^i = \sum_{s=0}^{t-1} \mathbb{1}\{|E_s| = k^*,\, u_s = i\}, \quad i \in \{0, 1\}.$$

Thus, the empirical frequency of choosing action $u_t = i$, $i \in \{0, 1\}$, in states $\{-k^*, k^*\}$ is $a_t^i/(a_t^0 + a_t^1)$. A steering strategy compares these empirical frequencies with the desired randomization probabilities $\theta^0 = 1 - \theta^*$ and $\theta^1 = \theta^*$ and chooses an action that *steers* the empirical frequency closer to the desired randomization probability. More formally, at states $\{-k^*, k^*\}$, the steering transmission strategy chooses the action

$$\arg\min_i \left\{ \theta^i - \frac{a_t^i + 1}{a_t^0 + a_t^1 + 1} \right\}$$

in states $\{-k^*, k^*\}$ and chooses deterministic actions according to f^* (given in (3.71)) in states except $\{-k^*, k^*\}$. Note that the above strategy is deterministic (non-randomized) but depends on the history of visits to states $\{-k^*, k^*\}$. Such strategies were proposed in [46], where it was shown that the steering strategy described above achieves the same performance as the randomized strategy f^* and hence is optimal for Problem 2.3.2 for $\beta = 1$. Variations of such steering strategies have been proposed in [47, 48], where the adaptation was done by comparing the sample path average cost with the expected value (rather than by comparing empirical frequencies).

Time-sharing strategies

Define a cycle to be the period of time between consecutive visits of process $\{E_t\}_{t=0}^{\infty}$ to state zero. A time-sharing strategy is defined by a series $\{(a_m, b_m)\}_{m=0}^{\infty}$ and uses strategy $f^{(k^*)}$ for the first a_0 cycles, uses strategy $f^{(k^*+1)}$ for the next b_0 cycles, and continues to alternate between using strategy $f^{(k^*)}$ for a_m cycles and strategy $f^{(k^*+1)}$ for b_m cycles. In particular, if $(a_m, b_m) = (a, b)$ for all m, then the time-sharing strategy is a periodic strategy that uses $f^{(k^*)}$ a cycles and $f^{(k^*+1)}$ for b cycles.

The performance of such time-sharing strategies was evaluated in [49], where it was shown that if the cycle-lengths of the time-sharing strategy are chosen such that,

$$\begin{aligned} \lim_{M\to\infty} \frac{\sum_{m=0}^{M} a_m}{\sum_{m=0}^{M}(a_m + b_m)} &= \frac{\theta^* N_1^{(k^*)}}{\theta^* N_1^{(k^*)} + (1-\theta^*) N_1^{(k^*+1)}} \\ &= \frac{\theta^* N_1^{(k^*)}}{\alpha}, \end{aligned}$$

then the time-sharing strategy $\{(a_m, b_m)\}_{m=0}^{\infty}$ achieves the same performance as the randomized strategy f^* and hence, is optimal for Problem 2.3.2 for $\beta = 1$.

2.6 Proof of the structural result: Theorem 2.4.1

2.6.1 Finite horizon setup

A finite horizon version of Problem 2.3.1 has been investigated in [10] (for Model A) and in [9, 11] (for Model B), where the structure of the optimal transmission and estimation strategy was established.

Theorem 2.6.1 *[9–11] For both Models A and B, for a finite horizon version of Problem 2.3.1, we have the following.*

1. Structure of optimal estimation strategy*: the estimation strategy defined in Theorem 2.4.1 is optimal.*
2. Structure of optimal transmission strategy*: define E_t as in Theorem 2.4.1. Then there exist thresholds $\{k_t\}_{t=1}^{T}$ such that the transmission strategy*

$$U_t := f_t(E_t) = \begin{cases} 1, & \text{if } |E_t| \geq k_t; \\ 0, & \text{if } |E_t| < k_t \end{cases} \tag{2.25}$$

   is optimal.

The above structural results were obtained in [10, Theorems 2 and 3] for Model A and in [9, Theorem 1] and [11, Lemmas 1, 3 and 4] of Model B.

Remark 6 The results in [10] were derived under the assumption that $\{W_t\}$ has finite support. These results can be generalized for $\{W_t\}$ having countable support using ideas from [50]. For that reason, we state Theorem 2.6.1 without any restriction on the support of $\{W_t\}$. See Appendix A.1 for the generalization of [10, Theorems 2 and 3] to $\{W_t\}$ with countable support.

2.6.2 Infinite horizon setup

In a general real-time communication system, the optimal estimation strategy depends on the choice of the transmission strategy and vice-versa. Theorem 2.6.1 shows that when the noise process and the distortion function satisfy appropriate symmetry assumptions, the optimal estimation strategy can be specified in closed form. Consequently, we can fix the estimation strategy to be of the above form and consider the optimization problem of identifying the best transmission strategy. This optimization problem has a single decision maker—the transmitter—and we use techniques from centralized stochastic control to solve it. Since the optimal estimation strategy is time-homogeneous, one expects the optimal transmission strategy (i.e., the choice of the optimal thresholds $\{k_t\}_{t=0}^{\infty}$) to be time-homogeneous as well. The technical difficulty in establishing such a result is that the state space is not compact and the distortion function may be unbounded.

To prove Theorem 2.4.1, we proceed as follows:

1. We show that the result of the theorem is true for $\beta \in (0,1)$ and the optimal strategy is given by an appropriate dynamic program.

2. We show that for the discounted setup, the value function of the dynamic program is even and increasing on $\mathbb{X}$.

3. For $\beta = 1$, we use the vanishing discount approach to show that the optimal strategy for the long-term average cost setup may be determined as a limit to the optimal strategy for the discounted cost setup is the discount factor $\beta \uparrow 1$.

The discounted setup

Lemma 2.6.1 *In Model A. an optimal transmission strategy is given by the unique and bounded solution of the following dynamic program: for all $e \in \mathbb{Z}$,*

$$
\begin{aligned}
V_\beta(e;\lambda) = \min\Big[&(1-\beta)\lambda + \beta \sum_{w\in\mathbb{Z}} p_w V_\beta(w;\lambda), \\
&(1-\beta)d(e) + \beta \sum_{w\in\mathbb{Z}} p_w V_\beta(ae+w;\lambda)\Big].
\end{aligned}
\tag{2.26}
$$

Proof When $d(\cdot)$ is bounded, the per-step cost $c(e,u) := (1-\beta)[\lambda u + d(e)(1-u)]$, $u \in \{0,1\}$, for a given λ is also bounded and hence according to [51, Proposition 4.7.1, Theorem 4.6.3], there exists the unique and bounded solution $V_\beta(e;\lambda)$ of the dynamic program (2.26).

When $d(\cdot)$ is unbounded, then for any communication cost λ, we first define $e_0 \in \mathbb{Z}_{\geq 0} <$

∞ as:

$$e_0 := \min\left\{e \,:\, d(e) \geq \frac{\lambda}{1-\beta}\right\}.$$

Now, for any state e, $|e| > e_0$, the per-step cost $(1-\beta)d(e)$ of not transmitting is greater then the cost of transmitting at each step in the future, which is given by $(1-\beta)\sum_{t=0}^{\infty}\beta^t\lambda = \lambda$. Thus, the optimal action is to transmit, i.e., $f^*(e) = 1$. Hence, the dynamic program can be written as

$$V_\beta(e;\lambda) = \min\{V_\beta^0(e;\lambda), V_\beta^1(e;\lambda)\},$$

where

$$\begin{aligned}
V_\beta^0(e;\lambda) &= (1-\beta)d(e) + \beta\sum_{w\in\mathbb{Z}} p_w V_\beta(ae+w;\lambda),\\
V_\beta^1(e;\lambda) &= (1-\beta)\lambda + \beta\sum_{w\in\mathbb{Z}} p_w V_\beta(w;\lambda).
\end{aligned}$$

Let $\mathcal{E}^* := \{e \,:\, |e| \geq e_0\}$. Then, for all $e \in \mathcal{E}^*$, $V_\beta(e;\lambda)$ is constant. Thus, (2.26) is equivalent to a finite-state Markov decision process with state space $\{-e_0+1, \cdots, e_0-1\}\cup e^*$ (where e^* is a generic state for all states in the set $\mathcal{E}^*$). Since the state space is now finite, the dynamic program (2.26) has a unique and bounded time-homogeneous solution by the argument given for bounded $d(\cdot)$.

Lemma 2.6.2 *In Model B, an optimal transmission strategy is given by the unique and bounded solution of the following dynamic program: for all $e \in \mathbb{R}$,*

$$\begin{aligned}
V_\beta(e;\lambda) = \min\Big[&(1-\beta)\lambda + \beta\int_{\mathbb{R}}\phi(w)V_\beta(w;\lambda)dw,\\
&(1-\beta)d(e) + \beta\int_{\mathbb{R}}\phi(w)V_\beta(ae+w;\lambda)dw\Big]. \qquad (2.27)
\end{aligned}$$

Proof When $d(\cdot)$ is bounded, the per-step cost $c(e,u)$, as defined in part (a), for a given λ is also bounded. Let $K = (1-\beta)\sup_{e\in\mathbb{R}}\{d(e)\}$. Then, the strategy 'always transmit' satisfies [52, Assumption 4.2.2] with $V_\beta(e;\lambda) \leq K/(1-\beta)$. Also, λ, $d(\cdot)$ and $\phi(\cdot)$ satisfy [52, Assumption 4.2.1]. Hence, the above dynamic program has a unique and bounded solution due to [52, Theorem 4.2.3].

When $d(\cdot)$ is unbounded, define e_0 and e^* as in the proof of Lemma 2.6.1. By an argument similar to that in the proof of Lemma 2.6.1, we can restrict the state space of (2.27) to $[-e_0, e_0] \cup e^*$. Hence, the state space is compact and on this state space $d(\cdot)$ is bounded. Thus, the dynamic program (2.27) has a unique and bounded solution by the argument given for bounded $d(\cdot)$.

Proof (Proof of Theorem 2.4.1 for $\boldsymbol{\beta \in (0,1)}$) The structure of the optimal strategies follows from Theorem 2.6.1. The optimal thresholds are time invariant because the corresponding dynamic programs (2.26) and (2.27) have a unique fixed point.

Properties of the value function

Proposition 2.6.1 *For any $a \in \mathbb{X}_{>0}$, consider the two Markov processes $\{X_t^{(+)}\}_{t=0}^{\infty}$ and $\{X_t^{(-)}\}_{t=0}^{\infty}$ such that $X_0^{(+)} = X_0^{(-)} = 0$ and*

$$X_{t+1}^{(+)} = aX_t^{(+)} + W_t \quad \textit{and} \quad X_{t+1}^{(-)} = -aX_t^{(-)} + W_t.$$

Let $V_\beta^{(+)}$ and $V_\beta^{(-)}$ be the value functions corresponding to $\{X_t^{(+)}\}_{t=0}^{\infty}$ and $\{X_t^{(-)}\}_{t=0}^{\infty}$. Then

$$V_\beta^{(+)}(e) = V_\beta^{(-)}(e), \quad \forall e.$$

Therefore, if k is an optimal threshold for $\{X_t^{(+)}\}_{t=0}^{\infty}$ then k is also optimal for $\{X_t^{(-)}\}_{t=0}^{\infty}$.

See Appendix A.4 for the proof.

Remark 7 As a consequence of the above proposition, we can restrict attention to $a > 0$ while proving the properties of the value function $V_\beta(\cdot)$.

Proposition 2.6.2 *For any $\lambda > 0$ and $\beta \in (0,1)$, the value functions $V_\beta(\cdot;\lambda)$ given by (2.26) and (2.27) are even and increasing on $\mathbb{X}_{\geq 0}$.*

See Appendix A.4 for the proof.

The long-term average setup

Proposition 2.6.3 *For any $\lambda \geq 0$, the value function $V_\beta(\cdot;\lambda)$ for Models A and B, as given by (2.26) and (2.27) respectively, satisfy the following SEN conditions of [51, 52]:*

(S1) There exists a reference state $e_0 \in \mathbb{X}$ and a non-negative scalar M_λ such that $V_\beta(e_0, \lambda) < M_\lambda$ for all $\beta \in (0,1)$.

(S2) Define $h_\beta(e;\lambda) = (1-\beta)^{-1}[V_\beta(e;\lambda) - V_\beta(e_0;\lambda)]$*. There exists a function* $K_\lambda : \mathbb{Z} \to \mathbb{R}$ *such that* $h_\beta(e;\lambda) \le K_\lambda(e)$ *for all* $e \in \mathbb{X}$ *and* $\beta \in (0,1)$.

(S3) There exists a non-negative (finite) constant L_λ *such that* $-L_\lambda \le h_\beta(e;\lambda)$ *for all* $e \in \mathbb{X}$ *and* $\beta \in (0,1)$.

Therefore, if f_β *denotes an optimal strategy for* $\beta \in (0,1)$*, and* f_1 *is any limit point of* $\{f_\beta\}$ *for any increasing sequence of* β*, then* f_1 *is optimal for* $\beta = 1$.

Proof Let $V_\beta^{(0)}(e;\lambda)$ denote the value function of the 'always transmit' strategy. Since $V_\beta(0;\lambda) \le V_\beta^{(0)}(0;\lambda)$ and $V_\beta^{(0)}(0;\lambda) = \lambda$, (S1) is satisfied with $e_0 = 0$ and $M_\lambda = \lambda$.

We show (S2) for Model B, but a similar argument works for Model A as well. Since not transmitting is optimal at state 0, we have

$$V_\beta(0;\lambda) = \beta \int_{-\infty}^{\infty} \phi(w) V_\beta(w;\lambda) dw.$$

Let $V_\beta^{(1)}(e,\lambda)$ denote the value function of the strategy that transmits at time 0 and follows the optimal strategy from then on. Then

$$\begin{aligned} V_\beta^{(1)}(e;\lambda) &= (1-\beta)\lambda + \beta \int_{-\infty}^{\infty} \phi(w) V_\beta(w;\lambda) dw \\ &= (1-\beta)\lambda + \beta V_\beta(0;\lambda) \end{aligned} \tag{2.28}$$

Since $V_\beta(e;\lambda) \le V_\beta^{(1)}(e;\lambda)$ and $V_\beta(0;\lambda) \ge 0$, from (2.28) we get that $(1-\beta)^{-1}[V_\beta(e;\lambda) - V_\beta(0,\lambda)] \le \lambda$. Hence (S2) is satisfied with $K_\lambda(e) = \lambda$.

By Proposition 2.6.2, $V_\beta(e;\lambda) \ge V_\beta(0;\lambda)$, hence (S3) is satisfied with $L_\lambda = 0$.

Proof (Proof of Theorem 2.4.1 for $\beta = 1$) Since the value function $V_\beta(\cdot;\lambda)$ satisfies the SEN conditions for reference state $e_0 = 0$, the optimality of the threshold strategy for long-term average setup follows from [51, Theorem 7.2.3] for Model A and [52, Theorem 5.4.3] for Model B, respectively.

2.7 Proof of Theorem 2.4.2

2.7.1 Preliminary results

Define operator $\mathcal{B}$ as follows:

- **Model A**: For any $v : \mathbb{Z} \to \mathbb{R}$, define operator $\mathcal{B}$ as

$$[\mathcal{B}v](e) \coloneqq \sum_{w=-\infty}^{\infty} p_w v(ae+w), \quad \forall e \in \mathbb{Z}.$$

  Or, equivalently,

$$[\mathcal{B}v](e) \coloneqq \sum_{n=-\infty}^{\infty} p_{n-ae} v(n), \quad \forall e \in \mathbb{Z}.$$

- **Model B**: For any bounded $v : \mathbb{R} \to \mathbb{R}$, define operator $\mathcal{B}$ as

$$[\mathcal{B}v](e) \coloneqq \int_{\mathbb{R}} \phi(w) v(ae+w) dw, \quad \forall e \in \mathbb{R}.$$

  Or, equivalently,

$$[\mathcal{B}v](e) \coloneqq \int_{\mathbb{R}} \phi(n-ae) v(n) dn, \quad \forall e \in \mathbb{R}.$$

As discussed in Remark 4, the error process $\{E_t\}_{t=0}^{\infty}$ is a controlled Markov process. Therefore, the functions $D_\beta^{(k)}$ and $N_\beta^{(k)}$ may be thought as value functions when strategy $f^{(k)}$ is used. Thus, they satisfy the following fixed point equations: for $\beta \in (0,1)$,

$$D_\beta^{(k)}(e) = \begin{cases} \beta[\mathcal{B}D_\beta^{(k)}](0), & \text{if } |e| \geq k \\ (1-\beta)d(e) + \beta[\mathcal{B}D_\beta^{(k)}](e), & \text{if } |e| < k, \end{cases} \tag{2.29}$$

$$N_\beta^{(k)}(e) = \begin{cases} (1-\beta) + \beta[\mathcal{B}N_\beta^{(k)}](0), & \text{if } |e| \geq k \\ \beta[\mathcal{B}N_\beta^{(k)}](e), & \text{if } |e| < k. \end{cases} \tag{2.30}$$

Lemma 2.7.1 *For $\beta \in (0,1]$,* (B.1) *and* (B.2) *have unique and bounded solutions $D_\beta^{(k)}(e)$ and $N_\beta^{(k)}(e)$ that*

1. *are even and increasing (on $\mathbb{X}_{\geq 0}$) in e for all k,*

2. *satisfy the SEN conditions (see Proposition 2.6.3) and therefore*

$$D_1^{(k)}(e) = \lim_{\beta \uparrow 1} D_\beta^{(k)}(e) \quad \text{and} \quad N_1^{(k)}(e) = \lim_{\beta \uparrow 1} N_\beta^{(k)}(e).$$

3. *$D_\beta^{(k)}(e)$ is increasing in k for all e and $N_\beta^{(k)}(e)$ is strictly decreasing in k for all e.*

The proofs of 1) and 2) follow from the arguments similar to those of Section 2.6 and are therefore omitted. The proof of 3) is given in Appendix A.5.

2.7.2 Proof of Theorem 2.4.2

We prove the result for the discounted cost setup, $\beta \in (0, 1)$. The result extends to the long-term average cost setup, $\beta = 1$, by using the vanishing discount approach similar to the argument given in Section 2.6.

We first consider the case $k = 0$. In this case, the recursive definition of $D_\beta^{(k)}$ and $N_\beta^{(k)}$, given by (B.1) and (B.2), simplify to the following:

$$D_\beta^{(0)}(e) = \beta[\mathcal{B} D_\beta^{(0)}](0);$$

and

$$N_\beta^{(0)}(e) = (1 - \beta) + \beta[\mathcal{B} N_\beta^{(0)}](0).$$

It can be easily verified that $D_\beta^{(0)}(e) = 0$ and $N_\beta^{(0)}(e) = 1$, $e \in \mathbb{X}$, satisfy the above equations. Also, $C_\beta^{(0)}(e; \lambda) = C_\beta(f^{(0)}, g^*; \lambda) = \lambda$. This proves the first part of the proposition.

For $k > 0$, let $\tau^{(k)}$ denote the stopping time when the Markov process in both Model A and B starting at state 0 at time $t = 0$ leaves the set $S^{(k)}$. Note that $\tau^{(0)} = 1$ and $\tau^{(\infty)} = \infty$.

Then,

$$L_\beta^{(k)}(0) = \mathbb{E}\Big[\sum_{t=0}^{\tau^{(k)}-1} \beta^t d(E_t) \;\Big|\; E_0 = 0 \Big] \tag{2.31}$$

$$M_\beta^{(k)}(0) = \mathbb{E}\Big[\sum_{t=0}^{\tau^{(k)}-1} \beta^t \;\Big|\; E_0 = 0 \Big] = \frac{1 - \mathbb{E}[\beta^{\tau^{(k)}} \mid E_0 = 0]}{1 - \beta} \tag{2.32}$$

$$D_\beta^{(k)}(0) = \mathbb{E}\Big[(1 - \beta) \sum_{t=0}^{\tau^{(k)}-1} \beta^t d(E_t) + \beta^{\tau^{(k)}} D_\beta^{(k)}(0) \;\Big|\; E_0 = 0 \Big] \tag{2.33}$$

$$N_\beta^{(k)}(0) = \mathbb{E}\Big[\beta^{\tau^{(k)}} \big((1 - \beta) + N_\beta^{(k)}(0) \big) \;\Big|\; E_0 = 0 \Big]. \tag{2.34}$$

Substituting (B.3) and (B.4) in (B.6) we get

$$D_\beta^{(k)}(0) = (1-\beta)L_\beta^{(k)}(0) + [1-(1-\beta)M_\beta^{(k)}(0)]D_\beta^{(k)}(0).$$

Rearranging, we get that

$$D_\beta^{(k)}(0) = \frac{L_\beta^{(k)}(0)}{M_\beta^{(k)}(0)}.$$

Similarly, substituting (B.3) and (B.4) in (B.8) we get

$$N_\beta^{(k)}(0) = [1-(1-\beta)M_\beta^{(k)}(0)][(1-\beta) + N_\beta^{(k)}(0)].$$

Rearranging, we get that

$$N_\beta^{(k)}(0) = \frac{1}{M_\beta^{(k)}(0)} - (1-\beta).$$

The expression for $C_\beta^{(k)}(0;\lambda)$ follows from the definition.

2.8 Proofs of results for Model A

2.8.1 Proof of Theorem 2.4.3

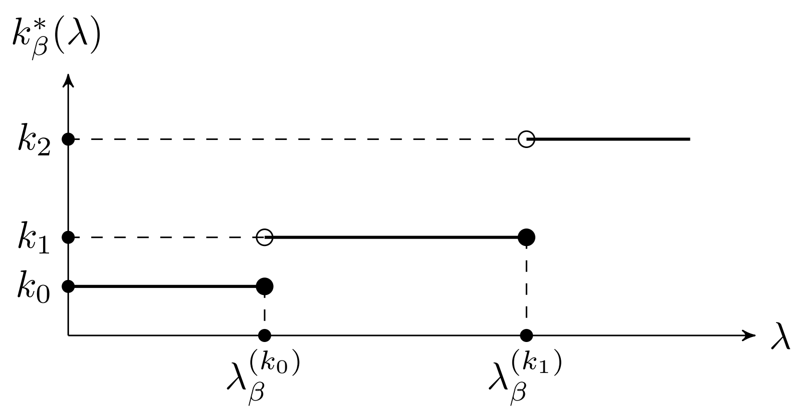


Fig. 2.8 Plot of $k_\beta^*(\lambda)$ for Model A.

By Proposition B.4.1, $k_\beta^*(\lambda) = \arg\inf_{k\geq 0} C_\beta^{(k)}(0;\lambda)$ is increasing in λ. Let $\mathbb{K}$ denote the set of all possible values of $k_\beta^*(\lambda)$. Since k is integer-valued, the plot of k_β^* vs λ must be a staircase function as shown in Fig. 2.8. In particular, there exists an increasing sequence $\{\lambda_\beta^{(k_n)}\}_{k_n \in \mathbb{K}}$ such that for $\lambda \in (\lambda_\beta^{(k_{n-1})}, \lambda_\beta^{(k_n)}]$, $k^*_\beta(\lambda) = k_n$. We will show that for any k_n,

$$C_\beta^{(k_n)}(0; \lambda_\beta^{(k_n)}) = C_\beta^{(k_{n+1})}(0; \lambda_\beta^{(k_n)}). \tag{2.35}$$

Simplifying (B.11), we get that $\lambda_\beta^{(k_n)}$ is given by (3.68).

Proof of (B.11)

For any $\lambda \in (\lambda_\beta^{(k_{n-1})}, \lambda_\beta^{(k_n)}]$, $C_\beta^{(k_n)}(0; \lambda) \leq C_\beta^{(k_{n+1})}(0; \lambda)$. In particular, for $\lambda = \lambda_\beta^{(k_n)}$,

$$C_\beta^{(k_n)}(0; \lambda_\beta^{(k_n)}) \leq C_\beta^{(k_{n+1})}(0; \lambda_\beta^{(k_n)}). \tag{2.36}$$

Similarly, for any $\lambda \in (\lambda_\beta^{(k_n)}, \lambda_\beta^{(k_{n+1})}]$, $C_\beta^{(k_{n+1})}(0; \lambda) \leq C_\beta^{(k_n)}(0; \lambda)$. Since both terms are continuous in λ, taking limit as $\lambda \downarrow \lambda_\beta^{(k_n)}$, we get

$$C_\beta^{(k_{n+1})}(0; \lambda_\beta^{(k_n)}) \leq C_\beta^{(k_n)}(0; \lambda_\beta^{(k_n)}). \tag{2.37}$$

Eq. (B.11) follows from combining (B.12) and (2.37).

Proof of Part 1)

By definition of $\lambda_\beta^{(k_n)}$, the strategy $f^{(k_n)}$ is optimal for $\lambda \in (\lambda_\beta^{(k_{n-1})}, \lambda_\beta^{(k_n)}]$.

Proof of Part 2)

Recall $C^*_\beta(\lambda) = \inf_{k \geq 0} C_\beta^{(k)}(0; \lambda)$. By definition, for $\lambda \geq 0$, $C_\beta^{(k)}(0; \lambda)$, is increasing and affine in λ. Therefore, its point-wise minimum (over k) is increasing and concave in λ.

As shown in part 1), for $\lambda \in (\lambda_\beta^{(k_n)}, \lambda_\beta^{(k_{n+1})}]$, $C^*_\beta(\lambda) = C_\beta^{(k_{n+1})}(0; \lambda)$, which is linear (and continuous) in λ; hence, $C^*_\beta(\lambda)$ is piecewise linear. Finally, by (B.11), $C_\beta^{(k_n)}(0; \lambda^{(k_n)}) = C_\beta^{(k_{n+1})}(0; \lambda^{(k_n)})$. Therefore, at the corner points, $\lim_{\lambda \uparrow \lambda_\beta^{(k_{n+1})}} C^*_\beta(\lambda) = \lim_{\lambda \downarrow \lambda_\beta^{(k_{n+1})}} C^*_\beta(\lambda)$. Hence, $C^*_\beta(\lambda)$ is continuous in λ.

2.8.2 Proof of Theorem 2.4.4

Note that by definition, $\theta^* \in [0,1]$ and

$$\theta^* N_\beta(f^{(k^*)}, g^*) + (1-\theta^*) N_\beta(f^{(k^*+1)}, g^*) = \alpha. \tag{2.38}$$

Proof of Part 1)

The optimality of (f^*, g^*) relies on the following characterization of the optimal strategy stated in [53, Proposition 1.2]. The characterization was stated for the long-term average setup but a similar result can be shown for the discounted case as well, for example, by using the approach of [54]. Also, see [55, Theorem 8.4.1] for a similar sufficient condition for general constrained optimization problem.

A (possibly randomized) strategy (f°, g°) is optimal for a constrained optimization problem with $\beta \in (0,1]$ if the following conditions hold:

(C1) $N_\beta(f^\circ, g^\circ) = \alpha$,

(C2) There exists a $\lambda^\circ \geq 0$ such that (f°, g°) is optimal for $C_\beta(f, g; \lambda^\circ)$.

We will show that the strategies (f^*, g^*) satisfy (C1) and (C2) with $\lambda^\circ = \lambda_\beta^{(k^*)}$.

(f^*, g^*) satisfy (C1) due to (2.38). For $\lambda = \lambda_\beta^{(k^*)}$, both $f^{(k^*)}$ and $f^{(k^*+1)}$ are optimal for $C_\beta(f, g; \lambda)$. Hence, any strategy randomizing between them, in particular f^*, is also optimal for $C_\beta(f, g; \lambda)$. Hence (f^*, g^*) satisfies (C2). Therefore, by [53, Proposition 1.2], (f^*, g^*) is optimal for Problem 2.3.2.

Proof of Part 2)

The expression of k^* and θ^* follow directly from (3.69) and (3.70). The form of $D_\beta^*(\alpha)$ given in (3.72) follows immediately from the fact that (f^*, g^*) is a Bernoulli randomized simple strategy.

$D_\beta^*(\alpha)$ is the solution to a constrained optimization problem with the constraint set $\{(f, g) : N_\beta(f, g) \leq \alpha\}$. Therefore, it is decreasing and convex in the constraint α. The optimality of (f^*, g^*) implies (3.72). Piecewise linearity of $D_\beta^*(\alpha)$ follows from (3.72). Finally, by definition of $\alpha^{(k)}$ and θ, $\lim_{a \uparrow \alpha^{(k)}} D_\beta^*(\alpha) = D_\beta^{(k)}(0) = \lim_{a \downarrow \alpha^{(k)}} D_\beta^*(\alpha)$. Hence, $D_\beta^*(\alpha)$ is continuous in α.

2.9 Proofs of results for Model B

Lemma 2.9.1 *In Model B, for $\beta \in (0,1]$,*

1. *$D_\beta^{(k)}$ and $N_\beta^{(k)}$ are continuous in k,*
2. *$N_\beta^{(k)}$ is* strictly *decreasing in k,*
3. *$D_\beta^{(k)}$, $N_\beta^{(k)}$ and $C_\beta^{(k)}$ are differentiable in k.*

Proof The proof follows from Lemma 2.4.1 and Theorem 2.4.2.

2.9.1 Proof of Theorem 3.9.3

Proof of Part 1)

The choice of λ implies that $\partial_k C_\beta^{(k)}(0;\lambda) = 0$. Hence strategy $(f^{(k)}, g^*)$ is optimal for the given λ.

Note that, (2.22) can also be written as $\lambda = \big((M_\beta^{(k)}(0))^2 \partial_k D_\beta^{(k)}(0)\big)/\partial_k M_\beta^{(k)}(0)$. By Lemma 2.4.1, $\partial_k M_\beta^{(k)}(0) > 0$ and by Lemma 2.7.1, $\partial_k D_\beta^{(k)}(0) \geq 0$. Hence, for any $k > 0$, λ given by (2.22) is positive. This completes the first part of the proof.

Proof of Part 2)

The monotonicity and concavity of $C_\beta^*(\lambda)$ follows from the same argument as in Model A.

Note that $k_\beta^*(\lambda) = \arg\inf_{k\geq 0} C_\beta^{(k)}(0;\lambda)$ can take a value ∞ (which corresponds to the strategy 'never communicate'). Thus, the domain of k is $\mathbb{X}_{\geq 0} \cup \{\infty\}$, which is a compact set. Now, $C_\beta^*(\lambda) = \min_{k\in[0,\infty]} C_\beta^{(k)}(0;\lambda)$, where $C_\beta^{(k)}(0;\lambda)$ is continuous in both λ and k. Since, $C_\beta^*(\lambda)$ is point-wise minimum of bounded continuous functions, where the minimization is over a compact set, it is continuous.

2.9.2 Proof of Theorem 3.9.4

Proof of Part 1)

Recall conditions (C1), (C2), given in Section 2.8.2, for a strategy to be optimal for a constrained optimization problem. We will show that for a given α, there exists a $k_\beta^*(\alpha) \in \mathbb{R}_{\geq 0}$ such that $(f^{(k_\beta^*(\alpha))}, g^*)$ satisfy conditions (C1) and (C2).

By Lemma 2.9.1, $N_\beta^{(k)}(0)$ is continuous and strictly decreasing in k. It is easy to see that $\lim_{k\to 0} N_\beta^{(k)}(0) = 1$ and $\lim_{k\to\infty} N_\beta^{(k)}(0) = 0$. Hence, for a given $\alpha \in (0,1)$, there exists a $k^*_\beta(\alpha)$ such that $N_\beta^{(k^*_\beta(\alpha))}(0) = N_\beta(f^{(k^*_\beta(\alpha))}, g^*) = \alpha$. Thus, $(f^{(k^*_\beta(\alpha))}, g^*)$ satisfies (C1).

Now, for $k^*_\beta(\alpha)$, we can find a λ satisfying (2.22) and hence we have by Theorem 3.9.3 that strategy $(f^{(k^*_\beta(\alpha))}, g^*)$ is optimal for $C_\beta(f, g; \lambda)$, and therefore satisfies (C2); and is consequently optimal for Problem 2.3.2.

Proof of Part 2)

By Lemma 2.9.1, $\tilde{N}(k) \coloneqq N_\beta^{(k)}(0)$ is strictly decreasing and continuous in k. Therefore, $\tilde{N}^{-1}$ exists and is continuous. Now,

$$D^*_\beta(\alpha) = \min_{\{k : k \le \tilde{N}^{-1}(\alpha)\}} D_\beta^{(k)}(0),$$

where, by Lemma 2.9.1, $D_\beta^{(k)}(0)$ is continuous in k. Thus, by Berge's maximum theorem, $D^*_\beta(\alpha)$ is continuous in α.

2.9.3 Proof of Theorem 2.4.7

To prove the theorem, we first need to prove the following lemma.

Lemma 2.9.2 *For Gauss-Markov model (a special case of Model B), let $L_\sigma^{(k)}$ and $M_\sigma^{(k)}$ be the solutions of* (2.11) *and* (2.12) *respectively, when the variance of W_t is σ^2. Then*

$$L_\sigma^{(k)}(e) = \sigma^2 L_1^{(k/\sigma)}\Big(\frac{e}{\sigma}\Big), \quad M_\sigma^{(k)}(e) = M_1^{(k/\sigma)}\Big(\frac{e}{\sigma}\Big), \tag{2.39}$$

$$D_\sigma^{(k)}(e) = \sigma^2 D_1^{(k/\sigma)}\Big(\frac{e}{\sigma}\Big), \quad N_\sigma^{(k)}(e) = N_1^{(k/\sigma)}\Big(\frac{e}{\sigma}\Big). \tag{2.40}$$

Proof Define $\hat{L}_\sigma^{(k)}(e) \coloneqq \sigma^2 L_1^{(k/\sigma)}\big(\frac{e}{\sigma}\big)$. Now consider,

$$\begin{aligned}
[\mathcal{B}_\sigma^{(k)} \hat{L}_\sigma^{(k)}](e) &= \int_{-k}^{k} \phi(n - ae)\hat{L}_\sigma^{(k)}(n)dn, \quad \forall e \in \mathbb{R} \\
&\overset{(a)}{=} \sigma^2 \int_{-k/\sigma}^{k/\sigma} \phi(z - ae/\sigma) L_1^{(k/\sigma)}(z)dz \\
&= \sigma^2 [\mathcal{B}_1^{(k/\sigma)} L_1^{(k/\sigma)}](e/\sigma),
\end{aligned}$$

where (a) uses a change of variables $n = \sigma z$. Therefore,

$$\begin{aligned}\left[\hat{L}_\sigma^{(k)} - \beta \mathcal{B}_\sigma^{(k)} \hat{L}_\sigma^{(k)}\right](e) &= \sigma^2 \left[L_1^{(k/\sigma)} - \beta \mathcal{B}_1^{(k/\sigma)} L_1^{(k/\sigma)}\right]\left(\frac{e}{\sigma}\right) \\ &= \sigma^2 \frac{e^2}{\sigma^2} = e^2.\end{aligned}$$

But, by Lemma 2.4.1, the above equation has a unique solution $L_\sigma^{(k)}$. Therefore $L_\sigma^{(k)} = \hat{L}_\sigma^{(k)}$.

A similar argument may be used to prove the scaling of $M_\sigma^{(k)}$. The scaling of $D_\sigma^{(k)}$ and $N_\sigma^{(k)}$ follow from Theorem 2.4.2.

Proof of Theorem 2.4.7

The theorem follows from Lemma 2.9.2, Theorem 2.4.2 and elementary algebra.

2.10 Proofs of results for Example 2.4.1

Lemma 2.10.1 *Define for $\beta \in (0, 1]$*

$$K_\beta = -2 - \frac{(1-\beta)}{\beta p} \quad \textit{and} \quad m_\beta = \cosh^{-1}(-K_\beta/2)$$

Then,

$$[Q_\beta^{(k)}]_{ij} = \frac{1}{\beta p} \frac{[A_\beta^{(k)}]_{ij}}{b_\beta^{(k)}}, \quad i, j \in S^{(k)},$$

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

$$\begin{aligned}[A_\beta^{(k)}]_{ij} &= \cosh((2k - |i-j|)m_\beta) - \cosh((i+j)m_\beta), \\ b_\beta^{(k)} &= \sinh(m_\beta)\sinh(2km_\beta);\end{aligned}$$

and for $\beta = 1$,

$$\begin{aligned}[A_1^{(k)}]_{ij} &= (k - \max\{i,j\})(k + \min\{i,j\}), \\ b_1^{(k)} &= 2k.\end{aligned}$$

In particular, the elements $[Q_\beta^{(k)}]_{0j}$ are given as follows. For $\beta \in (0,1)$,

$$[Q_\beta^{(k)}]_{0j} = \frac{1}{\beta p}\frac{\cosh((2k-|j|)m_\beta) - \cosh(jm_\beta)}{2\sinh(m_\beta)\sinh(2km_\beta)}, \tag{2.41}$$

and for $\beta = 1$,

$$[Q_1^{(k)}]_{0j} = \frac{k-|j|}{2p}. \tag{2.42}$$

Proof The matrix $I_{2k-1} - \beta P^{(k)}$ is a symmetric tridiagonal matrix given by

$$I_{2k-1} - \beta P^{(k)} = -\beta p \begin{bmatrix} K_\beta & 1 & 0 & \cdots & \cdots & 0 \\ 1 & K_\beta & 1 & 0 & \cdots & 0 \\ 0 & 1 & K_\beta & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & K_\beta & 1 \\ 0 & 0 & \cdots & 0 & 1 & K_\beta \end{bmatrix}.$$

$Q_\beta^{(k)}$ is the inverse of the above matrix. The inverse of the tridiagonal matrix in the above form with $K_\beta \le -2$ are computed in closed form in [56]. The result of the lemma follows from these results.

2.10.1 Proof of Lemma 2.4.2

By substituting the expression for $Q_\beta^{(k)}$ from Lemma 2.10.1 in the expressions for $L_\beta^{(k)}$ and $M_\beta^{(k)}$ from Proposition 2.4.2, we get that

1. For $\beta \in (0,1)$,

$$L_\beta^{(k)}(0) = \frac{\sinh(km_\beta) - k\sinh(m_\beta)}{4\beta p \sinh^2(m_\beta/2)\sinh(m_\beta)\cosh(km_\beta)},$$
$$M_\beta^{(k)}(0) = \frac{\sinh^2(km_\beta/2)}{2\beta p \sinh^2(m_\beta/2)\cosh(km_\beta)}.$$

2. For $\beta = 1$,

$$L_1^{(k)}(0) = k(k^2-1)/(6p), \quad M_1^{(k)}(0) = k^2/(2p).$$

The results of the lemma follow using the above expressions and Theorem 2.4.2. The expression for $\lambda_1^{(k)}$ is obtained by plugging the expressions of $D_1^{(k+1)}$, $D_1^{(k)}$, $N_1^{(k+1)}$, and $N_1^{(k)}$ in (3.68).

2.11 Conclusion

We characterize two fundamental limits of remote estimation of autoregressive Markov processes under communication constraints. First, when each transmission is costly, we characterize the minimum achievable cost of communication plus estimation error. Second, when there is a constraint on the average number of transmissions, we characterize the minimum achievable estimation error.

We also identify transmission and estimation strategies that achieve these fundamental limits. The structure of these optimal strategies had been previously identified by using dynamic programming for decentralized stochastic control systems. In particular, the optimal transmission strategy is to transmit when the estimation error process exceeds a threshold and the optimal estimation strategy is to select the transmitted state as the estimate, whenever there is a transmission. We use ideas based on renewal theory to identify the performance of a generic strategy that has such a structure. For the case of costly communication, we identify the value of communication cost for which a particular threshold-based strategy is optimal; for the case of constrained communication, we identify (possibly randomized) threshold-based strategies that achieve the communication constraint.

These results are derived under idealized assumptions on the communication channel: communication is noiseless and without any constraint on the transmission rate or the transmission bandwidth. Under these assumptions, the error process resets after each transmission (see Remark 4). This reset property is critical to derive the structure of optimal transmission and estimation strategies (Theorems 2.4.1 and 2.6.1). In the absence of such a structural result, the solution methodology developed in this chapter does not work and the optimal transmission and estimation strategies have to be identified by numerically solving the (decentralized) dynamic programs described in [6, 8].

Having said that, the transmission and estimation strategies described in Theorems 2.4.1 and 2.6.1 may be used as heuristic sub-optimal strategies when the communication channel does not satisfy the idealized assumptions described above. In that case, it may be possible to use the solution methodology developed in this chapter to obtain performance bounds on such strategies.

A similar remark holds for multi-dimensional autoregressive processes. It is reasonable to expect (although we are not aware of a proof of this statement) that for multi-dimensional autoregressive processes, the optimal estimation strategy will be similar to that described in Theorems 2.4.1 and 2.6.1 while the optimal transmission strategy will be to transmit when the error process lies outside a (multi-dimensional) ellipsoid. The performance of such strategies can be evaluated using the solution methodology developed in this chapter. The renewal relationships derived in Theorem 2.4.2 also hold for multi-dimensional autoregressive processes. The only difference is that $L_\beta^{(k)}(0)$ and $M_\beta^{(k)}(0)$ are computed by solving multi-dimensional Fredholm integral equations of the second kind. The optimal transmission strategies can then be computed by solving multi-dimensional versions of (2.22) (for costly communication) and (3.74) (for constrained communication). However, it is not immediately clear whether these equations will have a unique solution. Further investigation is required to obtain algorithms that identify the optimal transmission ellipsoid.

Finally, the solution methodology developed in this chapter to identify optimal thresholds is also of independent interest. In various applications of Markov decision processes threshold strategies are optimal. The approach developed in this chapter is directly applicable to such models.

Chapter 3

Remote estimation with packet drops

3.1 Introduction

3.1.1 Motivation and literature overview

In this chapter we consider an RE system in which a sensor/transmitter observes a first-order Markov process and causally decides which observations to transmit to a remotely located receiver/estimator. As in the previous chapter, here too we consider no source-coding, i.e., when the transmitter transmits, it transmits the entire source symbol and the size of the data-packet does not matter. Communication is expensive and takes place over a lossy channel. The channel has two states: OFF state and ON state. When the channel is in the OFF state, a packet transmitted from the sensor to the receiver is dropped. When the channel is in the ON state, a packet transmitted from the sensor to the receiver is received without error. The reception of the source symbol is acknowledged to the transmitter by a noiseless ACK/NACK feedback. The block diagram of the communication system is given in Fig. 3.1.

In this chapter we investigate the optimization problem of two erasure channels separately in two parts. In the first part we discuss the structural results for a Gilbert-Elliott channel, i.e., a channel with Markovian packet drops (also called a burst erasure channel). In the second part, we present an i.i.d. packet drop channel as a special case of the burst erasure channel. We analyze the optimization results and completely characterize the optimal thresholds and the optimal performances of the costly and constrained communication problems. In both cases, we consider a model where the communication takes place with acknowledgment; so either the transmitted packet is delivered without any error to the receiver or the packet is dropped. We assume that the channel state is causally observed at the receiver and is fed back to the transmitter with one-unit delay. Whenever there is a successful reception, the receiver sends an acknowledgment to the transmitter. The feedback is assumed to be noiseless.

At the time instances when the receiver does not receive a packet (either because the sensor did not transmit or because the transmitted packet was dropped), it needs to estimate the state of the source process. There is a fundamental trade-off between communication cost and estimation accuracy. Transmitting all the time minimizes the estimation error but incurs a high communication cost; not transmitting at all minimizes the communication cost but incurs a high estimation error.

As pointed out in Chapter 1, the existing literature on remote-estimation with a model similar to ours considers either channels with no or i.i.d. packet drops. In the cases where the threshold-based strategies are optimal, one would expect the structure of the threshold to be dependent on the channel model as well. In this chapter, in a fairly unified manner, we consider two cases where packet dropping channels are with i.i.d. or Markovian memory. We identify sufficient statistics at the transmitter and the receiver. When the source is a first-order autoregressive process, we show that threshold-based strategies (where the threshold depends on the previous channel-state) and the Kalman-like estimation strategies are optimal.

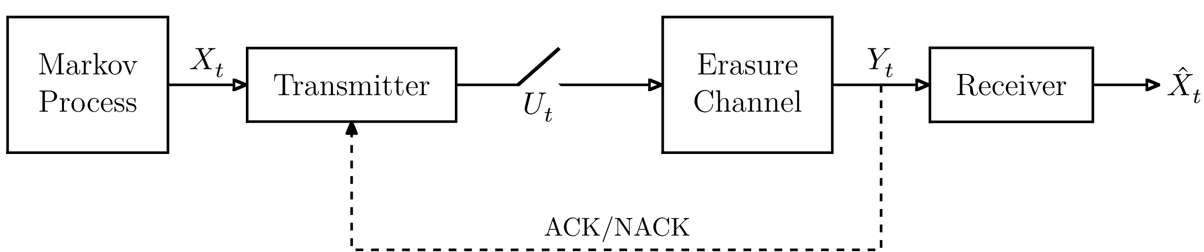


Fig. 3.1 Block diagram of a remote estimation setup with erasure channel.

3.2 Preliminary discussion on the proof approach

The remote estimation problem is essentially a decentralized control problem where there are two decision makers, a transmitter (sensor) and an estimator (receiver), who have access to different information. Based on the information available to them, they generate their strategy to minimize a common objective function. This set up fits into the realm of *Team theory*. We solve the team problem in two steps:

- The *person-by-person* approach - following [3,7], we arbitrarily fix the strategy of one agent (decision maker) and find the best strategy of the other. This approach helps to find out the information state of the transmitter for a fixed estimation strategy. Also, it is argued that the structure of the optimal estimator is independent of the transmission strategy.

- The *common-information approach* - Following [57], we split the information at the transmitter and the receiver into two parts: *common information* (which is the data known to all future decision makers) and *local information* (which is the total data minus the common information). Next we consider a centralized stochastic control problem, which we call the *coordinated system*, where a virtual decision maker observes the common information at each agent at time t and chooses a *prescription* (a function which based on the local information generates the action) according to a *coordination strategy* that is a function of the common information. The strategy of the agents are then generated based on that coordination strategy.

Since the coordinated system is centralized, an optimal coordinated strategy may be identified from an appropriate dynamic program. The detailed steps are discussed in the respective sections.

3.3 Remote estimation with Markovian packet drops

In the first part of this chapter, we consider the erasure channel to have Markovian packet drops. We first investigate the finite horizon problems and then extend the results to the infinite horizon. We establish the structure of the optimal communication strategies for a generic model. Then, for a stylized model with first order autoregressive sources, we characterize the optimal performances. The communication channel with i.i.d. packet drops is then expressed as a special case of the Markovian channel and we derive the computable expressions for the optimal performances.

3.3.1 The communication system

Fig. 3.1 shows the block diagram of the communication system. Each component is described in the following sections.

Source model

The source is a first-order time-homogeneous Markov process $\{X_t\}_{t\geq 0}$, $X_t \in \mathbb{X}$. For ease of exposition, in the first part of the chapter we assume that $\mathbb{X}$ is a finite set. We will later argue that a similar argument works when $\mathbb{X}$ is a general measurable space. The transition probability matrix of the source is denoted by P, i.e., for any $x, y \in \mathbb{X}$,

$$P_{xy} \coloneqq \mathbb{P}(X_{t+1} = y \mid X_t = x).$$

Channel model

The channel is a Gilbert-Elliott channel [58, 59]. The channel state $\{S_t\}_{t\geq 0}$ is a binary-valued first-order time-homogeneous Markov process. We use the convention that $S_t = 0$ denotes that the channel is in the OFF state and $S_t = 1$ denotes that the channel is in the ON state. The transition probability matrix of the channel state is denoted by Q, i.e., for $r, s \in \{0, 1\}$,

$$Q_{rs} \coloneqq \mathbb{P}(S_{t+1} = s | S_t = r).$$

The input alphabet $\bar{\mathbb{X}}$ of the channel is $\mathbb{X} \cup \{\mathfrak{E}\}$, where $\mathfrak{E}$ denotes the event that there is no transmission. The channel output alphabet $\mathbb{Y}$ is $\mathbb{X} \cup \{\mathfrak{E}_0, \mathfrak{E}_1\}$, where the symbols $\mathfrak{E}_0$ and $\mathfrak{E}_1$ are explained below. At time t, the channel input is denoted by $\bar{X}_t$ and the channel output is denoted by Y_t.

The channel is a channel with state. In particular, for any realization $(\bar{x}_{0:T}, s_{0:T}, y_{0:T})$ of $(\bar{X}_{0:T}, S_{0:T}, Y_{0:T})$, we have

$$\mathbb{P}(Y_t = y_t \mid \bar{X}_{0:t} = \bar{x}_{0:t}, S_{0:t} = s_{0:t}) = \mathbb{P}(Y_t = y_t \mid \bar{X}_t = \bar{x}_t, S_t = s_t) \tag{3.1}$$

and

$$\mathbb{P}(S_t = s_t \mid \bar{X}_{0:t} = \bar{x}_{0:t}, S_{0:t-1} = s_{0:t-1}) = \mathbb{P}(S_t = s_t \mid S_{t-1} = s_{t-1}) = Q_{s_{t-1}s_t} \tag{3.2}$$

Note that the channel output Y_t is a deterministic function of the input $\bar{X}_t$ and the state S_t. In particular, for any $\bar{x} \in \bar{\mathbb{X}}$ and $s \in \{0, 1\}$, the channel output y is given as follows:

$$y = \begin{cases} \bar{x}, & \text{if } \bar{x} \in \mathbb{X} \text{ and } s = 1 \\ \mathfrak{E}_1, & \text{if } \bar{x} = \mathfrak{E} \text{ and } s = 1 \\ \mathfrak{E}_0, & \text{if } s = 0 \end{cases}$$

This means that if there is a transmission (i.e., $\bar{x} \in \mathbb{X}$) and the channel is on (i.e., $s = 1$), then the receiver observes $\bar{x}$. However, if there is no transmission (i.e., $\bar{x} = \mathfrak{E}$) and the channel is on (i.e., $s = 1$), then the receiver observes $\mathfrak{E}_1$. If the channel is off, then the receiver observes $\mathfrak{E}_0$.

The transmitter

There is no need for channel coding in a RE setup. Instead, the role of the transmitter is to determine which source realizations need to be transmitted. Let $U_t \in \{0, 1\}$ denote the transmitter's decision. We use the convention that $U_t = 0$ denotes that there is no transmission (i.e., $\bar{X}_t = \mathfrak{E}$) and $U_1 = 1$ denotes that there is transmission (i.e., $\bar{X}_t = X_t$).

Transmission is costly. Each time the transmitter transmits (i.e., $U_t = 1$), it incurs a cost of λ.

The receiver

At time t, the receiver generates an estimate $\hat{X}_t \in \mathbb{X}$ of X_t. The quality of the estimate is determined by a distortion function $d \colon \mathbb{X} \times \mathbb{X} \to \mathbb{R}_{\geq 0}$.

3.3.2 Information structure and problem formulation for finite horizon

It is assumed that the receiver observes the channel state causally. Thus, the information available at the receiver[1] is

$$I_t^2 = \{S_{0:t}, Y_{0:t}\}.$$

[1]We use superscript 1 to denote variables at the transmitter and superscript 2 to denote variables at the receiver.

The estimate $\hat{X}_t$ is chosen according to

$$\hat{X}_t = g_t(I_t^2) = g_t(S_{0:t}, Y_{0:t}), \tag{3.3}$$

where g_t is called the *estimation rule* at time t. The collection $g \coloneqq (g_1, \ldots, g_T)$ for all time is called the *estimation strategy.*

It is assumed that there is one-step delayed feedback from the receiver to the transmitter.[2] Thus, the information available at the transmitter is

$$I_t^1 = \{X_{0:t}, U_{0:t-1}, S_{0:t-1}, Y_{0:t-1}\}.$$

The transmission decision U_t is chosen according to

$$U_t = f_t(I_t^1) = f_t(X_{0:t}, U_{0:t-1}, S_{0:t-1}, Y_{0:t-1}), \tag{3.4}$$

where f_t is called the *transmission rule* at time t. The collection $f \coloneqq (f_1, \ldots, f_T)$ for all time is called the *transmission strategy.*

The collection (f, g) is called a *communication strategy.* The performance of any communication strategy (f, g) over a finite horizon $T < \infty$ is given by

$$C_T(f, g; \lambda) = \mathbb{E}^{(f,g)}\left[\sum_{t=0}^{T} \lambda U_t + d(X_t, \hat{X}_t)\right] \tag{3.5}$$

where the expectation is taken with respect to the joint measure on all system variables induced by the choice of (f, g).

We are interested in the following optimization problem.

Problem 3.3.1 *In the model described above, identify a communication strategy* (f^*, g^*) *that minimizes the cost* $C_T(f, g; \lambda)$ *defined in* (3.5).

Remark 8 The initial sections of this chapter provide the detailed derivations for the structural results for the finite horizon setup (i.e., $T < \infty$). The results can be naturally extended under certain technical conditions [60]. In the infinite horizon setup, we consider two cases: *discounted* and *long-term average* and consider two optimization problems, *costly*

[2]Note that feedback requires two bits: the channel state S_t is binary and the channel output Y_t can be communicated by indicating whether $Y_t \in \mathbb{X}$ or not (i.e., transmitting an ACK or a NACK).

communication and *constrained communication* for both cases. The structural results are discussed in Section 3.7.

3.4 Main results for the finite horizon

3.4.1 Structure of optimal communication strategies

Two types of structural results are established in the real-time communication literature: (i) establishing that part of the data at the transmitter is irrelevant and can be dropped without any loss of optimality; (ii) establishing that the common information between the transmitter and the receiver can be "compressed" using a belief state. The first structural results were first established by Witsenhausen [3] while the second structural results were first established by Walrand-Varaiya [6].

We establish both types of structural results for RE. First, we show that $(X_{0:t-1}, U_{0:t-1})$ is irrelevant at the transmitter (Lemma 3.4.1); then, we use the common information approach of [57] and establish a belief-state for the common information $(S_{0:t}, Y_{0:t})$ between the transmitter and the receiver (Theorem 3.4.1).

Lemma 3.4.1 *For any estimation strategy of the form* (3.3)*, there is no loss of optimality in restricting attention to transmission strategies of the form*

$$U_t = f_t(X_t, S_{0:t-1}, Y_{0:t-1}). \tag{3.6}$$

The proof idea is similar to [7]. We show that $\{X_t, S_{0:t-1}, Y_{0:t-1}\}_{t\geq 0}$ is a controlled Markov process controlled by $\{U_t\}_{t\geq 0}$. See Section 3.5 for proof.

Now, following [57], for any transmission strategy f of the form (3.6) and any realization $(s_{0:T}, y_{0:T})$ of $(S_{0:T}, Y_{0:T})$, define $\varphi_t \colon \mathbb{X} \to \{0, 1\}$ as

$$\varphi_t(x) = f_t(x, s_{0:t-1}, y_{0:t-1}), \quad \forall x \in \mathbb{X}.$$

Furthermore, define conditional probability measures π_t^1 and π_t^2 on $\mathbb{X}$ as follows: for any

$x \in \mathbb{X}$,

$$\pi_t^1(x) \coloneqq \mathbb{P}^f(X_t = x \mid S_{0:t-1} = s_{0:t-1}, Y_{0:t-1} = y_{0:t-1}),$$
$$\pi_t^2(x) \coloneqq \mathbb{P}^f(X_t = x \mid S_{0:t} = s_{0:t}, Y_{0:t} = y_{0:t}).$$

We call π_t^1 the *pre-transmission belief* and π^2 the *post-transmission belief*. Note that when $(S_{0:T}, Y_{0:T})$ are random variables, then π_t^1 and π_t^2 are also random variables which we denote by Π_t^1 and Π_t^2.

For the ease of notation, for any $\varphi\colon \mathbb{X} \to \{0,1\}$ and $i \in \{0,1\}$, define the following:

- $B_i(\varphi) = \{x \in \mathbb{X} : \varphi(x) = i\}$.
- For any probability distribution π on $\mathbb{X}$ and any subset $\mathbb{A}$ of $\mathbb{X}$, $\pi(\mathbb{A})$ denotes $\sum_{x\in\mathbb{A}} \pi(x)$.
- For any probability distribution π on $\mathbb{X}$, $\xi = \pi|_\varphi$ means that $\xi(x) = \mathbb{1}_{\{\varphi(x)=0\}}\pi(x)/\pi(B_0(\varphi))$.

Lemma 3.4.2 *Given any transmission strategy f of the form* (3.6)*:*

1. *there exists a function F^1 such that*

$$\pi_{t+1}^1 = F^1(\pi_t^2) = \pi_t^2 P. \tag{3.7}$$

2. *there exists a function F^2 such that*

$$\pi_t^2 = F^2(\pi_t^1, \varphi_t, y_t) = \begin{cases} \delta_{y_t} & \text{if } y_t \in \mathbb{X} \\ \pi_t^1|_{\varphi_t}, & \text{if } y_t = \mathfrak{E}_1 \\ \pi_t^1, & \text{if } y_t = \mathfrak{E}_0. \end{cases} \tag{3.8}$$

Note that in (3.7), we are treating π_t^2 as a row-vector and in (3.8), δ_{y_t} denotes a Dirac measure centered at y_t. The update equations (3.7) and (3.8) are standard non-linear filtering equations.

See Section 3.5 for proof of Lemma 3.4.2.

Theorem 3.4.1 *In Problem 3.3.1, we have that:*

1. Structure of optimal strategies: *There is no loss of optimality in restricting attention to optimal transmission and estimation strategies of the form:*

$$U_t = f_t^*(X_t, S_{t-1}, \Pi_t^1), \tag{3.9}$$
$$\hat{X}_t = g_t^*(\Pi_t^2). \tag{3.10}$$

2. Dynamic program: *Let $\Delta(\mathbb{X})$ denote the space of probability distributions on $\mathbb{X}$. Define value functions $V_t^1 \colon \{0,1\} \times \Delta(\mathbb{X}) \to \mathbb{R}$ and $V_t^2 \colon \{0,1\} \times \Delta(\mathbb{X}) \to \mathbb{R}$ as follows.*

$$V_{T+1}^1(s, \pi^1) = 0, \tag{3.11}$$

and for $t \in \{T, \dots, 0\}$

$$\begin{aligned} V_t^1(s, \pi^1) = \min_{\varphi \colon \mathbb{X} \to \{0,1\}} \Big\{ &\lambda \pi^1(B_1(\varphi)) \\ &+ W_t^0(\pi^1, \varphi)\pi^1(B_0(\varphi)) + \sum_{x \in B_1(\varphi)} W_t^1(\pi^1, x)\pi^1(x) \Big\} \end{aligned} \tag{3.12}$$

$$V_t^2(s, \pi^2) = \min_{\hat{x} \in \mathbb{X}} \sum_{x \in \mathbb{X}} d(x, \hat{x})\pi^2(x) + V_{t+1}^1(s, \pi^2 P), \tag{3.13}$$

where,

$$\begin{aligned} W_t^0(\pi^1, \varphi) &= Q_{s0} V_t^2(0, \pi^1) + Q_{s1} V_t^2(1, \pi^1|_\varphi), \\ W_t^1(\pi^1, x) &= Q_{s0} V_t^2(0, \pi^1) + Q_{s1} V_t^2(1, \delta_x). \end{aligned}$$

Let $\Psi_t(s, \pi^1)$ denote the arg min of the right hand side of (3.12). *Then, the optimal transmission strategy of the form* (3.9) *is given by*

$$f_t^*(\cdot, s, \pi^1) = \Psi_t(s, \pi^1).$$

Furthermore, the optimal estimation strategy of the form (3.10) *is given by*

$$g_t^*(\pi^2) = \arg\min_{\hat{x} \in \mathbb{X}} \sum_{x \in \mathbb{X}} d(x, \hat{x})\pi^2(x). \tag{3.14}$$

The proof idea is as follows. Once we restrict attention to transmission strategies of the form (3.6), the information structure is partial history sharing [57]. Thus, one can use the common information approach of [57] and obtain the structure of optimal strategies.

See Section 3.5 for proof of Theorem 3.4.1.

Remark 9 The first term in (3.12) is the expected communication cost, the second term is the expected cost-to-go when the transmitter does not transmit, and the third term is the expected cost-to-go when the transmitter transmits. The first term in (3.13) is the expected distortion and the second term is the expected cost-to-go.

Remark 10 Although the above model and result are stated for sources with finite alphabets, they extend naturally to general state spaces (including Euclidean spaces) under standard technical assumptions. See [61] for details.

3.4.2 Optimality of threshold-based strategies for autoregressive source for finite horizon case

In this section, we consider a first-order autoregressive source $\{X_t\}_{t\geq 0}$, $X_t \in \mathbb{X}$, where the state space $\mathbb{X} \in \{\mathbb{R}, \mathbb{Z}\}$. We assume that the initial state $X_0 = 0$ and for $t \geq 0$, we have that

$$X_{t+1} = aX_t + W_t, \tag{3.15}$$

where $a \in \mathbb{X}$ and $W_t \in \mathbb{X}$ is distributed according to a symmetric and unimodal distribution. For $\mathbb{X} = \mathbb{R}$, let us denote the corresponding probability density function μ. For $\mathbb{X} = \mathbb{Z}$, μ is the corresponding probability mass function. Furthermore, the per-step distortion is given by $d(X_t - \hat{X}_t)$, where $d(\cdot)$ is a even function that is increasing on $\mathbb{X}_{\geq 0}$. The rest of the model is the same as before.

For the above model, we can further simplify the result of Theorem 3.4.1 for the finite horizon, as given by Theorem 3.4.2. See Section 3.6 for the proof.

Theorem 3.4.2 *For Problem 3.3.1 with the state dynamics* (3.15)*,*

1. Structure of optimal estimation strategy: *The optimal estimation strategy is given as follows:* $\hat{X}_0 = 0$*, and for* $t \geq 0$*,*

$$\hat{X}_t = \begin{cases} a\hat{X}_{t-1}, & \text{if } Y_t \in \{\mathfrak{E}_0, \mathfrak{E}_1\} \\ Y_t, & \text{if } Y_t \in \mathbb{X} \end{cases} \tag{3.16}$$

2. Structure of optimal transmission strategy: *There exist threshold functions* $k_t \colon \{0,1\} \to \mathbb{R}_{\geq 0}$ *such that the following transmission strategy is optimal:*

$$f_t(X_t, S_{t-1}, \Pi_t^1) = \begin{cases} 1, & \textit{if } |X_t - a\hat{X}_{t-1}| \geq k_t(S_{t-1}) \\ 0, & \textit{otherwise.} \end{cases} \tag{3.17}$$

Remark 11 As long as the receiver can distinguish between the events $\mathfrak{E}_0$ (i.e., $S_t = 0$) and $\mathfrak{E}_1$ (i.e., $U_t = 0$ and $S_t = 1$), the structure of the optimal estimator does not depend on the channel state information at the receiver.

Remark 12 It can be shown that under the optimal strategy, Π_t^2 is symmetric and unimodal (SU) (defined in Section 3.6.2) around $\hat{X}_t$ and, therefore, Π_t^1 is SU around $a\hat{X}_{t-1}$. Thus, the transmission and estimation strategies in Theorem 3.4.2 depend on the pre- and post-transmission beliefs only through their means.

Remark 13 Recall that the distortion function is even and increasing (in the states in $\mathbb{X}_{\geq 0}$). Therefore, the condition $|X_t - a\hat{X}_{t-1}| \geq k_t(S_{t-1})$ can be written as $d(X_t - a\hat{X}_{t-1}) \geq \tilde{k}_t(S_{t-1}) := d(k_t(S_{t-1}))$. Thus, the optimal strategy is to transmit if the per-step distortion due to not transmitting is greater than a threshold.

Remark 14 As noted in Remark 8, the structural results for the infinite horizon are discussed in Section 3.7. Moreover, in its subsequent sections we characterize the optimal performance for the infinite horizon setup.

3.5 Proof of the structural results for finite horizon case

3.5.1 Proof of Lemma 3.4.1

Arbitrarily fix the estimation strategy g and consider the *best response* strategy at the transmitter. We will show that $\tilde{I}_t^1 := (X_t, S_{0:t-1}, Y_{0:t-1})$ is an information state at the transmitter.

Given any realization $(x_{0:T}, s_{0:T}, y_{0:T}, u_{0:T})$ of the system variables $(X_{0:T}, S_{0:T}, Y_{0:T}, U_{0:T})$, define $i_t^1 = (x_{0:t}, s_{0:t-1}, y_{0:t-1}, u_{0:t-1})$ and $\tilde{i}_t^1 = (x_t, s_{0:t-1}, y_{0:t-1})$. Now, for any $\breve{i}_{t+1}^1 = (\breve{x}_{t+1}, \breve{s}_{0:t}, \breve{y}_{0:t}) = (\breve{x}_{t+1}, \breve{s}_t, \breve{y}_t, \breve{i}_t^1)$, we use the shorthand $\mathbb{P}(\tilde{i}_{t+1}^1 | \tilde{i}_{0:t}^1, u_{0:t})$ to denote $\mathbb{P}(\tilde{I}_{t+1}^1 =$

$\breve{i}^1_{t+1} | \tilde{I}^1_{0:t} = \tilde{i}^1_{0:t}, U_{0:t} = u_{0:t})$. Then,

$$\begin{aligned}
\mathbb{P}(\breve{i}^1_{t+1} | i^1_t, u_t) &= \mathbb{P}(\breve{x}_{t+1}, \breve{s}_t, \breve{y}_t, \breve{i}^1_t | x_{0:t}, s_{0:t-1}, y_{0:t-1}, u_{0:t}) \\
&\stackrel{(a)}{=} \mathbb{P}(\breve{x}_{t+1}, \breve{s}_t, \breve{y}_t, \breve{i}^1_t | x_{0:t}, \bar{x}_{0:t}, s_{0:t-1}, y_{0:t-1}, u_{0:t}) \\
&\stackrel{(b)}{=} \mathbb{P}(\breve{x}_{t+1} | x_t) \mathbb{P}(\breve{y}_t | \bar{x}_t, \breve{s}_t) \mathbb{P}(\breve{s}_t | s_{t-1}) \mathbb{1}_{\{\breve{i}^1_t = \tilde{i}^1_t\}} \\
&= \mathbb{P}(\breve{i}^1_{t+1} | \tilde{i}^1_t, u_t)
\end{aligned} \tag{3.18}$$

where we have added $\bar{x}_{0:t}$ in the conditioning in (a) because $\bar{x}_{0:t}$ is a deterministic function of $(x_{0:t}, u_{0:t})$ and (b) follows from the source and the channel models. By marginalizing (3.18), we get that for any $\breve{i}^2_t = (\breve{s}_t, \breve{y}_t, \breve{i}^1_t)$, we have

$$\mathbb{P}(\breve{i}^2_t | i^1_t, u_t) = \mathbb{P}(\breve{i}^2_t | \tilde{i}^1_t, u_t) \tag{3.19}$$

Now, let $c(X_t, U_t, \hat{X}_t) = \lambda U_t + d(X_t, \hat{X}_t)$ denote the per-step cost. Recall that $\hat{X}_t = g_t(I^2_t)$. Thus, by (3.19), we get that

$$\mathbb{E}[c(X_t, U_t, \hat{X}_t) | i^1_t, u_t] = \mathbb{E}[c(X_t, U_t, \hat{X}_t) | \tilde{i}^1_t, u_t]. \tag{3.20}$$

Eq. (3.18) shows that $\{\tilde{I}^1_t\}_{t \geq 0}$ is a controlled Markov process controlled by $\{U_t\}_{t \geq 0}$. Eq. (3.20) shows that $\tilde{I}^1_t$ is sufficient for performance evaluation. Hence, by Markov decision theory [62], there is no loss of optimality in restricting attention to transmission strategies of the form (3.6).

3.5.2 Proof of Lemma 3.4.2

Consider

$$\begin{aligned}
\pi^1_{t+1}(x_{t+1}) &= \mathbb{P}(x_{t+1} | s_{0:t}, y_{0:t}) \\
&= \sum_{x_t \in \mathbb{X}} \mathbb{P}(x_{t+1} | x_t) \mathbb{P}(x_t | s_{0:t}, y_{0:t}) \\
&= \sum_{x_t \in \mathbb{X}} P_{x_t x_{t+1}} \pi^2_t(x_t) = \pi^2_t P
\end{aligned} \tag{3.21}$$

which is the expression for $F^1(\cdot)$.

For F^2, we consider the three cases separately. For $y_t \in \mathbb{X}$, we have

$$\pi_t^2(x) = \mathbb{P}(X_t = x|s_{0:t}, y_{0:t}) = \mathbb{1}_{\{x=y_t\}}. \tag{3.22}$$

For $y_t \in \{\mathfrak{E}_0, \mathfrak{E}_1\}$, we have

$$\begin{aligned}\pi_t^2(x) &= \mathbb{P}(X_t = x|s_{0:t}, y_{0:t}) \\ &= \frac{\mathbb{P}(X_t = x, y_t, s_t|s_{0:t-1}, y_{0:t-1})}{\mathbb{P}(y_t, s_t|s_{0:t-1}, y_{0:t-1})}\end{aligned} \tag{3.23}$$

Now, when $y_t = \mathfrak{E}_0$, we have that

$$\begin{aligned}\mathbb{P}(x_t, y_t, s_t|s_{0:t-1}, y_{0:t-1}) &= \mathbb{P}(y_t|x_t, \varphi_t(x_t), s_t)Q_{s_{t-1}s_t}\pi_t^1(x_t) \\ &\stackrel{(a)}{=} \begin{cases} Q_{s_{t-1}1}\pi_t^1(x_t), & \text{if } \varphi_t(x_t) = 0 \text{ and } s_t = 1 \\ 0, & \text{otherwise} \end{cases}\end{aligned} \tag{3.24}$$

where (a) is obtained from the channel model. Substituting (3.24) in (3.23) and canceling $Q_{s_{t-1}1}\mathbb{1}_{\{s_t=1\}}$ from the numerator and the denominator, we get (recall that this is for the case when $y_t = \mathfrak{E}_0$),

$$\pi_t^2(x) = \frac{\mathbb{1}_{\{\varphi_t(x)=0\}}\pi_t^1(x)}{\pi_t^1(B_0(\varphi))}. \tag{3.25}$$

Similarly, when $y_t = \mathfrak{E}_1$, we have that

$$\begin{aligned}\mathbb{P}(x_t, y_t, s_t|s_{0:t-1}, y_{0:t-1}) &= \mathbb{P}(y_t|x_t, \varphi_t(x_t), s_t)Q_{s_{t-1}s_t}\pi_t^1(x_t) \\ &\stackrel{(b)}{=} \begin{cases} Q_{s_{t-1}0}\pi_t^1(x_t), & \text{if } s_t = 0 \\ 0, & \text{otherwise} \end{cases}\end{aligned} \tag{3.26}$$

where (b) is obtained from the channel model. Substituting (3.26) in (3.23) and canceling $Q_{s_{t-1}0}\mathbb{1}_{\{s_t=0\}}$ from the numerator and the denominator, we get (recall that this is for the case when $y_t = \mathfrak{E}_1$),

$$\pi_t^2(x) = \pi_t^1(x). \tag{3.27}$$

By combining (3.22), (3.25) and (3.27), we get (3.8).

3.5.3 Proof of Theorem 3.4.1

Once we restrict attention to transmission strategies of the form (3.6), the information structure is partial history sharing [57]. Thus, one can use the common information approach of [57] and obtain the structure of optimal strategies.

Following [57], we split the information available at each agent into a "common information" and "local information". Common information is the information available to all decision makers in the future; the remaining data at the decision maker is the local information. Thus, at the transmitter, the common information is $C^1_t \coloneqq \{S_{0:t-1}, Y_{0:t-1}\}$ and the local information is $L^1_t \coloneqq X_t$. Similarly, at the receiver, the common information is $C^2_t \coloneqq \{S_{0:t}, Y_{0:t}\}$ and the local information is $L^2_t \coloneqq \emptyset$. When the transmitter makes a decision, the state (sufficient for input-output mapping) of the system is (X_t, S_{t-1}); when the receiver makes a decision, the state of the system is (X_t, S_t). By [57, Proposition 1], we get that the sufficient statistic Θ^1_t for the common information at the transmitter is

$$\Theta^1_t(x, s) = \mathbb{P}(X_t = x, S_{t-1} = s | S_{0:t-1}, Y_{0:t-1}),$$

and the sufficient statistic Θ^2_t for the common information at the receiver is

$$\Theta^2_t(x, s) = \mathbb{P}(X_t = x, S_t = s | S_{0:t}, Y_{0:t}).$$

Note that Θ^1_t is equivalent to (Π^1_t, S_{t-1}) and Θ^2_t is equivalent to (Π^2_t, S_t). Therefore, by [57, Theorem 2], there is no loss of optimality in restricting attention to transmission strategies of the form (3.9) and estimation strategies of the form

$$\hat{X}_t = g_t(S_t, \Pi^2_t). \tag{3.28}$$

Furthermore, the dynamic program of 3.4.1 follows from [57, Theorem 3].

Note that the right hand side of (3.13) implies that $\hat{X}_t$ does not depend on S_t. Thus, instead of (3.28), we can restrict attention to estimation strategy of the form (3.10). Furthermore, the optimal estimation strategy is given by (3.14).

3.6 Proof of Theorem 3.4.2

We prove the result for $\mathbb{X} = \mathbb{R}$. Similar argument holds for $\mathbb{X} = \mathbb{Z}$.

3.6.1 A change of variables

Define a process $\{Z_t\}_{t \ge 0}$ as follows: $Z_0 = 0$ and for $t \ge 0$,

$$Z_t = \begin{cases} aZ_{t-1}, & \text{if } Y_t \in \{\mathfrak{E}_0, \mathfrak{E}_1\} \\ Y_t, & \text{if } Y_t \in \mathbb{X} \end{cases}$$

Note that Z_t is a function of $Y_{0:t-1}$. Next, define processes $\{E_t\}_{t \ge 0}$, $\{E_t^+\}_{t \ge 0}$, and $\{\hat{E}_t\}_{t \ge 0}$ as follows:

$$E_t \coloneqq X_t - aZ_{t-1}, \quad E_t^+ \coloneqq X_t - Z_t, \quad \hat{E}_t \coloneqq \hat{X}_t - Z_t$$

The processes $\{E_t\}_{t \ge 0}$ and $\{E_t^+\}_{t \ge 0}$ are related as follows: $E_0 = 0$, $E_0^+ = 0$, and for $t \ge 0$

$$E_t^+ = \begin{cases} E_t, & \text{if } Y_t \in \{\mathfrak{E}_0, \mathfrak{E}_1\} \\ 0, & \text{if } Y_t \in \mathbb{X} \end{cases} \quad \text{and} \quad E_{t+1} = aE_t^+ + W_t.$$

Since $X_t - \hat{X}_t = E_t^+ - \hat{E}_t$, we have that $d(X_t - \hat{X}_t) = d(E_t^+ - \hat{E}_t)$.

It turns out that it is easier to work with the processes $\{E_t\}_{t \ge 0}$, $\{E_t^+\}_{t \ge 0}$, and $\{\hat{E}_t\}_{t \ge 0}$ rather than $\{X_t\}_{t \ge 0}$ and $\{\hat{X}_t\}_{t \ge 0}$.

Next, redefine the pre- and post-transmission beliefs in terms of the error process. With a slight abuse of notation, we still denote the probability density of the pre- and post-transmission beliefs as π_t^1 and π_t^2. In particular, π_t^1 is the conditional pdf (probability density function) of E_t given $(s_{0:t-1}, y_{0:t-1})$ and π_t^2 is the conditional pdf of E_t^+ given $(s_{0:t}, y_{0:t})$.

Let $H_t \in \{\mathfrak{E}_0, \mathfrak{E}_1, 1\}$ denote the event whether the transmission was successful or not. In particular,

$$H_t = \begin{cases} \mathfrak{E}_0, & \text{if } Y_t = \mathfrak{E}_0 \\ \mathfrak{E}_1, & \text{if } Y_t = \mathfrak{E}_1 \\ 1, & \text{if } Y_t \in \mathbb{X}. \end{cases}$$

Note that H_t is a deterministic function of U_t and S_t. We use h_t to denote the realization of H_t.

The time-evolutions of π_t^1 and π_t^2 are similar to Lemma 3.4.2. In particular, we have

Lemma 3.6.1 *Given any transmission strategy f of the form* (3.4)*:*

1. *there exists a function F^1 such that*

$$\pi_{t+1}^1 = F^1(\pi_t^2). \tag{3.29}$$

   In particular,

$$\pi_{t+1}^1 = \begin{cases} \tilde{\pi}_t^2 \star \mu, & \text{if } y_t \in \{\mathfrak{E}_0, \mathfrak{E}_1\} \\ \mu, & \text{if } y_t \in \mathbb{X}, \end{cases} \tag{3.30}$$

   where $\tilde{\pi}_t^2$ given by $\tilde{\pi}_t^2(e) := (1/|a|)\pi_t^2(e/a)$ is the conditional probability density of aE_t^+, μ is the probability density function of W_t and $\star$ is the convolution operation.

2. *there exists a function F^2 such that*

$$\pi_t^2 = F^2(\pi_t^1, \varphi_t, h_t). \tag{3.31}$$

   In particular,

$$\pi_t^2 = \begin{cases} \delta_0, & \text{if } h_t = 1 \\ \pi_t^1|_{\varphi_t}, & \text{if } h_t = \mathfrak{E}_1 \\ \pi_t^1, & \text{if } h_t = \mathfrak{E}_0. \end{cases} \tag{3.32}$$

The key difference between Lemmas 3.4.2 and 3.6.1 (and the reason that we work with the error process $\{E_t\}_{t\geq 0}$ rather than $\{X_t\}_{t\geq 0}$) is that the function F^2 in (3.31) depends on h_t rather than y_t. Consequently, the dynamic program of Theorem 3.4.1 is now given by

$$V_{T+1}^1(s, \pi^1) = 0, \tag{3.33}$$

and for $t \in \{T, \dots, 0\}$

$$\begin{aligned} V_t^1(s, \pi^1) = \min_{\varphi \colon \mathbb{R} \to \{0,1\}} \Big\{ \lambda \pi^1(B_1(\varphi)) \\ + W_t^0(\pi^1, \varphi)\pi^1(B_0(\varphi)) + W_t^1(\pi^1)\pi^1(B_1(\varphi)) \Big\}, \end{aligned} \tag{3.34}$$

$$V_t^2(s, \pi^2) = D(\pi^2) + V_{t+1}^1(s, F^1(\pi^2)), \tag{3.35}$$

where,

$$\begin{aligned} W_t^0(\pi^1, \varphi) &= Q_{s0} V_t^2(0, \pi^1) + Q_{s1} V_t^2(1, \pi^1|_\varphi), \\ W_t^1(\pi^1) &= Q_{s0} V_t^2(0, \pi^1) + Q_{s1} V_t^2(1, \delta_0), \\ D(\pi^2) &= \begin{cases} \min_{\hat{e} \in \mathbb{X}} \int_{\mathbb{X}} d(e - \hat{e})\pi^2(e) de, & \text{if } \mathbb{X} = \mathbb{R} \\ \min_{\hat{e} \in \mathbb{X}} \sum_{\mathbb{X}} d(e - \hat{e})\pi^2(e), & \text{if } \mathbb{X} = \mathbb{Z}. \end{cases} \end{aligned}$$

Note that due to the change of variables, the expression for W_t^1 does not depend on the transmitted symbol. Consequently, the expression for V_t^1 is simpler than that in Theorem 3.4.1.

3.6.2 Symmetric unimodal distributions and their properties

For ease of exposition, we state the results in this section for $\mathbb{X} = \mathbb{R}$. The results for $\mathbb{X} = \mathbb{Z}$ hold analogously.

A probability density function π on reals is said to be *symmetric and unimodal (*SU*)* around $c \in \mathbb{R}$ if for any $x \in \mathbb{R}$, $\pi(c - x) = \pi(c + x)$ and π is non-decreasing in the interval $(-\infty, c]$ and non-increasing in the interval $[c, \infty)$.

Given $c \in \mathbb{R}$, a prescription $\varphi \colon \mathbb{X} \to \{0, 1\}$ is called *threshold based around* c if there exists $k \in \mathbb{X}$ such that

$$\varphi(e) = \begin{cases} 1, & \text{if } |e - c| \geq k \\ 0, & \text{if } |e - c| < k. \end{cases}$$

Let $\mathcal{F}(c)$ denote the family of all threshold-based prescription around c.

Now, we state some properties of symmetric and unimodal distributions.

Property 3.6.1 *If π is* SU(c)*, then*

$$c \in \arg\min_{\hat{e} \in \mathbb{R}} \int_{\mathbb{X}} d(e - \hat{e})\pi(e)de.$$

For $c = 0$, the above property is a special case of [9, Lemma 12]. The result for general c follows from a change of variables.

Property 3.6.2 *If π^1 is* SU(0) *and $\varphi \in \mathcal{F}(0)$, then for any $h \in \{\mathfrak{E}_0, \mathfrak{E}_1, 1\}$, $F^2(\pi^1, \varphi, h)$ is* SU(0)*.*

Proof We prove the result for each $h \in \{\mathfrak{E}_0, \mathfrak{E}_1, 1\}$ separately. Recall the update of π^1 given by (3.32). For $h_t = \mathfrak{E}_0$, $\pi^2 = \pi^1$ and hence π^2 is SU(0). For $h_t = \mathfrak{E}_1$, $\pi^2 = \pi^1|_{\varphi}$; if $\varphi \in \mathcal{F}(0)$, then $\pi^1(x)\mathbb{1}_{\{\varphi(x)=0\}}$ is SU(0) and hence π^1 is SU(0). For $h_t = 1$, $\pi^2 = \delta_0$, which is SU(0).

Property 3.6.3 *If π^2 is* SU(0)*, then $F^1(\pi^2)$ is also* SU(0)*.*

Proof Recall that F^1 is given by (3.30). The property follows from the fact that convolution of symmetric and unimodal distributions is symmetric and unimodal.

3.6.3 SU majorization and its properties

For any set $\mathbb{A}$, let $\mathcal{I}_{\mathbb{A}}$ denote its indicator function, i.e., $\mathcal{I}_{\mathbb{A}}(x)$ is 1 if $x \in \mathbb{X}$, else 0.

Let $\mathbb{A}$ be a measurable set of finite Lebesgue measure, its *symmetric rearrangement* $\mathbb{A}^{\sigma}$ is the open interval centered around origin whose Lebesgue measure is same as $\mathbb{A}$.

Given a function $\ell \colon \mathbb{R} \to \mathbb{R}$, its super-level set at level ρ, $\rho \in \mathbb{R}$, is $\{x \in \mathbb{R} : \ell(x) > \rho\}$. The *symmetric decreasing rearrangement* ℓ^{σ} of ℓ is a symmetric and decreasing function whose level sets are the same as ℓ, i.e.,

$$\ell^{\sigma}(x) = \int_0^{\infty} \mathcal{I}_{\{z \in \mathbb{R} : \ell(z) > \rho\}^{\sigma}}(x) d\rho.$$

Given two probability density functions ξ and π over $\mathbb{R}$, ξ *majorizes* π, which is denoted by $\xi \succeq_m \pi$, if for all $\rho \geq 0$,

$$\int_{|x| \geq \rho} \xi^{\sigma}(x)dx \geq \int_{|x| \geq \rho} \pi^{\sigma}(x)dx.$$

Given two probability density functions ξ and π over $\mathbb{R}$, ξ *SU majorizes* π, which we denote by $\xi \succeq_a \pi$, if ξ is SU and ξ majorizes π.

Now, we state some properties of SU majorization from [9].

Property 3.6.4 *For any $\xi \succeq_a \pi$, where ξ is $\mathrm{SU}(c)$ and for any prescription φ, let $\theta \in \mathcal{F}(c)$ be a threshold-based prescription such that for $i \in \{0, 1\}$, $\xi(B_i(\theta)) = \pi(B_i(\varphi))$. Then, $\xi|_\theta \succeq_a \pi|_\varphi$. Consequently, for any $h \in \{\mathfrak{E}_0, \mathfrak{E}_1, 1\}$,*

$$F^2(\xi, \theta, h) \succeq_a F^2(\pi, \varphi, h).$$

For $c = 0$, the result follows from [9, Lemma 7 and 8]. The result for general c follows from change of variables.

Property 3.6.5 *For any $\xi \succeq_a \pi$, $F^1(\xi) \succeq_a F^1(\pi)$.*

This follows from [9, Lemma 10].

Recall the definition of $D(\pi^2)$ given after (3.35).

Property 3.6.6 *If $\xi \succeq_a \pi$, then*

$$D(\pi) \geq D(\pi^\sigma) \geq D(\xi^\sigma) = D(\xi).$$

This follows from [9, Lemma 11].

3.6.4 Qualitative properties of the value function and optimal strategy

Lemma 3.6.2 *The value functions V_t^1 and V_t^2 of (3.33)–(3.35) satisfy the following property.*

(P1) For any $i \in \{1, 2\}$, $s \in \{0, 1\}$, $t \in \{0, \dots, T\}$, and pdfs (probability density functions) ξ^i and π^i such that $\xi^i \succeq_a \pi^i$, we have that $V_t^i(s, \xi^i) \leq V_t^i(s, \pi^i)$.

Furthermore, the optimal strategy satisfies the following properties. For any $s \in \{0, 1\}$ and $t \in \{0, \dots, T\}$:

(P2) if π^1 is $\mathrm{SU}(c)$, then there exists a prescription $\varphi_t \in \mathcal{F}(c)$ that is optimal. In general, φ_t depends on π^1.

(P3) if π^2 is $\mathrm{SU}(c)$*, then the optimal estimate $\hat{E}_t$ is c.*

Proof We proceed by backward induction. $V^1_{T+1}(s,\pi^1)$ trivially satisfies the (P1). This forms the basis of induction. Now assume that $V^1_{t+1}(s,\pi^1)$ also satisfies (P1). For $\xi^2 \succeq_a \pi^2$, we have that

$$\begin{aligned} V^2_t(s,\pi^2) &= D(\pi^2) + V^1_{t+1}(s, F^1(\pi^2)) \\ &\overset{(a)}{\geq} D(\xi^2) + V^1_{t+1}(s, F^1(\xi^2)) \\ &= V^2_t(s,\xi^2), \end{aligned} \tag{3.36}$$

where (a) follows from Properties 3.6.5 and 3.6.6 and the induction hypothesis. Eq. 3.36 implies that V^2_t also satisfies (P1).

Now, consider $\xi^1 \succeq_a \pi^1$. Let φ be the optimal prescription at π^1. Let θ be the threshold-based prescription corresponding to φ as defined in Property 3.6.3. By construction,

$$\pi^1(B_0(\varphi)) = \xi^1(B_0(\theta)) \quad \text{and} \quad \pi^1(B_1(\varphi)) = \xi^1(B_1(\theta)).$$

Moreover, from Property 3.6.3 and (3.36),

$$W^0_t(\pi^1,\varphi) \geq W^0_t(\xi^1,\theta) \quad \text{and} \quad W^1_t(\pi^1,\varphi) \geq W^1_t(\xi^1,\theta).$$

Combining the above two equations with (3.34), we get

$$\begin{aligned} V^1_t(s,\pi^1) &= \lambda\pi^1(B_1(\varphi)) + W^0(\pi^1,\varphi)\pi^1(B_0(\varphi)) \\ &\quad + W^1(\pi^1,\varphi)\pi^1(B_1(\varphi)) \\ &\geq \lambda\xi^1(B_1(\theta)) + W^0(\xi^1,\theta)\xi^1(B_0(\theta)) \\ &\quad + W^1(\xi^1,\theta)\xi^1(B_0(\theta)) \\ &\geq V^1_t(s,\xi^1) \end{aligned} \tag{3.37}$$

where the last inequality follows by minimizing over all θ. Eq. (3.37) implies that V^1_t also satisfies (P1). Hence, by the principle of induction, (P1) is satisfied for all time.

The argument in (3.37) also implies (P2). Furthermore, (P3) follows from Property 3.6.1.

3.6.5 Proof of Theorem 3.4.2

We first prove a weaker version of the structure of optimal transmission strategies. In particular, there exist threshold functions $\tilde{k}_t \colon \{0,1\} \times \Delta(\mathbb{R}) \to \mathbb{R}_{\geq 0}$ such that the following transmission strategy is optimal:

$$f_t(X_t, S_{t-1}, \Pi^1_t) = \begin{cases} 1, & \text{if } |X_t - aZ_{t-1}| \geq \tilde{k}_t(S_{t-1}, \Pi^1_t) \\ 0, & \text{otherwise} \end{cases} \tag{3.38}$$

or, equivalently, in terms of the $\{E_t\}_{t\geq 0}$ process:

$$f_t(E_t, S_{t-1}, \Pi^1_t) = \begin{cases} 1, & \text{if } |E_t| \geq \tilde{k}_t(S_{t-1}, \Pi^1_t) \\ 0, & \text{otherwise.} \end{cases} \tag{3.39}$$

We prove (3.39) by induction. Note that $\pi^1_0 = \delta_0$ which is SU(0). Therefore, by (P2), there exists a threshold-based prescription $\varphi_0 \in \mathcal{F}(0)$ that is optimal. This forms the basis of induction. Now assume that until time $t-1$, all prescriptions are in $\mathcal{F}(0)$. By Properties 3.6.2 and 3.6.3, Π^1_t is SU(0). Therefore, by (P2), there exists a threshold-based prescription $\varphi_t \in \mathcal{F}(0)$ that is optimal. This proves the induction step and, hence, by the principle of induction, threshold-based prescriptions of the form (3.39) are optimal for all time. Translating the result back to $\{X_t\}_{t\geq 0}$, we get that threshold-based prescriptions of the form (3.38) are optimal.

Observe that Properties 3.6.2 and 3.6.3 also imply that for all t, Π^2_t is SU(0). Therefore, by Property 3.6.1, the optimal estimate $\hat{E}_t = 0$. Recall that $\hat{E}_t = \hat{X}_t - Z_t$. Thus, $\hat{X}_t = Z_t$. This proves the first part of Theorem 3.4.2.

To prove that there exist optimal transmission strategies where the thresholds do not depend on Π^1_t, we fix the estimation strategy to be of the form (3.16) and consider the problem of finding the best transmission strategy at the sensor. This is a single-agent (centralized) stochastic control problem and the optimal solution is given by the following dynamic program:

$$J_{T+1}(e, s) = 0 \tag{3.40}$$

and for $t \in \{T, \dots, 0\}$

$$J_t(e,s) = \min\{J_t^0(e,s), J_t^1(e,s)\} \tag{3.41}$$

where

$$\begin{aligned} J_t^0(e,s) &= d(e) + Q_{s0}\mathbb{E}_W[J_{t+1}(ae+W,0)] \\ &\quad + Q_{s1}\mathbb{E}_W[J_{t+1}(ae+W,1)], \end{aligned} \tag{3.42}$$

$$\begin{aligned} J_t^1(e,s) &= \lambda + Q_{s0}d(e) + Q_{s0}\mathbb{E}_W[J_{t+1}(ae+W,0)] \\ &\quad + Q_{s1}\mathbb{E}_W[J_{t+1}(W,1)], \end{aligned} \tag{3.43}$$

We now use the results of Chapter 6 to show that the value function is even and quasi-convex on $\mathbb{R}_{\geq 0}$.

The results of Chapter 6 rely on stochastic dominance. Given two probability density functions ξ and π over $\mathbb{R}_{\geq 0}$, ξ *stochastically dominates* π, which we denote by $\xi \succeq_s \pi$, if

$$\int_{x \geq y} \xi(x)dx \geq \int_{x \geq y} \pi(x)dx, \quad \forall y \in \mathbb{R}_{\geq 0}.$$

Now, we show that dynamic program (3.40)–(3.43) satisfies conditions (C1)–(C3) of Theorem 1 in Chapter 6. In particular, we have: Condition (C1) is satisfied because the per-step cost functions $d(e)$ and $\lambda + Q_{s0}d(e)$ are even and quasi-convex. Condition (C2) is satisfied because the probability density μ of W_t is even, which implies that for any $e \in \mathbb{R}_{\geq 0}$,

$$\int_{w \in \mathbb{R}} \mu(ae+w)dw = \int_{w \in \mathbb{R}} \mu(-ae+w)dw.$$

Now, to check condition (C3), define for $e \in \mathbb{R}$ and $y \in \mathbb{R}_{\geq 0}$,

$$\begin{aligned} M^0(y|e) &= \int_y^{\infty} \mu(ae+w)dw + \int_{-\infty}^{-y} \mu(ae+w)dw \\ &= 1 - \int_{-y}^{y} \mu(ae+w)dw, \\ M^1(y|e) &= \int_y^{\infty} \mu(w)dw + \int_{-\infty}^{-y} \mu(w)dw. \end{aligned}$$

$M^1(y|e)$ does not depend on e and is thus trivially even and quasi-convex in e. Since μ is even, $M^0(y|e)$ is even in e. We show that $M^0(y|e)$ is increasing in e for $e \in \mathbb{R}_{\geq 0}$ later (see Lemma 3.6.3).

Since conditions (C1)–(C3) of Theorem 6.1.1 in Chapter 6 are satisfied, we have that for any $s \in \{0,1\}$, $J_t(e,s)$ is even in e and increasing for $e \in \mathbb{R}_{\geq 0}$. Now, observe that

$$J^0(e,s) - J^1(e,s) = (1-Q_{s0})d(e) + Q_{s1}\mathbb{E}_W[J_{t+1}(ae+W,1)] - \lambda - Q_{s1}\mathbb{E}_W[J_{t+1}(W,1)]$$

which is even in e and increasing in $e \in \mathbb{R}_{\geq 0}$.

Therefore, for any fixed $s \in \{0,1\}$, the set A of e in which $J_t^0(e,s) - J_t^1(e,s) \leq 0$ is convex and symmetric around the origin, i.e., a set of the form $[-k_t(s), k_t(s)]$. Thus, there exist a $k_t(\cdot)$ such that the action $u_t = 0$ is optimal for $e \in [-k_t(s), k_t(s)]$. This proves the structure of the optimal transmission strategy.

Now we prove Lemma 3.6.3. Note that a more generalized result is given by Claim 1 in Chapter 6 and is proved in Appendix E.1. Here we prove the result for binary actions for the continuity of the reading.

Lemma 3.6.3 *For any $y \in \mathbb{R}_{\geq 0}$, $M^0(y|e)$ is increasing in e, $e \in \mathbb{R}_{\geq 0}$.*

Proof To show that $M^0(y|e)$ is increasing in e for $e \in \mathbb{R}_{\geq 0}$, it sufficies to show that $1 - M^0(y|e) = \int_{-y}^{y} \mu(ae+w)dw$ is decreasing in e for $e \in \mathbb{R}_{\geq 0}$. Consider a change of variables $x = ae + w$. Then,

$$1 - M^0(y|e) = \int_{-y}^{y} \mu(ae+w)dw = \int_{-y-ae}^{y-ae} \mu(x)dx \tag{3.44}$$

Taking derivative with respect to e, we get that

$$\frac{\partial M^0(y|e)}{\partial e} = a[\mu(y-ae) - \mu(-y-ae)] \tag{3.45}$$

Now consider the following cases:

- If $a > 0$ and $y > ae > 0$, then the right hand side of (3.45) equals $a[\mu(y-ae) - \mu(y+ae)]$, which is positive.
- If $a > 0$ and $ae > y > 0$, then the right hand side of (3.45) equals $a[\mu(ae-y) - \mu(ae+y)]$, which is positive.

- If $a < 0$ and $y > |a|e > 0$, then the right hand side of (3.45) equals $|a|\,[\mu(y - |a|e) - \mu(y + |a|e)]$, which is positive.

- If $a < 0$ and $|a|e > y > 0$, then the right hand side of (3.45) equals $|a|\,[\mu(|a|e - y) - \mu(|a|e + y)]$, which is positive.

Thus, in all cases, $M^0(y|e)$ is increasing in e, $e \in \mathbb{R}_{\geq 0}$.

3.7 Structural results for infinite horizon optimization problem

Under some technical assumptions [60], the structural results for the finite horizon can be extended to the infinite horizon. In the rest of this chapter, we discuss the results for the infinite horizon. Let us first define the performances in the infinite horizon setup. In this context, we consider two cases: *discounted* and *long-term average*.

For the above two cases, similar to the finite horizon setup, the *costly* performance of any communication strategy (f, g) for the infinite horizon is given by the following:

- **Discounted cost**: Consider a discount factor $\beta \in (0, 1)$. Then

$$C_\beta(f, g; \lambda) = (1 - \beta)\mathbb{E}^{(f,g)}\left[\sum_{t=0}^{\infty} \beta^t(\lambda U_t + d(X_t, \hat{X}_t))\right] \tag{3.46}$$

  where the expectation is taken with respect to the joint measure on all system variables induced by the choice of (f, g).

- **Long-term average cost**:

$$C_1(f, g; \lambda)^3 = \lim_{T\to\infty} \frac{1}{T}\mathbb{E}^{(f,g)}\left[\sum_{t=0}^{T-1} \lambda U_t + d(X_t, \hat{X}_t)\right] \tag{3.47}$$

3.7.1 Performance metrics for the first order autoregressive process

Given a transmission and estimation strategy (f, g) and a discount factor $\beta \in (0, 1]$, we define the expected distortion and the expected number of transmissions as follows. For

[3] For consistency of notation, here we consider $\beta = 1$ as we can get the results for long-term average by taking $\beta \uparrow 1$ using *vanishing discount* approach.

$\beta \in (0, 1)$, the expected *discounted* distortion is given by

$$D_\beta(f, g) \coloneqq (1-\beta)\mathbb{E}^{(f,g)}\Big[\sum_{t=0}^{\infty} \beta^t d(X_t - \hat{X}_t) \;\Big|\; X_0 = 0\Big] \tag{3.48}$$

and for $\beta = 1$, the expected *long-term average* distortion is given by

$$D_1(f, g) \coloneqq \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}^{(f,g)}\Big[\sum_{t=0}^{T-1} d(X_t - \hat{X}_t) \;\Big|\; X_0 = 0\Big]. \tag{3.49}$$

Note that here we used with an abuse of notation $d(X_t - \hat{X}_t)$ to indicate the per step distortion.

Similarly, for $\beta \in (0, 1)$, the expected *discounted* number of transmissions is given by

$$N_\beta(f, g) \coloneqq (1-\beta)\mathbb{E}^{(f,g)}\Big[\sum_{t=0}^{\infty} \beta^t U_t \;\Big|\; X_0 = 0\Big] \tag{3.50}$$

and for $\beta = 1$, the expected *long-term average* number of transmissions is given by

$$N_1(f, g) \coloneqq \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}^{(f,g)}\Big[\sum_{t=0}^{T-1} U_t \;\Big|\; X_0 = 0\Big]. \tag{3.51}$$

Note that for $\beta \in (0, 1]$, $C_\beta(f, g; \lambda) = D_\beta(f, g) + \lambda N_\beta(f, g)$.

Remark 15 Similar to Chapter 2, we use a normalizing factor of $(1-\beta)$ to have a unified scaling for both discounted and long-term average setups. In particular, we will show that for any strategy (f, g)

$$C_1(f, g; \lambda) = \lim_{\beta\uparrow 1} C_\beta(f, g; \lambda), \quad \text{and} \quad D_1(f, g) = \lim_{\beta\uparrow 1} D_\beta(f, g).$$

Similar notation is used in [42].

3.7.2 Costly communication

Now we state the optimization problem for *costly communication* with the infinite horizon.

Problem 3.7.1 (Costly communication) *In the model described in* (3.15)*, for a discount factor* $\beta \in (0,1]$*, identify a communication strategy* (f^*, g^*) *that minimizes the cost* $C_\beta(f, g; \lambda)$ *defined in* (3.46)–(3.47)*.*

The following is the main result for Problem 3.7.1.

Theorem 3.7.1 *For a first-order autoregressive source with symmetric and unimodal disturbance,*

1. Structure of optimal estimation strategy: *for Problem 3.7.1, the optimal estimation strategy same as given by Theorem 3.4.2.*

2. Structure of optimal transmission strategy: *for Problem 3.7.1, there exist time-homogeneous threshold functions* $k\colon \{0,1\} \to \mathbb{X}_{\geq 0}$ *such that the following transmission strategy is optimal:*

$$f_t^*(X_t, S_{t-1}, \Pi_t^1) = \begin{cases} 1, & \text{if } |X_t - a\hat{X}_{t-1}| \geq k(S_{t-1}) \\ 0, & \text{otherwise.} \end{cases} \tag{3.52}$$

The proof technique of Theorem 3.7.1 for the discounted case, i.e., for $\beta \in (0,1)$ is similar to that of Theorem 3.4.2 and therefore is omitted. The only difference being that the threshold functions here are time-homogeneous. This is because when we fix the estimator to be of the form given in Theorem 3.7.1 and find the *best* transmitter, the dynamic program of the corresponding centralized Markov Decision Process can be shown to be a contraction and hence has a time-homogeneous unique solution [10]. For the long-term average case, i.e., for $\beta = 1$, we apply the *vanishing discount* approach, similar to Chapter 2.

3.7.3 Implication of Theorem 3.7.1 and dynamic programming decomposition

The implication of Theorem 3.7.1 is the following. In general, in RE problems, the structure of optimal estimation strategy depends on that of the optimal transmission strategy. However, as is shown in Theorem 3.7.1, the optimal estimation strategy can be characterized in closed form, independent of that of optimal transmission strategy. Thus, we can fix an estimation strategy of the form (3.16) and consider the optimization problem of finding the *best* transmission strategy corresponding the fixed estimation strategy. Since there is only one decision-maker (the transmitter), this optimization problem is centralized in nature. Since the optimal estimation strategy given by (3.16) is time-homogeneous, it can be shown that the optimal transmission strategy for discounted-case infinite horizon is time-homogeneous and is given by the following dynamic program:

$$V_\beta(e,s) = \min\{V_\beta^1(e,s), V_\beta^0(e,s)\} \tag{3.53}$$

where

$$\begin{aligned} V_\beta^1(e,s) = (1-Q_{s0})\big(\lambda + \beta\mathbb{E}[V_\beta(E_{t+1},S_t) \,|\, E_t = e, S_{t-1} = s, U_t = 1, C_t = 1] \\ + Q_{s0}(\lambda + d(e) + \beta\mathbb{E}[V_\beta(E_{t+1},S_t) \,|\, E_t = e, S_{t-1} = s, U_t = 1, C_t = 0]\big) \end{aligned}$$

and

$$V_\beta^0(e,s) = d(e) + \beta\mathbb{E}[V_\beta(E_{t+1},S_t) \,|\, E_t = e, S_{t-1} = s, U_t = 0].$$

Let D_β^1 denote the performance of a strategy in which we transmit all the time. We assume that D_β^1 is uniformly bounded[4], say by $\bar{D}_\beta^1 < \infty$.

The above dynamic program has a unique solution due to the following reasons. Let us consider $\mathbb{X} = \mathbb{Z}$. When the per-step distortion $d(\cdot)$ is bounded, the existence of a unique and bounded solution follows from [51, Proposition 4.7.1, Theorem 4.6.3]. When $d(\cdot)$ is unbounded, then for any communication cost λ, we first define $e_0 \in \mathbb{Z}_{\geq 0} < \infty$ as:

$$e_0 \coloneqq \min\Big\{e \,:\, d(e) \geq \frac{\bar{D}_\beta^1}{1-\beta}\Big\}.$$

Now, for any state e, $|e| > e_0$, the per-step cost $(1-\beta)d(e)$ of not transmitting is greater then the cost of transmitting at each step in the future, which is given by $\bar{D}_\beta^1$. Thus, the optimal action is to transmit, i.e., $f^*(e) = 1$.

Let $\mathcal{E}^* \coloneqq \{e \,:\, |e| \geq e_0\}$. Then the countable-state state-process is equivalent to a finite-state Markov chain with state space $\{-e_0+1, \cdots, e_0-1\} \cup e^*$ (where e^* is a generic state for all states in the set $\mathcal{E}^*$). Since the state space is now finite, the dynamic program (3.53) has a unique and bounded time-homogeneous solution by the argument given for bounded $d(\cdot)$. The argument goes through for $\mathbb{X} = \mathbb{R}$, with the exception that in that case the

[4]If D_β^1 is not uniformly bounded, then the performance of every strategy is infinite and seeking an optimal strategy is meaningless.

effective state space $[-e_0, e_0] \cup e^*$ is compact and hence the dynamic program (3.53) has a unique and bounded time-homogeneous solution.

The results for the long-term average case can be obtained by using vanishing discount approach. Similar to the technique adopted in Chapter 2, one can show that the value function satisfies Proposition 2.6.3.

Proof Let $V_\beta^{(0)}(e, s; \lambda)$ denote the value function of the 'always transmit' strategy. Since $V_\beta(0, s; \lambda) \le V_\beta^{(0)}(0, s; \lambda)$ and $V_\beta^{(0)}(0, s; \lambda) = \lambda$, (S1) is satisfied with $e_0 = 0$ and $M_\lambda = \lambda$.

We show (S2) for $\mathbb{X} = \mathbb{R}$, but a similar argument works for $\mathbb{X} = \mathbb{Z}$ as well. Since not transmitting is optimal at state 0, we have

$$V_\beta(0, s; \lambda) = \beta[Q_{s0} \int_{-\infty}^{\infty} \mu(w) V_\beta(w, 0; \lambda) dw + Q_{s1} \int_{-\infty}^{\infty} \mu(w) V_\beta(w, 1; \lambda) dw].$$

Let $V_\beta^{(1)}(e, s; \lambda)$ denote the value function of the strategy that transmits at time 0 and follows the optimal strategy from then on. Then

$$\begin{aligned} V_\beta^{(1)}(e, s; \lambda) &= (1-\beta)[\lambda + Q_{s0} d(e)] + \beta[Q_{s0} \int_{-\infty}^{\infty} \mu(w) V_\beta(w, 0; \lambda) dw + Q_{s1} \int_{-\infty}^{\infty} \mu(w) V_\beta(w, 1; \lambda) dw] \\ &= (1-\beta)[\lambda Q_{s0} d(e)] + \beta V_\beta(0, s; \lambda) \end{aligned} \tag{3.54}$$

Since $V_\beta(e, s; \lambda) \le V_\beta^{(1)}(e; \lambda)$ and $V_\beta(0, s; \lambda) \ge 0$, from (3.54) we get that $(1-\beta)^{-1}[V_\beta(e, s; \lambda) - V_\beta(0, s; \lambda)] \le \lambda + Q_{s0} d(e)$. Hence (S2) is satisfied with $K_\lambda(e) = \lambda + Q_{s0} d(e)$.

As shown in Section 3.6.5, the value function is even and quasi-convex and hence $V_\beta(e, s; \lambda) \ge V_\beta(0, s; \lambda)$. Hence (S3) is satisfied with $L_\lambda = 0$.

Now, we are ready to show the proof of Theorem 3.7.1 for long-term average set-up.

Proof (Proof of Theorem 3.7.1 for $\boldsymbol{\beta = 1}$) Since the value function $V_\beta(\cdot; \lambda)$ satisfies the SEN conditions for reference state $e_0 = 0$, the optimality of the threshold strategy for long-term average setup follows from [51, Theorem 7.2.3] for $\mathbb{X} = \mathbb{Z}$ and [52, Theorem 5.4.3] for $\mathbb{X} = \mathbb{R}$, respectively.

3.8 Computation of the performances of a generic threshold based strategy

Recall the process E_t introduced in Section 3.6.1. We call this process the *error process*.

Let $f^{(k)}$ denote the threshold-based transmission strategy:

$$f^{(k)}(E_t, S_{t-1}) \coloneqq \begin{cases} 1, & \text{if } |E_t| \geq k(S_{t-1}) \\ 0, & \text{if } |E_t| < k(S_{t-1}). \end{cases} \tag{3.55}$$

With a slight abuse of notation, to show the dependence of the thresholds on the channel state S_{t-1}, we use k_s, for any realization $s \in \{0, 1\}$ of S_{t-1}. Note that (3.55) can be written as:

$$f^{(k_0,k_1)}(E_t, S_{t-1}) \coloneqq \begin{cases} 1, & \text{if } S_{t-1} = 0 \text{ and } |E_t| \geq k_0 \\ 0, & \text{if } S_{t-1} = 0 \text{ and } |E_t| < k_0 \\ 1, & \text{if } S_{t-1} = 1 \text{ and } |E_t| \geq k_1 \\ 0, & \text{if } S_{t-1} = 1 \text{ and } |E_t| < k_1. \end{cases} \tag{3.56}$$

Denote the set $\mathcal{S}^{(k_s)}$ as follows:

$$\mathcal{S}^{(k_s)} = \begin{cases} (-k_s, k_s), & \text{if } \mathbb{X} = \mathbb{R} \\ \{-k_s + 1, \cdots, k_s - 1\}, & \text{if } \mathbb{X} = \mathbb{Z}. \end{cases}$$

Note that (3.55) implies that when the error state $e \in \mathcal{S}^{(k_s)}$, the transmitter does not transmit. For this reason, we call the set $\mathcal{S}^{(k_s)}$ the *silent set*. Denote by $\mu(\mathcal{S}^{(k_s)})$ the following:

$$\mu(\mathcal{S}^{(k_s)}) \coloneqq \begin{cases} \int_{e \in \mathcal{S}^{(k_s)}} \mu(e) de, & \text{if } \mathbb{X} = \mathbb{R} \\ \sum_{e \in \mathcal{S}^{(k_s)}} \mu(e), & \text{if } \mathbb{X} = \mathbb{Z}. \end{cases}$$

Define the functions $\mu^{(k_s)}$ and $d^{(k_s)}$ as follows:

$$\mu^{(k_s)}(e) \coloneqq \begin{cases} \mu(e), & \text{if } e \in \mathcal{S}^{(k_s)} \\ \frac{\mu(e)}{\mu(\mathcal{S}^{(k_s)})}, & \text{if } e \in \mathbb{X} \setminus \mathcal{S}^{(k_s)}. \end{cases} \tag{3.57}$$

Note that $\mu^{(k_s)}$ is the posterior π_t^2 mentioned in Lemma 3.4.2 and the process noise density

μ refers π_t^1 in the same lemma.

$$d^{(k_s)}(e) \coloneqq \begin{cases} d(e), & \text{if } e \in \mathcal{S}^{(k_s)} \\ Q_{s0} d(e), & \text{if } e \in \mathbb{X} \setminus \mathcal{S}^{(k_s)}, \end{cases} \tag{3.58}$$

where $Q_{s0} \coloneqq \mathbb{P}(S_t = 0 \,|\, S_{t-1} = s)$.

For any $v : \mathbb{X} \times \{0,1\} \to \mathbb{R}$, define operator $\mathcal{B}^{(k_0,k_1)}$ as

$$[\mathcal{B}^{(k_0,k_1)} v](e,s) \coloneqq \tag{3.59}$$
$$\begin{cases} \sum_{s' \in \{0,1\}} Q_{ss'} \int_{n \in \mathbb{X}} \mu^{(k_s)}(n - ae) v(n, s') dn, & \text{if } \mathbb{X} = \mathbb{R} \\ \sum_{s' \in \{0,1\}} Q_{ss'} \sum_{n \in \mathbb{X}} \mu^{(k_s)}(n - ae) v(n, s'), & \text{if } \mathbb{X} = \mathbb{Z}, \end{cases}$$

where $e \in \mathbb{X}$ and $s, s' \in \{0,1\}$. For $\beta \in (0,1]$, the source-state $e \in \mathbb{X}$ and the channel-state $s \in \{0,1\}$, define the following for a system that starts in state $(e,s) \in \mathbb{X} \times \{0,1\}$ and follows strategy $f^{(k_0,k_1)}$:

- $L_\beta^{(k_0,k_1)}(e,s)$: the expected distortion until the first successful reception
- $M_\beta^{(k_0,k_1)}(e,s)$: the expected time until the first successful reception
- $K_\beta^{(k_0,k_1)}(e,s)$: the expected number of transmissions until the first successful reception
- $D_\beta^{(k_0,k_1)}(e,s)$: the expected distortion
- $N_\beta^{(k_0,k_1)}(e,s)$: the expected number of transmissions
- $C_\beta^{(k_0,k_1)}(e,s;\lambda)$: the expected total cost, i.e.,

$$C_\beta^{(k_0,k_1)}(e,s;\lambda) \coloneqq D_\beta^{(k_0,k_1)}(e,s) + \lambda N_\beta^{(k_0,k_1)}(e,s), \quad \lambda \geq 0.$$

Note that under $f^{(k_0,k_1)}$, $\{E_t\}_{t \geq 0}$ is a Markov process. From the balance equations, we get: for all $a, e \in \mathbb{X}$, $s \in \{0,1\}$ and $k_s \in \mathbb{X}_{\geq 0}$

$$L_\beta^{(k_0,k_1)}(e,s) = \begin{cases} Q_{s0}\Big[d(e) + \beta[\mathcal{B}^{(k_0,k_1)} L_\beta^{(k_0,k_1)}](e,s)\Big], & \text{if } |e| \geq k_s \\ d(e) + \beta[\mathcal{B}^{(k_0,k_1)} L_\beta^{(k_0,k_1)}](e,s), & \text{if } |e| < k_s, \end{cases} \tag{3.60}$$

$$M_\beta^{(k_0,k_1)}(e,s) = \begin{cases} Q_{s0}\Big[1+\beta[\mathcal{B}^{(k_0,k_1)}M_\beta^{(k_0,k_1)}](e,s)\Big], & \text{if } |e| \geq k_s \\ 1+\beta[\mathcal{B}^{(k_0,k_1)}M_\beta^{(k_0,k_1)}](e,s), & \text{if } |e| < k_s, \end{cases} \tag{3.61}$$

and

$$K_\beta^{(k_0,k_1)}(e,s) = \begin{cases} 1+Q_{s0}\beta[\mathcal{B}^{(k_0,k_1)}K_\beta^{(k_0,k_1)}](e,s), & \text{if } |e| \geq k_s \\ \beta[\mathcal{B}^{(k_0,k_1)}K_\beta^{(k_0,k_1)}](e,s), & \text{if } |e| < k_s. \end{cases} \tag{3.62}$$

Following the proof technique adopted in Chapter 2, one can show the following (see Appendix B.1 for the detailed proof).

Lemma 3.8.1 *Equations* (3.60) *and* (3.61) *have unique solutions* $L_\beta^{(k_0,k_1)}$ *and* $M_\beta^{(k_0,k_1)}$ *that are strictly increasing in* k_s *for* $s \in \{0,1\}$ *and Equation* (3.62) *has a unique solution that is decreasing in* k_s*.*

Proposition 3.8.1 *For any* $\beta \in (0,1]$*, the performance of strategy* $f^{(k_0,k_1)}$ *for costly communication is given as follows: For* $k_s \in \mathbb{X}_{>0}$ *and* $s \in \{0,1\}$*,*

$$D_\beta^{(k_0,k_1)}(0,s) := D_\beta(f^{(k_0,k_1)}, g^*) = \frac{L_\beta^{(k_0,k_1)}(0,s)}{M_\beta^{(k_0,k_1)}(0,s)},$$

$$N_\beta^{(k_0,k_1)}(0,s) = \frac{K_\beta^{(k_0,k_1)}(0,s)}{M_\beta^{(k_0,k_1)}(0,s)},$$

and

$$C_\beta^{(k_0,k_1)}(0,s;\lambda) := C_\beta(f^{(k_0,k_1)}, g^*;\lambda) = \frac{L_\beta^{(k_0,k_1)}(0,s) + \lambda K_\beta^{(k_0,k_1)}(0,s)}{M_\beta^{(k_0,k_1)}(0,s)}.$$

See Appendix B.2 for the proof.

Note that by Proposition 3.8.1, one can compute $D_\beta^{(k_0,k_1)}$ and $N_\beta^{(k_0,k_1)}$ from the knowledge of $L_\beta^{(k_0,k_1)}$, $M_\beta^{(k_0,k_1)}$ and $K_\beta^{(k_0,k_1)}$.

3.8.1 Computation of $L_\beta^{(k_0,k_1)}$, $M_\beta^{(k_0,k_1)}$ and $K_\beta^{(k_0,k_1)}$ for $\mathbb{X} = \mathbb{Z}$

Definition 3.8.1 *The Hadamard product* $\odot$ *for vectors and matrices is defined as follows:*

- *For two vectors of same dimension,* v, w*,* $v \odot w$ *denotes the element-wise product, i.e., the* i*-th element of the vector* $v \odot w$ *is given by* $(v \odot w)_i = v_i w_i$*,*

- *For a vector v and a matrix W with the same number of rows, $v \odot W$ is a matrix given by*

$$[v \odot W]_{i,j} = v(i)W_{i,j}.$$

Note that $\odot$ is associative in the following sense. For vectors u, v and matrix W with compatible dimensions, $v \odot (Wu) = (v \odot W)u$.

Define the vector $h^{(k_s)}$ as follows:

$$h_e^{(k_s)} := \begin{cases} Q_{s0}, & \text{if } |e| \geq k_s, \\ 1, & \text{if } |e| < k_s. \end{cases}$$

Denote by $h^{(k_0,k_1)} \in \mathbb{Z}^2 \times 1 := [[h^{(k_0)}]^\intercal, [h^{(k_1)}]^\intercal]^\intercal$. Let us now consider the augmented states $z := (e, s)$ and $z' := (e', s')$. Define matrix $\tilde{P}$ on $(\mathbb{Z} \times \{0, 1\})^2$ as follows:

$$\tilde{P}_{zz'} = Q_{ss'}\mu_{|e'-ae|}, \quad \forall z, z' \in \mathbb{Z} \times \{0, 1\}.$$

Let P^e denote the transition probability matrix of the error process E_t and recall the same for the channel given by matrix Q. Note that, since E_t and S_t are statistically independent, we can write $\tilde{P}$ as $\tilde{P} = P^e \otimes Q$, where $\otimes$ denotes the Kronecker product of two matrices.

Define the operator $\mathcal{B}^{(k_0,k_1)}$ as given by the following

$$\mathcal{B}^{(k_0,k_1)}v := h^{(k_0,k_1)} \odot (\tilde{P}v) = (h^{(k_0,k_1)} \odot \tilde{P})v. \tag{3.63}$$

Note that (3.60) can be written as

$$L_\beta^{(k_0,k_1)} = h^{(k_0,k_1)} \odot \tilde{d} + \beta\mathcal{B}^{(k_0,k_1)}L_\beta^{(k_0,k_1)}, \tag{3.64}$$

where $\tilde{d} := [[d^{(k_0)}]^\intercal, [d^{(k_1)}]^\intercal]^\intercal$ with $d^{(k_0)}, d^{(k_1)}$ defined in (3.58).

Following Lemma 3.8.1, (3.64) has a unique fixed point solution given by

$$L_\beta^{(k_0,k_1)} = \left[I - \beta\mathcal{B}^{(k_0,k_1)}\right]^{-1}(h^{(k_0,k_1)} \odot \tilde{d}), \tag{3.65}$$

where I is the identity matrix of compatible dimension. Proceeding similarly, one can show that $M_\beta^{(k_0,k_1)}$ and $K_\beta^{(k_0,k_1)}$ (defined in a similar way as in (3.65)) can be computed by the following:

$$M_\beta^{(k_0,k_1)} = \left[I - \beta\mathcal{B}^{(k_0,k_1)}\right]^{-1} h^{(k_0,k_1)}, \tag{3.66}$$

$$K_\beta^{(k_0,k_1)} = \left[I - \beta\mathcal{B}^{(k_0,k_1)}\right]^{-1} \mathsf{1}^{(k_0,k_1)}, \tag{3.67}$$

where where I is the identity matrix of compatible dimension. $\mathsf{1}^{(k_0,k_1)} := [[\mathsf{1}^{(k_0)}]^\intercal, [\mathsf{1}^{(k_1)}]^\intercal]^\intercal$ and $\mathsf{1}^{(k_s)}$, $s \in \{0,1\}$ is given by:

$$\mathsf{1}^{(k_s)}(e,s) := \begin{cases} 0, & \text{if } e \in \mathcal{S}^{(k_s)} \\ 1, & \text{if } e \in \mathbb{X} \setminus \mathcal{S}^{k_s}. \end{cases}$$

3.8.2 Computation of $L_\beta^{(k_0,k_1)}$, $M_\beta^{(k_0,k_1)}$ and $K_\beta^{(k_0,k_1)}$ for $\mathbb{X} = \mathbb{R}$

The results for the countable Markov state process ($\mathbb{X} = \mathbb{Z}$) extend naturally for the uncountable case ($\mathbb{X} = \mathbb{R}$), where the process noise takes values in $\mathbb{R}$ and the optimal thresholds can be computed by solving Fredholm-like integral equations of second kind. For a generic threshold based strategy $f^{(k_0,k_1)}$, one can theoretically compute the performances $L_\beta^{(k_0,k_1)}$, $M_\beta^{(k_0,k_1)}$ and $K_\beta^{(k_0,k_1)}$ can be computed and using the renewal relationships in Proposition 3.8.1, one can compute $D_\beta^{(k_0,k_1)}$ and $N_\beta^{(k_0,k_1)}$. The optimal threshold can then be computed by a binary search. Although theoretically feasible to solve the Fredholm-like integral equations of second kind, it is not straightforward to solve them numerically because the integration kernel is discontinuous and the integration domain is $(-\infty, \infty)$. For this reason, we investigate an alternative computational approach in Chapter 4. The main idea behind our proposed solution is to replace the exact policy evaluation by a Monte Carlo based approximate policy evaluation and to replace the binary search for the optimal threshold by a SA iteration.

3.9 Results for i.i.d. packet drops in the channel

In this section we consider i.i.d. packet drops in the channel. We recognize that this can be expressed as a memoryless channel, which is a special case of the burst erasure channel with identical rows in the probability transition matrix Q. Let us denote $\mathbb{P}(S_t = 0 \,|\, S_{t-1} = s) = \mathbb{P}(S_t = 0) := \varepsilon$.

Evidently, the structural results given in Theorems 3.4.1, 3.4.2 and 3.7.1 hold for a communication channel with an i.i.d. erasure channel. As a consequence of the i.i.d. packet drops, the thresholds $k \in \mathbb{X}_{\geq 0}$ in Theorem 3.7.1 are scalar, independent of S_{t-1} and hence $k_0 = k_1$. For this reason, in the subsequent analysis, we substitute the threshold k_s, $s \in \{0, 1\}$, in the notations of all relevant parameters for the Markov erasure channel by k. Consequently, the state space of the controlled Markov process reduces to $\mathbb{X}$ from $\mathbb{X} \times \{0, 1\}$.

In this section we discuss in details the computational results for the integer state-process. Note that using Lemma 3.8.1 and Proposition 3.8.1, we can show the following:

Lemma 3.9.1 *For any $\beta \in (0, 1)$, $D^{(k)}_\beta(0)$ is increasing in k and $N^{(k)}_\beta(0)$ is strictly decreasing in k.*

See Appendix B.3 for the proof with $\mathbb{X} = \mathbb{Z}$. The proof for $\mathbb{X} = \mathbb{R}$ follows similarly.

When there is a constraint in the expected number of transmissions, $N_\beta(f, g)$, $\beta \in (0, 1]$, say by a given $\alpha \in (0, 1)$, then the *constrained communication* problem is stated as follows:

Problem 3.9.1 (Constrained communication) *In the model described in* (3.15) *, for a discount factor $\beta \in (0, 1]$ and a given $\alpha \in (0, 1)$, identify a communication strategy (f^*, g^*) that computes the following:*

$$D^*(f, g) \coloneqq \min_{N_\beta(f,g) \leq \alpha} D_\beta(f, g).$$

3.9.1 Optimal thresholds for costly and constrained communication for i.i.d. erasure channel for $\mathbb{X} = \mathbb{Z}$

Finally, we characterize the optimal strategies and optimal performances for Problems 3.7.1 and 3.9.1.

Definition 3.9.1 (Bernoulli randomized simple strategy) *Given two (non–randomized) time–homogeneous strategies f_1 and f_2 and a randomization parameter $\theta \in (0, 1)$, the* Bernoulli randomized strategy (f_1, f_2, θ) *is a strategy that randomizes between f_1 and f_2 at each stage; choosing f_1 with probability θ and f_2 with probability $(1 - \theta)$. Such a strategy is called a Bernoulli randomized* simple *strategy if f_1 and f_2 differ on exactly one state i.e. there exists a state e_0 such that for all $e \neq e_0$, $f_1(e) = f_2(e)$.*

The next two theorems characterize the performances for costly and constrained communication for infinite-horizon setup under the optimal communication strategies as given by Theorem 3.7.1.

Theorem 3.9.1 *(Characterization of optimal costly performance) For $\beta \in (0,1]$, let $\mathbb{K}$ denote $\{k \in \mathbb{Z}_{\geq 0} : D_\beta^{(k+1)}(0) > D_\beta^{(k)}(0)\}$. For $k_n \in \mathbb{K}$, define:*

$$\lambda_\beta^{(k_n)} := \frac{D_\beta^{(k_n+1)}(0) - D_\beta^{(k_n)}(0)}{N_\beta^{(k_n)}(0) - N_\beta^{(k_n+1)}(0)}. \tag{3.68}$$

Then, we have the following.

1. *For any $k_n \in \mathbb{K}$ and any $\lambda \in (\lambda_\beta^{(k_{n-1})}, \lambda_\beta^{(k_n)}]$, the strategy $f^{(k_n)}$ is optimal for Problem 3.7.1 with communication cost λ.*

2. *The optimal performance $C_\beta^*(\lambda)$ is continuous, concave, increasing and piecewise linear in λ. The corner points of $C_\beta^*(\lambda)$ are given by $\{(\lambda_\beta^{(k_n)}, D_\beta^{(k_n)}(0) + \lambda_\beta^{(k_n)} N_\beta^{(k_n)}(0))\}_{k_n \in \mathbb{K}}$ (see Fig. 3.2).*

See Appendix B.4 for the proof.

Theorem 3.9.2 *(Characterization of optimal constrained performance) For any $\beta \in (0,1)$ and $\alpha \in (0,1)$, define*

$$k_\beta^*(\alpha) = \sup\left\{k \in \mathbb{Z}_{\geq 0} : \alpha M_\beta^{(k)} \geq K_\beta^{(k)}\right\} \tag{3.69}$$

$$\theta_\beta^*(\alpha) = \frac{M_\beta^{(k^*)}\left(\alpha M_\beta^{(k^*+1)} - K_\beta^{(k^*+1)}\right)}{K_\beta^{(k^*)} M_\beta^{(k^*+1)} - K_\beta^{(k^*+1)} M_\beta^{(k^*)}}. \tag{3.70}$$

For ease of notation, we use $k^* = k_\beta^*(\alpha)$ and $\theta^* = \theta_\beta^*(\alpha)$.

Let f^* be the Bernoulli randomized simple strategy $(f^{(k^*)}, f^{(k^*+1)}, \theta^*)$, i.e.,

$$f^*(e) = \begin{cases} 0, & \text{if } |e| < k^*; \\ 0, & \text{w.p. } 1-\theta^*, \text{ if } |e| = k^*; \\ 1, & \text{w.p. } \theta^*, \text{ if } |e| = k^*; \\ 1, & \text{if } |e| > k^*. \end{cases} \tag{3.71}$$

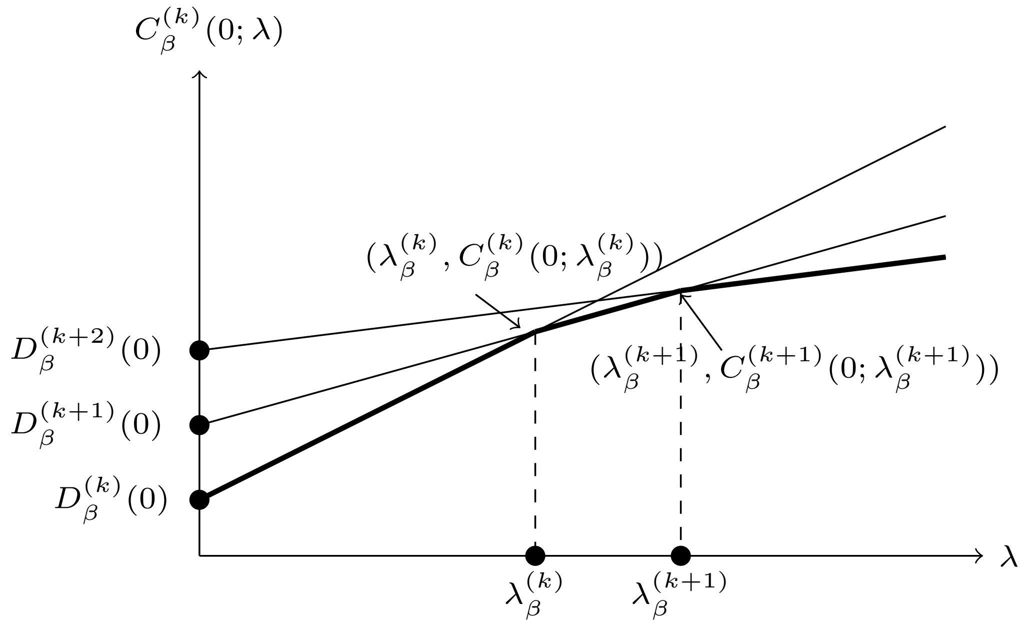


Fig. 3.2 The optimal costly performance as a function of λ.

Then,

1. (f^*, g^*) *is optimal for Problem 3.9.1 with constraint* α.
2. *Let* $\alpha^{(k)} = N_\beta(f^{(k)}, g^*)$. *Then, for* $\alpha \in (\alpha^{(k+1)}, \alpha^{(k)})$, $k^* = k$ *and* $\theta^* = (\alpha - \alpha^{(k+1)})/(\alpha^{(k)} - \alpha^{(k+1)})$, *and the distortion-transmission function is given by*

$$D_\beta^*(\alpha) = \theta^* D_\beta^{(k)} + (1-\theta^*) D_\beta^{(k+1)}. \tag{3.72}$$

Moreover, the distortion-transmission function is continuous, convex, decreasing and piecewise linear in α. *Thus, the corner points of* $D_\beta^*(\alpha)$ *are given by* $\{(N_\beta^{(k)}(0), D_\beta^{(k)}(0))\}_{k=1}^\infty$ *(see Fig. 3.3).*

The proof is similar to that of Theorem 2.4.4 and hence is omitted here.

3.9.2 An example: symmetric birth-death Markov chain with i.i.d. erasure channel

In this section, we verify with a numerical example the main results for Problem 3.9.1 and analyze the variation of the distortion-transmission function with the packet-drop

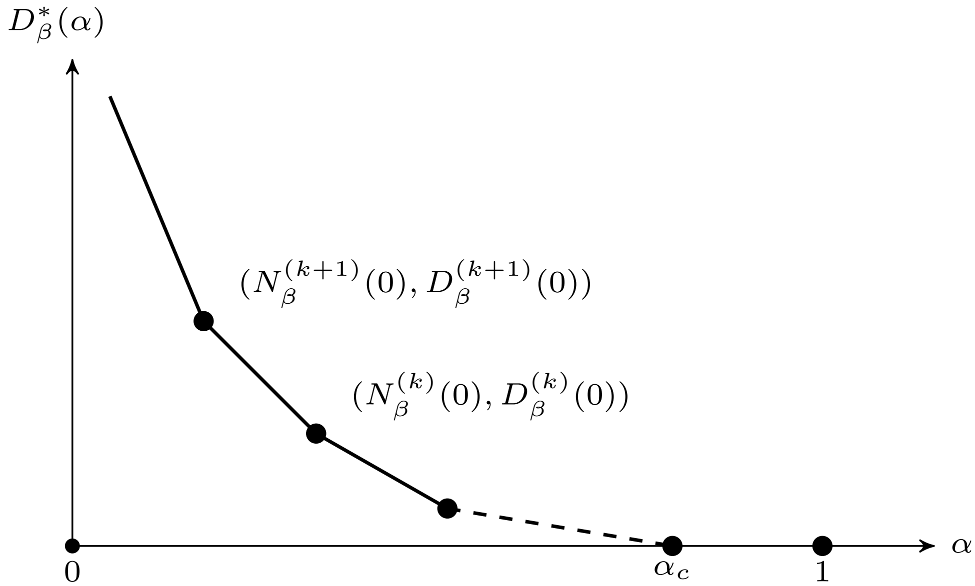


Fig. 3.3 $D^*_\beta(\alpha)$ as a function of α.

probability, ε. Consider an aperiodic, symmetric, birth-death Markov chain defined over $\mathbb{Z}$ with the transition probability matrix as given by:

$$P_{ij} = \begin{cases} p, & \text{if } |i-j| = 1; \\ 1-2p, & \text{if } i = j; \\ 0, & \text{otherwise,} \end{cases}$$

where we assume that $p \in (0, \frac{1}{3})$. Let the distortion function be $d(e) = |e|$. The model satisfies (3.15) with $a = 1$. We verify the main results for $p = 0.3$, $\beta = 0.9$. Fig. 3.4 shows the distortion-transmission function as a function of α for $\varepsilon \in \{0, 0.3, 0.7\}$. We see from the plots that the optimal distortion increases with increase in the value of ε, which is in consistent with the intuition.

3.9.3 Optimal thresholds for costly and constrained communication for i.i.d. erasure channel for $\mathbb{X} = \mathbb{R}$

The following two theorems characterize the optimal performances under costly and constrained communication, the results of which essentially rely on the monotonicity properties

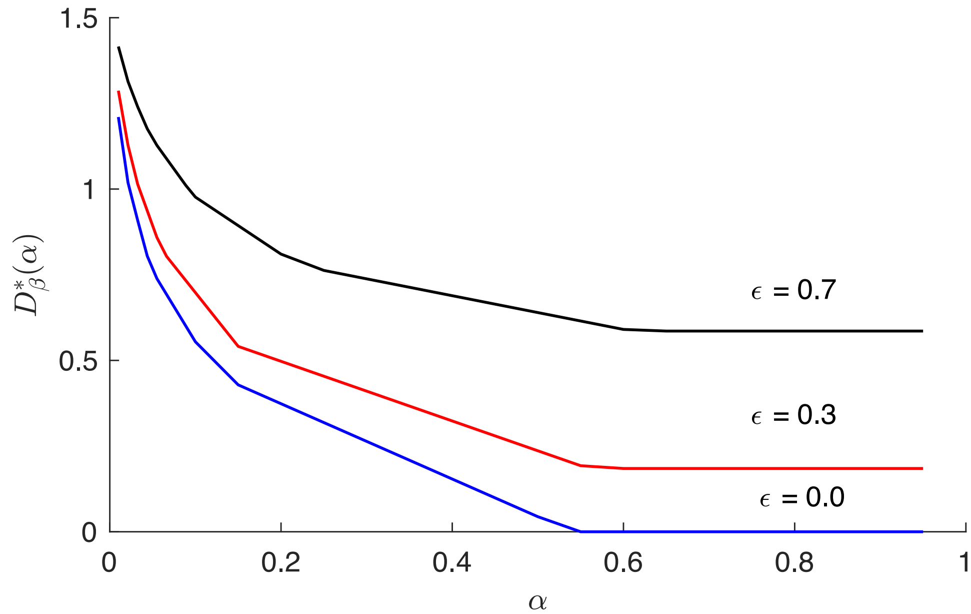


Fig. 3.4 Plots of $D^*_\beta(\alpha)$ versus α, for $\beta = 0.9$ and $\varepsilon \in \{0, 0.3, 0.7\}$.

of $D^{(k)}_\beta$, $N^{(k)}_\beta$ and $C^{(k)}_\beta$, which follow from Lemma 3.8.1 and Proposition 3.8.1. The proof is similar to what is given in Chapter 2 and Appendix B.1 ([13, 14]).

Let $\partial_k D^{(k)}_\beta$, $\partial_k N^{(k)}_\beta$ and $\partial_k C^{(k)}_\beta$ denote the derivative[5] of $D^{(k)}_\beta$, $N^{(k)}_\beta$ and $C^{(k)}_\beta$ with respect to k.

The following two theorems characterize the optimal performance for Problems 3.7.1 and 3.9.1. The proof is similar to [14].

Theorem 3.9.3 *For $\beta \in (0, 1]$, we have the following.*

1. *If the pair (λ, k) satisfies the following*

$$\lambda \partial_k N^{(k)}_\beta(0) + \partial_k D^{(k)}_\beta(0) = 0, \tag{3.73}$$

then, the strategy $(f^{(k)}, g^*)$ is optimal for Problem 3.7.1 with communication cost λ. Furthermore, for any $k > 0$, there exists a $\lambda \geq 0$ that satisfies (3.73).

[5] Following the argument given for Lemma 2.4.1, one can show that $D^{(k)}_\beta$, $N^{(k)}_\beta$ and $C^{(k)}_\beta$ are differentiable in k.

2. *The optimal performance $C^*_\beta(\lambda)$ is continuous, concave and increasing function of λ.*

Theorem 3.9.4 *For any $\beta \in (0,1]$ and $\alpha \in (0,1)$, let $k^*_\beta(\alpha) \in \mathbb{R}_{\geq 0}$ be such that*

$$N_\beta^{(k^*_\beta(\alpha))}(0) = \alpha. \tag{3.74}$$

Such a $k^*_\beta(\alpha)$ exists and we have the following:

1. *The strategy $(f^{(k^*_\beta(\alpha))}, g^*)$ is optimal for Problem 3.9.1 with constraint α.*

2. *The distortion-transmission function $D^*_\beta(\alpha)$ is continuous, convex and decreasing in α and is given by*
$$D^*_\beta(\alpha) = D_\beta^{(k^*_\beta(\alpha))}(0). \tag{3.75}$$

As we mentioned in the case for Markov erasure channel, although the analytical computation of $L^{(k)}_\beta$, $M^{(k)}_\beta$ and $K^{(k)}_\beta$ requires solving the Fredholm-like integration of second kind, here too the computation becomes challenging due to infinite limits of integration and the discontinuity of the integrand kernel. So, we investigate the applicability of the simulation-based approaches to find the optimal thresholds and the optimal performances.

3.9.4 A special case for $\mathbb{X} = \mathbb{R}$: Gaussian process noise: scaling with variance

In this section, similar to the ideal channel as described in [14], we derive the scaling of the optimal threshold and the optimal performance with the variance of the process noise.

First we state the following lemma:

Lemma 3.9.2 *For Gauss-Markov model (a special case of the first-order autoregressive model with $\mathbb{X} = \mathbb{R}$), let $L^{(k)}_\sigma$, $M^{(k)}_\sigma$ and $K^{(k)}_\sigma$ be the solutions of* (3.60), (3.61) *and* (3.62) *respectively, when the variance of W_t is σ^2. Let the per-step distortion is given by $d(e) = e^2$ for all $e \in \mathbb{R}$. Then*

$$L^{(k)}_\sigma(e) = \sigma^2 L^{(k/\sigma)}_1\Big(\frac{e}{\sigma}\Big), \quad M^{(k)}_\sigma(e) = M^{(k/\sigma)}_1\Big(\frac{e}{\sigma}\Big), \tag{3.76}$$
$$K^{(k)}_\sigma(e) = K^{(k/\sigma)}_1\Big(\frac{e}{\sigma}\Big),$$
$$D^{(k)}_\sigma(e) = \sigma^2 D^{(k/\sigma)}_1\Big(\frac{e}{\sigma}\Big), \quad N^{(k)}_\sigma(e) = N^{(k/\sigma)}_1\Big(\frac{e}{\sigma}\Big). \tag{3.77}$$

The proof is given in Appendix B.5.

Define $k^*(\lambda) = \arg\min_k C^{(k)}_\beta(0; \lambda)$ and $k^*(\alpha) = \arg\min_{k : N^{(k)}_\beta(0) \le \alpha} D^{(k)}_\beta(0)$. Then,

Theorem 3.9.5 *For the Gauss-Markov model for Problems 3.3.1–3.7.1, $k^*_\sigma(\lambda) = k^*_1(\lambda/a^2\sigma^2)$ and $C^*_\sigma(\lambda) = \sigma^2 C^*_1(\lambda/\sigma^2)$. For Problem 3.9.1, $k^*_\sigma(\alpha) = \sigma k^*_1(\alpha)$ and $D^*_\sigma(\alpha) = \sigma^2 D^*_1(\alpha)$.*

Proof The theorem follows from Lemma 3.9.2, Proposition 3.8.1 and elementary algebra.

3.10 Conclusion

In this chapter, we studied remote estimation over a Gilbert-Elliott channel with feedback. We assume that the channel state is observed by the receiver and fed back to the transmitter with one unit delay. In addition, the transmitter gets ACK/NACK feedback for successful/unsuccessful transmission. Using ideas from team theory, we establish the structure of optimal transmission and estimation strategies and identify a dynamic program to determine optimal strategies with that structure. We then consider first-order autoregressive sources where the noise process has unimodal and symmetric distribution. Using ideas from majorization theory, we show that the optimal transmission strategy has a threshold structure and the optimal estimation strategy is Kalman-like. Furthermore, we characterize the optimal *costly* and *constrained* performances for a first-order autoregressive model with i.i.d. packet drop, which is a special case of the Gilbert-Elliott channel.

Chapter 4

Stochastic approximation based approaches to compute optimal thresholds in remote estimation

4.1 Introduction

Previous chapters discuss the stochastic dynamic programming formulation to find the globally optimal solution of a team problem with non-classical information structure. In contrast to that, in this chapter, we employ two types of Stochastic Approximation (SA) approaches to find the optimal thresholds of the RE problem introduced in the previous chapters. For the costly communication, we focus on estimating the local minima using one type of SA methods, called the Stochastic Gradient (SG) method to find the optimal solution. For the constrained communication, we use fixed point iteration method to find the optimal solution. In the context of RE, we discuss the SA approaches to find the optimal solution.

Although the SA algorithms typically converge to locally optimal solution, they find their relevance in practice due to certain advantages over the stochastic dynamic programming formulation. First, SA can be implemented without the knowledge of the probability transition function of an Markov Decision Processes (MDP), whereas the dynamic programming formulation needs the complete knowledge of it. This in fact lays the ground for a field of *learning*-based treatments of stochastic optimization, called *Reinforcement Learn-*

ing. Second, often in scenarios where the state and action spaces of a MDP are large, the stochastic dynamic programming proves to be quite expensive, whereas the SA approaches can be comparatively much cheaper by dint of proper choice of lower-dimensional *features*. A term abundantly used in the field of *Machine Learning*, the features are essentially parameterization of the control strategy, which appear in formulating the *Approximate Dynamic Program*.

4.2 Motivation for applying SA approaches to RE problem

Recall the remote estimation problem with packet drops in the communication channel, as discussed in Chapter 3. It is shown that the optimal estimator is Kalman-like and the optimal transmitter is threshold based. In this chapter we consider the infinite horizon optimization problem, for which the thresholds are time-homogeneous. Furthermore, we restrict our discussion to i.i.d. erasure channels (with packet drop probability ε). As is explained in last chapter, the task to characterize the optimal strategy and optimal performances follows two steps: 1) compute the performance of a generic threshold-based strategy and 2) find the optimal threshold, which leads to the optimal performance. The first step is typically called the *policy evaluation* and the second step is called the *policy improvement*.

Exact policy evaluation: Given a policy $f^{(k)}$, compute $L_\beta^{(k)}(0)$, $M_\beta^{(k)}(0)$ and $K_\beta^{(k)}(0)$ by solving (3.60), (3.61), and (3.62), and compute the performances $D_\beta^{(k)}(f^{(k)}, g^*)$, $N_\beta^{(k)}(f^{(k)}, g^*)$, and $C_\beta^{(k)}(f^{(k)}, g^*; \lambda)$ using the expressions in Proposition 3.8.1.

According to Proposition 3.8.1, computing $L_\beta^{(k)}(0)$, $K_\beta^{(k)}(0)$ and $M_\beta^{(k)}(0)$ is sufficient to compute $D_\beta^{(k)}(0)$ and $N_\beta^{(k)}(0)$ (and therefore, compute the performance of strategy $f^{(k)}$). In Chapter 2, which considers the case of no packet drops (i.e., $\varepsilon = 0$), $L_\beta^{(k)}(0)$ and $M_\beta^{(k)}(0)$ were computed by solving the balance equations for the truncated Markov chain. These balance equations corresponded to Fredholm integral equations of the second kind. Using this exact policy evaluation, the optimal thresholds were identified by a binary search over k.

When $\varepsilon \neq 0$, the balance equations for the truncated Markov process still correspond to Fredholm integral equations of the second kind, but it is not straightforward to solve them numerically because the integration kernel is discontinuous and the integration domain is $(-\infty, \infty)$. For this reason, we investigate an alternative computational approach. The main idea behind our proposed solution is to replace the exact policy evaluation by a Monte

Carlo based approximate policy evaluation and to replace the binary search for the optimal threshold by a stochastic approximation iteration. In particular, we use KW algorithm [63] and SF [64] to solve (3.73) and RM algorithm [65] to solve (3.74). The details are presented in the next section.

Using the results of Proposition 3.8.1, it is possible to evaluate the performance of any strategy as follows. Using an approach similar to that introduced in Chapter 2, when there are no packet drops (i.e., $\varepsilon = 0$), one can derive the following standard Fredholm integral equations of the second kind: for $e \in [-k, k]$,

$$L_\beta^{(k)}(e) = d(e) + \beta \int_{-k}^{k} \mu(w - ae) L_\beta^{(k)}(w) dw,$$
$$M_\beta^{(k)}(e) = 1 + \beta \int_{-k}^{k} \mu(w - ae) M_\beta^{(k)}(w) dw,$$
$$K_\beta^{(k)}(e) = \beta \int_{-k}^{k} \mu(w - ae) K_\beta^{(k)}(w) dw.$$

This relationship was exploited in Chapter 2 to propose the following algorithms to compute optimal thresholds for both costly and constrained communication.

1. For costly communication, for a given strategy $f^{(k)}$, $\partial_k D_\beta^{(k)}(0)/\partial_k N_\beta^{(k)}(0)$ is approximated as $\big(D_\beta^{(k+\delta)}(0) - D_\beta^{(k)}(0)\big)/\big(N_\beta^{(k+\delta)}(0) - N_\beta^{(k)}(0)\big)$, where δ is a small number. Each term in the above expression is computed using exact policy evaluation. Then, an δ-optimal strategy is obtained by using binary search to identify a threshold k that satisfies (3.73).

2. For constrained communication, for a given strategy $f^{(k)}$, $N_\beta^{(k)}(0)$ is computed using exact policy evaluation. Then, an δ-optimal strategy is obtained by using binary search to identify a threshold k that satisfies (3.74).

In this chapter we propose stochastic approximation based algorithms to compute the optimal thresholds for both costly and constrained communication. Our motivation for considering an alternative computational approach is two-fold.

First, when $\varepsilon \neq 0$, Eqs. (3.60) reduces to the following (the expression for (3.61) is similar): for $e \in \mathbb{R}$,

$$L_\beta^{(k)}(e) = h^{(k)}(e) d(e) + \beta \int_{-\infty}^{\infty} \mu_1(w - ae) h^{(k)}(e) L_\beta^{(k)}(w) dw,$$

where $h^{(k)}(e)$ is 1 for $e \in (-k, k)$ and is ε otherwise. Although, this is a Fredholm integral equation of the second kind, it is not straightforward to solve it numerically because the kernel $\mu_1(w - ae)h^{(k)}(e)$ is discontinuous and the domain is $(-\infty, \infty)$.

Second, the idea of solving the Fredholm integral equations does not scale to higher dimensions. Although, it is not known whether or not the structural results of Theorem 3.7.1 hold when the source is a vector Gauss-Markov process, yet it is appealing to use threshold based transmission strategies due to their simplicity. Solving a multi-dimensional Fredholm integral equation involves iteratively solving multi-dimensional integrals, and suffers from the usual curse of dimensionality. Furthermore, even if the Fredholm integral equation is solved approximately, searching for thresholds satisfying (3.73) or (3.74) in higher dimensions is more complicated than the one dimensional binary search.

For these reasons, we investigate alternative computational approaches. The main idea behind our proposed solution is to replace the exact policy evaluation by a Monte Carlo based approximate policy evaluation and to replace the binary search for the optimal threshold by a stochastic approximation iteration. In particular, we use KW algorithm [63] and SF algorithm [64] to solve (3.73) and RM algorithm [65] to solve (3.74). The details are presented in the next section.

4.3 Stochastic approximation algorithms

4.3.1 Noisy policy evaluation

The first step to develop a stochastic approximation algorithm to identify the optimal thresholds is to replace the exact policy evaluation by an approximate policy evaluation. The simplest way to do so is to use sample path average. In particular, let $\{(E_t^{(k)}, U_t^{(k)})\}_{t\geq 0}$ denote the sample paths of the error process and the transmission process under policy $f^{(k)}$ and T be a large number. Then, $D_\beta^{(k)}(0) \approx (1-\beta)\sum_{t=0}^{T} \beta^t d(E_t^{(k)})$, $N_\beta^{(k)}(0) \approx (1-\beta)\sum_{t=0}^{T} \beta^t U_t^{(k)}$, and $C_\beta^{(k)}(0;\lambda) = D_\beta^{(k)}(0) + \lambda N_\beta^{(k)}(0)$.

For the discounted case, using naive approach leads to numerical difficulties as one needs to compute β^t for large t, which makes the term very small. To circumvent this, we estimate $L_\beta^{(k)}(0)$, $M_\beta^{(k)}(0)$ and $K_\beta^{(k)}(0)$ by Monte Carlo evaluations until renewal instance and then use the renewal relationship of Proposition 3.8.1 to approximate $D_\beta^{(k)}(0)$ and $N_\beta^{(k)}(0)$. We call this method the Renewal Monte Carlo (RMC). It is perhaps worth mentioning here that RMC is a low-bias method since it is inherently a Monte Carlo method, which uses no intial guess of the *value-action function* and at the same time it is a lower-variance method compared to a *naive* Monte Carlo method as the length of the *episodes* are shorter due to renewal.

The Monte Carlo evaluations are done by averaging over K episodes. In each episode, the error process starts at $E_0 = 0$ and evolves under strategy $f_t^{(k)}$. The episode ends at the stopping time $\tau^{(k)}$ of the first successful reception. Let $\{(E_{n,t}^{(k)}, U_{n,t}^{(k)})\}_{t\geq 0}$ denote the sample path of the error process and the transmission process in episode n. Then,

$$L_\beta^{(k)}(0) \approx \frac{1}{\mathsf{K}} \sum_{n=1}^{\mathsf{K}} \sum_{t=0}^{\tau^{(k)}-1} \beta^t d(E_{n,t}^{(k)}), \tag{4.1}$$

$$M_\beta^{(k)}(0) \approx \frac{1}{\mathsf{K}} \sum_{n=1}^{\mathsf{K}} \sum_{t=0}^{\tau^{(k)}-1} \beta^t, \tag{4.2}$$

$$K_\beta^{(k)}(0) \approx \frac{1}{\mathsf{K}} \sum_{n=1}^{\mathsf{K}} \sum_{t=0}^{\tau^{(k)}} \beta^t U_{n,t}^{(k)}. \tag{4.3}$$

Then, $D_\beta^{(k)}(0)$, $N_\beta^{(k)}(0)$, and $C_\beta^{(k)}(0;\lambda)$ can be computed using the expressions in Proposition 3.8.1. The complete details for this evaluation are shown in Algorithm 3.

Let $\hat{L}_\beta^{(k,\mathsf{K})}$, $\hat{M}_\beta^{(k,\mathsf{K})}$ and $\hat{K}_\beta^{(k,\mathsf{K})}$ denote the right hand sides of (4.1), (4.2) and (4.3). Then, (4.1)–(4.3) can be written as

$$\begin{aligned} &L_\beta^{(k)}(0) = \hat{L}_\beta^{(k,\mathsf{K})} + \xi_{\mathsf{K}}^L, \quad M_\beta^{(k)}(0) = \hat{M}_\beta^{(k,\mathsf{K})} + \xi_{\mathsf{K}}^M, \\ &K_\beta^{(k)}(0) = \hat{K}_\beta^{(k,\mathsf{K})} + \xi_{\mathsf{K}}^K, \end{aligned}$$

where ξ_{K}^L, ξ_{K}^M and ξ_{K}^K are approximation errors that go to zero as K goes to infinity. Define estimates $\hat{D}_\beta^{(k,\mathsf{K})}$, $\hat{N}_\beta^{(k,\mathsf{K})}$, and $\hat{C}_\beta^{(k,\mathsf{K})}(\lambda)$ for $D_\beta^{(k)}(0)$, $N_\beta^{(k)}(0)$, and $C_\beta^{(k)}(0;\lambda)$ in terms of $\hat{L}_\beta^{(k)}$, $\hat{M}_\beta^{(k)}$ and $\hat{K}_\beta^{(k)}$ using renewal expressions given in Proposition 3.8.1.

Note that the stochastic approximation algorithms that we describe next work under mild assumptions on ξ_{K}^L, ξ_{K}^M and ξ_{K}^K. Therefore, the number K of episodes need not be large. In our experiments that we report later, we choose K as 1000.

Algorithm 3: Algorithm for noisy policy evaluation

---

function *MonteCarloEvaluation*(k, K)
  input : Threshold $k \in \mathbb{R}_{>0}$
    Number of episodes $\mathsf{K} \in \mathbb{Z}_{>0}$
  output : Estimate $\hat{L}_\beta^{(k,\mathsf{K})}$ of $L_\beta^{(k)}(0)$
    Estimate $\hat{M}_\beta^{(k,\mathsf{K})}$ of $M_\beta^{(k)}(0)$
    Estimate $\hat{K}_\beta^{(k,\mathsf{K})}$ of $K_\beta^{(k)}(0)$
  initialize: $\hat{L} = 0,\ \hat{M} = 0,\ \hat{K} = 0$
  for *iteration* $i = 1$ *upto* K **do**
    Set $t = 0,\ \ell = 0,\ m = 0,\ \Bbbk = 0,\ E_0 = 0$
    while *true* **do**
      $S_{t+1} \sim \text{Bernoulli}(\varepsilon)$
      if $|E_t| < k$ *or* $S_{t+1} = 0$ **then**
        $\ell \leftarrow \ell + \beta^t E_t^2$
        $m \leftarrow m + \beta^t$
        $\Bbbk \leftarrow \Bbbk + \beta^t \mathbb{1}_{\{|E_t| \geq k\}}$
      else
        $\Bbbk \leftarrow \Bbbk + \beta^t$
        break
      $E_{t+1} = aE_t + W_t$, where $W_t \sim \mathcal{N}(0, \sigma^2)$
      $t \leftarrow t + 1$
    $\hat{L} \leftarrow \hat{L} + \ell,\ \hat{M} \leftarrow \hat{M} + m,\ \hat{K} \leftarrow \hat{K} + \Bbbk$
  return $(\hat{L}/\mathsf{K}, \hat{M}/\mathsf{K}, \hat{K}/\mathsf{K})$

4.3.2 Computing thresholds for costly communication using stochastic approximation

In our subsequent discussion, we assume the following:

(A1) $C_\beta^{(k)}(0; \lambda)$ is convex in k.

(A2) $\mathbb{E}[\mathsf{C}_\beta^{(k,\mathsf{K})}(\lambda)] = C_\beta^{(k)}(0; \lambda)$.

We verified through simulation that (A1) holds. (A2) holds if $\mathsf{C}_\beta^{(k,\mathsf{K})}(\lambda)$ is an unbiased estimator of $C_\beta^{(k)}(0; \lambda)$, which we verified through simulations.

According to Theorem 3.9.3, a threshold k is optimal if $\partial_k C_\beta^{(k)}(0; \lambda) = 0$. Using Algorithm 3, we can obtain a noisy "measurement" $\hat{C}_\beta^{(k,\mathsf{K})}(\lambda)$ of $C_\beta^{(k)}(0; \lambda)$. Using this noisy

Algorithm 4: Algorithm for costly communication using KW algorithm

input : Initial guess $\mathsf{k}_{\mathsf{init}} \in \mathbb{R}_{>0}$;
Number of episodes $\mathsf{K} \in \mathbb{Z}_{>0}$
Number of iterations $\mathsf{N}_{\mathsf{iterations}} \in \mathbb{Z}_{>0}$
output : Optimal threshold k°
initialize: $k^\circ = \mathsf{k}_{\mathsf{init}}$
for *iteration* $i = 1$ *upto* $\mathsf{N}_{\mathsf{iterations}}$ **do**
  Pick δ as a small non-negative real
  $k^\circ_+ \leftarrow k^\circ + \delta$
  $k^\circ_- \leftarrow k^\circ - \delta$
  $[\hat{L}_+, \hat{M}_+, \hat{K}_+] = \text{MonteCarloEvaluation}\big(k^\circ_+, \mathsf{K}\big)$
  $[\hat{L}_-, \hat{M}_-, \hat{K}_-] = \text{MonteCarloEvaluation}\big(k^\circ_-, \mathsf{K}\big)$
  Compute C_+, C_- using Proposition 3.8.1
  $\partial_k \mathsf{C} \leftarrow (\mathsf{C}_+ - \mathsf{C}_-)/2\delta$
  Compute γ_i using Adaptive Moments (AdaM) [66]
  $k^\circ \leftarrow k^\circ - \gamma_i \partial_k \mathsf{C}$
return k°

measurement, it is possible to search for the optimal threshold using the KW algorithm [63], which is a first-order stochastic gradient descent algorithm. In addition, as an alternative to the KW algorithm, we also apply SF algorithm [64], which we discuss in Section 4.4.

The KW algorithm works as follows. We start with an initial guess k°_0 of the optimal threshold. Let k°_i denote our guess at the beginning of iteration i. During iteration i, we obtain a noisy measurement of the gradient $\partial_k C^{(k)}_\beta(0;\lambda)$ using the finite difference $\Delta_i^{(k^\circ_i,\mathsf{K})} = \hat{C}_\beta^{(k^\circ_i+\delta,\mathsf{K})}(\lambda) - \hat{C}_\beta^{(k^\circ_i-\delta,\mathsf{K})}(\lambda)$ and update our guess as follows:

$$k^\circ_{i+1} = k^\circ_i - \gamma_i \Delta_i^{(k^\circ_i,\mathsf{K})}, \tag{4.4}$$

where γ_i are learning rates that satisfy $\sum_{i=1}^\infty \gamma_i = \infty$ and $\sum_{i=1}^\infty \gamma_i^2 < \infty$. See Algorithm 4 for details.

Theorem 4.3.1 *Under assumptions (A1)–(A2), the threshold iterates k°_i of Algorithm 4 converge almost surely to the optimal threshold, i.e.,* $\lim_{i\to\infty} k^\circ_i = k^*(\lambda)$, *a.s., where* $k^*(\lambda)$ *is optimal threshold for Problem 3.7.1.*

The proof follows immediately from [63].

Table 4.1 Comparative results for costly communication using SA (KW) and FIE for $a = 1$ and $\varepsilon = 0$.

(a) $\beta = 0.9$

λ	Threshold k^* SA	FIE	Error (Absolute)	Performance $C^*_\beta(\lambda)$ SA	FIE	Error (Absolute)
100	4.9355	4.9298	5.7×10^{-3}	5.2511	5.2510	1.1×10^{-4}
200	6.3221	6.3086	1.4×10^{-2}	6.5221	6.5219	2.3×10^{-5}
300	7.3421	7.3289	1.3×10^{-2}	7.2208	7.2205	3.0×10^{-4}
400	8.2118	8.1764	3.5×10^{-2}	7.6654	7.6652	1.5×10^{-4}
500	8.9469	8.9177	2.9×10^{-2}	7.9700	7.9699	9.2×10^{-5}
600	9.5830	9.5854	2.5×10^{-3}	8.1886	8.1886	2.5×10^{-5}
700	10.0803	10.1984	1.2×10^{-1}	8.3515	8.3507	8.3×10^{-4}

(b) $\beta = 1.0$

λ	Threshold k^* SA	FIE	Error (Absolute)	Performance $C^*_\beta(\lambda)$ SA	FIE	Error (Absolute)
100	4.3438	4.3446	7.9×10^{-4}	7.8540	7.8539	1.2×10^{-4}
200	5.283	5.2841	8.3×10^{-4}	11.2327	11.2324	3.0×10^{-4}
300	5.9340	5.9136	2.0×10^{-2}	13.8265	13.8257	7.8×10^{-4}
400	6.4079	6.4004	7.5×10^{-3}	16.0131	16.0124	7.6×10^{-4}
500	6.8028	6.8028	4.4×10^{-5}	17.9399	17.9390	9.5×10^{-4}
600	7.1487	7.1485	1.1×10^{-4}	19.6810	19.6809	6.7×10^{-5}
700	7.4569	7.4534	3.5×10^{-3}	21.2829	21.2828	8.0×10^{-5}

Table 4.2 Comparative results for constrained communication using SA (RM) and FIE for $a = 1$ and $\varepsilon = 0$.

(a) $\beta = 0.9$

α	Threshold k^* SA	FIE	Error (Absolute)	Performance $D^*_\beta(\alpha)$ SA	FIE	Error (Absolute)
0.1	2.2230	2.2217	1.3×10^{-3}	0.9293	0.9283	9.9×10^{-4}
0.2	1.4416	1.4404	1.2×10^{-3}	0.3954	0.3947	7.0×10^{-4}
0.3	1.0586	1.0620	3.4×10^{-3}	0.1974	0.1989	1.5×10^{-3}
0.4	0.8014	0.8057	4.3×10^{-3}	0.0989	0.1003	1.4×10^{-3}
0.5	0.6017	0.5981	3.5×10^{-3}	0.0460	0.0453	7.4×10^{-4}
0.6	0.4357	0.4395	3.7×10^{-3}	0.0186	0.0190	4.6×10^{-4}
0.7	0.2823	0.2808	1.5×10^{-3}	0.0052	0.0052	8.1×10^{-5}
0.8	0.1396	0.1465	6.8×10^{-3}	0.0006	0.0007	9.9×10^{-5}

(b) $\beta = 1.0$

α	Threshold k^* SA	FIE	Error (Absolute)	Performance $D^*_\beta(\alpha)$ SA	FIE	Error (Absolute)
0.1	2.5396	2.5391	5.7×10^{-4}	1.3677	1.3671	5.8×10^{-4}
0.2	1.6020	1.5991	2.9×10^{-3}	0.5485	0.5464	2.0×10^{-3}
0.3	1.1713	1.1719	6.2×10^{-4}	0.2767	0.2770	3.4×10^{-4}
0.4	0.9014	0.9033	1.9×10^{-3}	0.1477	0.1485	7.7×10^{-4}
0.5	0.6994	0.6958	3.6×10^{-3}	0.0767	0.0756	1.0×10^{-3}
0.6	0.5334	0.5371	3.7×10^{-3}	0.0365	0.0373	7.2×10^{-4}
0.7	0.3884	0.3906	2.2×10^{-3}	0.0148	0.1500	2.4×10^{-4}
0.8	0.2540	0.2563	2.3×10^{-3}	0.0043	0.0044	1.2×10^{-4}

Table 4.3 Comparative results for costly communication using SA (SF) and FIE for $a = 1$ and $\varepsilon = 0$.

(a) $\beta = 0.9$

λ	Threshold k^* SA	FIE	Error (Absolute)	Performance $C^*_\beta(\lambda)$ SA	FIE	Error (Absolute)
100	4.9318	4.9298	2.01×10^{-3}	5.2511	5.2510	9.8×10^{-5}
200	6.3074	6.3086	1.19×10^{-3}	6.5221	6.5219	2.0×10^{-4}
300	7.3200	7.3289	8.89×10^{-3}	7.2208	7.2205	2.8×10^{-4}
400	8.1858	8.1764	9.4×10^{-3}	7.6652	7.6652	2.8×10^{-5}
500	8.9218	8.9177	4.14×10^{-3}	7.9700	7.9699	2.2×10^{-5}
600	9.5868	9.5854	1.45×10^{-3}	8.1886	8.1886	2.5×10^{-5}
700	10.1753	10.1984	2.31×10^{-2}	8.3507	8.3507	5.6×10^{-5}

(b) $\beta = 1.0$

λ	Threshold k^* SA	FIE	Error (Absolute)	Performance $C^*_\beta(\lambda)$ SA	FIE	Error (Absolute)
100	4.3558	4.3446	1.1×10^{-2}	7.8541	7.8539	2.0×10^{-4}
200	5.2846	5.2841	4.9×10^{-4}	11.2327	11.2324	3.0×10^{-4}
300	5.9204	5.9136	6.8×10^{-3}	13.8263	13.8257	5.4×10^{-4}
400	6.4003	6.4004	1.1×10^{-4}	16.0131	16.0124	7.2×10^{-4}
500	6.8113	6.8028	8.5×10^{-3}	17.9400	17.9390	1.0×10^{-3}
600	7.1531	7.1485	4.6×10^{-3}	19.6810	19.6809	8.1×10^{-5}
700	7.4513	7.4534	2.1×10^{-3}	21.2829	21.2828	8.4×10^{-5}

The rate of convergence of the KW algorithm is sensitive to the choice of learning rates. We use AdaM algorithm [66] to adaptively tune the learning rate based on the "measurements" $\Delta_i^{(k_i^\circ, \mathsf{K})}$.

4.3.3 Computing thresholds for constrained communication using stochastic approximation

First, we note the following facts:

(F1) $\mathsf{M}_\beta^{(k,\mathsf{K})}$ is increasing with k and $\mathsf{K}_\beta^{(k,\mathsf{K})}$ is decreasing with k.

(F2) $\mathbb{E}[\mathsf{M}_\beta^{(k,\mathsf{K})}] = M_\beta^{(k)}(0)$ and $\mathbb{E}[\mathsf{K}_\beta^{(k,\mathsf{K})}] = K_\beta^{(k)}(0)$.

(F1) can be proved using an argument similar to the one used in Chapter 2. (F2) holds by definition.

According to Theorem 3.9.2, a threshold k is optimal if $\alpha M_\beta^{(k)}(0) = K_\beta^{(k)}(0)$. Using Algorithm 3, we can obtain noisy "measurements" of $M_\beta^{(k)}(0)$ and $K_\beta^{(k)}(0)$. Using these noisy measurements, it is possible to search for the optimal threshold using the RM algorithm [65], which is a first-order stochastic root-finding algorithm that works as follows.

We start with an initial guess k_0° of the optimal threshold. Let k_i° denote our guess at the beginning of iteration i. During iteration i, we obtain a noisy measurement $\hat{M}_\beta^{(k,\mathsf{K})}$ of $M_\beta^{(k)}(0)$ and $\hat{K}_\beta^{(k,\mathsf{K})}$ of $K_\beta^{(k)}(0)$ and update our guess as follows:

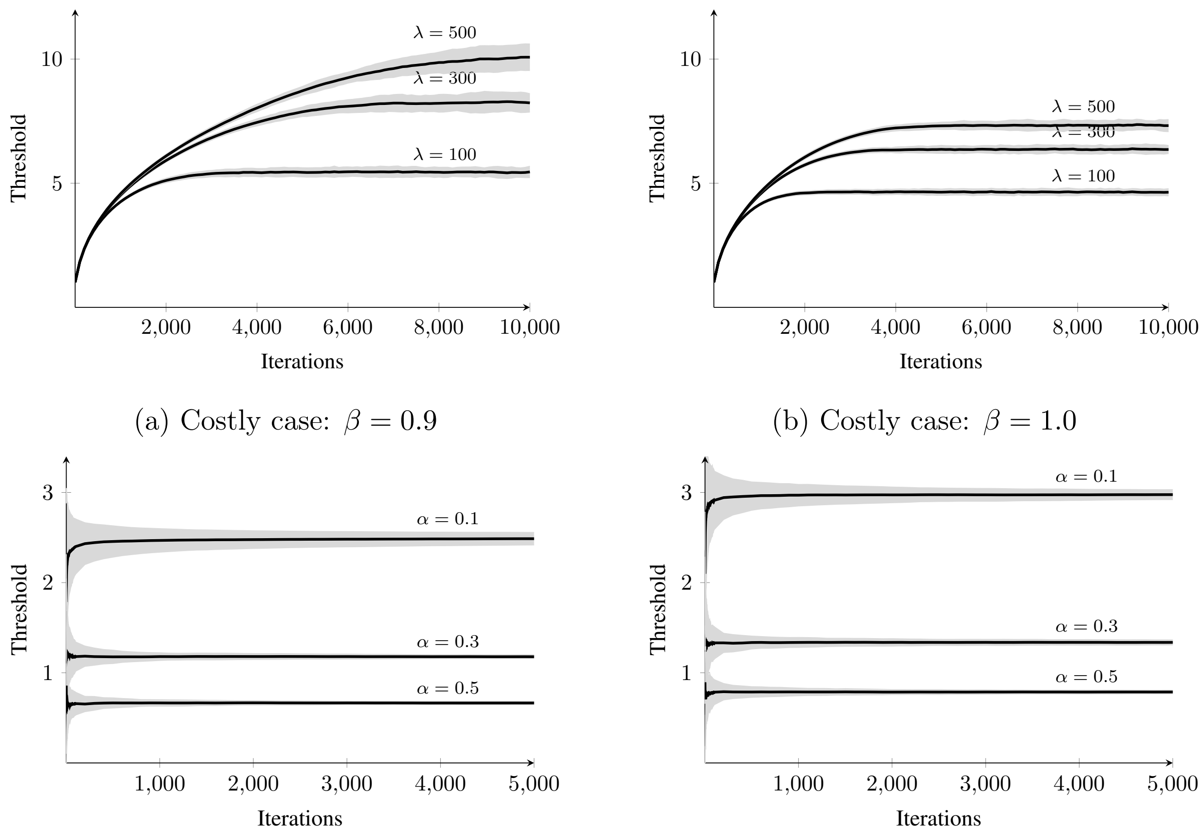
$$k_{i+1}^\circ = k_i^\circ - \gamma_i \left(\alpha \hat{M}_i^{(k_i^\circ,\mathsf{K})} - \hat{K}_i^{(k_i^\circ,\mathsf{K})} \right), \tag{4.5}$$

where γ_i are learning rates that satisfy $\sum_{i=1}^\infty \gamma_i = \infty$ and $\sum_{i=1}^\infty \gamma_i^2 < \infty$. See Algorithm 5 for details.

Algorithm 5: Algorithm for constrained communication

---

input : Initial guess $\mathsf{k}_{\mathsf{init}} \in \mathbb{R}_{>0}$;
Number of episodes $\mathsf{K} \in \mathbb{Z}_{>0}$
Number of iterations $\mathsf{N}_{\mathsf{iterations}} \in \mathbb{Z}_{>0}$
output : Optimal threshold k°
initialize: $k^\circ = \mathsf{k}_{\mathsf{init}}$
for *iteration* $i = 1$ *upto* $\mathsf{N}_{\mathsf{iterations}}$ **do**
  $\gamma_i \leftarrow 1/i$
  $[\hat{L}, \hat{M}, \hat{K}] = \text{MonteCarloEvaluation}\big(k^\circ, \mathsf{K}\big)$
  $k^\circ \leftarrow k^\circ - \gamma_i \big(\alpha \hat{M} - \hat{K}\big)$
return k°



(a) Costly case: $\beta = 0.9$

(b) Costly case: $\beta = 1.0$

(c) Constrained case: $\beta = 0.9$

(d) Constrained case: $\beta = 1.0$

Fig. 4.2 The sample paths for costly and constrained cases for $\varepsilon = 0.3$ with KW and RM algorithms. Here the bold lines represent the sample means for 100 runs and the shaded regions correspond to mean $\pm$ twice the standard deviation across the runs (i.e., the 95% confidence interval).

Theorem 4.3.2 *The threshold iterates k_i° of Algorithm 5 converge almost surely to the optimal thresholds, i.e., $\lim_{i\to\infty} k_i^\circ = k^*(\alpha)$, a.s., where $k^*(\alpha)$ is optimal threshold for Problem 3.9.1.*

The proof follows immediately from [65].

Here we found that using the learning rates $\gamma_i = 1/i$ yields fast convergence and hence we did not use AdaM to adapt the learning rates.

In the next section, we introduce an SA approach, alternative to KW, namely the Smoothed Functional algorithm, and investigate its performance in the costly communication problem.

4.4 Stochastic approximation for costly communication using Smoothed Functional algorithm

Stochastic approximation algorithms scale well to multi-dimensional setup, where the KW algorithm can be replaced by Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm [67] which requires only two random samples to estimate the gradient.

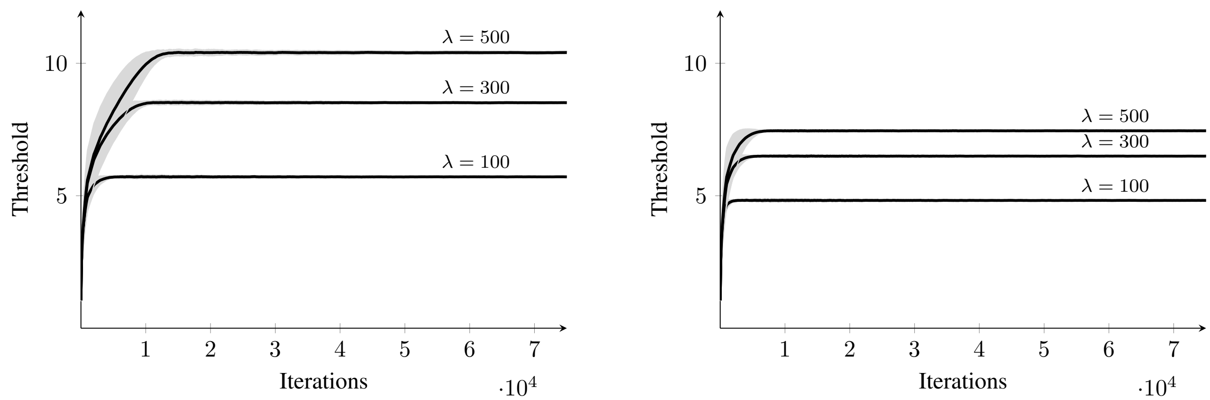
Like SPSA, Smoothed Functional (SF) algorithms, originally introduced in [64], also belong to the class of simultaneous perturbation methods, because they update the gradient/Hessian of the objective using function measurements involving parameter updates that are perturbed simultaneously in all component directions. [68, 69] explore the SF algorithm with Gaussian perturbations for the long-term average cost function with the underlying MDP being ergodic for any given parameter value.

The key idea of SF algorithms is to approximate the gradient/Hessian of the expected performance by its convolution with a multivariate smooth distribution function (most commonly a Gaussian distribution). This results in *smoothing* of the objective function which in turn helps the algorithm to converge to a global minimum or to a point close to it. The main advantage of SF over SPSA algorithm is that in SF one "convexifies" the objective function by convolving it with a convex function, which ensures that the optimal solution thus obtained is close to the globally optimal solution. Thus, it is expected that the variance in the sample paths in the SF algorithm will be smaller compared to that obtained with an SPSA algorithm.

In our work we utilize a variant of SF algorithm with Gaussian smoothing, namely two measurement Gaussian SF algorithm, that has the advantage of a lower estimation bias in comparison to the one-sided form [70]. The basic Gaussian SF algorithm is essentially a stochastic gradient descent algorithm, which works as follows. We start with an initial guess k_0° of the optimal threshold. Let k_i° denote our guess at the beginning of iteration i. During iteration i, we obtain a noisy measurement $\hat{C}_\beta^{(k_i^\circ, \mathsf{K})}(\lambda)$ of the objective function. Let η_i^{K} be a sequence of independent Gaussian variables distributed as $\mathcal{N}(0,1)$. We update our guess of optimal threshold according to the following gradiend descent rule:

$$k_{i+1}^\circ = k_i^\circ - \gamma_i \frac{\eta_i^{\mathsf{K}}}{2\tilde{\beta}} \Big(\hat{C}_i^{(k_i^\circ + \tilde{\beta}\eta_i^{\mathsf{K}}, \mathsf{K})} - \hat{C}_i^{(k_i^\circ - \tilde{\beta}\eta_i^{\mathsf{K}}, \mathsf{K})} \Big), \tag{4.6}$$

where γ_i are learning rates that satisfy $\sum_{i=1}^\infty \gamma_i = \infty$ and $\sum_{i=1}^\infty \gamma_i^2 < \infty$. $\tilde{\beta} > 0$ is tuning



(a) Costly case: $\beta = 0.9$, $\varepsilon = 0.3$ (b) Costly case: $\beta = 1.0$, $\varepsilon = 0.3$

Fig. 4.4 The sample paths for costly case for $\varepsilon = 0.3$ with SF algorithm. Here the bold lines represent the sample means for 100 runs and the shaded regions correspond to mean $\pm$ twice the standard deviation across the runs (i.e., the 95% confidence interval). Note that the sample paths are asymptotically converging to the mean value, which is not the case with KW algorithms (Fig. 4.2).

Algorithm 6: Algorithm for costly communication using SA (SF) algorithm

---

input : Initial guess $\mathsf{k}_{\mathsf{init}} \in \mathbb{R}_{>0}$;
Number of episodes $\mathsf{K} \in \mathbb{Z}_{>0}$
Number of iterations $\mathsf{N}_{\mathsf{iterations}} \in \mathbb{Z}_{>0}$
output : Optimal threshold k°
initialize: $k^\circ = \mathsf{k}_{\mathsf{init}}$
for *iteration* $i = 1$ *upto* $\mathsf{N}_{\mathsf{iterations}}$ **do**
  Pick $\tilde{\beta}$ as a small non-negative real
  Pick η a Gaussian sample with mean 0 and variance 1
  $[\hat{L}^\eta_+, \hat{M}^\eta_+, \hat{K}^\eta_+] = \text{MonteCarloEvaluation}\big(k^\circ + \tilde{\beta}\eta,\, \mathsf{K}\big)$
  $[\hat{L}^\eta_-, \hat{M}^\eta_-, \hat{K}^\eta_-] = \text{MonteCarloEvaluation}\big(k^\circ - \tilde{\beta}\eta,\, \mathsf{K}\big)$
  Compute C^η_+ and C^η_-, using Proposition 3.8.1
  Compute γ_i using AdaM [66]
  $k^\circ \leftarrow k^\circ - \gamma_i \frac{\eta}{2\tilde{\beta}}(\mathsf{C}^\eta_+ - \mathsf{C}^\eta_-)$
return k°

---

parameter. See Algorithm 6 for details.

It is shown in [70, Proposition 6.7], that the expected value of the term in the parenthesis in (4.6) asymptotically converges to the gradient of the performance (w.r.t. the threshold

k), $\nabla_k C_\beta^{(k)}$, as $\tilde{\beta} \to 0$. Furthermore, given $\delta > 0$, let $\mathbb{K}^\delta$ denote the set of points that are in an open δ-neighborhood of the set of global optimal solutions $\mathbb{K}$. Then, under certain assumptions, there exists a $\tilde{\beta}_0 > 0$ such that for all $\tilde{\beta} \in (0, \tilde{\beta}_0]$ corresponding optimal threshold, k° converges almost surely to $\mathbb{K}^\delta$ as the number of iterations $i \to \infty$ [70, Theorem 6.8].

4.5 Numerical results

In all the results reported below, $a = 1$ and $\sigma^2 = 1$. The code for the experiments is available at [71].

4.5.1 Channels with no packet drop (for validation)

We start by comparing the proposed stochastic approximation algorithms with the exact algorithm of Chapter 2.

For costly communication, we consider $\beta \in \{0.9, 1.0\}$ and $\lambda \in \{100, 200, \dots, 700\}$. We set the number of episodes in Algorithm 3 to 1000 and number of iterations in Algorithm 4 to 10,000. The corresponding thresholds are shown in Tables 4.1–4.3.

The optimal thresholds obtained by Fredholm integral equations (as proposed in Chapter 2) are also shown in Tables 4.1–4.3. The thresholds obtained by stochastic approximation are within 10^{-2} of the optimal for most cases. We also compute the total cost $C_\beta^{(k)}(0; \lambda)$ (by solving Fredholm integral equation) for both cases. The cost of the thresholds obtained by stochastic approximation is less than 10^{-3} from the optimal cost. For the SF algorithm, we tune the hyper-parameter $\tilde{\beta}$. We observed that $\tilde{\beta} \in \{0.1, 0.8\}$ yields good results.

For constrained communication, we consider $\beta \in \{0.9, 1.0\}$ and $\alpha \in \{0.1, 0.2, \dots, 0.8\}$. The number of episodes in Algorithm 3 is set to 1. The corresponding thresholds are shown in Table 4.2.

As in the case of costly communication, we compare the thresholds and the performance $D_\beta^{(k)}(0)$ obtained by stochastic approximation with those obtained by Fredholm integral equations. The thresholds obtained by stochastic approximation are within 10^{-3} or less of the optimal.

These results show that the results obtained by stochastic approximation algorithms are accurate.

4.5.2 Channel with packet drops

We repeat the experiments of the previous section with $\varepsilon = 0.3$. To understand the variability of stochastic approximation across different runs, we run each experiment 100 times and plot the mean and standard deviation of the thresholds versus the number of iterations for the KW (which is essentially the SPSA algorithm in 1-D) in Fig. 4.2–4.4. For ease of visualization, we only show the results for a subset of values of λ and α. For both costly and constrained communication, there is very little variation across multiple runs. For costly communication, it takes about 9000 iterations to converge with KW algorithm and around 20,000 iterations to converge with SF algorithm. Note that the asymptotic convergence of the sample paths mentioned in Theorem 4.3.1 is very slow in case of KW (see Fig. 4.2), whereas that is much faster in case of SF algorithm. For constrained communication it takes around 3000 iterations for convergence. We repeat the simulations for the costly performance with the SF algorithms as explained in Algorithm 6. The results obtained are shown in Figs. 4.2–4.4.

4.6 Conclusion

We present stochastic approximation algorithms to compute optimal thresholds for remote state estimation over communication channels with packet drops. The inner loops of these algorithms use Monte Carlo evaluation to get a noisy estimate of the performance of a threshold-based strategy. We embed the renewal feature of the error process to compute the performance using Monte Carlo and call the process Renewal Monte Carlo, which gives satisfactory result within moderately large time for convergence.

Stochastic approximation algorithms scale well to multi-dimensional setup, where the KW algorithm can be replaced by SPSA algorithm [67], SF algorithm [64] among other simultaneous perturbation algorithms, which requires only two random samples to estimate the gradient.

Chapter 5

Two-user interactive communication

In this chapter we discuss *two-user interactive communication* which is not quite in the line of RE problem discussed so far, but is of considerable interest. We consider a stylized framework with two users and analyze the optimality of threshold-based strategies.

5.1 Introduction and literature overview

In recent years, there has been increasing interest in interactive communication in the context interactive computing, interactive source coding, and interactive channel coding.

Communication complexity, which is coding-efficiency for function computation, is considered in [72], where the focus is on establishing order-of-magnitude upper and lower bounds for the communication complexity. The worst-case communication complexities of all Boolean functions are provided in the discussion of the computation of vector-valued functions in the communication complexity framework in [73].

The problem of interactive source reproduction is studied in [74,75] (where a distributed block source coding formulation, for discrete memoryless stationary sources taking values in finite alphabets is considered with the focus on the source reproduction), [76, 77](where two-terminal source-coding for lossless reproduction of a stationary nonergodic source with decoder side-information is considered. Here, the code termination criterion depended on the sources and previous messages). The problem of interactive function (Boolean) computation satisfying an expected per-sample Hamming distortion criterion is considered [78], which is similar to Wyner-Ziv source coding with a decoder side-information. A three-terminal problem with dicrete memoryless stationary sources taking values in finite alphabets is considered in [79], where the sources are observed at one terminal and the other two terminals try to compute the samplewise function of the sources losslessly. Two-terminal distributed source-coding with alternating messages for function computation at both teminals is studied in [80], where the authors show that while interaction is useless in terms of the minimum sum-rate for lossless source reproduction at one or both locations, the gains can be arbitrarily large for function computation even when the sources are independent. In the survey paper [81], the authors address four fundamental organizational and operational issues related to large sensor networks: connectivity, capacity, clocks, and function computation. In a two-user interactive communication setup [82], where a sender communicates with a receiver who wishes to reliably evaluate a function of their combined data, the authors show that if only the sender can transmit, the number of bits required is a conditional entropy of a naturally defined graph. They also determine the number of bits needed when the communicators exchange two messages.

In [83], the author considers an interactive protocol by which he investigates if the channel is noisy, what is the effect upon the number of transmissions needed in order to solve the computation problem reliably and provides a simulation protocol using *explicit tree code*. In [84], the authors describe new ways to simulate 2-party communication protocols and provide the first compression schemes for general randomized protocols and the first direct sum results in the general setting of randomized and distributional communication complexity, without requiring bound on the number of rounds in the protocol or that the distribution of inputs is independent.

In [85], the authors study the interactive channel capacity of an ε-noisy channel and compute the upper bound, which compares with Shannon's non-interactive channel capacity. For a small enough ε, their result gives the first separation between interactive and non-interactive channel capacity. Similar results are discussed in [86], where the first capacity approaching coding schemes are computed that robustly simulate any interactive protocol over an adversarial channel that corrupts any ε fraction of the transmitted symbols.

In this chapter we consider interactive source coding under a zero-delay or real-time constraint. In particular, we consider a model in which two users sequentially obtain noisy observations of a static random variable. At each time, after making its observation, user 1 sends a quantized symbol to user 2; after receiving user 1's symbol and its own observation, user 2 sends a quantized symbol to user 1. Then both users generate an estimate of the underlying static random variable. This processes repeats over a finite time horizon. At each stage, the users quantize and estimate based on the history of their source observations and the quantized symbols from the other user. The per-step cost consists of two parts: a cost associated with each quantized symbol and a distortion cost between the underlying random variable and the estimate made by the two users at that time. The objective is to minimize the total expected cost over a finite horizon.

The above model may also be considered to be a generalization of real-time communication models of [3, 6–8]. The real-time communication problem was first formulated in [3] where the real-time source coding problem was investigated and the structure of optimal encoding and decoding strategies was identified. These results were genralized to real-time joint source-channel coding with noiseless feedback in [6] where, in addition, to the structure of optimal encoding and decoding strategies, a dynamic program to determine the optimal strategies was also identified. The structure of optimal encoding and decoding strategies for communication over noisy channel without feedback was considered in [7]; a dynamic programming decomposition for this setup was obtained in [8]. These results were generalized to real-time communication over noisy channels with noisy feedback in [87].

Some multi-user real-time communication problems have also been investigated in the literature.

5.2 Model and problem formulation

Consider an interactive communication system as shown in Fig. 5.1. The system consists of two users that observe correlated sources $\{X^i_t\}_{t=1}^\infty$, $X^i_t \in \mathbb{X}^i$, $i \in \{1, 2\}$. The sources are generated according to

$$X^i_t = h^i_t(Z, W^i_t), \tag{5.1}$$

where h^i is a known function, $Z \in \mathbb{Z}$ is a random variable of interest and $\{W^1_t\}_{t=1}^\infty$, $\{W^2_t\}_{t=1}^\infty$ are i.i.d. sequences that are independent of each other and also independent of Z. $\{X^1_t\}_{t=1}^\infty$, $\{X^2_t\}_{t=1}^\infty$ are correlated across time and also correlated with each other. For ease of exposition, we assume that the alphabets $\mathbb{Z}$, $\mathbb{X}^1$, and $\mathbb{X}^2$ are finite.

The users sequentially quantize their observations and send a symbol to the other user over a finite-rate noiseless channel. In particular, during time slot t, first user 1 sends a symbol $U^1_t \in \mathbb{U}^1$ to user 2, then user 2 sends a symbol $U^2_t \in \mathbb{U}^2$ to user 1. Both $\mathbb{U}^1$ and

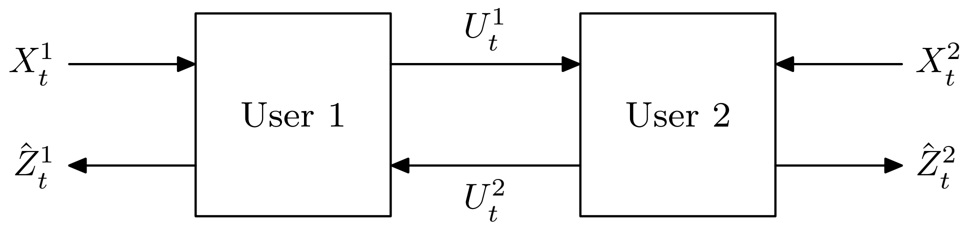


Fig. 5.1 Block diagram of a tewo-user interactive communication system.

$\mathbb{U}^2$ are finite sets and the quantized symbols are generated based on all the information available to users, i.e.,

$$U_t^1 = f_t^1(X_{1:t}^1, U_{1:t-1}^1, U_{1:t-1}^2),\ U_t^2 = f_t^2(X_{1:t}^2, U_{1:t}^1, U_{1:t-1}^2).$$

where f_t^i is called the *encoding rule* of user i at time t. Cost functions $c^i \colon \mathbb{U}^i \to \mathbb{R}_{\geq 0}$ measure the cost of transmission.[1]

During time slot t, after observing the quantized symbol from user 1, user 2 generates an estimate $\hat{Z}_t^2 \in \mathbb{Z}$; after observing the quantized symbol from user 2, user 1 generates an estimate $\hat{Z}_t^1 \in \mathbb{Z}$. These estimates are generated based on all the information available to the users, i.e.,

$$\hat{Z}_t^1 = g_t^1(X_{1:t}^1, U_{1:t}^1, U_{1:t-1}^2),\ \hat{Z}_t^2 = g_t^2(X_{1:t}^2, U_{1:t}^1, U_{1:t}^2), \tag{5.2}$$

where g_t^i is called the *decoding rule* of user i at time t. Distortion functions $d_t^i \colon \mathbb{Z}\times\mathbb{Z} \to \mathbb{R}_{\geq 0}$ measure the fidelity of reconstruction at time t

The sequence $\mathbf{f}^i \coloneqq (f_1^i, \cdots, f_T^i)$, $i \in \{1, 2\}$ is called the *encoding strategy* of user i. Similarly, the sequence $\mathbf{g}^i \coloneqq (g_1^i, \cdots, g_T^i)$, $i \in \{1, 2\}$ is called the *decoding strategy* of user i. The tuple $(\mathbf{f}^1, \mathbf{f}^2, \mathbf{g}^1, \mathbf{g}^2)$ is called the *communication strategy.*

The performance $J(\mathbf{f}^1, \mathbf{f}^2, \mathbf{g}^1, \mathbf{g}^2)$ of a communication strategy $(\mathbf{f}^1, \mathbf{f}^2, \mathbf{g}^1, \mathbf{g}^2)$ is given by

[1] Assuming a transmission cost allows us to model different scenarios. For example, in variable rate communication, the cost function $c^i(u^i) = \log|u^i|$ is used (see [88]). Even in fixed rate communication, a user may not transmit at each time and the transmission cost is zero for not transmitting and a constant for transmitting.

the expected total transmission cost and distortion under that strategy, i.e.,

$$J(\mathbf{f}^1, \mathbf{f}^2, \mathbf{g}^1, \mathbf{g}^2) = \mathbb{E}\Big[\sum_{t=1}^{T}\sum_{i=1}^{2}\big[c^i(U_t^i) + d_t^i(Z, \hat{Z}_t^i)\big]\Big], \tag{5.3}$$

where the expectation is with respect to a joint measure on all system variables induced by the choice of $(\mathbf{f}^1, \mathbf{f}^2, \mathbf{g}^1, \mathbf{g}^2)$.

We are interested in the following optimization problem.

Problem 5.2.1 *For the interactive communication system described above, choose a communication strategy* $(\mathbf{f}^1, \mathbf{f}^2, \mathbf{g}^1, \mathbf{g}^2)$ *that minimizes total expected cost* $J(\mathbf{f}^1, \mathbf{f}^2, \mathbf{g}^1, \mathbf{g}^2)$ *defined in* (5.3).

A key feature of the above model is that both users must generate an estimate of Z at each step. This feature makes our model different from the standard model of interactive communication, where there are multiple rounds of communication and each user generates a single estimate at the end of communication.

Due to this sequential nature of estimation, the standard information theoretic arguments cannot be used. Instead, we directly analyze the optimization problem. The above optimization problem has two decision makers—user 1 and user 2—that have access to different information but need to cooperate and coordinate their actions to minimize a common objective. Therefore, it belongs to the category of *dynamic team* problems [89].

The main conceptual difficulty in solving the above optimization problem is that the information available at both users is increasing with time, and hence, so is the domain of their stratgies. For example, suppose all alphabets are binary. Then there are $2^{2^{3t-2}}$ possibilities for encoding and decoding strategies at each user at time t. Thus, even for a horizon of 3, there are about 10^{175} possible communication strategies (with the dominant term being $(2^{2^7})^4$ possibilities at stage 3). Thus, a brute force search is computationally intractable.

In single agent muti-stage optimization problems, such a difficulty is resolved by identifying a time-homogeneous information state at the decision maker. It is difficult to identify such information states in multi-agent multi-stage decision problems because the different decision makers have access to different information.

We resolve this difficulty in two steps using ideas from team theory. In the first step, we take a person-by-person approach. We arbitrarily fix the strategy of one user, say user 2, and search for the *best response* strategy at user 1. By showing that $X^1_{1:t}$ and $X^2_{1:t}$ are conditionally independent given $(Z, U^1_{1:t}, U^2_{1:t})$, we identify a sufficient statistic $\xi^i_{t|t-1}$ (to be defined later) of $x^i_{1:t}$. This means that there is no loss of optimality in restricting attention to encoders of the form:

$$U^1_t = \hat{f}^1_t(\Xi^1_{t|t-1}, U^1_{1:t-1}, U^2_{1:t-1}), U^2_t = \hat{f}^2_t(\Xi^2_{t|t-1}, U^1_{1:t}, U^2_{1:t-1}).$$

A similar structure for the decoders is also identified.

In the second step, we use the common-information approach of [57] and identify a sufficient statistic π^1_t (to be defined later) of $(u^1_{1:t-1}, u^2_{1:t-1})$ at user 1 and a sufficient statistic π^2_t (to be defined later) of $(u^1_{1:t}, u^2_{1:t-1})$ at user 2. This means that there is no loss of optimality in restricting attention to encoders of the form:

$$U^1_t = \tilde{f}^1_t(\Xi^1_{t|t-1}, \Pi^1_t), \quad U^2_t = \tilde{f}^2_t(\Xi^2_{t|t-1}, \Pi^2_t).$$

We also identify a dynamic program that determines optimal encoding and decoding strategies of the above form.

Remark 16 In this work, we consider static Z. A natural extension could be to discuss the results for a Markovian Z_t process. The main difficulty in fitting a dynamic process is the following. In order to establish the structural results, the key steps are to (i) show the conditional independence of the sources X^i_t given the information $Z_{1:t}$ and (ii) find a sufficient statistic for the process Z_t, which does not grow with time. Although (i) holds for a dynamic Z_t, it is quite tricky to satisfy (ii).

5.2.1 A conditional independence result

The sources are conditionally independent given Z. Our main results rely on the fact that the sources remain conditionally independent when conditioned on Z and the communicated symbols. For ease of notation, $\mathbb{P}(X_{1:t} = x^1_{1:t} \,|\, Z = z, U^1_{1:t} = u^1_{1:t}, U^2_{1:t} = u^2_{1:t})$ is denoted by $\mathbb{P}(x^1_{1:t} \,|\, z, u^1_{1:t}, u^2_{1:t})$. We use similar notation for other probability expressions as well.

Lemma 5.2.1 *For any arbitrary encoding strategies* $(\mathbf{f}^1, \mathbf{f}^2)$ *and any realization* z *of* Z, $x^i_{1:t}$ *of* $X^i_{1:t}$, *and* $u^i_{1:t} \in U^i_{1:t}$, $i \in \{1, 2\}$, *we have the following:*

$$\mathbb{P}(x^1_{1:t}, x^2_{1:t} \mid z, u^1_{1:t}, u^2_{1:t}) = \mathbb{P}(x^1_{1:t} \mid z, u^1_{1:t}, u^2_{1:t})\mathbb{P}(x^2_{1:t} \mid z, u^1_{1:t}, u^2_{1:t}) \tag{5.4}$$

and

$$\mathbb{P}(x^1_{1:t}, x^2_{1:t} \mid z, u^1_{1:t-1}, u^2_{1:t-1}) = \mathbb{P}(x^1_{1:t} \mid z, u^1_{1:t-1}, u^2_{1:t-1})\mathbb{P}(x^2_{1:t} \mid z, u^1_{1:t-1}, u^2_{1:t-1}). \qquad (5.5)$$

Lemma 5.2.1 can be proved using algebraic calculations involving chain rule of probability and total probability. See Appendix D.1 for details. Similar results on conditional independence are discussed in [90] (for decentralized control systems with control sharing), in [91, 92] (for secret key argument and secure computing) and in [93] (for CEO problems).

5.2.2 Belief states and their update

For ease of notation, define $U_t = (U^1_t, U^2_t)$.

Definition 5.2.1 *For any realization $x^i_{1:t}$ of $X^i_{1:t}$ and $u^i_{1:t}$ of $U^i_{1:t}$, $i \in \{1, 2\}$ define belief states $\xi^i_{t|t-1}, \xi^i_{t|t} \in \Delta(\mathbb{Z})$ as follows: for any $z \in \mathbb{Z}$,*

$$\begin{aligned}\xi^i_{t|t-1}(z) &= \mathbb{P}(Z = z \,|\, X^i_{1:t} = x^i_{1:t}, U_{1:t-1} = u_{1:t-1}),\\ \xi^i_{t|t}(z) &= \mathbb{P}(Z = z \,|\, X^i_{1:t} = x^i_{1:t}, U_{1:t} = u_{1:t}),\end{aligned}$$

where $u_t = (u^1_t, u^2_t)$.

$\xi^i_{t|t-1}$ denotes user i's belief on Z after it has observed the source realization of time t but before the communication of that time slot takes place; $\xi^i_{t|t}$ denotes the belief after the communication has taken place. For a specific realization of $(x^i_{1:t}, u_{1:t})$, $\xi^i_{t|t-1}$ and $\xi^i_{t|t}$ are probability distributions. When the conditioning is on random variables $(X^i_{1:t}, U_{1:t})$, the beliefs are $\Delta(\mathbb{Z})$ valued random variables that we denote by the corresponding uppercase letters $\Xi^i_{t|t-1}$ and $\Xi^i_{t|t}$.

In order to derive the structural results, it is important to identify how these beliefs depend on the strategy. To do so, we determine how the beliefs evolve with time. In the sequel, we use $-i$ to denote the user different from user i.

Lemma 5.2.2 *There exist functions $F^i_{t|t}$, $F^i_{t+1|t}$, $i \in \{1, 2\}$, such that*

$$\xi^i_{t|t} = F^i_{t|t}\big(\xi^i_{t|t-1}, u_{1:t}, \mathbf{f}^{-i}\big),\ \xi^i_{t+1|t} = F^i_{t+1|t}\big(\xi^i_{t|t}, u_{1:t}, x^i_{t+1}\big). \qquad (5.6)$$

By combining these two, we get that there exists a function F^i_t such that

$$\xi^i_{t+1|t+1} = F^i_t(\xi^i_{t|t}, u_{1:t}, x^i_{t+1}, \mathbf{f}^{-i}). \tag{5.7}$$

The proof of Lemma 5.2.2 is given in Appendix D.2.

5.2.3 Step 1: The person-by-person approach

As explained earlier, we follow a two-step approach to derive the structure of optimal strategies. In the first step, we follow a person-by-person approach. We arbitrarily fix the strategy of one user and then investigate the best response strategy at the other user.

First, we identify the structure of optimal decoding strategies. Since decoding is a filtering problem, we have:

Proposition 5.2.1 (Structure of optimal decoding strategies) *There is no loss of optimality to restrict the attention to decoding strategies of the form:*

$$\hat{Z}^i_t = \hat{g}^i(\Xi^i_{t|t}), \quad i \in \{1, 2\}, \tag{5.8}$$

where $\hat{g}^i_t$ is given by

$$\hat{g}^i(\xi^i) = \arg\min_{\hat{z}^i \in \mathbb{Z}} \sum_{z \in \mathbb{Z}} d^i(z, \hat{z}^i)\xi^i(z).$$

Now, we fix the decoders at both users according to (5.8) and find the *best response* encoder. By combining Lemmas 5.2.1 and 5.2.2, one can show the following:

Lemma 5.2.3 *Fix decoding strategies $\mathbf{g}^1, \mathbf{g}^2$ to be of the form* (5.8)*. Arbitrarily fix the communication strategy $\mathbf{f}^2$ of user 2. Then, $R^1_t = (\Xi^i_{t|t-1}, U_{1:t-1})$ is an information state for the encoder at user 1. In particular, R^1_t satisfies the following properties:*

1. *R^1_t is a function of the information $(X^1_{1:t}, U_{1:t-1})$ available at user 1.*

2. *The conditional distribution of R^1_{t+1} given all the available information $(X^1_{1:t}, U_{1:t-1})$ and the current action U^1_t depends only on R^1_t and U^1_t, i.e.,*

$$\mathbb{P}(R^1_{t+1} \,|\, X^1_{1:t}, U_{1:t-1}, U^1_t) = \mathbb{P}(R^1_{t+1} \,|\, R^1_t, U^1_t). \tag{5.9}$$

3. R^1_t is a sufficient statistic for the current cost. In particular,

$$\mathbb{E}\Big[\sum_{i\in\{1,2\}}(c^i(U^i_t)+d^i(Z,\hat{Z}^i_t))\mid X^1_{1:t},U_{1:t-1},U^1_t\Big]=\mathbb{E}\Big[\sum_{i\in\{1,2\}}(c^i(U^i_t)+d^i(Z,\hat{Z}^i_t))\mid R^1_t,U^1_t\Big]. \tag{5.10}$$

A similar result holds if $(\mathbf{f}^1,\mathbf{g}^1)$ is fixed and we consider the best response at user 2.

Lemma 5.2.3 implies that $\{R^1_t\}_{t\geq 1}$ is a controlled Markov process with control action U^1_t. Therefore, there is no loss of optimality to restrict attention to *Markov strategies*

$$U^1_t=\hat{f}^1_t(\Xi^1_{t|t-1},U_{1:t-1}).$$

By repeating the argument at user 2, we get the following:

Proposition 5.2.2 (Structure of optimal encoding strategies) *There is no loss of optimality to restrict the attention to encoding strategies of the form:*

$$U^1_t=\hat{f}^1_t(\Xi^1_{t|t-1},U_{1:t-1}),\,U^2_t=\hat{f}^2_t(\Xi^2_{t|t-1},U_{1:t-1},U^1_t). \tag{5.11}$$

5.2.4 Step 2: The common-information approach

We have identified the structure of optimal decoders in closed form and simplified the structure of optimal encoders. In this section, we refine the structural result of Proposition 5.2.2 by following the common-information approach of [57].

We fix the decoding strategies as specified in Proposition 5.2.1 and consider the problem of optimally selecting encoding strategies that are of the form (5.11). Following [57], define the *common information* to be the data that is observed by all future decision makers, i.e., define the common information C^i_t at user i at time t as:

$$C^1_t=U_{1:t-1},\quad C^2_t=(U_{1:t-1},U^1_t).$$

Define the remaining information at user i as local information L^i_t, i.e., $L^i_t=\xi^i_{t|t-1}$. Thus, we can say

$$U^i_t=f^i_t(L^i_t,C^i_t).$$

The main idea of [57] is to consider Problem 5.2.1 from the point of view of a virtual decision maker that observes C^i_t and chooses prescriptions $\phi^i_t:L^i_t\mapsto U^i_t$ that map local information to actions. The encoders simply use these mappings and their local information to generate U^i_t.

It is shown in [57] that the above *coordinated system* is equivalent to the original system. Since the coordinated system has only one decision maker, it can be solved using tools from Markov decision theory. To describe the results, we first note that:

Lemma 5.2.4 *For the encoders of the form given in Proposition 5.2.2, the update of Lemma 5.2.2 can be written as*

$$\xi^i_{t|t} = F^i_{t|t}\big(\xi^i_{t|t-1}, u^{-i}_t, \phi^{-i}_t\big). \tag{5.12}$$

Definition 5.2.2 *For any realization $x^i_{1:t}$ of $X^i_{1:t}$ and $u^i_{1:t}$ of $U^i_{1:t}$, $i \in \{1,2\}$ define belief states $\pi^i_t \in \Delta(\Delta(\mathbb{X}^1) \times \Delta(\mathbb{X}^2))$ as follows: for any $\xi^1_{t|t-1}, \xi^2_{t|t-1} \in \Delta(\mathbb{Z})$,*

$$\begin{aligned}
\pi^1_t(\xi^1,\xi^2) &= \mathbb{P}(\Xi^1_{t|t-1} = \xi^1, \Xi^2_{t|t-1} = \xi^2 \,|\, U_{1:t-1} = u_{1:t-1}),\\
\pi^2_t(\xi^1,\xi^2) &= \mathbb{P}(\Xi^1_{t|t-1} = \xi^1, \Xi^2_{t|t-1} = \xi^2 \,|\, U_{1:t-1} = u_{1:t-1}, U^1_t = u^1_t),
\end{aligned}$$

where $u_t = (u^1_t, u^2_t)$.

Then, similar to Lemma 5.2.4, we can show the following

Lemma 5.2.5 *There exist functions $\tilde{F}^i_t$, $i \in \{1,2\}$, such that*

$$\pi^1_{t+1} = \tilde{F}^1_t\big(\pi^2_t, U^2_t, \phi^2_t\big), \quad \pi^2_t = \tilde{F}^2_t\big(\pi^1_t, U^1_t, \phi^1_t\big). \tag{5.13}$$

According to the discussion above, we fix the decoding strategy to be of the form Proposition 5.2.1 and restrict encoding strategy to be of the form Proposition 5.2.2. The optimization problem then satisfies the partial history sharing model of [57]. Therefore, from [57], we get the following:

Theorem 5.2.1 *There is no loss of optimality in restricting attention to encoding strategies of the form:*

$$U^1_t = \tilde{f}^1_t(\xi^1_{t|t-1}, \Pi^1_t), \quad U^2_t = \tilde{f}^2_t(\xi^2_{t|t-1}, \Pi^2_t). \tag{5.14}$$

Moreover, optimal strategies of this form may be determined from the following dynamic program. Define

$$D^i_t(\xi^i_{t|t}) = \sum_{z\in\mathbb{Z}} d^i_t(z, \hat{g}^i(\xi^i_{t|t}))\xi^i_{t|t}(z),\ i \in \{1,2\}.$$

Then, recursively define value functions $\{V^1_t\}_{t\geq 1}$ *and* $\{V^2_t\}_{t\geq 1}$ *as follows:*

$$V^2_{T+1}(\pi^2) = 0 \tag{5.15}$$

and for $t = T,\ T-1,\ \dots,\ 1$

$$V^2_t(\pi^2) = \min_{\phi^2_t\,:\,\Delta(\mathbb{Z})\to\mathbb{U}^2} \mathbb{E}[c^2(U^2_t) + D^2_t(\Xi^2_{t|t}) + V^1_{t+1}(\Pi^1_{t+1}) \mid \Pi^2_t = \pi^2, U^2_t = \phi^2_t(\Xi^2_{t|t-1})], \tag{5.16}$$

and

$$V^1_t(\pi^1) = \min_{\phi^1_t\,:\,\Delta(\mathbb{Z})\to\mathbb{U}^1} \mathbb{E}[c^1(U^1_t) + D^1_t(\Xi^1_{t|t}) + V^2_t(\Pi^2_t) \mid \Pi^1_t = \pi^1, U^1_t = \phi^1_t(\Xi^1_{t|t-1})]. \tag{5.17}$$

Let $\psi^2_t(\pi^2)$ *denote the arg min of* (5.16) *and* $\psi^1_t(\pi^1)$ *denote the arg min of* (5.17). *Then, the optimal strategy* $\tilde{\mathbf{f}}^1, \tilde{\mathbf{f}}^2$ *is given by*

$$\tilde{f}^i_t(\xi^i_{t|t-1}, \pi^i_t) = \psi^i_t(\pi^i_t)(\xi^i_{t|t-1}). \tag{5.18}$$

Note that the expectations in (5.16) and (5.17) can be computed using the update rules in Lemmas 5.2.2 and 5.2.5.

5.3 Discussion and conclusion

Theorem 5.2.1 identifies a sufficient statistic at the encoder and the decoder; the domain of which does not depend on time. Moreover, the dynamic program provides a way to identify optimal (or sub-optimal) strategies. As a consequence, the search complexity increases linearly with time horizon (rather than double exponentially, as for brute force search). If $\mathbb{Z}$ is finite, say of cardinality n, then $\Delta(\mathbb{X}^i)$ may be viewed as an element of $\mathbb{R}^{n-1}$; and hence the belief space is the space of probability distributions on $\mathbb{R}^{2n-2}$.

Note that the dynamic program is similar to the dynamic programs for partially observable Markov decision processes (POMDP). So, it is possible to use point-based algorithms for continuous state POMDPs to numerically solve the resultant dynamic program. Another option is to use discretization based algorithms developed for real-time communication [94]. Following [9], it may be possible to establish that threshold-based strategies are optimal when all random variables are Gaussian and the transmitter has the option of not transmitting.

Since the domain of the encoding and decoding strategies is not changing with time, the result of Theorem 1 naturally extends to infinite horizon setups as well. We expect that under appropriate regularity conditions, the optimal strategy is time homogeneous and given by the fixed point of a dynamic program. It may be possible to use such a dynamic program to find bounds on time average distortion.

Although the results of this chapter are derived for a two-user interactive communication system, they generalize to the following multi-terminal setup. Consider n users with observations similar to (5.1). During time-slot t, first user 1 broadcasts a symbol U_t^1 to all users. Then user 2 broadcasts U_t^2 to all users, and so on, until user n broadcasts U_t^n to all users. All users generate an estimate of Z and the process repeats at $t+1$. Such a multi-user setup can be analyzed using the same approach as presented in this work.

Chapter 6

Sufficient conditions for evenness and quasi-convexity of the value function and the optimal strategies

In this chapter we discuss a topic which is not a Remote Estimation (RE) problem by itself, but the main result discussed here is one of the central ideas in establishing the structural results in the previous chapters. In all of the scenarios we have mentioned so far, we recognize the role of symmetry and monotonicity for the optimality of the threshold-based strategies. This motivates us to analyze in a more general setup the sufficient conditions for the value function and the optimal strategies to be symmetric and quasi-convex.

6.1 Motivation

Markov decision theory is often used to identify structural or qualitative properties of optimal strategies. Examples include control limit strategies in machine maintenance [95, 96], threshold-based strategies for executing call options [97, 98], and monotone strategies in queueing systems [99, 100]. In all of these models, the optimal strategy is *monotone* in the state, i.e., if $x > y$ then the action chosen at x is greater (or less) than or equal to the action chosen at y. Motivated by this, general conditions under which the optimal strategy is monotone in scalar-valued states are identified in [60, 101–105]. Similar conditions for vector-valued states are identified in [106–108]. General conditions under which the value function is increasing and convex are established in [109].

Most of the above results are motivated by queueing models where the state (i.e., the queue length) takes non-negative values. However, for typical applications in systems and control, the state takes both positive and negative values. Often, the system behavior is symmetric for positive and negative values, so one expects the optimal strategy to be even. Thus, for such systems, a natural counterpart of monotone functions are even and quasi-convex (or quasi-concave) functions. In this chapter, we identify sufficient conditions under which the value function and optimal strategy are even and quasi-convex.

As a motivating example, consider a remote estimation system in which a sensor observes a Markov process and decides whether to transmit the current state of the Markov process to a remote estimator. There is a cost or constraint associated with transmission. When the transmitter does not transmit (or when the transmitted packet is dropped due to interference), the estimator generates an estimate of the state of the Markov process based on the previously received states. The objective is to choose transmission and estimation strategies that minimize either the expected distortion and cost of communication or minimize expected distortion under the transmission constraint. Variations of such models have been considered in [9–11, 13, 14, 26, 110].

In such models the optimal transmission and estimation strategies are identified in two steps. In the first step, the joint optimization of transmission and estimation strategies is investigated and it is established that there is no loss of optimality in restricting attention to estimation strategies of a specific form. In the second step, estimation strategies are restricted to the form identified in the first step and the structure of the best response transmission strategies is established. In particular, it is shown that the optimal transmission strategies are even and quasi-convex.[1] Currently, in the literature these results are established on a case by case basis. For example, see [9, Theorem 1], [10, Theorem 3], [28, Theorem 1], [14, Theorem 1] among others.

In this chapter, we identify sufficient conditions for the value functions and optimal strategy of a Markov decision process to be even and quasi-convex. We then consider a general model of remote estimation and verify these sufficient conditions.

[1] When the action space is binary—as is the case in most of the models of remote estimation—an even and quasi-convex strategy is equivalent to one in which the action zero is chosen whenever the absolute value of the state is less than a threshold.

6.1.1 Model and problem formulation

Consider a Markov decision process (MDP) with state space $\mathbb{X}$ (which is either $\mathbb{R}$, the real line, or a symmetric subset of the form $[-a, a]$) and action space $\mathbb{U}$ (which is either a countable set or a compact subset of reals).

Let $X_t \in \mathbb{X}$ and $U_t \in \mathbb{U}$ denote the state and action at time t. The initial state X_1 is distributed according to the probability density function μ and the state evolves in a controlled Markov manner, i.e., for any Borel measurable subset A of $\mathbb{X}$,

$$\mathbb{P}(X_{t+1} \in A \mid X_{1:t} = x_{1:t}, U_{1:t} = u_{1:t}) = \mathbb{P}(X_{t+1} \in A \mid X_t = x_t, U_t = u_t),$$

where $x_{1:t}$ is a short hand notation for $(x_1, \dots, x_t)$ and a similar interpretation holds of $u_{1:t}$. We assume that there exists a (time-homogeneous) controlled transition density $p(y|x; u)$ such that for any Borel measurable subset A of $\mathbb{X}$,

$$\mathbb{P}(X_{t+1} \in A \mid X_t = x, U_t = u) = \int_A p(y|x; u) dy.$$

We use $p(u)$ to transition density corresponding to action $u \in \mathbb{U}$.

The system operates for a finite horizon T. For any time $t \in \{1, \dots, T-1\}$, a measurable function $c_t \colon \mathbb{X} \times \mathbb{U} \to \mathbb{R}$ denotes the instantaneous cost at time t and at the terminal time T a measurable function $c_T \colon \mathbb{X} \to \mathbb{R}$ denotes the terminal cost.

The actions at time t are chosen according to a Markov strategy f_t, i.e.,

$$U_t = f_t(X_t), \quad t \in \{1, \dots, T-1\}.$$

The objective is to choose a decision strategy $\mathbf{f} \coloneqq (f_1, \dots, f_{T-1})$ to minimize the expected total cost

$$J(\mathbf{f}) \coloneqq \mathbb{E}^f \Big[\sum_{t=1}^{T-1} c_t(X_t, U_t) + c_T(X_T) \Big]. \tag{6.1}$$

We denote such an MDP by $(\mathbb{X}, \mathbb{U}, p, c)$.

From Markov decision theory [111], we know that an optimal strategy is given by the solution of the following dynamic program. Recursively define value functions $V_t \colon \mathbb{X} \to \mathbb{R}$

and value-action functions $Q_t \colon \mathbb{X} \times \mathbb{U} \to \mathbb{R}$ as follows: for all $x \in \mathbb{X}$ and $u \in \mathbb{U}$,

$$V_T(x) = c_T(x), \tag{6.2}$$

and for $t \in \{T-1, \dots, 1\}$,

$$\begin{aligned} Q_t(x,u) &= c_t(x,u) + \mathbb{E}[V_{t+1}(X_{t+1}) \mid X_t = x, U_t = u] \\ &= c_t(x,u) + \int_{\mathbb{X}} p(y|x;u) V_{t+1}(y) dy, \qquad (6.3) \\ V_t(x) &= \min_{u \in \mathbb{U}} Q_t(x,u). \qquad (6.4) \end{aligned}$$

Then, a strategy $\mathbf{f}^* = (f_1^*, \dots, f_{T-1}^*)$ defined as

$$f_t^*(x) \in \arg\min_{u \in \mathbb{U}} Q_t(x,u)$$

is optimal. To avoid ambiguity when the arg min is not unique, we pick

$$f_t^*(x) = \begin{cases} \max \left\{ v \in \arg\min_{u \in \mathbb{U}} Q_t(x,u) \right\}, & \text{if } x \geq 0 \\ \min \left\{ v \in \arg\min_{u \in \mathbb{U}} Q_t(x,u) \right\}, & \text{if } x < 0. \end{cases} \tag{6.5}$$

Let $\mathbb{X}_{\geq 0}$ and $\mathbb{X}_{>0}$ denote the sets $\{x \in \mathbb{X} : x \geq 0\}$ and $\{x \in \mathbb{X} : x > 0\}$. We say that a function $g \colon \mathbb{X} \to \mathbb{R}$ is *even and quasi-convex* if it is even and for $x, x' \in \mathbb{X}_{\geq 0}$ such that $x < x'$, we have that $g(x) \leq g(x')$. The main contribution of this chapter is to identify sufficient conditions under which V_t and f_t^* are even and quasi-convex.

6.1.2 Main result

For a given $u \in \mathbb{U}$, we say that a controlled transition density $p(u)$ on $\mathbb{X} \times \mathbb{X}$ is even if for all $x, y \in \mathbb{X}$, $p(y|x;u) = p(-y|-x;u)$.

Our main result is the following.

Theorem 6.1.1 *Given an MDP* $(\mathbb{X}, \mathbb{U}, p, c)$*, define for* $x, y \in \mathbb{X}_{\geq 0}$ *and* $u \in \mathbb{U}$*,*

$$S(y|x;u) = 1 - \int_{A_y} [p(z|x;u) + p(-z|x;u)] dz, \tag{6.6}$$

where $A_y = \{x \in \mathbb{X} : x < y\}$*. Consider the following conditions:*

(C1) $c_T(\cdot)$ *is even and increasing and for* $t \in \{1, \dots, T-1\}$ *and* $u \in \mathbb{U}$, $c_t(\cdot, u)$ *is even and quasi-convex.*

(C2) *For all* $u \in \mathbb{U}$, $p(u)$ *is even.*

(C3) *For all* $u \in \mathbb{U}$ *and* $y \in \mathbb{X}_{\geq 0}$, $S(y|x;u)$ *is increasing for* $x \in \mathbb{X}_{\geq 0}$.

(C4) *For* $t \in \{1, \dots, T-1\}$, $c_t(x,u)$ *is submodular*[2] *in* (x,u) *on* $\mathbb{X}_{\geq 0} \times \mathbb{U}$.

(C5) *For all* $y \in \mathbb{X}_{\geq 0}$, $S(y|x;u)$ *is submodular in* (x,u) *on* $\mathbb{X}_{\geq 0} \times \mathbb{U}$.

Then, under (C1)–(C3), $V_t(\cdot)$ *is even and quasi-convex for all* $t \in \{1, \dots, T\}$ *and under* (C1)–(C5), $f_t^*(\cdot)$ *is even and quasi-convex for all* $t \in \{1, \dots, T-1\}$.

The main idea of the proof is as follows. First, we identify conditions under which the value function and optimal strategy of an MDP are even. Next, we show that if we construct an MDP by "folding" the transition density, then the "folded MDP" has the same value function and optimal strategy as the original MDP for non-negative values of the state. Finally, we show that if we take the sufficient conditions under which the value function and the optimal strategy of the folded MDP are increasing and "unfold" these conditions back to the original model, we get conditions (C1)–(C5) above. The details are given in the next two sections.

6.2 Even MDPs and folded representations

We say that *an MDP is even* if for every t and every $u \in \mathbb{U}$, $V_t(x)$, $Q_t(x,u)$ and $f_t^*(x)$ are even in x. We start by identifying sufficient conditions for an MDP to be even.

6.2.1 Sufficient condition for MDP to be even

Proposition 6.2.1 *Suppose an MDP* $(\mathbb{X}, \mathbb{U}, p, c)$ *satisfies the following properties:*

(A1) $c_T(\cdot)$ *is even and for every* $t \in \{1, \cdots, T-1\}$ *and* $u \in \mathbb{U}$, $c_t(\cdot, u)$ *is even.*

(A2) *For every* $u \in \mathbb{U}$, *the transition density* $p(u)$ *is even.*

[2]Submodularity is defined in Sec. 6.3.2

Then, the MDP is even.

Proof We proceed by backward induction. $V_T(x) = c_T(x)$ which is even by (A1). This forms the basis of induction. Now assume that $V_{t+1}(x)$ is even in x. For any $u \in \mathbb{U}$, we show that $Q_t(x, u)$ is even in x. Consider,

$$\begin{aligned} Q_t(-x, u) &= c_t(-x, u) + \int_{\mathbb{X}} p(y|-x; u) V_{t+1}(y) dy \\ &\overset{(a)}{=} c_t(x, u) + \int_{\mathbb{X}} p(-z|-x; u) V_{t+1}(-z) dz \\ &\overset{(b)}{=} c_t(x, u) + \int_{\mathbb{X}} p(z|x; u) V_{t+1}(z) dz = Q_t(x, u) \end{aligned}$$

where (a) follows from (A1), a change of variables $y = -z$, and the fact that $\mathbb{X}$ is a symmetric interval; and (b) follows from (A2) and the induction hypothesis that $V_{t+1}(\cdot)$ is even. Hence, $Q_t(\cdot, u)$ is even.

Since $Q_t(\cdot, u)$ is even, Eqs. (6.4) and (6.5) imply that V_t and f_t^* are also even. Thus, the result is true for time t and, by induction, true for all time t.

6.2.2 Folding operator for distributions

We now show that if the value function is even, we can construct a "folded" MDP with state-space $\mathbb{X}_{\geq 0}$ such that the value function and optimal strategy of the folded MDP match that of the original MDP on $\mathbb{X}_{\geq 0}$. For that matter, we first define the following:

Definition 6.2.1 (Folding Operator) *Given a probability density π on $\mathbb{X}$, the folding operator $\mathcal{F}\pi$ gives a density $\tilde{\pi}$ on $\mathbb{X}_{\geq 0}$ such that for any $x \in \mathbb{X}_{\geq 0}$, $\tilde{\pi}(x) = \pi(x) + \pi(-x)$.*

As an immediate implication, we have the following:

Lemma 6.2.1 *If $f : \mathbb{X} \to \mathbb{R}$ is even, then for any probability distribution π on $\mathbb{X}$ and $\tilde{\pi} = \mathcal{F}\pi$, we have*

$$\int_{x \in \mathbb{X}} f(x) \pi(x) dx = \int_{x \in \mathbb{X}_{\geq 0}} f(x) \tilde{\pi}(x) dx.$$

Now, we generalize the folding operator to transition densities.

Definition 6.2.2 *Given a transition density p on $\mathbb{X} \times \mathbb{X}$, the folding operator $\mathcal{F}p$ gives a transition density $\tilde{p}$ on $\mathbb{X}_{\geq 0} \times \mathbb{X}_{\geq 0}$ such that for any $x, y \in \mathbb{X}_{\geq 0}$, $\tilde{p}(y|x) = p(y|x) + p(-y|x)$.*

Definition 6.2.3 (Folded MDP) *Given an MDP* $(\mathbb{X}, \mathbb{U}, p, c_t)$*, define the* folded MDP *as* $(\mathbb{X}_{\geq 0}, \mathbb{U}, \tilde{p}, c_t)$*, where for all* $u \in \mathbb{U}$*,* $\tilde{p}(u) = \mathcal{F}p(u)$*.*

Let $\tilde{Q}_t$ and $\tilde{V}_t$ and $\tilde{f}_t^*$ denote respectively the value-action function, the value function, and the optimal strategy of the folded MDP. Then, we have the following.

Proposition 6.2.2 *If the MDP* $(\mathbb{X}, \mathbb{U}, p, c_t)$ *is even, then for any* $x \in \mathbb{X}$ *and* $u \in \mathbb{U}$*,*

$$Q_t(x, u) = \tilde{Q}_t(|x|, u), \quad V_t(x) = \tilde{V}_t(|x|), \quad f_t^*(x) = \tilde{f}_t^*(|x|). \tag{6.7}$$

Proof We proceed by backward induction. For $x \in \mathbb{X}$ and $\tilde{x} \in \mathbb{X}_{\geq 0}$, $V_T(x) = c_T(x)$ and $\tilde{V}_T(\tilde{x}) = c_T(\tilde{x})$. Since $V_T(\cdot)$ is even, $V_T(x) = V_T(|x|) = \tilde{V}_T(|x|)$. This is the basis of induction. Now assume that for all $x \in \mathbb{X}$, $V_{t+1}(x) = \tilde{V}_{t+1}(|x|)$. Consider $x \in \mathbb{X}_{\geq 0}$ and $u \in \mathbb{U}$. Then we have

$$\begin{aligned} Q_t(x, u) &= c_t(x, u) + \int_{\mathbb{X}} p(y|x; u) V_{t+1}(y) dy \\ &\stackrel{(a)}{=} c_t(x, u) + \int_{\mathbb{X}_{\geq 0}} \tilde{p}(y|x; u) V_{t+1}(y) dy \\ &\stackrel{(b)}{=} c_t(x, u) + \int_{\mathbb{X}_{\geq 0}} \tilde{p}(y|x; u) \tilde{V}_{t+1}(y) dy = \tilde{Q}_t(x, u), \end{aligned}$$

where (a) uses Lemma 6.2.1 and that V_{t+1} is even and (b) uses the induction hypothesis.

Since the Q-functions match for $x \in \mathbb{X}_{\geq 0}$, equations (6.4) and (6.5) imply that the value functions and the optimal strategies also match on $\mathbb{X}_{\geq 0}$, i.e., for $x \in \mathbb{X}_{\geq 0}$,

$$V_t(x) = \tilde{V}_t(x) \quad \text{and} \quad f_t^*(x) = \tilde{f}_t^*(x).$$

Since V_t and f_t^* are even, we get that (6.7) is true at time t. Hence, by principle of induction, it is true for all t.

6.3 Monotonicity of the value function and the optimal strategy

We have shown that under (A1) and (A2) the original MDP is equivalent to a folded MDP with state-space $\mathbb{X}_{\geq 0}$. Thus, we can use standard conditions to determine when the value function and the optimal strategy of the folded MDP are monotone. Translating these conditions back to the original model, we get the sufficient conditions for the original model.

6.3.1 Monotonicity of the value function

The results on monotonicity of value functions rely on the notion of stochastic monotonicity.

Given a transition density p defined on $\mathbb{X}$, the cumulative transition density function P is defined as

$$P(y|x) = \int_{A_y} p(z)dz, \quad \text{where } A_y = \{x \in \mathbb{X} : x < y\}.$$

Definition 6.3.1 (Stochastic Monotonicity) *A transition density p on $\mathbb{X}$ is said to be* stochastically monotone *if for every $y \in \mathbb{Y}$, the cumulative density function $P(y|x)$ corresponding to p is decreasing in x.*

Proposition 6.3.1 *Suppose the folded MDP $(\mathbb{X}_{\geq 0}, \mathbb{U}, \tilde{p}, c)$ satisfies the following:*

(B1) *$c_T(x)$ is increasing in x for $x \in \mathbb{X}_{\geq 0}$; for any $t \in \{1, \dots, T-1\}$ and $u \in \mathbb{U}$, $c_t(x, u)$ is increasing in x for $x \in \mathbb{X}_{\geq 0}$.*

(B2) *For any $u \in \mathbb{U}$, $\tilde{p}(u)$ is stochastically monotone.*

Then, for any $t \in \{1, \dots, T\}$, $\tilde{V}_t(x)$ is increasing in x for $x \in \mathbb{X}_{\geq 0}$.

A version of this proposition when $\mathbb{X}$ is a subset of integers is given in [60, Theorem 4.7.3]. The same proof argument also works when $\mathbb{X}$ is a subset of reals.

Recall the definition of S given in (6.6). (B2) is equivalent to the following:

(B2') For every $u \in \mathbb{U}$ and $x, y \in \mathbb{X}_{\geq 0}$, $S(y|x, u)$ is increasing in x.

An immediate consequence of Propositions 6.2.1, 6.2.2, and 6.3.1 is the following:

Corollary 6.3.1 *Under* (A1)*,* (A2)*,* (B1)*, and* (B2) *(or* (B2')*), the value functions $V_t(\cdot)$ is even and quasi-convex.*

Remark 17 Note that (A1) and (B1) are equivalent to (C1), (A2) is same as (C2), and (A2) and (B2) (or equivalently, (A2) and (B2')) are equivalent to (C3). Thus, Corollary 6.3.1 proves the first part of Theorem 6.1.1.

6.3.2 Monotonicity of the optimal strategy

Now we state sufficient conditions under which the optimal strategy is increasing. These results rely on the notion of submodularity.

Definition 6.3.2 (Submodular function) *A function $g\colon \mathbb{X} \times \mathbb{U} \to \mathbb{R}$ is called submodular if for any $x, y \in \mathbb{X}$ and $u, v \in \mathbb{U}$ such that $x \geq y$ and $u \geq v$, we have*

$$g(x,u) + g(y,v) \leq g(x,v) + g(y,u).$$

An equivalent characterization of submodularity is that

$$\begin{aligned} & g(y,u) - g(y,v) \geq g(x,u) - g(x,v), \\ \implies\ & g(x,v) - g(y,v) \geq g(x,u) - g(y,u), \end{aligned}$$

which implies that the differences are decreasing.

Proposition 6.3.2 *Suppose that in addition to* (B1) *and* (B2) *(or* (B2')*), the folded MDP* $(\mathbb{X}_{\geq 0}, \mathbb{U}, \tilde{p}, c_t)$ *satisfies the following property:*

(B3) *For all $t \in \{1, \dots, T-1\}$, $c_t(x,u)$ is submodular in (x,u) on $\mathbb{X}_{\geq 0} \times \mathbb{U}$.*

(B4) *For all $y \in \mathbb{X}_{\geq 0}$, $S(y|x;u)$ is submodular in (x,u) on $\mathbb{X}_{\geq 0} \times \mathbb{U}$, where $S(y|x;u)$ is defined in* (6.6)*.*

Then, for every $t \in \{1, \cdots, T-1\}$, the optimal strategy $\tilde{f}_t^*(x)$ is increasing in x for $x \in \mathbb{X}_{\geq 0}$.

A version of this proposition when $\mathbb{X}$ is a subset of integers is given in [60, Theorem 4.7.4]. The same proof argument also works when $\mathbb{X}$ is a subset of reals.

An immediate consequence of Propositions 6.2.1, 6.2.2, and 6.3.1 is the following:

Corollary 6.3.2 *Under* (A1), (A2), (B1), (B2) *(or* (B2')*),* (B3)*, and* (B4) *the optimal strategy $f_t^*(\cdot)$ is even and quasi-convex.*

Remark 18 As argued in Remark 17, (A1), (A2), (B1), (B2) are equivalent to (C1)–(C3). Note that (B3), (B4) is the same as (C4), (C5). Thus, Corollary 6.3.2 proves the second part of Theorem 6.1.1.

6.4 Remarks about discrete $\mathbb{X}$

So far we assumed that $\mathbb{X}$ was a subset of the real line. Now suppose $\mathbb{X}$ is discrete (either the set $\mathbb{Z}$ of integers or a symmetric subset of the form $\{-a, \dots, a\}$). With a slight abuse of notation, let $p(y|x;u)$ denote $\mathbb{P}(X_{t+1} = y | X_t = x, U_t = u)$.

Theorem 6.4.1 *The result of Theorem 6.1.1 is true for discrete $\mathbb{X}$ with S defined as*

$$S(y|x,u) = 1 - \sum_{z \in A_y} \big[p(z|x;u) + p(-z|x;u) \big]$$

where $A_y = \{x \in \mathbb{X} : x < y\}$.

The proof proceeds along the same lines as the proof of Theorem 6.1.1. In particular,

- Proposition 6.2.1 is also true for discrete $\mathbb{X}$.
- Given a probability mass function π on $\mathbb{X}$, define the folding operator $\mathcal{F}$ as follows: $\tilde{\pi} = \mathcal{F}\pi$ means that $\tilde{\pi}(0) = \pi(0)$ and for any $x \in \mathbb{X}_{>0}$, $\tilde{\pi}(x) = \pi(x) + \pi(-x)$.
- Use this definition of the folding operator to define the folded MDP, as in Definition 6.2.3. Proposition 6.2.2 remains true with this modified definition.
- A discrete state Markov chain with transition function p is stochastically monotone if for every $y \in \mathbb{X}$,

  $$P(y|x) = \sum_{z \in \mathbb{A}_y} p(z), \quad \text{where } A_y = \{x \in \mathbb{X} : x < y\}$$

  is decreasing in x.
- Propositions 6.3.1 and 6.3.2 are also true for discrete $\mathbb{X}$.
- The result of Theorem 6.4.1 follows from Corollaries 6.3.1 and 6.3.2.

6.4.1 Monotone dynamic programming

Under (C1)–(C4), the even and quasi-convex property of the optimal strategy strategy can be used to simplify the dynamic program given by (6.2)–(6.4). For conciseness, assume that the state space $\mathbb{X}$ is a set of integers form $\{-a, -a+1, \cdots, a-1, a\}$ and the action space $\mathbb{U}$ is a set of integers of the form $\{\underline{u}, \underline{u}+1, \cdots, \bar{u}-1, \bar{u}\}$.

Initialize $V_T(x)$ as in (6.2). Now, suppose $V_{t+1}(\cdot)$ has been calculated. Instead of computing $Q_t(x,u)$ and $V_t(x)$ according to (appropriately modified versions of) (6.3) and (6.4), we proceed as follows:

1. Set $x = 0$ and $w_x = \underline{u}$.

2. For all $u \in [w_x, \bar{u}]$, compute $Q_t(x,u)$ according to (6.3).

3. Instead of (6.4), compute

$$V_t(x) = \min_{u \in [w_x, \bar{u}]} Q_t(x,u), \quad \text{and set}$$
$$f_t(x) = \max\{v \in [w_x, \bar{u}] \text{ s.t. } V_t(x) = Q_t(x,v)\}.$$

4. Set $V_t(-x) = V_t(x)$ and $f_t(-x) = f_t(x)$.

5. If $x = a$, then stop. Otherwise, set $w_{x+1} = f_t(x)$ and $x = x+1$. Go to step 2.

6.4.2 A remark on randomized actions

Suppose $\mathbb{U}$ is a discrete set of the form $\{\underline{u}, \underline{u}+1, \ldots, \bar{u}\}$. In constrained optimization problems, it is often useful to consider the action space $\mathbb{W} = [\underline{u}, \bar{u}]$, where for $u, u+1 \in \mathbb{U}$, an action $w \in (u, u+1)$ corresponds to a randomization between the "pure" actions u and $u+1$. More precisely, let transition probability $\breve{p}$ corresponding to $\mathbb{W}$ be given as follows: for any $x, y \in \mathbb{X}$ and $w \in (u, u+1)$,

$$\breve{p}(y|x;w) = (1-\theta(w))p(y|x;u) + \theta(w)p(y|x;u+1)$$

where $\theta : \mathbb{W} \to [0,1]$ is such that for any $u \in \mathbb{U}$,

$$\lim_{w \downarrow u} \theta(w) = 0, \quad \text{and} \quad \lim_{w \uparrow u+1} \theta(w) = 1. \tag{6.8}$$

Thus, $\breve{p}(w)$ is continuous at all $u \in \mathbb{U}$.

Theorem 6.4.2 *If $p(u)$ satisfies* (C2), (C3), *and* (C5) *then so does $\breve{p}(w)$.*

Proof Since $\breve{p}(w)$ is linear in $p(u)$ and $p(u+1)$, both of which satisfy (C2) and (C3), so does $\breve{p}(w)$.

To prove that $\breve{p}(w)$ satisfies (C5), note that

$$\breve{S}(y|x,w) = S(y|x,u) + \theta(w)[S(y|x,u+1) - S(y|x,u)].$$

So, for $v, w \in (u, u+1)$ such that $v > w$, we have that

$$\breve{S}(y|x,v) - \breve{S}(y|x,w) = \big(\theta(v) - \theta(w)\big)[S(y|x,u+1) - S(y|x,u)]$$

Since $\theta(\cdot)$ is increasing, $\theta(v) - \theta(w) \geq 0$. Moreover, since $S(y|x,u)$ is submodular in (x,u), $S(y|x,u+1) - S(y|x,u)$ is decreasing in x, and, therefore, so is $\breve{S}(y|x,v) - \breve{S}(y|x,w)$. Hence, $\breve{S}(y|x,w)$ is submodular in (x,w) on $\mathbb{X} \times (u, u+1)$. Due to (6.8), $\breve{S}(y|x;w)$ is continuous in w. Hence, $\breve{S}(y|x;w)$ is submodular in (x,w) on $\mathbb{X} \times [u, u+1]$. By piecing intervals of the form $[u, u+1]$ together, we get that $\breve{S}(y|x;w)$ is submodular on $\mathbb{X} \times \mathbb{W}$.

6.5 An example: Optimal power allocation strategies in remote estimation

Consider a remote estimation system that consists of a sensor and an estimator. The sensor observes a first order autoregressive process $\{X_t\}_{t\geq 1}$, $X_t \in \mathbb{X}$, where $X_1 = 0$ and for $t > 1$,

$$X_{t+1} = aX_t + W_t,$$

where $a \in \mathbb{X}$ is a constant and $\{W_t\}_{t\geq 1}$ is an i.i.d. noise process. We consider two cases:

1. **Case A:** The state space is continuous, i.e., $\mathbb{X} = \mathbb{R}$. In this case we assume that $\{W_t\}_{t\geq 1}$ is distributed according to probability density function φ.

2. **Case B:** The state space is discrete, i.e., $\mathbb{X} = \mathbb{Z}$. In this case we assume that $\{W_t\}_{t\geq 1}$ is distributed according to probability mass function φ.

At each time step, the sensor uses power U_t to send a packet containing X_t to the remote estimator. U_t takes values in $[0, u_{\max}]$, where $U_t = 0$ denotes that no packet is sent. The packet is received with probability $q(U_t)$, where q is an increasing function with $q(0) = 0$ and $q(u_{\max}) \leq 1$.

Let Y_t denote the received symbol. $Y_t = X_t$ if the packet is received and $Y_t = \mathfrak{E}$ if the packet is dropped. Packet reception is acknowledged, so the sensor knows Y_t with one unit delay. At each stage, the receiver generates an estimate $\hat{X}_t$ as follows. $\hat{X}_0$ is 0 and for $t > 0$,

$$\hat{X}_t = \begin{cases} a\hat{X}_{t-1}, & \text{if } Y_t = \mathfrak{E} \\ Y_t, & \text{if } Y_t \neq \mathfrak{E}. \end{cases}$$

Under some conditions, such an estimation rule is known to be optimal [9, 10, 12, 13, 112].

There are two types of costs: (i) a communication cost $\lambda(U_t)$, where λ is an increasing function with $\lambda(0) = 0$; and (ii) an estimation cost $d(X_t - \hat{X}_t)$, where d is an even and quasi-convex function with $d(0) = 0$.

Define the error process $\{E_t\}_{t\geq 0}$ as $E_t = X_t - a\hat{X}_{t-1}$. The error process $\{E_t\}_{t\geq 0}$ evolves in a controlled Markov manner as follows:

$$E_{t+1} = \begin{cases} aE_t + W_t, & \text{if } Y_t = \mathfrak{E} \\ W_t, & \text{if } Y_t \neq \mathfrak{E} \end{cases} \tag{6.9}$$

Due to packet acknowledgments, E_t is measurable at the sensor at time t. If a packet is received, then $\hat{X}_t = X_t$ and the estimation cost is 0. If the packet is dropped, $X_t - \hat{X}_t = E_t$ and an estimation cost of $d(E_t)$ is incurred.

The objective is to choose a transmission strategy $f = (f_1, \dots, f_T)$ of the form $U_t = f_t(E_t)$ to minimize

$$\mathbb{E}\bigg[\sum_{t=1}^{T} \big[\lambda(U_t) + (1 - q(U_t)d(E_t)\big]\bigg].$$

The above model is Markov decision process with state $E_t \in \mathbb{X}$, control action $U_t \in [0, u_{\max}]$, per-step cost

$$c(e, u) = \lambda(u) + (1 - q(u))d(e), \tag{6.10}$$

and transition density/mass function

$$p(e_+|e; u) = q(u)\varphi(e_+) + (1 - q(u))\varphi(e_+ - ae). \tag{6.11}$$

For ease of reference, we restate the assumptions imposed on the cost:

(M0) $q(0) = 0$ and $q(u_{\max}) \leq 1$.

(M1) $\lambda(\cdot)$ is increasing with $\lambda(0) = 0$.

(M2) $q(\cdot)$ is increasing.

(M3) $d(\cdot)$ is even and quasi-convex with $d(0) = 0$.

In addition, we impose the following assumptions on the probability density/mass function of the i.i.d. process $\{W_t\}_{t\geq 1}$:

(M4) $\varphi(\cdot)$ is even.

(M5) $\varphi(\cdot)$ is unimodal (i.e., quasi-concave).

Claim 1 We have the following:

1. under assumptions (M0) and (M3), the per step cost function given by (6.10) satisfies (C1).

2. under assumptions (M0), (M2) and (M3), the per step cost function given by (6.10) satisfies (C4).

3. under assumption (M4), the transition density $p(u)$ given by (6.11) satisfies (C2).

4. under assumptions (M0), (M2), (M4) and (M5), the transition density $p(u)$ satisfies (C3) and (C5).

The proof is given in Appendix E.1.

An immediate consequence of Theorem 6.1.1 and Claim 1 is the following:

Theorem 6.5.1 *Under assumptions* (M0), (M2)–(M4), *the value function for the remote estimation model is even and quasi-convex. Under the additional assumption* (M5), *the optimal strategy is also even and quasi-convex.*

Remark 19 Note that the result does not depend on (M1). This is for the following reason. Suppose there are two power levels $u_1, u_2 \in [0, u_{\max}]$ such that $u_1 < u_2$ but $\lambda(u_1) \geq \lambda(u_2)$, then for any $e \in \mathbb{X}$, $c(e, u_1) \geq c(e, u_2)$. Thus, action u_1 is dominated by action u_2 and is, therefore, never optimal and can be eliminated.

Remark 20 Although Theorem 6.5.1 is derived for continuous action space, it is also true when the action space is a discrete set. In particular, if we take the action space to be $\{0, 1\}$ and $q(1) = 1$, we get the results of [9, Theorem 1], [26, Proposition 1], [10, Theorem 3], [14, Theorem 1]; if we take the action space to be $\{0, 1\}$ and $q(1) = \varepsilon$, we get the result of [13, Theorem 1], [110, Theorem 2].

6.6 Conclusion

In this chapter we consider a Markov decision process with continuous or discrete state and action spaces and analyze the monotonicity of the optimal solutions. In particular, we identify sufficient conditions under which the value function and the optimal strategy are even and quasi-convex. The proof relies on a folded representation of the Markov decision process and uses stochastic monotonicity and submodularity. We present an example of optimal power allocation in remote estimation and show that the sufficient conditions are easily verified.

This page is intentionally left blank.

Chapter 7

Conclusion and future directions

7.1 Summary of the results presented in this thesis

In this thesis we investigated the dynamic team (i.e. team with decentralized control) that appears in a remote estimation problem. We restrict our attention to two-agent system—with one transmitter and one estimator–connected by an erasure communication channel which is either one-way and noiseless or two-way with noisy forward and noiseless feedback. The noise in the channel is considered to take place in the form of loss of packets. For the ideal channel and an erasure channel with and without memory, we have investigated the structure of the optimal communication strategy and characterized optimal performance in the context of costly or constrained communicationthe, where the constraint is on the expected number of transmissions. The Lagrange relaxation of the constrained problem is posed as a costly communication. For an arbitrary but fixed estimation strategy, we use a person-by-person approach to find the structure of the best performing transmitter. The estimator solves a filtering problem. Although in general the structure of the estimation strategy depends on the transmission strategy, for a stylized model with some simplifying assumptions, the structure of the optimal strategy can be computed irrespective of the transmission strategy. This trick converts the decentralized problem into an equivalent centralized problem with only one decision maker–the transmitter. The dynamic program for the infinite horizon optimization problem is formulated, which is similar to that of a POMDP, with the minimization taken over a functional space.

For a stylized model with some simplifying assumptions, we further simplify the computation of the optimal performance. It is argued that some sufficient conditions such as symmetry in the per-step cost function and the distribution of the process noise along with submodularity and stochastic monotonicity, the value function and the optimal strategy are even and quasi-convex. Thus, the optimality of a class of simple transmission strategies called the *threshold-based* strategies is established. The optimal estimator is shown to be Kalman-like.

After establishing the structures of the optimal communication strategies, the focus of this thesis centers around characterizing completely the optimal costly and constrained performances. By a change of variable, we proceed our discussion with a regenerative process called the *error process*. This enables us to compute certain parameters until the time of reset and apply *Renewal Theory* to compute the performance a generic threshold-based strategy. Computation of the optimal thresholds involve the submodularity property of the costly performance and some sufficient conditions for the optimality of a constrained performance.

7.1.1 Discussions on the generality of the results

The structural results in the RE problem with ideal and erasure channel have been previously established on a case-by-case basis. For example, [9] takes into account the first-order autoregressive process with Gaussian noise and finite horizon. [10] proves the optimality results along with a dynamic programming decomposition for finite horizon and discrete as well as continuous state space. Both of these works consider ideal communication channel. In contrast, i.i.d. packet drops in the channel in long-term average setup is considered in [12]. In this thesis, we have tried to provide a unified framework to discuss the optimality results for different scenarios occurring in the RE problem, which is summarized below.

- **Discrete and continuous state space:** We investigate the optimization problems for both discrete state space (we call it Model A in Chapter 2) and continuous state space (we call it Model B in Chapter 2). Although most of the results are generated for a first-order autoregressive Markov process, we talk about the structural results for a generic communication model without the stylized structure (for example, see Chapter 3.3).

- **Symmetry and monotonicity in optimal solutions:** We analyze the sufficient conditions for the value function and the optimal strategies to be even and quasi-convex in states. This leads to a class of optimal solutions which are of threshold-type. In Chapter 6 we have generalized the binary action-space (whether to transmit or not) that occurs in a remote estimation problem to a finite (for discrete actions) or a compact (for continuous actions) action-space and elucidate the results with an example of the power-allocation in remote estimation.

- **Infinite horizon:** In most of the thesis, we have considered the infinite horizon optimization problem, where we consider both the *discounted* and the *long-term average* cases. Under certain technical conditions to show that effectively the state space is compact and the distortion is bounded (for e.g. see Chapters 2–3), the structural results obtained with finite horizon setup can be extended to the infinite horizon case. The value function in infinite horizon is a contraction and consequently the optimal strategies are time-homogeneous. We have unified the results for the long-term average setup and those for the discounted case by invoking the *vanishing discount* approach.

- **Costly and constrained optimization:** In this research we have addressed the constrained optimization problem, where the constraint is on the expected number of transmissions. In order to find the optimal solutions, we first solved its Lagrange relaxation, which we call the costly optimization problem.

- **Erasure communication channel:** We consider a one-way noiseless channel (Chapter 2) as well as a two-way channel which is noisy in the forward path and the feedback is noiseless (Chapter 3). The noise in the channel is in terms of packet drop. Instead of a binary communication channel, we consider a more generic version of *erasure* channel, where the input alphabet to the channel is the observation made by the transmitter/sensor, and the output of the channel is the input alphabet or an erasure symbol, which is the input to the estimator/receiver.

- **Learning in RE:** In this work, in the context of the numerical methods to find a global optimal solution of a remote estimation problem, we explore the applicability of learning-based methods (stochastic approximation). The limitations in terms of computational complexity that arises in the analytical formulation triggered our curiosity to delve into the domain of the numerical methods to find globally optimal strategy. We focus on certain stochastic gradient methods that yield satisfactory results for one-dimensional problem. In case of Markov erasure channels, the need for a proper interpretation of optimality results in two-dimensional space is required. In one dimensional problem, the optimality of the threshold-based strategies is established. In higher dimension, although the optimality of threshold-based strategies is not known, it is perhaps still worth investigating the performance of the best threshold-based strategy due to the simplicity of implementation.This motivated us to investigate stochastic gradient algorithms using simultaneous perturbation methods, which has its merits in the face of curse of dimensionality. In order to shed some light on the effectiveness of such algorithms, we compare the results yielded by two such algorithms. In all simulations, we have integrated a variant of Monte Carlo simulation, which we call the *Renewal Monte Carlo* simulation, with the renewal relationships found in Chapters 2–3. As discussed in Chapter 4, for the discounted case, this method has the advantage of circumventing numerical issues arising with very small values of the β^t for a large value of t.

7.2 Some discussion on the results

7.2.1 Comments on the assumptions

The main assumptions that enable us to get an optimal estimator that is indifferent to transmission strategies and to analyze the optimality of the threshold-based strategy are as follows:

- **(A1) Unimodality and symmetry in the state-dynamics:** We assume that the process noise is unimodal and symmetric. This results in the pre- and post-transmission beliefs of the estimator to be symmetric and unimodal, which is crucial for our analysis, as this brings about the stochastic monotonic distributions and plays the key role in the structure of the optimal estimation strategy.

- **(A2) Even and quasi-convexity of the cost function:** We assume that the per-step cost function, which involves the transmission cost and the distortion due to estimation, to be even and quasi-convex in the state. This, together with the stochastic monotonicity of the state process, leads to the evenness and quasi-convexity of the value function.

- **(A3) The parameter a:** We assume that the state-dynamic parameter a is known to the agents.

(A1) is fairly mild since we are not restricting to any particular distribution (it could be of finite or infinite support). At the same time, it includes distributions, which occur in practice, e.g. Gaussian, Cauchy distribution etc. (A2) is a reasonable assumption since in most scenarios we consider the distortion to be an even function of the error. We assume (A3) for simplicity of the analysis. If a is not known to the agents, then the transmitter sends its estimate of a along with its observation of the state and the estimator uses this input to generate its output. The rest of the results remain the same.

7.2.2 Some issues relevant to RE

Given below is a brief synopsis of some aspects that might be interesting to analyze, which we left out in our research. These lay the course for future research in the context of RE.

- **Partial observation of the source symbol by the transmitter.** In this thesis we assume that the transmitter fully observes the state process and takes a decision of whether or not to transmit. This is a fairly reasonable in the context of battery-powered transmitters fully observing the source and sending data-packet over a packet-switched network, where the size of data-packet is cheap compared to the transmission of the data. When the transmitter does not observe the source completely, it can generate its own estimate of X_t based on its past decisions (solving a filtering problem), and send that estimate to the receiver. The rest of the analysis remains same for the optimal communication strategies.

- **On the noise present in the channel.** In this thesis we assume that the noise present in the forward path of the channel (i.e., from the transmitter to the estimator) is due to the loss of packets. There has been few recent results on the presence of noise in the communication channel. In the finite horizon setup, [113] shows the optimization results for a communication channel with i.i.d. noise with Gamma distribution. The transmitter transmits its observation of the source symbol X_t and the sign of the source symbol as a side-information to the receiver. It would be an interesting extension to our current framework to add noise in addition to the packet drop in the channel and to analyze the optimality results.

We assume that the feedback (ACK/NACK) from the receiver is noiseless. The situation of noisy feedback is tricky in the sense that although we can evaluate the performance of the noisy feedback in the same light of the noisy forward channel, it is not straightforward to establish the optimality results.

- **Quantization of the source.** In this thesis we assume that the transmission of the source is much costlier compared to the sensing and the size of the data-packet. This situation is commensurate with the applications in networked systems, where the transmitters are often battery-powered devices sending its observations over a packet-switched network to a remotely placed estimator. However, if the size of data-packet is also costly, the transmitter needs to quantize the data. Such a variant of the remote estimation problem can be taken care of by *vector quantization* of the source as a part of a *lossy source coding* and use our scheme to to come up with an approximation of the source symbol.

7.3 Future directions

We believe that there are several possible directions which may lead to interesting results in the framework of the optimality of the monotone strategies, the threshold-based strategy being one of them. In the current framework with two-agent dynamic team problem, the effects of the scenarios mentioned in Section 7.2.2 will be quite interesting.

A straightforward extension of the first-order autoregressive problems that we discussed in this thesis would be to consider a controlled state dynamics. For a class of problems with such a dynamics, such as *inventory control*, the optimality of the threshold-based strategies is established in literature [114]. We can use the SA approach developed in this thesis to compute the optimal threshold and optimal performances.

It would be interesting to utilize the framework for analyzing the structure of the optimal strategies and the optimal performance developed in the current work to control problems where there is a cost of communication along with the classical certainty-equivalent optimal controller (as in a classical Linear Quadratic Gaussian (LQG) control problem). The fundamental trade-off lies between the controller performance and the communication cost. It would be worthwhile to investigate the optimal solutions for one-step delayed sharing when there is a cost pertaining to sharing of the state.

Another way open for research is the extension of these results in a multidimensional framework. When the source symbols are vector-valued, the performances of threshold-based transmission strategies along with Kalman-like estimator can be evaluated. Such an extension could be worth investigating due to the simple implementation of the threshold-based strategies. Although the optimality of such strategies is not yet established, finding the best performing threshold-based strategy has the potential of finding significance practical applicability. For a non-diagonal system matrix in the vector-valued state dynamics, the transmission strategy faces the thresholds embedded in an Euclidean space (of dimension higher than one). The notion of monotonicity is trickier in such a scenario, since total ordering is not defined in higher dimensional Euclidean space. One needs to define a proper notion of ordering in order to define the optimal threshold.

Last but not the least, the scope of learning in remote estimation problem seems to have its own merits. In an even more generic framework, for a broader class of dynamic team problems having the state evolution with controlled restarts (which turns the *error* process regenerative), the application of RMC may prove to be advantageous over naive Monte Carlo or the temporal difference methods, which tend to have high variance and high bias respectively. With some additional knowledge of the symmetry and monotonicity property of the state dynamics and the optimal strategies, one could successfully employ functional approximation in the approximate dynamic programming. This aspect of research increases the feasibility of finding the optimal solutions with fewer assumptions on the source model and when the analytical computations become too expensive (e.g. the difficulty with continuous state processes that is mentioned in Chapter 4). Furthermore, this area of research opens the scope of analyzing the implication in general decentralized control problem with big data, where the decision of sharing the state of the source with the controller becomes harder due to high processing cost.

This page is intentionally left blank.

Appendix A

Proofs of Chapter 2

A.1 Proof of the structural results

The results of [10] relied on the notion of ASU (almost symmetric and unimodal) distributions introduced in [115].

Definition A.1.1 (Almost symmetric and unimodal distribution) *A probability distribution μ on $\mathbb{Z}$ is almost symmetric and unimodal (ASU) about a point $a \in \mathbb{Z}$ if for every $n \in \mathbb{Z}_{\geq 0}$,*

$$\mu_{a+n} \geq \mu_{a-n} \geq \mu_{a+n+1}.$$

A probability distribution that is ASU around 0 and even (i.e., $\mu_n = \mu_{-n}$) is called ASU and even. Note that the definition of ASU and even is equivalent to even and decreasing on $\mathbb{Z}_{\geq 0}$.

Definition A.1.2 (ASU Rearrangement) *The ASU rearrangement of a probability distribution μ, denoted by μ^+, is a permutation of μ such that for every $n \in \mathbb{Z}_{\geq 0}$,*

$$\mu^+_n \geq \mu^+_{-n} \geq \mu^+_{n+1}.$$

We now introduce the notion of majorization for distributions supported over $\mathbb{Z}$, as defined in [50].

Definition A.1.3 (Majorization) *Let μ and ν be two probability distributions defined over $\mathbb{Z}$. Then μ is said to majorize ν, which is denoted by $\mu \succeq_m \nu$, if for all $n \in \mathbb{Z}_{\geq 0}$,*

$$\sum_{i=-n}^{n} \mu_i^+ \leq \sum_{i=-n}^{n} \nu_i^+,$$
$$\sum_{i=-n}^{n+1} \mu_i^+ \leq \sum_{i=-n}^{n+1} \nu_i^+.$$

The structure of optimal estimator in Theorem 2.6.1 were proved in two steps in [10]. The first step relied on the following two results.

Lemma A.1.1 *Let μ and ν be probability distributions with finite support defined over $\mathbb{Z}$. If μ is ASU and even and ν is ASU about a, then the convolution $\mu * \nu$ is ASU about a.*

Lemma A.1.2 *Let μ, ν, and ξ be probability distributions with finite support defined over $\mathbb{Z}$. If μ is ASU and even, ν is ASU, and ξ is arbitrary, then $\nu \succeq_m \xi$ implies that $\mu * \nu \succeq_m \mu * \xi$.*

These results were originally proved in [115] and were stated as Lemmas 5 and 6 in [10].

The second step (in the proof of structure of optimal estimator in Theorem 2.6.1) in [10] relied on the following result.

Lemma A.1.3 *Let μ be a probability distribution with finite support defined over $\mathbb{Z}$ and $f\colon \mathbb{Z} \to \mathbb{R}_{\geq 0}$. Then,*

$$\sum_{n=-\infty}^{\infty} f(n)\mu_n \leq \sum_{n=-\infty}^{\infty} f^+(n)\mu_n^+.$$

We generalize the results of Lemmas A.1.1, A.1.2, and A.1.3 to distributions over $\mathbb{Z}$ with possibly countable support. With these generalizations, we can follow the same two-step approach of [10] to prove the structure of optimal estimator as given in Theorem 2.6.1.

The structure of optimal transmitter in Theorem 2.6.1 in [10] only relied on the structure of optimal estimator. The exact same proof works in our model as well.

A.1.1 Generalization of Lemma A.1.1 to distributions supported over $\mathbb{Z}$

The proof argument is similar to that presented in [115, Lemma 6.2]. We first prove the results for $a = 0$. Assume that ν is ASU and even. For any $n \in \mathbb{Z}_{\geq 0}$, let $r^{(n)}$ denote the rectangular function from $-n$ to n, i.e.,

$$r^{(n)}(e) = \begin{cases} 1, & \text{if } |e| \leq n, \\ 0, & \text{otherwise.} \end{cases}$$

Note that any ASU and even distribution μ may be written as a sum of rectangular functions as follows:

$$\mu = \sum_{n=0}^{\infty} (\mu_n - \mu_{n+1}) r^{(n)}.$$

It should be noted that $\mu_n - \mu_{n+1} \geq 0$ because μ is ASU and even. ν may also be written in a similar form.

The convolution of any two rectangular functions $r^{(n)}$ and $r^{(m)}$ is ASU and even. Therefore, by the distributive property of convolution, the convolution of μ and ν is also ASU and even.

The proof for the general $a \in \mathbb{Z}$ follows from the following facts:

1. Shifting a distribution is equivalent to convolution with a shifted delta function.

2. Convolution is commutative and associative.

A.1.2 Generalization of Lemma A.1.2 to distributions supported over $\mathbb{Z}$

We follow the proof idea of [50, Theorem II.1]. For any probability distribution μ, we can find distinct indices i_j, $|j| \leq n$ such that $\mu(i_j)$, $|j| \leq n$, are the $2n+1$ largest values of μ. Define

$$\mu_n(i_j) = \mu(i_j),$$

for $|j| \leq n$ and 0 otherwise. Clearly, $\mu_n \uparrow \mu$ and if μ is ASU and even, so is μ_n.

Now consider the distributions μ, ν, and ξ from Lemma A.1.2 but without the restriction that they have finite support. For every $n \in \mathbb{Z}_{\geq 0}$, define μ_n, ν_n, and ξ_n as above. Note that all distributions have finite support and μ_n is ASU and even and ν_n is ASU. Furthermore, since the definition of majorization remain unaffected by truncation described above, $\nu_n \succeq_m \xi_n$. Therefore, by Lemma A.1.2,

$$\mu_n * \nu_n \succeq_m \mu_n * \xi_n.$$

By taking limit over n and using the monotone convergence theorem, we get

$$\mu * \nu \succeq_m \mu * \xi.$$

A.1.3 Generalization of Lemma A.1.3 to distributions supported over $\mathbb{Z}$

This is an immediate consequence of [50, Theorem II.1].

A.2 Proof of Lemma 2.4.1

Let $\|\cdot\|_\infty$ denote the sup-norm, i.e., for any $v : S^{(k)} \to \mathbb{R}$,

$$\|v\|_\infty = \sup_{e \in S^{(k)}} |v(e)|.$$

To prove the lemma, let us first prove the following:

Lemma A.2.1 *For $\beta \in (0,1)$, for both Models A and B, the operator $\beta\mathcal{B}^{(k)}$ is a contraction, i.e., for any $v : S^{(k)} \to \mathbb{R}$,*

$$\|\beta\mathcal{B}^{(k)} v\|_\infty \leq \beta \|v\|_\infty.$$

Thus, for any bounded $h : S^{(k)} \to \mathbb{R}$, the equation

$$v = h + \beta\mathcal{B}^{(k)} v \tag{A.1}$$

has a unique bounded solution v. In addition, if h is continuous, then v is continuous.

Proof We state the proof for Model B. The proof for Model A is similar. By the definition of sup-norm, we have that for any bounded v

$$\begin{aligned}
\|\beta\mathcal{B}^{(k)} v\|_\infty &= \beta \sup_{e \in (-k,k)} \int_{-k}^{k} \phi(w - ae) v(w) dw \\
&\leq \beta \sup_{e \in (-k,k)} \|v\|_\infty \int_{-k}^{k} \phi(w - ae) dw \\
&\leq \beta \|v\|_\infty, \quad \text{(since } \phi \text{ is a pdf).}
\end{aligned}$$

Hence, $\beta\mathcal{B}^{(k)}$ is a contraction.

Now, consider the operator $\mathcal{B}'$ given as: $\mathcal{B}'v = h + \beta\mathcal{B}^{(k)}v$. Then we have,

$$\|\mathcal{B}'(v_1 - v_2)\|_\infty = \beta\|\mathcal{B}^{(k)}(v_1 - v_2)\|_\infty \leq \beta\|v_1 - v_2\|_\infty.$$

Since $\beta \in (0,1)$ and the space of bounded real-valued functions is complete, by Banach fixed point theorem, $\mathcal{B}'$ has a unique fixed point.

If h is continuous, we can define $\mathcal{B}^{(k)}$ and $\mathcal{B}'$ as operators on the space of continuous and bounded real-valued function (which is complete). Hence, the continuity of the fixed point follows also from Banach fixed point theorem.

A.2.1 Proof of (b) of Lemma 2.4.1

Note that for any bounded v, $\|\mathcal{B}^{(k)}v\|_\infty$ is bounded and increasing in k. We show that $L_\beta^{(k)}(e)$ is continuous and differentiable in k. Similar argument holds for $M_\beta^{(k)}(e)$.

We show the differentiability in k. Continuity follows from the fact that differentiable functions are continuous. Note that $L_\beta^{(k)}(e)$ and $M_\beta^{(k)}(e)$ are even functions of e. Now, for any $\varepsilon > 0$ we have

$$\begin{aligned}
&L_\beta^{(k+\varepsilon)}(e) - L_\beta^{(k)}(e) \\
&\quad = \beta\int_{-k}^{k}\phi(w-ae)[L_\beta^{(k+\varepsilon)}(w) - L_\beta^{(k)}(w)]dw + 2\beta\int_{k}^{k+\varepsilon}\phi(w-ae)L_\beta^{(k+\varepsilon)}(w)dw \\
&\quad = \beta\int_{-k}^{k}\phi(w-ae)[L_\beta^{(k+\varepsilon)}(w) - L_\beta^{(k)}(w)]dw + 2\beta\phi(k-ae)L_\beta^{(k+\varepsilon)}(k+\varepsilon)\varepsilon + O(\varepsilon^2)
\end{aligned}$$

Let $R_\beta^{(k)}(e,w;a)$ be the resolvent of ϕ, as given in (16). Then,

$$L_\beta^{(k+\varepsilon)}(e) - L_\beta^{(k)}(e) = 2\beta\int_{-k}^{k}R_\beta^{(k)}(e,w;a)\phi(k-ae)L_\beta^{(k+\varepsilon)}(w)\varepsilon dw + O(\varepsilon^2)$$

This implies that

$$\left|\ \frac{L_\beta^{(k+\varepsilon)}(e) - L_\beta^{(k)}(e)}{\varepsilon}\ \right| \leq 2\|\phi\|_\infty\|L_\beta^{(k)}\|_\infty\left|\int_{-k}^{k}\beta R_\beta^{(k)}(e,w;a)dw\right| + O(\varepsilon).$$

Since $\beta\mathcal{B}^{(k)}$ is a contraction, the value of the integral in the first term on the right hand side of the above inequality is less than 1 and the result follows from the definition of differtiability.

Proof of Lemma 2.4.1

The solutions of equations (2.11) and (2.12) exist due to Lemma A.2.1.

(a) Consider k, $l \in \mathbb{X}_{\geq 0}$ such that $k < l$. A sample path starting from $e \in S^{(k)}$ must escape $S^{(k)}$ before it escapes $S^{(l)}$. Thus $L_\beta^{(l)}(e) \geq L_\beta^{(k)}(e)$. In addition, the above inequality is strict because W_t has a unimodal distribution. Similar argument holds for $M_\beta^{(k)}$.

(b) The continuity and differentiability is shown in Section A.2.1.

(c) The limit holds since $L_\beta^{(k)}(e)$ and $M_\beta^{(k)}(e)$ are continuous functions of β.

A.3 Proof of Proposition 2.4.1

1. $C_\beta^{(l)}(0;\lambda) - C_\beta^{(k)}(0;\lambda) = (D_\beta^{(l)}(0) - D_\beta^{(k)}(0)) - \lambda(N_\beta^{(k)}(0) - N_\beta^{(l)}(0))$. By Lemma **??** and Theorem 2.4.2, $N_\beta^{(k)}(0) - N_\beta^{(l)}(0)$ is positive, hence $C_\beta^{(l)}(0;\lambda) - C_\beta^{(k)}(0;\lambda)$ is decreasing in λ. Hence $C_\beta^{(k)}(0;\lambda)$ is submodular.

2. Note that $k^*_\beta(\lambda) = \arg\inf_{k\geq 0} C_\beta^{(k)}(0;\lambda)$ can take a value ∞ (which corresponds to the strategy 'never communicate'). Thus, the domain of k is $\mathbb{X}_{\geq 0} \cup \{\infty\}$, which is compact. Hence, by [116, Theorem 2.8.2], k^*_β is increasing in λ.

A.4 Proofs of Propositions 2.6.1 and 2.6.2

We prove the results for Model A when the horizon T is finite. The results then follow by taking limits as $T \to \infty$. The proofs for Model B are almost identical.

The value function for the finite horizon setup for $\beta \in (0,1]$is given by $V_{\beta,T+1} = 0$ and for $t = T, \cdots, 1$

$$V_{\beta,t}(e;\lambda) = \min\Big\{(1-\beta)\lambda + \beta \sum_{n=-\infty}^{\infty} p_n V_{\beta,t+1}(n;\lambda),$$
$$(1-\beta)d(e) + \beta \sum_{n=-\infty}^{\infty} p_{n-ae} V_{\beta,t+1}(n;\lambda)\Big\}. \tag{A.2}$$

The value functions $V_t^{(+)}$ and $V_t^{(-)}$ are defined similarly.

For ease of notation, we drop β and λ in the rest of the discussion in this Appendix.

Lemma A.4.1 *The value functions $V_t(\cdot)$, $V_t^{(+)}(\cdot)$ and $V_t^{(-)}(\cdot)$ are even.*

Proof For all $a \in \mathbb{X}$, the per-step costs $d(e)$ and λ are even and the transition probabilities $P_{en}(0) = p_{n-ae}$ and $P_{en}(1) = p_n$ satisfy $P_{en}(u) = P_{(-e)(-n)}(u)$ for $u \in \{0,1\}$. Therefore, $V_t(e)$ is even [117, Theorem 1]. A similar argument holds for $V_t^{(+)}(e)$ and $V_t^{(-)}(e)$.

Lemma A.4.2 *For the finite horizon setup, $V_t^{(+)}(e) = V_t^{(-)}(e)$.*

Proof We prove the result by backward induction. The result is trivially true for $T+1$ as $V_{T+1}^{(+)}(e) = V_{T+1}^{(-)}(e) = 0$, which forms the basis of the induction. Assume $V_{t+1}^{(+)}(e) = V_{t+1}^{(-)}(e)$ for all $e \in \mathbb{X}$. Define

$$\hat{V}_t^{(+)}(e) = \sum_{n=-\infty}^{\infty} p_{n-ae} V_{t+1}^{(+)}(n), \quad \hat{V}_t^{(-)}(e) = \sum_{n=-\infty}^{\infty} p_{n+ae} V_{t+1}^{(-)}(n).$$

Then

$$\hat{V}_t^{(+)}(e) = \sum_{n=-\infty}^{\infty} p_{n-ae} V_{t+1}^{(+)}(n) = \sum_{-n=-\infty}^{\infty} p_{-n-ae} V_{t+1}^{(+)}(-n)$$
$$\overset{(a)}{=} \sum_{n=-\infty}^{\infty} p_{n+ae} V_{t+1}^{(+)}(n) \overset{(b)}{=} \sum_{n=-\infty}^{\infty} p_{n+ae} V_{t+1}^{(-)}(n) = \hat{V}_t^{(-)}(e),$$

where (a) uses p and $V_{t+1}^{(+)}$ are even and (b) uses the induction hypothesis. Substituting this back in the definition of $V_t^{(+)}(e)$ and $V_t^{(-)}(e)$, we get that $V_t^{(+)}(e) = V_t^{(-)}(e)$. Therefore, the result is true by induction.

Lemma A.4.3 *For $m, e \in \mathbb{X}_{\geq 0}$, define*

$$Q(m|e,0) = \sum_{n:|n|\geq m} p_{n-ae} \quad \text{and} \quad Q(m|e,1) = \sum_{n:|n|\geq m} p_n.$$

Then, for all $e, m \in \mathbb{X}_{\geq 0}$ and $a > 0$, $Q(m|e,0)$ and $Q(m|e,1)$ are increasing in e.

We will prove this Lemma later.

Definition A.4.1 *A function $f\colon \mathbb{X} \to \mathbb{R}$ is called even and increasing on $\mathbb{X}_{\geq 0}$ if for all $x \in \mathbb{X}_{\geq 0}$, $f(x) = f(-x)$ and $f(x) \leq f(x+1)$.*

Lemma A.4.4 *The value function $V_t(e)$ is even and increasing on $\mathbb{X}_{\geq 0}$.*

Proof We have already shown that $V_t(e)$ is even. For $a > 0$, the properties described in the proof of Lemma A.4.1 and the statement Lemma A.4.3 imply that $V_t(e)$ is even and increasing as shown in Theorem 6.1.1, Chapter 6. Now, Lemma A.4.2 implies that $V_t(e)$ is also even and increasing for $a < 0$.

Proof (Proofs of Propositions 2.6.1 and 2.6.2) The result follows from Lemmas A.4.2 and A.4.4 by taking the limit $T \to \infty$, since equality is preserved under limits.

Proof (Proof of Lemma A.4.3) $Q(m|e,1)$ is independent of e. Define $R(m|e) = \sum_{n:|n|\leq m} p_{n-e}$. Then, $Q(m|e,0) = 1 - R(m|ae)$. To show $Q(m|e,0)$ is increasing in e, it suffices to show that $R(m|ae) \geq R(m|ae+1)$ (which implies that $R(m|ae) \geq R(m|ae+a)$).

Now consider

$$R(m|ae) - R(m|ae+1) = p_{m-ae} - p_{-m-ae-1} = p_{m-ae} - p_{m+ae+1}.$$

If $m \geq ae$, then $0 \leq m - ae < m + ae + 1$, hence, $p_{m-ae} \geq p_{m+ae+1}$. If $m < ae$, then $0 < ae - m < m + ae + 1$, hence $p_{m-ae} = p_{ae-m} \geq p_{m+ae+1}$. Thus, in both cases, $R(m|ae) \geq R(m|ae+1)$.

A.5 Proof of Part 3) of Lemma 2.7.1

By Lemma 2.4.1, $M_\beta^{(k)}(e)$ is strictly increasing in k; therefore, by Theorem 2.4.2, $N_\beta^{(k)}(e)$ is strictly decreasing in k.

We prove the monotonicity of $D_\beta^{(k)}$ in k for Model A for $\beta \in (0,1)$. The result for $\beta = 1$ follows by taking limit $\beta \uparrow 1$. The result for Model B is similar. Based on Lemma A.4.2, we restrict attention to $a > 0$.

For any $\beta \in (0,1)$ and $k \in \mathbb{Z}_{\geq 0}$, define the operator $\mathcal{T}^{(k)} : (\mathbb{Z} \to \mathbb{R}) \to (\mathbb{Z} \to \mathbb{R})$ as follows. For any $D : \mathbb{Z} \to \mathbb{R}$,

$$[\mathcal{T}^{(k)} D](e) = \begin{cases} \beta[\mathcal{B}D](0), & \text{if } |e| \geq k \\ (1-\beta)d(e) + \beta[\mathcal{B}D](e) & \text{if } |e| < k. \end{cases} \tag{A.3}$$

This operator is the Bellman operator for evaluating strategy $f^{(k)}$. Hence, it is a contraction and $D^{(k)}$ is the unique fixed point of $\mathcal{T}^{(k)}$.

Define $D_\beta^{(k,0)} = D_\beta^{(k)}$, and for $m \in \mathbb{Z}_{>0}$, $D_\beta^{(k,m)} = \mathcal{T}^{(k+1)} D_\beta^{(k,m-1)}$.

From Lemma A.4.3 and [117, Lemma 2], we get that for any $e \in \mathbb{Z}_{\geq 0}$,

$$\sum_{n=-\infty}^{\infty} p_{n-ae} D_\beta^{(k)}(n) \geq \sum_{n=-\infty}^{\infty} p_n D_\beta^{(k)}(n),$$

or equivalently, $[\mathcal{B}D_\beta^{(k)}](e) \geq [\mathcal{B}D_\beta^{(k)}](0)$.

For $|e| = k$, $D_\beta^{(k,1)}(e) = (1-\beta)d(e) + \beta[\mathcal{B}D_\beta^{(k)}](e)$ and $D_\beta^{(k)}(e) = \beta[\mathcal{B}D_\beta^{(k)}](0)$; hence, $D_\beta^{(k,1)}(e) > D_\beta^{(k)}(e)$. For $|e| \neq k$, $D_\beta^{(k,1)}(e) = D_\beta^{(k)}(e)$ because both terms have the same expression. Hence, for all $e \in \mathbb{Z}$,

$$D_\beta^{(k,1)}(e) \geq D_\beta^{(k)}(e), \quad \text{or} \quad D_\beta^{(k,1)} \geq D_\beta^{(k)}.$$

If we apply the operator $\mathcal{T}^{(k+1)}$ to both sides, the monotonicity of $\mathcal{T}^{(k+1)}$ implies that $D_\beta^{(k,2)} \geq D_\beta^{(k,1)} \geq D_\beta^{(k)}$. Proceeding this way, we get that for any $m > 0$,

$$D_\beta^{(k,m)} \geq D_\beta^{(k)}. \tag{A.4}$$

Note that $\lim_{m\to\infty} D_\beta^{(k,m)} = D_\beta^{(k+1)}$, because $D_\beta^{(k+1)}$ is the unique fixed point of the operator $\mathcal{T}^{(k+1)}$. Thus, taking limit $m \to \infty$ in (B.10), we get that $D_\beta^{(k+1)} \geq D_\beta^{(k)}$.

This page is intentionally left blank.

Appendix B

Proofs of Chapter 3

B.1 Proof of Lemma 3.8.1

In the proof we consider $\mathbb{X} = \mathbb{Z}$. The proof for $\mathbb{X} = \mathbb{R}$ follows similar steps. Recall the notations $\mathcal{B}^{(k_0,k_1)}$, $L_\beta^{(k_0,k_1)}$, $M_\beta^{(k_0,k_1)}$ and $K_\beta^{(k_0,k_1)}$ in Chapter 3.8. We prove the result for $L_\beta^{(k_0,k_1)}$. Similar argument holds for $M_\beta^{(k_0,k_1)}$ and $K_\beta^{(k_0,k_1)}$.

Similar to the argument for the value function given in Section 3.7.3 , we can argue that $L_\beta^{(k_0,k_1)}$ is bounded. $\beta\mathcal{B}^{(k_0,k_1)}$ is a contraction operator from the space of bounded continuous functions on integers to itself.

Lemma B.1.1 *For $\beta, Q_{s0} \in (0,1)$ and $k_s \in \mathbb{Z}_{\geq 0}$, $k_s < \infty$, $s \in \{0,1\}$, the operator $\beta\mathcal{B}^{(k_0,k_1)}$ is a contraction.*

Proof

$$\begin{aligned}
\|\beta\mathcal{B}^{(k_0,k_1)}v\|_\infty &= \beta \sup_{z\in\mathbb{Z}\times\{0,1\}} \sum_{z'\in\mathbb{Z}\times\{0,1\}} h_z^{(k_0,k_1)}\tilde{P}_{zz'}v_{z'} \\
&\overset{(a)}{<} \beta \sup_{z\in\mathbb{Z}\times\{0,1\}} \sum_{z'\in\mathbb{Z}\times\{0,1\}} \tilde{P}_{zz'}v_{z'} \\
&\leq \beta\|v\|_\infty\Big(\sup_{z\in\mathbb{Z}\times\{0,1\}} \sum_{z'\in\mathbb{Z}\times\{0,1\}} \tilde{P}_{zz'}\Big) \overset{(b)}{=} \beta\|v\|_\infty,
\end{aligned}$$

where $z = (e,s)$, $z' = (e',s')$. (a) holds as $Q_{s0} < 1$ and $k_s < \infty$ and (b) holds since $\tilde{P}$ is a stochastic matrix.

Thus, by Banach fixed point theorem, (3.64) has a unique fixed point solution. This proves the existence of the unique solution, which is the first part of the lemma.

To prove the monotonicity of the solutions, we first consider the monotonicity in k_0. By a similar argument, the monotonicity in k_1 can be shown. Note that $f^{(m,k_1)} \geq f^{(\ell,k_1)}$ for any $\ell > m$. This implies that one would wait longer to transmit under $f^{(\ell,k_1)}$. Hence, the expected time till the first successful reception under $f^{(m,k_1)}$ is less than that for $f^{(\ell,k_1)}$, leading to a larger expected distortion under $f^{(\ell,k_1)}$ compared to that incurred with $f^{(m,k_1)}$. Thus $L_\beta^{(\ell,k_1)}(e) \geq L_\beta^{(m,k_1)}(e)$. In addition, the above inequality is strict because W_t has a unimodal distribution. This completes the proof.

B.2 Proof of Proposition 3.8.1

We start by recalling the operator $\mathcal{B}^{(k_0,k_1)}$ as introduced in Section 3.8. Note that the error process $\{E_t\}_{t=0}^\infty$ is a controlled Markov process. Therefore, the functions $D_\beta^{(k_0,k_1)}$ and $N_\beta^{(k_0,k_1)}$ may be thought as value functions when strategy $f^{(k_0,k_1)}$ is used. Thus, they satisfy the following fixed point equations: for $\beta \in (0,1)$,

$$D_\beta^{(k_0,k_1)}(e,s) = \begin{cases} Q_{s0}\big((1-\beta)d(e) + \beta[\mathcal{B}^{(k_0,k_1)} D_\beta^{(k_0,k_1)}](e,s)\big), & \text{if } |e| \geq k_s \\ (1-\beta)d(e) + \beta[\mathcal{B}^{(k_0,k_1)} D_\beta^{(k_0,k_1)}](e,s), & \text{if } |e| < k_s, \end{cases} \tag{B.1}$$

$$N_\beta^{(k_0,k_1)}(e,s) = \begin{cases} Q_{s0}\big((1-\beta) + \beta[\mathcal{B}^{(k_0,k_1)} N_\beta^{(k_0,k_1)}](e,s)\big), & \text{if } |e| \geq k_s \\ (1-\beta) + \beta[\mathcal{B}^{(k_0,k_1)} N_\beta^{(k_0,k_1)}](e,s), & \text{if } |e| < k_s \end{cases} \tag{B.2}$$

where $k_s = k(s)$ for $s \in \{0,1\}$. For $k_s > 0$, $s \in \{0,1\}$, let $\tau^{(k_0,k_1)}$ denote the stopping time of first successful reception when the Markov process starting at state $(0,s)$, follows strategy $f^{(k_0,k_1)}$.

Thus we have,

$$L_\beta^{(k_0,k_1)}(0,s) = \mathbb{E}\Big[\sum_{t=0}^{\tau^{(k_0,k_1)}-1} \beta^t d(E_t) \;\Big|\; E_0 = 0, S_0 = s\Big] \tag{B.3}$$

$$M_\beta^{(k_0,k_1)}(0,s) = \mathbb{E}\Big[\sum_{t=0}^{\tau^{(k_0,k_1)}-1} \beta^t \;\Big|\; E_0 = 0, S_0 = s\Big] = \frac{1 - \mathbb{E}[\beta^{\tau^{(k_0,k_1)}} \mid E_0 = 0, S_0 = s]}{1-\beta} \tag{B.4}$$

$$K_\beta^{(k_0,k_1)}(0,s) = \mathbb{E}\Big[\sum_{t=0}^{\tau^{(k_0,k_1)}} \beta^t U_t \;\Big|\; E_0 = 0, S_0 = s\Big] \tag{B.5}$$

$$D_\beta^{(k_0,k_1)}(0,s) = \mathbb{E}\Big[(1-\beta)\sum_{t=0}^{\tau^{(k_0,k_1)}-1} \beta^t d(E_t) + \beta^{\tau^{(k_0,k_1)}} D_\beta^{(k_0,k_1)}(0,S_t) \;\Big|\; E_0 = 0, S_0 = s\Big] \tag{B.6}$$

$$N_\beta^{(k_0,k_1)}(0,s) = \mathbb{E}\Big[(1-\beta)\sum_{t=0}^{\tau^{(k_0,k_1)}} \beta^t U_t + \beta^{\tau^{(k_0,k_1)}} N_\beta^{(k_0,k_1)}(0,S_t) \;\Big|\; E_0 = 0, S_0 = s\Big] \tag{B.7}$$

$$= (1-\beta)K_\beta^{(k_0,k_1)}(0,s) + \mathbb{E}[\beta^{\tau^{(k_0,k_1)}} \;\Big|\; E_0 = 0, S_0 = s] N_\beta^{(k_0,k_1)}(0,s). \tag{B.8}$$

Substituting (B.3) and (B.4) in (B.6) we get

$$D_\beta^{(k_0,k_1)}(0,s) = (1-\beta)L_\beta^{(k_0,k_1)}(0,s) + [1-(1-\beta)M_\beta^{(k_0,k_1)}(0,s)]D_\beta^{(k_0,k_1)}(0,s).$$

Rearranging, we get that

$$D_\beta^{(k_0,k_1)}(0,s) = \frac{L_\beta^{(k_0,k_1)}(0,s)}{M_\beta^{(k_0,k_1)}(0,s)}.$$

Similarly, substituting (B.4) in (B.8) we get

$$N_\beta^{(k_0,k_1)}(0,s) = (1-\beta)K_\beta^{(k_0,k_1)}(0,s) + [1-(1-\beta)M_\beta^{(k_0,k_1)}(0,s)]N_\beta^{(k_0,k_1)}(0,s).$$

Rearranging, we get that

$$N_\beta^{(k_0,k_1)}(0,s) = \frac{K_\beta^{(k_0,k_1)}(0,s)}{M_\beta^{(k_0,k_1)}(0,s)}.$$

The expression for $C_\beta^{(k_0,k_1)}(0,s;\lambda)$ follows from the definition.

B.3 Proof of Lemma 3.9.1

In the proof we consider $\mathbb{X} = \mathbb{Z}$. The proof for $\mathbb{X} = \mathbb{R}$ follows similar steps.

By Lemma 3.8.1 and Proposition 3.8.1, we have that $N_\beta^{(k)}(0)$ is strictly decreasing in k.

To prove the monotonicity of $D_\beta^{(k)}(0)$ in k, we restrict our attention to $a \in \mathbb{Z}_{>0}$. The result holds for any $a \in \mathbb{Z}$ due to [14, Lemma 11].

For any $v : \mathbb{Z} \to \mathbb{R}$, define operator $\mathcal{B}$ as

$$[\mathcal{B}v](e) \coloneqq \sum_{w=-\infty}^{\infty} p_w v(ae + w), \quad \forall e \in \mathbb{Z}.$$

Or, equivalently,

$$[\mathcal{B}v](e) \coloneqq \sum_{n=-\infty}^{\infty} p_{n-ae} v(n), \quad \forall e \in \mathbb{Z}.$$

For any $\beta \in (0,1)$, $\varepsilon \in (0,1)$ and $k \in \mathbb{Z}_{\geq 0}$, define the operator $\mathcal{T}^{(k)} : (\mathbb{Z} \to \mathbb{R}) \to (\mathbb{Z} \to \mathbb{R})$ as follows. For any $D : \mathbb{Z} \to \mathbb{R}$,

$$[\mathcal{T}^{(k)}D](e) = \begin{cases} \varepsilon\big((1-\beta)d(e) + \beta[\mathcal{B}D](e)\big), & \text{if } |e| \geq k \\ (1-\beta)d(e) + \beta[\mathcal{B}D](e) & \text{if } |e| < k. \end{cases} \tag{B.9}$$

This operator is the Bellman operator for evaluating strategy $f^{(k)}$. Hence, it is a contraction and D is the unique fixed point of $\mathcal{T}^{(k)}$.

Define $D_\beta^{(k,0)} = D_\beta^{(k)}$, and for $m \in \mathbb{Z}_{>0}$, $D_\beta^{(k,m)} = \mathcal{T}^{(k+1)} D_\beta^{(k,m-1)}$.

For $|e| = k$, $D_\beta^{(k,1)}(e) = (1-\beta)d(e) + \beta[\mathcal{B}D_\beta^{(k)}](e)$ and $D_\beta^{(k)}(e) = \varepsilon((1-\beta)d(e) + \beta[\mathcal{B}D_\beta^{(k)}](e))$; hence, $D_\beta^{(k,1)}(e) > D_\beta^{(k)}(e)$, since $\varepsilon \in (0,1)$. For $|e| \neq k$, $D_\beta^{(k,1)}(e) = D_\beta^{(k)}(e)$ because both terms have the same expression. Hence, for all $e \in \mathbb{Z}$,

$$D_\beta^{(k,1)}(e) \geq D_\beta^{(k)}(e), \quad \text{or} \quad D_\beta^{(k,1)} \geq D_\beta^{(k)}.$$

If we apply the operator $\mathcal{T}^{(k+1)}$ to both sides, the monotonicity of $\mathcal{T}^{(k+1)}$ implies that $D_\beta^{(k,2)} \geq D_\beta^{(k,1)} \geq D_\beta^{(k)}$. Proceeding this way, we get that for any $m > 0$,

$$D_\beta^{(k,m)} \geq D_\beta^{(k)}. \tag{B.10}$$

Note that $\lim_{m\to\infty} D_\beta^{(k,m)} = D_\beta^{(k+1)}$, because $D_\beta^{(k+1)}$ is the unique fixed point of the operator $\mathcal{T}^{(k+1)}$. Thus, taking limit $m \to \infty$ in (B.10), we get that $D_\beta^{(k+1)} \geq D_\beta^{(k)}$.

B.4 Proofs of Theorems 3.9.1 and 3.9.2

B.4.1 Proof of Theorem 3.9.1

We first characterize the structure of $C_\beta^{(k)}(0;\lambda)$.

Proposition B.4.1 *For the model given* (3.15)*,*

1. *$C_\beta^{(k)}(0;\lambda)$ is submodular in (k,λ), i.e., for $l > k$, $C_\beta^{(l)}(0;\lambda) - C_\beta^{(k)}(0;\lambda)$ is decreasing in λ.*

2. *Let $k_\beta^*(\lambda) = \arg\inf_{k\geq 0} C_\beta^{(k)}(0;\lambda)$ be the optimal k for a fixed λ. Then $k_\beta^*(\lambda)$ is increasing in λ.*

Proof

1. $C_\beta^{(l)}(0;\lambda) - C_\beta^{(k)}(0;\lambda) = (D_\beta^{(l)}(0) - D_\beta^{(k)}(0)) - \lambda(N_\beta^{(k)}(0) - N_\beta^{(l)}(0))$. By Lemma 3.9.1, $N_\beta^{(k)}(0) - N_\beta^{(l)}(0)$ is positive, hence $C_\beta^{(l)}(0;\lambda) - C_\beta^{(k)}(0;\lambda)$ is decreasing in λ. Hence $C_\beta^{(k)}(0;\lambda)$ is submodular.

2. Note that $k_\beta^*(\lambda) = \arg\inf_{k\geq 0} C_\beta^{(k)}(0;\lambda)$ can take a value ∞ (which corresponds to the strategy 'never communicate'). Thus, the domain of k is $\mathbb{X}_{\geq 0} \cup \{\infty\}$, which is compact. Hence, by [116, Theorem 2.8.2], k_β^* is increasing in λ.

By Proposition B.4.1, $k_\beta^*(\lambda) = \arg\inf_{k\geq 0} C_\beta^{(k)}(0;\lambda)$ is increasing in λ. Let $\mathbb{K}$ denote the set of all possible values of $k_\beta^*(\lambda)$. Since k is integer-valued, the plot of k_β^* vs λ must be a staircase function. In particular, there exists an increasing sequence $\{\lambda_\beta^{(k_n)}\}_{k_n\in\mathbb{K}}$ such that for $\lambda \in (\lambda_\beta^{(k_{n-1})}, \lambda_\beta^{(k_n)}]$, $k_\beta^*(\lambda) = k_n$. We will show that for any k_n,

$$C_\beta^{(k_n)}(0;\lambda_\beta^{(k_n)}) = C_\beta^{(k_{n+1})}(0;\lambda_\beta^{(k_n)}),$$

or, equivalently,

$$\frac{L_\beta^{(k_n)}(0) + \lambda_\beta^{(k_n)} K_\beta^{(k_n)}(0)}{M_\beta^{(k_n)}(0)} = \frac{L_\beta^{(k_{n+1})}(0) + \lambda_\beta^{(k_n)} K_\beta^{(k_{n+1})}(0)}{M_\beta^{(k_{n+1})}(0)}, \tag{B.11}$$

from which we get that $\lambda_\beta^{(k_n)}$ is given by (3.68) by using the relations given in Proposition 3.8.1.

Proof of (B.11)

For any $\lambda \in (\lambda_\beta^{(k_{n-1})}, \lambda_\beta^{(k_n)}]$, $C_\beta^{(k_n)}(0;\lambda) \le C_\beta^{(k_{n+1})}(0;\lambda)$. In particular, for $\lambda = \lambda_\beta^{(k_n)}$,

$$C_\beta^{(k_n)}(0;\lambda_\beta^{(k_n)}) \le C_\beta^{(k_{n+1})}(0;\lambda_\beta^{(k_n)}). \tag{B.12}$$

Similarly, for any $\lambda \in (\lambda_\beta^{(k_n)}, \lambda_\beta^{(k_{n+1})}]$, $C_\beta^{(k_{n+1})}(0;\lambda) \le C_\beta^{(k_n)}(0;\lambda)$. Since both terms are continuous in λ, taking limit as $\lambda \downarrow \lambda_\beta^{(k_n)}$, we get

$$C_\beta^{(k_{n+1})}(0;\lambda_\beta^{(k_n)}) \le C_\beta^{(k_n)}(0;\lambda_\beta^{(k_n)}). \tag{B.13}$$

Eq. (B.11) follows from combining (B.12) and (B.13).

Proof of Part 1)

By definition of $\lambda_\beta^{(k_n)}$, the strategy $f^{(k_n)}$ is optimal for $\lambda \in (\lambda_\beta^{(k_{n-1})}, \lambda_\beta^{(k_n)}]$.

Proof of Part 2)

Recall $C_\beta^*(\lambda) = \inf_{k\ge 0} C_\beta^{(k)}(0;\lambda)$. By definition, for $\lambda \ge 0$, $C^{(k)}(0;\lambda)$, is increasing and affine in λ. Therefore, its pointwise minimum (over k) is increasing and concave in λ.

As shown in part 1), for $\lambda \in (\lambda_\beta^{(k_n)}, \lambda_\beta^{(k_{n+1})}]$, $C_\beta^*(\lambda) = C_\beta^{(k_{n+1})}(0;\lambda)$, which is linear (and continuous) in λ; hence, $C_\beta^*(\lambda)$ is piecewise linear. Finally, by (B.11), $C_\beta^{(k_n)}(0;\lambda^{(k_n)}) = C_\beta^{(k_{n+1})}(0;\lambda^{(k_n)})$. Therefore, at the corner points,
$\lim_{\lambda\uparrow\lambda_\beta^{(k_{n+1})}} C_\beta^*(\lambda) = \lim_{\lambda\downarrow\lambda_\beta^{(k_{n+1})}} C_\beta^*(\lambda)$. Hence, $C_\beta^*(\lambda)$ is continuous in λ.

B.5 Proof of Lemma 3.9.2

Following the definition given by (3.57)–(3.58), we define the function $\mu^{(k)}$ for the i.i.d. erasure channel as follows:

$$\mu^{(k)}(e) \coloneqq \begin{cases} \mu(e), & \text{if } e \in (-k,k) \\ \varepsilon\mu(e), & \text{if } e \in \mathbb{R}\setminus(-k,k), \end{cases}$$

where $\varepsilon = \mathbb{P}(S_t = 0)$ is the probability of packet-drop.

Similarly, define the function $d^{(k)}$ as follows:

$$d^{(k)}(e) \coloneqq \begin{cases} d(e), & \text{if } e \in (-k, k) \\ \varepsilon d(e), & \text{if } e \in \mathbb{R} \setminus (-k, k). \end{cases} \tag{B.14}$$

To show the dependence on the variance σ^2, let us use a subscript σ in the above equation, i.e.,

$$\mu_\sigma^{(k)}(e) \coloneqq \begin{cases} \mu_\sigma(e), & \text{if } e \in (-k, k) \\ \varepsilon \mu_\sigma(e), & \text{if } e \in \mathbb{R} \setminus (-k, k), \end{cases} \tag{B.15}$$

Note that $e \in (-k, k)$ implies $e/\sigma \in (-k/\sigma, k/\sigma)$ and $e \in \mathbb{R} \setminus (-k, k)$ implies $e/\sigma \in \mathbb{R} \setminus (-k/\sigma, k/\sigma)$. Now, for any $e \in (-k, k)$,

$$\mu_\sigma^{(k)}(e) = \mu_\sigma(e) \overset{(a)}{=} (1/\sigma)\mu_1(e/\sigma) \overset{(b)}{=} (1/\sigma)\mu_1^{(k/\sigma)}(e/\sigma),$$

where (a) holds by algebraic calculation on a Gaussian pdf with zero mean and variance σ^2 and (b) holds due to (B.15). Similarly, for $e \in \mathbb{R} \setminus (-k, k)$,

$$\mu_\sigma^{(k)}(e) = \varepsilon \mu_\sigma(e) = \frac{\varepsilon}{\sigma}\mu_1(e/\sigma) = \frac{1}{\sigma}\mu_1^{(k/\sigma)}(e/\sigma).$$

Now, define $\hat{L}_\sigma^{(k)}(e) \coloneqq \sigma^2 L_1^{(k/\sigma)}\left(\frac{e}{\sigma}\right)$ and the operator $\mathcal{B}^{(k)}$ as introduced in Appendix B.1. Then,

$$\begin{aligned} [\mathcal{B}_\sigma^{(k)} \hat{L}_\sigma^{(k)}](e) &= \int_{n \in \mathbb{R}} \mu_\sigma^{(k)}(n - ae)\hat{L}_\sigma^{(k)}(n) dn, \quad \forall e \in \mathbb{R} \\ &\overset{(a)}{=} \sigma^2 \int_{z \in \mathbb{R}} \mu_1^{(k/\sigma)}(z - ae/\sigma) L_1^{(k/\sigma)}(z) dz \\ &= \sigma^2 [\mathcal{B}_1^{(k/\sigma)} L_1^{(k/\sigma)}](e/\sigma), \end{aligned}$$

where (a) uses a change of variables $n = \sigma z$. Therefore, for $e \in (-k, k)$,

$$\begin{aligned}\left[\hat{L}_\sigma^{(k)} - \beta \mathcal{B}_\sigma^{(k)} \hat{L}_\sigma^{(k)}\right](e) &= \sigma^2 \left[L_1^{(k/\sigma)} - \beta \mathcal{B}_1^{(k/\sigma)} L_1^{(k/\sigma)}\right]\left(\frac{e}{\sigma}\right) \\ &= \sigma^2 \frac{e^2}{\sigma^2} = e^2 = d^{(k)}(e).\end{aligned}$$

Similarly, for $e \in \mathbb{R} \setminus (-k, k)$,

$$\begin{aligned}\left[\hat{L}_\sigma^{(k)} - \beta \mathcal{B}_\sigma^{(k)} \hat{L}_\sigma^{(k)}\right](e) &= \sigma^2 \left[L_1^{(k/\sigma)} - \beta \mathcal{B}_1^{(k/\sigma)} L_1^{(k/\sigma)}\right]\left(\frac{e}{\sigma}\right) \\ &= \sigma^2 \varepsilon \frac{e^2}{\sigma^2} = \varepsilon e^2 = d^{(k)}(e).\end{aligned}$$

Appendix C

Adaptive Moment algorithm for computing the learning rates in SA

C.1 Adaptive Moment approximation - AdaM

AdaM is an efficient way of computing adaptive learning rates in SA algorithms for different parameters from estimates of first and second moments of the gradients. This algorithm combines the advantages of two previously developed algorithms–AdaGrad [118] and RMSProp [119]. The first works well with sparse gradients and the second works efficiently in on-line and non-stationary frames. Some of the advantages that makes AdaM a good choice for SA are the following, (i) the magnitudes of parameter updates are invariant to rescaling of the gradient, (ii) its step-sizes are approximately bounded by the step-size hyper-parameter, (iii) it does not require a stationary objective, it works with sparse gradients, and it naturally performs a form of step size annealing.

C.1.1 The algorithm

In this section we briefly describe the algorithm, which is given in Algorithm 7. Let the parameterized stochastic objective function to be minimized be denoted by $J_t(k_t)$, where k_t is the parameter at time t. Denote by k° the initial parameter vector in n-dimensional Euclidean space $\mathbb{R}^n$. Denote by m_t and v_t the first and second moment vectors at time t and their initial values by m_0 and v_0. Let the gradient of the objective function at time t be denoted by g_t. Note that with a slight abuse of notations, all operations on the vectors

Algorithm 7: AdaM Algorithm for computation of optimal parameter using SA.

input : Step-size $\gamma \in \mathbb{R}_{>0}$, exponential decay rates $\beta_1, \beta_2 \in [0, 1)$, $\epsilon \in \mathbb{R}_{>0}$
output: k_t
Initial guess of parameter: $k_0 \leftarrow k^\circ$
Initial value of the first moment vector: $m_0 \leftarrow 0$
Initial value of the second moment vector: $v_0 \leftarrow 0$
Initial time: $t \leftarrow 0$
while k_t *not converged* **do**
  $t \leftarrow t + 1$
  $g_t \leftarrow \nabla_k J_t(k_{t-1})$
  $m_t \leftarrow \beta_1 m_{t-1} + (1 - \beta_1) g_t$
  $v_t \leftarrow \beta_2 v_{t-1} + (1 - \beta_2) g_t \odot g_t$
  Bias-corrected first moment estimate: $\hat{m}_t \leftarrow m_t/(1 - \beta_1^t)$
  Bias-corrected second moment estimate: $\hat{v}_t \leftarrow v_t/(1 - \beta_2^t)$
  $k_t \leftarrow k_{t-1} - \gamma(\hat{m}_t/(\sqrt{\hat{v}_t}) + \epsilon)$
return k_t

mean element-wise operations. For details, please refer to [66].

Algorithm 7 updates exponential moving averages of the gradient and the squared gradient where the hyper-parameters control the exponential decay rates of these moving averages. The moving averages themselves are estimates of the first moment (the mean) and the second raw moment (the uncentered variance) of the gradient of the objective function. The initialization of the moments are taken to be zero, which may create a bias towards zeros, especially especially during the initial steps and with low decay rates. This is counterbalanced by computing *bias-corrected* estimates. The update rule of the algorithm relies heavily on the careful choice of step-sizes. [66, Theorem 4.1] provides an upper bound of the *regret*, i.e., total approximation error for the objective value.

In is shown in [66] that AdaM works at least as well as AdaGrad and RMSProp in terms of speed of convergence, when a noisy observation of the gradient of the objective function is available. In high-dimensional optimization problems such as *Neural Networks* or *Logistic Regression* this algorithm works well. However, in our algorithms we are using a simultaneous perturbation method to compute an estimate of the gradient of the objective function. So, the effect of computing the learning rates involving the random perturbation vector on the speed of convergence is worth investigating.

Appendix D

Proofs of Chapter 5

D.1 Proof of Lemma 5.2.1

Define the functions A_t^i, B_t^i, $i \in \{1, 2\}$ as follows:

$$\begin{aligned} A_t^i(x_{1:t}^i, z) &\coloneqq \mathbb{P}(x_1^i \mid z) \cdots \mathbb{P}(x_t^i \mid z) \\ B_t^1(x_{1:t}^1, u_{1:t}) &\coloneqq \mathbb{P}(u_1^1 \mid x_1^1) \cdots \mathbb{P}(u_t^1 \mid x_{1:t}^1, u_{1:t-1}^2) \\ B_t^2(x_{1:t}^2, u_{1:t}) &\coloneqq \mathbb{P}(u_1^2 \mid x_1^2, u_1^1) \cdots \mathbb{P}(u_t^1 \mid x_{1:t}^2, u_{1:t}^1). \end{aligned}$$

Then, we have by the chain rule of the probability,

$$\begin{aligned} &\mathbb{P}(z, x_{1:t}^1, x_{1:t}^2, u_{1:t-1}^1, u_{1:t-1}^2) \\ &= \mathbb{P}(z) A_t^1(x_{1:t}^1, z) B_t^1(x_{1:t}^1, u_{1:t}) A_t^2(x_{1:t}^2, z) B^2(x_{1:t}^2, u_{1:t}). \end{aligned} \tag{D.1}$$

Now, by total probability we have,

$$\begin{aligned} \mathbb{P}(z, u_{1:t}^1, u_{1:t}^2) = \mathbb{P}(z) \sum_{x_{1:t}^1, x_{1:t}^2} \Big(& A_t^1(x_{1:t}^1, z) B_t^1(x_{1:t}^1, u_{1:t}) \\ & A_t^2(x_{1:t}^2, z) B^2(x_{1:t}^2, u_{1:t}) \Big) \end{aligned} \tag{D.2}$$

$$= \mathbb{P}(z) \sum_{x^1} A_t^1(x_{1:t}^1, z) B_t^1(x_{1:t}^1, u_{1:t}) \sum_{x^2} A_t^2(x_{1:t}^2, z) B^2(x_{1:t}^2, u_{1:t}). \tag{D.3}$$

Therefore,

$$\begin{aligned}\mathbb{P}(x^1_{1:t}, x^2_{1:t} \mid z, u_{1:t}) &= \frac{\mathbb{P}(z, x^1_{1:t}, x^2_{1:t}, u^1_{1:t}, u^2_{1:t})}{\mathbb{P}(z, u^1_{1:t}, u^2_{1:t})}\\ &\overset{(a)}{=} \frac{A^1_t(x^1_{1:t}, z)B^1_t(x^1_{1:t}, u_{1:t})}{\sum_{x^1} A^1_t(x^1_{1:t}, z)B^1_t(x^1_{1:t}, u_{1:t})} \frac{A^2_t(x^2_{1:t}, z)B^2_t(x^2_{1:t}, u_{1:t})}{\sum_{x^1} A^2_t(x^2_{1:t}, z)B^2_t(x^2_{1:t}, u_{1:t})}\\ &= \mathbb{P}(x^1_{1:t} \mid z, u_{1:t})\mathbb{P}(x^2_{1:t} \mid z, u_{1:t}),\end{aligned}$$

where (a) follows from (D.1) and (D.3). Following the similar steps we get $\mathbb{P}(x^1_{1:t}, x^2_{1:t} \mid z, u_{1:t-1}) = \mathbb{P}(x^1_{1:t} \mid z, u_{1:t-1})\mathbb{P}(x^2_{1:t} \mid z, u_{1:t-1})$.

D.2 Proof of Lemma 5.2.2

Let us arbitrarily fix the strategy of user 2. Consider the following:

$$\xi^1_{t|t}(z) = \mathbb{P}(z \mid x^1_{1:t}, u_{1:t}) = \frac{\mathbb{P}(z, x^1_{1:t}, u_{1:t})}{\mathbb{P}(x^1_{1:t}, u_{1:t})}.$$

Now, by total probability, we have $\mathbb{P}(z, x^1_{1:t}, u_{1:t}) = \sum_{x^2_{1:t}} \mathbb{P}(z, x_{1:t}, u_{1:t})$, where $x_{1:t} = (x^1_{1:t}, x^2_{1:t})$. Also, by chain rule we have

$$\begin{aligned}\mathbb{P}(z, x_{1:t}, u_{1:t}) &= \mathbb{P}(u^1_t \mid x^1_{1:t}, u^1_{1:t-1}, u^2_{1:t-1})\mathbb{P}(u^2_t \mid x^2_{1:t}, u^1_{1:t}, u^2_{1:t-1})\mathbb{P}(x^2_{1:t} \mid z, u_{1:t-1})\\ &\qquad \mathbb{P}(z \mid x^1_{1:t}, u_{1:t-1})\mathbb{P}(x^1_{1:t}, u_{1:t-1})\\ &= \mathbb{P}(u^1_t \mid x^1_{1:t}, u^1_{1:t-1}, u^2_{1:t-1})\mathbb{P}(u^2_t \mid x^2_{1:t}, u^1_{1:t}, u^2_{1:t-1})\mathbb{P}(x^2_{1:t} \mid z, u_{1:t-1})\\ &\qquad \xi^1_{t|t-1}(z)\mathbb{P}(x^1_{1:t}, u_{1:t-1}).\end{aligned} \tag{D.4}$$

Now, it is shown in the proof of Lemma 5.2.1 that,

$$\mathbb{P}(x^2_{1:t} \mid z, u_{1:t-1}) = \frac{A^2_t(x^2_{1:t}, z)B^2_t(x^2_{1:t}, u_{1:t})}{\sum_{x^1} A^2_t(x^2_{1:t}, z)B^2_t(x^2_{1:t}, u_{1:t})},$$

which depends only on $\mathbf{f}^2$. Let us denote it as $\ell(\mathbf{f}^2)$. Substituting this in the expression for $\xi^1_{t|t}$, we have

$$\begin{aligned}\xi^1_{t|t}(z) &= \frac{\xi^1_{t|t-1}(z)\sum_{x^2_{1:t}}\left(\mathbb{1}(u^2_t = f^2_t(x^2_{1:t}, u^1_{1:t}, u^2_{1:t-1}))\ell(\mathbf{f}^2)\right)}{\sum_z\left(\xi^1_{t|t-1}(z)\sum_{x^2_{1:t}}\left(\mathbb{1}(u^2_t = f^2_t(x^2_{1:t}, u^1_{1:t}, u^2_{1:t-1}))\ell(\mathbf{f}^2)\right)\right)}\\ &=: F^1_{t|t}(\xi^1_{t|t-1}, u^2_t, \mathbf{f}^2).\end{aligned}$$

Note that $F^1_{t|t}$ is of the above form since there is no dependence of $F^1_{t|t}$ on $\mathbf{f}^1$ through $\xi^1_{t|t-1}$.

Now consider the following,

$$\xi^1_{t+1|t}(z) = \mathbb{P}(z \mid x^1_{1:t+1}, u_{1:t}) = \frac{\mathbb{P}(z, x^1_{t+1} \mid x^1_{1:t}, u_{1:t})}{\mathbb{P}(x^1_{t+1} \mid x^1_{1:t}, u_{1:t})}. \tag{D.5}$$

Also, it can be shown by similar calculation that $\mathbb{P}(z, x^1_{t+1} \mid x^1_{1:t}, u_{1:t}) = \mathbb{P}(x^1_{t+1} \mid z)\xi^1_{t|t}(z)$. Substituting back in (D.5), we have

$$\xi^1_{t+1|t}(z) = \frac{\mathbb{P}(x^1_{t+1} \mid z)\xi^1_{t|t}(z)}{\sum_z \mathbb{P}(x^1_{t+1} \mid z)\xi^1_{t|t}(z)} =: F^1_{t+1|t}(\xi^1_{t|t}, x^1_{t+1}).$$

This completes the proof for user 1. The results for user 2 can be derived similarly.

This page is intentionally left blank.

Appendix E

Proofs of Chapter 6

E.1 Proof of Claim 1

We first prove some intermediate results:

Lemma E.1.1 *Under* (M4) *and* (M5)*, for any* $x, y \in \mathbb{X}_{\geq 0}$*, we have that*

$$\varphi(y-x) \geq \varphi(y+x)$$

Proof We consider two cases: $y \geq x$ and $y < x$.

1. If $y \geq x$, then $y + x \geq y - x \geq 0$. Thus, (M5) implies that $\varphi(y+x) \geq \varphi(y-x)$.
2. If $y < x$, then $y+x \geq x-y$. Thus, (M5) implies that $\varphi(y+x) \geq \varphi(x-y) = \varphi(y-x)$, where the last equality follows from (M4).

Some immediate implications of Lemma E.1.1 are the following.

Lemma E.1.2 *Under* (M4) *and* (M5)*, for any* $a \in \mathbb{X}$ *and* $x, y \in \mathbb{X}_{\geq 0}$*, we have that*

$$a\big[\varphi(y-ax) - \varphi(y+ax)\big] \geq 0.$$

Proof For $a \in \mathbb{X}_{\geq 0}$, from Lemma E.1.1 we get that $\varphi(y-ax) \geq \varphi(y+ax)$. For $a \in \mathbb{X}_{<0}$, from Lemma E.1.1 we get that $\varphi(y+ax) \geq \varphi(y-ax)$.

Lemma E.1.3 *Under* (M4) *and* (M5)*, for any* $a, b, x, y \in \mathbb{X}_{\geq 0}$*, we have that*

$$\varphi(y-ax-b) \geq \varphi(y+ax+b) \geq \varphi(y+ax+b+1).$$

Proof By taking $y = y - b$ and $x = ax$ in Lemma E.1.1, we get

$$\varphi(y - b - ax) \geq \varphi(y - b + ax).$$

Now, by taking $y = y + ax$ and $x = b$ in Lemma E.1.1, we get

$$\varphi(y + ax - b) \geq \varphi(y + ax + b).$$

By combining these two inequalities, we get

$$\varphi(y - ax - b) \geq \varphi(y + ax + b).$$

The last inequality in the result follows from (M5).

Lemma E.1.4 *Under* (M4) *and* (M5)*, for $a \in \mathbb{Z}$ and $x, y \in \mathbb{Z}_{\geq 0}$,*

$$\Phi(y + ax) + \Phi(y - ax) \geq \Phi(y + ax + a) + \Phi(y - ax - a).$$

Proof The statement holds trivially for $a = 0$. Furthermore, the statement does not depend on the sign of a. So, without loss of generality, we assume that $a > 0$.

Now consider the following series of inequalities (which follow from Lemma E.1.3)

$$\begin{aligned} \varphi(y - ax) &\geq \varphi(y + ax + 1), \\ \varphi(y - ax - 1) &\geq \varphi(y + ax + 2), \\ \cdots &\geq \cdots \\ \varphi(y - ax - a + 1) &\geq \varphi(y + ax + a). \end{aligned}$$

Adding these inequalities, we get

$$\Phi(y - ax) - \Phi(y - ax - a) \geq \Phi(y + ax + a) - \Phi(y + ax),$$

which proves the result.

Proof of Claim 1

First, let's assume that $\mathbb{X} = \mathbb{R}$. We prove each part separately.

1. Fix $u \in [0, u_{\max}]$. $c(\cdot, u)$ is even because $d(\cdot)$ is even (from (M3)). $c(\cdot, u)$ is quasi-convex because $1 - q(u) \geq 0$ (from (M0)) and $d(\cdot)$ is quasi-convex (from (M3)).

2. Consider $e_1, e_2 \in \mathbb{R}_{\geq 0}$ and $u_1, u_2 \in [0, u_{\max}]$ such that $e_1 \geq e_2$ and $u_1 \geq u_2$. The per-step cost is submodular on $\mathbb{R}_{\geq 0} \times [0, u_{\max}]$ because

$$\begin{aligned} c(e_1, u_2) - c(e_2, u_2) &= (1 - q(u_2))(d(e_1) - d(e_2)) \\ &\overset{(a)}{\geq} (1 - q(u_1))(d(e_1) - d(e_2)) \\ &= c(e_1, u_1) - c(e_2, u_1), \end{aligned}$$

where (a) is true because $d(e_1) - d(e_2) \geq 0$ (from (M3)) and $1 - q(u_2) \geq 1 - q(u_1) \geq 0$ (from (M0) and (M2)).

3. Fix $u \in [0, u_{\max}]$ and consider $e, e_+ \in \mathbb{R}$. Then, $p(u)$ is even because

$$\begin{aligned} p(-e_+ | -e; u) &= q(u)\varphi(e_+) + (1 - q(u))\varphi(-e_+ + ae) \\ &\overset{(b)}{=} q(u)\varphi(e_+) + (1 - q(u))\varphi(e_+ - ae) \\ &= p(e_+ | e; u), \end{aligned}$$

where (b) is true because φ is even (from (M4)).

4. First note that

$$\begin{aligned} S(y|x; u) &= 1 - \int_{-\infty}^{y} \big[p(z|x; u) + p(-z|x; u)\big] dz \\ &= 1 - \int_{-\infty}^{y} q(u)\big[\varphi(z) + \varphi(-z)\big] dz \\ &\quad - \int_{-\infty}^{y} (1 - q(u))\big[\varphi(z - ax) + \varphi(-z - ax)\big] dz \\ &\overset{(c)}{=} 1 - 2q(u)\Phi(y) \\ &\quad - (1 - q(u))\big[\Phi(y - ax) + \Phi(y + ax)\big] \end{aligned}$$

where Φ is the cumulative density of φ and (c) uses the fact that φ is even (condition (M4)).

Let $S_x(y|x;u)$ denote $\partial S/\partial x$. Then

$$S_x(y|x;u) = (1-q(u))a\big[\varphi(y-ax) - \varphi(y+ax)\big].$$

From (M0) and Lemma E.1.2, we get that $S_x(y|x;u) \geq 0$ for any $x, y \in \mathbb{R}_{\geq 0}$ and $u \in [0, u_{\max}]$. Thus, $S(y|x;u)$ is increasing in x.

Furthermore, from (M2) $S_x(y|x,u)$ is decreasing in u. Thus, $S(y|x,u)$ is submodular in (x,u) on $\mathbb{R}_{\geq 0} \times [0, u_{\max}]$.

Now, let's assume that $\mathbb{X} = \mathbb{Z}$. The proof of the first three parts remains the same. Now, in part 4), it is still the case that

$$S(y|x;u) = 1 - 2q(u)\Phi(y) - (1-q(u))\big[\Phi(y-ax) + \Phi(y+ax)\big]$$

However, since $\mathbb{X}$ is discrete, we cannot take the partial derivative with respect to x. Nonetheless, following the same intuition, for any $x, y \in \mathbb{Z}_{\geq 0}$, consider

$$\begin{aligned} S(y|x+1;u) - S(y|x;u) = (1-q(u))\big[&\Phi(y+ax) \\ &- \Phi(y+ax+a) + \Phi(y-ax) - \Phi(y-ax-a)\big] \end{aligned} \tag{E.1}$$

Now, by Lemma E.1.4, the term in the square bracket is positive, and hence $S(y|x;u)$ is increasing in x. Moreover, since $(1-q(u))$ is decreasing in u, so is $S(y|x+1;u) - S(x|x;u)$. Hence, $S(y|x;u)$ is submodular in $\mathbb{Z}_{\geq 0} \times [0, u_{\max}]$.

References

[1] J. Marschak, “Towards an economic theory of organization and information,” *Decision Processes*, vol. 3, no. 1, pp. 187–220, 1954.

[2] R. Radner, “Team decision problems,” *Annals of Mathmatical Statistics*, vol. 33, pp. 857–881, 1962.

[3] H. S. Witsenhausen, “On the structure of real-time source coders,” *Bell System Technical Journal*, vol. 58, pp. 1437–1451, July-August 1979.

[4] C. E. Shannon, “A mathematical theory of communication,” *Bell System Technical Journal*, vol. 22, pp. 379–423, July 1948.

[5] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, “Block-fading channels at finite blocklength,” in *Wireless Communication Systems (ISWCS 2013), Proceedings of the Tenth International Symposium on*, pp. 1–4, VDE, 2013.

[6] J. C. Walrand and P. Varaiya, “Optimal causal coding-decoding problems,” *IEEE Transactions on Information Theory*, vol. 29, pp. 814–820, Nov. 1983.

[7] D. Teneketzis, “On the structure of optimal real-time encoders and decoders in noisy communication,” *IEEE Transactions on Information Theory*, pp. 4017–4035, Sept. 2006.

[8] A. Mahajan and D. Teneketzis, “Optimal design of sequential real-time communication systems,” *IEEE Transactions on Information Theory*, vol. 55, pp. 5317–5338, Nov. 2009.

[9] G. M. Lipsa and N. Martins, “Remote state estimation with communication costs for first-order LTI systems,” *IEEE Transactions on Automatic Control*, vol. 56, pp. 2013–2025, Sept. 2011.

[10] A. Nayyar, T. Basar, D. Teneketzis, and V. Veeravalli, “Optimal strategies for communication and remote estimation with an energy harvesting sensor,” *IEEE Transactions on Automatic Control*, vol. 58, no. 9, pp. 2246–2260, 2013.

[11] A. Molin and S. Hirche, "An iterative algorithm for optimal event-triggered estimation," in *4th IFAC Conference on Analysis and Design of Hybrid Systems (ADHS'12)*, pp. 64–69, 2012.

[12] X. Ren, J. Wu, K. H. Johansson, G. Shi, and L. Shi, "Infinite Horizon Optimal Transmission Power Control for Remote State Estimation over Fading Channels," *arXiv: 1604.08680*, Apr. 2016.

[13] J. Chakravorty and A. Mahajan, "Remote-state estimation with packet drop," *IFAC-PapersOnLine*, vol. 49, no. 22, pp. 7–12, 2016. 6th {IFAC} Workshop on Distributed Estimation and Control in Networked Systems {NECSYS} 2016Tokyo, Japan, 8?9 September 2016.

[14] J. Chakravorty and A. Mahajan, "Fundamental limits of remote estimation of autoregressive markov processes under communication constraints," *IEEE Transactions on Automatic Control*, vol. 62, pp. 1109–1124, March 2017.

[15] T. Linder and G. Lagosi, "A zero-delay sequential scheme for lossy coding of individual sequences," *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2533–2538, 2001.

[16] T. Weissman and N. Merhav, "On limited-delay lossy coding and filtering of individual sequences," vol. 48, no. 3, pp. 721–733, 2002.

[17] A. György, T. Linder, and G. Lugosi, "Efficient adaptive algorithms and minimax bounds for zero-delay lossy source coding," vol. 52, no. 8, pp. 2337–2347, 2004.

[18] S. Matloub and T. Weissman, "Universal zero-delay joint source-channel coding," vol. 52, pp. 5240–5250, Dec. 2006.

[19] Y. Kaspi and N. Merhav, "Structure theorems for real-time variable rate coding with and without side information," vol. 58, no. 12, pp. 7135–7153, 2012.

[20] H. Asnani and T. Weissman, "Real-time coding with limited lookahead," vol. 59, no. 6, pp. 3582–3606, 2013.

[21] H. Kushner, "On the optimum timing of observations for linear control systems with unknown initial state," *IEEE Transactions on Automatic Control*, vol. 9, pp. 1844–150, Apr 1964.

[22] E. Skafidas and A. Nerode, "Optimal measurement scheduling in linear quadratic gaussian control problems," in *IEEE Proceedings of International Conference on Control Applications*, pp. 1225–1229, Sep 1998.

[23] H. S. Witsenhausen, “Separation of estimation and control for discrete time systems,” vol. 59, pp. 1557–1566, Nov. 1971.

[24] Y.-C. Ho and K.-C. Chu, “Team decision theory and information structures in optimal control problems–Part I,” vol. 17, no. 1, pp. 15–22, 1972.

[25] K. J. Åström and B. M. Bernhardsson, “Comparison of Riemann and Lebesgue sampling for first order stochastic system,” in *IEEE Proceedings of Conference on Decision and Control*, pp. 2011–2016, 2002.

[26] O. C. Imer and T. Basar, “Optimal estimation with limited measurements,” *Joint 44the IEEE Conference on Decision and Control and European Control Conference*, vol. 29, pp. 1029 – 1034, 2005.

[27] M. Rabi, G. V. Moustakides, and J. S. Baras, “Adaptive sampling for linear state estimation,” *SIAM Journal on Control and Optimization*, vol. 50, no. 2, pp. 672–702, 2012.

[28] Y. Xu and J. P. Hespanha, “Optimal communication logics in networked control systems,” in *Proceedings of 43rd IEEE Conference on Decision and Control*, vol. 4, pp. 3527–3532, 2004.

[29] C. Rago, P. Willett, and Y. Bar-Shalom, “Censoring sensors: A low-communication rate scheme for distributed detection,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, pp. 554–568, April 1996.

[30] S. Appadwedula, V. V. Veeravalli, and D. L. Jones, “Decentralized detection with censoring sensors,” *IEEE Transactions on Signal Processing*, vol. 56, pp. 1362–1373, April 2008.

[31] M. Athans, “On the determination of optimal costly measurement strategies for linear stochastic systems,” *Automatica*, vol. 8, no. 4, pp. 397–412, 1972.

[32] J. Geromel, “Global optimization of measurement strategies for linear stochastic systems,” *Automatica*, vol. 25, no. 2, pp. 293–300, 1989.

[33] W. Wu, A. Araposthathis, and V. V. Veeravalli, “Optimal sensor querying: General Markovian and LQG models with controlled observations,” *IEEE Transactions on Automatic Control*, vol. 53, no. 6, pp. 1392–1405, 2008.

[34] D. Shuman and M. Liu, “Optimal sleep scheduling for a wireless sensor network node,” in *Proceedings of the Asilomar Conference on Signals, Systems, and Computers*, pp. 1337–1341, October 2006.

[35] M. Sarkar and R. L. Cruz, “Analysis of power managemnet for energy and delay trade-off in a WLAN,” in *Proceedings ofthe Conference on Information Sciences and Systems,*, March 2004.

[36] M. Sarkar and R. L. Cruz, “An adaptive sleep algorithm for efficient power management in WLANs,” in *Proceedings of the Vehicular Technology Conference*, May 2005.

[37] A. Federgruen and K. C. So, “Optimality ofthreshold policies in single- server queueing systems with server vacations,” *Adv. Appl. Prob.*, vol. 23, no. 2, pp. 388–405, 1991.

[38] K. J. Åström, *Analysis and Design of Nonlinear Control Systems*, ch. Event based control. Berlin, Heidelberg: Springer, 2008.

[39] X. Meng and T. Chen, “Optimal sampling and performance comparison of periodic and event based impulse control,” *IEEE Transactions of Automatic Control*, vol. 57, no. 12, pp. 3252–3259, 2012.

[40] D. Shi, L. Shi, and T. Chen, “Event-based state estimation–a stochastic perspective, ser,” *Studies in Systems, Decision and Control. Springer International Publishing*, vol. 41, 2016.

[41] H. S. Witsenhausen, “A counterexample in stochastic optimum control,” *SIAM Journal on Control*, vol. 6, no. 1, pp. 131–147, 1968.

[42] E. Altman, *Constrained Markov decision processes.* Stochastic Modeling, Chapman and Hall/CRC, 1998.

[43] A. D. Polyanin and A. V. Manzhirov, *Handbook of integral equations.* Chapman and Hall/CRC Press, second ed., 2008.

[44] K. Atkinson and L. F. Shampine, “Solving Fredholm integral equations of the second kind in Matlab,” *ACM Trans. Math. Software*, 2008.

[45] L. I. Sennott, “Constrained discounted Markov decision chains,” *Probability in the Engineering and Informational Sciences*, vol. 6, pp. 463–475, Oct 1991.

[46] E. Feinberg, “Optimality of deterministic policies for certain stochastic control problems with multiple criteria and constraints,” in *Mathematical Control Theory and Finance* (A. Sarychev, A. Shiryaev, M. Guerra, and M. Grossinho, eds.), pp. 137–148, Springer Berlin Heidelberg, 2008.

[47] A. Shwartz and A. M. Makowski, “An optimal adaptive scheme for two competing queues with constraints,” in *Analysis and optimization of systems*, pp. 515–532, Springer Berlin Heidelberg, 1986.

[48] D. J. Ma, A. M. Makowski, and A. Shwartz, "Stochastic approximations for finite-state Markov chains," *Stochastic Processes and Their Applications*, vol. 35, no. 1, pp. 27–45, 1990.

[49] E. Altman and A. Shwartz, "Time-sharing policies for controlled Markov chains," *Operations Research*, vol. 41, no. 6, pp. 1116–1124, 1993.

[50] L. Wang, J. Woo, and M. Madiman, "A lower bound on Rényi entropy of convolutions in the integers," in *Proceedings of the 2014 IEEE International Symposium on Information Theory*, pp. 2829–2833, July 2014.

[51] L. I. Sennott, *Stochastic dynamic programming and the control of queueing systems*. New York, NY, USA: Wiley, 1999.

[52] O. H. Lerma and J. B. Lasserre, *Discrete-time Markov control processes : basic optimality criteria*. Applications of mathematics 30, Springer, 1996.

[53] L. I. Sennott, "Computing average optimal constrained policies in stochastic dynamic programming," *Probability in the Engineering and Informational Sciences*, vol. 15, pp. 103–133, 2001.

[54] V. Borkar, "A convex analytic approach to Markov decision processes," *Probability Theory and Related Fields*, vol. 78, no. 4, pp. 583–602, 1988.

[55] D. Luenberger, *Optimization by Vector Space Methods*. Professional Series, Wiley, 1968.

[56] G. Hu and R. F. O'Connell, "Analytical inversion of symmetric tridiagonal matrices," *Journal of Physics A: Mathematical and General*, vol. 29, no. 7, p. 1511, 1996.

[57] A. Nayyar, A. Mahajan, and D. Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach," vol. 58, pp. 1644–1658, July 2013.

[58] E. N. Gilbert, "Capacity of a burst-noise channel," *Bell System Technical Journal*, vol. 39, no. 5, pp. 1253–1265, 1960.

[59] E. O. Elliott, "Estimates of error rates for codes on burst-noise channels," *Bell System Technical Journal*, vol. 42, no. 5, pp. 1977–1997, 1963.

[60] M. Puterman, *Markov decision processes: Discrete Stochastic Dynamic Programming*. John Wiley and Sons, 1994.

[61] S. Yuksel, "On optimal causal coding of partially observed Markov sources in single and multiterminal settings," vol. 59, no. 1, pp. 424–437, 2013.

[62] P. R. Kumar and P. Varaiya, *Stochastic Systems: Estimation Identification and Adaptive Control*. Prentice Hall, 1986.

[63] J. Kiefer and J. Wolfowitz, "Stochastic estimation of the maximum of a regression function," *Ann. Math. Statist.*, vol. 23, pp. 462–466, Sep 1952.

[64] V. Y. Katkovnik and K. OY, "Convergence of a class of random search algorithms," *Automation and Remote Control*, vol. 33, no. 8, pp. 1321–1326, 1972.

[65] H. Robbins and S. Monro, "A stochastic approximation method," *Ann. Math. Statist.*, vol. 22, pp. 400–407, Sep 1951.

[66] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arxiv:1412.6980*, Jan 2017.

[67] J. C. Spall, "Multivariate stochastic approximation using a simultaneous perturbation gradient approximation," vol. 37, pp. 332–341, Mar 1992.

[68] S. Bhatnagar and V. S. Borkar, "Multiscale chaotic spsa and smoothed functional algorithms for simulation optimization," *Simulation*, vol. 79, no. 10, pp. 568–580, 2003.

[69] S. Bhatnagar, "Adaptive newton-based multivariate smoothed functional algorithms for simulation optimization," *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, vol. 18, no. 1, p. 2, 2007.

[70] S. Bhatnagar, H. Prasad, and L. Prashanth, *Stochastic recursive algorithms for optimization: simultaneous perturbation methods*, vol. 434. Springer, 2012.

[71] `https://github.com/adityam/remote-estimation-packet-drops`.

[72] E. Kushilevitz and N. Nisan, *Communication complexity*. Cambridge, U.K.: Cambridge University Press, 1997.

[73] R. Ahlswede and N. Cai, "On communication complexity of vector-valued functions," *Information Theory, IEEE Transactions on*, vol. 40, pp. 2062–2067, Sep 1994.

[74] A. Kaspi, "Two-way source coding with a fidelity criterion," *Information Theory, IEEE Transactions on*, vol. 31, pp. 735–740, Nov 1985.

[75] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley series in Telecommunication, John Wiley and Sons, 1991.

[76] E. Yang and D. He, "On interactive encoding and decoding for lossless source coding with decoder only side information," in *2008 IEEE International Symposium on Information Theory, ISIT 2008, Toronto, ON, Canada, July 6-11, 2008*, pp. 419–423, 2008.

[77] E.-H. Yang and D.-K. He, "Interactive encoding and decoding for one way learning: Near lossless recovery with side information at the decoder," *IEEE Trans. Inf. Theor.*, vol. 56, pp. 1808–1824, Apr. 2010.

[78] H. Yamamoto, "Wyner - ziv theory for a general function of the correlated sources (corresp.)," *Information Theory, IEEE Transactions on*, vol. 28, pp. 803–807, Sep 1982.

[79] T. S. Han and K. Kobayashi, "A dichotomy of functions f(x, Y) of correlated sources (x, Y)," *IEEE Transactions on Information Theory*, vol. 33, no. 1, pp. 69–76, 1987.

[80] N. Ma and P. Ishwar, "Some results on distributed source coding for interactive function computation," *Information Theory, IEEE Transactions on*, vol. 57, no. 9, pp. 6180–6195, 2011.

[81] N. M. Freris, H. Kowshik, and P. Kumar, "Fundamentals of large sensor networks: Connectivity, capacity, clocks, and computation," *Proceedings of the IEEE*, vol. 98, no. 11, pp. 1828–1846, 2010.

[82] A. Orlitsky and J. R. Roche, "Coding for computing," *IEEE Trans. Inf. Theor.*, vol. 47, pp. 903–917, Sept. 2006.

[83] L. J. Schulman, "Coding for interactive communication," in *In Proceedings of the 25th Annual Symposium on Theory of Computing*, pp. 747–756, 1996.

[84] B. Barak, M. Braverman, X. Chen, and A. Rao, "How to compress interactive communication," *SIAM Journal on Computing*, vol. 42, no. 3, pp. 1327–1363, 2013.

[85] G. Kol and R. Raz, "Interactive channel capacity," in *Proceedings of the Forty-fifth Annual ACM Symposium on Theory of Computing*, STOC '13, (New York, NY, USA), pp. 715–724, ACM, 2013.

[86] B. Haeupler, "Interactive channel capacity revisited," *CoRR*, vol. abs/1408.1467, 2014.

[87] A. Mahajan and D. Teneketzis, "On the design of globally optimal communication strategies for real-time noisy communication systems with noisy feedback," vol. 26, pp. 580–595, May 2008.

[88] Y. Kaspi and N. Merhav, “Structure theorem for real-time variable-rate lossy source encoders and memory-limited decoders with side information,” in *proceedings of the IEEE Symposium on Information Theory*, (Austin, TX), 2010.

[89] J. Marschak and R. Radner, *Economic Theory of Teams*. New Haven: Yale University Press, 1972.

[90] A. Mahajan, “Optimal decentralized control of coupled subsystems with control sharing,” vol. 58, pp. 2377–2382, Sept. 2013.

[91] H. Tyagi and S. Watanabe, “Converses for secret key agreement and secure computing,” *Information Theory, IEEE Transactions on*, vol. 61, no. 9, pp. 4809–4827, 2015.

[92] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Orlando, FL, USA: Academic Press, Inc., 1982.

[93] V. Prabhakaran, K. Ramchandran, and D. Tse, “On the role of interaction between sensors in the ceo problem,” in *Allerton Conference on Communication, Control, and Computing*, September-October 2004.

[94] R. Wood, T. Linder, and S. Yuksel, “Optimality of Walrand-Varaiya type policies and approximation results for zero delay coding of markov sources,” in *Information Theory (ISIT), 2015 IEEE International Symposium on*, pp. 1382–1386, June 2015.

[95] C. Derman, “On optimal replacement rules when changes of state are Markovian,” *Mathematical optimization techniques*, vol. 396, 1963.

[96] P. Kolesar, “Minimum cost replacement under Markovian deterioration,” *Management Science*, vol. 12, no. 9, pp. 694–706, 1966.

[97] H. M. Taylor, “Evaluating a call option and optimal timing strategy in the stock market,” *Management Science*, vol. 14, no. 1, pp. 111–120, 1967.

[98] R. C. Merton, “Theory of rational option pricing,” *The Bell Journal of Economics and Management Science*, vol. 4, no. 1, pp. 141–183, 1973.

[99] M. J. Sobel, “Optimal operation of queues,” in *Mathematical methods in queueing theory*, pp. 231–261, Springer, 1974.

[100] S. Stidham Jr and R. R. Weber, “Monotonic and insensitive optimal policies for control of queues with undiscounted costs,” *Operations Research*, vol. 37, no. 4, pp. 611–625, 1989.

[101] R. F. Serfozo, *Stochastic Systems: Modeling, Identification and Optimization, II*, ch. Monotone optimal policies for Markov decision processes, pp. 202–215. Berlin, Heidelberg: Springer, 1976.

[102] C. C. White, "Monotone control laws for noisy, countable-state Markov chains," *European Journal of Operational Research*, vol. 5, no. 2, pp. 124–132, 1980.

[103] S. M. Ross, *Introduction to Stochastic Dynamic Programming: Probability and Mathematical*. Orlando, FL, USA: Academic Press, Inc., 1983.

[104] D. P. Heyman and M. J. Sobel, *Stochastic Models in Operations Research*. New York, USA: McGraw Hill, 1984.

[105] N. L. Stokey and J. Lucas, Robert E, *Recursive methods in economic dynamics*. Harvard University Press, 1989.

[106] D. M. Topkis, "Minimizing a submodular function on a lattice," *Operations research*, vol. 26, no. 2, pp. 305–321, 1978.

[107] D. M. Topkis, *Supermodularity and Complementarity*. Princeton University Press, 1998.

[108] K. Papadaki and W. B. Powell, "Monotonicity in multidimensional Markov decision processes for the batch dispatch problem," *Operations research letters*, vol. 35, no. 2, pp. 267–272, 2007.

[109] J. E. Smith and K. F. McCardle, "Structural properties of stochastic dynamic programs," *Operations Research*, vol. 50, pp. 796–809, Sep.–Oct. 2002.

[110] J. Chakravorty and A. Mahajan, "Structure of optimal strategies for remote estimation over Gilbert-Elliott channel with feedback," *arXiv: 1701.05943 [math.OC]*, Jan 20 2017.

[111] O. Hernández-Lerma and J. Lasserre, *Discrete-Time Markov Control Processes*. Springer-Verlag, 1996.

[112] L. Shi and L. Xie, "Optimal sensor power scheduling for state estimation of Gauss-Markov systems over a packet-dropping network," vol. 60, pp. 2701–2705, May 2012.

[113] X. Gao, E. Akyol, and T. Başar, "Optimal sensor scheduling and remote estimation over an additive noise channel," in *American Control Conference (ACC), 2015*, pp. 2723–2728, IEEE, 2015.

[114] R. Bellman, I. Glicksberg, and O. Gross, "On the optimal inventory equation," *Management Science*, vol. 2, no. 1, pp. 83–104, 1955.

[115] B. Hajek, K. Mitzel, and S. Yang, "Paging and registration in cellular networks: Jointly optimal policies and an iterative algorithm," vol. 64, pp. 608–622, Feb. 2008.

[116] D. M. Topkis, *Supermodularity and complementarity*. Frontiers of economic research, Princeton, NJ, USA: Princeton University Press, 1998.

[117] J. Chakravorty and A. Mahajan, "Sufficient conditions for the value function and optimal strategy to be even and quasi-convex," *arXiv: 1703.10746*, Mar 2017.

[118] J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization," *Journal of Machine Learning Research*, vol. 12, no. Jul, pp. 2121–2159, 2011.

[119] T. Tieleman and G. Hinton, "Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude," *COURSERA: Neural networks for machine learning*, vol. 4, no. 2, 2012.