

# UC San Diego

## UC San Diego Electronic Theses and Dissertations

**Title**

Active Learning and Hypothesis Testing

**Permalink**

<https://escholarship.org/uc/item/27j491d9>

**Author**

Naghshvar, Mohammad

**Publication Date**

2013

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

**Active Learning and Hypothesis Testing**

A dissertation submitted in partial satisfaction of the  
requirements for the degree  
Doctor of Philosophy

in

Electrical Engineering (Communication Theory and Systems)

by

Mohammad Naghshvar

Committee in charge:

Professor Tara Javidi, Chair  
Professor Young-Han Kim  
Professor Bhaskar Rao  
Professor Jason Schweinsberg  
Professor Angela Yu

2013

Copyright  
Mohammad Naghshvar, 2013  
All rights reserved.

The dissertation of Mohammad Naghshvar is approved,  
and it is acceptable in quality and form for publication  
on microfilm and electronically:

---

---

---

---

---

Chair

University of California, San Diego

2013

## DEDICATION

*To my lovely parents, Faranak and Farshad, whose support, encouragement, and endless love have nurtured me throughout my life, and to the loving memory of my grandfather, Ezatollah, who would have been proud of me.*

## TABLE OF CONTENTS

|                                                                  |      |
|------------------------------------------------------------------|------|
| Signature Page . . . . .                                         | iii  |
| Dedication . . . . .                                             | iv   |
| Table of Contents . . . . .                                      | v    |
| List of Figures . . . . .                                        | viii |
| List of Tables . . . . .                                         | ix   |
| Acknowledgements . . . . .                                       | x    |
| Vita . . . . .                                                   | xii  |
| Abstract of the Dissertation . . . . .                           | xiii |
| Chapter 1    Introduction . . . . .                              | 1    |
| Chapter 2    Sequentiality and Adaptivity Gains . . . . .        | 6    |
| 2.1    Introduction . . . . .                                    | 7    |
| 2.2    Problem Setup . . . . .                                   | 8    |
| 2.2.1    Problem Formulation . . . . .                           | 8    |
| 2.2.2    Types of Policies . . . . .                             | 10   |
| 2.2.3    Information State as Sufficient Statistic . . . . .     | 10   |
| 2.2.4    Overview of the Results and Literature Survey . . . . . | 11   |
| 2.3    Analytic Results . . . . .                                | 14   |
| 2.3.1    Assumptions and Basic Definitions . . . . .             | 14   |
| 2.3.2    Main Theorems . . . . .                                 | 16   |
| 2.4    Consequences of the Bounds . . . . .                      | 18   |
| 2.4.1    Sequentiality and Adaptivity Gains . . . . .            | 19   |
| 2.4.2    Reliability and Error Exponent . . . . .                | 21   |
| 2.5    Special Case: Binary Hypothesis Testing . . . . .         | 22   |
| 2.5.1    Analytical Results . . . . .                            | 22   |
| 2.5.2    Numerical Example . . . . .                             | 24   |
| 2.6    Discussion . . . . .                                      | 25   |
| 2.7    Proofs . . . . .                                          | 25   |
| 2.7.1    Proof of Theorem 2.3.1 . . . . .                        | 26   |
| 2.7.2    Proof of Theorem 2.3.2 . . . . .                        | 27   |
| 2.7.3    Proof of Theorem 2.3.3 . . . . .                        | 30   |
| 2.7.4    Proof of Proposition 2.3.1 . . . . .                    | 32   |
| 2.7.5    Technical Background . . . . .                          | 35   |

|           |                                                                                          |     |
|-----------|------------------------------------------------------------------------------------------|-----|
| Chapter 3 | Dynamic Programming and Optimal Notion of Information . .                                | 41  |
| 3.1       | Introduction . . . . .                                                                   | 42  |
| 3.2       | Characterization of an Optimal Policy . . . . .                                          | 43  |
| 3.2.1     | Reduction to Passive Sensing . . . . .                                                   | 45  |
| 3.3       | Heuristic Markov Policies . . . . .                                                      | 47  |
| 3.3.1     | Symmetric Divergences . . . . .                                                          | 48  |
| 3.3.2     | Heuristic Policies: Maximizing Divergence . . . .                                        | 52  |
| 3.4       | Numerical Example . . . . .                                                              | 53  |
| 3.5       | Proofs . . . . .                                                                         | 56  |
| 3.5.1     | Proof of Theorem 3.2.1 . . . . .                                                         | 56  |
| 3.5.2     | Proof of Theorem 3.2.2 . . . . .                                                         | 58  |
| Chapter 4 | Non-asymptotic and Asymptotic Analysis of Active Sequential Hypothesis Testing . . . . . | 60  |
| 4.1       | Introduction . . . . .                                                                   | 61  |
| 4.2       | Upper Bounds and Achievability Analysis . . . . .                                        | 62  |
| 4.3       | Lower Bounds for the Optimal Performance . . . . .                                       | 66  |
| 4.4       | Improved Stopping Rule: Numerical Example . . . . .                                      | 69  |
| 4.5       | Asymptotic Analysis and Consequences . . . . .                                           | 72  |
| 4.5.1     | Order and Asymptotic Optimality in $L$ . . . . .                                         | 72  |
| 4.5.2     | Order and Asymptotic Optimality in $L$ and $M$ . .                                       | 73  |
| 4.5.3     | Information Acquisition Rate and Reliability . . .                                       | 76  |
| 4.6       | Discussion: Technical Assumptions . . . . .                                              | 79  |
| 4.6.1     | Assumption 4.1.1 . . . . .                                                               | 79  |
| 4.6.2     | Assumption 4.1.2 . . . . .                                                               | 80  |
| 4.7       | Proofs . . . . .                                                                         | 82  |
| 4.7.1     | Preliminary Lemmas . . . . .                                                             | 82  |
| 4.7.2     | Proof of Theorem 4.2.1 . . . . .                                                         | 89  |
| 4.7.3     | Proof of Theorem 4.3.1 . . . . .                                                         | 93  |
| 4.7.4     | Proof of Theorem 4.3.2 . . . . .                                                         | 97  |
| 4.7.5     | Proof of Theorem 4.3.3 . . . . .                                                         | 98  |
| 4.7.6     | Proof of Corollaries 4.3.1 and 4.5.2 . . . . .                                           | 104 |
| 4.7.7     | Proof of Proposition 4.2.2 . . . . .                                                     | 105 |
| 4.7.8     | Proof of Proposition 4.3.1 . . . . .                                                     | 109 |
| 4.7.9     | Proof of Lemma 4.5.1 and Corollaries 4.5.4–4.5.6 .                                       | 111 |
| 4.7.10    | Proof of Proposition 4.6.1 . . . . .                                                     | 117 |
| 4.7.11    | Proof of Proposition 4.6.2 . . . . .                                                     | 120 |
| Chapter 5 | Channel Coding with Feedback . . . . .                                                   | 123 |
| 5.1       | Introduction . . . . .                                                                   | 124 |
| 5.2       | Coding over DMC with Noiseless Feedback . . . . .                                        | 126 |
| 5.2.1     | Problem Setup . . . . .                                                                  | 126 |
| 5.2.2     | Asymptotic Bounds on Minimum Expected Length                                             | 128 |

|              |       |                                                                    |     |
|--------------|-------|--------------------------------------------------------------------|-----|
|              | 5.2.3 | Stochastic Control View . . . . .                                  | 130 |
| 5.3          |       | Main Result and Applications . . . . .                             | 132 |
|              | 5.3.1 | Variable-Length Posterior Matching . . . . .                       | 134 |
|              | 5.3.2 | MaxEJS Coding . . . . .                                            | 134 |
| 5.4          |       | Coding for Binary-Input Channels . . . . .                         | 135 |
|              | 5.4.1 | Generalized Horstein-Burnashev-Zigangirov Scheme                   | 136 |
|              | 5.4.2 | Optimal Binary Variable-Length Codes . . . . .                     | 138 |
| 5.5          |       | Proofs . . . . .                                                   | 141 |
|              | 5.5.1 | Proof of Proposition 5.3.1 . . . . .                               | 141 |
|              | 5.5.2 | Proof of Proposition 5.3.2 . . . . .                               | 143 |
|              | 5.5.3 | Proof of Proposition 5.4.1 . . . . .                               | 144 |
|              | 5.5.4 | Proof of Proposition 5.4.2 . . . . .                               | 145 |
|              | 5.5.5 | Proof of Proposition 5.4.3 . . . . .                               | 146 |
| Chapter 6    |       | Two-Dimensional Visual Search . . . . .                            | 148 |
|              | 6.1   | Introduction . . . . .                                             | 149 |
|              | 6.2   | Problem Formulation . . . . .                                      | 150 |
|              | 6.3   | Heuristics . . . . .                                               | 152 |
|              | 6.4   | Main Results . . . . .                                             | 153 |
|              | 6.4.1 | Lower and Upper Bounds . . . . .                                   | 153 |
|              | 6.4.2 | Inspection Constraints and Computational Com-<br>plexity . . . . . | 156 |
|              | 6.5   | Numerical Example . . . . .                                        | 157 |
|              | 6.6   | Proofs . . . . .                                                   | 158 |
|              | 6.6.1 | Proof of Theorem 6.4.1 . . . . .                                   | 158 |
|              | 6.6.2 | Proof of Theorem 6.4.2 . . . . .                                   | 159 |
| Chapter 7    |       | Noisy Bayesian Active Learning . . . . .                           | 162 |
|              | 7.1   | Introduction . . . . .                                             | 163 |
|              | 7.2   | Problem Formulation . . . . .                                      | 165 |
|              | 7.2.1 | Related Work . . . . .                                             | 166 |
|              | 7.3   | Main Results . . . . .                                             | 167 |
|              | 7.3.1 | Preliminaries . . . . .                                            | 167 |
|              | 7.3.2 | Lower and Upper Bounds . . . . .                                   | 169 |
|              | 7.4   | Special Case: Noisy Generalized Binary Search . . . . .            | 171 |
|              | 7.4.1 | Comparison to Known Results . . . . .                              | 174 |
|              | 7.4.2 | Numerical Example . . . . .                                        | 175 |
|              | 7.5   | Proofs . . . . .                                                   | 178 |
|              | 7.5.1 | Proof of Proposition 7.3.1 . . . . .                               | 178 |
|              | 7.5.2 | Proof of Proposition 7.3.2 . . . . .                               | 179 |
|              | 7.5.3 | Finding $\alpha$ and $\beta$ for Different Function Classes . .    | 180 |
| Bibliography |       | . . . . .                                                          | 184 |

## LIST OF FIGURES

|             |                                                                                                                                                                                                                                                                                                                                                                                         |     |
|-------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 2.1: | Active binary hypothesis testing with additive Gaussian noise. .                                                                                                                                                                                                                                                                                                                        | 24  |
| Figure 3.1: | Active binary hypothesis testing with additive Gaussian noise. .                                                                                                                                                                                                                                                                                                                        | 53  |
| Figure 3.2: | JS divergence for sensing actions $a$ and $b$ . . . . .                                                                                                                                                                                                                                                                                                                                 | 54  |
| Figure 3.3: | EJS divergence for sensing actions $a$ and $b$ . . . . .                                                                                                                                                                                                                                                                                                                                | 54  |
| Figure 3.4: | Expected total cost of the proposed policies for different values<br>of $L$ . . . . .                                                                                                                                                                                                                                                                                                   | 55  |
| Figure 4.1: | Active binary hypothesis testing with additive Gaussian noise. .                                                                                                                                                                                                                                                                                                                        | 70  |
| Figure 4.2: | Expected total cost of the proposed policies for different values<br>of $L$ . . . . .                                                                                                                                                                                                                                                                                                   | 71  |
| Figure 4.3: | Lower and upper bounds on the optimal reliability function $E(R)$ .                                                                                                                                                                                                                                                                                                                     | 78  |
| Figure 4.4: | Computing $K'_1$ for $M = 2$ , $L = 5$ , and $D_{\min}(M) = 0.75$ . In<br>this example, the derivative of $f_1$ with respect to $\rho_1$ is equal to<br>$L$ at $\rho_1 = 0.35$ and $K'_1 \geq 0.56$ ensures that $f_1(L, \boldsymbol{\rho}) - K'_1 \leq$<br>$\min\{\rho_1, 1 - \rho_1\}L$ . We have $\alpha = 2.67$ , $\beta = 0.07$ , $L^* = 206.06$ ,<br>and $K'_1 = 13.16$ . . . . . | 96  |
| Figure 5.1: | A noisy memoryless channel with a noiseless causal feedback link.                                                                                                                                                                                                                                                                                                                       | 126 |
| Figure 5.2: | Two-agent problem with common and private observations from<br>the point of view of the fictitious agent. . . . .                                                                                                                                                                                                                                                                       | 130 |
| Figure 5.3: | Binary-input ternary-output channel with capacity-achieving<br>input distribution $\pi_0^* = \pi_1^* = 1/2$ . . . . .                                                                                                                                                                                                                                                                   | 140 |
| Figure 6.1: | Two-dimensional visual search under different resolutions. . . .                                                                                                                                                                                                                                                                                                                        | 150 |
| Figure 6.2: | $3 \times 3$ Grid. The inspection space of all squares in this grid<br>contains 9 single segments, 4 squares of size $2 \times 2$ (one of which<br>is shown in the figure), and one square of size $3 \times 3$ . . . . .                                                                                                                                                               | 157 |
| Figure 6.3: | Expected total cost for different values of $L$ and $M$ . . . . .                                                                                                                                                                                                                                                                                                                       | 158 |
| Figure 7.1: | Comparison of $C(p)$ , $C_1(p)$ , $(\frac{1}{2} - p)^2$ , and $\lambda(p)$ , for $p \in (0, 1/2)$ .                                                                                                                                                                                                                                                                                     | 175 |
| Figure 7.2: | Expected number of samples as $\epsilon$ varies for $M = 5$ and $p = 0.2$ .                                                                                                                                                                                                                                                                                                             | 176 |
| Figure 7.3: | Expected number of samples as $M$ varies for $\epsilon = 0.01$ and $p = 0.2$ .                                                                                                                                                                                                                                                                                                          | 176 |
| Figure 7.4: | Average probability of error as $N$ varies for $M = 5$ and $p = 0.2$ .                                                                                                                                                                                                                                                                                                                  | 177 |
| Figure 7.5: | Average probability of error as $M$ varies for $N = 10$ and $p = 0.2$ .                                                                                                                                                                                                                                                                                                                 | 177 |

# LIST OF TABLES

|            |                                                                                                |     |
|------------|------------------------------------------------------------------------------------------------|-----|
| Table 2.1: | Hypothesis Testing Literature . . . . .                                                        | 12  |
| Table 2.2: | Comparison of performance bounds for the example of Fig. 2.1. .                                | 25  |
| Table 4.1: | Different stopping thresholds for the example of Fig. 4.1. . . . .                             | 70  |
| Table 4.2: | Summary of Notations . . . . .                                                                 | 72  |
| Table 6.1: | Computational Complexity of $\pi_{EJS}$ . . . . .                                              | 156 |
| Table 7.1: | $\alpha$ and $\beta$ for different function classes. . . . .                                   | 172 |
| Table 7.2: | Performance comparison of NGBS, MSGBS, and $\pi_{EJS}$ for different function classes. . . . . | 174 |

## ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Professor Tara Javidi, for her mentorship and support. Her vision and enthusiasm to tackle new projects have been a constant source of inspiration and guidance throughout my study.

I wish to express my profound thankfulness and appreciation to my dissertation committee members, Professor Angela Yu, Professor Jason Schweinsberg, Professor Young-Han Kim, and Professor Bhaskar Rao. I owe an immense debt of gratitude to Professor Yu whose intriguing work on *perceptual decision making and active sensory processing* was the main inspiration behind this dissertation and who has supported my research during the last four years by providing me with valuable discussions and feedback and welcoming me to her group meetings. I truly enjoyed Professor Schweinsberg's super organized classes in the Department of Mathematics and I owe him my knowledge of probability theory. I was lucky to be a negligible part of Professor Kim's journey in writing the amazing book of network information theory. I am honored to have Professor Rao on my committee. In addition, I would like to thank Professor Michèle Wigger and Professor Kamalika Chaudhuri whom I had the great opportunity to collaborate with.

I am grateful to my former and current colleagues Sheeraz Ahmad, Rathinakumar Appuswamy, Fatemeh Arbabjolfai, Bernd Bandemer, Abhijeet Bhorkar, Lorenzo Coviello, Chiao-Yi Chen, He Heung, Nikhil Karamchandani, Ali Karimian, Hessam MahdaviFar, Paolo Minero, Bakhtiyar Neyman, Matthew Pugh, Anand Sarwate, Ofer Shayevitz, Lele Wang, and Yu Xiang for many interesting research-related discussions and their warm friendship. I am especially thankful to Ehsan Ardestanizadeh for his invaluable friendship and support since the very first days I came to the United States.

Finally, it is my greatest honor to thank my family for their unconditional love and endless support throughout my entire life. My mother, father, and brother have been a source of encouragement and inspiration without which I could not have achieved any of my most proud accomplishments. Most importantly, I would like to thank my fiancé, Parastoo, for sharing with me the ups and downs of life and the hardships of long-distance relationship.

Chapter 2, in full, is a reprint of the material as it is scheduled to appear in M. Naghshvar and T. Javidi, “Sequentiality and adaptivity gains in active hypothesis testing,” *IEEE Journal of Selected Topics in Signal Processing*, October 2013, available on arXiv:1211.2291. The dissertation author was the primary investigator and author of this paper. Chapter 3, in part, has been submitted for publication as M. Naghshvar and T. Javidi, “Active sequential hypothesis testing,” available on arXiv:1203.4626. The dissertation author was the primary investigator and author of this paper. Chapter 4, in part, has been submitted for publication as M. Naghshvar and T. Javidi, “Active sequential hypothesis testing,” available on arXiv:1203.4626. The dissertation author was the primary investigator and author of this paper. Chapter 5, in full, is currently being prepared for submission for publication as M. Naghshvar, T. Javidi, and M. Wigger, “Extrinsic Jensen-Shannon divergence: applications to variable-length coding.” The dissertation author was the primary investigator and author of this material. Chapter 7, in full, is currently being prepared for submission for publication as M. Naghshvar, T. Javidi, and K. Chaudhuri, “Noisy Bayesian active learning.” The dissertation author was the primary investigator and author of this paper.

## VITA

|      |                                                                                                         |
|------|---------------------------------------------------------------------------------------------------------|
| 2007 | B.S. in Electrical Engineering, Sharif University of Technology                                         |
| 2009 | M.S. in Electrical Engineering (Communication Theory and Systems), University of California, San Diego  |
| 2013 | Ph.D. in Electrical Engineering (Communication Theory and Systems), University of California, San Diego |

## PUBLICATIONS

A. Bhorkar, M. Naghshvar, T. Javidi, and B. Rao, “Adaptive opportunistic routing for wireless ad-hoc networks,” *IEEE/ACM Transactions on Networking*, vol. 20, no. 1, pp. 243–256, February 2012.

M. Naghshvar, H. Zhuang, and T. Javidi, “A general class of throughput optimal routing policies in multi-hop wireless networks,” *IEEE Transactions on Information Theory*, vol. 58, no. 4, pp. 2175–2193, April 2012.

M. Naghshvar and T. Javidi, “Sequentiality and adaptivity gains in active hypothesis testing,” to appear in *IEEE Journal of Selected Topics in Signal Processing*, October 2013. Available on arXiv:1211.2291.

M. Naghshvar and T. Javidi, “Active sequential hypothesis testing,” submitted for publication. Available on arXiv:1203.4626.

M. Naghshvar, T. Javidi, and M. Wigger, “Extrinsic Jensen-Shannon divergence: applications to variable-length coding,” in preparation.

M. Naghshvar, T. Javidi, and K. Chaudhuri, “Noisy Bayesian active learning,” in preparation.

ABSTRACT OF THE DISSERTATION

**Active Learning and Hypothesis Testing**

by

Mohammad Naghshvar

Doctor of Philosophy in Electrical Engineering  
(Communication Theory and Systems)

University of California, San Diego, 2013

Professor Tara Javidi, Chair

This dissertation considers a generalization of the classical hypothesis testing problem. Suppose there are  $M$  hypotheses of interest among which only one is true. A Bayesian decision maker is responsible to collect observation samples so as to enhance his information about the true hypothesis in a speedy manner while accounting for the penalty of wrong declaration. In contrast to the classical hypothesis testing problem, at any given time, the decision maker can choose one of the available sensing actions and hence, exert some control over the collected samples' "information content." This generalization, referred to as the *active* hypothesis testing, naturally arises in a broad spectrum of applications such as medical diagnosis, cognition, communication, sensor management, image inspection,

generalized search, and group testing.

The first part of the dissertation provides a theoretical analysis of the problem of active hypothesis testing. Using results in sequential analysis and dynamic programming, lower bounds for the optimal performance are established. The lower bounds are complementary for various values of the parameters of the problem, and characterize the fundamental limits on the maximum achievable information acquisition rate and the optimal reliability. Moreover, upper bounds are obtained via an analysis of the proposed heuristic policies for dynamic selection of actions. From the obtained bounds, sufficient conditions are provided under which the maximum information acquisition rate and reliability are achieved, establishing the asymptotic optimality of the proposed heuristics.

The second part of the dissertation investigates the applications of the first part for three important special cases of the active hypothesis testing. Chapter 5 considers the problem of conveying a message over discrete memoryless channels with noiseless feedback. Chapter 6 studies the problem of two-dimensional search to locate a target in an image against a background of distractors. Finally, in Chapter 7, the problem of active learning for multiclass classification is investigated where the outcomes of label queries are corrupted by noise. In each of these chapters, the results in the first part of the dissertation are specialized, new results are obtained, and many of the known results are recovered with concise proofs.

# Chapter 1

## Introduction

This dissertation considers a generalization of the classical hypothesis testing problem due to Wald [1]. Suppose there are  $M$  hypotheses among which only one is true. A Bayesian decision maker is responsible to enhance his information about the true hypothesis in a speedy manner with a small number of samples while accounting for the penalty of wrong declaration. In contrast to the classical  $M$ -ary hypothesis testing problem [2–4], at any given time, our decision maker can choose one of  $K$  available actions and hence, exert some control over the collected samples’ “information content.” We refer to this generalization, originally tackled by Chernoff [5], as the *active* hypothesis testing problem. The active hypothesis testing problem naturally arises in a broad spectrum of applications such as medical diagnosis [6], cognition [7], sensor management [8], underwater inspection [9], generalized search [10], group testing [11], and channel coding with feedback [12].

The most well known instance of our problem is the case of binary hypothesis testing with passive sensing ( $M = 2$ ,  $K = 1$ ), first studied by Wald [1]. In this instance of the problem, the optimal action at any given time is provided by a sequential probability ratio test (SPRT). There are numerous studies on the generalizations to  $M > 2$  ( $K = 1$ ) and the performance of known simple and practical heuristic tests such as MSPRT [2–4]. The generalization to the active testing case was considered by Chernoff in [5] where a heuristic randomized policy was proposed and whose asymptotic performance was analyzed. More specifically, under a certain technical assumption on uniformly distinguishable hypotheses, the pro-

posed heuristic policy is shown to achieve asymptotic optimality; where the notion of asymptotic optimality [5] implies that the relative difference between the total cost associated with the proposed policy and the optimal total cost approaches zero as the penalty of wrong declaration (hence the number of collected samples) increases.

The policies under which the decision maker selects sensing actions can be categorized based on the following two factors: i) sequential vs. non-sequential; ii) adaptive vs. non-adaptive. Non-sequential policies collect a fixed number of observation samples and make the final decision afterwards; while under sequential policies, the sample size is not fixed a priori and is determined by the observation outcomes. Under adaptive policies, the decision maker relies on the previous collected samples to select the next sensing action; while under non-adaptive policies, the actions are selected independent of the past observation outcomes. In Chapter 2, following Chernoff’s approach and his notion of asymptotic optimality, we provide performance bounds for the policies in each category in terms of the penalty of wrong declaration. Using these bounds, *sequentiality gain* and *adaptivity gain*, i.e., the gains of sequential and adaptive selection of actions are characterized. In particular, these findings show a logarithmic sequentiality gain in all cases and an additional logarithmic adaptivity gain in a large class of practically relevant cases.

The problem of active hypothesis testing also generalizes another classic problem in the literature: the comparison of experiments first introduced by Blackwell [13]. This is a single-shot version of the active hypothesis testing problem in which the decision maker can choose one of several (usually two) actions/experiments to collect a single observation sample before making the final decision. There have been extensive studies [13–19] on comparing the actions and identifying conditions under which one action is “sufficient” for others, i.e., by using the observations from a particular sensing action, the decision maker can do at least as well as by using the observations from other actions.

The problem above becomes more challenging when the decision maker has to select a sequence of actions and make the final decision afterwards. In this case, each sensing action has a dual role: 1) it immediately reduces the uncertainty

in the decision maker’s belief; and 2) it shapes the future belief via Bayes’ rule. To achieve the best performance, it is essential that the decision maker, in each step, selects an action that provides the “highest amount of information,” where an appropriate notion of information must capture the influence that every action has over the entire decision making horizon. Inspired by this view of the problem which coincides with that promoted by DeGroot [20], in Chapter 3, we derive the optimal notion of information in the context of active hypothesis testing based on the optimal value function of the corresponding dynamic programming. Although this result provides a general and structural characterization of the (Markov and deterministic) optimal policy, in the absence of a closed-form for the optimal value function, it is of little use in identifying achievable schemes. To address this issue, we consider alternative notions of information namely Jensen-Shannon (JS) divergence and Extrinsic Jensen-Shannon (EJS) divergence, and propose two heuristic policies  $\pi_{JS}$  and  $\pi_{EJS}$  based on greedy maximization of these divergences. Via numerical and theoretical analysis, the performance of the proposed policies and the relevance of these alternative notions of information in the context of active hypothesis testing are investigated.

Chapter 4 provides a detailed analysis of the schemes proposed in Chapter 3 and discusses their advantage and disadvantage compared to those of Chapter 2. Unlike the schemes proposed in Chapter 2, policy  $\pi_{EJS}$  is shown to achieve asymptotic optimality (as the penalty of wrong declaration increases) only in a limited setting. However, this policy has a provable advantage for large number of hypotheses,  $M$ , over those proposed in Chapter 2 as well as other solutions in the literature. More specifically, this policy can provide, under a technical condition, reliability and speedy declaration simultaneously. In information theoretic terms, this policy can be shown to achieve non-zero information acquisition rate and error exponent simultaneously.

The second part of the dissertation investigates the applications of the results obtained in the first part. In particular, the second part of the dissertation consists of three chapters focusing on three important special cases of the active hypothesis testing.

Chapter 5 considers the problem of variable-length coding over a discrete memoryless channel (DMC) with noiseless feedback. It is shown that this problem is a special case of the active hypothesis testing and using the results of Chapter 4, a rate–reliability test is provided. Any variable-length coding scheme that satisfies the conditions of the test is guaranteed to achieve the capacity and optimal error exponent. Furthermore, it is shown that policy  $\pi_{EJS}$ , when specialized to channel coding with feedback, provides a sequential deterministic coding scheme that satisfies the conditions of the rate–reliability test and hence, achieves the capacity and optimal error exponent of any DMC. This scheme has only one phase, in contrast to all previous coding schemes in the literature which required two different phases of operation to achieve the capacity and optimal error exponent. In Chapter 5, we also prove that variable-length posterior matching achieves the capacity. In case of a class of symmetric channels with binary inputs, we generalize the schemes of Horstein [21] and Burnashev–Zigangirov [22], and we propose simple deterministic one-phase schemes to achieve the capacity and optimal error exponent.

Chapter 6 considers the problem of sequentially searching for a single target in an image where the goal is to find the target quickly and accurately. In each step, the player can visually inspect an allowable combination of the segments, and the outcome of the inspection is noisy. This problem is closely related to the problems of *fault detection*, *whereabouts search*, and *group testing*. One possible search strategy for these problems is the maximum likelihood policy which inspects a segment with the highest probability of having the target. However, as the number of segments increases, the scheme becomes inefficient. In such a case, it is more intuitive to initially inspect larger areas and narrow down the search to single segments only after we have collected sufficient information supporting the presence of the target in those segments. Following this intuition, policy  $\pi_{EJS}$  is considered as a candidate search strategy and its performance is compared against that of the maximum likelihood policy via numerical and asymptotic analysis.

In Chapter 7, we consider the problem of noisy Bayesian active learning given a sample space, a finite label set, and a finite set of label generating functions from the sample space to the label set, also known as the function class. The

objective is to identify the function in the function class that generates the labels using as few label queries as possible and with low probability of error despite possible corruption by independent noise. The key to achieving this objective relies on the selection of queries in a strategic and adaptive manner. The problem generalizes the problem of noisy generalized binary search [10]. We explore the connection between the above Bayesian active learning problem and the problem of active hypothesis testing. This view of the problem allows for developing a general lower bound on the expected number of queries needed to identify the function that generates the labels in terms of the observation noise statistics, the desired probability of error, and the cardinality of the function class. Furthermore, we compare the performance of  $\pi_{EJS}$  with the state of the art strategies for noisy generalized binary search and for different function classes. In the case where the function class is *sample-rich*, it is shown that  $\pi_{EJS}$  is better than existing results in the literature and, in particular, matches the earlier proposed lower bound asymptotically.

Notations: Let  $[x]^+ = \max\{x, 0\}$ . The indicator function  $\mathbf{1}_{\{A\}}$  takes the value 1 whenever event  $A$  occurs, and 0 otherwise. The  $i^{\text{th}}$  element of vector  $\mathbf{v}$  is denoted by  $v_i$ . For any set  $\mathcal{S}$ ,  $|\mathcal{S}|$  denotes the cardinality of  $\mathcal{S}$ . For a set  $\mathcal{A}$ ,  $\mathbb{P}(\mathcal{A})$  denotes the collection of all probability distributions on elements of  $\mathcal{A}$ , i.e.,  $\mathbb{P}(\mathcal{A}) = \{\boldsymbol{\lambda} \in [0, 1]^{|\mathcal{A}|} : \sum_{a \in \mathcal{A}} \lambda_a = 1\}$ . Little-o notation  $o(f(x))$  represents positive terms such that  $\frac{o(f(x))}{f(x)} \rightarrow 0$  as  $x \rightarrow \infty$ . Big-O notation  $O(f(x))$  represents positive terms such that  $O(f(x)) \leq cf(x)$ ,  $\forall x > x_0$ , for some constants  $c$  and  $x_0$ . All logarithms are in base 2. The *Kullback-Leibler (KL) divergence* between two probability density functions  $q(\cdot)$  and  $q'(\cdot)$  on space  $\mathcal{Z}$  is defined as  $D(q\|q') = \int_{\mathcal{Z}} q(z) \log \frac{q(z)}{q'(z)} dz$ , with the convention  $0 \log \frac{a}{0} = 0$  and  $b \log \frac{b}{0} = \infty$  for  $a, b \in [0, 1]$  with  $b \neq 0$ . The *Rényi divergence of order  $\alpha$* ,  $\alpha \in [0, 1]$ , between two probability density functions  $q(\cdot)$  and  $q'(\cdot)$  on space  $\mathcal{Z}$  is denoted by  $D_{\alpha}(q\|q')$  where  $D_{\alpha}(q\|q') = \frac{-1}{1-\alpha} \log \int_{\mathcal{Z}} q^{\alpha}(z) q'^{1-\alpha}(z) dz$  for  $\alpha \in [0, 1)$ , and  $D_{\alpha}(q\|q') = D(q\|q')$  for  $\alpha = 1$ . Finally, let  $N(m, \sigma^2)$  denote a normal distribution with mean  $m$  and variance  $\sigma^2$ .

## Chapter 2

# Sequentiality and Adaptivity Gains

Consider the problem of active hypothesis testing introduced in Chapter 1 where a decision maker is responsible to collect observations so as to enhance his information about an underlying phenomena of interest in a speedy manner. The policies under which the decision maker selects sensing actions can be categorized based on the following two factors: i) sequential vs. non-sequential; ii) adaptive vs. non-adaptive. Non-sequential policies collect a fixed number of observation samples and make the final decision afterwards; while under sequential policies, the sample size is not known initially and is determined by the observation outcomes. Under adaptive policies, the decision maker relies on the previous collected samples to select the next sensing action; while under non-adaptive policies, the actions are selected independent of the past observation outcomes.

In this chapter, performance bounds are provided for the policies in each category. Using these bounds, *sequentiality gain* and *adaptivity gain*, i.e., the gains of sequential and adaptive selection of actions are characterized.

## 2.1 Introduction

Consider the problem of active hypothesis testing introduced in Chapter 1 where a decision maker is responsible to collect observations so as to enhance his information in a speedy manner about an underlying phenomena of interest. The sample size and the sensing actions can be selected either based on the past observation outcomes (on-line or closed loop) or independent from them (off-line or open loop). According to this fact, the solutions are divided into four categories based on the following two factors: i) sequential vs. non-sequential; ii) adaptive vs. non-adaptive. Non-sequential schemes collect a fixed number of observation samples and make the final decision afterwards; while under sequential schemes, the sample size is not set in advance and instead is determined by the specific observations made. Under adaptive policies, the decision maker relies on the previous collected samples to select the next sensing action; while under non-adaptive policies, the actions are selected independent of the past observation outcomes. A question of both theoretical and practical significance is the characterization of the benefits of making sequential and adaptive decisions relative to the non-sequential and non-adaptive solutions.

Due to the importance of the question, such gains have been characterized for many special cases of the active hypothesis testing problem [9, 23, 24]. For instance, in [23] and [24], simple sequential and adaptive high dimensional reconstruction and sparse recovery are shown to significantly outperform the best non-sequential non-adaptive solutions. In contrast, [9] identifies scenarios where the gain in practice is insignificant. In this chapter, we consider the problem of active hypothesis testing in its full generality and provide upper and lower bounds on the expected cost of the optimal sensing selection strategies in sequential and non-sequential as well as adaptive and non-adaptive classes of policies. Furthermore, the bounds are shown to be asymptotically tight and logarithmically increasing in the penalty of wrong declaration (or equivalently logarithmically decreasing in the error probability).

As simple corollaries, we provide a full characterization of the sequentiality and adaptivity gains in the general active hypothesis testing framework. These

findings generalize and extend those of [23] and [24] by showing a logarithmic sequentiality gain in all cases and an additional logarithmic adaptivity gain in a large class of practically relevant cases. Furthermore, the results prove, as a corollary, the conjecture given in [9] on the insignificance of adaptivity gain when there exists a “most informative” sensing action which is independent of the Bayesian prior. Finally, we specialize our results in the active binary hypothesis testing case and state a simple necessary and sufficient condition for a logarithmic adaptivity gain.

This work and analysis is closely related and complimentary to a growing body of literature on hypothesis testing [1–3, 5, 25–30]. We discuss the specific contributions and connections in Section 2.2.4.

The remainder of this chapter is organized as follows. In Section 2.2, we formulate the problem and define various types of policies for selecting actions. Section 2.3 provides the main results of this chapter which are upper and lower bounds on the optimal expected cost under sequential and non-sequential as well as adaptive and non-adaptive classes of policies. Section 2.4 discusses the advantages of sequential and adaptive selection of actions. In Section 2.5, active binary hypothesis testing is investigated as a special case, and a necessary and sufficient condition for a logarithmic adaptivity gain is provided.

## 2.2 Problem Setup

In Section 2.2.1, we formulate the problem of active hypothesis testing. Section 2.2.2 discusses different types of policies for selecting actions. Section 2.2.3 explains why active hypothesis testing is a partially observable Markov decision problem (POMDP) and provides the sufficient statistic for this problem. Finally, in Section 2.2.4, we state our main contributions and provide a summary of related works.

### 2.2.1 Problem Formulation

Here, we provide a precise formulation for the active  $M$ -ary hypothesis testing problem.

Let  $\Omega = \{1, 2, \dots, M\}$ . Let  $H_i$ ,  $i \in \Omega$ , denote  $M$  hypotheses of interest among which only one holds true. Let  $\theta$  be the random variable that takes the value  $\theta = i$  on the event that  $H_i$  is true for  $i \in \Omega$ . We consider a Bayesian scenario with a given prior (belief) about  $\theta$ , i.e., initially  $P(\theta = i) = \rho_i(0) > 0$  for all  $i \in \Omega$ .  $\mathcal{A}$  is the set of all sensing actions and is assumed to be finite with  $|\mathcal{A}| = K < \infty$ .  $\mathcal{Z}$  is the *observation space*. For all  $a \in \mathcal{A}$ , the observation kernel  $q_i^a(\cdot)$  (on  $\mathcal{Z}$ ) is the probability density function for observation  $Z$  when action  $a$  has been taken and  $H_i$  is true. We assume that observation kernels  $\{q_i^a(\cdot)\}_{i \in \Omega, a \in \mathcal{A}}$  are known, and the observations are conditionally independent over time. Let  $L$ ,  $L > 1$ , denote the penalty for a wrong declaration, i.e., the penalty of selecting  $H_j$ ,  $j \neq i$ , when  $H_i$  is true.<sup>1</sup> Let  $\tau$  be the (stopping) time at which the decision maker retires. The objective is to find a stopping time  $\tau$ , a sequence of sensing actions  $A(0), A(1), \dots, A(\tau - 1)$ , and a declaration rule  $d : \mathcal{A}^\tau \times \mathcal{Z}^\tau \rightarrow \Omega$  that collectively minimize the expected total cost

$$\mathbb{E} \left[ \tau + L \mathbf{1}_{\{d(A^{\tau-1}, Z^{\tau-1}) \neq \theta\}} \right], \quad (2.1)$$

where  $A^{\tau-1} = [A(0), \dots, A(\tau-1)]$ ,  $Z^{\tau-1} = [Z(0), \dots, Z(\tau-1)]$ , and the expectation is taken with respect to the initial belief on  $\theta$  as well as the distributions of action sequence, observation sequence, and the stopping time.

Note that in the above problem, the cost of a test is stated in terms of minimizing the expected sample size plus the expected penalty of wrong declaration. We are interested in the characterization of this cost as a function of penalty  $L$ . It is easy to show that under the optimal selection rule, the probability of error approaches zero as  $L$  approaches infinity. In particular, it can be shown that the above problem is (asymptotically) equivalent to the problem of minimizing the (expected) number of samples subject to a constraint  $\epsilon = (L \log L)^{-1}$  on the probability of error. This connection will be discussed in more detail in Section 4.5.3.

---

<sup>1</sup>In general, we can define a loss matrix  $[L_{ij}]_{i,j \in \Omega}$ , where  $L_{ij}$  denotes the penalty (loss) of selecting  $H_j$  when  $H_i$  is true.

### 2.2.2 Types of Policies

A *policy* is a rule based on which stopping time  $\tau$  and sensing actions  $A(t)$ ,  $t = 0, 1, \dots, \tau - 1$  are selected. We assume that sensing actions are selected according to a randomized rule  $\boldsymbol{\lambda} \in \mathbb{P}(\mathcal{A})$  whose element  $\lambda_a$  indicates the probability of selecting sensing action  $a$  and in general might change with time. The sensing actions and the stopping time can be selected either based on the past observation outcomes or independent from them. According to this fact, policies are divided into four categories based on the following two factors: i) sequential vs. non-sequential; ii) adaptive vs. non-adaptive. Non-sequential policies collect a fixed number of observation samples and make the final decision afterwards; while under sequential policies, the sample size is not known initially and is determined by the observation outcomes. More precisely, under non-sequential policies,  $\tau = N$  for some  $N \in \mathbb{N}$ ; while for sequential policies,  $\tau$  is a random stopping time. Under adaptive policies, the decision maker relies on the previous collected samples to select the next sensing action; while under non-adaptive policies, the actions are selected independent of the past observation outcomes and according to a fixed randomized rule.

### 2.2.3 Information State as Sufficient Statistic

The problem of active  $M$ -ary hypothesis testing is a partially observable Markov decision problem (POMDP) where the state is static and observations are noisy. It is known that any POMDP is equivalent to an MDP with a compact yet uncountable state space, for which the belief of the decision maker about the underlying state becomes an information state [31]. In our setup, thus, the information state at time  $t$  is a belief vector specified by the conditional probability of hypotheses  $H_1, H_2, \dots, H_M$  to be true given the initial belief and all the previous actions and observations. Let  $\boldsymbol{\rho}(t) := [\rho_1(t), \rho_2(t), \dots, \rho_M(t)]$  denote the posterior belief after  $t$  observations where  $\rho_i(t) := P(\{\theta = i\} | A^{t-1}, Z^{t-1})$ . Accordingly, the information state space is defined as  $\mathbb{P}(\Omega) = \{\boldsymbol{\rho} \in [0, 1]^M : \sum_{i=1}^M \rho_i = 1\}$ . In one sensing step, the evolution of the belief vector follows Bayes' rule and the expected

total cost (2.1) can be rewritten as<sup>2</sup>

$$\mathbb{E}[\tau] + L\text{Pe}, \quad (2.2)$$

where  $\text{Pe} = \mathbb{E}[1 - \max_{i \in \Omega} \rho_i(\tau)]$  is the probability of wrong declaration, and the expectations are taken with respect to the initial prior distribution on  $\theta$  as well as the distributions of action sequence, observation sequence, and the stopping time.

Let  $V_{NN}(\boldsymbol{\rho})$ ,  $V_{SN}(\boldsymbol{\rho})$ ,  $V_{SA}(\boldsymbol{\rho})$ , and  $V_{NA}(\boldsymbol{\rho})$  denote the minimum expected total cost (2.2) for prior belief  $\boldsymbol{\rho}$  under non-sequential non-adaptive, sequential non-adaptive, sequential adaptive, and non-sequential adaptive policies, respectively.

## 2.2.4 Overview of the Results and Literature Survey

Active hypothesis testing generalizes the passive (classical) hypothesis testing problem where the number of sensing actions is limited to one, both in the non-sequential (fixed sample size) case [26, 27] and the sequential (variable sample size) one [1–3]. While the fixed sample size studies have primarily focused on the asymptotic analysis in form of identifying error exponents for various error types [26, 27], the study of sequential hypothesis testing has come in form of identifying the expected optimal sample size to achieve a given error probability [1–3].

The generalization to the active testing case was considered by Chernoff in [5] in which a decision maker controls sensing actions to optimize the expected total cost (2.2) in a sequential (variable sample size) setting. In particular, in [5] and its extensions [25, 32–38], heuristic sequential adaptive randomized policies were proposed and were shown to be asymptotically optimal as  $L \rightarrow \infty$  where the notion of asymptotic optimality [5] denotes the relative tightness of the performance upper bound associated with the proposed policy and the lower bound associated with the optimal policy.<sup>3</sup>

---

<sup>2</sup>The probability of error is minimized under the maximum a posteriori (MAP) declaration.

<sup>3</sup>In [5], the objective was to minimize  $c\mathbb{E}[\tau] + \text{Pe}$  and the proposed policy was shown to be asymptotically optimal as  $c \rightarrow 0$ . It is straightforward to show that for  $L = \frac{1}{c}$ , this problem coincides with the active hypothesis testing problem defined here. However, we have chosen  $\mathbb{E}[\tau] + L\text{Pe}$  as an objective function here because of its interpretation as the Lagrangian relaxation of an information acquisition problem in which the objective is to minimize  $\mathbb{E}[\tau]$  subject to  $\text{Pe} \leq \epsilon$  where  $\epsilon > 0$  denotes the desired probability of error.

The general active binary hypothesis testing problem was recently studied in [28, 29] where full characterization of the error exponent corresponding to the class of non-sequential adaptive and non-sequential non-adaptive policies was provided. In particular, the error exponent corresponding to these two classes was shown to be equal, hence establishing zero adaptivity gain among non-sequential policies. The generalization to  $M > 2$  was considered in [30]. Note that while [30] fully characterizes the error exponent corresponding to non-sequential non-adaptive policies; it provides only a partial characterization of (i.e., loose upper and lower bounds on) the error exponent corresponding to non-sequential adaptive policies.

Table 2.1 provides a visual summary of the literature on hypothesis testing.

**Table 2.1:** Hypothesis Testing Literature

| Type                               | $M = 2$     | $M > 2$                |
|------------------------------------|-------------|------------------------|
| Sequential Passive ( $K = 1$ )     | [1]         | [2, 3]                 |
| Sequential Non-adaptive            | [29, 37]    | see below <sup>4</sup> |
| Sequential Adaptive                | [5, 29, 37] | [5, 25, 38]            |
| Non-sequential Passive ( $K = 1$ ) | [26]        | [27]                   |
| Non-sequential Non-adaptive        | [28, 29]    | [30]                   |
| Non-sequential Adaptive            | [28, 29]    | [30]                   |

We close our literature survey with an overview of the main contributions:

- We provide asymptotically tight lower and upper bounds on  $V_{NN}(\boldsymbol{\rho})$ ,  $V_{SN}(\boldsymbol{\rho})$ , and  $V_{SA}(\boldsymbol{\rho})$  which hold uniformly for all prior  $\boldsymbol{\rho} \in \mathbb{P}(\Omega)$ .
  - The asymptotic tight bounds on  $V_{NN}(\boldsymbol{\rho})$  relies on the analysis of [26, 27], and the realization that in order to minimize the total cost, we have to decrease the error probabilities of various types with the same exponent

<sup>4</sup>The class of sequential non-adaptive policies studied here has received little attention in the literature. To the best of our knowledge, the result on sequential non-adaptive policies is new and has not been established before.

among the worst pair of hypotheses. Since unlike the passive case studied in [26, 27], the non-adaptive policies produce non-i.i.d. observation samples, the final step is to characterize the relationship between the error exponent of a fixed block length and one-step error exponent.

- The asymptotic tight bounds on  $V_{SN}(\boldsymbol{\rho})$  extend the results obtained by [3] to the Bayesian context while allowing for randomized non-adaptive policies. More specifically, the result of [3] is obtained via the law of large numbers and only holds if the observations are i.i.d. Since observations are not identical (although they are independent), a different proof technique is required (note that unlike the non-sequential case of extending the work of [26, 27], the random nature of sample size in the sequential case does not allow for a predetermined relationship between the error exponent of a fixed block and the one-step error exponent).
- The asymptotic tight bounds on  $V_{SA}(\boldsymbol{\rho})$  extend those obtained by [5, 25] to the Bayesian context while relaxing the assumption on uniform discrimination of hypotheses or the need for the infinitely often reliance on a randomized action to ensure sufficient discrimination among hypotheses. We will study the class of sequential adaptive policies in more details in the next chapters. In particular, we will discuss the drawbacks of Chernoff’s notion of asymptotic optimality in  $L$ , and characterize the expected total cost under sequential adaptive policies in terms of  $L$  and  $M$  simultaneously.
- In addition, we partially characterize a lower bound for  $V_{NA}(\boldsymbol{\rho})$ . This is, in the Bayesian context, similar to the partial characterization of error exponent of [30].
- As corollaries to the above performance bounds, we characterize the sequentiality gain and the adaptivity gain in terms of  $L$ . In particular, it is shown that the sequentiality gain grows logarithmically as the penalty  $L$  increases. We also state a simple necessary and sufficient condition ensuring a logarithmic adaptivity gain in  $L$  for the active binary hypothesis testing case.

- Furthermore, primarily as a sanity check, in Section 2.4.2 we derive the maximum achievable error exponents  $E_{NN}$ ,  $E_{SN}$ , and  $E_{SA}$  in the Bayesian context. In particular, our result regarding  $E_{NN}$  coincides with that of [28–30]; while the result regarding  $E_{SA}$  coincides with that of [5, 25] in the Bayesian context. To the best of our knowledge, the result on  $E_{SN}$  is new and has not been established before; while our upper bound on  $E_{NA}$  is subsumed by the analysis in [30].

## 2.3 Analytic Results

In this section, we provide our main results regarding the asymptotic characterization (in  $L$ ) of  $V_{NN}(\boldsymbol{\rho})$ ,  $V_{SN}(\boldsymbol{\rho})$ ,  $V_{SA}(\boldsymbol{\rho})$ , and  $V_{NA}(\boldsymbol{\rho})$ .

### 2.3.1 Assumptions and Basic Definitions

We have the following technical Assumptions.

**Assumption 2.3.1.** *For any two hypotheses  $i, j \in \Omega$ ,  $i \neq j$ , there exists an action  $a$ ,  $a \in \mathcal{A}$ , such that  $D(q_i^a \| q_j^a) > 0$ .*

**Assumption 2.3.2.** *There exists  $\xi < \infty$  such that*

$$\max_{i,j \in \Omega} \max_{a \in \mathcal{A}} \sup_{z \in \mathcal{Z}} \log \frac{q_i^a(z)}{q_j^a(z)} \leq \xi.$$

Assumption 2.3.1 ensures the possibility of discrimination between any two hypotheses. Assumption 2.3.2 implies that no two hypotheses are fully distinguishable using a single observation sample.

To continue with our analysis, we need the following definitions and notations.

**Definition 2.3.1.** For all  $i \in \Omega$  and  $\boldsymbol{\lambda} \in \mathbb{P}(\mathcal{A})$ , the *optimized discrimination* of hypothesis  $i$  under randomized rule  $\boldsymbol{\lambda}$  is defined as

$$D_{\alpha^*}(i, \boldsymbol{\lambda}) := \min_{j \neq i} \max_{\alpha \in [0,1]} (1 - \alpha) \sum_{a \in \mathcal{A}} \lambda_a D_{\alpha}(q_i^a \| q_j^a).$$

**Definition 2.3.2.** For all  $i \in \Omega$  and  $\boldsymbol{\lambda} \in \mathbb{P}(\mathcal{A})$ , the *reliability* function of hypothesis  $i$  with regard to randomized rule  $\boldsymbol{\lambda}$  is defined as

$$R(i, \boldsymbol{\lambda}) := \min_{j \neq i} \sum_{a \in \mathcal{A}} \lambda_a D(q_i^a \| q_j^a),$$

and the maximal randomized rule for hypothesis  $i$  is denoted by

$$\boldsymbol{\lambda}_i^* := \arg \max_{\boldsymbol{\lambda} \in \mathbb{P}(\mathcal{A})} R(i, \boldsymbol{\lambda}).$$

For  $\boldsymbol{\lambda} \in \mathbb{P}(\mathcal{A})$ , let  $\bar{R}(\boldsymbol{\lambda})$  denote the harmonic mean of  $\{R(i, \boldsymbol{\lambda})\}_{i \in \Omega}$ , i.e.,

$$\bar{R}(\boldsymbol{\lambda}) := \frac{M}{\sum_{i=1}^M \frac{1}{R(i, \boldsymbol{\lambda})}},$$

and let  $\bar{R}^*$  denote the harmonic mean of  $\{R(i, \boldsymbol{\lambda}_i^*)\}_{i \in \Omega}$ , i.e.,

$$\bar{R}^* := \frac{M}{\sum_{i=1}^M \frac{1}{R(i, \boldsymbol{\lambda}_i^*)}}.$$

These notions of discrimination and reliability, as we will see, are natural (and Bayesian) extensions of reliability in classical detection [39] where the reliability function for hypothesis  $i$  is related to type  $i$  error probability. The following lemma enables a concrete relationship between these notions.

**Lemma 2.3.1.** *For two probability density functions  $q(\cdot)$  and  $q'(\cdot)$  with the same support and for all  $\alpha \in [0, 1]$  we have*

$$(1 - \alpha)D_\alpha(q \| q') \leq \min \{(1 - \alpha)D(q \| q'), \alpha D(q' \| q)\}.$$

*Proof.* For probability distributions defined over a finite alphabet, Lemma 2.3.1 is a consequence of Theorem 1 in [40]; while for general distributions, it is a consequence of Theorem 27 in [41]. Here we provide a simple proof.

Let  $q(\cdot)$  and  $q'(\cdot)$  be two probability density functions on space  $\mathcal{Z}$  with the same support. The inequality holds trivially for  $\alpha = 1$ . For  $\alpha \in [0, 1)$ , we have

$$\begin{aligned} D_\alpha(q \| q') &= \frac{-1}{1 - \alpha} \log \int_{\mathcal{Z}} q(z) \left( \frac{q'(z)}{q(z)} \right)^{1-\alpha} dz \\ &\stackrel{(a)}{\leq} \frac{-1}{1 - \alpha} \int_{\mathcal{Z}} q(z) \log \left( \frac{q'(z)}{q(z)} \right)^{1-\alpha} dz \\ &= \int_{\mathcal{Z}} q(z) \log \frac{q(z)}{q'(z)} dz = D(q \| q'), \end{aligned} \tag{2.3}$$

where (a) follows from the Jensen's inequality. Similarly, we can show that

$$\begin{aligned}
D_\alpha(q\|q') &= \frac{-1}{1-\alpha} \log \int_{\mathcal{Z}} \left( \frac{q(z)}{q'(z)} \right)^\alpha q'(z) dz \\
&\leq \frac{-1}{1-\alpha} \int_{\mathcal{Z}} q'(z) \log \left( \frac{q(z)}{q'(z)} \right)^\alpha dz \\
&= \frac{\alpha}{1-\alpha} \int_{\mathcal{Z}} q'(z) \log \frac{q'(z)}{q(z)} dz = \frac{\alpha}{1-\alpha} D(q'\|q).
\end{aligned} \tag{2.4}$$

Combining (2.3) and (2.4), we have the assertion of the Lemma.  $\square$

### 2.3.2 Main Theorems

In this section, we provide upper and lower bounds on the minimum expected total cost (2.2) under different types of policies defined in Section 2.2.2. These bounds will then be used in Section 2.4 to characterize the gains of sequential and adaptive selection of actions.

**Theorem 2.3.1** (Non-sequential non-adaptive policy). *Under Assumptions 2.3.1 and 2.3.2,*

$$V_{NN}(\boldsymbol{\rho}) \leq \frac{\log L}{\hat{D}} + o(\log L), \tag{2.5}$$

$$V_{NN}(\boldsymbol{\rho}) \geq \frac{\log L}{\hat{D}} - o(\log L), \tag{2.6}$$

where

$$\hat{D} := \max_{\boldsymbol{\lambda} \in \mathbb{P}(\mathcal{A})} \min_{i \in \Omega} D_{\alpha^*}(i, \boldsymbol{\lambda}). \tag{2.7}$$

*Proof.* The detailed proof is provided in Section 2.7.1. Here we provide an overview.

The proof of the lower bound relies on a generalization of Theorem 10 in [26]; while the upper bound is achieved via  $\pi_{NN}$ , a randomized, non-sequential, and non-adaptive policy which deterministically collects  $\hat{n} = (\log L + o(\log L))/\hat{D}$  samples and selects sensing actions according to the randomized rule  $\hat{\boldsymbol{\lambda}} \in \mathbb{P}(\mathcal{A})$  that achieves the maximum in (2.7).  $\square$

**Theorem 2.3.2** (Sequential non-adaptive policy). *Under Assumptions 2.3.1 and 2.3.2,*

$$V_{SN}(\boldsymbol{\rho}) \leq \min_{\boldsymbol{\lambda} \in \mathbb{P}(\mathcal{A})} \sum_{i=1}^M \rho_i \frac{\log L}{R(i, \boldsymbol{\lambda})} + o(\log L), \quad (2.8)$$

$$V_{SN}(\boldsymbol{\rho}) \geq \min_{\boldsymbol{\lambda} \in \mathbb{P}(\mathcal{A})} \sum_{i=1}^M \rho_i \frac{\log L}{R(i, \boldsymbol{\lambda})} - o(\log L). \quad (2.9)$$

*Proof.* The detailed proof is provided in Section 2.7.2. Here we provide an overview.

Suppose  $\hat{\boldsymbol{\lambda}} \in \mathbb{P}(\mathcal{A})$  achieves the minimum in (2.8). The upper bound (2.8) is achieved by policy  $\pi_{SN}$  that selects sensing actions according to  $\hat{\boldsymbol{\lambda}}$  and stops sampling at

$$\tau := \min\{n : \max_{i \in \Omega} \rho_i(n) \geq 1 - L^{-1}\}.$$

From upper bound (2.8) we know that the total cost under the optimal policy is  $O(\log L)$ . This implies that the error probability  $\text{Pe}$  of the optimal policy is  $O(\frac{\log L}{L})$ . Hence, without loss of generality in our proof of the lower bound, we can restrict the set of sequential and non-adaptive policies to those whose probability of making an error is  $O(\frac{\log L}{L})$ . Conditioning on the true hypothesis and considering the dynamic of pairwise likelihoods, we then compute the minimum expected number of samples necessary to achieve this target error probability.  $\square$

**Theorem 2.3.3** (Sequential adaptive policy). *Under Assumptions 2.3.1 and 2.3.2,*

$$V_{SA}(\boldsymbol{\rho}) \leq \sum_{i=1}^M \rho_i \frac{\log L}{R(i, \boldsymbol{\lambda}_i^*)} + o(\log L), \quad (2.10)$$

$$V_{SA}(\boldsymbol{\rho}) \geq \sum_{i=1}^M \rho_i \frac{\log L}{R(i, \boldsymbol{\lambda}_i^*)} - o(\log L). \quad (2.11)$$

*Proof.* The detailed proof is provided in Section 2.7.3. Here we provide an overview.

The proof of the lower bound relies on a generalization of Theorem 2 in [5]. The upper bound is achieved via  $\pi_{SA}$ , a heuristic two-phase policy which in its first phase, selects actions in a way that all pairs of hypotheses can be distinguished from each other; while its second phase coincides with Chernoff's scheme [5] where only the pairs including the most likely hypothesis are considered. The second phase of  $\pi_{SA}$  ensures its asymptotic optimality in  $L$ ; while its first phase in a

very natural manner relaxes the technical assumption in [5] where all actions are assumed to discriminate between all hypotheses pairs or the need for the infinitely often reliance on a suboptimal randomized action in order to ensure sufficient discrimination among hypotheses.  $\square$

We close this section by a note on the class of non-sequential adaptive policies even though they seem rather unnatural to us (since it is more reasonable to control the sample size using the observation outcomes if they are already being used to select sensing actions). The next proposition provides a lower bound on the minimum expected total cost under non-sequential adaptive policies, denoted by  $V_{NA}$ .

**Proposition 2.3.1** (Non-sequential adaptive policy). *Under Assumptions 2.3.1 and 2.3.2,*

$$V_{NA}(\boldsymbol{\rho}) \geq \frac{\log L}{\min_{i \in \Omega} R(i, \boldsymbol{\lambda}_i^*)} - o(\log L). \quad (2.12)$$

Next we state and discuss the consequences of the bounds proposed above. In Section 2.4.1, we focus on the advantages of causally selecting the retire-declare time as well as the adaptive selecting of sensing actions. In Section 2.4.2, we derive the error exponent corresponding to different types of policies.

## 2.4 Consequences of the Bounds

In this section, we first specialize and simplify the results provided in Section 2.3 for uniform prior. In particular, assume that the hypotheses, initially, are equally likely, i.e.,  $\rho_i(0) = \frac{1}{M}$  for all  $i \in \Omega$ . Let  $V_{NN}$ ,  $V_{SN}$ , and  $V_{SA}$ , respectively, denote the minimum expected total cost under non-sequential non-adaptive, sequential non-adaptive, and sequential adaptive policies for uniform prior, i.e.,  $V_x := V_x([\frac{1}{M}, \frac{1}{M}, \dots, \frac{1}{M}])$  where  $x$  denotes the class of policies  $NN$ ,  $SN$ , and  $SA$ .

From Fact 2.3.1, we know that

$$\hat{D} \leq 0.5 \max_{\boldsymbol{\lambda} \in \mathbb{P}(\mathcal{A})} \min_{i \in \Omega} \min_{j \neq i} \sum_{a \in \mathcal{A}} \lambda_a D(q_i^a \| q_j^a). \quad (2.13)$$

Theorem 2.3.1 together with (2.13) implies that:

**Corollary 2.4.1** (Non-sequential non-adaptive policy). *Under Assumptions 2.3.1 and 2.3.2,*

$$\begin{aligned} V_{NN} &= \frac{\log L}{\hat{D}} \pm o(\log L) \\ &\geq \frac{2 \log L}{\max_{\lambda \in \mathbb{P}(\mathcal{A})} \min_{i \in \Omega} R(i, \lambda)} - o(\log L). \end{aligned} \quad (2.14)$$

**Corollary 2.4.2** (Sequential non-adaptive policy). *Under Assumptions 2.3.1 and 2.3.2,*

$$V_{SN} = \frac{\log L}{\max_{\lambda \in \mathbb{P}(\mathcal{A})} \bar{R}(\lambda)} \pm o(\log L). \quad (2.15)$$

**Corollary 2.4.3** (Sequential adaptive policy). *Under Assumptions 2.3.1 and 2.3.2,*

$$V_{SA} = \frac{\log L}{\bar{R}^*} \pm o(\log L). \quad (2.16)$$

From the results above, it is evident that the minimum expected total cost under all classes of policies grows logarithmically in  $L$ . However, the coefficient of the  $\log L$  term is not the same in general and we have

$$\bar{R}^* \geq \max_{\lambda \in \mathbb{P}(\mathcal{A})} \bar{R}(\lambda) \geq \max_{\lambda \in \mathbb{P}(\mathcal{A})} \min_{i \in \Omega} R(i, \lambda) \geq \hat{D}. \quad (2.17)$$

## 2.4.1 Sequentiality and Adaptivity Gains

In this section, we discuss the advantage of causally selecting the retire-declare time, i.e.,  $\tau$  as well as the sensing actions.

First, we show that the performance gap between the sequential and non-sequential policy,  $V_{NN} - V_{SN}$ , grows logarithmically as the penalty  $L$  increases. We refer to this performance gap as the *sequentiality gain*.

**Corollary 2.4.4.** *Under Assumptions 2.3.1 and 2.3.2, the sequentiality gain is characterized as*

$$V_{NN} - V_{SN} \geq \log L \left( \frac{2}{\max_{\lambda \in \mathbb{P}(\mathcal{A})} \min_{i \in \Omega} R(i, \lambda)} - \frac{1}{\max_{\lambda \in \mathbb{P}(\mathcal{A})} \bar{R}(\lambda)} \right) - o(\log L).$$

**Remark 2.4.1.** The sequentiality gain grows logarithmically in  $L$  and from (2.17),

$$V_{NN} - V_{SN} \geq \frac{\log L}{\max_{\lambda \in \mathbb{P}(\mathcal{A})} \bar{R}(\lambda)} - o(\log L).$$

Next, the advantage of adaptively selecting the sensing actions is discussed. In particular, it is shown that the performance gap between the adaptive and non-adaptive policy,  $V_{SN} - V_{SA}$ , grows logarithmically as the penalty  $L$  increases. We refer to this performance gap as the *adaptivity gain*.

**Corollary 2.4.5.** *Under Assumptions 2.3.1 and 2.3.2, the adaptivity gain is characterized as*

$$V_{SN} - V_{SA} = \log L \left( \frac{1}{\max_{\lambda \in \mathbb{P}(\mathcal{A})} \bar{R}(\lambda)} - \frac{1}{\bar{R}^*} \right) \pm o(\log L).$$

**Remark 2.4.2.** Note that the simple two phase structure of  $\pi_{SA}$ , the policy that achieves the upper bound in (2.10), implies that the adaptivity gain can be obtained via coarse level adaptation.

**Remark 2.4.3.** Unless there exists a fixed randomized rule  $\tilde{\lambda} \in \mathbb{P}(\mathcal{A})$  such that,

$$R(i, \tilde{\lambda}) = R(i, \lambda_i^*) \text{ for all } i \in \Omega,$$

the adaptivity gain grows logarithmically with  $L$ .

A sufficient condition under which there is no adaptivity gain is that of stochastic dominance/degradation [13], i.e., if there exists a *stochastic transformation*  $W$  from  $\mathcal{Z}$  to  $\mathcal{Z}$  and<sup>5</sup> a sensing action  $a^*$  such that for all other sensing actions  $a \in \mathcal{A}$ ,

$$q_i^a(z) = \int q_i^{a^*}(y) W(y; z) dy, \quad \forall i \in \Omega. \quad (2.18)$$

As shown by Sakaguchi [42], (2.18) implies that

$$D(q_i^a \| q_j^a) \leq D(q_i^{a^*} \| q_j^{a^*}), \quad \forall a \in \mathcal{A}, \quad \forall i, j \in \Omega,$$

hence, ensuring zero adaptivity gain when observations obtained by all actions are stochastically degraded versions of the observation under sensing action  $a^*$ . This formalizes the notion of informativeness and confirms the conjecture provided in [9].

---

<sup>5</sup>Function  $W : \mathcal{Z} \times \mathcal{Z} \rightarrow \mathbb{R}_+$  is called a *stochastic transformation* from  $\mathcal{Z}$  to  $\mathcal{Z}$  if it satisfies  $\int_{\mathcal{Z}} W(y; z) dz = 1$  for all  $y \in \mathcal{Z}$  and  $\int_{\mathcal{Z}} W(y; z) dy < \infty$  for all  $z \in \mathcal{Z}$ .

### 2.4.2 Reliability and Error Exponent

Let  $\mathbb{E}_\pi[\tau]$  denote the expected stopping time (or equivalently the expected number of collected samples) under policy  $\pi$ . Policy  $\pi$  is said to achieve error exponent  $E > 0$  if

$$\lim_{t \rightarrow \infty} \frac{-1}{t} \log \text{Pe}_\pi(t, M) = E, \quad (2.19)$$

where  $\text{Pe}_\pi(t, M)$  is the minimum probability of error that policy  $\pi$  can guarantee for  $M$  hypotheses with the constraint  $\mathbb{E}_\pi[\tau] \leq t$  (Note that  $\tau$  is deterministic for non-sequential policies).

Next we use the bounds obtained in Section 2.3 to characterize the maximum achievable error exponent for different types of policies. Let  $E_{NN}$ ,  $E_{SN}$ ,  $E_{SA}$ , and  $E_{NA}$  denote the maximum achievable error exponent under non-sequential non-adaptive, sequential non-adaptive, sequential adaptive, and non-sequential adaptive policies.

**Corollary 2.4.6.** *Under Assumptions 2.3.1 and 2.3.2, we have*

$$\begin{aligned} E_{NN} &= \hat{D}, \\ E_{SN} &= \max_{\boldsymbol{\lambda} \in \mathbb{P}(\mathcal{A})} \bar{R}(\boldsymbol{\lambda}), \\ E_{SA} &= \bar{R}^*. \end{aligned}$$

**Remark 2.4.4.** The above characterizations of maximum achievable error exponent are nothing but the Bayesian and  $M$ -ary version of the results in the literature (see Table 2.1). In fact, as discussed in Section 2.2.4, these results provide a sanity check viz a viz the prior work:  $E_{NN}$  coincides with that of [28–30]; while  $E_{SA}$  coincides with that of [5, 25]. To the best of our knowledge, the result on  $E_{SN}$  is new and has not been established before.

**Remark 2.4.5.** The above corollary provides alternative means to underline and characterize the sequentiality and adaptivity gains. In particular, sequentiality always results in an improvement in the maximum achievable error exponent since  $E_{NN} \leq 0.5 \max_{\boldsymbol{\lambda} \in \mathbb{P}(\mathcal{A})} \min_{i \in \Omega} R(i, \boldsymbol{\lambda}) < E_{SN}$ . In contrast, adaptive selection of actions

results in an improvement in the maximum achievable error exponent only if  $\max_{\lambda \in \mathbb{P}(\mathcal{A})} \bar{R}(\lambda) \neq \bar{R}^*$ .

We can also find an upper bound on the maximum achievable error exponent of any non-sequential yet adaptive policy (tight lower bounds are necessary for full characterization, however).

**Corollary 2.4.7.** *Under Assumptions 2.3.1 and 2.3.2, we have*

$$E_{NA} \leq \min_{i \in \Omega} R(i, \lambda_i^*).$$

**Remark 2.4.6.** Our upper bound on  $E_{NA}$  is subsumed by [30, Theorem 2].

## 2.5 Special Case: Binary Hypothesis Testing

In this section, we consider active binary hypothesis testing ( $M = 2$ ) as a special case.

### 2.5.1 Analytical Results

The performance bounds provided in Section 2.3 are simplified by substituting the following equations into the denominators of the bounds.

$$\begin{aligned} D_{\alpha^*}(1, \lambda) &= \max_{\alpha \in [0,1]} (1 - \alpha) \sum_{a \in \mathcal{A}} \lambda_a D_{\alpha}(q_1^a \| q_2^a) = D_{\alpha^*}(2, \lambda), \\ \hat{D} &= \max_{a \in \mathcal{A}} \max_{\alpha \in [0,1]} (1 - \alpha) D_{\alpha}(q_1^a \| q_2^a), \\ R(1, \lambda) &= \sum_{a \in \mathcal{A}} \lambda_a D(q_1^a \| q_2^a), \quad R(2, \lambda) = \sum_{a \in \mathcal{A}} \lambda_a D(q_2^a \| q_1^a), \\ R(1, \lambda_1^*) &= \max_{a \in \mathcal{A}} D(q_1^a \| q_2^a), \quad R(2, \lambda_2^*) = \max_{a \in \mathcal{A}} D(q_2^a \| q_1^a), \\ \bar{R}(\lambda) &= \left( \frac{0.5}{\sum_{a \in \mathcal{A}} \lambda_a D(q_1^a \| q_2^a)} + \frac{0.5}{\sum_{a \in \mathcal{A}} \lambda_a D(q_2^a \| q_1^a)} \right)^{-1}, \\ \bar{R}^* &= \left( \frac{0.5}{\max_{a \in \mathcal{A}} D(q_1^a \| q_2^a)} + \frac{0.5}{\max_{a \in \mathcal{A}} D(q_2^a \| q_1^a)} \right)^{-1}. \end{aligned}$$

Next we state a simple necessary and sufficient condition for a logarithmic adaptivity gain in the active binary hypothesis testing case.

**Corollary 2.5.1.** *In the active binary hypothesis testing case, the adaptivity gain grows logarithmically in  $L$  if and only if*

$$\arg \max_{a \in \mathcal{A}} D(q_1^a \| q_2^a) \neq \arg \max_{a \in \mathcal{A}} D(q_2^a \| q_1^a).$$

The problem of passive binary hypothesis testing ( $K = 1$ ,  $M = 2$ ) with fixed-length (non-sequential) as well as variable-length (sequential) sample size has been studied by [1, 26, 39, 43]. Our sequentiality gain, in this case, is the manifestation of the fact that sequential tests are superior in ensuring that both error probabilities decrease at the best possible exponential rates [43].

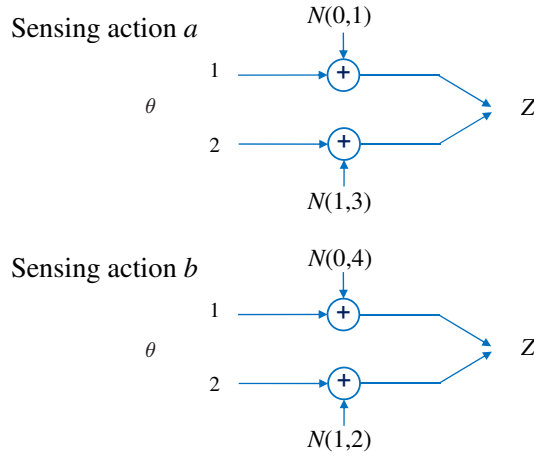
Recently, the authors in [28] and [29] have studied the problem of active binary hypothesis testing for fixed-length and variable-length sample size, respectively. Our work complements the findings in [29] by providing an asymptotic optimal solution in a total cost (and Bayesian) sense as well as establishing a non-zero sequentiality and potentially non-zero adaptivity gain. In [28], the error exponent corresponding to the class of  $NN$  and  $NA$  policies were fully characterized for the problem of active binary hypothesis testing with fixed sample size. In the Bayesian context, the result of [28] regarding the error exponent of the class of  $NN$  policies coincides with our Corollary 2.4.6, while the full characterization of the error exponent corresponding to the class of  $NA$  policies in [28], strengthens Corollary 2.4.7 in the binary case. In particular, it is shown that in the binary hypothesis testing setup  $E_{NN} = E_{NA}$ , hence, establishing zero adaptivity gain among non-sequential policies. For the special case of channel coding with feedback<sup>6</sup> with two messages, the above result, i.e., the zero adaptivity gain among non-sequential policies, was established in [44, 45].

---

<sup>6</sup>The problem of channel coding with feedback can be interpreted as a special case of active hypothesis testing (see Chapter 5 for more details).

### 2.5.2 Numerical Example

Consider the active binary hypothesis testing problem with additive Gaussian noisy observations under two actions  $a$  and  $b$  shown in Fig. 2.1. In this example, the observation noise associated with actions  $a$  and  $b$  is such that it adds unequal noise to the hypotheses. In the remainder of this section, we compare the performance of all considered policies for this example.



**Figure 2.1:** Active binary hypothesis testing with additive Gaussian noise.

The KL and Rényi divergences between two normal distributions  $q = N(m_1, \sigma_1^2)$  and  $q' = N(m_2, \sigma_2^2)$  can be computed using the following closed-form expressions [46]:

$$D(q||q') = \frac{(m_2 - m_1)^2 + \sigma_1^2 - \sigma_2^2}{2\sigma_2^2} \log e + \log \frac{\sigma_2}{\sigma_1}, \quad (2.20)$$

$$D_\alpha(q||q') = \frac{\frac{1}{2}\alpha(m_2 - m_1)^2}{(1 - \alpha)\sigma_1^2 + \alpha\sigma_2^2} \log e - \frac{1}{1 - \alpha} \log \frac{\sigma_1^{1-\alpha}\sigma_2^\alpha}{\sqrt{(1 - \alpha)\sigma_1^2 + \alpha\sigma_2^2}}. \quad (2.21)$$

Table 2.2 compares the performance bounds of the considered policies for the example of Fig. 2.1.

**Remark 2.5.1.** In practice, non-sequential and non-adaptive policies are simpler to implement compared to sequential and adaptive policies. To select the most

**Table 2.2:** Comparison of performance bounds for the example of Fig. 2.1.

|              | Sequential    | Non-sequential        |
|--------------|---------------|-----------------------|
| Adaptive     | $\log L/2.26$ | $\gtrsim \log L/1.66$ |
| Non-adaptive | $\log L/1.87$ | $\log L/0.49$         |

appropriate policy, one needs to trade off between the computational complexity and the performance. The results of this chapter can be applied to measure and compare the performance of different types of policies. For instance, in the numerical example given above and for large values of  $L$ , sequential selection of actions reduces the expected total cost by 73.7%, and adaptivity reduces the expected total cost by an additional 17.2%. In light of the complexity of Bayesian updates associated with tracking the posterior, this will provide the system designer with an (asymptotic) computation–sampling complexity tradeoff.

## 2.6 Discussion

In this chapter, we considered the problem of active hypothesis testing and we analyzed the gain of sequential and adaptive selection of actions. In our analysis, however, we only investigated asymptotic performance in  $L$ , and the complementary role of asymptotic analysis in  $M$  was neglected. In particular, we have only identified the zero-rate characterization of error exponent; while for a full characterization in which error exponent is traded off with information acquisition rate, we would need an asymptotic characterization of the problem both in  $L$  and  $M$ . We will address this issue for the class of sequential policies in Chapter 4.

## 2.7 Proofs

In this section, we provide the proof of the main results of this chapter, i.e., Theorems 2.3.1, 2.3.2, 2.3.3, and Proposition 2.3.1. Before we get into details, we give a high-level explanation on how being sequential/non-sequential and adaptive/non-adaptive changes the analysis and the proofs.

To analyze the performance of non-sequential policies, we derive upper and lower bounds on the probability of error  $\text{Pe}$  in terms of the number of collected observation samples  $n$ . However, for sequential policies, we find bounds on the expected number of collected samples  $\mathbb{E}[\tau]$  in terms of the target probability of error  $\epsilon$ .

The class of adaptive policies provides the opportunity to select the sensing actions in favor of the most likely hypothesis, i.e., the actions that discriminate this hypothesis the best from each and every element in the set of alternative hypotheses. However, non-adaptive policies use a fixed randomized rule and hence, have to take into account the worst case scenario.

### 2.7.1 Proof of Theorem 2.3.1

In this section, we show that

$$V_{NN}(\boldsymbol{\rho}) \leq \frac{\log L}{\hat{D}} + o(\log L), \quad (2.22)$$

$$V_{NN}(\boldsymbol{\rho}) \geq \frac{\log L}{\hat{D}} - o(\log L), \quad (2.23)$$

where

$$\hat{D} = \max_{\boldsymbol{\lambda} \in \mathbb{P}(\mathcal{A})} \min_{i \in \Omega} \min_{j \neq i} \max_{\alpha \in [0,1]} \sum_{a \in \mathcal{A}} \lambda_a (1 - \alpha) D_\alpha(q_i^a \| q_j^a). \quad (2.24)$$

Suppose  $\hat{\boldsymbol{\lambda}} \in \mathbb{P}(\mathcal{A})$  achieves the maximum in (2.24). Let  $\pi_{NN}$  be a non-sequential non-adaptive policy that collects  $\hat{n}$  observation samples, and selects sensing actions according to the randomized rule  $\hat{\boldsymbol{\lambda}}$ . The expected total cost under this policy is  $\hat{n} + L\text{Pe}$ . Next we find an upper bound for  $\text{Pe}$  as a function of  $\hat{n}$ . Let  $\mathcal{Z}_i(n) := \cap_{j \neq i} \{\rho_i(n) \geq \rho_j(n)\}$ ,  $\mathcal{Z}_i^c(n) := \cup_{j \neq i} \{\rho_i(n) < \rho_j(n)\}$ , and  $e_{ij}(n) := P(\{\rho_i(n) < \rho_j(n)\} | \theta = i)$ .

$$\begin{aligned} \text{Pe} &= \mathbb{E}[1 - \max_{i \in \Omega} \rho_i(\hat{n})] = \sum_{i=1}^M \rho_i P(\mathcal{Z}_i^c(\hat{n}) | \theta = i) \\ &\leq \sum_{i=1}^M \rho_i \sum_{j \neq i} e_{ij}(\hat{n}) \leq (M-1) \max_{i,j \in \Omega} e_{ij}(\hat{n}). \end{aligned} \quad (2.25)$$

From (2.25) and Lemma 2.7.1 in Section 2.7.5, we obtain

$$\text{Pe} \leq (M-1) 2^{\left(-\hat{n} \min_{i,j \in \Omega} \max_{\alpha \in [0,1]} (1-\alpha) \sum_{a \in \mathcal{A}} \hat{\lambda}_a D_\alpha(q_i^a \| q_j^a) + o(\hat{n})\right)}.$$

We can select  $\hat{n}$  as

$$\hat{n} = (\log L + o(\log L)) / \hat{D}$$

such that  $\text{Pe} = O(\frac{1}{L})$ , and hence,

$$V_{NN} \leq \hat{n} + L\text{Pe} \leq \hat{n} + O(1) = \frac{\log L}{\hat{D}} + o(\log L).$$

This completes the proof of upper bound. Next the proof of lower bound is given.

Consider a non-sequential non-adaptive policy  $\pi_{n,\boldsymbol{\lambda}}$  that collects  $n$  observation samples according to the randomized rule  $\boldsymbol{\lambda} \in \mathbb{P}(\mathcal{A})$ . We have

$$\begin{aligned} \text{Pe} &= \sum_{i=1}^M \rho_i P(\cup_{j \neq i} \{\rho_i(n) < \rho_j(n)\} | \theta = i) \\ &\geq \rho_i e_{ij}(n) + \rho_j e_{ji}(n) \quad \text{for any } i, j \in \Omega, i \neq j. \end{aligned} \quad (2.26)$$

From (2.26) and Lemma 2.7.1 in Section 2.7.5, a lower bound is obtained for the expected total cost under policy  $\pi_{n,\boldsymbol{\lambda}}$ . The lower bound for  $V_{NN}$  is obtained by minimizing over the choices of  $n$  and  $\boldsymbol{\lambda}$ .

## 2.7.2 Proof of Theorem 2.3.2

In this section, we show that

$$V_{SN}(\boldsymbol{\rho}) \leq \min_{\boldsymbol{\lambda} \in \mathbb{P}(\mathcal{A})} \sum_{i=1}^M \rho_i \frac{\log L}{R(i, \boldsymbol{\lambda})} + o(\log L), \quad (2.27)$$

$$V_{SN}(\boldsymbol{\rho}) \geq \min_{\boldsymbol{\lambda} \in \mathbb{P}(\mathcal{A})} \sum_{i=1}^M \rho_i \frac{\log L}{R(i, \boldsymbol{\lambda})} - o(\log L). \quad (2.28)$$

In contrast to the passive case, the observations in the active case (either adaptive or non-adaptive) are not necessarily identical over time. Therefore, the analysis of [3] for sequential passive hypothesis testing (which is based on the law

of large numbers and results for random walks) is not applicable to the problem of sequential non-adaptive hypothesis testing.

Suppose  $\hat{\boldsymbol{\lambda}} \in \mathbb{P}(\mathcal{A})$  achieves the minimum in (2.27). The upper bound (2.27) is achieved by policy  $\pi_{SN}$  that selects sensing actions according to  $\hat{\boldsymbol{\lambda}}$  and stops sampling at

$$\tau := \min\{n : \max_{i \in \Omega} \rho_i(n) \geq 1 - L^{-1}\}. \quad (2.29)$$

Let  $\tau_i$ ,  $i \in \Omega$ , be Markov stopping times defined as follows:

$$\tau_i := \min\left\{n : \min_{j \neq i} \frac{\rho_j(n)}{\rho_j(n)} \geq \frac{1 - L^{-1}}{L^{-1}/(M-1)}\right\}. \quad (2.30)$$

Note that by construction (2.30),

$$\begin{aligned} \rho_i(\tau_i) &\geq \frac{1}{M-1} \sum_{j \neq i} \rho_j(\tau_i) \frac{1 - L^{-1}}{L^{-1}/(M-1)} \\ &= (1 - \rho_i(\tau_i)) \frac{1 - L^{-1}}{L^{-1}}. \end{aligned}$$

This implies that  $\rho_i(\tau_i) \geq 1 - L^{-1}$  and hence,  $\tau \leq \tau_i$  for all  $i \in \Omega$ . From (2.2), the total cost under  $\pi_{SN}$  can be written as

$$\begin{aligned} V(\boldsymbol{\rho}) &= \mathbb{E}[\tau] + L\mathbb{E}[1 - \max_{i \in \Omega} \rho_i(\tau)] \\ &\stackrel{(a)}{\leq} \mathbb{E}[\tau] + 1 \\ &\stackrel{(b)}{=} \sum_{i=1}^M \rho_i \mathbb{E}[\tau | \theta = i] + 1 \\ &\stackrel{(c)}{\leq} \sum_{i=1}^M \rho_i \mathbb{E}[\tau_i | \theta = i] + 1, \end{aligned} \quad (2.31)$$

where (a) follows from construction (2.29); (b) holds since given the prior belief  $\boldsymbol{\rho} = [\rho_1, \rho_2, \dots, \rho_M]$ ,  $P(\{\theta = i\}) = \rho_i$  for all  $i \in \Omega$ ; and (c) follows from the fact that  $\tau \leq \tau_i$  for all  $i \in \Omega$ .

Next we find an upper bound for  $\mathbb{E}[\tau_i | \theta = i]$ ,  $i \in \Omega$ . Before we proceed, we introduce the following notation to facilitate the proof:

$$\bar{T}(i, \boldsymbol{\lambda}) := \frac{\left[ \log \frac{1-L^{-1}}{L^{-1}/(M-1)} - \min_{k \neq i} \log \frac{\rho_i}{\rho_k} \right]^+}{R(i, \boldsymbol{\lambda})}. \quad (2.32)$$

Let  $\iota := (\log L)^{-\frac{1}{3}}$ . We have

$$\begin{aligned}
\mathbb{E}[\tau_i | \theta = i] &= \sum_{n=0}^{\infty} P(\{\tau_i > n\} | \theta = i) \\
&\leq 1 + (1 + \iota) \overline{T}(i, \hat{\lambda}) + \sum_{n: n > (1+\iota) \overline{T}(i, \hat{\lambda})} P(\{\tau_i > n\} | \theta = i) \\
&\stackrel{(a)}{\leq} 1 + (1 + \iota) \overline{T}(i, \hat{\lambda}) + M \frac{e^{-b(\iota) \overline{T}(i, \hat{\lambda})}}{1 - e^{-b(\iota)}} \\
&\stackrel{(b)}{\leq} 1 + (1 + \iota) \overline{T}(i, \hat{\lambda}) + M \frac{1 + \max\{1, \frac{1}{b(\iota)}\}}{e^{b(\iota) \overline{T}(i, \hat{\lambda})}} \\
&\stackrel{(c)}{=} \frac{\log L}{R(i, \hat{\lambda})} + o(\log L), \tag{2.33}
\end{aligned}$$

where (a) follows by Lemma 2.7.2 in Section 2.7.5, and  $b(\iota)$  is given in (2.49); (b) follows from the fact that  $\frac{1}{1-e^{-x}} \leq 1 + \max\{1, 1/x\}$  for  $x > 0$ ; and (c) holds since  $\iota = (\log L)^{-\frac{1}{3}}$ . Now from (2.31) and (2.33), we have the assertion of the theorem.

Next we provide the proof of lower bound (2.28) which follows closely the proof of Theorem 2 in [5].

From upper bound (2.27) we know that the total cost under the optimal policy is  $O(\log L)$ . This implies that the probability of error corresponding to the optimal policy is  $O(\frac{\log L}{L})$ . Hence, without loss of generality in our computation of the lower bound, we can restrict the set of policies to those whose probability of making an error is  $O(\frac{\log L}{L})$ .

Let  $\pi_{\epsilon, \lambda}$  denote a sequential non-adaptive policy that collects samples according to the randomized rule  $\lambda \in \mathbb{P}(\mathcal{A})$  and selects the stopping time  $\tau$  such that  $\text{Pe} \leq \epsilon$ . For all  $i \in \Omega$  and arbitrary  $\delta \in (0, 1)$ , let

$$\underline{T}(i, \lambda) := \frac{\left[ (1 - \delta) \log \frac{1}{\epsilon} - \max_{k \neq i} \log \frac{\rho_i}{\rho_k} \right]^+}{R(i, \lambda) + \delta}. \tag{2.34}$$

Under policy  $\pi_{\epsilon, \lambda}$ ,

$$\begin{aligned}
P(\{\tau < \underline{T}(i, \lambda)\} | \theta = i) &= P\left( \{\tau < \underline{T}(i, \lambda)\} \cap \bigcap_{j \neq i} \left\{ \frac{\rho_i(\tau)}{\rho_j(\tau)} \geq \left(\frac{1}{\epsilon}\right)^{1-\delta} \right\} | \theta = i \right) \\
&\quad + P\left( \{\tau < \underline{T}(i, \lambda)\} \cap \bigcup_{j \neq i} \left\{ \frac{\rho_i(\tau)}{\rho_j(\tau)} < \left(\frac{1}{\epsilon}\right)^{1-\delta} \right\} | \theta = i \right)
\end{aligned}$$

$$\begin{aligned}
&\stackrel{(a)}{\leq} \frac{\xi^2}{\underline{T}(i, \boldsymbol{\lambda})\delta^2} + \sum_{j \neq i} P \left( \left\{ \frac{\rho_i(\tau)}{\rho_j(\tau)} < \left(\frac{1}{\epsilon}\right)^{1-\delta} \right\} \mid \theta = i \right) \\
&\stackrel{(b)}{\leq} \frac{\xi^2}{\underline{T}(i, \boldsymbol{\lambda})\delta^2} + (M-1) \frac{2\epsilon^\delta}{\rho_i},
\end{aligned} \tag{2.35}$$

where (a) follows from Lemma 2.7.4 in Section 2.7.5 and the union bound; and (b) follows from Lemma 2.7.3 in Section 2.7.5.

From (2.35), the expected total cost under policy  $\pi_{\epsilon, \boldsymbol{\lambda}}$  is lower bounded as

$$\begin{aligned}
\mathbb{E}[\tau] + LPe &\geq \sum_{i=1}^M \rho_i \mathbb{E}[\tau \mid \theta = i] \\
&= \sum_{i=1}^M \rho_i \mathbb{E}[\tau \mathbf{1}_{\{\tau \geq \underline{T}(i, \boldsymbol{\lambda})\}} + \tau \mathbf{1}_{\{\tau < \underline{T}(i, \boldsymbol{\lambda})\}} \mid \theta = i] \\
&\geq \sum_{i=1}^M \rho_i \underline{T}(i, \boldsymbol{\lambda}) P(\{\tau \geq \underline{T}(i, \boldsymbol{\lambda})\} \mid \theta = i) \\
&\geq \sum_{i=1}^M \rho_i \underline{T}(i, \boldsymbol{\lambda}) \left( 1 - \frac{\xi^2}{\underline{T}(i, \boldsymbol{\lambda})\delta^2} - \frac{2\epsilon^\delta M}{\rho_i} \right).
\end{aligned}$$

For  $\delta = (\log \frac{1}{\epsilon})^{-\frac{1}{3}}$ , the lower bound simplifies to

$$\begin{aligned}
\mathbb{E}[\tau] + LPe &\geq \sum_{i=1}^M \rho_i \underline{T}(i, \boldsymbol{\lambda}) - \xi^2 (\log \frac{1}{\epsilon})^{\frac{2}{3}} - o(1) \\
&\geq \sum_{i=1}^M \rho_i \frac{\log L}{R(i, \boldsymbol{\lambda})} - o(\log L),
\end{aligned}$$

where  $o(1) \rightarrow 0$  as  $\epsilon \rightarrow 0$  ( $L \rightarrow \infty$ ), and the last inequality follows from the fact that for an optimal policy,  $\epsilon = O(\frac{\log L}{L})$ . The lower bound for  $V_{SN}$  is obtained by minimizing over  $\boldsymbol{\lambda}$ .

### 2.7.3 Proof of Theorem 2.3.3

In this section, we show that

$$V_{SA}(\boldsymbol{\rho}) \leq \sum_{i=1}^M \rho_i \frac{\log L}{R(i, \boldsymbol{\lambda}_i^*)} + o(\log L), \tag{2.36}$$

$$V_{SA}(\boldsymbol{\rho}) \geq \sum_{i=1}^M \rho_i \frac{\log L}{R(i, \boldsymbol{\lambda}_i^*)} - o(\log L). \tag{2.37}$$

The upper bound is achieved via  $\pi_{SA}$ , a heuristic two-phase policy which is described next. Consider a threshold  $\check{\rho}$ ,  $\check{\rho} > \frac{1}{2}$ . The two-phase policy  $\pi_{SA}$  is defined as follows. The policy stops sampling at

$$\tau := \min\{n : \max_{i \in \Omega} \rho_i(n) \geq 1 - L^{-1}\}.$$

For any time  $t < \tau$ , sensing actions are selected according to the randomized rule  $\lambda_i^*$  if  $\rho_i(t) \geq \check{\rho}$ ,  $i \in \Omega$ ; otherwise, the actions are selected according to  $\lambda_0^*$  where

$$\lambda_0^* := \arg \max_{\lambda \in \mathbb{P}(\mathcal{A})} \min_{i \in \Omega} \min_{j \neq i} \sum_{a \in \mathcal{A}} \lambda_a D(q_i^a \| q_j^a). \quad (2.38)$$

In other words, policy  $\pi_{SA}$ , in its first phase, selects actions in a way that all pairs of hypotheses can be distinguished from each other; while in its second phase, i.e., when the belief about one of the hypotheses becomes larger than  $\check{\rho}$ , only the pairs including the most likely hypothesis are considered. The second phase of  $\pi_{SA}$  coincides with Chernoff's scheme [5]; while its first phase in a very natural manner weakens the technical assumption in [5] in which all actions are assumed to discriminate between all hypotheses pairs, i.e.,  $D(q_i^a \| q_j^a) > 0$ ,  $\forall i, j \in \Omega$ ,  $i \neq j$ ,  $\forall a \in \mathcal{A}$  (which is very restrictive and does not hold in many problems of interest).

Similar to Lemma 2.7.2, we can show that under the policy  $\pi_{SA}$  and for  $n > \bar{T}(i, \lambda_i^*)(1 + o(1))$ ,  $P(\{\tau_i > n\} | \theta = i)$  is exponentially decaying in  $n$  (see Section 4.7.7 for details). Following similar lines as those in (2.31) and (2.33), we obtain the bound (2.36).

The proof of the lower bound relies on a generalization of Theorem 2 in [5] and is provided next.

From upper bound (2.36) we know that the total cost under the optimal policy is  $O(\log L)$ . This implies that the error probability corresponding to the optimal policy is  $O(\frac{\log L}{L})$ . Hence, without loss of generality in our computation of the lower bound, we can restrict the set of policies to those whose probability of making an error is  $O(\frac{\log L}{L})$ .

Let  $\pi_\epsilon$  denote a sequential policy that selects the stopping time  $\tau$  such that  $\text{Pe} \leq \epsilon$ . Under policy  $\pi_\epsilon$ ,

$$P(\{\tau < \underline{T}(i, \lambda_i^*)\} | \theta = i) = P\left(\{\tau < \underline{T}(i, \lambda_i^*)\} \cap \bigcap_{j \neq i} \left\{ \frac{\rho_i(\tau)}{\rho_j(\tau)} \geq \left(\frac{1}{\epsilon}\right)^{1-\delta} \right\} | \theta = i\right)$$

$$\begin{aligned}
& + P\left(\{\tau < \underline{T}(i, \boldsymbol{\lambda}_i^*)\} \cap \bigcup_{j \neq i} \left\{ \frac{\rho_i(\tau)}{\rho_j(\tau)} < \left(\frac{1}{\epsilon}\right)^{1-\delta} \right\} \mid \theta = i\right) \\
& \stackrel{(a)}{\leq} \frac{M\xi^2}{\underline{T}(i, \boldsymbol{\lambda}_i^*)\delta^2} + \sum_{j \neq i} P\left(\left\{ \frac{\rho_i(\tau)}{\rho_j(\tau)} < \left(\frac{1}{\epsilon}\right)^{1-\delta} \right\} \mid \theta = i\right) \\
& \stackrel{(b)}{\leq} \frac{M\xi^2}{\underline{T}(i, \boldsymbol{\lambda}_i^*)\delta^2} + (M-1)\frac{2\epsilon^\delta}{\rho_i}, \tag{2.39}
\end{aligned}$$

where (a) follows from Lemma 2.7.5 in Section 2.7.5 and the union bound; and (b) follows from Lemma 2.7.3 in Section 2.7.5.

From (2.39), the expected total cost under policy  $\pi_\epsilon$  is lower bounded as

$$\begin{aligned}
\mathbb{E}[\tau] + LPe & \geq \sum_{i=1}^M \rho_i \mathbb{E}[\tau \mid \theta = i] \\
& = \sum_{i=1}^M \rho_i \mathbb{E}[\tau \mathbf{1}_{\{\tau \geq \underline{T}(i, \boldsymbol{\lambda}_i^*)\}} + \tau \mathbf{1}_{\{\tau < \underline{T}(i, \boldsymbol{\lambda}_i^*)\}} \mid \theta = i] \\
& \geq \sum_{i=1}^M \rho_i \underline{T}(i, \boldsymbol{\lambda}_i^*) P(\{\tau \geq \underline{T}(i, \boldsymbol{\lambda}_i^*)\} \mid \theta = i) \\
& \geq \sum_{i=1}^M \rho_i \underline{T}(i, \boldsymbol{\lambda}_i^*) \left(1 - \frac{M\xi^2}{\underline{T}(i, \boldsymbol{\lambda}_i^*)\delta^2} - \frac{2\epsilon^\delta M}{\rho_i}\right).
\end{aligned}$$

For  $\delta = (\log \frac{1}{\epsilon})^{-\frac{1}{3}}$ , the lower bound simplifies to

$$\begin{aligned}
\mathbb{E}[\tau] + LPe & \geq \sum_{i=1}^M \rho_i \underline{T}(i, \boldsymbol{\lambda}_i^*) - M\xi^2 (\log \frac{1}{\epsilon})^{\frac{2}{3}} - o(1) \\
& \geq \sum_{i=1}^M \rho_i \frac{\log L}{R(i, \boldsymbol{\lambda}_i^*)} - o(\log L),
\end{aligned}$$

where  $o(1) \rightarrow 0$  as  $\epsilon \rightarrow 0$  ( $L \rightarrow \infty$ ), and the last inequality follows from the fact that for an optimal policy,  $\epsilon = O(\frac{\log L}{L})$ .

## 2.7.4 Proof of Proposition 2.3.1

In this section, we show that

$$V_{NA}(\boldsymbol{\rho}) \geq \frac{\log L}{\min_{i \in \Omega} R(i, \boldsymbol{\lambda}_i^*)} - o(\log L).$$

*Proof.* We first consider the set of all non-sequential adaptive policies that collect  $n$  observation samples, and find a lower bound on the expected total cost under these policies. The lower bound for  $V_{NA}$  is obtained by minimizing over  $n$ .

Consider an arbitrary  $\delta > 0$  and define

$$n' := \max \left\{ n, \frac{2\xi^2}{\delta^2} \log M \right\}, \quad (2.40)$$

$$\epsilon_i := \frac{1}{2^{(n'(R(i, \lambda_i^*) + \delta) + \max_{k, l \in \Omega} \log \frac{\rho_k}{\rho_l})} + 1}, \quad \forall i \in \Omega. \quad (2.41)$$

The error probability can be bounded as

$$\begin{aligned} \text{Pe} &= \sum_{i=1}^M \rho_i \mathbb{E}[1 - \max_{k \in \Omega} \rho_k(n) | \theta = i] \\ &\geq \sum_{i=1}^M \rho_i \mathbb{E}[(1 - \max_{k \in \Omega} \rho_k(n)) \mathbf{1}_{\{\max_{k \in \Omega} \rho_k(n) \leq 1 - \epsilon_i\}} | \theta = i] \\ &\geq \sum_{i=1}^M \rho_i \epsilon_i P(\{\max_{k \in \Omega} \rho_k(n) \leq 1 - \epsilon_i\} | \theta = i). \end{aligned} \quad (2.42)$$

Next we find a lower bound for the right-hand side of (2.42). Note that

$$\begin{aligned} &P(\{\max_{k \in \Omega} \rho_k(n) > 1 - \epsilon_i\} | \theta = i) \\ &\leq \sum_{k=1}^M P(\{\rho_k(n) > 1 - \epsilon_i\} | \theta = i) \\ &= \sum_{k=1}^M P\left(\left\{\log \frac{\rho_k(n)}{1 - \rho_k(n)} > \log \frac{1 - \epsilon_i}{\epsilon_i}\right\} | \theta = i\right) \\ &\leq \sum_{k=1}^M P\left(\bigcap_{j \neq k} \left\{\log \frac{\rho_k(n)}{\rho_j(n)} > \log \frac{1 - \epsilon_i}{\epsilon_i}\right\} | \theta = i\right). \end{aligned} \quad (2.43)$$

To find an upper bound for the right-hand side of (2.43), we consider the following two cases:

**Case I.**  $k \neq i$ . Given  $\{\theta = i\}$  and for any sequence of actions  $\{A(t)\}_{t=0}^{n-1}$ ,

$$\begin{aligned} \mathbb{E}[\log \frac{\rho_k(n)}{\rho_i(n)}] &= \log \frac{\rho_k}{\rho_i} + \sum_{t=0}^{n-1} \mathbb{E}[\log \frac{\rho_k(t+1)}{\rho_i(t+1)} - \log \frac{\rho_k(t)}{\rho_i(t)}] \\ &= \log \frac{\rho_k}{\rho_i} + \sum_{t=0}^{n-1} \mathbb{E}[\log \frac{q_k^{A(t)}(Z)}{q_i^{A(t)}(Z)}] \end{aligned}$$

$$\begin{aligned}
&= \log \frac{\rho_k}{\rho_i} - \sum_{t=0}^{n-1} \sum_{a \in \mathcal{A}} P(A(t) = a) D(q_i^a \| q_k^a) \\
&\leq \log \frac{\rho_k}{\rho_i}.
\end{aligned} \tag{2.44}$$

From (2.44) and for any  $k \neq i$ ,

$$\begin{aligned}
&P\left(\cap_{j \neq k} \left\{ \log \frac{\rho_k(n)}{\rho_j(n)} > \log \frac{1 - \epsilon_i}{\epsilon_i} \right\} | \theta = i\right) \\
&\leq P\left(\left\{ \log \frac{\rho_k(n)}{\rho_i(n)} - \mathbb{E}\left[\log \frac{\rho_k(n)}{\rho_i(n)}\right] > \log \frac{1 - \epsilon_i}{\epsilon_i} - \log \frac{\rho_k}{\rho_i} \right\} | \theta = i\right) \\
&\stackrel{(a)}{\leq} P\left(\left\{ \log \frac{\rho_k(n)}{\rho_i(n)} - \mathbb{E}\left[\log \frac{\rho_k(n)}{\rho_i(n)}\right] > n' \delta \right\} | \theta = i\right) \\
&\stackrel{(b)}{\leq} \exp(-n' \delta^2 / 2\xi^2),
\end{aligned} \tag{2.45}$$

where (a) follows from (2.41); and (b) follows from Fact 2.7.2 in Section 2.7.5 and since  $n \leq n'$ .

**Case II.**  $k = i$ . Consider a sequence of actions  $\{A(t)\}_{t=0}^{n-1}$ . Given  $\{\theta = i\}$  and for  $\hat{j} = \arg \min_{j \neq i} \sum_{t=0}^{n-1} \mathbb{E}[\log \frac{q_i^{A(t)}(Z)}{q_j^{A(t)}(Z)}]$ ,

$$\begin{aligned}
\mathbb{E}[\log \frac{\rho_i(n)}{\rho_{\hat{j}}(n)}] &= \log \frac{\rho_i}{\rho_{\hat{j}}} + \sum_{t=0}^{n-1} \mathbb{E}[\log \frac{q_i^{A(t)}(Z)}{q_{\hat{j}}^{A(t)}(Z)}] \\
&\leq \log \frac{\rho_i}{\rho_{\hat{j}}} + n \min_{j \neq i} \sum_{a \in \mathcal{A}} \lambda_{ia}^* D(q_i^a \| q_j^a) \\
&\leq \max_{j \neq i} \log \frac{\rho_i}{\rho_j} + nR(i, \boldsymbol{\lambda}_i^*).
\end{aligned} \tag{2.46}$$

Using (2.46) and for  $k = i$ ,

$$\begin{aligned}
&P\left(\cap_{j \neq k} \left\{ \log \frac{\rho_k(n)}{\rho_j(n)} > \log \frac{1 - \epsilon_i}{\epsilon_i} \right\} | \theta = i\right) \\
&\leq P\left(\left\{ \log \frac{\rho_i(n)}{\rho_{\hat{j}}(n)} - \mathbb{E}[\log \frac{\rho_i(n)}{\rho_{\hat{j}}(n)}] > \log \frac{1 - \epsilon_i}{\epsilon_i} - \max_{j \neq i} \log \frac{\rho_i}{\rho_j} - nR(i, \boldsymbol{\lambda}_i^*) \right\} | \theta = i\right) \\
&\stackrel{(a)}{\leq} P\left(\left\{ \log \frac{\rho_i(n)}{\rho_{\hat{j}}(n)} - \mathbb{E}[\log \frac{\rho_i(n)}{\rho_{\hat{j}}(n)}] > n' \delta \right\} | \theta = i\right) \\
&\stackrel{(b)}{\leq} \exp(-n' \delta^2 / 2\xi^2),
\end{aligned} \tag{2.47}$$

where (a) follows from (2.40) and (2.41); and (b) follows from Fact 2.7.2 in Section 2.7.5.

Combining (2.43), (2.45), and (2.47),

$$P(\{\max_{k \in \Omega} \rho_k(n) > 1 - \epsilon_i\} | \theta = i) \leq M \exp(-n' \delta^2 / 2\xi^2) < 1,$$

which together with (2.42) establishes the following lower bound on the expected total cost,

$$n + LPe \geq n + L \sum_{i=1}^M \rho_i \epsilon_i (1 - M \exp(-n' \delta^2 / 2\xi^2)). \quad (2.48)$$

Setting  $\delta = (\log L)^{-\frac{1}{3}}$  and minimizing the bound over  $n$ , we have the assertion of the proposition.  $\square$

### 2.7.5 Technical Background

In this section, we provide some preliminary facts and lemmas which are technical and helpful in proving the main results.

**Fact 2.7.1** (Kolmogorov's Maximal Inequality [47]). *Suppose  $X_t$  for  $t = 1, 2, \dots$ , be independent random variables with  $\mathbb{E}[X_t] = 0$  and  $\text{Var}(X_t) < \infty$ . Let  $S_n = \sum_{t=1}^n X_t$ . Then*

$$P\left(\max_{0 \leq n \leq N} |S_n| > x\right) \leq \frac{\text{Var}(S_N)}{x^2} = \frac{\sum_{t=1}^N \text{Var}(X_t)}{x^2}.$$

**Fact 2.7.2** (McDiarmid's Inequality [48]). *Let  $\mathbf{X} = (X_1, \dots, X_n)$  be a family of independent random variables with  $X_k$  taking values in a set  $\mathcal{X}_k$  for each  $k$ . Suppose a real-valued function  $f$  defined on  $\prod_{k=1}^n \mathcal{X}_k$  satisfies  $|f(\mathbf{x}) - f(\mathbf{x}')| \leq c_k$ , whenever the vectors  $\mathbf{x}$  and  $\mathbf{x}'$  only differ in the  $k$ -th coordinate. Then for any  $\nu \geq 0$ ,*

$$\begin{aligned} P(f(\mathbf{X}) - \mathbb{E}[f(\mathbf{X})] \geq \nu) &\leq e^{-2\nu^2 / \sum_{k=1}^n c_k^2}, \\ P(f(\mathbf{X}) - \mathbb{E}[f(\mathbf{X})] \leq -\nu) &\leq e^{-2\nu^2 / \sum_{k=1}^n c_k^2}. \end{aligned}$$

**Lemma 2.7.1.** *Consider a policy that collects observation samples according to a randomized rule  $\lambda \in \mathbb{P}(\mathcal{A})$ . Under this policy and for all  $i, j \in \Omega$ , and  $\alpha \in [0, 1]$ ,*

$$\begin{aligned} \max\{e_{ij}(n), e_{ji}(n)\} &\leq 2^{\left(-n(1-\alpha) \sum_{a \in \mathcal{A}} \lambda_a D_\alpha(q_i^a \| q_j^a) + o(n)\right)}, \\ \max\{e_{ij}(n), e_{ji}(n)\} &\geq 2^{\left(-n(1-\alpha) \sum_{a \in \mathcal{A}} \lambda_a D_\alpha(q_i^a \| q_j^a) - o(n)\right)}. \end{aligned}$$

The proof of Lemma 2.7.1 follows closely the proof of Theorems 9 and 10 in [26].

**Lemma 2.7.2.** *Consider a sequential non-adaptive policy that selects sensing actions according to the randomized rule  $\hat{\lambda}$ . Under this policy and for any  $\iota > 0$  and  $n > (1 + \iota)\bar{T}(i, \hat{\lambda})$ , we have  $P(\{\tau_i > n\}|\theta = i) \leq (M - 1)e^{-b(\iota)n}$  where*

$$b(\iota) = \frac{2\iota^2}{(1 + \iota)^2} \left( \frac{R(i, \hat{\lambda})}{2\xi} \right)^2. \quad (2.49)$$

*Proof of Lemma 2.7.2.* Let  $B_{ij}(n)$  be an event in the probability space defined as follows:

$$B_{ij}(n) := \left\{ \log \frac{\rho_i(n)}{\rho_j(n)} < \log \frac{1 - L^{-1}}{L^{-1}/(M - 1)} \right\}.$$

By construction (2.30),

$$P(\{\tau_i > n\}|\theta = i) \leq P(\cup_{j \neq i} B_{ij}(n)|\theta = i) \leq \sum_{j \neq i} P(B_{ij}(n)|\theta = i). \quad (2.50)$$

Furthermore, we have

$$\begin{aligned} & P(B_{ij}(n)|\theta = i) \\ &= P\left(\left\{ \log \frac{\rho_i(n)}{\rho_j(n)} - \mathbb{E}\left[\log \frac{\rho_i(n)}{\rho_j(n)}\right] < \log \frac{1 - L^{-1}}{L^{-1}/(M - 1)} - \mathbb{E}\left[\log \frac{\rho_i(n)}{\rho_j(n)}\right] \right\}|\theta = i\right) \\ &= P\left(\left\{ \log \frac{\rho_i(n)}{\rho_j(n)} - \mathbb{E}\left[\log \frac{\rho_i(n)}{\rho_j(n)}\right] < \log \frac{1 - L^{-1}}{L^{-1}/(M - 1)} - \mathbb{E}\left[\log \frac{\rho_i}{\rho_j} + \sum_{t=0}^{n-1} \log \frac{q_i^{A(t)}}{q_j^{A(t)}}\right] \right\}|\theta = i\right) \\ &\leq P\left(\left\{ \log \frac{\rho_i(n)}{\rho_j(n)} - \mathbb{E}\left[\log \frac{\rho_i(n)}{\rho_j(n)}\right] < \log \frac{1 - L^{-1}}{L^{-1}/(M - 1)} - \min_{k \neq i} \log \frac{\rho_i}{\rho_k} - nR(i, \hat{\lambda}) \right\}|\theta = i\right) \\ &\leq P\left(\left\{ \log \frac{\rho_i(n)}{\rho_j(n)} - \mathbb{E}\left[\log \frac{\rho_i(n)}{\rho_j(n)}\right] < (\bar{T}(i, \hat{\lambda}) - n)R(i, \hat{\lambda}) \right\}|\theta = i\right). \end{aligned} \quad (2.51)$$

For any  $a, \hat{a} \in \mathcal{A}$  and  $i, j \in \Omega$ , we have  $\left| \log \frac{q_i^a}{q_j^a} - \log \frac{q_i^{\hat{a}}}{q_j^{\hat{a}}} \right| \leq 2\xi$ . For  $k = 1, 2, \dots, n$ , let  $X_k = \log \frac{q_i^{A(k-1)}}{q_j^{A(k-1)}}$  and  $\mathbf{X} = [X_1, X_2, \dots, X_n]$ . Define the function  $f(\mathbf{X}) = \log \frac{\rho_i}{\rho_j} + \sum_{k=1}^n X_k = \log \frac{\rho_i(n)}{\rho_j(n)}$ . From (2.50), (2.51), and Fact 2.7.2, and for  $n > (1 + \iota)\bar{T}(i, \hat{\lambda})$ , we have

$$P(\{\tau_i > n\}|\theta = i) \leq (M - 1) \exp \left( -2n \left( \frac{R(i, \hat{\lambda})}{2\xi} \right)^2 \left( 1 - \frac{1}{n} \bar{T}(i, \hat{\lambda}) \right)^2 \right)$$

$$\leq (M-1) \exp \left( -n \frac{2\iota^2}{(1+\iota)^2} \left( \frac{R(i, \hat{\lambda})}{2\xi} \right)^2 \right).$$

□

**Lemma 2.7.3.** *Consider a sequential policy  $\pi$  that selects the stopping time  $\tau$  such that  $\text{Pe} \leq \epsilon$ . For any  $i, j \in \Omega$  and  $\delta \in [0, 1]$ , we have*

$$P \left( \left\{ \frac{\rho_i(\tau)}{\rho_j(\tau)} < \left( \frac{1}{\epsilon} \right)^{1-\delta} \right\} \mid \theta = i \right) \leq \frac{2\epsilon^\delta}{\rho_i}.$$

*Proof.* The proof follows closely the proof of Lemma 4 in [5]. Let  $\hat{\theta} = d(A^\tau, Z^\tau)$  denote the final declaration under policy  $\pi$ . We have

$$\begin{aligned} & P \left( \left\{ \frac{\rho_i(\tau)}{\rho_j(\tau)} < \left( \frac{1}{\epsilon} \right)^{1-\delta} \right\} \mid \theta = i \right) \\ &= P \left( \left\{ \frac{\rho_i(\tau)}{\rho_j(\tau)} < \left( \frac{1}{\epsilon} \right)^{1-\delta} \right\} \cap \left\{ \hat{\theta} = i \right\} \mid \theta = i \right) \\ &\quad + P \left( \left\{ \frac{\rho_i(\tau)}{\rho_j(\tau)} < \left( \frac{1}{\epsilon} \right)^{1-\delta} \right\} \cap \left\{ \hat{\theta} \neq i \right\} \mid \theta = i \right) \\ &\stackrel{(a)}{\leq} \left( \frac{1}{\epsilon} \right)^{1-\delta} \frac{\rho_j}{\rho_i} P \left( \left\{ \hat{\theta} = i \right\} \mid \theta = j \right) + P \left( \left\{ \hat{\theta} \neq i \right\} \mid \theta = i \right) \\ &\leq \left( \frac{1}{\epsilon} \right)^{1-\delta} \frac{\rho_j}{\rho_i} P \left( \left\{ \hat{\theta} \neq j \right\} \mid \theta = j \right) + P \left( \left\{ \hat{\theta} \neq i \right\} \mid \theta = i \right) \\ &\stackrel{(b)}{\leq} \left( \frac{1}{\epsilon} \right)^{1-\delta} \frac{\rho_j}{\rho_i} \frac{\epsilon}{\rho_j} + \frac{\epsilon}{\rho_i} \leq \frac{2\epsilon^\delta}{\rho_i}, \end{aligned}$$

where (a) holds since under the event  $\left\{ \frac{\rho_i(\tau)}{\rho_j(\tau)} < \left( \frac{1}{\epsilon} \right)^{1-\delta} \right\}$ , we have

$$\rho_i \prod_{t=0}^{\tau-1} q_i^{A(t)}(Z(t)) \leq \left( \frac{1}{\epsilon} \right)^{1-\delta} \rho_j \prod_{t=0}^{\tau-1} q_j^{A(t)}(Z(t)),$$

hence, the ratio between the probability of the event conditioned on  $\theta = i$  and the probability conditioned on  $\theta = j$  is bounded by  $\left( \frac{1}{\epsilon} \right)^{1-\delta} \frac{\rho_j}{\rho_i}$ ; and (b) follows from the fact that under policy  $\pi$  and for all  $i \in \Omega$ ,

$$P \left( \left\{ \hat{\theta} \neq i \right\} \mid \theta = i \right) \leq \frac{1}{\rho_i} \sum_{k=1}^M \rho_k P \left( \left\{ \hat{\theta} \neq k \right\} \mid \theta = k \right) = \frac{1}{\rho_i} \text{Pe} \leq \frac{\epsilon}{\rho_i}.$$

□

**Lemma 2.7.4.** Consider a sequential non-adaptive policy  $\pi_{\epsilon, \lambda}$  that selects actions according to  $\lambda \in \mathbb{P}(\mathcal{A})$  and selects the stopping time  $\tau$  such that  $\mathbb{P}e \leq \epsilon$ . We have

$$P\left(\left\{\tau < \underline{T}(i, \lambda)\right\} \cap \bigcap_{j \neq i} \left\{\frac{\rho_i(\tau)}{\rho_j(\tau)} \geq \left(\frac{1}{\epsilon}\right)^{1-\delta}\right\} \mid \theta = i\right) \leq \frac{\xi^2}{\underline{T}(i, \lambda)\delta^2}.$$

*Proof.* Let  $\hat{j} = \arg \min_{j \neq i} \sum_{a \in \mathcal{A}} \lambda_a D(q_i^a \| q_j^a)$ . We have

$$\begin{aligned} & P\left(\left\{\tau < \underline{T}(i, \lambda)\right\} \cap \bigcap_{j \neq i} \left\{\frac{\rho_i(\tau)}{\rho_j(\tau)} \geq \left(\frac{1}{\epsilon}\right)^{1-\delta}\right\} \mid \theta = i\right) \\ & \leq P\left(\left\{\min_n : \log \frac{\rho_i(n)}{\rho_j(n)} > (1-\delta) \log \frac{1}{\epsilon}, \forall j \neq i\right\} < \underline{T}(i, \lambda) \mid \theta = i\right) \\ & = P\left(\left\{\exists n, 0 \leq n < \underline{T}(i, \lambda) \text{ s.t. } \log \frac{\rho_i(n)}{\rho_j(n)} > (1-\delta) \log \frac{1}{\epsilon}, \forall j \neq i\right\} \mid \theta = i\right) \\ & \stackrel{(a)}{\leq} P\left(\left\{\exists n, 0 \leq n < \underline{T}(i, \lambda) \text{ s.t. } \log \frac{\rho_i(n)}{\rho_{\hat{j}}(n)} - \mathbb{E}[\log \frac{\rho_i(n)}{\rho_{\hat{j}}(n)}] \right. \right. \\ & \quad \left. \left. > (1-\delta) \log \frac{1}{\epsilon} - \max_{k \neq i} \log \frac{\rho_i}{\rho_k} - nR(i, \lambda)\right\} \mid \theta = i\right) \\ & \stackrel{(b)}{\leq} P\left(\left\{\exists n, 0 \leq n < \underline{T}(i, \lambda) \text{ s.t. } \log \frac{\rho_i(n)}{\rho_{\hat{j}}(n)} - \mathbb{E}[\log \frac{\rho_i(n)}{\rho_{\hat{j}}(n)}] > \underline{T}(i, \lambda)\delta\right\} \mid \theta = i\right) \\ & \leq P\left(\max_{0 \leq n < \underline{T}(i, \lambda)} \left\{\log \frac{\rho_i(n)}{\rho_{\hat{j}}(n)} - \mathbb{E}[\log \frac{\rho_i(n)}{\rho_{\hat{j}}(n)}]\right\} > \underline{T}(i, \lambda)\delta \mid \theta = i\right) \\ & \stackrel{(c)}{\leq} \frac{\underline{T}(i, \lambda)\xi^2}{(\underline{T}(i, \lambda)\delta)^2}, \end{aligned}$$

where (a) follows from the fact that given  $\{\theta = i\}$ ,

$$\begin{aligned} \mathbb{E}[\log \frac{\rho_i(n)}{\rho_{\hat{j}}(n)}] &= \log \frac{\rho_i}{\rho_{\hat{j}}} + \sum_{t=0}^{n-1} \mathbb{E}[\log \frac{\rho_i(t+1)}{\rho_{\hat{j}}(t+1)} - \log \frac{\rho_i(t)}{\rho_{\hat{j}}(t)}] \\ &= \log \frac{\rho_i}{\rho_{\hat{j}}} + \sum_{t=0}^{n-1} \sum_{a \in \mathcal{A}} \mathbb{E}[\log \frac{q_i^a(Z)}{q_{\hat{j}}^a(Z)}] P(A(t) = a) \\ &\leq \max_{k \neq i} \log \frac{\rho_i}{\rho_k} + n \min_{j \neq i} \sum_{a \in \mathcal{A}} \lambda_a D(q_i^a \| q_j^a); \end{aligned}$$

inequality (b) follows from the definition of  $\underline{T}(i, \lambda)$  in (2.34) and the fact that  $0 \leq n < \underline{T}(i, \lambda)$ ; and (c) follows from Fact 2.7.1. □

**Lemma 2.7.5.** *Consider a sequential policy  $\pi_\epsilon$  that selects the stopping time  $\tau$  such that  $\text{Pe} \leq \epsilon$ . We have*

$$P\left(\{\tau < \underline{T}(i, \boldsymbol{\lambda}_i^*)\} \cap \bigcap_{j \neq i} \left\{ \frac{\rho_i(\tau)}{\rho_j(\tau)} \geq \left(\frac{1}{\epsilon}\right)^{1-\delta} \right\} | \theta = i\right) \leq \frac{(M-1)\xi^2}{\underline{T}(i, \boldsymbol{\lambda}_i^*)\delta^2}.$$

*Proof.* The proof follows closely the proof of Lemma 5 in [5]. We have

$$\begin{aligned} & P\left(\{\tau < \underline{T}(i, \boldsymbol{\lambda}_i^*)\} \cap \bigcap_{j \neq i} \left\{ \frac{\rho_i(\tau)}{\rho_j(\tau)} \geq \left(\frac{1}{\epsilon}\right)^{1-\delta} \right\} | \theta = i\right) \\ & \leq P\left(\left\{ \min_n : \log \frac{\rho_i(n)}{\rho_j(n)} > (1-\delta) \log \frac{1}{\epsilon}, \forall j \neq i \right\} < \underline{T}(i, \boldsymbol{\lambda}_i^*) | \theta = i\right) \\ & = P\left(\left\{ \exists n, 0 \leq n < \underline{T}(i, \boldsymbol{\lambda}_i^*) \text{ s.t. } \log \frac{\rho_i(n)}{\rho_j(n)} > (1-\delta) \log \frac{1}{\epsilon}, \forall j \neq i \right\} | \theta = i\right) \\ & \stackrel{(a)}{\leq} P\left(\bigcup_{j \neq i} \left\{ \exists n, 0 \leq n < \underline{T}(i, \boldsymbol{\lambda}_i^*) \text{ s.t. } \log \frac{\rho_i(n)}{\rho_j(n)} - \mathbb{E}[\log \frac{\rho_i(n)}{\rho_j(n)}] \right. \right. \\ & \quad \left. \left. > (1-\delta) \log \frac{1}{\epsilon} - \max_{k \neq i} \log \frac{\rho_i}{\rho_k} - nR(i, \lambda_i^*) \right\} | \theta = i\right) \\ & \stackrel{(b)}{\leq} P\left(\bigcup_{j \neq i} \left\{ \exists n, 0 \leq n < \underline{T}(i, \boldsymbol{\lambda}_i^*) \text{ s.t. } \log \frac{\rho_i(n)}{\rho_j(n)} - \mathbb{E}[\log \frac{\rho_i(n)}{\rho_j(n)}] > \underline{T}(i, \boldsymbol{\lambda}_i^*)\delta \right\} | \theta = i\right) \\ & \leq \sum_{j \neq i} P\left(\max_{0 \leq n < \underline{T}(i, \boldsymbol{\lambda}_i^*)} \left\{ \log \frac{\rho_i(n)}{\rho_j(n)} - \mathbb{E}[\log \frac{\rho_i(n)}{\rho_j(n)}] \right\} > \underline{T}(i, \boldsymbol{\lambda}_i^*)\delta | \theta = i\right) \\ & \stackrel{(c)}{\leq} (M-1) \frac{\underline{T}(i, \boldsymbol{\lambda}_i^*)\xi^2}{(\underline{T}(i, \boldsymbol{\lambda}_i^*)\delta)^2}, \end{aligned}$$

where (a) holds since

$$\begin{aligned} \mathbb{E}[\log \frac{\rho_i(n)}{\rho_j(n)}] &= \log \frac{\rho_i}{\rho_j} + \sum_{t=0}^{n-1} \mathbb{E}[\log \frac{q_i^{A(t)}(Z)}{q_j^{A(t)}(Z)}] \\ &= \log \frac{\rho_i}{\rho_j} + n \sum_{a \in \mathcal{A}} \frac{\sum_{t=0}^{n-1} P(A(t) = a)}{n} D(q_i^a \| q_j^a), \end{aligned}$$

and hence there exists  $j \in \Omega$  for which

$$\mathbb{E}[\log \frac{\rho_i(n)}{\rho_j(n)}] \leq \max_{k \neq i} \log \frac{\rho_i}{\rho_k} + nR(i, \lambda_i^*);$$

inequality (b) follows from the definition of  $\underline{T}(i, \boldsymbol{\lambda}_i^*)$  and the fact that  $0 \leq n < \underline{T}(i, \boldsymbol{\lambda}_i^*)$ ; and (c) follows from Fact 2.7.1.  $\square$

Chapter 2, in full, is a reprint of the material as it is scheduled to appear in M. Naghshvar and T. Javidi, “Sequentiality and adaptivity gains in active hypothesis testing,” *IEEE Journal of Selected Topics in Signal Processing*, October 2013, available on arXiv:1211.2291. The dissertation author was the primary investigator and author of this paper.

## Chapter 3

# Dynamic Programming and Optimal Notion of Information

In Chapter 2, the expected total cost of various classes of policies were characterized in terms of the penalty of wrong declaration  $L$ . It was shown that sequential (variable sample size) policies are superior to non-sequential (fixed sample size) ones, and the gain of selecting actions in a sequential manner is logarithmic in  $L$ . The class of sequential policies will be the focus of the rest of this dissertation.

This chapter studies the problem of active sequential hypothesis testing from a different point of view. Suppose there exists an optimal notion of information that captures the influence of every action over the entire decision making horizon. In this case, the active hypothesis testing reduces to a sequence of one-shot problems in each of which an optimal action is the one that provides the highest amount of information. In this chapter, it is shown that this optimal notion of information exists for the active sequential hypothesis testing and can be derived from the solution to the corresponding dynamic programming (DP) equation. However, finding the closed-form solution to the DP equation is not feasible in general. In lieu of numerically solving the DP equation, alternative notions of information are considered and heuristic policies are proposed based on greedy maximization of these notions. Via numerical and theoretical analysis, the performance of the proposed policies and the relevance of the corresponding notions of information in the context of active hypothesis testing are investigated.

### 3.1 Introduction

Following Chernoff's approach [5], we analyzed the classes of sequential non-adaptive and sequential adaptive policies in Chapter 2. We proposed two policies  $\pi_{SN}$  and  $\pi_{SA}$  which were shown to achieve asymptotic optimality in  $L$ ; where the notion of asymptotic optimality in  $L$ , due to Chernoff [5], denotes the relative tightness of the performance upper bound associated with the proposed policy and the lower bound associated with the optimal policy.

In this chapter, we tackle the problem of active sequential hypothesis testing from a different point of view which coincides with that promoted by DeGroot [20]. Suppose the decision maker's amount of uncertainty about the true hypothesis at any belief state can be characterized using an *uncertainty* function. In [20], the *information* of a sensing action is defined as the expected difference between the uncertainty of the prior belief and the uncertainty of the posterior belief. To achieve the best performance, it is essential that the decision maker, in each step, selects an action that provides the highest amount of information. This raises the question as what information and uncertainty function are appropriate in the context of active hypothesis testing. It is intuitive that an optimal notion of information should capture the influence that every action has over the entire decision making horizon. In Section 3.2, the optimal notion of information is derived based on the optimal value function of the corresponding dynamic programming. Although this result provides a general and structural characterization of the (Markov and deterministic) optimal policy, in the absence of a closed-form for the optimal value function, it has no practical use. To address this issue, in Section 3.3, we introduce alternative notions of information namely Jensen-Shannon (JS) divergence and Extrinsic Jensen-Shannon (EJS) divergence, and propose two heuristic policies  $\pi_{JS}$  and  $\pi_{EJS}$  based on greedy maximization of these divergences. The performance of these policies as well as that of the optimal policy are compared in a numerical example in Section 3.4. Theoretical analysis of these policies are given in Chapter 4 where we discuss in detail the advantage and disadvantage of the new approach applied in the current chapter with that of Chernoff and the results obtained in Chapter 2.

### 3.2 Characterization of an Optimal Policy

Recall from Section 2.2.3 that active  $M$ -ary hypothesis testing is equivalent to an MDP with state space  $\mathbb{P}(\Omega) = \{\boldsymbol{\rho} \in [0, 1]^M : \sum_{i=1}^M \rho_i = 1\}$ . The belief state at time  $t$  is given by the vector  $\boldsymbol{\rho}(t) := [\rho_1(t), \rho_2(t), \dots, \rho_M(t)]$  where  $\rho_i(t) := P(\{\theta = i\} | A^{t-1}, Z^{t-1})$ . The objective is to find a stopping time  $\tau$  and a sequence of sensing actions  $A(0), A(1), \dots, A(\tau - 1)$  that collectively minimize the expected total cost

$$\mathbb{E}[\tau] + L\text{Pe}, \quad (3.1)$$

where  $\text{Pe} = \mathbb{E}[1 - \max_{i \in \Omega} \rho_i(\tau)]$  is the probability of wrong declaration, and the expectations are taken with respect to the initial prior distribution on  $\theta$  as well as the distributions of action sequence, observation sequence, and the stopping time.

**Definition 3.2.1.** For all  $\boldsymbol{\rho} \in \mathbb{P}(\Omega)$ , let functional  $V^*(\boldsymbol{\rho})$ , hereafter referred to as the *optimal value function*, denote the optimal expected total cost (3.1) of the active hypothesis testing problem given the Bayesian prior  $\boldsymbol{\rho}$ . In other words,  $V^*(\boldsymbol{\rho}) := \min \{\mathbb{E}[\tau] + L\text{Pe}\}$  given the initial belief  $\boldsymbol{\rho}$ , where the minimization is taken over the stopping time  $\tau$  and the sequence of actions.

A general approach to solving the active hypothesis testing problem is to provide a functional characterization of  $V^*$ : given  $V^*$  in its functional form, the optimal expected total cost for the active hypothesis testing problem can be obtained by a simple evaluation of  $V^*$  at the initial belief  $\boldsymbol{\rho}(0)$ . Next we state a dynamic programming equation which characterizes  $V^*$ .

To obtain the dynamic programming equation, consider a single step of the problem. In one sensing step, the evolution of the belief vector follows Bayes' rule and is given by  $\boldsymbol{\Phi}^a$ , a measurable function from  $\mathbb{P}(\Omega) \times \mathcal{Z}$  to  $\mathbb{P}(\Omega)$  for all  $a \in \mathcal{A}$ :

$$\boldsymbol{\Phi}^a(\boldsymbol{\rho}, z) := \left[ \rho_1 \frac{q_1^a(z)}{q_{\boldsymbol{\rho}}^a(z)}, \rho_2 \frac{q_2^a(z)}{q_{\boldsymbol{\rho}}^a(z)}, \dots, \rho_M \frac{q_M^a(z)}{q_{\boldsymbol{\rho}}^a(z)} \right], \quad (3.2)$$

where  $q_{\boldsymbol{\rho}}^a(z) = \sum_{i=1}^M \rho_i q_i^a(z)$ , and  $\boldsymbol{\Phi}^a(\boldsymbol{\rho}, z) = \boldsymbol{\rho}$  if  $q_{\boldsymbol{\rho}}^a(z) = 0$ . In other words, if  $\boldsymbol{\rho} \in \mathbb{P}(\Omega)$  is an apriori distribution,  $\boldsymbol{\Phi}^a(\boldsymbol{\rho}, z)$  gives the posteriori distribution when sensing action  $a$  has been taken and  $z$  has been observed.

We define Markov operator  $\mathbb{T}^a$ ,  $a \in \mathcal{A}$ , such that for any measurable function  $g : \mathbb{P}(\Omega) \rightarrow \mathbb{R}$ ,

$$(\mathbb{T}^a g)(\boldsymbol{\rho}) := \int g(\boldsymbol{\Phi}^a(\boldsymbol{\rho}, z)) q_{\boldsymbol{\rho}}^a(z) dz. \quad (3.3)$$

Note that at any given information state  $\boldsymbol{\rho}$ , taking sensing action  $a \in \mathcal{A}$  followed by the optimal policy results in expected total cost  $1 + (\mathbb{T}^a V^*)(\boldsymbol{\rho})$  where 1 denotes the one unit of time spent to take the sensing action and collect the corresponding observation sample, and  $(\mathbb{T}^a V^*)(\boldsymbol{\rho})$  is the expected value of  $V^*$  on the space of posterior beliefs; while declaration  $j$  results in expected cost  $(1 - \rho_j)L$  where  $(1 - \rho_j)$  is the probability that hypothesis  $H_j$  is not true, and  $L$  is the penalty of making a wrong declaration. This intuition, while relying on the compactness of  $\mathbb{P}(\Omega)$  to treat various measurability issues, can be formalized in the following dynamic programming equation.

**Theorem 3.2.1.** *The optimal value function  $V^* : \mathbb{P}(\Omega) \rightarrow \mathbb{R}_+$  is the unique solution to the following fixed point DP equation:*

$$V^*(\boldsymbol{\rho}) = \min \left\{ 1 + \min_{a \in \mathcal{A}} (\mathbb{T}^a V^*)(\boldsymbol{\rho}), \min_{j \in \Omega} (1 - \rho_j)L \right\}. \quad (3.4)$$

*Proof.* From [49, Proposition 9.8], we know that value function  $V^*$  satisfies the DP equation (3.4). We prove that  $V^*$  is the unique solution to (3.4) by showing that any functional  $V$  that satisfies (3.4) with  $\leq$  ( $\geq$ ) is a lower (upper) bound for the optimal value function  $V^*$ . The detailed proof is provided in Section 3.5.1.  $\square$

**Definition 3.2.2.** A Markov stationary *policy* is a stochastic kernel from the information state space  $\mathbb{P}(\Omega)$  to  $\mathcal{A} \cup \{d\}$  describing the conditional distribution on sensing actions  $A(t)$ ,  $t = 0, 1, \dots, \tau - 1$  and stopping time  $\tau$  (the choice of declaration  $d$  marks the stopping time  $\tau$ ). In other words, under policy  $\pi$ , the probability that action  $a$  is selected at belief state  $\boldsymbol{\rho}$  is given by  $\pi(a|\boldsymbol{\rho})$ .

Equation (3.4) provides a characterization of an optimal Markov stationary deterministic policy  $\pi^*$  for the active hypothesis testing problem as follows: Sensing action  $a^* = \arg \min_{a \in \mathcal{A}} (\mathbb{T}^a V^*)(\boldsymbol{\rho})$  is the least costly sensing action, resulting in  $1 + \min_{a \in \mathcal{A}} (\mathbb{T}^a V^*)(\boldsymbol{\rho})$ , hence is the optimal action to take unless wrongly declaring

$H_{i^*}$ , where  $i^* = \arg \min_{j \in \Omega} (1 - \rho_j)L$ , is even less costly in which case it is optimal to retire and declare  $H_{i^*}$  as the true hypothesis.

**Remark 3.2.1.** It follows from (3.4) that if  $\min_{j \in \Omega} (1 - \rho_j)L \leq 1$ , then we have a full characterization of  $V^*(\boldsymbol{\rho}) = \min_{j \in \Omega} (1 - \rho_j)L$  and the optimal policy. Therefore, the region of interest in our analysis is restricted to  $L > 1$  and  $\mathbb{P}_L(\Omega) := \{\boldsymbol{\rho} \in \mathbb{P}(\Omega) : \min_{j \in \Omega} (1 - \rho_j)L > 1\}$ .

In [20], DeGroot defined the information of an action  $a$  as the difference between the uncertainty prior to observing its corresponding outcome and the expected uncertainty after having observed the outcome. More precisely:

**Definition 3.2.3.** The *information* of action  $a$  at belief state  $\boldsymbol{\rho}$  relative to the (uncertainty) function  $V : \mathbb{P}(\Omega) \rightarrow \mathbb{R}$  is defined as

$$\mathcal{I}(\boldsymbol{\rho}, a, V) := V(\boldsymbol{\rho}) - (\mathbb{T}^a V)(\boldsymbol{\rho}). \quad (3.5)$$

Policy  $\pi^*$  can also be interpreted from DeGroot's point of view as the one that greedily maximizes the information relative to  $V^*$ .

Although optimal policy  $\pi^*$  can be characterized using the optimal value function  $V^*$ , finding a closed-form formulation for  $V^*$ , in general, might not be feasible. A possible solution is to compute  $V^*$  numerically using the value iteration technique [49, Chapter 9.5] or to estimate it using basis functions. The former guarantees convergence to  $V^*$  [49, Proposition 9.17] but it becomes intractable as the number of hypotheses grows; while the latter is easier to implement but provides no convergence guarantee. In lieu of numerical approximation of or derivation of a closed-form for  $V^*$ , we introduce alternative notions of information in Section 3.3 maximizing which gives rise to simple deterministic and Markov heuristic policies. Next we discuss a sufficient condition which reduces the active hypothesis test to a passive one.

### 3.2.1 Reduction to Passive Sensing

In this section, we provide a sufficient condition under which there exists an action whose performance is superior to all other actions and hence, the problem

of active hypothesis testing reduces to passive testing. Before we proceed, we need the following definition:

**Definition 3.2.4** (Blackwell Ordering [13]). Given two groups of conditional probability densities  $q^a = \{q_i^a\}_{i \in \Omega}$  and  $q^b = \{q_i^b\}_{i \in \Omega}$  (on space  $\mathcal{Z}$ ), we say that  $q^b$  is *less informative than*  $q^a$  ( $q^b \leq_B q^a$ ) if there exists a *stochastic transformation*  $W$  from  $\mathcal{Z}$  to  $\mathcal{Z}$  such that<sup>1</sup>

$$q_i^b(z) = \int q_i^a(y) W(y; z) dy \quad \forall i \in \Omega. \quad (3.6)$$

The following fact is an important outcome of Blackwell ordering.

**Fact 3.2.1** (see [16] ch. 14.17). *Let  $q^a = \{q_i^a\}_{i \in \Omega}$  and  $q^b = \{q_i^b\}_{i \in \Omega}$  be two groups of observation kernels. If  $q^b \leq_B q^a$ , then  $(\mathbb{T}^a g)(\boldsymbol{\rho}) \leq (\mathbb{T}^b g)(\boldsymbol{\rho})$  for all  $\boldsymbol{\rho} \in \mathbb{P}(\Omega)$  and for any concave function  $g : \mathbb{P}(\Omega) \rightarrow \mathbb{R}$ .*

**Theorem 3.2.2.** *The value function  $V^*(\cdot)$  is concave.*

The proof of Theorem 3.2.2 is provided in Section 3.5.2.

Combining Theorem 3.2.2 and Fact 3.2.1, the following result is obtained:

**Corollary 3.2.1.** *Consider the problem of active hypothesis testing such that there exists a sensing action  $a^*$  satisfying  $q^a \leq_B q^{a^*}$  for all  $a \in \mathcal{A}$ . It is always optimal to take sensing action  $a^*$  (independently of belief vector) up to the stopping time.*

Corollary 3.2.1 provides a sufficient condition to reduce the active hypothesis test to a passive test. In other words, under the condition of Corollary 3.2.1, the optimal solution to the active hypothesis testing problem remains the same for any  $\mathcal{A}' \subseteq \mathcal{A}$  such that  $a^* \in \mathcal{A}'$ .

It is also interesting to analyze the behavior of greedy policies when the sufficient condition provided by Corollary 3.2.1 holds. The next corollary of Fact 3.2.1 shows that in this case, any policy  $\pi$  that greedily maximizes the information relative to a concave uncertainty function always selects sensing action  $a^*$ .

---

<sup>1</sup>Function  $W : \mathcal{Z} \times \mathcal{Z} \rightarrow \mathbb{R}_+$  is called a *stochastic transformation* from  $\mathcal{Z}$  to  $\mathcal{Z}$  if it satisfies  $\int_{\mathcal{Z}} W(y; z) dz = 1$  for all  $y \in \mathcal{Z}$  and  $\int_{\mathcal{Z}} W(y; z) dy < \infty$  for all  $z \in \mathcal{Z}$ .

**Corollary 3.2.2.** *Consider the problem of active hypothesis testing such that there exists a sensing action  $a^*$  satisfying  $q^a \leq_B q^{a^*}$  for all  $a \in \mathcal{A}$ . Let  $\pi$  be a policy that maximizes in each step the information relative to a concave (uncertainty) function. Policy  $\pi$  always selects sensing action  $a^*$  independently from the belief state prior to its retire-declare decision.*

### 3.3 Heuristic Markov Policies

In this section, we consider heuristic Markov policies based on the following two principles:

1. If  $\rho_i(t) \geq 1 - L^{-1}$ , retire and declare  $H_i$  as the true hypothesis;
2. At each time  $t$  prior to the stopping time, select action  $\arg \max_{a \in \mathcal{A}} \mathcal{I}(\boldsymbol{\rho}(t), a, V)$  where  $V : \mathbb{P}(\Omega) \rightarrow \mathbb{R}$  is a concave uncertainty function.<sup>2</sup>

The first principle follows from the fact that when  $L(1 - \rho_i(t)) \leq 1$ , then  $LPe \leq 1$  and further reduction in  $Pe$  is not worth taking one more sensing action and hence increasing the stopping time  $\tau$  by 1. The intuition behind the second principle is that, by choosing an appropriate notion of information, the problem of active sequential hypothesis testing can be reduced to a sequence of one-shot problems in each of which an optimal sensing action is selected deterministically so as to provide the highest amount of information.

In Section 3.3.1, we give the definition of *Jensen–Shannon (JS) divergence*, and we introduce *Extrinsic Jensen–Shannon (EJS) divergence*. In Section 3.3.2, we show that JS and EJS divergences are equal to information relative to the Shannon entropy and average log-likelihood function, respectively. Furthermore, we propose two heuristic policies based on greedy maximization of these divergences.

---

<sup>2</sup>Intuitively, information of a sensing action should always be non-negative, i.e., an action can, at worst, contain no information. It is known that  $\mathcal{I}(\boldsymbol{\rho}, a, V) \geq 0$  for all  $\boldsymbol{\rho} \in \mathbb{P}(\Omega)$  and all  $a \in \mathcal{A}$  if and only if  $V$  is concave [20, Theorem 2.1].

### 3.3.1 Symmetric Divergences

We first recall some well known divergences. The *Kullback–Leibler (KL) divergence* between two probability distributions  $P_Z$  and  $P'_Z$  over a finite set  $\mathcal{Z}$  is defined as  $D(P_Z \| P'_Z) := \sum_{z \in \mathcal{Z}} P_Z(z) \log \frac{P_Z(z)}{P'_Z(z)}$  with the convention  $0 \log \frac{a}{0} = 0$  and  $b \log \frac{b}{0} = \infty$  for  $a, b \in [0, 1]$  with  $b \neq 0$ . The KL divergence satisfies the following lemma.

**Lemma 3.3.1.** *For any two distributions  $P$  and  $Q$  on a set  $\mathcal{Z}$  and  $\alpha \in [0, 1]$ ,  $D(P \| \alpha P + (1 - \alpha)Q)$  is decreasing in  $\alpha$ .*

*Proof.* Let  $\beta \in [0, 1]$  satisfy  $\beta \leq \alpha$ . Then,

$$\alpha P + (1 - \alpha)Q = \gamma(\beta P + (1 - \beta)Q) + (1 - \gamma)P,$$

where  $\gamma = \frac{1-\alpha}{1-\beta} \leq 1$ . By Jensen's inequality and the convexity of the KL divergence:

$$\begin{aligned} D(P \| \alpha P + (1 - \alpha)Q) &\leq \gamma D(P \| \beta P + (1 - \beta)Q) + (1 - \gamma)D(P \| P) \\ &\leq D(P \| \beta P + (1 - \beta)Q), \end{aligned} \quad (3.7)$$

where the last inequality follows because  $D(P \| P) = 0$  and  $\gamma \leq 1$ .  $\square$

The KL divergence is *not* symmetric, i.e., in general

$$D(P_Z \| P'_Z) \neq D(P'_Z \| P_Z).$$

The *J divergence* [50] and *L divergence* [51] symmetrize the KL divergence:

$$J(P_1, P_2) := D(P_1 \| P_2) + D(P_2 \| P_1), \quad (3.8)$$

$$L(P_1, P_2) := D\left(P_1 \| \frac{1}{2}P_1 + \frac{1}{2}P_2\right) + D\left(P_2 \| \frac{1}{2}P_1 + \frac{1}{2}P_2\right). \quad (3.9)$$

The L divergence can be related to the *Jensen difference* with respect to the Shannon entropy function [52]:

$$\frac{1}{2}L(P_1, P_2) = H\left(\frac{1}{2}P_1 + \frac{1}{2}P_2\right) - \left(\frac{1}{2}H(P_1) + \frac{1}{2}H(P_2)\right). \quad (3.10)$$

The *Jensen–Shannon (JS) divergence* [51, 52] is defined as an  $M$ -dimensional generalization of the L divergence. Given  $M$  distributions  $P_1, P_2, \dots, P_M$  over a set

$\mathcal{Z}$  and a vector of a priori weights  $\boldsymbol{\rho} = [\rho_1, \rho_2, \dots, \rho_M]$ , where  $\boldsymbol{\rho} \in [0, 1]^M$  and  $\sum_{i=1}^M \rho_i = 1$ , the JS divergence is defined as [51, 52]:

$$\begin{aligned} JS(\boldsymbol{\rho}; P_1, \dots, P_M) &:= \sum_{i=1}^M \rho_i D\left(P_i \parallel \sum_{j=1}^M \rho_j P_j\right) \\ &= H\left(\sum_{i=1}^M \rho_i P_i\right) - \sum_{i=1}^M \rho_i H(P_i). \end{aligned} \quad (3.11)$$

Let  $\theta$  be a random variable that takes values in  $\{1, 2, \dots, M\}$  and has probability mass function  $\boldsymbol{\rho}$  and  $Z \sim P_\theta$  (which implies that  $P(Z = z) = \sum_{i=1}^M \rho_i P_i(z)$ ). From (3.11),

$$JS(\boldsymbol{\rho}; P_1, \dots, P_M) = H(Z) - H(Z|\theta) = I(\theta; Z), \quad (3.12)$$

where  $I(\theta; Z)$  is the mutual information between  $\theta$  and  $Z$ .

From (3.12) and the fact that  $I(\theta; Z) = H(\theta) - H(\theta|Z)$ , the JS divergence can also be expressed as:

$$JS(\boldsymbol{\rho}; P_1, \dots, P_M) = H(\boldsymbol{\rho}) - \sum_{z \in \mathcal{Z}} P_\rho(z) H\left(\left[\frac{\rho_1 P_1(z)}{P_\rho(z)}, \dots, \frac{\rho_M P_M(z)}{P_\rho(z)}\right]\right), \quad (3.13)$$

where  $P_\rho = \sum_{i=1}^M \rho_i P_i$ .

## A New Divergence: Extrinsic Jensen–Shannon (EJS) Divergence

We introduce the *Extrinsic Jensen–Shannon (EJS) divergence* as an  $M$ -dimensional generalization of J divergence as

$$EJS(\boldsymbol{\rho}; P_1, \dots, P_M) := \sum_{i=1}^M \rho_i D\left(P_i \parallel \sum_{j \neq i} \frac{\rho_j}{1 - \rho_i} P_j\right), \quad (3.14a)$$

when  $\rho_i < 1$  for all  $i \in \{1, \dots, M\}$ , and as

$$EJS(\boldsymbol{\rho}; P_1, \dots, P_M) := \max_{j \neq i} D(P_i \parallel P_j) \quad (3.14b)$$

when  $\rho_i = 1$  for some  $i \in \{1, \dots, M\}$ .

Let  $U(\cdot)$  denote the average log-likelihood function:

$$U(\boldsymbol{\rho}) := \sum_{i=1}^M \rho_i \log \frac{1 - \rho_i}{\rho_i}. \quad (3.15)$$

**Lemma 3.3.2** (Properties of EJS Divergence). *The EJS divergence defined in (3.14) satisfies the following three properties.*

1. *It is lower bounded by the JS divergence:*

$$EJS(\boldsymbol{\rho}; P_1, \dots, P_M) \geq JS(\boldsymbol{\rho}; P_1, \dots, P_M). \quad (3.16)$$

2. *It can be expressed as*

$$EJS(\boldsymbol{\rho}; P_1, \dots, P_M) = U(\boldsymbol{\rho}) - \sum_{z \in \mathcal{Z}} P_{\boldsymbol{\rho}}(z) U\left(\left[\frac{\rho_1 P_1(z)}{P_{\boldsymbol{\rho}}(z)}, \dots, \frac{\rho_M P_M(z)}{P_{\boldsymbol{\rho}}(z)}\right]\right). \quad (3.17)$$

3. *It is convex in the distributions  $P_1, \dots, P_M$ .*

*Proof of Lemma 3.3.2.* Property 1 is proved as follows:

$$\begin{aligned} JS(\boldsymbol{\rho}; P_1, \dots, P_M) &= \sum_{i=1}^M \rho_i D\left(P_i \parallel \sum_{j=1}^M \rho_j P_j\right) \\ &= \sum_{i=1}^M \rho_i D\left(P_i \parallel \rho_i P_i + (1 - \rho_i) \sum_{j \neq i} \frac{\rho_j}{1 - \rho_i} P_j\right) \\ &\stackrel{(a)}{\leq} \sum_{i=1}^M \left[ \rho_i^2 D(P_i \parallel P_i) + \rho_i (1 - \rho_i) D\left(P_i \parallel \sum_{j \neq i} \frac{\rho_j}{1 - \rho_i} P_j\right) \right] \\ &= EJS(\boldsymbol{\rho}; P_1, \dots, P_M) - \sum_{i=1}^M \rho_i^2 D\left(P_i \parallel \sum_{j \neq i} \frac{\rho_j}{1 - \rho_i} P_j\right) \\ &\stackrel{(b)}{\leq} EJS(\boldsymbol{\rho}; P_1, \dots, P_M), \end{aligned}$$

where (a) and (b) follow respectively because KL divergence is convex (in both arguments) and nonnegative.

The proof of property 2 is provided next.

$$\begin{aligned} EJS(\boldsymbol{\rho}; P_1, \dots, P_M) &= \sum_{i=1}^M \rho_i D\left(P_i \parallel \sum_{j \neq i} \frac{\rho_j}{1 - \rho_i} P_j\right) \\ &= \sum_{i=1}^M \rho_i \sum_{z \in \mathcal{Z}} P_i(z) \log \frac{P_i(z)}{\sum_{j \neq i} \frac{\rho_j}{1 - \rho_i} P_j(z)} \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^M \rho_i \log \frac{1 - \rho_i}{\rho_i} + \sum_{i=1}^M \sum_{z \in \mathcal{Z}} \rho_i P_i(z) \log \frac{\rho_i P_i(z)}{\sum_{j \neq i} \rho_j P_j(z)} \\
&= U(\boldsymbol{\rho}) + \sum_{z \in \mathcal{Z}} P_{\boldsymbol{\rho}}(z) \sum_{i=1}^M \frac{\rho_i P_i(z)}{P_{\boldsymbol{\rho}}(z)} \log \frac{\frac{\rho_i P_i(z)}{P_{\boldsymbol{\rho}}(z)}}{1 - \frac{\rho_i P_i(z)}{P_{\boldsymbol{\rho}}(z)}} \\
&= U(\boldsymbol{\rho}) - \sum_{z \in \mathcal{Z}} P_{\boldsymbol{\rho}}(z) U\left(\left[\frac{\rho_1 P_1(z)}{P_{\boldsymbol{\rho}}(z)}, \dots, \frac{\rho_M P_M(z)}{P_{\boldsymbol{\rho}}(z)}\right]\right).
\end{aligned}$$

Property 3 is proved as follows:

Let  $P_1, P_2, \dots, P_M$  and  $Q_1, Q_2, \dots, Q_M$  be two sets of distributions. For any  $\lambda \in [0, 1]$  and  $\bar{\lambda} = 1 - \lambda$ ,

$$\begin{aligned}
&EJS(\boldsymbol{\rho}; \lambda P_1 + \bar{\lambda} Q_1, \dots, \lambda P_M + \bar{\lambda} Q_M) \\
&= \sum_{i=1}^M \rho_i D\left(\lambda P_i + \bar{\lambda} Q_i \parallel \sum_{j \neq i} \frac{\rho_j}{1 - \rho_i} \lambda P_j + \sum_{j \neq i} \frac{\rho_j}{1 - \rho_i} \bar{\lambda} Q_j\right) \\
&\stackrel{(a)}{\leq} \sum_{i=1}^M \rho_i \left[ \lambda D\left(P_i \parallel \sum_{j \neq i} \frac{\rho_j}{1 - \rho_i} P_j\right) + \bar{\lambda} D\left(Q_i \parallel \sum_{j \neq i} \frac{\rho_j}{1 - \rho_i} Q_j\right) \right] \\
&= \lambda EJS(\boldsymbol{\rho}; P_1, \dots, P_M) + \bar{\lambda} EJS(\boldsymbol{\rho}; Q_1, \dots, Q_M)
\end{aligned}$$

where (a) follows because KL divergence is convex in both arguments.  $\square$

**Remark 3.3.1.** The EJS divergence defined in this section is not the unique generalization of the J divergence. There exist other  $M$ -dimensional generalizations of the J divergence such as  $\sum_{i=1}^M \rho_i \sum_{j=1}^M \rho_j J(P_i, P_j)$  which was studied in [53]. However, as will be discussed in details later, properties of EJS such as the one provided by (3.17) above makes it a proper notion of information for our applications of interest.

**Remark 3.3.2.** The EJS divergence is equivalent to the *full anthropic correction* proposed in the context of mutual information estimation [54]. In particular, the authors in [54] used the notion of anthropic correction as an estimator of the mutual information between signals acquired in neurophysiological experiments where only a small number of stimuli can be tested.

### 3.3.2 Heuristic Policies: Maximizing Divergence

In this section, we show that JS and EJS divergences are proper notions of information, and propose deterministic Markov policies based on greedy maximization of these divergences.

Given a belief vector  $\boldsymbol{\rho} \in \mathbb{P}(\Omega)$  and a sensing action  $a \in \mathcal{A}$ , we use the notations

$$JS(\boldsymbol{\rho}, a) := JS(\boldsymbol{\rho}; q_1^a, q_2^a, \dots, q_M^a), \quad (3.18)$$

$$EJS(\boldsymbol{\rho}, a) := EJS(\boldsymbol{\rho}; q_1^a, q_2^a, \dots, q_M^a). \quad (3.19)$$

Definition 3.2.3 together with properties (3.11) and (3.17) shows that the JS and EJS divergences are nothing but the information relative to the entropy and average log-likelihood functions, respectively, i.e.,

$$JS(\boldsymbol{\rho}, a) = \mathcal{I}(\boldsymbol{\rho}, a, H), \quad (3.20)$$

$$EJS(\boldsymbol{\rho}, a) = \mathcal{I}(\boldsymbol{\rho}, a, U). \quad (3.21)$$

We also use the following notations to denote the amount of information a Markov stationary policy  $\pi$  obtains in a single step:

$$JS(\boldsymbol{\rho}, \pi) := \sum_{a \in \mathcal{A}} \pi(a|\boldsymbol{\rho}) JS(\boldsymbol{\rho}, a), \quad (3.22)$$

$$EJS(\boldsymbol{\rho}, \pi) := \sum_{a \in \mathcal{A}} \pi(a|\boldsymbol{\rho}) EJS(\boldsymbol{\rho}, a). \quad (3.23)$$

We are now ready to introduce our heuristic policies.

Policy  $\pi_{JS}$  is defined as follows:

- if  $\rho_i(t) \geq 1 - L^{-1}$ , retire and declare  $H_i$  as the true hypothesis;
- otherwise, select  $\arg \max_{a \in \mathcal{A}} JS(\boldsymbol{\rho}(t), a)$ .

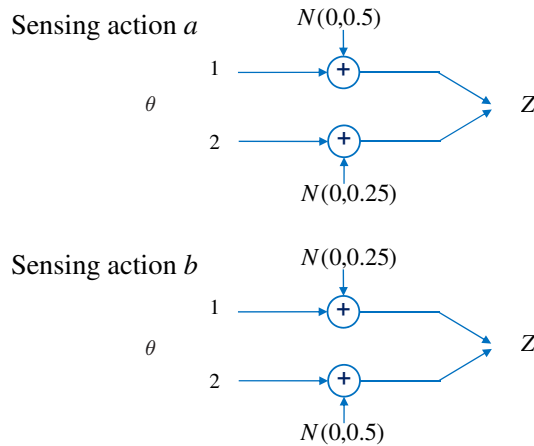
Similarly, policy  $\pi_{EJS}$  is defined as follows:

- if  $\rho_i(t) \geq 1 - L^{-1}$ , retire and declare  $H_i$  as the true hypothesis;
- otherwise, select  $\arg \max_{a \in \mathcal{A}} EJS(\boldsymbol{\rho}(t), a)$ .

**Remark 3.3.3.** Note that as the belief about one of the hypotheses, say  $\rho_i$ , approaches 1,  $D(q_i^a \| \sum_{j=1}^M \rho_j q_j^a)$  converges to  $D(q_i^a \| q_i^a) = 0$  for any action  $a \in \mathcal{A}$ ; and consequently, independently of the observation kernels  $q_1^a, q_2^a, \dots, q_M^a$ , divergence  $JS(\boldsymbol{\rho}, a)$  approaches 0. In contrast, as  $\rho_i$  becomes large,  $EJS(\boldsymbol{\rho}, a)$  approaches  $D(q_i^a \| \sum_{j \neq i} \frac{\rho_j}{1-\rho_i} q_j^a)$  and hence,  $\pi_{EJS}$  selects action  $a$  such that  $D(q_i^a \| \sum_{j \neq i} \frac{\rho_j}{1-\rho_i} q_j^a)$  is maximized, i.e., it selects an action that distinguishes  $H_i$  from the collection of alternate hypotheses the most. As we will see in the next section, these different philosophies result in significant performance difference.

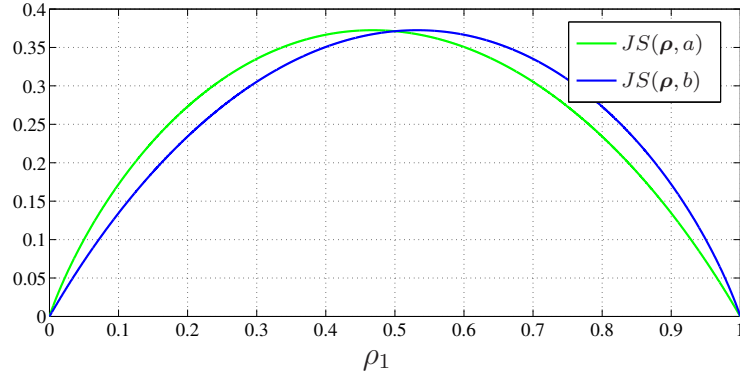
### 3.4 Numerical Example

Consider the active binary hypothesis testing problem with additive Gaussian noisy observations under two actions  $a$  and  $b$  shown in Fig. 3.1. In this example, the observation noise associated with actions  $a$  and  $b$  is such that it adds unequal noise to the hypotheses. In the remainder of this section, we compare the performance of policies  $\pi^*$ ,  $\pi_{JS}$ , and  $\pi_{EJS}$  for this example. In our performance comparisons, we also consider the sequential non-adaptive policy  $\pi_{SN}$  which prior to its stopping time selects actions  $a$  and  $b$  uniformly at random, i.e.,  $\pi_{SN}(a|\boldsymbol{\rho}) = \pi_{SN}(b|\boldsymbol{\rho}) = 0.5$  for any  $\boldsymbol{\rho} \in \mathbb{P}_L(\Omega)$ .

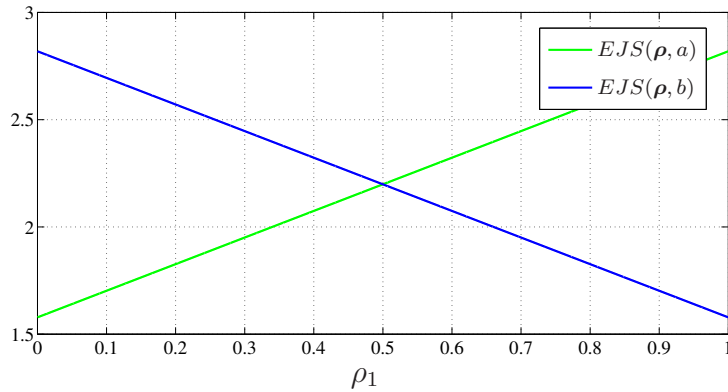


**Figure 3.1:** Active binary hypothesis testing with additive Gaussian noise.

Figures 3.2 and 3.3 show respectively the JS and EJS divergence for all  $\rho \in \mathbb{P}(\Omega)$  and for sensing actions  $a$  and  $b$ . It is clear from these figures that  $\pi_{JS}$  selects action  $a$  when  $\rho_1 \geq 0.5$  and selects action  $b$  otherwise; while  $\pi_{EJS}$  does exactly the opposite. These figures illustrate the fact mentioned in Remark 3.3.3.

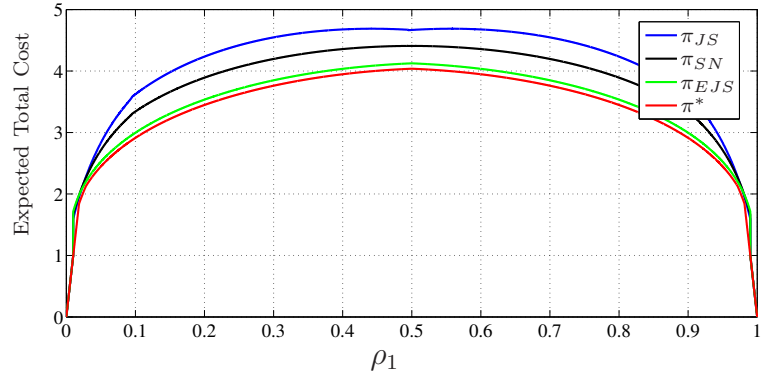
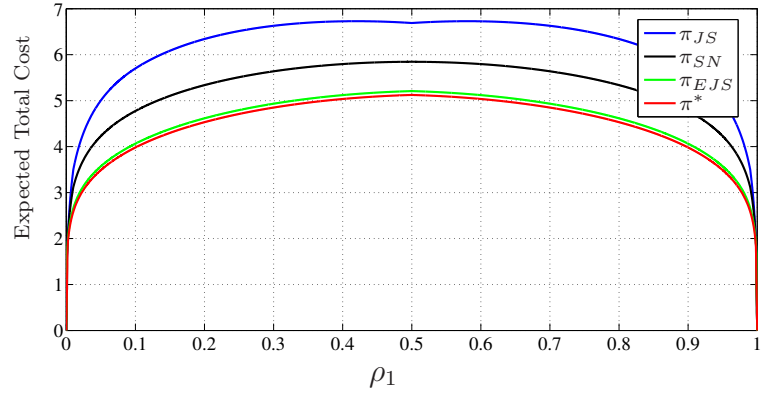
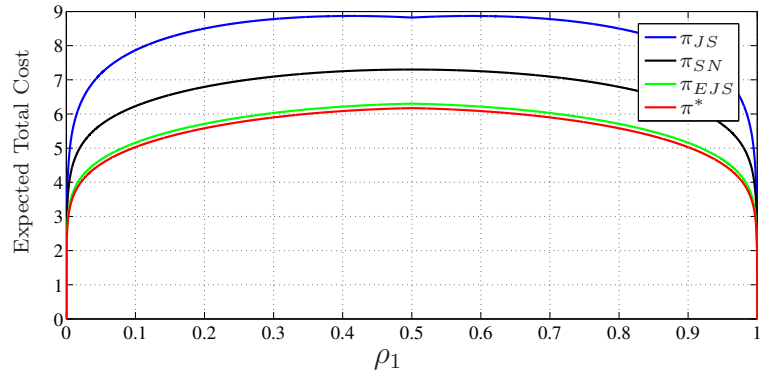


**Figure 3.2:** JS divergence for sensing actions  $a$  and  $b$ .



**Figure 3.3:** EJS divergence for sensing actions  $a$  and  $b$ .

Figure 3.4 plots the expected total cost of policies  $\pi^*$ ,  $\pi_{JS}$ ,  $\pi_{EJS}$ , and  $\pi_{SN}$  for different values of  $L$ . Figure 3.4 shows that  $\pi_{EJS}$  performs the best among the heuristic policies under investigation; while  $\pi_{JS}$  is the farthest from the optimal performance. Even more significant is the growth in performance gap between the optimal policy  $\pi^*$  and policies  $\pi_{JS}$  and  $\pi_{SN}$  as  $L$  increases. In particular, the growth in the performance gap between  $\pi^*$  and  $\pi_{SN}$  is an illustration of the adaptivity gain in active binary hypothesis testing and is consistent with Corollary 2.5.1.

(a)  $L = 100$ .(b)  $L = 1000$ .(c)  $L = 10000$ .**Figure 3.4:** Expected total cost of the proposed policies for different values of  $L$ .

Analytic results regarding the performance of the proposed heuristics are provided in the next chapter.

## 3.5 Proofs

### 3.5.1 Proof of Theorem 3.2.1

The proof follows by combining Fact 3.5.1, Lemma 3.5.1, and Lemma 3.5.2 as given below.

**Fact 3.5.1** (Proposition 9.8 in [49]). *The optimal value function  $V^* : \mathbb{P}(\Omega) \rightarrow \mathbb{R}_+$  satisfies the following fixed point equation:*

$$V^*(\boldsymbol{\rho}) = \min \left\{ 1 + \min_{a \in \mathcal{A}} (\mathbb{T}^a V^*)(\boldsymbol{\rho}), \min_{j \in \Omega} (1 - \rho_j) L \right\}. \quad (3.24)$$

**Lemma 3.5.1.** *Suppose there exist  $\beta > 0$  and a functional  $V : \mathbb{P}(\Omega) \rightarrow \mathbb{R}_+$  such that for all belief vectors  $\boldsymbol{\rho} \in \mathbb{P}(\Omega)$ ,*

$$V(\boldsymbol{\rho}) \leq \min \left\{ \beta + \min_{a \in \mathcal{A}} (\mathbb{T}^a V)(\boldsymbol{\rho}), \min_{j \in \Omega} (1 - \rho_j) \beta L \right\}. \quad (3.25)$$

*Then  $V^*(\boldsymbol{\rho}) \geq \frac{1}{\beta} V(\boldsymbol{\rho})$  for all  $\boldsymbol{\rho} \in \mathbb{P}(\Omega)$ .*

*Proof of Lemma 3.5.1.* To prove Lemma 3.5.1, we have to slightly modify the state space and introduce new notations. We assume that after taking the retire-declare action, the system goes to the termination state, denoted by  $F$ , and remains in that state for the rest of the time. The state space is modified to  $\mathcal{S} = \mathbb{P}(\Omega) \cup \{F\}$  to include the termination state. For  $a \in \mathcal{A} \cup \{d\}$ ,  $s \in \mathcal{S}$ , let

$$c^a(s) = \begin{cases} 1 & \text{if } s = \boldsymbol{\rho} \in \mathbb{P}(\Omega), a \in \mathcal{A} \\ \min_{j \in \Omega} (1 - \rho_j) L & \text{if } s = \boldsymbol{\rho} \in \mathbb{P}(\Omega), a = d \\ 0 & \text{if } s = F \end{cases}.$$

The Bayes operator is modified as follows:

$$\Phi^a(s, z) = \begin{cases} \Phi^a(\boldsymbol{\rho}, z) & \text{if } s = \boldsymbol{\rho} \in \mathbb{P}(\Omega), a \in \mathcal{A} \\ F & \text{if } s = \boldsymbol{\rho} \in \mathbb{P}(\Omega), a = d \\ F & \text{if } s = F \end{cases}.$$

Using the notations above, condition (3.25) is rewritten as

$$\begin{aligned} V(F) &= 0, \\ V(s) &\leq \min_{a \in \mathcal{A} \cup \{d\}} \{\beta c^a(s) + \mathbb{E}[V(\Phi^a(s, Z))]\}, \quad \forall s \in \mathcal{S} - \{F\}. \end{aligned} \quad (3.26)$$

Let  $S_0, S_1, S_2, \dots$  be a sequence of random variables denoting the belief states at times  $t = 0, 1, 2, \dots$  starting from belief state  $s$ , i.e.,

$$\begin{aligned} S_0 &= s, \\ S_n &= \Phi^{A(n-1)}(S_{n-1}, Z), \quad \forall n, n > 0. \end{aligned}$$

Using (3.26) iteratively for  $N$  times, we obtain

$$\begin{aligned} V(s) &\leq \beta \mathbb{E}_{\pi^*}[c^{A(0)}(s)] + \mathbb{E}_{\pi^*}[V(\Phi^{A(0)}(s, Z))] \\ &= \beta \mathbb{E}_{\pi^*}[c^{A(0)}(S_0)] + \mathbb{E}_{\pi^*}[V(S_1)] \\ &\leq \beta \mathbb{E}_{\pi^*}\left[\sum_{n=0}^1 c^{A(n)}(S_n)\right] + \mathbb{E}_{\pi^*}[V(S_2)] \\ &\leq \beta \mathbb{E}_{\pi^*}\left[\sum_{n=0}^{N-1} c^{A(n)}(S_n)\right] + \mathbb{E}_{\pi^*}[V(S_N)], \end{aligned}$$

where subscript  $\pi^*$  implies that actions are selected according to an optimal policy  $\pi^*$ .<sup>3</sup> Taking the limit as  $N \rightarrow \infty$ , we obtain

$$\begin{aligned} V(s) &\stackrel{(a)}{\leq} \beta \mathbb{E}_{\pi^*}\left[\sum_{n=0}^{\infty} c^{A(n)}(S_n)\right] + \lim_{N \rightarrow \infty} \mathbb{E}_{\pi^*}[V(S_N)] \\ &\stackrel{(b)}{=} \beta V^*(s) + \lim_{N \rightarrow \infty} \mathbb{E}_{\pi^*}[V(S_N)] \\ &= \beta V^*(s) + \lim_{N \rightarrow \infty} \mathbb{E}_{\pi^*}[V(F)\mathbf{1}_{\{S_N=F\}} + V(S_N)\mathbf{1}_{\{S_N \neq F\}}] \\ &= \beta V^*(s) + \lim_{N \rightarrow \infty} \mathbb{E}_{\pi^*}[V(S_N)\mathbf{1}_{\{S_N \neq F\}}] \\ &\stackrel{(c)}{\leq} \beta V^*(s) + L \lim_{N \rightarrow \infty} P_{\pi^*}(S_N \neq F) \\ &\stackrel{(d)}{=} \beta V^*(s), \end{aligned}$$

where (a) follows from the monotone convergence theorem, (b) follows from the definition of  $V^*$ , (c) follows from the fact that  $V(\boldsymbol{\rho}) \leq \min_{j \in \Omega} (1 - \rho_j)L \leq L$  for

---

<sup>3</sup>The existence of an optimal policy follows from [49, Corollary 9.12.1] and since  $|\mathcal{A}| < \infty$ .

any  $\boldsymbol{\rho} \in \mathbb{P}(\Omega)$ , and (d) holds since  $L \geq V^*(s) \geq \mathbb{E}_{\pi^*}[\tau] = \sum_{n=0}^{\infty} P_{\pi^*}(\tau > n) = \sum_{n=0}^{\infty} P_{\pi^*}(S_n \neq F)$ .  $\square$

**Lemma 3.5.2.** *Suppose there exist  $\beta > 0$  and a functional  $V : \mathbb{P}(\Omega) \rightarrow \mathbb{R}_+$  such that for all belief vectors  $\boldsymbol{\rho} \in \mathbb{P}(\Omega)$ ,*

$$V(\boldsymbol{\rho}) \geq \min\{\beta + \min_{a \in \mathcal{A}}(\mathbb{T}^a V)(\boldsymbol{\rho}), \min_{j \in \Omega}(1 - \rho_j)\beta L\}. \quad (3.27)$$

Then  $V^*(\boldsymbol{\rho}) \leq \frac{1}{\beta}V(\boldsymbol{\rho})$  for all  $\boldsymbol{\rho} \in \mathbb{P}(\Omega)$ .

*Proof of Lemma 3.5.2.* Following similar lines as those in the proof of Lemma 3.5.1, we get

$$V(s) \geq \beta \mathbb{E}_{\pi^*}[\sum_{n=0}^{N-1} c^{A(n)}(S_n)] + \mathbb{E}_{\pi^*}[V(S_N)].$$

Taking the limit as  $N \rightarrow \infty$ , we obtain

$$\begin{aligned} V(s) &\geq \beta \mathbb{E}_{\pi^*}[\sum_{n=0}^{\infty} c^{A(n)}(S_n)] + \lim_{N \rightarrow \infty} \mathbb{E}_{\pi^*}[V(S_N)] \\ &\stackrel{(a)}{\geq} \beta V^*(s), \end{aligned}$$

where (a) follows from the fact that  $V(\cdot)$  is non-negative.  $\square$

### 3.5.2 Proof of Theorem 3.2.2

We first consider the finite horizon scenario for which the decision maker has to take a retire-declare action by time  $N$ . For each belief vector  $\boldsymbol{\rho}$ , the expected total cost from time  $n$  onward, denoted by  $V_n^N(\boldsymbol{\rho})$ , solves the following iterative system of equations:

$$V_N^N(\boldsymbol{\rho}) = \min_{j \in \Omega}(1 - \rho_j)L, \quad (3.28)$$

$$V_n^N(\boldsymbol{\rho}) = \min \left\{ 1 + \min_{a \in \mathcal{A}}(\mathbb{T}^a V_{n+1}^N)(\boldsymbol{\rho}), V_N^N(\boldsymbol{\rho}) \right\} \quad \text{for } n = 1, 2, \dots, N-1. \quad (3.29)$$

It will be shown that for any  $N$  and  $n$ ,  $0 \leq n \leq N$ , the value function  $V_n^N(\cdot)$  is concave. Then using the fact that  $V^*(\boldsymbol{\rho}) = \lim_{N \rightarrow \infty} V_0^N(\boldsymbol{\rho})$  (Proposition 1.5 in [55]), we have the assertion of the theorem.

We prove the concavity of  $V_n^N(\cdot)$  by induction. From (3.28), it is clear that  $V_N^N(\cdot)$  is a concave function. Now suppose  $V_{n+1}^N(\cdot)$  is concave. In order to prove the concavity of  $V_n^N(\cdot)$ , it suffices to show that for every  $a \in \mathcal{A}$ ,  $(\mathbb{T}^a V_{n+1}^N)(\cdot)$  is concave: Let  $\boldsymbol{\rho}, \boldsymbol{\mu} \in \mathbb{P}(\Omega)$ . For any  $\alpha$ ,  $0 \leq \alpha \leq 1$ , and  $\bar{\alpha} = 1 - \alpha$  we have,

$$\begin{aligned}
& \alpha(\mathbb{T}^a V_{n+1}^N)(\boldsymbol{\rho}) + \bar{\alpha}(\mathbb{T}^a V_{n+1}^N)(\boldsymbol{\mu}) \\
&= \int \alpha V_{n+1}^N(\Phi^a(\boldsymbol{\rho}, z)) q_{\boldsymbol{\rho}}^a(z) dz + \int \bar{\alpha} V_{n+1}^N(\Phi^a(\boldsymbol{\mu}, z)) q_{\boldsymbol{\mu}}^a(z) dz \\
&= \int (\alpha q_{\boldsymbol{\rho}}^a(z) + \bar{\alpha} q_{\boldsymbol{\mu}}^a(z)) \times \\
&\quad \left\{ \frac{\alpha q_{\boldsymbol{\rho}}^a(z)}{\alpha q_{\boldsymbol{\rho}}^a(z) + \bar{\alpha} q_{\boldsymbol{\mu}}^a(z)} V_{n+1}^N(\Phi^a(\boldsymbol{\rho}, z)) + \frac{\bar{\alpha} q_{\boldsymbol{\mu}}^a(z)}{\alpha q_{\boldsymbol{\rho}}^a(z) + \bar{\alpha} q_{\boldsymbol{\mu}}^a(z)} V_{n+1}^N(\Phi^a(\boldsymbol{\mu}, z)) \right\} dz \\
&\leq \int (\alpha q_{\boldsymbol{\rho}}^a(z) + \bar{\alpha} q_{\boldsymbol{\mu}}^a(z)) V_{n+1}^N \left( \frac{\alpha q_{\boldsymbol{\rho}}^a(z) \Phi^a(\boldsymbol{\rho}, z) + \bar{\alpha} q_{\boldsymbol{\mu}}^a(z) \Phi^a(\boldsymbol{\mu}, z)}{\alpha q_{\boldsymbol{\rho}}^a(z) + \bar{\alpha} q_{\boldsymbol{\mu}}^a(z)} \right) dz \\
&= \int V_{n+1}^N(\Phi^a(\alpha \boldsymbol{\rho} + \bar{\alpha} \boldsymbol{\mu}, z)) q_{\alpha \boldsymbol{\rho} + \bar{\alpha} \boldsymbol{\mu}}^a(z) dz \\
&= (\mathbb{T}^a V_{n+1}^N)(\alpha \boldsymbol{\rho} + \bar{\alpha} \boldsymbol{\mu}).
\end{aligned}$$

Therefore,  $(\mathbb{T}^a V_{n+1}^N)(\cdot)$ ,  $\forall a \in \mathcal{A}$ , are concave. From (3.29),  $V_n^N(\cdot)$  is the minimum of concave functions and hence, it is concave.

Chapter 3, in part, has been submitted for publication as M. Naghshvar and T. Javidi, “Active sequential hypothesis testing,” available on arXiv:1203.4626. The dissertation author was the primary investigator and author of this paper.

## Chapter 4

# Non-asymptotic and Asymptotic Analysis of Active Sequential Hypothesis Testing

This chapter provides a detailed analysis of the schemes proposed in Chapter 3, in particular  $\pi_{EJS}$ , and discusses their advantage and disadvantage compared to those of Chapter 2. Unlike the schemes proposed in Chapter 2, policy  $\pi_{EJS}$  is shown to achieve asymptotic optimality (as the penalty of wrong declaration,  $L$ , increases) only in a limited setting. However, this policy has a provable advantage for large number of hypotheses,  $M$ , over those proposed in Chapter 2 as well as other solutions in the literature. More specifically, this policy can provide, under a technical condition, reliability and speedy declaration simultaneously. In information theoretic terms, this policy can be shown to achieve non-zero information acquisition rate and error exponent simultaneously.

Using dynamic programming techniques, three lower bounds are derived for the optimal performance,  $V^*$ . Unlike the lower bounds provided in Chapter 2, these new lower bounds, in general, are not asymptotically tight in  $L$ . However, they are non-asymptotic and complementary for various values of  $L$  and  $M$ . From the proposed upper and lower bounds, sufficient conditions to achieve the maximum information rate and reliability are obtained. Special cases of active hypothesis testing that satisfy these conditions will be provided in the subsequent chapters.

## 4.1 Introduction

One of the main drawbacks of Chernoff’s asymptotic optimality notion was his neglecting the complementary role of asymptotic analysis in  $M$ . In particular, the notion of asymptotic optimality in  $L$  falls short in showing the tension between using an (asymptotically) large number of samples to discriminate among a few hypotheses with (asymptotically) high accuracy or an (asymptotically) large number of hypotheses with a lower degree of accuracy. As a result, although the scheme proposed in [5] and its extensions [25, 30, 32–36] as well as two-phase policy  $\pi_{SA}$  introduced in Chapter 2 are asymptotically optimal in  $L$ , their provable information acquisition rate is restricted to zero. Intuitively, the rate of information acquisition under any heuristic policy relates to the ratio between  $\log M$  and the expected number of samples: the larger this ratio the faster information is acquired.

In Section 4.2, we analyze the performance of the proposed heuristics, and provide upper bounds on the expected total cost achieved by these policies in terms of both  $L$  and  $M$ . To compare the performance of the proposed heuristics with that of the optimal policy  $\pi^*$ , in Section 4.3, we use dynamic programming techniques, in particular Lemma 3.5.1, to find lower bounds for the optimal performance,  $V^*$ . Using the lower bounds obtained in Section 4.3, we propose a new stopping rule in Section 4.4 whose improvement over the one proposed in Chapter 3 is illustrated in a numerical example. Section 4.5 states the asymptotic consequence of the bounds obtained in Sections 4.2 and 4.3. In particular, the obtained bounds are used to 1) establish notions of order and asymptotic optimality for the proposed policies (generalizing that of [5]); and 2) characterize lower and upper bounds on the maximum achievable information acquisition rate and the optimal reliability. In Section 4.6, we discuss the technical assumptions made in our work and contrast them with the assumptions in the literature. More specifically, we show that our first technical assumption weakens significantly one of the assumptions made in [5]. On the other hand, our second technical assumption is significantly stronger than the corresponding assumptions in the literature. We show that while this assumption is critical in obtaining the non-asymptotic upper and lower bounds of Sections 4.2 and 4.3, it has no bearing on our asymptotic results in Section 4.5.

Throughout this chapter, for clarification and to emphasize the dependence of the parameters of the problem on the hypotheses set, hence the number of hypotheses, we stop suppressing the variable  $M$ : we denote the hypotheses set by  $\Omega_M := \{1, 2, \dots, M\}$ , and the action space by  $\mathcal{A}_M$  where  $\mathcal{A}_M$  may depend on  $M$  and is assumed to be finite with  $|\mathcal{A}_M| < \infty$ .

Similar to the analysis in Chapter 2, we have the following Assumptions.

**Assumption 4.1.1.** *For any two hypotheses  $i, j \in \Omega_M$ ,  $i \neq j$ , there exists an action  $a$ ,  $a \in \mathcal{A}_M$ , such that  $D(q_i^a \| q_j^a) > 0$ .*

**Assumption 4.1.2.** *There exists  $\xi_M < \infty$  such that*

$$\max_{i, j \in \Omega_M} \max_{a \in \mathcal{A}_M} \sup_{z \in \mathcal{Z}} \log \frac{q_i^a(z)}{q_j^a(z)} \leq \xi_M.$$

Assumption 4.1.1 ensures the possibility of discrimination between any two hypotheses, hence ensuring that the problem of active hypothesis testing has a meaningful solution. Assumption 4.1.2 implies that no two hypotheses are fully distinguishable using a single observation. Assumption 4.1.2 is a technical one which enables our non-asymptotic characterizations of the upper and lower bounds. In Section 4.6, we discuss the consequence of weakening this assumption in detail.

## 4.2 Upper Bounds and Achievability Analysis

In this section, we analyze the performance of the proposed heuristics, and provide upper bounds on the expected total cost achieved by these policies. Let  $V_\pi(\boldsymbol{\rho}) := \mathbb{E}_\pi[\tau] + L\text{Pe}_\pi$  denote the expected total cost (3.1) achieved by policy  $\pi$  given that the initial belief is  $\boldsymbol{\rho}$ .

**Theorem 4.2.1.** *Consider a policy  $\pi$  that selects the retire-declare action at*

$$\tilde{\tau}_{1/L} = \min\{t : \max_{i \in \Omega_M} \rho_i(t) \geq 1 - L^{-1}\}. \quad (4.1)$$

*Suppose policy  $\pi$  at each time  $t = 0, \dots, \tilde{\tau}_{1/L} - 1$  and given the posterior vector  $\boldsymbol{\rho}(t)$ , selects sensing actions in a way that  $EJS(\boldsymbol{\rho}(t), \pi) \geq \alpha$  for some  $\alpha > 0$ . Under*

Assumptions 4.1.1 and 4.1.2, and for  $L > 1$  and  $\boldsymbol{\rho} \in \mathbb{P}_L(\Omega_M)$ ,

$$V_\pi(\boldsymbol{\rho}) \leq \bar{V}_\alpha(\boldsymbol{\rho}) := \frac{H(\boldsymbol{\rho}) + \max\{\log \log M, \log L\} + 2^{\xi_M+2}}{\alpha} + 1. \quad (4.2)$$

Furthermore, if there exist positive values  $\alpha$  and  $\beta$  such that at each time  $t < \tilde{\tau}_{1/L}$  and given the posterior vector  $\boldsymbol{\rho}(t)$ ,

$$EJS(\boldsymbol{\rho}(t), \pi) \geq \begin{cases} \alpha & \text{if } \max_{i \in \Omega_M} \rho_i(t) < \tilde{\rho} \\ \beta & \text{otherwise} \end{cases}, \quad (4.3)$$

where

$$\tilde{\rho} := 1 - \frac{1}{1 + \max\{\log M, \log L\}}, \quad (4.4)$$

then the following bound is obtained

$$V_\pi(\boldsymbol{\rho}) \leq \bar{V}_{\alpha\beta}(\boldsymbol{\rho}) := \frac{H(\boldsymbol{\rho}) + \max\{\log \log M, \log \log L\}}{\alpha} + \frac{\log L}{\beta} + \frac{3 \times 2^{2\xi_M+4}}{\alpha\beta} + 1. \quad (4.5)$$

The proof of Theorem 4.2.1 is provided in Section 4.7.2.

Recall from Section 3.3.2 that policy  $\pi_{EJS}$  selects the retire-declare action at  $\tilde{\tau}_{1/L}$  and selects sensing actions in a way to maximize the EJS divergence, i.e.,

$$EJS(\boldsymbol{\rho}, \pi_{EJS}) = \max_{a \in \mathcal{A}_M} EJS(\boldsymbol{\rho}, a). \quad (4.6)$$

Theorem 4.2.1 together with (4.6) yields the following:

**Corollary 4.2.1.** *Suppose there exist positive values  $\alpha$  and  $\beta$  such that*

$$\max_{a \in \mathcal{A}_M} EJS(\boldsymbol{\rho}, a) \geq \begin{cases} \alpha & \text{if } \max_{i \in \Omega_M} \rho_i < \tilde{\rho} \\ \beta & \text{otherwise} \end{cases}, \quad \forall \boldsymbol{\rho} \in \mathbb{P}_L(\Omega_M). \quad (4.7)$$

Under Assumptions 4.1.1 and 4.1.2, and for  $L > 1$  and  $\boldsymbol{\rho} \in \mathbb{P}_L(\Omega_M)$ ,

$$V_{\pi_{EJS}}(\boldsymbol{\rho}) \leq \frac{H(\boldsymbol{\rho}) + \max\{\log \log M, \log \log L\}}{\alpha} + \frac{\log L}{\beta} + \frac{3 \times 2^{2\xi_M+4}}{\alpha\beta} + 1. \quad (4.8)$$

In Chapters 5–7, we will extensively use Theorem 4.2.1 and Corollary 4.2.1 to analyze the performance of  $\pi_{EJS}$  and other simple heuristics for important special cases of the active hypothesis testing problem.

Next we provide an upper bound for  $V_{\pi_{EJS}}$  for the general active hypothesis testing problem by characterizing  $\alpha$  and  $\beta$  in (4.7) and using Corollary 4.2.1. For notational simplicity, let

$$I_0(M) := \max_{\lambda \in \mathbb{P}(\mathcal{A}_M)} \min_{i \in \Omega_M} \min_{\hat{\rho} \in \mathbb{P}_L(\Omega_M)} \sum_{a \in \mathcal{A}_M} \lambda_a D(q_i^a \| \sum_{j \neq i} \frac{\hat{\rho}_j}{1 - \hat{\rho}_i} q_j^a), \quad (4.9)$$

$$D_i(M) := \max_{\lambda \in \mathbb{P}(\mathcal{A}_M)} \min_{\hat{\rho} \in \mathbb{P}_L(\Omega_M)} \sum_{a \in \mathcal{A}_M} \lambda_a D(q_i^a \| \sum_{j \neq i} \frac{\hat{\rho}_j}{1 - \hat{\rho}_i} q_j^a), \quad \forall i \in \Omega_M. \quad (4.10)$$

Note that for any  $\lambda \in \mathbb{P}(\mathcal{A}_M)$  and  $\rho \in \mathbb{P}_L(\Omega_M)$ ,

$$\max_{a \in \mathcal{A}_M} EJS(\rho, a) \geq \sum_{a \in \mathcal{A}_M} \lambda_a EJS(\rho, a) \geq \min_{\hat{\rho} \in \mathbb{P}_L(\Omega_M)} \sum_{a \in \mathcal{A}_M} \lambda_a EJS(\hat{\rho}, a),$$

which implies that

$$\max_{a \in \mathcal{A}_M} EJS(\rho, a) \geq I_0(M). \quad (4.11)$$

Similarly, we can show that for any  $\lambda \in \mathbb{P}(\mathcal{A}_M)$  and  $\rho \in \mathbb{P}_L(\Omega_M)$  such that  $\rho_i \geq \tilde{\rho}$ ,

$$\max_{a \in \mathcal{A}_M} EJS(\rho, a) \geq \sum_{a \in \mathcal{A}_M} \lambda_a EJS(\rho, a) \geq \min_{\tilde{\rho} \in \mathbb{P}_L(\Omega_M)} \sum_{a \in \mathcal{A}_M} \lambda_a \tilde{\rho} D(q_i^a \| \sum_{j \neq i} \frac{\tilde{\rho}_j}{1 - \tilde{\rho}_i} q_j^a),$$

which implies that

$$\max_{a \in \mathcal{A}_M} EJS(\rho, a) \geq \tilde{\rho} D_i(M) \quad \text{if } \rho_i \geq \tilde{\rho}. \quad (4.12)$$

Corollary 4.2.1 together with (4.11) and (4.12) provides the following upper bound on the expected total cost of policy  $\pi_{EJS}$ ,

$$\begin{aligned} V_{\pi_{EJS}}(\rho) &\leq \frac{H(\rho) + \max\{\log \log M, \log \log L\}}{I_0(M)} + \frac{\log L}{\tilde{\rho} \min_{i \in \Omega_M} D_i(M)} \\ &\quad + \frac{3 \times 2^{2\xi_M+4}}{\tilde{\rho} \min_{i \in \Omega_M} D_i(M) I_0(M)} + 1. \end{aligned} \quad (4.13)$$

Next we consider policy  $\pi_2$ , a two-phase policy which was introduced and analyzed in [38]. Let  $\eta_0$  and  $\eta_i$ ,  $i \in \Omega_M$ , be vectors in  $\mathbb{P}(\mathcal{A}_M)$  that achieve the maximum in (4.9) and (4.10), respectively, i.e.,

$$\eta_0 := \arg \max_{\lambda \in \mathbb{P}(\mathcal{A}_M)} \min_{i \in \Omega_M} \min_{\hat{\rho} \in \mathbb{P}_L(\Omega_M)} \sum_{a \in \mathcal{A}_M} \lambda_a D(q_i^a \| \sum_{j \neq i} \frac{\hat{\rho}_j}{1 - \hat{\rho}_i} q_j^a), \quad (4.14)$$

$$\eta_i := \arg \max_{\lambda \in \mathbb{P}(\mathcal{A}_M)} \min_{\hat{\rho} \in \mathbb{P}_L(\Omega_M)} \sum_{a \in \mathcal{A}_M} \lambda_a D(q_i^a \| \sum_{j \neq i} \frac{\hat{\rho}_j}{1 - \hat{\rho}_i} q_j^a), \quad \forall i \in \Omega_M. \quad (4.15)$$

Moreover, let  $\eta_{0a}$  and  $\eta_{ia}$  denote elements of  $\boldsymbol{\eta}_0$  and  $\boldsymbol{\eta}_i$  corresponding to  $a \in \mathcal{A}_M$ , respectively. Consider a threshold  $\check{\rho}$ ,  $\check{\rho} > \frac{1}{2}$ . Markov (randomized) policy  $\pi_2$  is defined as follows:

- If  $\rho_i \geq 1 - L^{-1}$ , retire and select  $H_i$  as the true hypothesis;
- If  $\rho_i \in [\check{\rho}, 1 - L^{-1})$ , then  $\pi_2(a|\boldsymbol{\rho}) = \eta_{ia}$ ,  $\forall a \in \mathcal{A}_M$ ;
- If  $\rho_i < \min\{\check{\rho}, 1 - L^{-1}\}$ , for all  $i \in \Omega_M$ , then  $\pi_2(a|\boldsymbol{\rho}) = \eta_{0a}$ ,  $\forall a \in \mathcal{A}_M$ .

Next theorem provides an upper bound on the expected total cost of  $\pi_2$ :

**Theorem 4.2.2** (see [38]). *Under Assumptions 4.1.1 and 4.1.2, and for  $L > 1$  and any  $\boldsymbol{\rho} \in \mathbb{P}_L(\Omega_M)$ ,*

$$V_{\pi_2}(\boldsymbol{\rho}) \leq \frac{H(\boldsymbol{\rho}) + \xi_M + K_2''}{I_0(M)} + \sum_{i=1}^M \rho_i \frac{\log L}{D_i(M)} + 1,$$

where  $K_2''$  is a constant independent of  $L$  and  $M$ .

The proof of Theorem 4.2.2 relies on the (non-Bayesian) analysis of the conditional expected cost. Following a similar approach and for large values of  $L$  and  $M$ , the upper bound (4.13) for  $\pi_{EJS}$  can be tightened as follows:

**Proposition 4.2.1.** *Under Assumptions 4.1.1 and 4.1.2, and for  $L > 1$  and  $\boldsymbol{\rho} \in \mathbb{P}_L(\Omega_M)$ ,*

$$V_{\pi_{EJS}}(\boldsymbol{\rho}) \leq \frac{H(\boldsymbol{\rho}) + \max\{\log \log M, \log \log L\}}{I_0(M)} + \sum_{i=1}^M \rho_i \frac{\log L + \frac{K_{EJS}'' \times 2^{2\xi_M}}{I_0(M)}}{D_i(M)} + 1,$$

where  $K_{EJS}''$  is a constant independent of  $L$  and  $M$ .

Recall that in Theorem 2.3.3, an upper bound was provided for the class of sequential adaptive policies by analyzing the performance of a heuristic two-phase policy  $\pi_{SA}$  which can be described as follows:

- If  $\rho_i \geq 1 - L^{-1}$ , retire and select  $H_i$  as the true hypothesis;
- If  $\rho_i \in [\check{\rho}, 1 - L^{-1})$ , then  $\pi_{SA}(a|\boldsymbol{\rho}) = \mu_{ia}$ ,  $\forall a \in \mathcal{A}_M$ ;
- If  $\rho_i < \min\{\check{\rho}, 1 - L^{-1}\}$ , for all  $i \in \Omega_M$ , then  $\pi_{SA}(a|\boldsymbol{\rho}) = \mu_{0a}$ ,  $\forall a \in \mathcal{A}_M$ ;

where  $\boldsymbol{\mu}_0$  and  $\boldsymbol{\mu}_i$ ,  $i \in \Omega_M$ , are vectors in  $\mathbb{P}(\mathcal{A}_M)$  that achieve the maximum in (4.16) and (4.17), respectively:

$$\tilde{I}_0(M) := \max_{\boldsymbol{\lambda} \in \mathbb{P}(\mathcal{A}_M)} \min_{i \in \Omega_M} \min_{j \neq i} \sum_{a \in \mathcal{A}_M} \lambda_a D(q_i^a \| q_j^a), \quad (4.16)$$

$$\tilde{D}_i(M) := \max_{\boldsymbol{\lambda} \in \mathbb{P}(\mathcal{A}_M)} \min_{j \neq i} \sum_{a \in \mathcal{A}_M} \lambda_a D(q_i^a \| q_j^a), \quad \forall i \in \Omega_M. \quad (4.17)$$

The next proposition generalizes the upper bound in Theorem 2.3.3 and is provided as a benchmark for comparison with  $\pi_{EJS}$ .

**Proposition 4.2.2.** *Under Assumptions 4.1.1 and 4.1.2, and for  $L > 1$ ,  $\boldsymbol{\rho} \in \mathbb{P}_L(\Omega_M)$ , and arbitrary  $\iota \in (0, 1)$ ,*

$$\begin{aligned} V_{\pi_{SA}}(\boldsymbol{\rho}) \leq & \frac{H(\boldsymbol{\rho}) + \log M + K''}{\tilde{I}_0(M)} (1 + \iota) + \sum_{i=1}^M \rho_i \frac{\log L}{\tilde{D}_i(M)} (1 + \iota) \\ & + M \left( 6 + \frac{1}{\left(\frac{\iota/2}{1+\iota}\right)^5 \left(\frac{\tilde{I}_0(M)}{2\xi_M}\right)^4} \right) \left( L(1 - \max_{j \in \Omega_M} \rho_j) \right)^{-\frac{\iota^3}{(1+\iota)^2} \frac{\tilde{I}_0^2(M)}{4\xi_M^3}} + 2, \end{aligned}$$

where  $K''$  is a constant independent of  $L$  and  $M$ .

The proof of Proposition 4.2.2 is provided in Section 4.7.7.

To compare the performance of the proposed heuristics with that of the optimal policy  $\pi^*$ , we require knowledge of the optimal value function  $V^*$ . In lieu of numerical approximation of or derivation of a closed-form for  $V^*$ , in the next section, we use Lemma 3.5.1 to find lower bounds for  $V^*$ .

## 4.3 Lower Bounds for the Optimal Performance

From Lemma 3.5.1, we know that if there exists a functional  $V : \mathbb{P}(\Omega_M) \rightarrow \mathbb{R}_+$  such that for all belief vectors  $\boldsymbol{\rho} \in \mathbb{P}(\Omega_M)$ ,

$$V(\boldsymbol{\rho}) \leq \min\{1 + \min_{a \in \mathcal{A}_M} (\mathbb{T}^a V)(\boldsymbol{\rho}), \min_{j \in \Omega_M} (1 - \rho_j)L\},$$

then  $V^*(\boldsymbol{\rho}) \geq V(\boldsymbol{\rho})$  for all  $\boldsymbol{\rho} \in \mathbb{P}(\Omega_M)$ . In this section, we use the above fact to find three lower bounds for  $V^*$ . These lower bounds are non-asymptotic and complementary for various values of the parameters of the problem.

**Theorem 4.3.1.** *Under Assumption 4.1.1 and for  $L > 1$  and  $\boldsymbol{\rho} \in \mathbb{P}_L(\Omega_M)$ ,*

$$V^*(\boldsymbol{\rho}) \geq \underline{V}_1(\boldsymbol{\rho}) := \left[ \sum_{i=1}^M \rho_i \max_{j \neq i} \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{\rho_i}{\rho_j}}{\max_{a \in \mathcal{A}_M} D(q_i^a \| q_j^a)} - K'_1 \right]^+,$$

where  $K'_1$  is a constant independent of  $L$  whose closed-form is given by (4.74).

The proof of Theorem 4.3.1 is provided in Section 4.7.3.

Let define

$$D_{\max}(M) := \max_{i,j \in \Omega_M} \max_{a \in \mathcal{A}_M} D(q_i^a \| q_j^a), \quad (4.18)$$

$$I_{\max}(M) := \max_{a \in \mathcal{A}_M} \max_{\hat{\boldsymbol{\rho}} \in \mathbb{P}(\Omega_M)} JS(\hat{\boldsymbol{\rho}}, a). \quad (4.19)$$

Next we provide another lower bound which is more appropriate for large values of  $M$  or small values of  $I_{\max}(M)$ .

**Theorem 4.3.2.** *Under Assumption 4.1.1 and for  $L > 1$  and  $\boldsymbol{\rho} \in \mathbb{P}_L(\Omega_M)$ ,*

$$V^*(\boldsymbol{\rho}) \geq \underline{V}_2(\boldsymbol{\rho}) := \left[ \frac{H(\boldsymbol{\rho}) - H([\alpha(L, M), 1 - \alpha(L, M)]) - \alpha(L, M) \log(M-1)}{I_{\max}(M)} + \alpha(L, M)L \right]^+,$$

where  $\alpha(L, M) := \frac{M-1}{M-1+2^L I_{\max}(M)}$ .

The proof of Theorem 4.3.2 is provided in Section 4.7.4.

**Remark 4.3.1.** The lower bounds in Theorems 4.3.1 and 4.3.2 can be explained by the following intuition: For any uncertainty function  $V : \mathbb{P}(\Omega_M) \rightarrow \mathbb{R}$ , the number of samples required to reduce the uncertainty down to a target level  $V_{\text{target}}$  has to be at least  $\frac{V(\boldsymbol{\rho}(0)) - V_{\text{target}}}{\Delta_{\max}(V)}$ , where  $\Delta_{\max}(V)$  is the maximum amount of reduction in  $V$  associated with a single sample, i.e.,  $\Delta_{\max}(V) = \max_{a \in \mathcal{A}_M} \max_{\boldsymbol{\rho} \in \mathbb{P}(\Omega_M)} \mathcal{I}(\boldsymbol{\rho}, a, V)$ . The lower bound in Theorem 4.3.1 is associated with such a lower bound when taking  $V$  to be the log-likelihood function, while the lower bound in Theorem 4.3.2 is associated with setting  $V$  to be the Shannon entropy.

Theorem 4.3.2 can be used to show that when  $L < \frac{\log M}{I_{\max}(M)}$ , the problem of active hypothesis testing will have a trivial solution. The precise statement is given by the following corollary.

**Corollary 4.3.1.** *Let  $L < \frac{\log M}{I_{\max}(M)}$ , and suppose the decision maker has a uniform prior belief about the hypotheses. For sufficiently large  $M$ , the optimal policy randomly guesses the true hypothesis without collecting any observation, hence,  $\text{Pe}$ , the probability of making a wrong declaration, approaches  $1 - \frac{1}{M}$ .*

The proof of Corollary 4.3.1 is provided in Section 4.7.6.

Next theorem combines the above lower bounds and is appropriate when  $L$  and  $M$  are both large.

**Theorem 4.3.3.** *Under Assumptions 4.1.1 and 4.1.2, and for  $L > \max\{1, \frac{\log M}{I_{\max}(M)}\}$  and arbitrary  $\delta \in (0, 0.5]$ ,*

$$V^*(\boldsymbol{\rho}) \geq \underline{V}_3(\boldsymbol{\rho}) := \left[ \frac{H(\boldsymbol{\rho}) - H([\delta, 1 - \delta]) - \delta \log(M - 1)}{I_{\max}(M)} + \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{1-\delta}{\delta} - \xi_M}{D_{\max}(M)} \mathbf{1}_{\{\max_{i \in \Omega_M} \rho_i \leq 1-\delta\}} - K'_3 \right]^+,$$

where  $K'_3$  is a constant independent of  $\delta$  and  $L$  whose closed-form is given by (4.92).<sup>1</sup>

The proof of Theorem 4.3.3 is provided in Section 4.7.5.

Recall that in Theorem 2.3.3, following Chernoff's approach [5, Theorem 2], we provided a lower bound for the class of sequential adaptive policies. The next proposition generalizes this lower bound and is provided as a benchmark for comparison with the lower bounds derived using dynamic programming techniques.

**Proposition 4.3.1.** *Under Assumptions 4.1.1 and 4.1.2, and for  $L > 1$ ,  $\boldsymbol{\rho} \in \mathbb{P}_L(\Omega_M)$ , and arbitrary  $\delta \in (0, 1)$ ,*

$$V^*(\boldsymbol{\rho}) \geq \left[ \sum_{i=1}^M \rho_i \frac{\left[ (1 - \delta) \log \frac{L}{K' \log 2L} - \max_{j \neq i} \log \frac{\rho_i}{\rho_j} \right]^+}{\tilde{D}_i(M) + \delta} \left( 1 - \frac{2M \left( \frac{K' \log 2L}{L} \right)^\delta}{\rho_i} \right) - \frac{M \xi_M^2}{\delta^2} \right]^+,$$

where  $K'$  is a constant independent of  $\delta$  and  $L$  whose closed-form is given by (4.107).

The proof of Proposition 4.3.1 is provided in Section 4.7.8.

---

<sup>1</sup>As it will be discussed in Section 4.5.2,  $K'_3$  can be selected independent of  $M$  as well if  $\sup_M \xi_M < \infty$ .

## 4.4 Improved Stopping Rule: Numerical Example

All the policies proposed so far apply the same stopping rule  $\tilde{\tau}_{1/L}$ , i.e., they all stop sampling as soon as the belief about one of the hypotheses passes the threshold  $1 - L^{-1}$ . One shortcoming of this rule is that it does not take into account the amount of information that samples provide. In this section, we address this insensitivity to the choice of the observation kernels, and we propose a modified threshold for stopping. For simplicity of exposition, we consider the case of active binary hypothesis testing; however, the results can be extended to  $M > 2$ .

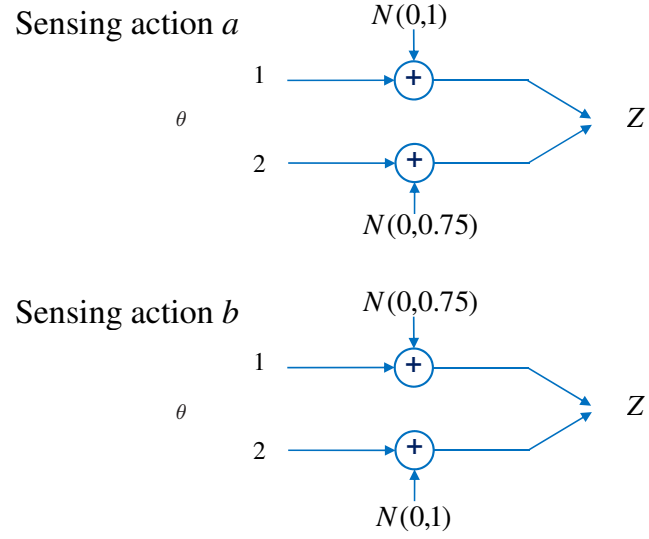
As shown in the proofs of Theorem 4.3.2 and Corollary 4.3.1, at belief vector  $\boldsymbol{\nu} = [1 - \alpha(L, 2), \alpha(L, 2)]$  where  $\alpha(L, 2) = \frac{1}{1 + 2^{LI_{\max}(2)}}$ , we have

$$\underline{V}_2(\boldsymbol{\nu}) = V^*(\boldsymbol{\nu}) = L \min\{\alpha(L, 2), 1 - \alpha(L, 2)\} = L\alpha(L, 2),$$

and hence the optimal action is to retire and declare  $H_1$  as the true hypothesis. This together with the facts that  $V^*$  is concave (see Theorem 3.2.2) and  $V^*([1, 0]) = 0$  implies that at any belief vector  $\boldsymbol{\rho} \in \mathbb{P}(\Omega_2)$  such that  $\rho_1 \geq 1 - \alpha(L, 2)$ , it is optimal to retire and declare  $H_1$  as the true hypothesis. Similar stopping threshold can be obtained for  $H_2$ . This shows that in scenarios where  $\alpha(L, 2) \gg L^{-1}$  (or equivalently  $I_{\max}(2) \ll \frac{\log(L-1)}{L}$ ), the stopping rule  $\tilde{\tau}_{1/L}$  is not a reasonable one. Similarly, we can use lower bound  $\underline{V}_1$  to find another stopping threshold. Let  $\hat{\pi}_{EJS}$  denote a modified version of  $\pi_{EJS}$  that selects the best stopping rule among the above candidates. Next we will compare the performance of  $\hat{\pi}_{EJS}$  against  $\pi_{EJS}$  as well as the optimal policy in a numerical example.

Consider the active binary hypothesis testing problem with additive Gaussian noise shown in Fig. 4.1. We have modified the variance of the noise to reduce  $I_{\max}(2)$  and  $D_{\max}(2)$ , hence the information that samples can provide.

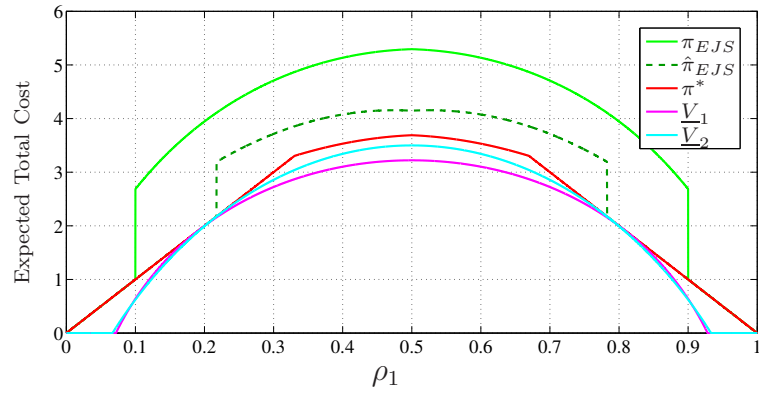
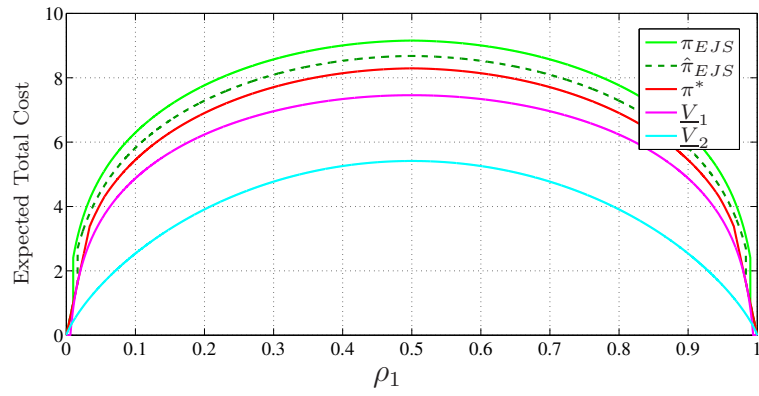
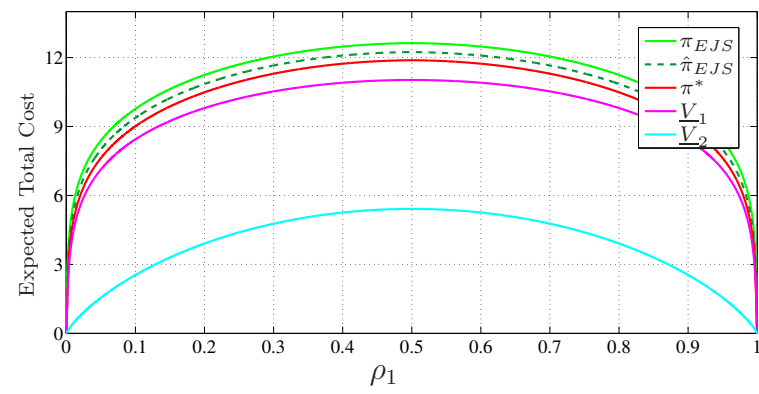
Figure 4.2 plots the expected total cost of policies  $\pi^*$ ,  $\pi_{EJS}$  and  $\hat{\pi}_{EJS}$  for different values of  $L$ . Additionally, Fig. 4.2 illustrates lower bounds  $\underline{V}_1$  and  $\underline{V}_2$  as benchmark. Table 4.1 compares the proposed stopping thresholds and the optimal one. The results confirm that  $\tilde{\tau}_{1/L}$  is a reasonable stopping rule when  $L$  is large; but it might be far from the optimal one if  $L$  is small.



**Figure 4.1:** Active binary hypothesis testing with additive Gaussian noise.

**Table 4.1:** Different stopping thresholds for the example of Fig. 4.1.

|            | $L^{-1}$ | using $\underline{V}_1$ | using $\underline{V}_2$ | optimal (using $V^*$ ) |
|------------|----------|-------------------------|-------------------------|------------------------|
| $L = 10$   | 0.1000   | 0.1930                  | 0.2175                  | 0.3304                 |
| $L = 100$  | 0.0100   | 0.0164                  | $\approx 0$             | 0.0338                 |
| $L = 1000$ | 0.0010   | 0.0015                  | $\approx 0$             | 0.0031                 |

(a)  $L = 10$ .(b)  $L = 100$ .(c)  $L = 1000$ .**Figure 4.2:** Expected total cost of the proposed policies for different values of  $L$ .

**Table 4.2:** Summary of Notations

| Notation         | Description                                                                                                                                                                                                                  |
|------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| $I_{\max}(M)$    | $\max_{a \in \mathcal{A}_M} \max_{\hat{\rho} \in \mathbb{P}(\Omega_M)} JS(\hat{\rho}, a)$                                                                                                                                    |
| $D_{\max}(M)$    | $\max_{i,j \in \Omega_M} \max_{a \in \mathcal{A}_M} D(q_i^a \  q_j^a)$                                                                                                                                                       |
| $I_0(M)$         | $\max_{\lambda \in \mathbb{P}(\mathcal{A}_M)} \min_{i \in \Omega_M} \min_{\hat{\rho} \in \mathbb{P}_L(\Omega_M)} \sum_{a \in \mathcal{A}_M} \lambda_a D(q_i^a \  \sum_{j \neq i} \frac{\hat{\rho}_j}{1-\hat{\rho}_i} q_j^a)$ |
| $D_i(M)$         | $\max_{\lambda \in \mathbb{P}(\mathcal{A}_M)} \min_{\hat{\rho} \in \mathbb{P}_L(\Omega_M)} \sum_{a \in \mathcal{A}_M} \lambda_a D(q_i^a \  \sum_{j \neq i} \frac{\hat{\rho}_j}{1-\hat{\rho}_i} q_j^a)$                       |
| $\tilde{I}_0(M)$ | $\max_{\lambda \in \mathbb{P}(\mathcal{A}_M)} \min_{i \in \Omega_M} \min_{j \neq i} \sum_{a \in \mathcal{A}_M} \lambda_a D(q_i^a \  q_j^a)$                                                                                  |
| $\tilde{D}_i(M)$ | $\max_{\lambda \in \mathbb{P}(\mathcal{A}_M)} \min_{j \neq i} \sum_{a \in \mathcal{A}_M} \lambda_a D(q_i^a \  q_j^a)$                                                                                                        |

## 4.5 Asymptotic Analysis and Consequences

In this section, we state and discuss the consequence of the bounds obtained in Sections 4.2 and 4.3 in asymptotically large  $L$  and  $M$ . Note that Table 4.2 provides a list of the notations introduced in Sections 4.2 and 4.3.

### 4.5.1 Order and Asymptotic Optimality in $L$

The lower and upper bounds provided in Sections 4.2 and 4.3 can be applied to establish the order optimality and asymptotic optimality of the proposed policies as defined below.

**Definition 4.5.1.** For fixed  $M$ , policy  $\pi$  is referred to as *order optimal* in  $L$  if for all  $\rho \in \mathbb{P}(\Omega_M)$ ,

$$\lim_{L \rightarrow \infty} \frac{V_\pi(\rho) - V^*(\rho)}{V_\pi(\rho)} < 1.$$

**Definition 4.5.2.** For fixed  $M$ , policy  $\pi$  is referred to as *asymptotically optimal* in  $L$  if for all  $\rho \in \mathbb{P}(\Omega_M)$ ,

$$\lim_{L \rightarrow \infty} \frac{V_\pi(\rho) - V^*(\rho)}{V_\pi(\rho)} = 0.$$

**Remark 4.5.1.** It is clear from the above definitions that order optimality is weaker than asymptotic optimality. If policy  $\pi$  is asymptotically optimal in  $L$ , then  $V_\pi$  and  $V^*$  will have the same dominating terms in  $L$ ; while order optimality of policy  $\pi$  only implies that dominating terms in  $V_\pi$  and  $V^*$  are the same up to a constant factor.

The following theorems establish order and asymptotic optimality of our proposed policies.

**Corollary 4.5.1.** *Under Assumptions 4.1.1 and 4.1.2, policy  $\pi_{SA}$  is asymptotically optimal in  $L$ . Furthermore, policy  $\pi_{EJS}$  attains asymptotic optimality in  $L$  if*

$$D_i(M) = \tilde{D}_i(M), \quad \forall i \in \Omega_M. \quad (4.20)$$

*Proof.* Using Proposition 4.3.1 and by setting  $\delta = (\log L)^{-\frac{1}{3}}$ , we obtain

$$V^*(\boldsymbol{\rho}) \geq \sum_{i=1}^M \rho_i \frac{\log L}{\tilde{D}_i(M)} + O\left((\log L)^{\frac{2}{3}}\right). \quad (4.21)$$

On the other hand, from Proposition 4.2.2 and by setting  $\iota = (\log L)^{-\frac{1}{4}}$ , we find

$$V_{\pi_{SA}}(\boldsymbol{\rho}) \leq \sum_{i=1}^M \rho_i \frac{\log L}{\tilde{D}_i(M)} + O\left((\log L)^{\frac{3}{4}}\right). \quad (4.22)$$

The proof of the first part of the corollary simply follows from Definition 4.5.2, inequality (4.21), and (4.22).

Similarly, the proof of the second part of the corollary follows from Definition 4.5.2, inequality (4.21), and Proposition 4.2.1.

□

## 4.5.2 Order and Asymptotic Optimality in $L$ and $M$

As mentioned in Section 4.1, one of the main drawbacks of Chernoff's asymptotic optimality notion was his neglecting the complementary role of the parameter  $M$ . In particular, the notion of asymptotic optimality in  $L$  falls short in showing the tension between using an (asymptotically) large number of samples to discriminate among a few hypotheses with (asymptotically) high accuracy or

an (asymptotically) large number of hypotheses with a lower degree of accuracy. In this section, we address this issue by analyzing the bounds when  $L$  and  $M$  are both asymptotically large. More specifically, we consider a sequence of problems indexed by the parameter  $M$  in which the set of actions and observation kernels grow monotonically as  $M$  increases, i.e., for all  $M < M'$ ,

$$\mathcal{A}_M \subseteq \mathcal{A}_{M'}, \text{ and } \{q_i^a(\cdot)\}_{i \in \Omega_M, a \in \mathcal{A}_M} \subseteq \{q_i^a(\cdot)\}_{i \in \Omega_{M'}, a \in \mathcal{A}_{M'}}. \quad (4.23)$$

Recall the notations listed in Table 4.2. Also let  $D_0(M)$  and  $\tilde{D}_0(M)$  denote respectively the harmonic mean of  $\{D_i(M)\}_{i \in \Omega_M}$  and  $\{\tilde{D}_i(M)\}_{i \in \Omega_M}$ , i.e.,

$$D_0(M) = M \left( \sum_{i=1}^M \frac{1}{D_i(M)} \right)^{-1}, \quad \tilde{D}_0(M) = M \left( \sum_{i=1}^M \frac{1}{\tilde{D}_i(M)} \right)^{-1}. \quad (4.24)$$

Moreover, let

$$\bar{I}_{\max} := \sup_M I_{\max}(M), \quad \bar{D}_{\max} := \sup_M D_{\max}(M), \quad (4.25)$$

$$\underline{I}_{\max} := \inf_M I_{\max}(M), \quad \underline{D}_{\max} := \inf_M D_{\max}(M), \quad (4.26)$$

$$\underline{I}_0 := \inf_M I_0(M), \quad \underline{D}_0 := \inf_M D_0(M). \quad (4.27)$$

By the definition and from (4.23),  $D_{\max}(M)$  and  $I_{\max}(M)$  are non-decreasing in  $M$ . Furthermore, from (4.19), (3.18), and Jensen's inequality,

$$\begin{aligned} I_{\max}(M) &= \max_{a \in \mathcal{A}_M} \max_{\hat{\rho} \in \mathbb{P}(\Omega_M)} \sum_{i=1}^M \hat{\rho}_i D(q_i^a \| \sum_{j=1}^M \hat{\rho}_j q_j^a) \\ &\leq \max_{a \in \mathcal{A}_M} \max_{\hat{\rho} \in \mathbb{P}(\Omega_M)} \sum_{i=1}^M \hat{\rho}_i \sum_{j=1}^M \hat{\rho}_j D(q_i^a \| q_j^a) \\ &\leq \max_{a \in \mathcal{A}_M} \max_{i,j \in \Omega_M} D(q_i^a \| q_j^a) = D_{\max}(M); \end{aligned} \quad (4.28)$$

and by Assumption 4.1.2, we have

$$D_{\max}(M) \leq \max_{i,j \in \Omega_M} \max_{a \in \mathcal{A}_M} \log \frac{q_i^a}{q_j^a} \leq \xi_M. \quad (4.29)$$

Similarly,  $I_0(M) \leq D_i(M) \leq \tilde{D}_i(M) \leq D_{\max}(M) \leq \xi_M$ ,  $\forall i \in \Omega_M$ , for all  $M$ . Since  $D_{\max}(M)$  and  $I_{\max}(M)$  are non-decreasing in  $M$ , we have  $\underline{D}_{\max} = D_{\max}(2)$ ,  $\bar{D}_{\max} = \lim_{M \rightarrow \infty} D_{\max}(M)$ ,  $\underline{I}_{\max} = I_{\max}(2)$ , and  $\bar{I}_{\max} = \lim_{M \rightarrow \infty} I_{\max}(M)$ .

Furthermore, to ensure that the distance between the observation kernels remains bounded as  $M$  increases (and  $\overline{D}_{\max} < \infty$ ), we consider the following assumption:

**Assumption 4.5.1.** *There exists  $\xi < \infty$  such that*

$$\sup_M \xi_M \leq \xi.$$

This assumption allows us to specialize Theorem 4.3.3 as follows.

**Corollary 4.5.2.** *Let  $\boldsymbol{\rho}_{u,M}$  denote a uniform prior on the set of hypotheses  $\Omega_M$ . Under Assumptions 4.1.1, 4.1.2, and 4.5.1, and for  $L > \max\{2, \frac{\log M}{I_{\max}(M)}\}$ ,*

$$\underline{V}_3(\boldsymbol{\rho}_{u,M}) \geq \left[ \frac{\log M - 2}{\overline{I}_{\max}} + \frac{\log \frac{1-L^{-1}}{L^{-1}}}{\overline{D}_{\max}} - \frac{\log \log LM + \xi}{\underline{D}_{\max}} - K'_3 \right]^+,$$

where  $K'_3$  is a positive constant independent of  $L$  and  $M$ .

The proof of Corollary 4.5.2 is provided in Section 4.7.6.

The next definition extends the notions of order and asymptotic optimality defined in Section 4.5.1 to the case where  $M$  increases as well.

**Definition 4.5.3.** Policy  $\pi$  is referred to as *order optimal* and *asymptotically optimal* in  $L$  and  $M$  if respectively,<sup>2</sup>

$$\lim_{L,M \rightarrow \infty} \frac{V_\pi(\boldsymbol{\rho}_{u,M}) - V^*(\boldsymbol{\rho}_{u,M})}{V_\pi(\boldsymbol{\rho}_{u,M})} < 1, \quad \lim_{L,M \rightarrow \infty} \frac{V_\pi(\boldsymbol{\rho}_{u,M}) - V^*(\boldsymbol{\rho}_{u,M})}{V_\pi(\boldsymbol{\rho}_{u,M})} = 0.$$

**Corollary 4.5.3.** *Under Assumptions 4.1.1, 4.1.2, and 4.5.1, for  $L > \frac{\log M}{I_{\max}(M)}$ , and if  $\underline{I}_0 > 0$ , policy  $\pi_{EJS}$  is order optimal in  $L$  and  $M$ . Furthermore, if  $\overline{I}_{\max} = \underline{I}_0$  and  $\overline{D}_{\max} = \underline{D}_0$ , policy  $\pi_{EJS}$  is asymptotically optimal in  $L$  and  $M$ .*

*Proof.* The proof simply follows from Definition 4.5.3, Proposition 4.2.1, and Corollary 4.5.2.  $\square$

---

<sup>2</sup>Note that unlike Definitions 4.5.1 and 4.5.2 where we considered the performance gap between policy  $\pi$  and the optimal policy  $\pi^*$  for all values of  $\boldsymbol{\rho} \in \mathbb{P}(\Omega_M)$ , here we consider the performance gap specifically at the uniform vector in the information state space.

### 4.5.3 Information Acquisition Rate and Reliability

In this section, we explain the primal (constrained) version of the problem of active hypothesis testing, referred to as the information acquisition problem, and use the obtained bounds in Sections 4.2 and 4.3 to extend the (information theoretic) notions of achievable communication rate and error exponent to the context of active sequential hypothesis testing.

#### Information Acquisition Problem:

Consider a sequence of active hypothesis testing problems indexed by  $M$  hypotheses of interest, action space  $\mathcal{A}_M$ , and observation kernels  $\{q_i^a\}_{i \in \Omega_M, a \in \mathcal{A}_M}$ : A Bayesian decision maker with uniform prior belief  $\boldsymbol{\rho}(0) = \boldsymbol{\rho}_{u,M}$  is responsible to find the true hypothesis with the objective to

$$\text{minimize } \mathbb{E}[\tau] \text{ subject to } \text{Pe} \leq \epsilon, \quad (4.30)$$

where  $\tau$  is the stopping time at which the decision maker retires,  $\text{Pe}$  is the probability of making a wrong declaration, and  $\epsilon > 0$  denotes the desired probability of error. Furthermore, let the set of actions and observation kernels grow monotonically as  $M$  increases, i.e., for all  $M < M'$

$$\mathcal{A}_M \subseteq \mathcal{A}_{M'}, \text{ and } \{q_i^a(\cdot)\}_{i \in \Omega_M, a \in \mathcal{A}_M} \subseteq \{q_i^a(\cdot)\}_{i \in \Omega_{M'}, a \in \mathcal{A}_{M'}}. \quad (4.31)$$

Let  $\mathbb{E}_\pi[\tau]$  and  $\text{Pe}_\pi$  denote respectively the expected stopping time (or equivalently the expected number of collected samples) and the probability of error under policy  $\pi$ . Following the notations in [56], we define  $M_\pi(t, \epsilon)$  as the maximum number of hypotheses among which policy  $\pi$  can find the true hypothesis with  $\mathbb{E}_\pi[\tau] \leq t$  and  $\text{Pe}_\pi \leq \epsilon$ . Policy  $\pi$  is said to achieve information acquisition rate  $R > 0$  with reliability (also known as error exponent)  $E > 0$  if

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log M_\pi(t, 2^{-Et}) = R. \quad (4.32)$$

As mentioned in Section 2.4.2, for fixed number of hypotheses  $M$ , hence at information acquisition rate  $R = 0$ , policy  $\pi$  is said to achieve reliability  $E > 0$

if

$$\lim_{t \rightarrow \infty} \frac{-1}{t} \log \text{Pe}_\pi(t, M) = E, \quad (4.33)$$

where  $\text{Pe}_\pi(t, M)$  is the minimum probability of error that policy  $\pi$  can guarantee for  $M$  hypotheses with the constraint  $\mathbb{E}_\pi[\tau] \leq t$ .

The reliability function  $E(R)$  is defined as the maximum achievable error exponent at information acquisition rate  $R$ .

Before we proceed with the upper and lower bounds on the maximum achievable information acquisition rate and the optimal reliability function, we refer the reader to Table 4.2 for the list of notations introduced in Sections 4.2 and 4.3. Also recall that  $D_0(M)$  and  $\tilde{D}_0(M)$  denote respectively the harmonic mean of  $\{D_i(M)\}_{i \in \Omega_M}$  and  $\{\tilde{D}_i(M)\}_{i \in \Omega_M}$ .

**Corollary 4.5.4.** *For any given fixed  $M$  (rate  $R = 0$ ), no policy can achieve reliability higher than  $\tilde{D}_0(M)$ . Also, no policy can achieve positive reliability  $E > 0$  at rates higher than  $\bar{I}_{\max}$ . Furthermore,*

$$E(R) \leq \bar{D}_{\max} \left( 1 - \frac{R}{\bar{I}_{\max}} \right), \quad R \in (0, \bar{I}_{\max}). \quad (4.34)$$

**Remark 4.5.2.** Corollary 4.5.4 establishes an upper bound,  $\bar{I}_{\max}$ , on maximum achievable information acquisition rate. As shown in Section 4.7.9, this result can be strengthened to show that no policy can achieve diminishing error probability at rates higher than  $\bar{I}_{\max}$ .

**Corollary 4.5.5.** *For fixed  $M$ , hence at rate  $R = 0$ , a policy  $\pi$  can achieve the maximum reliability, i.e.,  $E = \tilde{D}_0(M)$ , if and only if it is asymptotically optimal in  $L$ . Furthermore, a policy  $\pi$  can achieve a non-zero rate  $R > 0$  with non-zero reliability  $E > 0$  only if it is order optimal in  $L$  and  $M$ .*

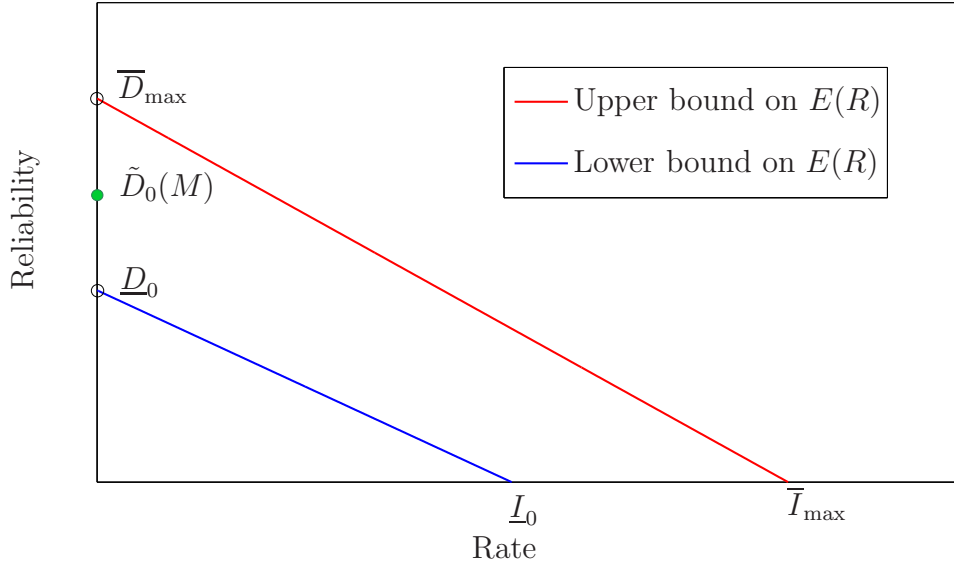
Corollary 4.5.5 implies that for fixed  $M$ , hence at  $R = 0$ , policies  $\pi_{SA}$  and  $\pi^*$  achieve the optimal error exponent; while policy  $\pi_{EJS}$  might or might not (depending on condition (4.20)). Furthermore, Corollary 4.5.5, in effect, underlines the deficiency of characterizing the solution to the problem of active hypothesis testing in terms of  $L$  in isolation from  $M$ , i.e., Chernoff's notion of asymptotic

optimality (solely in  $L$ ). In particular, an order optimal policy can achieve non-zero rate and reliability simultaneously, which is an improvement over  $\pi_{SA}$  (and all extensions of [5]).

**Corollary 4.5.6.** *Policy  $\pi_{EJS}$  achieves rate  $R \in [0, \underline{L}_0]$  with reliability  $E$  if*

$$E \leq \underline{D}_0 \left( 1 - \frac{R}{\underline{L}_0} \right). \quad (4.35)$$

Fig. 4.3 summarizes the above results. The upper bound on the reliability function is shown in red. Policy  $\pi_{SA}$  achieves the optimal reliability  $\tilde{D}_0(M)$  for fixed  $M$  (at  $R = 0$ ) with no provable guarantee for  $R > 0$  (this point is shown in green); while policy  $\pi_{EJS}$  ensures an exponentially decaying error probability (the error exponent is shown in blue) for  $R \in [0, \underline{L}_0]$ .



**Figure 4.3:** Lower and upper bounds on the optimal reliability function  $E(R)$ .

**Remark 4.5.3.** It can be shown that optimal policy  $\pi^*$  for the problem of active hypothesis testing also achieves any rate  $R \in [0, \underline{L}_0]$  with reliability  $E$  satisfying (4.35) for the information acquisition problem. Moreover, any policy  $\pi$  that satisfies (4.3) can achieve rate  $R \in [0, \alpha]$  with reliability  $E$  if  $E \leq \beta \left( 1 - \frac{R}{\alpha} \right)$ .<sup>3</sup>

<sup>3</sup>In case parameters  $\alpha$  and  $\beta$  are dependent on  $M$ , they are replaced by  $\underline{\alpha} := \liminf_{M \rightarrow \infty} \alpha$  and  $\underline{\beta} := \liminf_{M \rightarrow \infty} \beta$ , respectively.

The proofs of all the results in this section are provided in Section 4.7.9 and are based on the fact that active hypothesis testing can be viewed as a Lagrangian relaxation of the information acquisition problem. It is somewhat intuitive that as  $L \rightarrow \infty$  the solution of active hypothesis testing is closely related to that of information acquisition problem when  $\epsilon \rightarrow 0$ . The following lemma makes this intuition precise.

**Lemma 4.5.1.** *Let  $\mathbb{E}[\tau_\epsilon^*]$  denote the minimum expected number of samples required to achieve  $\text{Pe} \leq \epsilon$ . We have*

$$\mathbb{E}[\tau_\epsilon^*] \geq (1 - \epsilon L) (V^*(\boldsymbol{\rho}(0)) - 1), \quad (4.36)$$

where  $V^*$  is the optimal value function of the problem of active hypothesis testing with penalty of wrong declaration  $L$ .

Given the above connection, Corollary 4.5.4 follows readily from the lower bounds obtained in Proposition 4.3.1 and Theorem 4.3.3 (in particular its Corollary 4.5.2); and Corollaries 4.5.5 and 4.5.6 follow from the upper bounds given by Propositions 4.2.1 and 4.2.2.

## 4.6 Discussion: Technical Assumptions

In this section we provide a discussion on our technical assumptions. In particular, we discuss the necessity of Assumptions 4.1.1 and 4.1.2, and compare them with the common assumptions in the literature. In contrast to Assumption 4.1.1, which is shown to be necessary for the problem of active hypothesis testing to have a meaningful solution, Assumption 4.1.2 can be relaxed to more general assumptions without affecting the asymptotic results.

### 4.6.1 Assumption 4.1.1

We first discuss the necessity of Assumption 4.1.1. If Assumption 4.1.1 does not hold, then there exist two hypotheses  $i, j \in \Omega_M$ ,  $i \neq j$ , such that for all  $a \in \mathcal{A}_M$ ,  $D(q_i^a \| q_j^a) = 0$ . In other words,  $q_i^a(\cdot) = q_j^a(\cdot)$  for all  $a \in \mathcal{A}_M$ , and hence

the decision maker is not capable of distinguishing these two hypotheses. In this sense, Assumption 4.1.1 is necessary for the problem of active hypothesis testing to be meaningful.

Next we compare Assumption 4.1.1 to its counterpart in [5]:

**Assumption 4.6.1.**  $D(q_i^a \| q_j^a) > 0, \forall i, j \in \Omega_M, i \neq j, \forall a \in \mathcal{A}_M$ .

This assumption assures *consistency* [5, Lemma 1], i.e.,  $\arg \max_{i \in \Omega_M} \rho_i(t)$  converges exponentially fast to the true hypothesis regardless of the way the sensing actions are selected. However, this assumption is very restrictive and does not hold in many problems of interest such as channel coding with feedback [12]. It was remarked in [5, Section 7] that the above restrictive assumption can be relaxed if the proposed scheme is modified to take a (possibly randomized) action capable of discriminating between all hypotheses pairs infinitely often (for example at any time  $t$  when  $t$  is a perfect square). In this dissertation, however, we took a different approach and constructed policy  $\pi_{SA}$ , a simple two-phase modification of Chernoff's original scheme in which testing for the maximum likely hypothesis is delayed and contingent on obtaining a certain level of confidence.

## 4.6.2 Assumption 4.1.2

We first discuss the necessity of Assumption 4.1.2. For observation kernels with bounded support, Assumption 4.1.2 is a necessary condition to ensure that the observation kernels are absolutely continuous with respect to each other and hence no observation is noise-free. Although this assumption might hold in many settings, it does not hold in general for observation kernels with unbounded support. Next we replace Assumption 4.1.2 by more general assumptions on the observation kernels, and discuss the consequences.

To the best of our knowledge, Assumption 4.6.2 below, given first by [5], is the weakest condition in the literature of hypothesis testing and sequential analysis, and is often interpreted to an assumption which limits the *excess over the boundary* at the stopping time [57].

**Assumption 4.6.2.** *There exists  $\xi_M < \infty$  such that*

$$\max_{i,j \in \Omega_M} \max_{a \in \mathcal{A}_M} \int_{\mathcal{Z}} q_i^a(z) \left| \log \frac{q_i^a(z)}{q_j^a(z)} \right|^2 dz \leq \xi_M. \quad (4.37)$$

The proof of Proposition 4.2.2 relies on Chernoff's approach [5], and the asymptotic behavior of the bound remains intact if Assumption 4.1.2 is replaced with Assumption 4.6.2. However, as shown in the proof of this proposition in Section 4.7.7, Assumption 4.1.2 allows us to give a precise non-asymptotic characterization of the bound by applying the *method of bounded differences* and in particular the McDiarmid's inequality [48]. Furthermore, Proposition 4.3.1 remains valid even if Assumption 4.1.2 is replaced with Assumption 4.6.2 (with the only minor change that  $\xi_M^2$  is replaced with  $\xi_M$  in the bound).

Next we consider the consequence of weakening Assumption 4.1.2 on Theorems 4.2.1 and 4.3.3. To do so, we consider an even weaker assumption than Assumption 4.6.2 as given below:

**Assumption 4.6.3.** *There exists  $\xi_M < \infty$  and  $\gamma > 0$  such that*

$$\max_{i,j \in \Omega_M} \max_{a \in \mathcal{A}_M} \int_{\mathcal{Z}} q_i^a(z) \left| \log \frac{q_i^a(z)}{q_j^a(z)} \right|^{1+\gamma} dz \leq \xi_M.$$

Define function  $\psi_M : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  as follows

$$\psi_M(b) := \max_{i,j \in \Omega_M} \max_{a \in \mathcal{A}_M} \int_{\mathcal{Z}} q_i^a(z) \left[ \log \frac{q_i^a(z)}{q_j^a(z)} \right]_b dz,$$

where  $[g]_b = g \mathbf{1}_{\{g > b\}}$ . Note that  $\psi_M(b)$  is in general non-increasing in  $b$ , and if Assumption 4.6.3 holds,  $\psi_M(b) \leq b^{-\gamma} \xi_M$ . Under the weaker Assumption 4.6.3 (and naturally Assumption 4.6.2), Theorem 4.3.3 can be replaced by:

**Proposition 4.6.1.** *Under Assumptions 4.1.1 and 4.6.3 and for  $L > \frac{\log M}{I_{\max}(M)}$ ,  $\boldsymbol{\rho} \in \mathbb{P}_L(\Omega_M)$ ,  $\delta \in (0, 0.5]$ , and  $b > 0$ ,*

$$V^*(\boldsymbol{\rho}) \geq \underline{V}_3'(\boldsymbol{\rho}) := \frac{1}{1 + \frac{\psi_M(b)}{D_{\max}(M)}} \left[ \frac{H(\boldsymbol{\rho}) - H([\delta, 1 - \delta]) - \delta \log(M - 1)}{I_{\max}(M)} + \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{1-\delta}{\delta} - b}{D_{\max}(M)} \mathbf{1}_{\{\max_{i \in \Omega_M} \rho_i \leq 1-\delta\}} - K_3' \right]^+,$$

where  $K_3'$  is a positive constant independent of  $L$ . In addition, if Assumption 4.5.1 holds as well, then  $K_3'$  can be selected independent of  $M$  as well.

The proof of Proposition 4.6.1 is provided in Section 4.7.10.

We can also obtain the following upper bound:

**Proposition 4.6.2.** *Consider a policy  $\pi$  with stopping rule  $\tilde{\tau}_{1/L}$ . Suppose policy  $\pi$  at each time  $t = 0, \dots, \tilde{\tau}_{1/L} - 1$  and given the posterior vector  $\boldsymbol{\rho}(t)$ , selects sensing actions in a way that  $EJS(\boldsymbol{\rho}(t), \pi) \geq \alpha$  for some  $\alpha > 0$ . Under Assumptions 4.1.1 and 4.6.3, and for  $L > 1$  and  $\boldsymbol{\rho} \in \mathbb{P}_L(\Omega_M)$ , there exists  $b' \in (0, \infty)$  such that  $\frac{b}{b-3}\psi_M(b-3) < \alpha$  for all  $b \geq b'$  and*

$$V_\pi(\boldsymbol{\rho}) \leq \bar{V}'_\alpha(\boldsymbol{\rho}) := \frac{H(\boldsymbol{\rho}) + \max\{\log \log M, \log L\} + b}{\alpha - \frac{b}{b-3}\psi_M(b-3)} + 1. \quad (4.38)$$

The proof of Proposition 4.6.2 is provided in Section 4.7.11. Following a similar approach, we can also generalize  $\bar{V}_{\alpha\beta}$  given in (4.5).

As we discussed,  $\psi_M(b) \leq b^{-\gamma}\xi_M$  under Assumption 4.6.3. Furthermore, if Assumption 4.5.1 holds, then  $\sup_M \psi_M(b) \leq b^{-\gamma}\xi$ . In other words, we can select  $b$  as a function of  $L$  and  $M$  (for instance  $b = \log \log LM$ ) such that  $\underline{V}'_3$  and  $\bar{V}'_\alpha$  have the same dominating terms (in  $L$  and  $M$ ) as  $\underline{V}_3$  and  $\bar{V}_\alpha$ , respectively.

In summary, the asymptotic results of this dissertation hold under the weaker Assumptions 4.6.2 and 4.6.3 replacing Assumption 4.1.2. Our choice to present the work under Assumption 4.1.2, however, significantly simplifies the presentation, and also enables a precise non-asymptotic characterization of the lower and upper bounds.

## 4.7 Proofs

### 4.7.1 Preliminary Lemmas

In this section, we provide some preliminary lemmas. These lemmas are technical and only helpful in proving the main results.

**Lemma 4.7.1.** *Under Assumption 4.1.2 and for any  $i \in \Omega_M$ ,*

$$\left| \log \frac{\rho_i(t+1)}{1 - \rho_i(t+1)} - \log \frac{\rho_i(t)}{1 - \rho_i(t)} \right| \leq \xi_M.$$

*Proof.*

$$\begin{aligned}
\left| \log \frac{\rho_i(t+1)}{1-\rho_i(t+1)} - \log \frac{\rho_i(t)}{1-\rho_i(t)} \right| &= \left| \log \frac{\rho_i(t) q_i^{A(t)}(Z(t))}{\sum_{j \neq i} \rho_j(t) q_j^{A(t)}(Z(t))} - \log \frac{\rho_i(t)}{1-\rho_i(t)} \right| \\
&= \left| \log \frac{q_i^{A(t)}(Z(t))}{\sum_{j \neq i} \frac{\rho_j(t)}{1-\rho_i(t)} q_j^{A(t)}(Z(t))} \right| \\
&\leq \max_{a \in \mathcal{A}_M} \sup_{z \in \mathcal{Z}} \log \frac{q_i^a(z)}{\min_{j \neq i} q_j^a(z)} \leq \xi_M.
\end{aligned}$$

□

**Lemma 4.7.2.** *Under Assumption 4.1.2 and for any  $i \in \Omega_M$ ,*

$$|\rho_i(t+1) - \rho_i(t)| \leq \min \{ \rho_i(t)(1 - \rho_i(t)), \rho_i(t+1)(1 - \rho_i(t+1)) \} 2^{\xi_M}.$$

*Proof.*

$$\begin{aligned}
|\rho_i(t+1) - \rho_i(t)| &= \rho_i(t) \left| \frac{q_i^{A(t)}(Z(t))}{\sum_{j=1}^M \rho_j(t) q_j^{A(t)}(Z(t))} - 1 \right| \\
&= \rho_i(t) \left| \frac{(1 - \rho_i(t)) q_i^{A(t)}(Z(t)) - \sum_{j \neq i} \rho_j(t) q_j^{A(t)}(Z(t))}{\sum_{j=1}^M \rho_j(t) q_j^{A(t)}(Z(t))} \right| \\
&\leq \rho_i(t)(1 - \rho_i(t)) \left( \max_{k, j \in \Omega_M} \max_{a \in \mathcal{A}_M} \sup_{z \in \mathcal{Z}} \frac{q_k^a(z)}{q_j^a(z)} \right) \\
&\leq \rho_i(t)(1 - \rho_i(t)) 2^{\xi_M}.
\end{aligned}$$

Similarly, we can show that

$$\begin{aligned}
|\rho_i(t+1) - \rho_i(t)| &= \rho_i(t+1) \left| 1 - \rho_i(t) - \frac{\sum_{j \neq i} \rho_j(t) q_j^{A(t)}(Z(t))}{q_i^{A(t)}(Z(t))} \right| \\
&= \rho_i(t+1)(1 - \rho_i(t+1)) \left| \frac{1 - \rho_i(t)}{1 - \rho_i(t+1)} - \frac{\sum_{j \neq i} \rho_j(t) q_j^{A(t)}(Z(t))}{(1 - \rho_i(t+1)) q_i^{A(t)}(Z(t))} \right|
\end{aligned}$$

$$\begin{aligned}
&= \rho_i(t+1)(1 - \rho_i(t+1)) \left| \frac{\sum_{j=1}^M \rho_j(t) q_j^{A(t)}(Z(t))}{\sum_{j \neq i} \frac{\rho_j(t)}{1 - \rho_i(t)} q_j^{A(t)}(Z(t))} - \frac{\sum_{j=1}^M \rho_j(t) q_j^{A(t)}(Z(t))}{q_i^{A(t)}(Z(t))} \right| \\
&\leq \rho_i(t+1)(1 - \rho_i(t+1)) \left( \max_{k,j \in \Omega_M} \max_{a \in \mathcal{A}_M} \sup_{z \in \mathcal{Z}} \frac{q_k^a(z)}{q_j^a(z)} \right) \\
&\leq \rho_i(t+1)(1 - \rho_i(t+1)) 2^{\xi_M}.
\end{aligned}$$

□

**Lemma 4.7.3.** *Under Assumption 4.1.2 and for any  $\delta \in (0, 1)$ , if*

$$\max_{i \in \Omega_M} \max \{\rho_i(t), \rho_i(t+1)\} \geq 1 - \delta,$$

*then*

$$|U(\boldsymbol{\rho}(t+1)) - U(\boldsymbol{\rho}(t))| \leq 2^{\xi_M} (3 + \delta \log(M-1)).$$

*Proof.* We first consider the case  $\max_{i \in \Omega_M} \rho_i(t) \geq 1 - \delta$ . Without loss of generality assume  $\hat{i} = \arg \max_{i \in \Omega_M} \rho_i(t)$  and  $\rho_i(t) \geq 1 - \delta$ . We obtain

$$\begin{aligned}
&|U(\boldsymbol{\rho}(t+1)) - U(\boldsymbol{\rho}(t))| \\
&= \left| \sum_{i=1}^M \rho_i(t+1) \log \frac{\rho_i(t+1)}{1 - \rho_i(t+1)} - \sum_{i=1}^M \rho_i(t) \log \frac{\rho_i(t)}{1 - \rho_i(t)} \right| \\
&= \left| \sum_{i=1}^M \rho_i(t+1) \left( \log \frac{\rho_i(t+1)}{1 - \rho_i(t+1)} - \log \frac{\rho_i(t)}{1 - \rho_i(t)} \right) \right. \\
&\quad \left. + \sum_{i=1}^M (\rho_i(t+1) - \rho_i(t)) \log \frac{\rho_i(t)}{1 - \rho_i(t)} \right| \\
&\leq \max_{i \in \Omega_M} \left| \log \frac{\rho_i(t+1)}{1 - \rho_i(t+1)} - \log \frac{\rho_i(t)}{1 - \rho_i(t)} \right| + \left| \sum_{i=1}^M (\rho_i(t+1) - \rho_i(t)) \log \frac{\rho_i(t)}{1 - \rho_i(t)} \right| \\
&\stackrel{(a)}{\leq} \xi_M + \sum_{i=1}^M |\rho_i(t+1) - \rho_i(t)| \cdot \left| \log \frac{\rho_i(t)}{1 - \rho_i(t)} \right| \\
&\stackrel{(b)}{\leq} \xi_M + 2^{\xi_M} \sum_{i=1}^M \rho_i(t)(1 - \rho_i(t)) \left| \log \frac{\rho_i(t)}{1 - \rho_i(t)} \right| \\
&\leq \xi_M + 2^{\xi_M} \rho_{\hat{i}}(t)(1 - \rho_{\hat{i}}(t)) \left| \log \frac{\rho_{\hat{i}}(t)}{1 - \rho_{\hat{i}}(t)} \right| + 2^{\xi_M} \sum_{i \neq \hat{i}} \rho_i(t) \log \frac{1}{\rho_i(t)}
\end{aligned}$$

$$\begin{aligned}
& \stackrel{(c)}{\leq} \xi_M + 2^{\xi_M} + 2^{\xi_M} \left( \sum_{i \neq \hat{i}} \rho_i(t) \right) \log \frac{M-1}{\sum_{i \neq \hat{i}} \rho_i(t)} \\
& \leq \xi_M + 2^{\xi_M} + 2^{\xi_M} (\delta \log(M-1) + 1) \\
& \stackrel{(d)}{\leq} 2^{\xi_M} (3 + \delta \log(M-1)),
\end{aligned}$$

where (a) and (b) follow respectively from Lemmas 4.7.1 and 4.7.2; (c) follows from Jensen's inequality and the fact that

$$x(1-x) \left| \log \frac{x}{1-x} \right| \leq 1, \quad x \in [0, 1];$$

and (d) holds since  $\xi_M \leq 2^{\xi_M}$ .

This completes the proof for the case  $\max_{i \in \Omega_M} \rho_i(t) \geq 1 - \delta$ . The proof for the case  $\max_{i \in \Omega_M} \rho_i(t+1) \geq 1 - \delta$  is done by following the similar lines and interchanging time indices  $(t)$  and  $(t+1)$ .  $\square$

**Lemma 4.7.4.** *Assume that the sequence  $\{\zeta(t)\}$ ,  $t = 0, 1, 2, \dots$  forms a submartingale with respect to a filtration  $\{\mathcal{F}(t)\}$ . Furthermore, assume there exist positive constants  $K_1$ ,  $K_2$ , and  $K_3$  such that<sup>4</sup>*

$$\mathbb{E}[\zeta(t+1)|\mathcal{F}(t)] \geq \zeta(t) + K_1 \text{ if } \zeta(t) < 0, \quad (4.39a)$$

$$\mathbb{E}[\zeta(t+1)|\mathcal{F}(t)] \geq \zeta(t) + K_2 \text{ if } \zeta(t) \geq 0, \quad (4.39b)$$

$$|\zeta(t+1) - \zeta(t)| \leq K_3 \text{ if } \max\{\zeta(t+1), \zeta(t)\} \geq 0. \quad (4.39c)$$

Consider the stopping time  $v = \min\{t : \zeta(t) \geq B\}$ ,  $B > 0$ . Then we have

$$\mathbb{E}[v] \leq \frac{B - \zeta(0)}{K_2} + \zeta(0) \mathbf{1}_{\{\zeta(0) < 0\}} \left( \frac{1}{K_2} - \frac{1}{K_1} \right) + \frac{3K_3^2}{K_1 K_2}. \quad (4.40)$$

Furthermore, if  $K_1 \geq K_2$ , and there exists  $K_4 \leq K_3$  such that

$$|\zeta(t+1) - \zeta(t)| \leq K_4 \text{ if } \max\{\zeta(t+1), \zeta(t)\} \geq B, \quad (4.41)$$

the bound is tightened as

$$\mathbb{E}[v] \leq \frac{B - \zeta(0)}{K_2} + \zeta(0) \mathbf{1}_{\{\zeta(0) < 0\}} \left( \frac{1}{K_2} - \frac{1}{K_1} \right) + \frac{K_4}{K_2}. \quad (4.42)$$

---

<sup>4</sup>It is implied from (4.39) that  $K_1, K_2 \leq K_3$ .

*Proof.* This lemma is a generalization of Lemma 1 in [58]. The proof is provided below.

Consider the sequence  $\{\eta(t)\}$  defined as follows

$$\eta(t) = \begin{cases} -A + \frac{\zeta(t)}{K_1} - t & \text{if } \zeta(t) < 0 \\ -Ae^{-\alpha\zeta(t)} + \frac{\zeta(t)}{K_2} - t & \text{if } \zeta(t) \geq 0 \end{cases},$$

where  $A = \left[ \frac{3K_3^2}{K_2} \left( \frac{1}{K_1} - \frac{1}{K_2} \right) \right]^+$  and  $\alpha = \frac{0.5K_2}{K_3^2}$ .

**Claim 4.7.1.** The sequence  $\{\eta(t)\}$  forms a submartingale with respect to the filtration  $\{\mathcal{F}(t)\}$ .

By Doob's Stopping Theorem,

$$\begin{aligned} \eta(0) &\leq \mathbb{E}[\eta(v)] \\ &\leq \mathbb{E} \left[ \frac{\zeta(v)}{K_2} - v \right] \\ &= \frac{\mathbb{E}[\zeta(v-1)] + \mathbb{E}[\zeta(v) - \zeta(v-1)]}{K_2} - \mathbb{E}[v] \\ &\stackrel{(a)}{<} \frac{B + K_3}{K_2} - \mathbb{E}[v], \end{aligned} \tag{4.43}$$

where (a) follows from (4.39c) and the facts that  $\zeta(v-1) < B$  and  $\zeta(v) \geq B > 0$ .

On the other hand, we have

$$\begin{aligned} \eta(0) &= \left( -A + \frac{\zeta(0)}{K_1} \right) \mathbf{1}_{\{\zeta(0) < 0\}} + \left( -Ae^{-\alpha\zeta(0)} + \frac{\zeta(0)}{K_2} \right) \mathbf{1}_{\{\zeta(0) \geq 0\}} \\ &\geq -A + \frac{\zeta(0)}{K_2} - \zeta(0) \mathbf{1}_{\{\zeta(0) < 0\}} \left( \frac{1}{K_2} - \frac{1}{K_1} \right). \end{aligned} \tag{4.44}$$

Combining inequalities (4.43) and (4.44), we obtain

$$\begin{aligned} \mathbb{E}[v] &\leq \frac{B + K_3}{K_2} - \eta(0) \\ &\leq \frac{B + K_3}{K_2} + A - \frac{\zeta(0)}{K_2} + \zeta(0) \mathbf{1}_{\{\zeta(0) < 0\}} \left( \frac{1}{K_2} - \frac{1}{K_1} \right) \\ &\leq \frac{B - \zeta(0)}{K_2} + \zeta(0) \mathbf{1}_{\{\zeta(0) < 0\}} \left( \frac{1}{K_2} - \frac{1}{K_1} \right) + \left[ \frac{3K_3^2}{K_2} \left( \frac{1}{K_1} - \frac{1}{K_2} \right) \right]^+ + \frac{K_3}{K_2} \\ &\stackrel{(a)}{\leq} \frac{B - \zeta(0)}{K_2} + \zeta(0) \mathbf{1}_{\{\zeta(0) < 0\}} \left( \frac{1}{K_2} - \frac{1}{K_1} \right) + \frac{3K_3^2}{K_1 K_2}, \end{aligned} \tag{4.45}$$

where (a) holds since by definition  $K_1, K_2 \leq K_3$  and hence,  $\frac{K_3}{K_2} \leq \min \left\{ \frac{3K_3^2}{K_1 K_2}, \frac{3K_3^2}{K_2^2} \right\}$ . Moreover, if  $K_1 \geq K_2$  and (4.41) holds, then  $A = \left[ \frac{3K_3^2}{K_2} \left( \frac{1}{K_1} - \frac{1}{K_2} \right) \right]^+ = 0$  and (4.43) is tightened as

$$\eta(0) < \frac{B + K_4}{K_2} - \mathbb{E}[v]. \quad (4.46)$$

This together with (4.44) yields the following bound

$$\mathbb{E}[v] \leq \frac{B - \zeta(0)}{K_2} + \zeta(0) \mathbf{1}_{\{\zeta(0) < 0\}} \left( \frac{1}{K_2} - \frac{1}{K_1} \right) + \frac{K_4}{K_2}. \quad (4.47)$$

□

*Proof of Claim 4.7.1.* We will show that  $\mathbb{E}[\eta(t+1)|\mathcal{F}(t)] \geq \eta(t)$ . There are two cases:

**Case I.**  $\zeta(t) < 0$ :

If  $\zeta(t+1) < 0$ , then

$$\eta(t+1) = -A + \frac{\zeta(t+1)}{K_1} - (t+1). \quad (4.48)$$

On the other hand, if  $\zeta(t+1) \geq 0$ , then by the assumption of Lemma 4.7.4,  $\zeta(t+1) \leq K_3$ , and we have

$$\begin{aligned} \eta(t+1) &= -Ae^{-\alpha\zeta(t+1)} + \frac{\zeta(t+1)}{K_2} - (t+1) \\ &\stackrel{(a)}{\geq} -A + \frac{\zeta(t+1)}{K_1} - (t+1), \end{aligned} \quad (4.49)$$

where (a) follows from the fact that 1) if  $K_1 \geq K_2$ , then by definition  $A = 0$ , and  $\frac{x}{K_2} \geq \frac{x}{K_1}$  for  $x \geq 0$ ; and 2) if  $K_1 < K_2$ , then  $-Ae^{-\alpha x} + \frac{x}{K_2}$  is concave in  $x$ ,  $-Ae^{-\alpha x} + \frac{x}{K_2} = -A + \frac{x}{K_1}$  for  $x = 0$ , and for  $x = K_3$

$$\begin{aligned} -Ae^{-\alpha K_3} + \frac{K_3}{K_2} &\geq -A(1 - \alpha K_3 + \frac{1}{2}(\alpha K_3)^2) + \frac{K_3}{K_2} \\ &= -A + A\alpha K_3(1 - \frac{1}{4}\frac{K_2}{K_3}) + \frac{K_3}{K_2} \\ &\geq -A + \frac{9}{8}K_3 \left( \frac{1}{K_1} - \frac{1}{K_2} \right) + \frac{K_3}{K_2} \\ &\geq -A + \frac{K_3}{K_1}. \end{aligned}$$

Combining (4.48) and (4.49), we obtain

$$\begin{aligned}
\mathbb{E}[\eta(t+1)|\mathcal{F}(t)] &\geq \mathbb{E}\left[-A + \frac{\zeta(t+1)}{K_1} - (t+1)|\mathcal{F}(t)\right] \\
&\geq -A + \frac{\zeta(t) + K_1}{K_1} - (t+1) \\
&= -A + \frac{\zeta(t)}{K_1} - t = \eta(t).
\end{aligned} \tag{4.50}$$

**Case II.**  $\zeta(t) \geq 0$ :

If  $\zeta(t+1) \geq 0$ , then

$$\eta(t+1) = -Ae^{-\alpha\zeta(t+1)} + \frac{\zeta(t+1)}{K_2} - (t+1). \tag{4.51}$$

On the other hand, if  $\zeta(t+1) < 0$ , then we have

$$\begin{aligned}
\eta(t+1) &= -A + \frac{\zeta(t+1)}{K_1} - (t+1) \\
&\stackrel{(a)}{\geq} -Ae^{-\alpha\zeta(t+1)} + \frac{\zeta(t+1)}{K_2} - (t+1),
\end{aligned} \tag{4.52}$$

where (a) follows from the fact that 1) if  $K_1 \geq K_2$ , then by definition  $A = 0$ , and  $\frac{x}{K_1} \geq \frac{x}{K_2}$  for  $x < 0$ ; and 2) if  $K_1 < K_2$ , then  $-Ae^{-\alpha x} + \frac{x}{K_2}$  is concave in  $x$ ,  $-Ae^{-\alpha x} + \frac{x}{K_2} = -A + \frac{x}{K_1}$  for  $x = 0$ , and  $-Ae^{-\alpha K_3} + \frac{K_3}{K_2} \geq -A + \frac{K_3}{K_1}$ .

Combining (4.51) and (4.52), we obtain

$$\begin{aligned}
&\mathbb{E}[\eta(t+1)|\mathcal{F}(t)] \\
&\geq \mathbb{E}\left[-Ae^{-\alpha\zeta(t+1)} + \frac{\zeta(t+1)}{K_2} - (t+1)|\mathcal{F}(t)\right] \\
&\geq \mathbb{E}\left[-Ae^{-\alpha\zeta(t+1)}|\mathcal{F}(t)\right] + \frac{\zeta(t) + K_2}{K_2} - (t+1) \\
&= \mathbb{E}\left[-Ae^{-\alpha\zeta(t+1)}|\mathcal{F}(t)\right] + Ae^{-\alpha\zeta(t)} + \eta(t) \\
&= \eta(t) - Ae^{-\alpha\zeta(t)}\mathbb{E}[e^{-\alpha(\zeta(t+1)-\zeta(t))} - 1|\mathcal{F}(t)] \\
&\stackrel{(a)}{\geq} \eta(t) - Ae^{-\alpha\zeta(t)}\mathbb{E}\left[-\alpha(\zeta(t+1) - \zeta(t)) + \frac{1}{2}\alpha^2(\zeta(t+1) - \zeta(t))^2 e^{\alpha K_3}|\mathcal{F}(t)\right] \\
&\geq \eta(t) + A\alpha e^{-\alpha\zeta(t)}\left[K_2 - \frac{1}{2}\alpha K_3^2 e^{\alpha K_3}\right] \\
&\stackrel{(b)}{\geq} \eta(t),
\end{aligned} \tag{4.53}$$

where (a) follows from the fact that for  $|x| \leq K$ ,

$$\begin{aligned} e^x &= 1 + \sum_{n=1}^{\infty} \frac{x^n}{n!} \\ &\leq 1 + x + \frac{x^2}{2} \left( 1 + \frac{K}{3} + \frac{K^2}{12} + \dots \right) \\ &\leq 1 + x + \frac{x^2}{2} e^K; \end{aligned}$$

and (b) holds since  $\frac{1}{2}\alpha K_3^2 e^{\alpha K_3} = \frac{1}{4}K_2 e^{\frac{0.5K_2}{K_3}} \leq \frac{e^{0.5}}{4}K_2 \leq K_2$ .  $\square$

#### 4.7.2 Proof of Theorem 4.2.1

We first prove inequality (4.2). From (3.1), the expected total cost under policy  $\pi$  can be written as

$$\begin{aligned} V_{\pi}(\boldsymbol{\rho}) &= \mathbb{E}_{\pi}[\tilde{\tau}_{1/L}] + L\mathbb{E}_{\pi}[1 - \max_{i \in \Omega_M} \rho_i(\tilde{\tau}_{1/L})] \\ &\stackrel{(a)}{\leq} \mathbb{E}_{\pi}[\tilde{\tau}_{1/L}] + 1, \end{aligned} \tag{4.54}$$

where (a) follows by construction (4.1). Next we find an upper bound for  $\mathbb{E}_{\pi}[\tilde{\tau}_{1/L}]$ .

Notice that for all  $i \in \Omega_M$ , upon selecting  $A(t) = a$  and observing  $Z(t) = z$ , the belief state evolves as

$$\rho_i(t+1) = \frac{\rho_i(t)q_i^a(z)}{\sum_{j=1}^M \rho_j(t)q_j^a(z)}.$$

Recall that  $U(\cdot)$  denotes the average log-likelihood function defined as

$$U(\boldsymbol{\rho}) = \sum_{i=1}^M \rho_i \log \frac{1 - \rho_i}{\rho_i}, \tag{4.55}$$

and let  $\mathcal{F}(t) = \sigma\{A(0), Z(0), \dots, A(t-1), Z(t-1)\}$  denote the history of sensing actions and observations up to time  $t$ . We have

$$\begin{aligned} &\mathbb{E}_{\pi}[U(\boldsymbol{\rho}(t+1))|\mathcal{F}(t)] \\ &= \sum_{a \in \mathcal{A}_M} P_{\pi}(A(t) = a) \mathbb{E}_{\pi} \left[ \sum_{i=1}^M \rho_i(t+1) \log \frac{1 - \rho_i(t+1)}{\rho_i(t+1)} | \mathcal{F}(t), A(t) = a \right] \end{aligned}$$

$$\begin{aligned}
&= \sum_{a \in \mathcal{A}_M} \pi(a|\boldsymbol{\rho}(t)) \int_{\mathcal{Z}} \sum_{i=1}^M \rho_i(t) q_i^a(z) \log \frac{\sum_{j \neq i} \rho_j(t) q_j^a(z)}{\rho_i(t) q_i^a(z)} dz \\
&= \sum_{i=1}^M \rho_i(t) \log \frac{1 - \rho_i(t)}{\rho_i(t)} + \sum_{a \in \mathcal{A}_M} \pi(a|\boldsymbol{\rho}(t)) \sum_{i=1}^M \int_{\mathcal{Z}} \rho_i(t) q_i^a(z) \log \frac{\sum_{j \neq i} \frac{\rho_j(t)}{1 - \rho_i(t)} q_j^a(z)}{q_i^a(z)} dz \\
&= U(\boldsymbol{\rho}(t)) - \sum_{a \in \mathcal{A}_M} \pi(a|\boldsymbol{\rho}(t)) \sum_{i=1}^M \rho_i(t) D(q_i^a \| \sum_{j \neq i} \frac{\rho_j(t)}{1 - \rho_i(t)} q_j^a) \\
&= U(\boldsymbol{\rho}(t)) - \sum_{a \in \mathcal{A}_M} \pi(a|\boldsymbol{\rho}(t)) EJS(\boldsymbol{\rho}(t), a) \\
&= U(\boldsymbol{\rho}(t)) - EJS(\boldsymbol{\rho}(t), \pi) \\
&\stackrel{(a)}{\leq} U(\boldsymbol{\rho}(t)) - \alpha,
\end{aligned} \tag{4.56}$$

where (a) follows from the assumption of Theorem 4.2.1.

Thus, the sequence  $\{-\frac{U(\boldsymbol{\rho}(t))}{\alpha} - t\}$  forms a submartingale with respect to the filtration  $\{\mathcal{F}(t)\}$ . Let define a stopping time<sup>5</sup>

$$v := \min \left\{ t : \max_{i \in \Omega_M} \rho_i(t) \geq 1 - \min \left\{ \frac{1}{\log 2M}, L^{-1} \right\} \right\}. \tag{4.57}$$

It is clear from (4.1) and (4.57) that  $\tilde{\tau}_{1/L} \leq v$  and hence,  $\mathbb{E}_\pi[\tilde{\tau}_{1/L}] \leq \mathbb{E}_\pi[v]$ . Next we find an upper bound on  $\mathbb{E}_\pi[v]$ . By Doob's Stopping Theorem,

$$\frac{-U(\boldsymbol{\rho}(0))}{\alpha} \leq \mathbb{E}_\pi \left[ \frac{-U(\boldsymbol{\rho}(v))}{\alpha} - v \right].$$

Rearranging the terms, we obtain

$$\begin{aligned}
\mathbb{E}_\pi[v] &\leq \frac{U(\boldsymbol{\rho}(0))}{\alpha} + \mathbb{E}_\pi \left[ \frac{-U(\boldsymbol{\rho}(v))}{\alpha} \right] \\
&\leq \frac{U(\boldsymbol{\rho}(0)) + \mathbb{E}_\pi [-U(\boldsymbol{\rho}(v-1)) + U(\boldsymbol{\rho}(v-1)) - U(\boldsymbol{\rho}(v))]}{\alpha} \\
&\stackrel{(a)}{\leq} \frac{U(\boldsymbol{\rho}(0)) + \max\{\log \log M, \log L\} + \mathbb{E}_\pi [U(\boldsymbol{\rho}(v-1)) - U(\boldsymbol{\rho}(v))]}{\alpha} \\
&\stackrel{(b)}{\leq} \frac{U(\boldsymbol{\rho}(0)) + \max\{\log \log M, \log L\} + 2^{\xi_M} \left( 3 + \frac{1}{\log 2M} \log(M-1) \right)}{\alpha} \\
&\leq \frac{U(\boldsymbol{\rho}(0)) + \max\{\log \log M, \log L\} + 4 \times 2^{\xi_M}}{\alpha}
\end{aligned}$$

---

<sup>5</sup>For this definition to make sense (and which we only need in the analysis) we assume that the decision maker continues sampling even after time  $\tilde{\tau}_{1/L}$  using the same rule as before.

$$\leq \frac{H(\boldsymbol{\rho}(0)) + \max\{\log \log M, \log L\} + 4 \times 2^{\xi_M}}{\alpha}, \quad (4.58)$$

where (a) holds since  $\rho_i(v-1) < 1 - \min\{\frac{1}{\log 2M}, L^{-1}\}$  for all  $i \in \Omega_M$  and hence,

$$\begin{aligned} -U(\boldsymbol{\rho}(v-1)) &= \sum_{i=1}^M \rho_i(v-1) \log \frac{\rho_i(v-1)}{1 - \rho_i(v-1)} \\ &< \log \frac{1 - \min\{\frac{1}{\log 2M}, L^{-1}\}}{\min\{\frac{1}{\log 2M}, L^{-1}\}} \\ &< \max\{\log \log M, \log L\}; \end{aligned}$$

and (b) follows from Lemma 4.7.3.

Inequality (4.58) together with (4.54) completes the proof of (4.2). Next we prove (4.5).

Recall that  $\tilde{\rho} = 1 - \frac{1}{1 + \max\{\log M, \log L\}}$ . Notice that if  $\rho_i(t) < \tilde{\rho}$  for all  $i \in \Omega_M$ ,

$$U(\boldsymbol{\rho}(t)) = \sum_{i=1}^M \rho_i(t) \log \frac{1 - \rho_i(t)}{\rho_i(t)} > \sum_{i=1}^M \rho_i(t) \log \frac{1 - \tilde{\rho}}{\tilde{\rho}} = \log \frac{1 - \tilde{\rho}}{\tilde{\rho}}; \quad (4.59)$$

which implies that if  $U(\boldsymbol{\rho}(t)) \leq \log \frac{1 - \tilde{\rho}}{\tilde{\rho}}$ , there exists  $i \in \Omega_M$  such that  $\rho_i(t) \geq \tilde{\rho}$ .

On the other hand, if  $\rho_i(t) \geq \tilde{\rho}$  for any  $i \in \Omega_M$ , then

$$\begin{aligned} U(\boldsymbol{\rho}(t)) &= \rho_i(t) \log \frac{1 - \rho_i(t)}{\rho_i(t)} + \sum_{j \neq i} \rho_j(t) \log \frac{1 - \rho_j(t)}{\rho_j(t)} \\ &\leq \rho_i(t) \log \frac{1 - \rho_i(t)}{\rho_i(t)} + (1 - \rho_i(t)) \log \frac{\sum_{j \neq i} (1 - \rho_j(t))}{1 - \rho_i(t)} \\ &\leq \tilde{\rho} \log \frac{1 - \tilde{\rho}}{\tilde{\rho}} + 1 + (1 - \tilde{\rho}) \log M \\ &\leq \log \frac{1 - \tilde{\rho}}{\tilde{\rho}} + 3; \end{aligned} \quad (4.60)$$

which implies that if  $U(\boldsymbol{\rho}(t)) > \log \frac{1 - \tilde{\rho}}{\tilde{\rho}} + 3$ , then  $\rho_i(t) < \tilde{\rho}$  for all  $i \in \Omega_M$ .

Following similar lines as those in (4.56) and by (4.3), (4.59), and (4.60),

$$\mathbb{E}_\pi [U(\boldsymbol{\rho}(t+1)) | \mathcal{F}(t)] \leq \begin{cases} U(\boldsymbol{\rho}(t)) - \alpha & \text{if } U(\boldsymbol{\rho}(t)) > \log \frac{1 - \tilde{\rho}}{\tilde{\rho}} + 3 \mathbf{1}_{\{\alpha > \beta\}} \\ U(\boldsymbol{\rho}(t)) - \beta & \text{if } U(\boldsymbol{\rho}(t)) \leq \log \frac{1 - \tilde{\rho}}{\tilde{\rho}} + 3 \mathbf{1}_{\{\alpha > \beta\}} \end{cases}. \quad (4.61)$$

The rest of the proof is divided into two cases:

- Case 1:  $L \leq \log 2M$ .

In this case,  $\log L \leq \log \log 2M \leq \log M$  and hence,  $\tilde{\rho} = 1 - \frac{1}{\log 2M} \geq 1 - L^{-1}$ . Following similar lines as those in the proof of inequality (4.2), we obtain

$$\mathbb{E}_\pi[\tilde{\tau}_{1/L}] \leq \frac{H(\boldsymbol{\rho}(0)) + \log \log M + 4 \times 2^{\xi_M}}{\alpha}. \quad (4.62)$$

- Case 2:  $L > \log 2M$ .

Define the sequence  $\{\tilde{U}(\boldsymbol{\rho}(t))\}_t$  as

$$\tilde{U}(\boldsymbol{\rho}(t)) := -U(\boldsymbol{\rho}(t)) + \log \frac{1 - \tilde{\rho}}{\tilde{\rho}} + 3 \mathbf{1}_{\{\alpha > \beta\}}. \quad (4.63)$$

Notice that if  $\max\{\tilde{U}(\boldsymbol{\rho}(t)), \tilde{U}(\boldsymbol{\rho}(t+1))\} \geq 3 \mathbf{1}_{\{\alpha > \beta\}}$ , then

$$\max_{i \in \Omega} \max\{\rho_i(t), \rho_i(t+1)\} \geq \tilde{\rho} = 1 - \frac{1}{1 + \max\{\log M, \log L\}}.$$

By Lemma 4.7.3 and from (4.61), the sequence  $\{\tilde{U}(\boldsymbol{\rho}(t))\}_t$  forms a submartingale with the following properties:

$$\begin{aligned} \mathbb{E}_\pi[\tilde{U}(\boldsymbol{\rho}(t+1)) | \mathcal{F}(t)] &\geq \tilde{U}(\boldsymbol{\rho}(t)) + \alpha \quad \text{if } \tilde{U}(\boldsymbol{\rho}(t)) < 0, \\ \mathbb{E}_\pi[\tilde{U}(\boldsymbol{\rho}(t+1)) | \mathcal{F}(t)] &\geq \tilde{U}(\boldsymbol{\rho}(t)) + \beta \quad \text{if } \tilde{U}(\boldsymbol{\rho}(t)) \geq 0, \\ |\tilde{U}(\boldsymbol{\rho}(t+1)) - \tilde{U}(\boldsymbol{\rho}(t))| &\leq 2^{\xi_M+2} \quad \text{if } \max\{\tilde{U}(\boldsymbol{\rho}(t+1)), \tilde{U}(\boldsymbol{\rho}(t))\} \geq 3 \mathbf{1}_{\{\alpha > \beta\}}. \end{aligned}$$

Let  $v := \min\{t : \tilde{U}(\boldsymbol{\rho}(t)) \geq \log L + 3 \mathbf{1}_{\{\alpha > \beta\}}\}$ . Note that by construction,  $\tilde{\tau}_{1/L} \leq v$ . Appealing to Lemma 4.7.4 and for  $\alpha \leq \beta$ ,

$$\begin{aligned} \mathbb{E}_\pi[\tilde{\tau}_{1/L}] &\leq \mathbb{E}_\pi[v] \\ &\leq \frac{\log L - \tilde{U}(\boldsymbol{\rho}(0))}{\beta} + \tilde{U}(\boldsymbol{\rho}(0)) \mathbf{1}_{\{\tilde{U}(\boldsymbol{\rho}(0)) < 0\}} \left( \frac{1}{\beta} - \frac{1}{\alpha} \right) + \frac{3(4 \times 2^{\xi_M})^2}{\alpha\beta} \\ &\leq \frac{\log L}{\beta} + \frac{U(\boldsymbol{\rho}(0)) - \log \frac{1-\tilde{\rho}}{\tilde{\rho}}}{\alpha} \mathbf{1}_{\{\tilde{U}(\boldsymbol{\rho}(0)) < 0\}} + \frac{3(4 \times 2^{\xi_M})^2}{\alpha\beta} \\ &\leq \frac{\log L}{\beta} + \frac{H(\boldsymbol{\rho}(0)) + \log \frac{\tilde{\rho}}{1-\tilde{\rho}}}{\alpha} + \frac{3(4 \times 2^{\xi_M})^2}{\alpha\beta} \\ &\stackrel{(a)}{\leq} \frac{H(\boldsymbol{\rho}(0)) + \max\{\log \log M, \log \log L\}}{\alpha} + \frac{\log L}{\beta} + \frac{3(4 \times 2^{\xi_M})^2}{\alpha\beta}, \end{aligned} \quad (4.64)$$

where (a) holds since  $\log \frac{\tilde{\rho}}{1-\tilde{\rho}} = \max\{\log \log M, \log \log L\}$ .

Similarly, from Lemma 4.7.4 and for  $\alpha > \beta$ ,

$$\mathbb{E}_\pi[\tilde{\tau}_{1/L}] \leq \frac{H(\boldsymbol{\rho}(0)) + \max\{\log \log M, \log \log L\}}{\alpha} + \frac{\log L + 3}{\beta} + \frac{4 \times 2^{\xi_M}}{\beta}. \quad (4.65)$$

Combining (4.62), (4.64), and (4.65) and from the fact that  $\max\{\alpha, \beta\} \leq \xi_M \leq 4 \times 2^{\xi_M}$ , we get

$$\mathbb{E}_\pi[\tilde{\tau}_{1/L}] \leq \frac{H(\boldsymbol{\rho}) + \max\{\log \log M, \log \log L\}}{\alpha} + \frac{\log L}{\beta} + \frac{3(4 \times 2^{\xi_M})^2}{\alpha\beta},$$

which together with (4.54) completes the proof of (4.5).

### 4.7.3 Proof of Theorem 4.3.1

Let  $\Gamma$  be the set of all mappings  $\gamma : \Omega_M \rightarrow \Omega_M$  such that  $\gamma(i) \neq i$  for  $i \in \Omega_M$ . Now associated with any  $\gamma \in \Gamma$  define

$$\underline{V}_1^\gamma(\boldsymbol{\rho}) = \left[ \sum_{i=1}^M \rho_i \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{\rho_i}{\rho_{\gamma(i)}}}{\max_{\hat{a} \in \mathcal{A}_M} D(q_i^{\hat{a}} \| q_{\gamma(i)}^{\hat{a}})} - K'_1 \right]^+.$$

Next we use Lemma 3.5.1 to show that  $V^* \geq \underline{V}_1^\gamma$  for all  $\gamma \in \Gamma$ . In particular, we show that for all  $\gamma \in \Gamma$  and for all  $\boldsymbol{\rho} \in \mathbb{P}(\Omega_M)$ ,  $\underline{V}_1^\gamma(\boldsymbol{\rho}) \leq \min\{1 + \min_{a \in \mathcal{A}_M} (\mathbb{T}^a \underline{V}_1^\gamma)(\boldsymbol{\rho}), \min_{j \in \Omega_M} (1 - \rho_j)L\}$ . For any  $\boldsymbol{\rho}$  such that  $\underline{V}_1^\gamma(\boldsymbol{\rho}) = 0$ , the inequality holds trivially. For  $\underline{V}_1^\gamma(\boldsymbol{\rho}) > 0$  and for any action  $a \in \mathcal{A}_M$  we have

$$\begin{aligned} (\mathbb{T}^a \underline{V}_1^\gamma)(\boldsymbol{\rho}) &\geq \sum_{i=1}^M \int \rho_i q_i^a(z) \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{\rho_i q_i^a(z)}{\rho_{\gamma(i)} q_{\gamma(i)}^a(z)}}{\max_{\hat{a} \in \mathcal{A}_M} D(q_i^{\hat{a}} \| q_{\gamma(i)}^{\hat{a}})} dz - K'_1 \\ &= \underline{V}_1^\gamma(\boldsymbol{\rho}) - \sum_{i=1}^M \rho_i \frac{D(q_i^a \| q_{\gamma(i)}^a)}{\max_{\hat{a} \in \mathcal{A}_M} D(q_i^{\hat{a}} \| q_{\gamma(i)}^{\hat{a}})} \\ &\geq \underline{V}_1^\gamma(\boldsymbol{\rho}) - 1. \end{aligned} \quad (4.66)$$

From (4.66), we have  $\underline{V}_1^\gamma(\boldsymbol{\rho}) \leq 1 + \min_{a \in \mathcal{A}_M} (\mathbb{T}^a \underline{V}_1^\gamma)(\boldsymbol{\rho})$ . What remains is to select  $K'_1$  sufficiently large such that  $\underline{V}_1^\gamma(\boldsymbol{\rho}) \leq \min_{j \in \Omega_M} (1 - \rho_j)L$  for all  $\gamma \in \Gamma$  and for

all  $\boldsymbol{\rho} \in \mathbb{P}(\Omega_M)$ . It is trivial that the desired result is obtained for any choice of  $K'_1$  satisfying

$$K'_1 \geq \max_{\gamma \in \Gamma} \max_{\boldsymbol{\rho} \in \mathbb{P}(\Omega_M)} \left\{ \sum_{i=1}^M \rho_i \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{\rho_i}{\rho_{\gamma(i)}}}{\max_{\hat{a} \in \mathcal{A}_M} D(q_i^{\hat{a}} \| q_{\gamma(i)}^{\hat{a}})} - \min_{j \in \Omega_M} (1 - \rho_j) L \right\}. \quad (4.67)$$

Using Claim 4.7.2 below and letting  $\underline{V}_1(\cdot) = \max_{\gamma \in \Gamma} \underline{V}_1^\gamma(\cdot)$ , we have the assertion of the theorem.

**Claim 4.7.2.** The constant  $K'_1$  can be selected independent of  $L$  such that the inequality  $\underline{V}_1^\gamma(\boldsymbol{\rho}) \leq \min_{j \in \Omega_M} (1 - \rho_j) L$  is satisfied for all  $\gamma \in \Gamma$ .

*Proof of Claim 4.7.2.* Let  $D_{\min}(M) = \min_{i,j \in \Omega_M} \max_{a \in \mathcal{A}_M} D(q_i^a \| q_j^a)$ . First we notice that, if  $\boldsymbol{\rho} \notin \mathbb{P}_L(\Omega_M)$ , then  $\underline{V}_1$  is bounded as:

$$\begin{aligned} \underline{V}_1(\boldsymbol{\rho}) &\leq \left[ \sum_{i=1}^M \rho_i \max_{j \neq i} \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{\rho_i}{1-\rho_i}}{\max_{a \in \mathcal{A}_M} D(q_i^a \| q_j^a)} - K'_1 \right]^+ \\ &\leq \left[ \sum_{\{i \in \Omega_M : \rho_i < 1-L^{-1}\}} \rho_i \frac{\log L + \log \frac{1}{\rho_i}}{D_{\min}(M)} - K'_1 \right]^+ \\ &\stackrel{(a)}{\leq} \left[ \left( \sum_{\{i \in \Omega_M : \rho_i < 1-L^{-1}\}} \rho_i \right) \frac{\log L + \log \frac{|\{i \in \Omega_M : \rho_i < 1-L^{-1}\}|}{\sum_{\{i \in \Omega_M : \rho_i < 1-L^{-1}\}} \rho_i}}{D_{\min}(M)} - K'_1 \right]^+ \\ &\stackrel{(b)}{\leq} \left[ \frac{2 + L^{-1} \log(M-1)}{D_{\min}(M)} - K'_1 \right]^+, \end{aligned} \quad (4.68)$$

where (a) follows by Jensen's inequality; and (b) holds since  $\sum_{\{i \in \Omega_M : \rho_i < 1-L^{-1}\}} \rho_i < L^{-1} < 1$  for any  $\boldsymbol{\rho} \notin \mathbb{P}_L(\Omega_M)$ , and  $x \log \frac{1}{x} \leq 1$  for  $x \in [0, 1]$ . In other words, for  $K'_1 \geq \frac{2+L^{-1} \log(M-1)}{D_{\min}(M)}$ ,

$$\underline{V}_1(\boldsymbol{\rho}) = 0 \leq \min_{j \in \Omega_M} (1 - \rho_j) L \quad \forall \boldsymbol{\rho} \notin \mathbb{P}_L(\Omega_M).$$

On the other hand, for all  $\boldsymbol{\rho} \in \mathbb{P}_L(\Omega_M)$ , we have

$$\begin{aligned} \underline{V}_1(\boldsymbol{\rho}) &\leq \left[ \sum_{i=1}^M \rho_i \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{\rho_i}{1-\rho_i}}{D_{\min}(M)} - K'_1 \right]^+ \\ &= \left[ \frac{\log(L-1)}{D_{\min}(M)} + \frac{\sum_{i=1}^M \rho_i \log \frac{1-\rho_i}{\rho_i}}{D_{\min}(M)} - K'_1 \right]^+. \end{aligned} \quad (4.69)$$

Now let

$$\begin{aligned}
 f_1(L, \boldsymbol{\rho}) &:= \frac{\log(L-1)}{D_{\min}(M)} + \frac{\sum_{i=1}^M \rho_i \log \frac{1-\rho_i}{\rho_i}}{D_{\min}(M)} \\
 &\stackrel{(a)}{=} \frac{\log(L-1)}{D_{\min}(M)} + \frac{\sum_{i=1}^{M-1} \rho_i \log \frac{1-\rho_i}{\rho_i} + (1 - \sum_{i=1}^{M-1} \rho_i) \log \frac{(\sum_{i=1}^{M-1} \rho_i)}{(1 - \sum_{i=1}^{M-1} \rho_i)}}{D_{\min}(M)},
 \end{aligned} \tag{4.70}$$

where (a) holds since  $\sum_{i=1}^M \rho_i = 1$ . Furthermore, let  $\alpha = \max\{1, \frac{2}{D_{\min}(M)}\}$ ,  $\beta = \frac{1}{3} (D_{\min}(M) - \frac{\log e}{\alpha})$ , and  $L^* = \max\{2\alpha, \frac{4}{\beta}, \frac{1}{\beta^2}, \frac{\log(M-1)}{\beta}\}$ . Next we show that for all  $L \geq L^*$  and at belief vector  $\boldsymbol{\rho} = [\frac{\alpha}{(M-1)L}, \dots, \frac{\alpha}{(M-1)L}, 1 - \frac{\alpha}{L}]$ , function  $f_1(L, \cdot)$  has an upper bound independent of  $L$  while its partial derivatives with respect to  $\rho_i$ ,  $i = 1, 2, \dots, M-1$  are less than  $L$ :

$$\begin{aligned}
 &f_1(L, \boldsymbol{\rho}) \Big|_{\boldsymbol{\rho} = [\frac{\alpha}{(M-1)L}, \dots, \frac{\alpha}{(M-1)L}, 1 - \frac{\alpha}{L}]} \\
 &= \frac{\log(L-1)}{D_{\min}(M)} + \frac{\sum_{i=1}^{M-1} \frac{\alpha}{(M-1)L} \log \frac{1 - \frac{\alpha}{(M-1)L}}{\frac{\alpha}{(M-1)L}} + (1 - \frac{\alpha}{L}) \log \frac{\frac{\alpha}{L}}{1 - \frac{\alpha}{L}}}{D_{\min}(M)} \\
 &= \frac{\log \frac{\alpha(L-1)}{L-\alpha}}{D_{\min}(M)} + \frac{\alpha \log \frac{1 - \frac{\alpha}{(M-1)L}}{\frac{\alpha}{(M-1)L}} + \log \frac{1 - \frac{\alpha}{L}}{\frac{\alpha}{L}}}{L D_{\min}(M)} \\
 &= \frac{\log \frac{\alpha(L-1)}{L-\alpha}}{D_{\min}(M)} + \frac{\alpha \log(M-1) + 2 \log \frac{\alpha}{L}}{L D_{\min}(M)} \\
 &\stackrel{(a)}{\leq} \frac{\log \frac{\alpha(L-1)}{L-\alpha}}{D_{\min}(M)} + \frac{\alpha\beta + 2}{D_{\min}(M)} \\
 &= \frac{\log \frac{\alpha}{1 - \frac{\alpha-1}{L-1}}}{D_{\min}(M)} + \frac{\alpha\beta + 2}{D_{\min}(M)} \\
 &\stackrel{(b)}{\leq} \frac{\log(2\alpha - 1) + \alpha\beta + 2}{D_{\min}(M)},
 \end{aligned} \tag{4.71}$$

where (a) follows from the fact that  $x \log \frac{1}{x} \leq 1$  for  $x \in [0, 1]$  and  $L \geq \frac{\log(M-1)}{\beta}$ ; and (b) holds since  $L \geq 2\alpha$ , and

$$\begin{aligned}
 &\frac{\partial f_1}{\partial \rho_i}(L, \boldsymbol{\rho}) \Big|_{\boldsymbol{\rho} = [\frac{\alpha}{(M-1)L}, \dots, \frac{\alpha}{(M-1)L}, 1 - \frac{\alpha}{L}]} \\
 &= \left( \log \frac{1 - \rho_i}{\rho_i} - \frac{\log e}{1 - \rho_i} - \log \frac{1 - \rho_M}{\rho_M} + \frac{\log e}{1 - \rho_M} \right) \frac{1}{D_{\min}(M)} \Big|_{\boldsymbol{\rho} = [\frac{\alpha}{(M-1)L}, \dots, \frac{\alpha}{(M-1)L}, 1 - \frac{\alpha}{L}]} \\
 &= \left( \log \frac{(M-1)L - \alpha}{\alpha} - \frac{\log e}{1 - \frac{\alpha}{(M-1)L}} - \log \frac{\alpha}{L - \alpha} + \frac{\log e}{\frac{\alpha}{L}} \right) \frac{1}{D_{\min}(M)}
 \end{aligned}$$

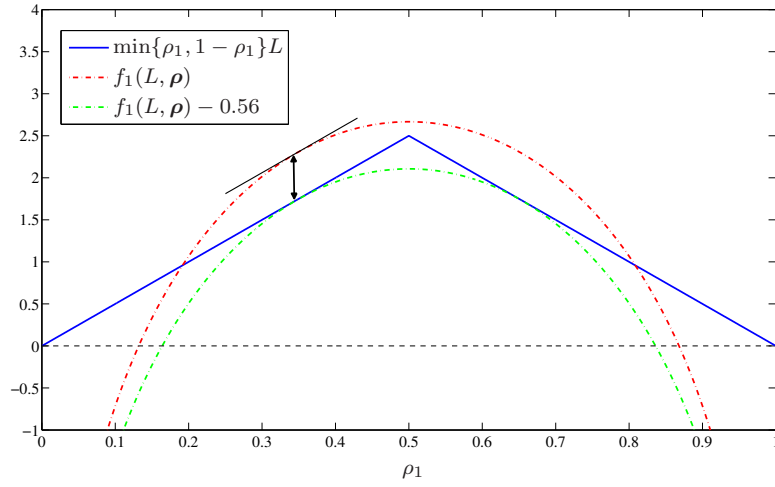
$$\begin{aligned}
&\leq \left( \log(M-1) + 2\log L + \frac{\log e}{\alpha} L \right) \frac{1}{D_{\min}(M)} \\
&= L + \left( \log(M-1) + 2\log L - \left( D_{\min}(M) - \frac{\log e}{\alpha} \right) L \right) \frac{1}{D_{\min}(M)} \\
&\leq L + (\log(M-1) + 2\log L - 3\beta L) \frac{1}{D_{\min}(M)} \\
&\leq L + (\log L - \beta L) \frac{2}{D_{\min}(M)} \\
&\stackrel{(a)}{\leq} L + (\log L - \log(\beta L)^2) \frac{2}{D_{\min}(M)} \\
&\leq L,
\end{aligned} \tag{4.72}$$

where (a) follows from the fact that  $\log x^2 \leq x$  for  $x \geq 4$ .

From symmetry and concavity of  $f_1(L, \cdot)$  and inequalities (4.71) and (4.72), it is clear that for  $L \geq L^*$ , and for all  $\boldsymbol{\rho} \in \mathbb{P}_L(\Omega_M)$ ,

$$\left[ f_1(L, \boldsymbol{\rho}) - \frac{\log(2\alpha - 1) + \alpha\beta + 2}{D_{\min}(M)} \right]^+ \leq \min_{j \in \Omega_M} (1 - \rho_j) L. \tag{4.73}$$

Fig. 4.4 shows this for  $M = 2$ .



**Figure 4.4:** Computing  $K'_1$  for  $M = 2$ ,  $L = 5$ , and  $D_{\min}(M) = 0.75$ . In this example, the derivative of  $f_1$  with respect to  $\rho_1$  is equal to  $L$  at  $\rho_1 = 0.35$  and  $K'_1 \geq 0.56$  ensures that  $f_1(L, \boldsymbol{\rho}) - K'_1 \leq \min\{\rho_1, 1 - \rho_1\}L$ . We have  $\alpha = 2.67$ ,  $\beta = 0.07$ ,  $L^* = 206.06$ , and  $K'_1 = 13.16$ .

Furthermore, for all  $L < L^*$ ,

$$f_1(L, \boldsymbol{\rho}) \leq f_1(L^*, \boldsymbol{\rho}) \leq \max_{\hat{\boldsymbol{\rho}}} f_1(L^*, \hat{\boldsymbol{\rho}}) \leq \frac{\log L^* + \log M}{D_{\min}(M)}.$$

This together with (4.68), (4.69), and (4.73) implies the assertion of the claim for

$$\begin{aligned} K'_1 &= \frac{1 + \alpha\beta + \log L^* + \log M}{D_{\min}(M)} \\ &\geq \frac{\max\{2 + L^{-1} \log(M-1), \log L^* + \log M, \log(2\alpha - 1) + \alpha\beta + 2\}}{D_{\min}(M)}. \end{aligned} \quad (4.74)$$

□

#### 4.7.4 Proof of Theorem 4.3.2

We show that for all  $\boldsymbol{\rho} \in \mathbb{P}(\Omega_M)$ ,

$$\begin{aligned} V^*(\boldsymbol{\rho}) &\geq \left[ \frac{H(\boldsymbol{\rho}) - H([\alpha(L, M), 1 - \alpha(L, M)]) - \alpha(L, M) \log(M-1)}{I_{\max}(M)} \right. \\ &\quad \left. + \alpha(L, M)L \right]^+. \end{aligned} \quad (4.75)$$

Note that the right-hand side of (4.75) can be written as

$$G(\boldsymbol{\rho}) := \left[ \frac{H(\boldsymbol{\rho}) - H(\boldsymbol{\nu})}{I_{\max}(M)} + \alpha(L, M)L \right]^+, \quad (4.76)$$

where

$$\boldsymbol{\nu} = \left[ \frac{\alpha(L, M)}{M-1}, \dots, \frac{\alpha(L, M)}{M-1}, 1 - \alpha(L, M) \right]. \quad (4.77)$$

Next we show that  $G(\boldsymbol{\rho}) \leq \min\{1 + \min_{a \in \mathcal{A}_M} (\mathbb{T}^a G)(\boldsymbol{\rho}), \min_{j \in \Omega_M} (1 - \rho_j)L\}$  for all  $\boldsymbol{\rho} \in \mathbb{P}(\Omega_M)$ . For any  $\boldsymbol{\rho}$  such that  $G(\boldsymbol{\rho}) = 0$ , the inequality holds trivially. For  $G(\boldsymbol{\rho}) > 0$  and for any action  $a \in \mathcal{A}_M$  we have

$$\begin{aligned} (\mathbb{T}^a G)(\boldsymbol{\rho}) &= \frac{\int H(\Phi^a(\boldsymbol{\rho}, z)) q_{\boldsymbol{\rho}}^a(z) dz - H(\boldsymbol{\nu})}{I_{\max}(M)} + \alpha(L, M)L \\ &= \frac{H(\boldsymbol{\rho}) - I(\boldsymbol{\rho}; q_{\boldsymbol{\rho}}^a) - H(\boldsymbol{\nu})}{I_{\max}(M)} + \alpha(L, M)L \end{aligned}$$

$$\begin{aligned}
&= G(\boldsymbol{\rho}) - \frac{I(\boldsymbol{\rho}; q_{\boldsymbol{\rho}}^a)}{I_{\max}(M)} \\
&\geq G(\boldsymbol{\rho}) - 1,
\end{aligned} \tag{4.78}$$

where the last inequality follows from the fact that

$$I(\boldsymbol{\rho}; q_{\boldsymbol{\rho}}^a) \leq \max_{\hat{\mathbf{a}} \in \mathcal{A}_M} \max_{\hat{\boldsymbol{\rho}} \in \mathbb{P}(\Omega_M)} I(\hat{\boldsymbol{\rho}}; q_{\hat{\boldsymbol{\rho}}}^{\hat{\mathbf{a}}}) = I_{\max}(M).$$

Therefore,

$$G(\boldsymbol{\rho}) \leq 1 + \min_{a \in \mathcal{A}_M} (\mathbb{T}^a G)(\boldsymbol{\rho}).$$

What remains is to show that  $G(\boldsymbol{\rho}) \leq \min_{j \in \Omega_M} (1 - \rho_j)L$ . Rewriting  $G$  as

$$G(\boldsymbol{\rho}) = \left[ \frac{\sum_{i=1}^{M-1} \rho_i \log \frac{1}{\rho_i} + (1 - \sum_{i=1}^{M-1} \rho_i) \log \frac{1}{1 - \sum_{i=1}^{M-1} \rho_i} - H(\boldsymbol{\nu})}{I_{\max}(M)} + \alpha(L, M)L \right]^+,$$

we can compute the gradient at  $\boldsymbol{\nu}$ . For all  $i = 1, 2, \dots, M-1$ ,

$$\begin{aligned}
\frac{\partial G}{\partial \rho_i}(\boldsymbol{\nu}) &= \left( \log \frac{1}{\rho_i} - \log e - \log \frac{1}{1 - \sum_{i=1}^{M-1} \rho_i} + \log e \right) / I_{\max}(M) \Big|_{\boldsymbol{\rho}=\boldsymbol{\nu}} \\
&= \left( \log \frac{\rho_M}{\rho_i} \right) / I_{\max}(M) \Big|_{\boldsymbol{\rho}=\boldsymbol{\nu}} = \left( \log \frac{1 - \alpha(L, M)}{\frac{\alpha(L, M)}{M-1}} \right) / I_{\max}(M) = L.
\end{aligned}$$

Furthermore,  $G(\boldsymbol{\nu}) = \alpha(L, M)L = (1 - \nu_M)L$ . Without loss of generality and since both functions  $G(\boldsymbol{\rho})$  and  $\min_{j \in \Omega_M} (1 - \rho_j)L$  are symmetric, let us focus on  $\mathbb{P}_M(\Omega_M) := \{\boldsymbol{\rho} \in \mathbb{P}(\Omega_M) : \rho_M \geq \rho_i, \forall i \in \Omega_M - \{M\}\}$ . In this case,  $\min_{j \in \Omega_M} (1 - \rho_j)L = (1 - \rho_M)L = \sum_{i=1}^{M-1} \rho_i L$  and hence  $\min_{j \in \Omega_M} (1 - \rho_j)L$  is the tangent hyperplane to  $G(\boldsymbol{\rho})$  at  $\boldsymbol{\nu}$ . This along with concavity of function  $G$  implies  $G(\boldsymbol{\rho}) \leq \min_{j \in \Omega_M} (1 - \rho_j)L$ . Using Lemma 3.5.1, we have the assertion of the theorem.

#### 4.7.5 Proof of Theorem 4.3.3

We need to show that

$$\begin{aligned}
V^*(\boldsymbol{\rho}) \geq \underline{V}_3(\boldsymbol{\rho}) &= \left[ \frac{H(\boldsymbol{\rho}) - H([\delta, 1 - \delta]) - \delta \log(M-1)}{I_{\max}(M)} \right. \\
&\quad \left. + \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{1-\delta}{\delta} - \xi_M}{D_{\max}(M)} \mathbf{1}_{\{\max_{i \in \Omega_M} \rho_i \leq 1-\delta\}} - K'_3 \right]^+.
\end{aligned} \tag{4.79}$$

We show this in two steps. First we consider the following function:

$$J'(\boldsymbol{\rho}) := \left[ \sum_{i=1}^M \rho_i \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{\rho_i}{1-\rho_i}}{D_{\max}(M)} - K'_3 \right]^+. \quad (4.80)$$

We use Jensen's inequality to show that

**Claim 4.7.3.** For all  $\boldsymbol{\rho} \in \mathbb{P}(\Omega_M)$ ,  $J'(\boldsymbol{\rho}) \leq 1 + \min_{a \in \mathcal{A}_M} (\mathbb{T}^a J')(\boldsymbol{\rho})$ .

*Proof of Claim 4.7.3.* For any  $\boldsymbol{\rho}$  such that  $J'(\boldsymbol{\rho}) = 0$ , the inequality holds trivially.

For any  $\boldsymbol{\rho}$  such that  $J'(\boldsymbol{\rho}) > 0$  and for any  $a \in \mathcal{A}_M$  we have

$$\begin{aligned} (\mathbb{T}^a J')(\boldsymbol{\rho}) &\geq \sum_{i=1}^M \int \rho_i q_i^a(z) \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{\rho_i q_i^a(z)}{\sum_{k \neq i} \rho_k q_k^a(z)}}{D_{\max}(M)} dz - K'_3 \\ &= J'(\boldsymbol{\rho}) - \sum_{i=1}^M \rho_i \frac{\int q_i^a(z) \log \frac{q_i^a(z)}{\sum_{k \neq i} \frac{\rho_k}{1-\rho_i} q_k^a(z)} dz}{D_{\max}(M)} \\ &\stackrel{(a)}{\geq} J'(\boldsymbol{\rho}) - \sum_{i=1}^M \rho_i \frac{\sum_{k \neq i} \frac{\rho_k}{1-\rho_i} D(q_i^a \| q_k^a)}{D_{\max}(M)} \\ &\geq J'(\boldsymbol{\rho}) - 1, \end{aligned}$$

where (a) follows from Jensen's inequality. □

Next we define  $J(\boldsymbol{\rho}) = \max\{J'(\boldsymbol{\rho}), J''(\boldsymbol{\rho})\}$  where  $J''(\boldsymbol{\rho})$  is the right-hand side of (4.79), i.e.,

$$\begin{aligned} J''(\boldsymbol{\rho}) &= \left[ \frac{H(\boldsymbol{\rho}) - H([\delta, 1-\delta]) - \delta \log(M-1)}{I_{\max}(M)} \right. \\ &\quad \left. + \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{1-\delta}{\delta} - \xi_M}{D_{\max}(M)} \mathbf{1}_{\{\max_{i \in \Omega_M} \rho_i \leq 1-\delta\}} - K'_3 \right]^+. \end{aligned} \quad (4.81)$$

- **Case 1:** For all  $\boldsymbol{\rho}$  such that  $J(\boldsymbol{\rho}) = 0$  or  $J(\boldsymbol{\rho}) = J'(\boldsymbol{\rho})$ , it is trivial from Claim 4.7.3 that

$$J(\boldsymbol{\rho}) = J'(\boldsymbol{\rho}) \leq 1 + \min_{a \in \mathcal{A}_M} (\mathbb{T}^a J')(\boldsymbol{\rho}) \leq 1 + \min_{a \in \mathcal{A}_M} (\mathbb{T}^a J)(\boldsymbol{\rho}). \quad (4.82)$$

- **Case 2:** For all  $\boldsymbol{\rho}$  such that  $J(\boldsymbol{\rho}) = J''(\boldsymbol{\rho}) > 0$ , and for any action  $a \in \mathcal{A}_M$ ,

$$\begin{aligned}
(\mathbb{T}^a J)(\boldsymbol{\rho}) &= \int J(\Phi^a(\boldsymbol{\rho}, z)) q_{\boldsymbol{\rho}}^a(z) dz \\
&\stackrel{(a)}{\geq} \frac{\int H(\Phi^a(\boldsymbol{\rho}, z)) q_{\boldsymbol{\rho}}^a(z) dz - H([\delta, 1 - \delta]) - \delta \log(M - 1)}{I_{\max}(M)} \\
&\quad + \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{1-\delta}{\delta} - \xi_M}{D_{\max}(M)} \mathbf{1}_{\{\max_{i \in \Omega_M} \rho_i \leq 1-\delta\}} - K'_3 \\
&= J''(\boldsymbol{\rho}) - \frac{I(\boldsymbol{\rho}; q_{\boldsymbol{\rho}}^a)}{I_{\max}(M)} \\
&\geq J''(\boldsymbol{\rho}) - 1 \\
&\stackrel{(b)}{=} J(\boldsymbol{\rho}) - 1,
\end{aligned} \tag{4.83}$$

where (a) follows from Claim 4.7.4 below and (b) holds since  $\boldsymbol{\rho}$  is such that  $J(\boldsymbol{\rho}) = J''(\boldsymbol{\rho})$ .

**Claim 4.7.4.** Let  $\boldsymbol{\rho}$  be such that  $J(\boldsymbol{\rho}) = J''(\boldsymbol{\rho}) > 0$ . If Assumption 4.1.2 holds, then for all actions  $a \in \mathcal{A}_M$  and observations  $z \in \mathcal{Z}$ ,

$$\begin{aligned}
J(\Phi^a(\boldsymbol{\rho}, z)) &\geq \frac{H(\Phi^a(\boldsymbol{\rho}, z)) - H([\delta, 1 - \delta]) - \delta \log(M - 1)}{I_{\max}(M)} \\
&\quad + \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{1-\delta}{\delta} - \xi_M}{D_{\max}(M)} \mathbf{1}_{\{\max_{i \in \Omega_M} \rho_i \leq 1-\delta\}} - K'_3.
\end{aligned} \tag{4.84}$$

*Proof of Claim 4.7.4.* For all  $\boldsymbol{\rho}$  satisfying  $\max_{i \in \Omega_M} \rho_i > 1 - \delta$ ,

$$\begin{aligned}
H(\boldsymbol{\rho}) &< (1 - \delta) \log \frac{1}{1 - \delta} + (M - 1) \times \frac{\delta}{M - 1} \log \frac{1}{\delta / (M - 1)} \\
&= H([\delta, 1 - \delta]) + \delta \log(M - 1),
\end{aligned}$$

hence,  $J'' < 0$ . In other words,  $J(\boldsymbol{\rho}) = J''(\boldsymbol{\rho}) > 0$  implies that  $\max_{i \in \Omega_M} \rho_i \leq 1 - \delta$ .

Let  $\hat{\boldsymbol{\rho}} = \Phi^a(\boldsymbol{\rho}, z)$ . Inequality (4.84) holds trivially if  $\max_{i \in \Omega_M} \hat{\rho}_i \leq 1 - \delta$  since  $J(\hat{\boldsymbol{\rho}}) \geq J''(\hat{\boldsymbol{\rho}})$  and  $J''(\hat{\boldsymbol{\rho}})$  is greater than or equal to the right-hand side of (4.84).

If  $\max_{i \in \Omega_M} \hat{\rho}_i > 1 - \delta$ , we get

$$\begin{aligned}
J(\hat{\boldsymbol{\rho}}) &= J'(\hat{\boldsymbol{\rho}}) \\
&= \left[ \sum_{i=1}^M \hat{\rho}_i \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{\hat{\rho}_i}{1-\hat{\rho}_i}}{D_{\max}(M)} - K'_3 \right]^+
\end{aligned}$$

$$\begin{aligned}
&\stackrel{(a)}{\geq} \left[ \sum_{i=1}^M \hat{\rho}_i \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{1-\delta}{\delta} - \xi_M}{D_{\max}(M)} - K'_3 \right]^+ \\
&\geq \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{1-\delta}{\delta} - \xi_M}{D_{\max}(M)} - K'_3,
\end{aligned}$$

where (a) follows from the fact that under Assumption 4.1.2 and for all  $i \in \Omega_M$ ,

$$\begin{aligned}
\log \frac{\hat{\rho}_i}{1 - \hat{\rho}_i} &\leq \left| \log \frac{\hat{\rho}_i}{1 - \hat{\rho}_i} - \log \frac{\rho_i}{1 - \rho_i} \right| + \left| \log \frac{\rho_i}{1 - \rho_i} \right| \\
&\leq \left| \log \frac{\rho_i q_i^a(z)}{\sum_{j \neq i} \rho_j q_j^a(z)} - \log \frac{\rho_i}{1 - \rho_i} \right| + \log \frac{1 - \delta}{\delta} \\
&= \left| \log \frac{q_i^a(z)}{\sum_{j \neq i} \frac{\rho_i}{1 - \rho_j} q_j^a(z)} \right| + \log \frac{1 - \delta}{\delta} \\
&\leq \xi_M + \log \frac{1 - \delta}{\delta}.
\end{aligned}$$

□

Combining (4.82) and (4.83), we have that

$$J(\boldsymbol{\rho}) \leq 1 + \min_{a \in \mathcal{A}_M} (\mathbb{T}^a J)(\boldsymbol{\rho}). \quad (4.85)$$

Lemma 3.5.1, together with (4.85) and Claim 4.7.5 below, implies that  $V^* \geq J = \max\{J', J''\} \geq J'' = \underline{V}_3$ . This is a slightly stronger result than (4.79).

**Claim 4.7.5.** For  $L > \frac{\log M}{I_{\max}(M)}$ , constant  $K'_3$  can be selected independent of  $\delta$  and  $L$  such that  $J(\boldsymbol{\rho}) \leq \min_{j \in \Omega_M} (1 - \rho_j)L$ . Furthermore, if  $\sup_M \xi_M < \infty$ , then  $K'_3$  can be selected independent of  $M$  as well.

*Proof of Claim 4.7.5.* Following similar lines as the proof of Claim 4.7.2, we can select  $K'_3$  sufficiently large such that  $J'(\boldsymbol{\rho}) \leq \min_{j \in \Omega_M} (1 - \rho_j)L$ . Recall that  $J'(\boldsymbol{\rho}) = [f_2(L, \boldsymbol{\rho}) - K'_3]^+$ , where

$$f_2(L, \boldsymbol{\rho}) := \frac{\log(L-1)}{D_{\max}(M)} + \frac{\sum_{i=1}^M \rho_i \log \frac{1-\rho_i}{\rho_i}}{D_{\max}(M)}. \quad (4.86)$$

By the assumption of Claim 4.7.5, there exists  $\kappa > 0$  such that  $\log M \leq (I_{\max}(M) - \kappa)L$  for all  $M$ . Set  $\alpha = \max\{1, \frac{2}{\kappa}\}$ ,  $\beta = \frac{1}{2} \left( \kappa - \frac{\log e}{\alpha} \right)$ , and let  $L'_2 = \max\{2\alpha, \frac{4}{\beta}, \frac{1}{\beta^2}\}$ . For  $L \geq L'_2$  and for all  $i = 1, 2, \dots, M-1$ , we obtain

$$\begin{aligned}
f_2(L, \boldsymbol{\rho}) \Big|_{\boldsymbol{\rho} = [\frac{\alpha}{(M-1)L}, \dots, \frac{\alpha}{(M-1)L}, 1 - \frac{\alpha}{L}]} &= \frac{\log \frac{\alpha(L-1)}{L-\alpha}}{D_{\max}(M)} + \frac{\alpha \log(M-1) + 2 \log \frac{\alpha}{L}}{L D_{\max}(M)} \\
&\leq \frac{\log \frac{\alpha(L-1)}{L-\alpha}}{D_{\max}(M)} + \frac{\alpha(I_{\max}(M) - \kappa) + 2}{D_{\max}(M)} \\
&\leq \frac{\log(2\alpha - 1) + \alpha(I_{\max}(M) - \kappa) + 2}{D_{\max}(M)}, \quad (4.87)
\end{aligned}$$

and,

$$\begin{aligned}
\frac{\partial f_2}{\partial \rho_i}(L, \boldsymbol{\rho}) \Big|_{\boldsymbol{\rho} = [\frac{\alpha}{(M-1)L}, \dots, \frac{\alpha}{(M-1)L}, 1 - \frac{\alpha}{L}]} &\leq \left( \log(M-1) + 2 \log L + \frac{\log e}{\alpha} L \right) \frac{1}{D_{\max}(M)} \\
&\leq \left( 2 \log L + (I_{\max}(M) - \kappa + \frac{\log e}{\alpha}) L \right) \frac{1}{D_{\max}(M)} \\
&\leq L + (\log L - \beta L) \frac{2}{D_{\max}(M)} \\
&\leq L. \quad (4.88)
\end{aligned}$$

From symmetry and concavity of  $f_2(L, \cdot)$ , (4.87), (4.88), it is clear that for  $L \geq L'_2$ , and for all  $\boldsymbol{\rho} \in \mathbb{P}_L(\Omega_M)$ ,

$$\left[ f_2(L, \boldsymbol{\rho}) - \frac{\log(2\alpha - 1) + \alpha(I_{\max}(M) - \kappa) + 2}{D_{\max}(M)} \right]^+ \leq \min_{j \in \Omega_M} (1 - \rho_j) L. \quad (4.89)$$

Furthermore, for  $L < L'_2$ , we have

$$\begin{aligned}
f_2(L, \boldsymbol{\rho}) &\leq f_2(L'_2, \boldsymbol{\rho}) \leq \max_{\hat{\boldsymbol{\rho}}} f_2(L'_2, \hat{\boldsymbol{\rho}}) \\
&\leq \frac{\log L'_2 + \log M}{D_{\max}(M)} \leq \frac{\log L'_2 + L'_2(I_{\max}(M) - \kappa)}{D_{\max}(M)}.
\end{aligned}$$

In other words, selecting

$$K'_3 \geq \frac{\max\{\log L'_2 + L'_2(I_{\max}(M) - \kappa), \log(2\alpha - 1) + \alpha(I_{\max}(M) - \kappa) + 2\}}{D_{\max}(M)}, \quad (4.90)$$

satisfies  $J'(\boldsymbol{\rho}) \leq \min_{j \in \Omega_M} (1 - \rho_j) L$ .

Next we discuss on the selection of  $K'_3$  such that  $J''(\boldsymbol{\rho}) \leq \min_{j \in \Omega_M} (1 - \rho_j) L$  is also satisfied. Let

$$f_3(L, \boldsymbol{\rho}) := \frac{H(\boldsymbol{\rho}) - H([\delta, 1 - \delta]) - \delta \log(M-1)}{I_{\max}(M)},$$

and rewrite,

$$J''(\boldsymbol{\rho}) = \left[ f_3(L, \boldsymbol{\rho}) + \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{1-\delta}{\delta} - \xi_M}{D_{\max}(M)} \mathbf{1}_{\{\max_{i \in \Omega_M} \rho_i \leq 1-\delta\}} - K'_3 \right]^+.$$

We show that at belief vector  $\boldsymbol{\rho} = [\frac{0.5}{(M-1)L}, \dots, \frac{0.5}{(M-1)L}, 1 - \frac{0.5}{L}]$ , and for  $L \geq L''_2 := \max\{\frac{4I_{\max}(M)}{\kappa}, \frac{I_{\max}^2(M)}{\kappa^2}\}$ ,  $f_3(L, \boldsymbol{\rho})$  has an upper bound independent of  $L$  and its partial derivatives with respect to  $\rho_i$ ,  $i = 1, 2, \dots, M-1$  are less than  $L$ . In other words,

$$\begin{aligned} & f_3(L, \boldsymbol{\rho}) \Big|_{\boldsymbol{\rho} = [\frac{0.5}{(M-1)L}, \dots, \frac{0.5}{(M-1)L}, 1 - \frac{0.5}{L}]} \\ &= \frac{H([\frac{1}{2L}, 1 - \frac{1}{2L}]) + \frac{1}{2L} \log(M-1) - H([\delta, 1-\delta]) - \delta \log(M-1)}{I_{\max}(M)} \\ &\leq \frac{1 + \frac{1}{2}(I_{\max}(M) - \kappa)}{I_{\max}(M)}, \end{aligned}$$

and,

$$\begin{aligned} \frac{\partial f_3}{\partial \rho_i}(L, \boldsymbol{\rho}) \Big|_{\boldsymbol{\rho} = [\frac{0.5}{(M-1)L}, \dots, \frac{0.5}{(M-1)L}, 1 - \frac{0.5}{L}]} &= \frac{1}{I_{\max}(M)} \log \frac{\rho_M}{\rho_i} \Big|_{\boldsymbol{\rho} = [\frac{0.5}{(M-1)L}, \dots, \frac{0.5}{(M-1)L}, 1 - \frac{0.5}{L}]} \\ &= \frac{1}{I_{\max}(M)} \log \frac{1 - \frac{0.5}{L}}{\frac{0.5}{(M-1)L}} \\ &\leq \frac{1}{I_{\max}(M)} (\log(M-1) + \log L) \\ &\leq \frac{1}{I_{\max}(M)} (L(I_{\max}(M) - \kappa) + \log L) \\ &= L - \left( \frac{\kappa}{I_{\max}(M)} L - \log L \right) \\ &\leq L. \end{aligned}$$

Furthermore, for  $L < L''_2$ , we have

$$f_3(L, \boldsymbol{\rho}) \leq f_3(L''_2, \boldsymbol{\rho}) \leq \max_{\boldsymbol{\rho}} f_3(L''_2, \boldsymbol{\rho}) \leq \frac{\log M}{I_{\max}(M)} \leq \frac{L''_2(I_{\max}(M) - \kappa)}{I_{\max}(M)}.$$

We also note that for all  $K'_3$  satisfying (4.90),

$$\frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{1-\delta}{\delta} - \xi_M}{D_{\max}(M)} \mathbf{1}_{\{\max_{i \in \Omega_M} \rho_i \leq 1-\delta\}} - K'_3 \leq J'(\boldsymbol{\rho}) \leq \min_{j \in \Omega_M} (1 - \rho_j)L.$$

Thus, if  $K'_3$  satisfies (4.90) as well as the following condition,

$$K'_3 \geq \frac{\max\{L_2''(I_{\max}(M) - \kappa), 1 + \frac{1}{2}(I_{\max}(M) - \kappa)\}}{I_{\max}(M)}, \quad (4.91)$$

we have

$$\left[ f_3(L, \boldsymbol{\rho}) + \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{1-\delta}{\delta} - \xi_M}{D_{\max}(M)} \mathbf{1}_{\{\max_{i \in \Omega_M} \rho_i \leq 1-\delta\}} - K'_3 \right]^+ \leq \min_{j \in \Omega_M} (1 - \rho_j) L.$$

By selecting  $K'_3$  to be independent from  $\delta$  and  $L$  and larger than (4.90) and (4.91), we have the assertion of the claim. The following selection of  $K'_3$  satisfies the above conditions:

$$K'_3 = \max \left\{ \frac{\log L'_2 + L'_2(I_{\max}(M) - \kappa) + 2}{D_{\max}(M)}, \frac{4I_{\max}(M)}{\kappa} + \frac{I_{\max}^2(M)}{\kappa^2} + \frac{1}{I_{\max}(M)} \right\}. \quad (4.92)$$

If  $\sup_M \xi_M < \infty$ , then  $I_{\max}(M) \leq \bar{I}_{\max} < \infty$  for all  $M$ , and we can select  $K'_3$  to be independent from  $M$  as well:

$$K'_3 = \max \left\{ \frac{\log L'_2 + L'_2(\bar{I}_{\max} - \kappa) + 2}{\underline{D}_{\max}}, \frac{4\bar{I}_{\max}}{\kappa} + \frac{\bar{I}_{\max}^2}{\kappa^2} + \frac{1}{\underline{I}_{\max}} \right\}. \quad (4.93)$$

□

#### 4.7.6 Proof of Corollaries 4.3.1 and 4.5.2

*Proof of Corollary 4.3.1.* We first show that for all belief vectors  $\boldsymbol{\nu} \in \mathbb{P}(\Omega_M)$  for which  $\nu_i = 1 - \alpha(L, M)$  for an  $i \in \Omega_M$  and  $\nu_j = \frac{\alpha(L, M)}{M-1}$  for all  $j \in \Omega_M - \{i\}$ , the optimal action is to retire and declare  $H_i$  as the true hypothesis, i.e.,  $V^*(\boldsymbol{\nu}) = L(1 - \nu_i)$ . Without loss of generality, consider  $i = M$ , hence

$$\boldsymbol{\nu} = \left[ \frac{\alpha(L, M)}{M-1}, \dots, \frac{\alpha(L, M)}{M-1}, 1 - \alpha(L, M) \right].$$

In the proof of Theorem 4.3.2, we have seen that  $V^*(\boldsymbol{\nu}) \geq G(\boldsymbol{\nu}) = \alpha(L, M)L = (1 - \nu_M)L$ . Furthermore, DP equation (3.4) implies that  $V^*(\boldsymbol{\nu}) \leq \min_{j \in \Omega_M} (1 - \nu_j)L = (1 - \nu_M)L$ . Therefore,  $V^*(\boldsymbol{\nu}) = (1 - \nu_M)L$  and the proof is complete.

□

*Proof of Corollary 4.5.2.* From Theorem 4.3.3 and by setting  $\delta = \frac{1}{\log 2ML}$ , we obtain

$$\begin{aligned} \underline{V}_3(\boldsymbol{\rho}) = & \left[ \frac{H(\boldsymbol{\rho}) - H([\frac{1}{\log 2ML}, 1 - \frac{1}{\log 2ML}]) - \frac{\log(M-1)}{\log 2ML}}{I_{\max}(M)} \right. \\ & \left. + \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \log ML - \xi_M}{D_{\max}(M)} \mathbf{1}_{\{\max_{i \in \Omega_M} \rho_i \leq 1 - \frac{1}{\log 2ML}\}} - K'_3 \right]^+, \end{aligned} \quad (4.94)$$

where  $K'_3$  is a positive constant independent of  $L$  and  $M$  whose closed-form is given in (4.93).

Note that for  $M \geq 2$  and  $L > 1$  which is the region of interest for these parameters,  $\log 2ML \geq 2$  and  $1 - \frac{1}{\log 2ML} \geq \frac{1}{2}$ . Thus, for the uniform prior  $\boldsymbol{\rho} = \boldsymbol{\rho}_{u,M}$ , we have

$$\max_{i \in \Omega_M} \rho_i = \frac{1}{M} \leq \frac{1}{2} \leq 1 - \frac{1}{\log 2ML}.$$

Thus, under the uniform prior, (4.94) simplifies to

$$\begin{aligned} \underline{V}_3(\boldsymbol{\rho}_{u,M}) = & \left[ \frac{\log M - H([\frac{1}{\log 2ML}, 1 - \frac{1}{\log 2ML}]) - \frac{\log(M-1)}{\log 2ML}}{I_{\max}(M)} \right. \\ & \left. + \frac{\log \frac{1-L^{-1}}{L^{-1}}}{D_{\max}(M)} - \frac{\log \log ML + \xi_M}{D_{\max}(M)} - K'_3 \right]^+ \\ \geq & \left[ \frac{\log M - 2}{\bar{I}_{\max}} + \frac{\log \frac{1-L^{-1}}{L^{-1}}}{\bar{D}_{\max}} - \frac{\log \log LM + \xi}{\underline{D}_{\max}} - K'_3 \right]^+. \end{aligned}$$

□

#### 4.7.7 Proof of Proposition 4.2.2

Recall that  $\rho_i(n)$  denotes the posterior belief about hypothesis  $H_i$  after  $n$  observations. Let  $\tau, \tau_i, i \in \Omega_M$ , be Markov stopping times defined as follows:

$$\tau := \min \left\{ n : \min_{j \in \Omega_M} \{1 - \rho_j(n)\} \leq L^{-1} \right\}, \quad (4.95)$$

$$\tau_i := \min \left\{ n : \min_{j \neq i} \frac{\rho_i(n)}{\rho_j(n)} \geq \frac{1 - L^{-1}}{L^{-1}/(M-1)} \right\}. \quad (4.96)$$

From (3.1), total cost under policy  $\pi_{SA}$  can be written as

$$\begin{aligned}
V_{\pi_{SA}}(\boldsymbol{\rho}) &= \mathbb{E}_{\pi_{SA}}[\tau + \min_{j \in \Omega_M} (1 - \rho_j(\tau))L] \\
&\leq \mathbb{E}_{\pi_{SA}}[\tau] + 1 \\
&\leq \sum_{i=1}^M \rho_i \mathbb{E}_{\pi_{SA}}[\tau_i | \theta = i] + 1,
\end{aligned} \tag{4.97}$$

where the last inequality follows from the fact that  $\tau \leq \tau_i, \forall i \in \Omega_M$ . For notational simplicity, subscript  $\pi_{SA}$  is dropped for the rest of the proof.

Next we find an upper bound for  $\mathbb{E}[\tau_i | \theta = i]$ ,  $i \in \Omega_M$ . Before we proceed, however, we introduce the following notations to facilitate the proof. Let

$$\begin{aligned}
T^* &:= \log L - \log \frac{\check{\rho}}{1 - \check{\rho}}, \\
\tilde{T}_i(\boldsymbol{\rho}) &:= \left[ \log \frac{\check{\rho}}{(1 - \check{\rho})/(M - 1)} - \min_{k \neq i} \log \frac{\rho_i}{\rho_k} \right]^+.
\end{aligned}$$

For arbitrary  $\iota \in (0, 1)$ , we have

$$\begin{aligned}
&\mathbb{E}[\tau_i | \theta = i] \\
&= \sum_{n=0}^{\infty} P(\{\tau_i > n\} | \theta = i) \\
&\leq 1 + \left( \frac{\tilde{T}_i(\boldsymbol{\rho})}{\tilde{I}_0(M)} + \frac{T^*}{\tilde{D}_i(M)} \right) (1 + \iota) + \sum_{n: n > \left( \frac{\tilde{T}_i(\boldsymbol{\rho})}{\tilde{I}_0(M)} + \frac{T^*}{\tilde{D}_i(M)} \right) (1 + \iota)} P(\{\tau_i > n\} | \theta = i) \\
&\stackrel{(a)}{\leq} 1 + \left( \frac{\tilde{T}_i(\boldsymbol{\rho})}{\tilde{I}_0(M)} + \frac{T^*}{\tilde{D}_i(M)} \right) (1 + \iota) + \frac{MB(\iota) e^{-b(\iota) \left( \frac{\tilde{T}_i(\boldsymbol{\rho})}{\tilde{I}_0(M)} + \frac{T^*}{\tilde{D}_i(M)} \right) (1 + \iota)}}{1 - e^{-b(\iota)}} \\
&\stackrel{(b)}{\leq} 1 + \left( \frac{\tilde{T}_i(\boldsymbol{\rho})}{\tilde{I}_0(M)} + \frac{T^*}{\tilde{D}_i(M)} \right) (1 + \iota) \\
&\quad + MB(\iota) \left( 1 + \max \left\{ 1, \frac{1}{b(\iota)} \right\} \right) e^{-b(\iota) \left( \frac{\log L - \max_{j \in \Omega_M} \log \frac{\rho_j}{1 - \rho_j}}{\max_{j \in \Omega_M} \tilde{D}_{\mu_j}(M)} \right) (1 + \iota)} \\
&\leq 1 + \left( \frac{\tilde{T}_i(\boldsymbol{\rho})}{\tilde{I}_0(M)} + \frac{T^*}{\tilde{D}_i(M)} \right) (1 + \iota) \\
&\quad + M \left( 6 + \frac{4}{\frac{\iota^2}{2(1 + \iota)^2} \left( \frac{\tilde{I}_0(M)}{2\xi_M} \right)^2 b(\iota)} \right) \left( L(1 - \max_{j \in \Omega_M} \rho_j) \right)^{-b(\iota) \frac{1 + \iota}{\xi_M}}
\end{aligned}$$

$$\begin{aligned}
&\leq \left( \frac{\log M + \log \frac{1}{\rho_i} + K_1''}{\tilde{I}_0(M)} + \frac{\log L}{\tilde{D}_i(M)} \right) (1 + \iota) \\
&\quad + M \left( 6 + \frac{1}{\left( \frac{\iota/2}{1+\iota} \right)^5 \left( \frac{\tilde{I}_0(M)}{2\xi_M} \right)^4} \right) \left( L(1 - \max_{j \in \Omega_M} \rho_j) \right)^{-\frac{\iota^3}{(1+\iota)^2} \frac{I_1^2(M)}{4\xi_M^3}} + 1,
\end{aligned} \tag{4.98}$$

where  $K_1'' := \log \frac{\check{\rho}}{1-\check{\rho}}$  is independent of  $L$  and  $M$ ; inequality (a) follows from Lemma 4.7.5 below, and  $B(\iota)$  and  $b(\iota)$  are given in (4.99); and (b) holds since  $\frac{1}{1-e^{-x}} \leq 1 + \max\{1, 1/x\}$  for  $x > 0$ , and

$$\frac{\tilde{T}_i(\boldsymbol{\rho})}{\tilde{I}_0(M)} + \frac{T^*}{\tilde{D}_i(M)} \geq \frac{\left[ \log \frac{\check{\rho}}{1-\check{\rho}} - \log \frac{\rho_i}{1-\rho_i} \right]^+}{\tilde{I}_0(M)} + \frac{T^*}{\tilde{D}_i(M)} \geq \frac{\log L - \log \frac{\rho_i}{1-\rho_i}}{\tilde{D}_i(M)}.$$

Now from (4.97), (4.98), and the fact that  $\sum_{i=1}^M \rho_i \log \frac{1}{\rho_i} = H(\boldsymbol{\rho})$ , we have the assertion of the proposition.

**Lemma 4.7.5.** *Given any  $\iota \in (0, 1)$  and for  $n > \left( \frac{\tilde{T}_i(\boldsymbol{\rho})}{\tilde{I}_0(M)} + \frac{T^*}{\tilde{D}_i(M)} \right) (1 + \iota)$ , we have  $P(\{\tau_i > n\} | \theta = i) \leq MB(\iota)e^{-b(\iota)n}$  where*

$$B(\iota) = 2 + \max \left\{ 1, \frac{1}{\frac{\iota^2}{2(1+\iota)^2} \left( \frac{\tilde{I}_0(M)}{2\xi_M} \right)^2} \right\}, \quad b(\iota) = \frac{\iota^3}{4(1+\iota)^3} \left( \frac{\tilde{I}_0(M)}{2\xi_M} \right)^2. \tag{4.99}$$

*Proof of Lemma 4.7.5.* Let  $B_{ij}(n)$  and  $\tilde{B}_{ij}(n)$  be events in the probability space defined as follows:

$$\begin{aligned}
B_{ij}(n) &:= \left\{ \log \frac{\rho_i(n)}{\rho_j(n)} < \log \frac{1 - L^{-1}}{L^{-1}/(M-1)} \right\}, \\
\tilde{B}_{ij}(n) &:= \left\{ \log \frac{\rho_i(n)}{\rho_j(n)} < \log \frac{\check{\rho}}{(1-\check{\rho})/(M-1)} \right\}.
\end{aligned}$$

We have,

$$\begin{aligned}
&P(\tilde{B}_{ij}(n) | \theta = i) \\
&= P\left( \left\{ \log \frac{\rho_i(n)}{\rho_j(n)} - \mathbb{E}[\log \frac{\rho_i(n)}{\rho_j(n)}] < \log \frac{\check{\rho}}{(1-\check{\rho})/(M-1)} - \mathbb{E}[\log \frac{\rho_i(n)}{\rho_j(n)}] \right\} | \theta = i \right) \\
&= P\left( \left\{ \log \frac{\rho_i(n)}{\rho_j(n)} - \mathbb{E}[\log \frac{\rho_i(n)}{\rho_j(n)}] < \right. \right. \\
&\quad \left. \left. \log \frac{\check{\rho}}{(1-\check{\rho})/(M-1)} - \mathbb{E} \left[ \log \frac{\rho_i}{\rho_j} + \sum_{t=0}^{n-1} \log \frac{q_i^{A(t)}}{q_j^{A(t)}} \right] \right\} | \theta = i \right)
\end{aligned}$$

$$\begin{aligned}
&\leq P\left(\left\{\log \frac{\rho_i(n)}{\rho_j(n)} - \mathbb{E}[\log \frac{\rho_i(n)}{\rho_j(n)}] < \right. \right. \\
&\quad \left. \left. \log \frac{\check{\rho}}{(1-\check{\rho})/(M-1)} - \min_{k \neq i} \log \frac{\rho_i}{\rho_k} - n\tilde{I}_0(M) \right\} | \theta = i\right) \\
&= P\left(\left\{\log \frac{\rho_i(n)}{\rho_j(n)} - \mathbb{E}[\log \frac{\rho_i(n)}{\rho_j(n)}] < \tilde{T}_i(\boldsymbol{\rho}) - n\tilde{I}_0(M) \right\} | \theta = i\right). \tag{4.100}
\end{aligned}$$

Similarly, we can show that

$$P(B_{ij}(n)|\theta = i) \leq P\left(\left\{\log \frac{\rho_i(n)}{\rho_j(n)} - \mathbb{E}[\log \frac{\rho_i(n)}{\rho_j(n)}] < T^* - (n - \check{\tau}_i)\tilde{D}_i(M) \right\} | \theta = i\right), \tag{4.101}$$

where

$$\check{\tau}_i = \min \left\{ n : \min_{j \neq i} \frac{\rho_i(n')}{\rho_j(n')} \geq \frac{\check{\rho}}{(1-\check{\rho})/(M-1)} \quad \forall n' \geq n \right\}.$$

By construction (4.96),

$$\begin{aligned}
&P(\{\tau_i > n\} | \theta = i) \\
&\leq P(\cup_{j \neq i} B_{ij}(n) | \theta = i) \\
&\leq P\left(\cup_{j \neq i} B_{ij}(n) \cap \left\{\tilde{\tau}_i \leq \frac{\tilde{T}_i(\boldsymbol{\rho})}{\tilde{I}_0(M)} + n \frac{\iota/2}{1+\iota}\right\} | \theta = i\right) \\
&\quad + P\left(\left\{\tilde{\tau}_i > \frac{\tilde{T}_i(\boldsymbol{\rho})}{\tilde{I}_0(M)} + n \frac{\iota/2}{1+\iota}\right\} | \theta = i\right) \\
&\leq \sum_{j \neq i} \left( P\left(B_{ij}(n) \cap \left\{\tilde{\tau}_i \leq \frac{\tilde{T}_i(\boldsymbol{\rho})}{\tilde{I}_0(M)} + n \frac{\iota/2}{1+\iota}\right\} | \theta = i\right) + \sum_{m: m > \frac{\tilde{T}_i(\boldsymbol{\rho})}{\tilde{I}_0(M)} + n \frac{\iota/2}{1+\iota}} P(\tilde{B}_{ij}(m) | \theta = i) \right) \\
&= \sum_{j \neq i} \left( P\left(\left\{\log \frac{\rho_i(n)}{\rho_j(n)} - \mathbb{E}[\log \frac{\rho_i(n)}{\rho_j(n)}] < \tilde{D}_i(M) \left(\frac{\tilde{T}_i(\boldsymbol{\rho})}{\tilde{I}_0(M)} + \frac{T^*}{\tilde{D}_i(M)} - n \frac{1+\iota/2}{1+\iota}\right)\right\} | \theta = i\right) \right. \\
&\quad \left. + \sum_{m: m > \frac{\tilde{T}_i(\boldsymbol{\rho})}{\tilde{I}_0(M)} + n \frac{\iota/2}{1+\iota}} P\left(\left\{\log \frac{\rho_i(m)}{\rho_j(m)} - \mathbb{E}[\log \frac{\rho_i(m)}{\rho_j(m)}] < \tilde{T}_i(\boldsymbol{\rho}) - m\tilde{I}_0(M)\right\} | \theta = i\right) \right). \tag{4.102}
\end{aligned}$$

For any  $a, \hat{a} \in \mathcal{A}_M$  and  $i, j \in \Omega_M$ , we have  $\left| \log \frac{q_i^a}{q_j^a} - \log \frac{q_i^{\hat{a}}}{q_j^{\hat{a}}} \right| \leq 2\xi_M$ . For  $k = 1, 2, \dots, n$ , let  $X_k = \log \frac{q_i^{A(k-1)}}{q_j^{A(k-1)}}$  and  $\mathbf{X} = [X_1, X_2, \dots, X_n]$ . Define the function  $f(\mathbf{X}) = \log \frac{\rho_i}{\rho_j} + \sum_{k=1}^n X_k = \log \frac{\rho_i(n)}{\rho_j(n)}$ . From (4.102) and Fact 2.7.2, and for  $n > \left( \frac{\tilde{T}_i(\boldsymbol{\rho})}{\tilde{I}_0(M)} + \frac{T^*}{\tilde{D}_i(M)} \right) (1+\iota)$ , we have

$$\begin{aligned}
& P(\{\tau_i > n\} | \theta = i) \\
& \leq M \left[ \exp \left( -2n \left( \frac{\tilde{D}_i(M)/(1+\iota)}{2\xi_M} \right)^2 \left( 1 + \iota/2 - \frac{1}{n} \left( \frac{\tilde{T}_i(\boldsymbol{\rho})}{\tilde{I}_0(M)} + \frac{T^*}{\tilde{D}_i(M)} \right) (1+\iota) \right)^2 \right) \right. \\
& \quad \left. + \sum_{m: m > \frac{\tilde{T}_i(\boldsymbol{\rho})}{\tilde{I}_0(M)} + n \frac{\iota/2}{1+\iota}} \exp \left( -2m \left( \frac{\tilde{I}_0(M)/(1+\iota/2)}{2\xi_M} \right)^2 \left( 1 + \iota/2 - \frac{1}{m} \frac{\tilde{T}_i(\boldsymbol{\rho})}{\tilde{I}_0(M)} (1+\iota/2) \right)^2 \right) \right] \\
& \leq M \left[ \exp \left( -n \frac{\iota^2}{2} \left( \frac{\tilde{D}_i(M)/(1+\iota)}{2\xi_M} \right)^2 \right) \right. \\
& \quad \left. + \sum_{m: m > \frac{\tilde{T}_i(\boldsymbol{\rho})}{\tilde{I}_0(M)} + n \frac{\iota/2}{1+\iota}} \exp \left( -m \frac{\iota^2}{2} \left( \frac{\tilde{I}_0(M)/(1+\iota/2)}{2\xi_M} \right)^2 \right) \right] \\
& \leq M \left[ \exp \left( -n \frac{\iota^2}{2} \left( \frac{\tilde{D}_i(M)/(1+\iota)}{2\xi_M} \right)^2 \right) + \frac{\exp \left( -n \frac{\iota^3}{4(1+\iota)} \left( \frac{\tilde{I}_0(M)/(1+\iota/2)}{2\xi_M} \right)^2 \right)}{1 - \exp \left( -\frac{\iota^2}{2} \left( \frac{\tilde{I}_0(M)/(1+\iota/2)}{2\xi_M} \right)^2 \right)} \right] \\
& \leq M \left[ 1 + \frac{1}{1 - \exp \left( -\frac{\iota^2}{2} \left( \frac{\tilde{I}_0(M)/(1+\iota/2)}{2\xi_M} \right)^2 \right)} \right] \exp \left( -n \frac{\iota^3}{4(1+\iota)^3} \left( \frac{\tilde{I}_0(M)}{2\xi_M} \right)^2 \right) \\
& \stackrel{(a)}{\leq} M \left[ 2 + \max \left\{ 1, \frac{1}{\frac{\iota^2}{2(1+\iota)^2} \left( \frac{\tilde{I}_0(M)}{2\xi_M} \right)^2} \right\} \right] \exp \left( -n \frac{\iota^3}{4(1+\iota)^3} \left( \frac{\tilde{I}_0(M)}{2\xi_M} \right)^2 \right),
\end{aligned}$$

where (a) follows from the fact that  $\frac{1}{1-e^{-x}} \leq 1 + \max\{1, 1/x\}$  for  $x > 0$ .  $\square$

#### 4.7.8 Proof of Proposition 4.3.1

Consider a parameter  $\epsilon > 0$  (whose exact value will be determined later). To prove this proposition, we divide the policies into two classes: 1) the ones that achieve  $\text{Pe} > \epsilon$ ; and 2) the ones that achieve  $\text{Pe} \leq \epsilon$ . We find a lower bound on the expected total cost for the policies in these two classes. The optimal expected total cost is greater than the minimum of the obtained lower bounds.

**Class 1.** Consider a policy  $\pi$  that selects the stopping time  $\tau$  such that  $\text{Pe}_\pi > \epsilon$ . The expected total cost under policy  $\pi$  is lower bounded as

$$V_\pi(\boldsymbol{\rho}) = \mathbb{E}_\pi[\tau] + L\text{Pe}_\pi > L\epsilon. \quad (4.103)$$

**Class 2.** Consider a policy  $\pi$  that selects  $\tau$  such that  $\text{Pe}_\pi \leq \epsilon$ . Next we find a lower bound for  $\mathbb{E}_\pi[\tau]$ . For all  $i \in \Omega_M$  and arbitrary  $\delta \in (0, 1)$ , let

$$T_i^*(\boldsymbol{\rho}) := \frac{\left[ (1 - \delta) \log \frac{1}{\epsilon} - \max_{k \neq i} \log \frac{\rho_i}{\rho_k} \right]^+}{\tilde{D}_i(M) + \delta}, \quad (4.104)$$

where

$$\tilde{D}_i(M) = \max_{\boldsymbol{\lambda} \in \mathbb{P}(\mathcal{A}_M)} \min_{j \neq i} \sum_{a \in \mathcal{A}_M} \lambda_a D(q_i^a \| q_j^a).$$

Under policy  $\pi$ ,

$$\begin{aligned} P(\{\tau < T_i^*(\boldsymbol{\rho})\} | \theta = i) &= P\left(\{\tau < T_i^*(\boldsymbol{\rho})\} \cap \bigcap_{j \neq i} \left\{ \frac{\rho_i(\tau)}{\rho_j(\tau)} \geq \left(\frac{1}{\epsilon}\right)^{1-\delta} \right\} | \theta = i\right) \\ &\quad + P\left(\{\tau < T_i^*(\boldsymbol{\rho})\} \cap \bigcup_{j \neq i} \left\{ \frac{\rho_i(\tau)}{\rho_j(\tau)} < \left(\frac{1}{\epsilon}\right)^{1-\delta} \right\} | \theta = i\right) \\ &\stackrel{(a)}{\leq} \frac{M\xi_M^2}{T_i^*(\boldsymbol{\rho})\delta^2} + \sum_{j \neq i} P\left(\left\{ \frac{\rho_i(\tau)}{\rho_j(\tau)} < \left(\frac{1}{\epsilon}\right)^{1-\delta} \right\} | \theta = i\right) \\ &\stackrel{(b)}{\leq} \frac{M\xi_M^2}{T_i^*(\boldsymbol{\rho})\delta^2} + (M-1)\frac{2\epsilon^\delta}{\rho_i}, \end{aligned} \quad (4.105)$$

where (a) follows from Lemma 2.7.5 and the union bound; and (b) follows from Lemma 2.7.3.

The expected total cost under policy  $\pi$  is lower bounded as

$$\begin{aligned} V_\pi(\boldsymbol{\rho}) &= \mathbb{E}_\pi[\tau] + L\text{Pe}_\pi \\ &\geq \sum_{i=1}^M \rho_i \mathbb{E}_\pi[\tau | \theta = i] \\ &= \sum_{i=1}^M \rho_i \mathbb{E}_\pi[\tau \mathbf{1}_{\{\tau \geq T_i^*(\boldsymbol{\rho})\}} + \tau \mathbf{1}_{\{\tau < T_i^*(\boldsymbol{\rho})\}} | \theta = i] \\ &\geq \sum_{i=1}^M \rho_i T_i^*(\boldsymbol{\rho}) P(\{\tau \geq T_i^*(\boldsymbol{\rho})\} | \theta = i) \\ &= \sum_{i=1}^M \rho_i T_i^*(\boldsymbol{\rho}) (1 - P(\{\tau < T_i^*(\boldsymbol{\rho})\} | \theta = i)) \\ &\stackrel{(a)}{\geq} \sum_{i=1}^M \rho_i T_i^*(\boldsymbol{\rho}) \left( 1 - \frac{M\xi_M^2}{T_i^*(\boldsymbol{\rho})\delta^2} - \frac{2M\epsilon^\delta}{\rho_i} \right) \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^M \rho_i T_i^*(\boldsymbol{\rho}) \left(1 - \frac{2M\epsilon^\delta}{\rho_i}\right) - \frac{M\xi_M^2}{\delta^2} \\
&= \sum_{i=1}^M \rho_i \frac{\left[(1-\delta) \log \frac{1}{\epsilon} - \max_{j \neq i} \log \frac{\rho_i}{\rho_j}\right]^+}{\tilde{D}_i(M) + \delta} \left(1 - \frac{2M\epsilon^\delta}{\rho_i}\right) - \frac{M\xi_M^2}{\delta^2}, \quad (4.106)
\end{aligned}$$

where (a) follows from (4.105).

It is trivial that  $V^*(\boldsymbol{\rho})$  is lower bounded by the minimum of (4.103) and (4.106). Let  $\epsilon = \frac{K' \log 2L}{L}$  where

$$K' := \max \left\{ 1, \frac{\log M}{\min_{i \in \Omega_M} \tilde{D}_i(M)} \right\}. \quad (4.107)$$

For this value of  $\epsilon$ , the minimum of (4.103) and (4.106) is achieved by (4.106), i.e.,

$$\begin{aligned}
&\sum_{i=1}^M \rho_i \frac{\left[(1-\delta) \log \frac{1}{\epsilon} - \max_{j \neq i} \log \frac{\rho_i}{\rho_j}\right]^+}{\tilde{D}_i(M) + \delta} \left(1 - \frac{2M\epsilon^\delta}{\rho_i}\right) - \frac{M\xi_M^2}{\delta^2} \\
&\leq \sum_{i=1}^M \rho_i \frac{\left[(1-\delta) \log \frac{L}{K' \log 2L} - \max_{j \neq i} \log \frac{\rho_i}{\rho_j}\right]^+}{\tilde{D}_i(M) + \delta} \\
&\leq \sum_{i=1}^M \rho_i \frac{\log L + \log \frac{1}{\rho_i}}{\tilde{D}_i(M) + \delta} \leq \frac{\log L + \sum_{i=1}^M \rho_i \log \frac{1}{\rho_i}}{\min_{i \in \Omega_M} \tilde{D}_i(M)} \\
&\leq \frac{\log L + \log M}{\min_{i \in \Omega_M} \tilde{D}_i(M)} \leq \frac{\log M (\log L + 1)}{\min_{i \in \Omega_M} \tilde{D}_i(M)} \\
&\leq K' \log 2L = L\epsilon, \quad (4.108)
\end{aligned}$$

and hence,

$$V^*(\boldsymbol{\rho}) \geq \left[ \sum_{i=1}^M \rho_i \frac{\left[(1-\delta) \log \frac{L}{K' \log 2L} - \max_{j \neq i} \log \frac{\rho_i}{\rho_j}\right]^+}{\max_{\lambda \in \mathbb{P}(\mathcal{A}_M)} \min_{j \neq i} \sum_{a \in \mathcal{A}_M} \lambda_a D(q_i^a \| q_j^a) + \delta} \left(1 - \frac{2M(\frac{K' \log 2L}{L})^\delta}{\rho_i}\right) - \frac{M\xi_M^2}{\delta^2} \right]^+.$$

#### 4.7.9 Proof of Lemma 4.5.1 and Corollaries 4.5.4–4.5.6

*Proof of Lemma 4.5.1.* Recall that  $\tilde{\tau}_{1/L} = \min\{n : \max_{j \in \Omega_M} \rho_j(n) \geq 1 - L^{-1}\}$ , and let  $\mathbb{E}[\tilde{\tau}_{1/L}^*]$  be the minimum expected sample size that can be achieved by policies with

the stopping rule  $\tilde{\tau}_{1/L}$ . We have

$$\begin{aligned}
\mathbb{E}[\tau_\epsilon^*] &\geq \mathbb{E}[\tau_\epsilon^* | \max_{j \in \Omega_M} \rho_j(\tau_\epsilon^*) \geq 1 - L^{-1}] P(\{\max_{j \in \Omega_M} \rho_j(\tau_\epsilon^*) \geq 1 - L^{-1}\}) \\
&\stackrel{(a)}{\geq} \mathbb{E}[\tau_\epsilon^* | \max_{j \in \Omega_M} \rho_j(\tau_\epsilon^*) \geq 1 - L^{-1}] (1 - \mathbb{E}[1 - \max_{j \in \Omega_M} \rho_j(\tau_\epsilon^*)]L) \\
&\stackrel{(b)}{\geq} \mathbb{E}[\tau_\epsilon^* | \max_{j \in \Omega_M} \rho_j(\tau_\epsilon^*) \geq 1 - L^{-1}] (1 - \epsilon L) \\
&\geq \mathbb{E}[\tilde{\tau}_{1/L}^*] (1 - \epsilon L),
\end{aligned} \tag{4.109}$$

where (a) follows from Markov inequality and (b) follows from the definition of  $\tau_\epsilon^*$  which implies that  $\text{Pe} = \mathbb{E}[1 - \max_{j \in \Omega_M} \rho_j(\tau_\epsilon^*)] \leq \epsilon$ .

Let  $V_L : \mathbb{P}(\Omega_M) \rightarrow \mathbb{R}_+$  be the minimal solution to the following fixed point equation:

$$V_L(\boldsymbol{\rho}) = \min \left\{ 1 + \min_{a \in \mathcal{A}_M} \{(\mathbb{T}^a V_L)(\boldsymbol{\rho})\}, \mathfrak{L}(\boldsymbol{\rho}) \right\}, \tag{4.110}$$

where

$$\mathfrak{L}(\boldsymbol{\rho}) = \begin{cases} 0 & \text{if } \min_{j \in \Omega_M} (1 - \rho_j) \leq L^{-1} \\ \infty & \text{otherwise} \end{cases}. \tag{4.111}$$

It can easily be shown that

$$\mathbb{E}[\tilde{\tau}_{1/L}^*] = V_L(\boldsymbol{\rho}(0)) \geq V^*(\boldsymbol{\rho}(0)) - 1. \tag{4.112}$$

Combining (4.109) and (4.112) completes the proof.  $\square$

*Proof of Corollary 4.5.4.* Set  $\epsilon = 2^{-Et}$ ,  $L = \frac{1}{\epsilon \log \frac{1}{\epsilon}}$ , and  $\delta = \frac{1}{\log \frac{1}{\epsilon}}$ . For  $L > \frac{\log M}{I_{\max}(M)}$  and  $t > \frac{4}{E}$ , we can use Lemma 4.5.1 and Theorem 4.3.3 to obtain the following lower bound

$$\begin{aligned}
\mathbb{E}[\tau_\epsilon^*] &\geq (1 - \frac{1}{Et}) \left[ \frac{H(\boldsymbol{\rho}(0)) - \frac{1}{Et} \log M}{I_{\max}(M)} \right. \\
&\quad \left. + \frac{\log \frac{2^{Et}}{Et} - \log Et}{D_{\max}(M)} \mathbf{1}_{\{\max_{i \in \Omega_M} \rho_i(0) < 1 - \frac{1}{Et}\}} - O(1) \right]^+.
\end{aligned}$$

For uniform prior  $\boldsymbol{\rho}(0) = [1/M, \dots, 1/M]$ , the lower bound simplifies to

$$\mathbb{E}[\tau_\epsilon^*] \geq \left[ \frac{(1 - \frac{1}{Et})^2 \log M}{I_{\max}(M)} + (1 - \frac{1}{Et}) \frac{Et - 2 \log Et}{D_{\max}(M)} - O(1) \right]^+.$$

Therefore, for any policy  $\pi$ ,

$$\frac{(1 - \frac{1}{Et})^2 \log M_\pi(t, 2^{-Et})}{\bar{I}_{\max}} + (1 - \frac{1}{Et}) \frac{Et - 2 \log Et}{\bar{D}_{\max}} - O(1) \leq t,$$

and hence,

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{1}{t} \log M_\pi(t, 2^{-Et}) &\leq \lim_{t \rightarrow \infty} \frac{1}{t} \frac{t - (1 - \frac{1}{Et}) \frac{Et - 2 \log Et}{\bar{D}_{\max}} + O(1)}{(1 - \frac{1}{Et})^2} \bar{I}_{\max} \\ &= \lim_{t \rightarrow \infty} \frac{1 - \frac{E}{\bar{D}_{\max}} + \frac{2(1 - \frac{1}{Et}) \log Et + 1}{t \bar{D}_{\max}} + O(\frac{1}{t})}{(1 - \frac{1}{Et})^2} \bar{I}_{\max} \\ &= \bar{I}_{\max} \left( 1 - \frac{E}{\bar{D}_{\max}} \right) \end{aligned} \quad (4.113)$$

$$< \bar{I}_{\max}. \quad (4.114)$$

Inequality (4.114) implies that for  $E > 0$ , no policy can achieve rates greater than  $\bar{I}_{\max}$ . Furthermore, using (4.113), the reliability function can be bounded as

$$E(R) \leq \bar{D}_{\max} \left( 1 - \frac{R}{\bar{I}_{\max}} \right).$$

Next we show that for fixed  $M$ , hence at  $R = 0$ , no policy can achieve reliability higher than  $\tilde{D}_0(M)$ . From Lemma 4.5.1 and Proposition 4.3.1, and for uniform prior  $\boldsymbol{\rho}(0) = [1/M, \dots, 1/M]$ , we obtain the following lower bound

$$\mathbb{E}[\tau_\epsilon^*] \geq (1 - \epsilon L) \left( \frac{\log L}{\tilde{D}_0(M)} - o(\log L) \right).$$

Using the above inequality, we can find a lower bound on  $\epsilon$  such that  $\mathbb{E}[\tau_\epsilon^*] \leq t$  is satisfied. More precisely, for any policy  $\pi$  we obtain

$$\text{Pe}_\pi(t, M) \geq L^{-1} \left( 1 - \frac{t}{\frac{\log L}{\tilde{D}_0(M)} - o(\log L)} \right).$$

We can select  $L = 2^{\tilde{D}_0(M)t + o(t)}$  such that it satisfies

$$\frac{1}{\text{Pe}_\pi(t, M)} \leq 2^{\tilde{D}_0(M)t + o(t)} O(t), \quad (4.115)$$

and hence,

$$\begin{aligned}
\lim_{t \rightarrow \infty} \frac{-1}{t} \log \text{Pe}_\pi(t, M) &\leq \lim_{t \rightarrow \infty} \frac{1}{t} \log \left( 2^{\tilde{D}_0(M)t + o(t)} O(t) \right) \\
&= \lim_{t \rightarrow \infty} \left( \tilde{D}_0(M) + \frac{o(t)}{t} + O\left(\frac{\log t}{t}\right) \right) \\
&= \tilde{D}_0(M).
\end{aligned} \tag{4.116}$$

The result of Corollary 4.5.4 can be strengthened to show that no policy can achieve diminishing error probability at rates higher than  $\bar{I}_{\max}$ . Here we provide the sketch of the proof.

From (4.110) and a slight modification of Lemma 3.5.1 in which  $\min_{j \in \Omega_M} (1 - \rho_j)L$  is replaced by  $\mathfrak{L}(\boldsymbol{\rho})$  (as defined in (4.111)), we can find the following lower bound for  $V_L(\boldsymbol{\rho})$ :

$$\begin{aligned}
\underline{V}_L(\boldsymbol{\rho}) := & \left[ \frac{H(\boldsymbol{\rho}) - H([\delta, 1 - \delta]) - \delta \log(M - 1)}{I_{\max}(M)} \right. \\
& \left. + \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{1-\delta}{\delta} - \xi_M}{D_{\max}(M)} \mathbf{1}_{\{\max_{i \in \Omega_M} \rho_i \leq 1-\delta\}} - \hat{K} \right]^+,
\end{aligned} \tag{4.117}$$

where

$$\begin{aligned}
\hat{K} &= \max \left\{ \frac{L^{-1} \log(M - 1) + L^{-1} \log L}{I_{\max}(M)}, \frac{L^{-1} \log(M - 1) + L^{-1} \log L + 1}{D_{\max}(M)} \right\} \\
&\leq \frac{L^{-1} \log(M - 1) + H([L^{-1}, 1 - L^{-1}])}{I_{\max}(M)} + \frac{1}{D_{\max}(M)}.
\end{aligned}$$

Combining (4.109), (4.112), and (4.117), we get

$$\mathbb{E}[\tau_\epsilon^*] \geq (1 - \epsilon L) \underline{V}_L(\boldsymbol{\rho}(0)).$$

Let  $u(t)$  be a function such that  $u(t) \rightarrow 0$  as  $t \rightarrow \infty$  but for any  $E > 0$ ,  $u(t)2^{Et} \rightarrow \infty$ . In other words,  $\frac{\log \frac{1}{u(t)}}{t} \rightarrow 0$  as  $t \rightarrow \infty$ . Set  $\epsilon = u(t)$ ,  $L = \frac{1}{\sqrt{u(t)}}$ , and  $\delta = \sqrt{u(t)}$ . We obtain the following lower bound

$$\mathbb{E}[\tau_\epsilon^*] \geq (1 - \sqrt{u(t)}) \left[ \frac{\log M - 2\sqrt{u(t)} \log M}{I_{\max}(M)} - O(1) \right]^+. \tag{4.118}$$

Therefore, for any policy  $\pi$ ,

$$\frac{(1 - 3\sqrt{u(t)} + 2u(t)) \log M_\pi(t, u(t))}{\bar{I}_{\max}} - O(1) \leq t,$$

and hence,

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log M_\pi(t, u(t)) \leq \lim_{t \rightarrow \infty} \frac{1}{t} \frac{t + O(1)}{1 - 3\sqrt{u(t)} + 2u(t)} \bar{I}_{\max} = \bar{I}_{\max}. \quad (4.119)$$

□

*Proof of Corollary 4.5.5.* By Definition 4.5.2 and Propositions 4.3.1 and 4.2.2, if policy  $\pi$  is asymptotically optimal in  $L$ ,  $V_\pi(\boldsymbol{\rho}) = \mathbb{E}_\pi[\tau] + L\text{Pe}_\pi \leq \frac{\log L}{\tilde{D}_0(M)} + o(\log L)$ . This implies that policy  $\pi$  achieves  $\text{Pe}_\pi \leq L^{-1} \left( \frac{\log L}{\tilde{D}_0(M)} + o(\log L) \right)$  with  $\mathbb{E}_\pi[\tau] \leq \frac{\log L}{\tilde{D}_0(M)} + o(\log L)$ . To satisfy  $\mathbb{E}_\pi[\tau] \leq t$ ,  $L$  can be selected as  $L = 2^{\tilde{D}_0(M)t(1-o(1))}$  where  $o(1) \rightarrow 0$  as  $t \rightarrow \infty$ . For this selection of  $L$ ,  $\text{Pe}_\pi(t, M)$  is bounded as

$$\text{Pe}_\pi(t, M) \leq 2^{-\tilde{D}_0(M)t(1-o(1))} t,$$

and by definition, the error exponent of policy  $\pi$  is

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{-1}{t} \log \text{Pe}_\pi(t, M) &\geq \lim_{t \rightarrow \infty} \frac{-1}{t} \left( \log 2^{-\tilde{D}_0(M)t(1-o(1))} + \log t \right) \\ &= \lim_{t \rightarrow \infty} \left( \tilde{D}_0(M)(1-o(1)) - \frac{\log t}{t} \right) \\ &= \tilde{D}_0(M). \end{aligned} \quad (4.120)$$

Let  $\mathbb{E}_\pi[\tau_\epsilon]$  denote the expected number of samples that policy  $\pi$  requires to achieve  $\text{Pe}_\pi \leq \epsilon$ . If policy  $\pi$  is not asymptotically optimal in  $L$ , then there exists  $\delta > 0$  such that

$$\mathbb{E}_\pi[\tau_\epsilon] > (1 + \delta) \frac{\log \frac{1}{\epsilon}}{\tilde{D}_0(M)} + o(\log \frac{1}{\epsilon}). \quad (4.121)$$

Proof is done by contradiction. Suppose  $\mathbb{E}_\pi[\tau_\epsilon] \leq (1 + \delta) \frac{\log \frac{1}{\epsilon}}{\tilde{D}_0(M)} + o(\log \frac{1}{\epsilon})$  for all  $\delta > 0$ . Then

$$\begin{aligned} V_\pi(\boldsymbol{\rho}) &= \mathbb{E}_\pi[\tau] + L\text{Pe}_\pi \leq \mathbb{E}_\pi[\tau_\epsilon] + L\epsilon|_{\epsilon=L^{-1}} \\ &\leq \inf_{\delta > 0} (1 + \delta) \frac{\log L}{\tilde{D}_0(M)} + o(\log L) = \frac{\log L}{\tilde{D}_0(M)} + o(\log L), \end{aligned} \quad (4.122)$$

which is an asymptotically optimal solution to the problem of active hypothesis testing.

Using (4.121) we can find a lower bound on  $\epsilon$  such that  $\mathbb{E}_\pi[\tau_\epsilon] \leq t$  is satisfied. More precisely, we obtain

$$\text{Pe}_\pi(t, M) > 2^{-\frac{\tilde{D}_0(M)}{1+\delta}t - o(t)},$$

and by definition, the error exponent of policy  $\pi$  is

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{-1}{t} \log \text{Pe}_\pi(t, M) &< \lim_{t \rightarrow \infty} \frac{-1}{t} \log 2^{-\frac{\tilde{D}_0(M)}{1+\delta}t - o(t)} \\ &= \lim_{t \rightarrow \infty} \left( \frac{\tilde{D}_0(M)}{1+\delta} - \frac{o(t)}{t} \right) \\ &< \tilde{D}_0(M). \end{aligned} \tag{4.123}$$

Next we provide the proof of the second part of the corollary. We show that if policy  $\pi$  is not an order optimal solution to the problem of active hypothesis testing, then it cannot achieve non-zero rate with non-zero reliability.

Recall that  $\mathbb{E}_\pi[\tau_\epsilon]$  is the expected number of samples that policy  $\pi$  requires to achieve  $\text{Pe}_\pi \leq \epsilon$ . If policy  $\pi$  is not order optimal in  $L$  and  $M$ , then  $\mathbb{E}_\pi[\tau_\epsilon]$  is either 1)  $\omega(\log M) + O(\log \frac{1}{\epsilon})$  where  $\frac{\omega(\log M)}{\log M} \rightarrow \infty$  as  $M \rightarrow \infty$ ; 2)  $O(\log M) + \omega(\log \frac{1}{\epsilon})$ ; or 3)  $\omega(\log M) + \omega(\log \frac{1}{\epsilon})$ . Proof is done by contradiction. Suppose  $\mathbb{E}_\pi[\tau_\epsilon] = O(\log M) + O(\log \frac{1}{\epsilon})$ . Then

$$V_\pi(\boldsymbol{\rho}) = \mathbb{E}_\pi[\tau] + L\text{Pe}_\pi \leq \mathbb{E}_\pi[\tau_\epsilon] + L\epsilon|_{\epsilon=L^{-1}} = O(\log M) + O(\log L),$$

which is an order optimal solution to the problem of active hypothesis testing.

- **Case 1:**  $\mathbb{E}_\pi[\tau_\epsilon] = \omega(\log M) + O(\log \frac{1}{\epsilon})$ .

Setting  $\epsilon = 2^{-Et}$  for some  $E > 0$ , we obtain from the above condition that  $\log M_\pi(t, 2^{-Et}) = o(t)$ . By definition,

$$R = \lim_{t \rightarrow \infty} \frac{1}{t} \log M_\pi(t, 2^{-Et}) = \lim_{t \rightarrow \infty} \frac{o(t)}{t} = 0.$$

- **Case 2:**  $\mathbb{E}_\pi[\tau_\epsilon] = O(\log M) + \omega(\log \frac{1}{\epsilon})$ .

Setting  $M = 2^{Rt}$  for some  $R > 0$ , we obtain from the above condition that  $-\log \text{Pe}_\pi(t, 2^{Rt}) = o(t)$ . By definition,

$$E = \lim_{t \rightarrow \infty} \frac{-1}{t} \log \text{Pe}_\pi(t, 2^{Rt}) = \lim_{t \rightarrow \infty} \frac{o(t)}{t} = 0.$$

- **Case 3:**  $\mathbb{E}_\pi[\tau_\epsilon] = \omega(\log M) + \omega(\log \frac{1}{\epsilon})$ .

Proof follows similar lines as the proof of Case 1 and 2.

□

*Proof of Corollary 4.5.6.* From Proposition 4.2.1 and for the uniform prior  $\boldsymbol{\rho}_{u,M}$ ,  $V_{\pi_{EJS}}(\boldsymbol{\rho}_{u,M}) \leq (\frac{\log M}{I_0(M)} + \frac{\log L}{D_0(M)})(1 + o(1))$ . This implies that  $\pi_{EJS}$  can achieve  $\text{Pe}_{\pi_{EJS}} \leq L^{-1}(\frac{\log M}{I_0(M)} + \frac{\log L}{D_0(M)})(1 + o(1))$  with  $\mathbb{E}_{\pi_{EJS}}[\tau] \leq (\frac{\log M}{I_0(M)} + \frac{\log L}{D_0(M)})(1 + o(1))$ . Let  $L = t2^{Et}$ . We obtain

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{1}{t} \log M_{\pi_{EJS}}(t, 2^{-Et}) &\geq \lim_{t \rightarrow \infty} \frac{t(1 - o(1)) - \frac{\log t 2^{Et}}{\underline{D}_0}}{t} \underline{I}_0 \\ &= \underline{I}_0 \left(1 - \frac{E}{\underline{D}_0}\right). \end{aligned} \quad (4.124)$$

Therefore, for any rate  $R \in [0, \underline{I}_2]$ , there exists a reliability  $E$ ,  $E \geq \underline{D}_2 \left(1 - \frac{R}{\underline{I}_2}\right)$ , such that  $\pi^*$  can achieve rate  $R$  with reliability  $E$ .

□

#### 4.7.10 Proof of Proposition 4.6.1

The proof follows closely the proof of Theorem 4.3.3. We define  $J_b(\boldsymbol{\rho}) = \max\{J'(\boldsymbol{\rho}), J''_b(\boldsymbol{\rho})\}$  where as in (4.80),

$$J'(\boldsymbol{\rho}) := \left[ \sum_{i=1}^M \rho_i \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{\rho_i}{1-\rho_i}}{D_{\max}(M)} - K'_3 \right]^+, \quad (4.125)$$

and  $J''_b$  is a slightly modified version of (4.81):

$$\begin{aligned} J''_b(\boldsymbol{\rho}) = & \left[ \frac{H(\boldsymbol{\rho}) - H([\delta, 1 - \delta]) - \delta \log(M - 1)}{I_{\max}(M)} \right. \\ & \left. + \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{1-\delta}{\delta} - b}{D_{\max}(M)} \mathbf{1}_{\{\max_{i \in \Omega_M} \rho_i \leq 1-\delta\}} - K'_3 \right]^+. \end{aligned} \quad (4.126)$$

From (4.82), for all  $\boldsymbol{\rho}$  such that  $J_b(\boldsymbol{\rho}) = 0$  or  $J_b(\boldsymbol{\rho}) = J'(\boldsymbol{\rho})$ , we have that

$$J_b(\boldsymbol{\rho}) = J'(\boldsymbol{\rho}) \leq 1 + \min_{a \in \mathcal{A}_M} (\mathbb{T}^a J')(\boldsymbol{\rho}) \leq 1 + \min_{a \in \mathcal{A}_M} (\mathbb{T}^a J_b)(\boldsymbol{\rho}). \quad (4.127)$$

On the other hand, for all  $\boldsymbol{\rho}$  such that  $J_b(\boldsymbol{\rho}) = J_b''(\boldsymbol{\rho}) > 0$ , and for any action  $a \in \mathcal{A}_M$ , we have

$$\begin{aligned} (\mathbb{T}^a J_b)(\boldsymbol{\rho}) &= \int_{\mathcal{Z}} J_b(\boldsymbol{\Phi}^a(\boldsymbol{\rho}, z)) q_{\boldsymbol{\rho}}^a(z) dz \\ &\geq \int_{\mathcal{Z}(\boldsymbol{\rho}, a, \delta)} J_b''(\boldsymbol{\Phi}^a(\boldsymbol{\rho}, z)) q_{\boldsymbol{\rho}}^a(z) dz + \int_{\mathcal{Z}^c(\boldsymbol{\rho}, a, \delta)} J'(\boldsymbol{\Phi}^a(\boldsymbol{\rho}, z)) q_{\boldsymbol{\rho}}^a(z) dz, \end{aligned} \quad (4.128)$$

where  $\mathcal{Z}(\boldsymbol{\rho}, a, \delta) := \{z \in \mathcal{Z} : \max_{i \in \Omega_M} \rho_i \frac{q_i^a(z)}{q_{\boldsymbol{\rho}}^a(z)} \leq 1 - \delta\}$  and  $\mathcal{Z}^c(\boldsymbol{\rho}, a, \delta) = \mathcal{Z} - \mathcal{Z}(\boldsymbol{\rho}, a, \delta)$ . Note that  $J_b''(\boldsymbol{\rho}) > 0$  implies that  $\max_{i \in \Omega_M} \rho_i \leq 1 - \delta$ .

Next we find lower bounds for the terms on the right-hand side of (4.128). For notational simplicity, let  $\hat{\boldsymbol{\rho}}(z) = \boldsymbol{\Phi}^a(\boldsymbol{\rho}, z)$ . By definition,  $\max_{i \in \Omega_M} \hat{\rho}_i(z) \leq 1 - \delta$  for all  $z \in \mathcal{Z}(\boldsymbol{\rho}, a, \delta)$ , and  $\max_{i \in \Omega_M} \hat{\rho}_i(z) > 1 - \delta$  for all  $z \in \mathcal{Z}^c(\boldsymbol{\rho}, a, \delta)$ . For the first term on the right-hand side of (4.128), we have

$$\begin{aligned} \int_{\mathcal{Z}(\boldsymbol{\rho}, a, \delta)} J_b''(\hat{\boldsymbol{\rho}}(z)) q_{\boldsymbol{\rho}}^a(z) dz &\geq \int_{\mathcal{Z}(\boldsymbol{\rho}, a, \delta)} \left[ \frac{H(\hat{\boldsymbol{\rho}}(z)) - H([\delta, 1 - \delta]) - \delta \log(M - 1)}{I_{\max}(M)} \right. \\ &\quad \left. + \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{1-\delta}{\delta} - b}{D_{\max}(M)} - K'_3 \right] q_{\boldsymbol{\rho}}^a(z) dz. \end{aligned} \quad (4.129)$$

The second term on the right-hand side of (4.128) is lower bounded as

$$\begin{aligned} &\int_{\mathcal{Z}^c(\boldsymbol{\rho}, a, \delta)} J'(\hat{\boldsymbol{\rho}}(z)) q_{\boldsymbol{\rho}}^a(z) dz \\ &\geq \int_{\mathcal{Z}^c(\boldsymbol{\rho}, a, \delta)} \left[ \sum_{i=1}^M \hat{\rho}_i(z) \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{\hat{\rho}_i(z)}{1-\hat{\rho}_i(z)}}{D_{\max}(M)} - K'_3 \right] q_{\boldsymbol{\rho}}^a(z) dz \\ &= \int_{\mathcal{Z}^c(\boldsymbol{\rho}, a, \delta)} \left[ \sum_{i=1}^M \hat{\rho}_i(z) \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{\rho_i}{1-\rho_i} - b}{D_{\max}(M)} - K'_3 \right] q_{\boldsymbol{\rho}}^a(z) dz \\ &\quad - \int_{\mathcal{Z}^c(\boldsymbol{\rho}, a, \delta)} \left[ \sum_{i=1}^M \hat{\rho}_i(z) \frac{\log \frac{\hat{\rho}_i(z)}{1-\hat{\rho}_i(z)} - \log \frac{\rho_i}{1-\rho_i} - b}{D_{\max}(M)} \right] q_{\boldsymbol{\rho}}^a(z) dz \\ &\stackrel{(a)}{\geq} \int_{\mathcal{Z}^c(\boldsymbol{\rho}, a, \delta)} \left[ \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{1-\delta}{\delta} - b}{D_{\max}(M)} - K'_3 \right] q_{\boldsymbol{\rho}}^a(z) dz \\ &\quad - \sum_{i=1}^M \rho_i \int_{\mathcal{Z}^c(\boldsymbol{\rho}, a, \delta)} q_i^a(z) \frac{\log \frac{q_i^a(z)}{\sum_{j \neq i} \frac{\rho_j}{1-\rho_i} q_j^a(z)} - b}{D_{\max}(M)} dz \end{aligned}$$

$$\stackrel{(b)}{\geq} \int_{\mathcal{Z}^c(\boldsymbol{\rho}, a, \delta)} \left[ \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{1-\delta}{\delta} - b}{D_{\max}(M)} - K'_3 \right] q_{\boldsymbol{\rho}}^a(z) dz - \frac{\psi_M(b)}{D_{\max}(M)}, \quad (4.130)$$

where (a) follows from the fact that  $\max_{i \in \Omega_M} \rho_i \leq 1 - \delta$  and hence,  $\log \frac{\rho_i}{1-\rho_i} \leq \log \frac{1-\delta}{\delta}$  for all  $i \in \Omega_M$ ; and (b) holds since for all  $i \in \Omega_M$  and  $a \in \mathcal{A}_M$ ,

$$\begin{aligned} & \int_{\mathcal{Z}^c(\boldsymbol{\rho}, a, \delta)} q_i^a(z) \left( \log \frac{q_i^a(z)}{\sum_{j \neq i} \frac{\rho_j}{1-\rho_i} q_j^a(z)} - b \right) dz \\ & \leq \int_{\mathcal{Z}^c(\boldsymbol{\rho}, a, \delta)} q_i^a(z) \left[ \log \frac{q_i^a(z)}{\sum_{j \neq i} \frac{\rho_j}{1-\rho_i} q_j^a(z)} \right]_b dz \\ & \leq \int_{\mathcal{Z}} q_i^a(z) \left[ \log \frac{q_i^a(z)}{\sum_{j \neq i} \frac{\rho_j}{1-\rho_i} q_j^a(z)} \right]_b dz \\ & \stackrel{(a)}{\leq} \psi_M(b), \end{aligned}$$

where (a) follows from Lemma 5 in [59].

From (4.130), and noting that  $H(\hat{\boldsymbol{\rho}}(z)) \leq H([\delta, 1 - \delta]) + \delta \log(M - 1)$  for all  $z \in \mathcal{Z}(\boldsymbol{\rho}, a, \delta)$ , we get

$$\begin{aligned} \int_{\mathcal{Z}^c(\boldsymbol{\rho}, a, \delta)} J'(\hat{\boldsymbol{\rho}}(z)) q_{\boldsymbol{\rho}}^a(z) dz & \geq \int_{\mathcal{Z}^c(\boldsymbol{\rho}, a, \delta)} \left[ \frac{H(\hat{\boldsymbol{\rho}}(z)) - H([\delta, 1 - \delta]) - \delta \log(M - 1)}{I_{\max}(M)} \right. \\ & \quad \left. + \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{1-\delta}{\delta} - b}{D_{\max}(M)} - K'_3 \right] q_{\boldsymbol{\rho}}^a(z) dz - \frac{\psi_M(b)}{D_{\max}(M)}. \end{aligned} \quad (4.131)$$

Inequality (4.128), together with (4.129) and (4.131), implies that

$$\begin{aligned} (\mathbb{T}^a J_b)(\boldsymbol{\rho}) & \geq \frac{\int_{\mathcal{Z}} H(\boldsymbol{\Phi}^a(\boldsymbol{\rho}, z)) q_{\boldsymbol{\rho}}^a(z) dz - H([\delta, 1 - \delta]) - \delta \log(M - 1)}{I_{\max}(M)} \\ & \quad + \frac{\log \frac{1-L^{-1}}{L^{-1}} - \log \frac{1-\delta}{\delta} - b}{D_{\max}(M)} - K'_2 - \frac{\psi_M(b)}{D_{\max}(M)} \\ & = J_b''(\boldsymbol{\rho}) - \frac{I(\boldsymbol{\rho}; q_{\boldsymbol{\rho}}^a)}{I_{\max}(M)} - \frac{\psi_M(b)}{D_{\max}(M)} \\ & \geq J_b''(\boldsymbol{\rho}) - 1 - \frac{\psi_M(b)}{D_{\max}(M)} \end{aligned}$$

$$\stackrel{(a)}{=} J_b(\boldsymbol{\rho}) - 1 - \frac{\psi_M(b)}{D_{\max}(M)}, \quad (4.132)$$

where (a) holds since  $\boldsymbol{\rho}$  is such that  $J_b(\boldsymbol{\rho}) = J_b''(\boldsymbol{\rho})$ .

Combining (4.127) and (4.132), we have that

$$J_b(\boldsymbol{\rho}) \leq 1 + \frac{\psi_M(b)}{D_{\max}(M)} + \min_{a \in \mathcal{A}_M} (\mathbb{T}^a J_b)(\boldsymbol{\rho}). \quad (4.133)$$

Moreover, selecting  $K'_3$  as suggested in (4.92) in the proof of Claim 4.7.5, we obtain

$$J_b(\boldsymbol{\rho}) \leq \min_{j \in \Omega_M} (1 - \rho_j) L. \quad (4.134)$$

Lemma 3.5.1, together with (4.133) and (4.134), implies that

$$(1 + \frac{\psi_M(b)}{D_{\max}(M)}) V^* \geq J_b = \max\{J', J_b''\} \geq J_b''.$$

#### 4.7.11 Proof of Proposition 4.6.2

The proof follows closely the proof of Theorem 4.2.1. However, the challenge here is that when Assumption 4.1.2 is weakened, we cannot apply Lemma 4.7.3 to obtain (4.58). To resolve this issue, we start with a modification of  $\{U(\boldsymbol{\rho}(t))\}$  with bounded jumps. In particular, let  $b > 0$  be a scalar. Define  $U'(\boldsymbol{\rho}(0)) = U(\boldsymbol{\rho}(0)) - \log \frac{1-\tilde{\rho}'}{\tilde{\rho}'}$ , where  $\tilde{\rho}' = 1 - \frac{1}{1+\max\{\log M, L\}}$ , and for all  $t \geq 0$ :

$$U'(\boldsymbol{\rho}(t+1)) = \begin{cases} U'(\boldsymbol{\rho}(t)) + U(\boldsymbol{\rho}(t+1)) - U(\boldsymbol{\rho}(t)) \\ \quad \text{if } U(\boldsymbol{\rho}(t+1)) - U(\boldsymbol{\rho}(t)) > -b - U'(\boldsymbol{\rho}(t)) \mathbf{1}_{\{U'(\boldsymbol{\rho}(t)) > 0\}} \\ U'(\boldsymbol{\rho}(t)) \mathbf{1}_{\{U'(\boldsymbol{\rho}(t)) \leq 0\}} - b \\ \quad \text{if } U(\boldsymbol{\rho}(t+1)) - U(\boldsymbol{\rho}(t)) \leq -b - U'(\boldsymbol{\rho}(t)) \mathbf{1}_{\{U'(\boldsymbol{\rho}(t)) > 0\}} \end{cases}. \quad (4.135)$$

By construction,

$$U'(\boldsymbol{\rho}(t)) \geq U(\boldsymbol{\rho}(t)) - \log \frac{1-\tilde{\rho}'}{\tilde{\rho}'}. \quad (4.136)$$

Let  $U_{\min}(t) := \min \left\{ U(\boldsymbol{\rho}(t)), \log \frac{1-\tilde{\rho}'}{\tilde{\rho}'} \right\} \leq 0$ . From (4.136),

$$-U'(\boldsymbol{\rho}(t)) \mathbf{1}_{\{U'(\boldsymbol{\rho}(t)) > 0\}} + U(\boldsymbol{\rho}(t)) \leq U_{\min}(t). \quad (4.137)$$

We have

$$\begin{aligned}
& \mathbb{E}_\pi[U'(\boldsymbol{\rho}(t+1))|\mathcal{F}(t)] - U'(\boldsymbol{\rho}(t)) \\
&= \mathbb{E}_\pi[U(\boldsymbol{\rho}(t+1))|\mathcal{F}(t)] - U(\boldsymbol{\rho}(t)) \\
&\quad + \mathbb{E}_\pi \left[ [-b - U'(\boldsymbol{\rho}(t))\mathbf{1}_{\{U'(\boldsymbol{\rho}(t))>0\}} - U(\boldsymbol{\rho}(t+1)) + U(\boldsymbol{\rho}(t))]^+ | \mathcal{F}(t) \right] \\
&\stackrel{(a)}{\leq} -\alpha + \mathbb{E}_\pi \left[ [-U(\boldsymbol{\rho}(t+1)) + U_{\min}(t)]_b | \mathcal{F}(t) \right] \\
&= -\alpha + \sum_{a \in \mathcal{A}_M} \pi(a|\boldsymbol{\rho}(t)) \int_{\mathcal{Z}} q_{\boldsymbol{\rho}(t)}^a(z) \left[ \sum_{i=1}^M \frac{\rho_i(t)q_i^a(z)}{q_{\boldsymbol{\rho}(t)}^a(z)} \log \frac{\rho_i(t)q_i^a(z)}{\sum_{j \neq i} \rho_j(t)q_j^a(z)} + U_{\min}(t) \right]_b dz \\
&\stackrel{(b)}{\leq} -\alpha + \sum_{a \in \mathcal{A}_M} \pi(a|\boldsymbol{\rho}(t)) \int_{\mathcal{Z}} \sum_{i=1}^M \rho_i(t)q_i^a(z) \left[ \log \frac{\rho_i(t)q_i^a(z)}{\sum_{j \neq i} \rho_j(t)q_j^a(z)} + U_{\min}(t) \right]_b dz \\
&\leq -\alpha + \sum_{a \in \mathcal{A}_M} \pi(a|\boldsymbol{\rho}(t)) \sum_{i=1}^M \rho_i(t) \times \\
&\quad \int_{\mathcal{Z}} q_i^a(z) \left[ \log \frac{q_i^a(z)}{\sum_{j \neq i} \frac{\rho_j(t)}{1-\rho_i(t)} q_j^a(z)} + \log \frac{\rho_i(t)}{1-\rho_i(t)} + U_{\min}(t) \right]_b dz, \quad (4.138)
\end{aligned}$$

where (a) follows from (4.56) and (4.137), and (b) follows since if  $\log \frac{\rho_i(t)q_i^a(z)}{\sum_{j \neq i} \rho_j(t)q_j^a(z)} \geq 0$  for any  $i \in \Omega_M$ , then  $\log \frac{\rho_k(t)q_k^a(z)}{\sum_{j \neq k} \rho_j(t)q_j^a(z)}$  for all  $k \neq i$ .

Next we find an upper bound for the right-hand side of (4.138). Note that if there exists  $i \in \Omega_M$  such that  $\rho_i(t) \geq \tilde{\rho}'$ , then

$$\begin{aligned}
\log \frac{\rho_i(t)}{1-\rho_i(t)} + U(\boldsymbol{\rho}(t)) &= (1-\rho_i(t)) \log \frac{\rho_i(t)}{1-\rho_i(t)} + \sum_{j \neq i} \rho_j(t) \log \frac{1-\rho_j(t)}{\rho_j(t)} \\
&\leq 1 + (1-\rho_i(t)) \log \frac{\sum_{j \neq i} (1-\rho_j(t))}{1-\rho_i(t)} \\
&\leq 2 + (1-\tilde{\rho}') \log M \\
&\leq 3.
\end{aligned}$$

On the other hand, if  $\rho_i(t) < \tilde{\rho}'$  for all  $i \in \Omega_M$ ,

$$\log \frac{\rho_i(t)}{1-\rho_i(t)} + \log \frac{1-\tilde{\rho}'}{\tilde{\rho}'} \leq 0.$$

Therefore,

$$\max_{i \in \Omega_M} \log \frac{\rho_i(t)}{1-\rho_i(t)} + U_{\min}(t) \leq 3, \quad (4.139)$$

which together with (4.138) implies that

$$\begin{aligned}
& \mathbb{E}_\pi[U'(\boldsymbol{\rho}(t+1))|\mathcal{F}(t)] - U'(\boldsymbol{\rho}(t)) \\
& \leq -\alpha + \sum_{a \in \mathcal{A}_M} \pi(a|\boldsymbol{\rho}(t)) \sum_{i=1}^M \rho_i(t) \int_{\mathcal{Z}} q_i^a(z) \left[ \log \frac{q_i^a(z)}{\sum_{j \neq i} \frac{\rho_j(t)}{1-\rho_i(t)} q_j^a(z)} + 3 \right] dz \\
& \leq -\alpha + \frac{b}{b-3} \psi_M(b-3).
\end{aligned} \tag{4.140}$$

Recall that  $\psi_M(b)$  is decreasing in  $b$  and by Assumption 4.6.3,  $\psi_M(b) \rightarrow 0$  as  $b \rightarrow \infty$ . Therefore, there exists  $b' \in (0, \infty)$  such that for all  $b \geq b'$ , we have  $\frac{b}{b-3} \psi_M(b-3) < \alpha$ .

Let  $v'$  be a Markov stopping time defined as follows:

$$v' := \min \{t : U'(\boldsymbol{\rho}(t)) \leq 0\}. \tag{4.141}$$

By construction (4.141) and from (4.136),

$$\mathbb{E}_\pi[\tau] \leq \mathbb{E}_\pi[v']. \tag{4.142}$$

To obtain an upper bound on  $\mathbb{E}_\pi[v']$  we follow steps similar to those in (4.58). For all  $b \geq b'$ ,

$$\begin{aligned}
\mathbb{E}_\pi[v'] & \leq \frac{U'(\boldsymbol{\rho}(0)) + \mathbb{E}_\pi[-U'(\boldsymbol{\rho}(v'))]}{\alpha - \frac{b}{b-3} \psi_M(b-3)} \\
& \stackrel{(a)}{\leq} \frac{U(\boldsymbol{\rho}(0)) - \log \frac{1-\tilde{\rho}'}{\tilde{\rho}'} + b}{\alpha - \frac{b}{b-3} \psi_M(b-3)} \\
& \leq \frac{H(\boldsymbol{\rho}(0)) + \max\{\log \log M, \log L\} + b}{\alpha - \frac{b}{b-3} \psi_M(b-3)},
\end{aligned} \tag{4.143}$$

where (a) follows from (4.135) and (4.141).

Inequality (4.143) together with (4.54) completes the proof of (4.38).

Chapter 4, in part, has been submitted for publication as M. Naghshvar and T. Javidi, “Active sequential hypothesis testing,” available on arXiv:1203.4626. The dissertation author was the primary investigator and author of this paper.

## Chapter 5

# Channel Coding with Feedback

This chapter considers the problem of variable-length coding over a discrete memoryless channel (DMC) with noiseless feedback. It is shown that the problem of channel coding with feedback is a special case of the active hypothesis testing. Using the results of Chapter 4, a rate–reliability test is provided which imposes some conditions on the EJS divergence between the conditional channel output distributions. Any variable-length coding scheme that satisfies the conditions of the test is guaranteed to achieve the capacity and optimal error exponent. Furthermore, it is shown that policy  $\pi_{EJS}$ , when specialized to channel coding with feedback, provides a sequential deterministic coding scheme that satisfies the conditions of the rate–reliability test and hence, achieves the capacity and optimal error exponent of any DMC. This scheme has only one phase, in contrast to all previous coding schemes in the literature which require two different phases of operation to achieve the capacity and optimal error exponent. Additionally, in this chapter, it is proved that variable-length posterior matching achieves the capacity. Furthermore, in case of a class of symmetric channels with binary inputs, the schemes of Horstein [21] and Burnashev–Zigangirov [22] are generalized, and simple deterministic one-phase schemes are proposed to achieve the capacity and optimal error exponent.

## 5.1 Introduction

In his seminal paper [12], Burnashev provided upper and lower bounds on the minimum expected number of channel uses  $\mathbb{E}[\tau_\epsilon^*]$  that are needed to convey a message (from a fixed message set of size  $M$ ) with average probability of error smaller than some  $\epsilon$  over a discrete memoryless channel (DMC) with feedback. For all code rates below the capacity of the DMC, the ratio between the upper and lower bounds approaches 1 as  $\epsilon \rightarrow 0$ . Therefore, the bounds yield the optimal reliability function

$$E(R) := \lim_{\epsilon \rightarrow 0} \frac{-\log \epsilon}{\mathbb{E}[\tau_\epsilon^*]} = C_1 \left(1 - \frac{R}{C}\right), \quad (5.1)$$

where  $C$  denotes the capacity of the channel,  $R \in [0, C]$  is the expected rate of the code, and  $C_1$  is the maximum Kullback–Leibler (KL) divergence between the conditional output distributions given any two inputs.

Burnashev proved the upper bound using a two-phase coding scheme. In the first phase, referred to as the *communication* phase, the transmitter tries to increase the decoder’s belief about the true message using a randomized coding scheme. At the end of this phase, the message with the highest posterior probability is selected as a candidate. The second phase, referred to as the *confirmation* phase, serves to verify the correctness of the output of phase one. Subsequently, in [60, 61] alternative two-phase coding schemes attaining the optimal reliability function were provided, while it was shown in [62] that Burnashev’s communication phase can be replaced with *any* capacity achieving block code. In [63], Burnashev’s reliability function was shown to be attainable using a two-phase scheme for a binary symmetric channel (BSC) with an unknown crossover probability.

In [21, 22], see also [64], a sequential, one-phase scheme for transmission over a BSC with noiseless feedback was proposed. This scheme, first proposed in [21], is briefly explained next. Each message is represented as a subinterval of size  $\frac{1}{M}$  of the unit interval. After each transmission and given the channel output, the posterior probability of all subintervals are updated. In the next time slot, the transmitter sends 0 if the true message’s corresponding subinterval is below the current median, or 1 if it is above. If the current median lies within the true

message's subinterval, then the transmitter sends 0 and 1 randomly according to weights determined by the length of the portions of the subinterval above and below the median. As the rounds of transmission proceed, the posterior probability of the true message's subinterval most likely grows larger than  $\frac{1}{2}$ , which pushes the median within the message's subinterval and thus leads to a randomized encoding. This simple one-phase scheme is known to achieve the capacity of a BSC [22], and its posterior matching extension has recently been shown to achieve the capacity of general DMCs [64].

These previous results raise the question whether having two phases of operation and randomized encoding are necessary to achieve the optimal reliability function or not. To address this question, in this chapter, the problem of channel coding with feedback is considered from a control theoretic point of view and is modeled as a special case of active hypothesis testing. Having established this connection, the results obtained in the earlier chapters are specialized to channel coding with feedback. The main contributions of this chapter are:

- The EJS divergence between the conditional output distributions with respect to the receiver's posterior probability is considered as the key performance measure of any given coding scheme. It is shown that strictly positive lower bounds on the EJS divergence provide a non-asymptotic upper bound on the expected number of channel uses necessary for a coding scheme to obtain a given (arbitrarily small) error probability.
- As a corollary, a rate–reliability test for variable-length coding schemes is proposed. Any variable-length coding scheme that satisfies the conditions of the test is guaranteed to achieve Burnashev's optimal asymptotic performance, hence, the capacity and optimal error exponent of the channel.
- A deterministic one-phase coding scheme is proposed (by specializing policy  $\pi_{EJS}$  to the problem of channel coding with feedback) and it is proved that this scheme achieves the optimal reliability function. This, for the first time, shows that neither having two separate phases of operation nor randomized encoding are necessary to achieve the optimal reliability function.

The remainder of this chapter is organized as follows. In Section 5.2, we formulate the problem of channel coding with noiseless feedback. Section 5.3 provides the main results of the chapter for general DMCs: i) an EJS divergence based analysis of variable-length coding, ii) a specialization of this analysis to variable-length posterior matching, and iii) a specialization to a new deterministic one-phase coding scheme that is based on greedy maximization of the EJS divergence. In Section 5.4, we consider the special case of symmetric binary-input channels and propose simple deterministic schemes.

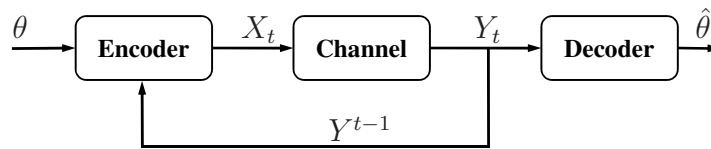
## 5.2 Coding over DMC with Noiseless Feedback

### 5.2.1 Problem Setup

Consider the problem of coding over a discrete memoryless channel (DMC) with noiseless feedback as depicted in Fig. 5.1. The DMC is described by finite input and output sets  $\mathcal{X}$  and  $\mathcal{Y}$ , and a collection of conditional probabilities  $P(Y|X)$ . To simplify notation, and without loss of generality, we assume that

$$\mathcal{X} = \{0, 1, \dots, |\mathcal{X}| - 1\}, \quad (5.2)$$

$$\mathcal{Y} = \{0, 1, \dots, |\mathcal{Y}| - 1\}. \quad (5.3)$$



**Figure 5.1:** A noisy memoryless channel with a noiseless causal feedback link.

Let  $C$  denote the *Shannon capacity* of the DMC  $P(Y|X)$  [65, p. 184]:

$$C = \max_{P_X} I(X; Y), \quad (5.4)$$

with capacity-achieving input distribution  $\pi_0^*, \pi_1^*, \dots, \pi_{|\mathcal{X}|-1}^*$ . The operational meaning of the Shannon capacity is discussed in Section 5.2.2.

Let  $C_1$  be the KL divergence between the two most distinguishable inputs of the DMC:

$$C_1 = \max_{x, x' \in \mathcal{X}} D(P(Y|X = x) \| P(Y|X = x')). \quad (5.5)$$

We also denote

$$C_2 = \max_{y \in \mathcal{Y}} \frac{\max_{x \in \mathcal{X}} P(Y = y|X = x)}{\min_{x \in \mathcal{X}} P(Y = y|X = x)}. \quad (5.6)$$

In this chapter, we assume  $C$ ,  $C_1$ ,  $C_2$  are positive and finite.<sup>1</sup>

Let  $\tau$  denote the total transmission time (or equivalently the total length of the code). The transmitter wishes to communicate a message  $\theta$  to the receiver, where the message is uniformly distributed over a message set

$$\Omega := \{1, 2, \dots, M\}. \quad (5.7)$$

To this end, it produces channel inputs  $X_t$  for  $t = 0, 1, \dots, \tau - 1$ , which it can compute as a function of the message  $\theta$  and (thanks to the feedback) also of the past channel outputs  $Y^{t-1} := [Y_0, Y_1, \dots, Y_{t-1}]$ :

$$X_t = e_t(\theta, Y^{t-1}), \quad t = 0, 1, \dots, \tau - 1, \quad (5.8)$$

for some encoding function  $e_t: \Omega \times \mathcal{Y}^t \rightarrow \mathcal{X}$ .

After observing the  $\tau$  channel outputs  $Y_0, Y_1, \dots, Y_{\tau-1}$ , the receiver guesses the message  $\theta$  as

$$\hat{\theta} = d(Y^{\tau-1}), \quad (5.9)$$

for some decoding function  $d: \mathcal{Y}^\tau \rightarrow \Omega$ . The probability of error of the scheme is thus

$$\text{Pe} := \Pr(\hat{\theta} \neq \theta).$$

In contrast to fixed-length coding where the total transmission time  $\tau$  is deterministic and known before the transmission starts, in this chapter, the focus is on variable-length coding, i.e., the case where  $\tau$  is a random stopping time

---

<sup>1</sup>It can be shown that  $C \leq C_1 \leq C_2$ . Furthermore, if  $C_1 < \infty$ , then the transition probability  $P(Y = y|X = x)$  is positive for all  $x \in \mathcal{X}$  and  $y \in \mathcal{Y}$ , which implies that  $C_2 < \infty$  as well. Therefore,  $C > 0$  and  $C_1 < \infty$  are sufficient to ensure that  $C$ ,  $C_1$ ,  $C_2$  are positive and finite.

decided at the receiver as a function of the observed channel outputs. Thanks to the noiseless feedback, the transmitter is also informed of the channel outputs and hence of the stopping time.

For a fixed DMC and for a given  $\epsilon > 0$ , the goal is to find encoding and decoding rules as in (5.8) and (5.9), and a stopping time  $\tau_\epsilon$  such that the probability of error satisfies  $\text{Pe} \leq \epsilon$  and the expected number of channel uses  $\mathbb{E}[\tau_\epsilon]$  is minimized. Let  $\mathbb{E}[\tau_\epsilon^*]$  be the minimum expected number of channel uses that can be achieved by coding schemes with the stopping rule  $\tau_\epsilon$ .

We shall often use the functions  $\{\gamma_{y^{t-1}}\}$  for  $y^{t-1} \in \mathcal{Y}^t$ ,  $t \in \{0, \dots, \tau - 1\}$ , where

$$\gamma_{y^{t-1}}: \Omega \rightarrow \mathcal{X} \quad (5.10a)$$

$$i \mapsto e_t(i, y^{t-1}) \quad (5.10b)$$

to describe the encoding process. To simplify notation and where it is clear from the context, we shall often omit the subscript  $y^{t-1}$  and simply write  $\gamma$ .

In some examples we also allow for *randomized* encoding rules. In this case the encoding is described by the *random* encoding functions  $\{\Gamma_{y^{t-1}}\}$  whose realizations  $\gamma_{y^{t-1}}$  are of the form in (5.10). Again, for notational convenience we shall omit the subscript  $y^{t-1}$  where it is clear from the context.

Note that a variable-length code is more than a single encoding function but instead is an adaptive and sequential rule that dictates the choice of (possibly random) encoding functions depending on the past channel observations and past selected encoding functions prior to the stopping time. In this paper, we refer to this adaptive and sequential rule as an encoding scheme,  $\mathfrak{c}$ , which together with the particular realization of channel outputs  $y_0, y_1, \dots, y_{\tau-2}$ , dictates the encoding functions  $\Gamma_{y^0}^{\mathfrak{c}}, \Gamma_{y^1}^{\mathfrak{c}}, \dots, \Gamma_{y^{\tau-2}}^{\mathfrak{c}}$ .

### 5.2.2 Asymptotic Bounds on Minimum Expected Length

In [12], Burnashev provided the following lower and upper bounds on the minimum expected number of channel uses  $\mathbb{E}[\tau_\epsilon^*]$  for a large class of DMCs and arbitrary  $\epsilon > 0$ .

**Fact 5.2.1** (Theorems 1 and 2 of [12]). *For any DMC with  $C > 0$  and  $C_1 < \infty$ :*

$$\mathbb{E}[\tau_\epsilon^*] \geq \left( \frac{\log M}{C} + \frac{\log \frac{1}{\epsilon}}{C_1} \right) (1 - o(1)), \quad (5.11)$$

and

$$\mathbb{E}[\tau_\epsilon^*] \leq \left( \frac{\log M}{C} + \frac{\log \frac{1}{\epsilon}}{C_1} \right) (1 + o(1)), \quad (5.12)$$

where  $o(1) \rightarrow 0$  as  $\epsilon \rightarrow 0$ .<sup>2</sup>

Inequality (5.11) was proved in [12] using a lengthy Martingale argument, and it was reproved in alternative ways in [56, 66]. In Section 5.2.3, we provide yet another alternative proof for Burnashev's lower bound utilizing dynamic programming techniques.

Burnashev proved the upper bound (5.12) using the following two-phase scheme [12]. While in the first phase (*communication phase*) the transmitter iteratively refines the receiver's belief about the true message, in the second phase (*confirmation phase*) it simply confirms whether the receiver's highest belief after the first phase corresponds to the true message. As shown in [60, 62], the specific sequential scheme in the first phase can be exchanged by any capacity-achieving block coding schemes.

Let  $\mathbb{E}_c[\tau_\epsilon]$  denote the expected number of channel uses under coding scheme  $\mathbf{c}$  to achieve  $\text{Pe}_c \leq \epsilon$ . Similar to Section 4.5.3, we define  $M_c(t, \epsilon)$  as the maximum number of messages among which coding scheme  $\mathbf{c}$  can convey the true message with  $\mathbb{E}_c[\tau_\epsilon] \leq t$ . Coding scheme  $\mathbf{c}$  is said to achieve (information) rate  $R > 0$  if there exists a function  $u(t)$  such that

$$\lim_{t \rightarrow \infty} u(t) = 0 \quad \text{and} \quad \lim_{t \rightarrow \infty} \frac{1}{t} \log M_c(t, u(t)) = R. \quad (5.13)$$

Furthermore, we say that coding scheme  $\mathbf{c}$  achieves error exponent  $E > 0$  at rate  $R > 0$  if

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log M_c(t, 2^{-Et}) = R. \quad (5.14)$$

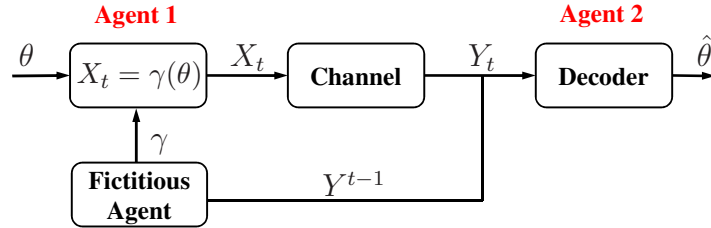
---

<sup>2</sup>If  $\epsilon \rightarrow 0$ , then  $o(1) \rightarrow 0$  regardless of whether  $M$  is fixed or  $M \rightarrow \infty$ . However, for fixed  $\epsilon$ ,  $\mathbb{E}[\tau_\epsilon^*] \approx \frac{(1-\epsilon) \log M}{C}$  and hence,  $o(1) \not\rightarrow 0$  even if  $M \rightarrow \infty$  (see [56] for more details).

The capacity of a DMC is defined as the largest rate  $R$  that is achievable over this channel. For a given rate  $R$  below capacity, the reliability function  $E(R)$  is defined as the maximum achievable error exponent at rate  $R$ . By (5.11) and (5.12), the capacity of a DMC is equal to the Shannon capacity  $C$  as defined in (5.4). Furthermore, the reliability function is given by

$$E(R) = C_1 \left( 1 - \frac{R}{C} \right), \quad R \in (0, C). \quad (5.15)$$

### 5.2.3 Stochastic Control View



**Figure 5.2:** Two-agent problem with common and private observations from the point of view of the fictitious agent.

The problem of variable-length coding with noiseless feedback is a decentralized team problem with two agents (the encoder and the decoder) and non-classical information structure [67]. Appealing to [68], the problem can be interpreted to be a special case of active hypothesis testing in which a (fictitious) Bayesian decision maker is responsible to enhance his information about the correct message in a speedy manner by sequentially sampling from conditionally independent observations at the output of the channel (given the input). Here the (fictitious) decision maker has access to the channel output symbols causally (common observations) and is responsible to control the conditional distribution of the observations given the true message (private observation) by selecting encoding functions for the encoder which map the message  $\theta$  to the input symbols of the channel. Let  $\mathcal{E} := \{\gamma(\cdot) : \Omega \rightarrow \mathcal{X}\}$  be the set of all encoding functions from  $\Omega$  to  $\mathcal{X}$ . From the discussion above, the problem of channel coding with feedback can be viewed as

an active hypothesis testing with hypotheses set  $\Omega$ , action space  $\mathcal{E}$ , observation space  $\mathcal{Y}$ , and observation kernels  $\{q_i^\gamma\}_{i \in \Omega, \gamma \in \mathcal{E}}$  where  $q_i^\gamma(y) = P(Y = y | X = \gamma(i))$  for all  $y \in \mathcal{Y}$ .

Let the decision maker's belief about each possible message  $i \in \Omega$ , updated after each channel use (observation) for  $t = 0, 1, \dots, \tau - 1$ , be

$$\rho_i(t) := \Pr(\theta = i | Y^{t-1}). \quad (5.16)$$

The decision maker's posteriors about the messages collectively,

$$\boldsymbol{\rho}(t) := [\rho_1(t), \rho_2(t), \dots, \rho_M(t)], \quad (5.17)$$

form a sufficient statistics for our (fictitious) Bayesian decision maker. Furthermore, this (fictitious) decision maker's posterior at any time  $t$  coincides with the receiver's posterior and, thanks to the perfect feedback, is available to the transmitter. (Notice that  $\rho_i(0) = \Pr(\theta = i) = \frac{1}{M}$  denotes the receiver's initial belief of  $\theta = i$  before the transmission starts.) In other words, the selection of encoding and decoding rules as a function of belief vector  $\boldsymbol{\rho}(t) := [\rho_1(t), \rho_2(t), \dots, \rho_M(t)]$  does not incur any loss of optimality [69]. In particular, the optimal receiver produces as its guess the message with the highest posterior at time  $\tau$ , i.e.,

$$\hat{\theta} = \arg \max_{i \in \Omega} \rho_i(\tau). \quad (5.18)$$

In this chapter we focus on the following (possibly suboptimal) stopping rule. For any given coding scheme  $\mathbf{c}$ , the transmission is only stopped when one of the posteriors becomes larger than  $1 - \epsilon$ , where  $\epsilon > 0$  is the desired probability of error:

$$\tilde{\tau}_\epsilon := \min\{t : \max_{i \in \Omega} \rho_i(t) \geq 1 - \epsilon\}. \quad (5.19)$$

Let  $\mathbb{E}[\tilde{\tau}_\epsilon^*]$  denote the optimal expected length of the code with the stopping rule as given in (5.19). From the described optimal decoding rule of (5.18), the constraint on the probability of error is satisfied by any coding scheme with the stopping rule (5.19):

$$\text{Pe} = \mathbb{E}[1 - \max_{i \in \Omega} \rho_i(\tilde{\tau}_\epsilon)] \leq \epsilon,$$

hence, by construction,

$$\mathbb{E}[\tau_\epsilon^*] \leq \mathbb{E}[\tilde{\tau}_\epsilon^*]. \quad (5.20)$$

Furthermore, from (4.109) and for any scalars  $\iota \geq \epsilon > 0$ , we have

$$\mathbb{E}[\tilde{\tau}_\iota^*] \left(1 - \frac{\epsilon}{\iota}\right) \leq \mathbb{E}[\tau_\epsilon^*]. \quad (5.21)$$

Note that combining (5.21) with the following lemma when  $\iota = \frac{\epsilon}{2} \log \frac{4}{\epsilon}$  and  $\delta = \frac{1}{\log \frac{4}{\epsilon}}$  provides an alternative proof for Burnashev's converse (5.11).

**Lemma 5.2.1.** *For any  $\iota \in (0, 1)$ , and for any  $\delta \in (0, 1/2)$ ,*

$$\mathbb{E}[\tilde{\tau}_\iota^*] \geq \left[ \frac{\log M - F_M(\delta) - F_M(\iota)}{C} + \frac{\log \frac{1-\iota}{\iota} - \log \frac{1-\delta}{\delta} - \log C_2 - 1}{C_1} \right]^+$$

where  $F_M(z) := H([z, 1-z]) + z \log(M-1)$  for  $0 \leq z \leq 1$ .

*Proof of Lemma 5.2.1.* The proof simply follows from (4.112), (4.117), and the fact that  $I_{\max}(M) = C$ ,  $D_{\max}(M) = C_1$ , and  $\xi_M = \log C_2$  for the problem of channel coding with feedback.  $\square$

## 5.3 Main Result and Applications

In this section, we provide the main results of the chapter. Before we proceed, we define the following notations.

We denote the conditional probability  $P(Y|X = x)$  by  $P_x$ , for  $x \in \mathcal{X}$ . Similar to (3.19) and (3.23), given a DMC  $P(Y|X)$  and a sequence of channel outputs  $y^{t-1}$  and posteriors  $\boldsymbol{\rho}(t)$ , we denote the EJS divergence associated with an encoding function  $\gamma$  and a (possibly) randomized encoding function  $\Gamma$  by

$$EJS(\boldsymbol{\rho}(t), \gamma) := EJS(\boldsymbol{\rho}(t); P_{\gamma(1)}, \dots, P_{\gamma(M)}), \quad (5.22)$$

$$EJS(\boldsymbol{\rho}(t), \Gamma) := \sum_{\gamma \in \mathcal{E}} \Pr(\{\Gamma = \gamma\} | Y^{t-1} = y^{t-1}) EJS(\boldsymbol{\rho}(t), \gamma), \quad (5.23)$$

where recall that  $\mathcal{E}$  denotes the set of all possible encoding functions.

We first specialize Theorem 4.2.1 to the problem of channel coding with feedback.

**Theorem 5.3.1.** *Let*

$$\tilde{\rho} := 1 - \frac{1}{1 + \max\{\log M, \log \frac{1}{\epsilon}\}}. \quad (5.24)$$

*Consider a (possibly randomized) encoding scheme  $\mathbf{c}$  under which at each time  $t = 0, \dots, \tilde{\tau}_\epsilon - 1$  and for each  $y^{t-1}$  the encoding function  $\Gamma^\mathbf{c}$  satisfies*

$$EJS(\boldsymbol{\rho}(t), \Gamma^\mathbf{c}) \geq \begin{cases} R_{\min} & \text{if } \max_{i \in \Omega} \rho_i(t) < \tilde{\rho} \\ \tilde{\rho} E_{\min} & \text{otherwise} \end{cases}, \quad (5.25)$$

*for some  $R_{\min}, E_{\min} > 0$ . Then,*

$$\mathbb{E}_\mathbf{c}[\tilde{\tau}_\epsilon] \leq \left( \frac{\log M}{R_{\min}} + \frac{\log \frac{1}{\epsilon}}{E_{\min}} \right) (1 + o(1)), \quad (5.26)$$

*where  $o(1) \rightarrow 0$  as  $\epsilon \rightarrow 0$  or  $M \rightarrow \infty$ .*

Consider a coding scheme  $\mathbf{c}$  under which  $\mathbb{E}_\mathbf{c}[\tau_\epsilon] \leq \left( \frac{\log M}{R_{\min}} + \frac{\log \frac{1}{\epsilon}}{E_{\min}} \right) (1 + o(1))$  where  $0 < R_{\min} \leq C$  and  $0 < E_{\min} \leq C_1$ . Coding scheme  $\mathbf{c}$  can achieve any error exponent  $E$  at rate  $R \in (0, R_{\min})$  if

$$E \leq E_{\min} \left( 1 - \frac{R}{R_{\min}} \right). \quad (5.27)$$

The above fact together with Theorem 5.3.1 provides the following rate–reliability test:

**Corollary 5.3.1.** *Consider a DMC with  $C > 0$  and  $C_1 < \infty$ . Any variable-length coding scheme  $\mathbf{c}$  that at each time  $t$  prior to the stopping time selects a (possibly randomized) encoding function  $\Gamma^\mathbf{c}$  such that*

$$EJS(\boldsymbol{\rho}(t), \Gamma^\mathbf{c}) \geq C \quad (5.28)$$

*achieves the capacity of the channel. Furthermore, if the scheme is such that*

$$EJS(\boldsymbol{\rho}(t), \Gamma^\mathbf{c}) \geq \begin{cases} C & \text{if } \max_{i \in \Omega} \rho_i(t) < \tilde{\rho} \\ \tilde{\rho} C_1 & \text{otherwise} \end{cases}, \quad (5.29)$$

*then it also achieves the optimal reliability function of the channel.*

### 5.3.1 Variable-Length Posterior Matching

We consider a variable-length version of the coding schemes in [21, 22, 64]. At each time  $t = 0, \dots, \tilde{\tau}_\epsilon - 1$ , if  $\theta = i$  and given the posterior vector  $\boldsymbol{\rho}(t)$  the input  $X(t)$  takes value in the set

$$\mathcal{X}_i(t) := \left\{ x \in \mathcal{X} : \sum_{i'=1}^{i-1} \rho_{i'}(t) < \sum_{x' \leq x} \pi_{x'}^* \text{ and } \sum_{x' < x} \pi_{x'}^* \leq \sum_{i'=1}^i \rho_{i'}(t) \right\}$$

where each value  $x \in \mathcal{X}_i(t)$  is taken with probability

$$\begin{aligned} & \Pr(\{X(t) = x\} | \theta = i, Y^{t-1} = y^{t-1}) \\ &= \frac{\min \left\{ \sum_{i'=1}^i \rho_{i'}(t), \sum_{x' \leq x} \pi_{x'}^* \right\} - \max \left\{ \sum_{i'=1}^{i-1} \rho_{i'}(t), \sum_{x' < x} \pi_{x'}^* \right\}}{\rho_i(t)}. \end{aligned} \quad (5.30)$$

**Proposition 5.3.1.** *Under the above variable-length posterior matching encoding, for each  $t = 0, \dots, \tilde{\tau}_\epsilon - 1$  and all possible output sequences  $y^{t-1}$ :*

$$EJS(\boldsymbol{\rho}(t), \Gamma^{\text{PM}}) \geq C.$$

The proof of Proposition 5.3.1 is provided in Section 5.5.1.

**Remark 5.3.1.** By Theorem 5.3.1 and Proposition 5.3.1, under the variable-length posterior matching encoding

$$\mathbb{E}_{\Gamma^{\text{PM}}}[\tilde{\tau}_\epsilon] \leq \left( \frac{\log M + \log \frac{1}{\epsilon}}{C} \right) (1 + o(1)). \quad (5.31)$$

### 5.3.2 MaxEJS Coding

In this section, we specialize  $\pi_{EJS}$  to the problem of channel coding with feedback. This results in a new sequential deterministic coding scheme which we refer to as MaxEJS. At each time  $t = 0, \dots, \tilde{\tau}_\epsilon - 1$  and given the posterior vector  $\boldsymbol{\rho}(t)$ , MaxEJS chooses the  $\gamma^*$  that maximizes the EJS divergence:

$$\gamma^* := \arg \max_{\gamma \in \mathcal{E}} EJS(\boldsymbol{\rho}(t), \gamma). \quad (5.32)$$

**Proposition 5.3.2.** *For every  $t = 0, \dots, \tilde{\tau}_\epsilon - 1$  and all possible output sequences  $y^{t-1}$ , MaxEJS encoding satisfies*

$$EJS(\boldsymbol{\rho}(t), \gamma^*) \geq \begin{cases} C & \text{if } \max_{i \in \Omega} \rho_i(t) < \tilde{\rho} \\ \tilde{\rho}C_1 & \text{otherwise} \end{cases}. \quad (5.33)$$

The proof of this proposition is given in Section 5.5.2.

**Remark 5.3.2.** By Theorem 5.3.1 and Proposition 5.3.2, MaxEJS encoding together with the decoding and stopping rules described in (5.18) and (5.19) achieves Burnashev's optimal asymptotic performance in (5.12).

**Remark 5.3.3.** The presented deterministic one-phase sequential scheme differs from the previous schemes achieving Burnashev's optimal asymptotic performance which are randomized and have two phases [12, 60–62].

The computational complexity of the MaxEJS coding scheme could be prohibitive. In Section 5.4.2, we propose simpler coding schemes for a class of binary-input channels that achieve Burnashev's optimal asymptotic performance in (5.12).

## 5.4 Coding for Binary-Input Channels

In this section, we focus on channels with binary inputs  $\mathcal{X} = \{0, 1\}$  and with the following property

$$P(Y = y|X = 0) = P(Y = z - y|X = 1), \quad \forall y \in \mathcal{Y} \quad (5.34)$$

for some  $z \in \mathbb{R}$ .

The first attempt to address the problem of coding over a symmetric binary-input channel goes back to Horstein's coding scheme [21] over a binary symmetric channel (BSC) with crossover probability  $p \in (0, 1/2)$ . Horstein considered the message to be a point in the interval  $[0, 1]$  and suggested that to achieve the capacity of the channel, at any given time, the transmitter selects the channel input such that it informs the receiver whether the message is smaller than the median of the posterior or larger. Later, Burnashev and Zigangirov [22] presented

a similar (randomized) coding scheme for discrete message sets as in (5.7) and proved that this scheme achieves capacity.

In Section 5.4.1, we generalize the schemes of Horstein [21] and Burnashev-Zigangirov [22] for discrete message sets and arbitrary symmetric binary-input channels satisfying (5.34). In Section 5.4.2, we then improve our scheme so that it achieves Burnashev's optimal asymptotic performance in (5.12) over this class of symmetric binary-input channels.

### 5.4.1 Generalized Horstein-Burnashev-Zigangirov Scheme

We propose a simple deterministic scheme as a counterpart to Horstein's scheme for discrete message sets. For each time  $t = 0, \dots, \tilde{\tau}_\epsilon - 1$  and given the posterior vector  $\boldsymbol{\rho}(t)$ , we choose the encoding function:

$$\gamma^{\text{GHBZ}}(i) = \begin{cases} 0 & 1 \leq i \leq k^* \\ 1 & k^* < i \leq M \end{cases} \quad (5.35)$$

where

$$k^* := \arg \min_{k \in \Omega} \left| \sum_{i=1}^k \rho_i(t) - \frac{1}{2} \right|. \quad (5.36)$$

**Proposition 5.4.1.** *Consider the deterministic scheme proposed above over a binary-input DMC that satisfies (5.34). For every  $t = 0, \dots, \tilde{\tau}_\epsilon - 1$  and all possible output sequences  $y^{t-1}$ ,*

$$EJS(\boldsymbol{\rho}(t), \gamma^{\text{GHBZ}}) \geq C. \quad (5.37)$$

The proof is given in Section 5.5.3.

**Remark 5.4.1.** By Theorem 5.3.1 and Proposition 5.4.1, the described encoding satisfies

$$\mathbb{E}[\tilde{\tau}_\epsilon] \leq \left( \frac{\log M + \log \frac{1}{\epsilon}}{C} \right) (1 + o(1)). \quad (5.38)$$

Note that, when specialized to binary-input channels with uniform capacity-achieving input distribution, the variable-length posterior matching scheme of Section 5.3.1, at each time  $t = 0, \dots, \tilde{\tau}_\epsilon - 1$  and given the posterior vector  $\boldsymbol{\rho}(t)$ , chooses

encoding function  $\gamma^{\text{GHBZ}}$  with probability

$$\lambda_{\gamma^{\text{GHBZ}}} = \frac{\delta_2(t)}{\delta_1(t) + \delta_2(t)} \quad (5.39)$$

where

$$\delta_1(t) := \left| \sum_{i=1}^{k^*} \rho_i(t) - \frac{1}{2} \right|, \quad \delta_2(t) := \left| \sum_{i=1}^{k_2^*} \rho_i(t) - \frac{1}{2} \right|, \quad (5.40)$$

and

$$k_2^* := k^* - \text{sign} \left( \sum_{i=1}^{k^*} \rho_i(t) - \frac{1}{2} \right); \quad (5.41)$$

and it chooses the encoding function

$$\bar{\gamma}^{\text{GHBZ}}(i) = \begin{cases} 0 & 1 \leq i \leq k_2^* \\ 1 & k_2^* < i \leq M \end{cases} \quad (5.42)$$

with probability  $\bar{\lambda}_{\gamma^{\text{GHBZ}}} = 1 - \lambda_{\gamma^{\text{GHBZ}}}$ .

Combining Proposition 5.4.1 and Proposition 5.3.1, we have that there exists a class (a continuum) of randomized schemes that satisfy (5.37):

**Corollary 5.4.1.** *Every (randomized) encoding function  $\Gamma$  that selects  $\gamma^{\text{GHBZ}}$  with probability  $\lambda \geq \frac{\delta_2(t)}{\delta_1(t) + \delta_2(t)}$  and selects  $\bar{\gamma}^{\text{GHBZ}}$  with probability  $\bar{\lambda} = 1 - \lambda$ , satisfies (5.25) with  $R_{\min} = E_{\min} = C$ .*

This corollary provides an alternative proof that the variable-length coding scheme of Burnashev and Zigangirov [22] satisfies (5.38) over the BSC with crossover probability  $p \in (0, 1/2)$ . In fact, their scheme selects  $\gamma^{\text{GHBZ}}$  and  $\bar{\gamma}^{\text{GHBZ}}$  with probabilities  $\lambda = \frac{\nu(\delta_2(t))}{\nu(\delta_1(t)) + \nu(\delta_2(t))}$  and  $\bar{\lambda} = 1 - \lambda$ , respectively, where  $\nu(x) = \log \frac{0.5 + (1-2p)x}{0.5 - (1-2p)x}$ . We next prove that  $\frac{\nu(\delta_2(t))}{\nu(\delta_1(t)) + \nu(\delta_2(t))} \geq \frac{\delta_2(t)}{\delta_1(t) + \delta_2(t)}$ , which by Corollary 5.4.1 establishes that the Burnashev-Zigangirov scheme indeed satisfies (5.38).

Notice that  $\nu(x) = \log \left( -1 + \frac{1}{0.5 - (1-2p)x} \right)$  is convex for all  $x$  because  $p \in (0, 1/2)$ . Since also  $f: x \mapsto \frac{\nu(x)}{\nu(\delta_2(t))}$  is convex and since  $f(0) = 0$  and  $f(\delta_2(t)) = 1$ , we conclude that  $\frac{\nu(x)}{\nu(\delta_2(t))} \leq \frac{x}{\delta_2(t)}$ , for all  $x \in [0, \delta_2(t)]$ . By (5.40) and (5.41),  $0 \leq \delta_1(t) \leq \delta_2(t)$  and hence  $\frac{\nu(\delta_1(t))}{\nu(\delta_2(t))} \leq \frac{\delta_1(t)}{\delta_2(t)}$ . This immediately establishes the desired inequality  $\frac{\nu(\delta_2(t))}{\nu(\delta_1(t)) + \nu(\delta_2(t))} \geq \frac{\delta_2(t)}{\delta_1(t) + \delta_2(t)}$ .

### 5.4.2 Optimal Binary Variable-Length Codes

Motivated by the analysis above, we strive to simplify our deterministic one-phase MaxEJS scheme for the simpler symmetric binary-input channels. We propose the following encoding scheme. At each time  $t = 0, \dots, \tilde{\tau}_\epsilon - 1$  and given the posterior vector  $\boldsymbol{\rho}(t)$ , we choose the encoding function  $\gamma$  in a way that

$$0 \leq \sum_{j \in \Omega: \gamma(j)=0} \rho_j(t) - \sum_{j \in \Omega: \gamma(j)=1} \rho_j(t) \leq \rho_i(t), \quad \forall i \in \{j \in \Omega: \gamma(j) = 0\}. \quad (5.43)$$

By condition (5.43), at each  $t$ , the probabilities of sending a 0 or a 1 are approximately  $(1/2, 1/2)$  when all posteriors  $\{\rho_i(t)\}_{i \in \Omega}$  are small, and they are  $(\max_{i \in \Omega} \rho_i(t), 1 - \max_{i \in \Omega} \rho_i(t))$  when  $\max_{i \in \Omega} \rho_i(t)$  is larger than  $1/2$ .

**Proposition 5.4.2.** *If for every  $t = 0, \dots, \tilde{\tau}_\epsilon - 1$  and all possible output sequences  $y^{t-1}$  the encoding function  $\gamma$  satisfies (5.43), then*

$$EJS(\boldsymbol{\rho}(t), \gamma) \geq \begin{cases} C & \text{if } \max_{i \in \Omega} \rho_i(t) < \tilde{\rho} \\ \tilde{\rho} C_1 & \text{otherwise} \end{cases}. \quad (5.44)$$

The proof is provided in Section 5.5.4.

**Remark 5.4.2.** Applying the encoding rule described above and the decoding and stopping rules in (5.18) and (5.19) achieves Burnashev's optimal asymptotic performance in (5.12).

Next we present two algorithms that at each time  $t = 0, \dots, \tilde{\tau}_\epsilon - 1$  and for a given posterior vector  $\boldsymbol{\rho}(t)$  implement encoding functions  $\gamma$  satisfying (5.43).

**Proposition 5.4.3.** *Both Algorithms 1 and 2 satisfy condition (5.43). Algorithm 1 has computational complexity of order  $O(2^M)$  for each encoding step while Algorithm 2 has complexity of order  $O(M^2)$ .<sup>3</sup>*

The proof is given in Appendix 5.5.5.

---

<sup>3</sup>The computational complexity of Algorithm 1 is of the same order as that of MaxEJS which in each step requires to find an encoding function (among  $2^M$  choices) that maximizes the EJS divergence between the conditional output distributions. However, implementation of Algorithm 1 is simpler since it only requires linear operations instead of computing the EJS divergence (which can be computationally intensive, especially for channels with large output alphabet set).

---

**Algorithm 1:**


---

```

1   $\delta = 1.$ 
2  for  $n = 1, 2, \dots, 2^M$  do
3       $v = \text{dec2bin}(n, M)$     % binary representation of  $n$  with  $M$  digits.
4       $z = (2v - 1) \times [\rho_1(t), \rho_2(t), \dots, \rho_M(t)]^\top.$ 
5      if  $z > 0$  &&  $z < \delta$  then
6           $\delta = z.$ 
7           $\hat{v} = v.$ 
8      end
9  end
10 for  $i = 1, 2, \dots, M$  do
11      $\gamma(i) = \hat{v}_i$     %  $\hat{v}_i$  denotes  $i^{\text{th}}$  bit of  $\hat{v}.$ 
12 end

```

---



---

**Algorithm 2:**


---

```

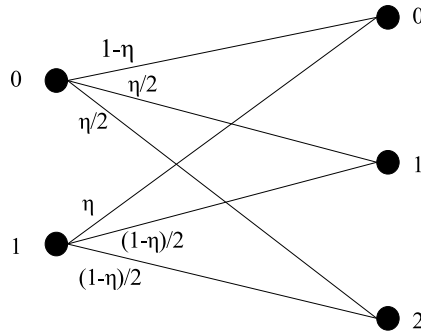
1   $S_0 = \{1, 2, \dots, M\}, S_1 = \emptyset, r_0 = 1, r_1 = 0, \rho_{\min} = 0,$  and  $\delta = 1.$ 
2  while  $\rho_{\min} < \delta$  do
3       $k = \arg \min_{i \in S_0} \rho_i(t).$ 
4       $S_0 = S_0 - \{k\}$  and  $S_1 = S_1 \cup \{k\}.$ 
5       $r_0 = r_0 - \rho_k(t)$  and  $r_1 = r_1 + \rho_k(t).$ 
6      if  $r_0 < r_1$  then
7          Swap  $S_0$  and  $S_1.$ 
8          Swap  $r_0$  and  $r_1.$ 
9      end
10      $\delta = r_0 - r_1.$ 
11      $\rho_{\min} = \min_{i \in S_0} \rho_i(t).$ 
12 end
13 for  $i = 1, 2, \dots, M$  do
14      $\gamma(i) = 0$  if  $i \in S_0,$  and  $\gamma(i) = 1$  if  $i \in S_1.$ 
15 end

```

---

**Remark 5.4.3.** In contrast to the previous one-phase sequential schemes in [21, 22, 64], the encoding processes described by Algorithms 1 and 2 here are completely deterministic. By insisting on a deterministic encoding, we can match our scheme's inputs only *approximately* to the capacity-achieving input distribution of  $(1/2, 1/2)$ . On the other hand, the proposed deterministic schemes are such that once a particular message's posterior passes a certain threshold, the transmitter assigns this message exclusively to one of the two inputs. This is critical to achieve the optimal  $E_{\min} = C_1$ .

**Remark 5.4.4.** The proofs of Propositions 5.4.2 and 5.4.3 continue to hold for those binary-input channels with uniform capacity-achieving input distribution  $\pi_0^* = \pi_1^* = 1/2$  where for ease of notation we assume that  $C_1 = D(P_0 \| P_1)$ . This class of channels includes the class of channels for which (5.34) holds, for example the binary symmetric channel (BSC) with crossover probability  $p \in (0, 1/2)$ , as well as the non-symmetric channel in Fig. 5.3 for  $\eta \in (0, 1/2)$ .



**Figure 5.3:** Binary-input ternary-output channel with capacity-achieving input distribution  $\pi_0^* = \pi_1^* = 1/2$ .

**Remark 5.4.5.** The results in Proposition 5.4.2 and Remark 5.4.2 can also be extended to the case of  $K$ -ary symmetric channel with alphabet sets  $\mathcal{X} = \mathcal{Y} = \{0, 1, \dots, K-1\}$  and transition probabilities of the form

$$P(Y = y | X = x) = \begin{cases} 1 - p & \text{if } x = y \\ \frac{p}{K-1} & \text{if } x \neq y \end{cases},$$

where  $p \in (0, \frac{K-1}{K})$ . Consider a coding scheme that at each time  $t$  prior to the stopping time chooses the encoding function  $\gamma$  in a way that if

$$\sum_{j \in \Omega: \gamma(j)=x} \rho_j(t) \geq \max \left\{ \frac{1}{K}, \sum_{j \in \Omega: \gamma(j)=x'} \rho_j(t) \right\} \text{ for any } x, x' \in \mathcal{X},$$

then

$$\sum_{j \in \Omega: \gamma(j)=x} \rho_j(t) - \sum_{j \in \Omega: \gamma(j)=x'} \rho_j(t) \leq \rho_i(t), \quad \forall i \in \{j \in \Omega: \gamma(j) = x\}.$$

This coding scheme under the decoding and stopping rules described in (5.18) and (5.19), achieves Burnashev's optimal asymptotic performance in (5.12) for the  $K$ -ary symmetric channel.

## 5.5 Proofs

The following result will be used in our proofs.

**Fact 5.5.1** (Theorem 4.5.1 in [70]). *Consider a DMC with capacity-achieving input distribution  $\pi_0^*, \pi_1^*, \dots, \pi_{|\mathcal{X}|-1}^*$ . If  $\pi_k^* > 0$ ,*

$$D\left(P(Y|X=k) \parallel \sum_{l=0}^{|\mathcal{X}|-1} \pi_l^* P(Y|X=l)\right) = C.$$

### 5.5.1 Proof of Proposition 5.3.1

Fix a time instant  $t$  and assume that  $Y^{t-1} = y^{t-1}$ . Let

$$\lambda_\gamma := \Pr(\{\Gamma^{\text{PM}} = \gamma\} | Y^{t-1} = y^{t-1}).$$

Define for each  $i \in \Omega$  and  $x \in \mathcal{X}$ :

$$\Lambda_{i,x} := \sum_{\gamma: \gamma(i)=x} \lambda_\gamma = \Pr(\{X=x\} | \theta=i, Y^{t-1}=y^{t-1}) \quad (5.45)$$

and

$$\hat{\rho}_{i,x}(t) := \rho_i(t) \Lambda_{i,x} = \Pr(\{X=x, \theta=i\} | Y^{t-1}=y^{t-1}). \quad (5.46)$$

Notice that for each  $i, j \in \Omega$ ,  $x, x' \in \mathcal{X}$ , and for a fixed posterior distribution, the various messages are mapped into inputs of the channel independently of each other and hence,

$$\sum_{\gamma: \substack{\gamma(i)=x \\ \gamma(j)=x'}} \lambda_\gamma = \Lambda_{i,x} \Lambda_{j,x'}. \quad (5.47)$$

Rearranging terms and using Jensen's inequality, we obtain

$$\begin{aligned} EJS(\boldsymbol{\rho}(t), \Gamma^{\text{PM}}) &= \sum_{\gamma \in \mathcal{E}} \lambda_\gamma \sum_{i=1}^M \rho_i(t) D\left(P_{\gamma(i)} \left\| \sum_{j \neq i} \frac{\rho_j(t)}{1 - \rho_i(t)} P_{\gamma(j)}\right.\right) \\ &= \sum_{i=1}^M \rho_i(t) \sum_{x \in \mathcal{X}} \sum_{\gamma: \gamma(i)=x} \lambda_\gamma D\left(P_x \left\| \sum_{j \neq i} \frac{\rho_j(t)}{1 - \rho_i(t)} P_{\gamma(j)}\right.\right) \\ &\geq \sum_{i=1}^M \sum_{x \in \mathcal{X}} \rho_i(t) \Lambda_{i,x} D\left(P_x \left\| \sum_{j \neq i} \frac{\rho_j(t)}{1 - \rho_i(t)} \sum_{\gamma: \gamma(i)=x} \frac{\lambda_\gamma}{\Lambda_{i,x}} P_{\gamma(j)}\right.\right) \\ &= \sum_{i=1}^M \sum_{x \in \mathcal{X}} \hat{\rho}_{i,x}(t) D\left(P_x \left\| \sum_{j \neq i} \frac{\rho_j(t)}{1 - \rho_i(t)} \sum_{x' \in \mathcal{X}} \sum_{\gamma: \substack{\gamma(i)=x \\ \gamma(j)=x'}} \frac{\lambda_\gamma}{\Lambda_{i,x}} P_{x'}\right.\right) \\ &\stackrel{(a)}{=} \sum_{i=1}^M \sum_{x \in \mathcal{X}} \hat{\rho}_{i,x}(t) D\left(P_x \left\| \frac{\sum_{j \neq i} \sum_{x' \in \mathcal{X}} \rho_j(t) \Lambda_{j,x'} P_{x'}}{1 - \rho_i(t)}\right.\right) \\ &= \sum_{i=1}^M \sum_{x \in \mathcal{X}} \hat{\rho}_{i,x}(t) D\left(P_x \left\| \frac{\sum_{x' \in \mathcal{X}} (\pi_{x'}^* P_{x'} - \hat{\rho}_{i,x'}(t) P_{x'})}{1 - \rho_i(t)}\right.\right) \\ &= \sum_{i=1}^M \sum_{x \in \mathcal{X}} \hat{\rho}_{i,x}(t) D\left(P_x \left\| \frac{\sum_{x' \in \mathcal{X}} (\pi_{x'}^* P_{x'} - \hat{\rho}_{i,x'}(t) P_{x'})}{1 - \rho_i(t)}\right.\right) \\ &\quad + \sum_{i=1}^M \sum_{x \in \mathcal{X}} \hat{\rho}_{i,x}(t) \frac{\rho_i(t)}{1 - \rho_i(t)} D\left(P_x \left\| \frac{\sum_{x'} \hat{\rho}_{i,x'}(t) P_{x'}}{\rho_i(t)}\right.\right) \\ &\quad - \sum_{i=1}^M \sum_{x \in \mathcal{X}} \hat{\rho}_{i,x}(t) \frac{\rho_i(t)}{1 - \rho_i(t)} D\left(P_x \left\| \frac{\sum_{x'} \hat{\rho}_{i,x'}(t) P_{x'}}{\rho_i(t)}\right.\right) \\ &\geq \sum_{i=1}^M \sum_{x \in \mathcal{X}} \frac{\hat{\rho}_{i,x}(t)}{1 - \rho_i(t)} D\left(P_x \left\| \sum_{x' \in \mathcal{X}} \pi_{x'}^* P_{x'}\right.\right) \\ &\quad - \sum_{i=1}^M \frac{(\rho_i(t))^2}{1 - \rho_i(t)} \sum_{x \in \mathcal{X}} \Lambda_{i,x} D\left(P_x \left\| \sum_{x' \in \mathcal{X}} \Lambda_{i,x'} P_{x'}\right.\right) \end{aligned}$$

$$\begin{aligned}
&\stackrel{(b)}{\geq} \sum_{i=1}^M \sum_{x \in \mathcal{X}} \frac{\hat{\rho}_{i,x}(t)}{1 - \rho_i(t)} C - \sum_{i=1}^M \frac{(\rho_i(t))^2}{1 - \rho_i(t)} C \\
&= \sum_{i=1}^M \frac{\rho_i(t)}{1 - \rho_i(t)} C - \sum_{i=1}^M \frac{(\rho_i(t))^2}{1 - \rho_i(t)} C \\
&= C,
\end{aligned} \tag{5.48}$$

where (a) follows from (5.47); and inequality (b) follows from Fact 5.5.1 and that  $\sum_{x \in \mathcal{X}} \Lambda_{i,x} D(P_x \| \sum_{x' \in \mathcal{X}} \Lambda_{i,x'} P_{x'})$  is the mutual information  $I(X; Y)$  between an input  $X$  with probability mass function  $\{\Lambda_{i,x}\}_{x \in \mathcal{X}}$  and the output produced by the channel (see property (3.12) of the JS divergence), and thus is smaller than the capacity  $C$ .

### 5.5.2 Proof of Proposition 5.3.2

Fix a time  $t$  and assume that  $Y^{t-1} = y^{t-1}$ . Recall that  $\Gamma^{\text{PM}}$  denotes the random encoding function of the variable-length posterior matching scheme in Section 5.3.1. By definition (5.32) and by Proposition 5.3.1,

$$EJS(\boldsymbol{\rho}(t), \gamma^*) \geq EJS(\boldsymbol{\rho}(t), \Gamma^{\text{PM}}) \geq C. \tag{5.49}$$

Now, assume that  $\max_{i \in \Omega} \rho_i(t) \geq \tilde{\rho}$  and define

$$\hat{i} := \arg \max_{i \in \Omega} \rho_i(t). \tag{5.50}$$

Then,

$$\rho_{\hat{i}}(t) \geq \tilde{\rho}. \tag{5.51}$$

Let  $x, x' \in \mathcal{X}$  be two inputs of the channel satisfying  $D(P_x \| P_{x'}) = C_1$ . Also, define the encoding function

$$\hat{\gamma}(i) := \begin{cases} x & \text{if } i = \hat{i} \\ x' & \text{otherwise.} \end{cases} \tag{5.52}$$

By definition (5.32), from (5.51), and by the selection of  $x, x'$ :

$$EJS(\boldsymbol{\rho}(t), \gamma^*) \geq EJS(\boldsymbol{\rho}(t), \hat{\gamma}) \geq \rho_{\hat{i}}(t) D(P_x \| P_{x'}) \geq \tilde{\rho} C_1.$$

### 5.5.3 Proof of Proposition 5.4.1

Let

$$\pi_x(t) := \sum_{i \in \Omega: \gamma^{\text{GHBZ}}(i)=x} \rho_i(t), \quad x \in \{0, 1\}. \quad (5.53)$$

Let

$$k_2^* := k^* - \text{sign}\left(\sum_{i=1}^{k^*} \rho_i(t) - \frac{1}{2}\right),$$

and define

$$\delta_1(t) := \left| \sum_{i=1}^{k^*} \rho_i(t) - \frac{1}{2} \right|, \quad \delta_2(t) := \left| \sum_{i=1}^{k_2^*} \rho_i(t) - \frac{1}{2} \right|.$$

Suppose  $\sum_{i=1}^{k^*} \rho_i(t) - \frac{1}{2} < 0$  which implies that  $k_2^* = k^* + 1$ . Note that by definition,  $\pi_0(t) = \frac{1}{2} - \delta_1(t)$ ,  $\rho_{k_2^*}(t) = \delta_1(t) + \delta_2(t)$ , and  $\pi_1(t) = \frac{1}{2} + \delta_1(t)$ . In this case, the EJS divergence is bounded as

$$\begin{aligned} EJS(\boldsymbol{\rho}(t), \gamma^{\text{GHBZ}}) &= \sum_{i=1}^{k^*} \rho_i(t) D\left(P_0 \left\| \frac{\pi_0(t) - \rho_i(t)}{1 - \rho_i(t)} P_0 + \frac{\pi_1(t)}{1 - \rho_i(t)} P_1 \right\| \right) \\ &\quad + \rho_{k_2^*}(t) D\left(P_1 \left\| \frac{\pi_0(t)}{1 - \rho_{k_2^*}(t)} P_0 + \frac{\pi_1(t) - \rho_{k_2^*}(t)}{1 - \rho_{k_2^*}(t)} P_1 \right\| \right) \\ &\quad + \sum_{i=k_2^*+1}^M \rho_i(t) D\left(P_1 \left\| \frac{\pi_0(t)}{1 - \rho_i(t)} P_0 + \frac{\pi_1(t) - \rho_i(t)}{1 - \rho_i(t)} P_1 \right\| \right) \\ &\stackrel{(a)}{\geq} \pi_0(t) D\left(P_0 \left\| \pi_0(t) P_0 + \pi_1(t) P_1 \right\| \right) \\ &\quad + \rho_{k_2^*}(t) D\left(P_1 \left\| \frac{1}{2} P_0 + \frac{1}{2} P_1 \right\| \right) \\ &\quad + (\pi_1(t) - \rho_{k_2^*}(t)) D\left(P_1 \left\| \pi_0(t) P_0 + \pi_1(t) P_1 \right\| \right) \\ &\stackrel{(b)}{=} \pi_0(t) D\left(P_0 \left\| \pi_0(t) P_0 + \pi_1(t) P_1 \right\| \right) \\ &\quad + \rho_{k_2^*}(t) D\left(P_0 \left\| \frac{1}{2} P_0 + \frac{1}{2} P_1 \right\| \right) \\ &\quad + (\pi_1(t) - \rho_{k_2^*}(t)) D\left(P_0 \left\| \pi_1(t) P_0 + \pi_0(t) P_1 \right\| \right) \\ &\stackrel{(c)}{\geq} D\left(P_0 \left\| \frac{1}{2} P_0 + \frac{1}{2} P_1 \right\| \right) \\ &= C, \end{aligned}$$

where (a) follows from the facts that  $\frac{\pi_0(t) - \rho_i(t)}{1 - \rho_i(t)} \leq \pi_0(t)$ ,  $\frac{\pi_1(t) - \rho_{k_2^*}(t)}{1 - \rho_{k_2^*}(t)} \leq \frac{1}{2}$ ,  $\frac{\pi_1(t) - \rho_i(t)}{1 - \rho_i(t)} \leq \pi_1(t)$ , and by Lemma 3.3.1; (b) holds because of condition (5.34); and (c) follows from the facts that KL divergence is convex,  $(\pi_0(t))^2 + \frac{1}{2}\rho_{k_2^*}(t) + (\pi_1(t) - \rho_{k_2^*}(t))\pi_1(t) = \frac{1}{2} + \delta_1(t)(\delta_1(t) - \delta_2(t)) \leq \frac{1}{2}$ , and by Lemma 3.3.1.

The proof for the case  $\sum_{i=1}^{k^*} \rho_i(t) - \frac{1}{2} \geq 0$  follows similarly.

#### 5.5.4 Proof of Proposition 5.4.2

Suppose  $\gamma$  is an encoding function that satisfies (5.43). Let

$$\pi_x(t) = \sum_{i \in \Omega: \gamma(i)=x} \rho_i(t) \quad \text{for } x \in \{0, 1\},$$

and define  $\delta(t) = \pi_0(t) - \pi_1(t)$ . From (5.43),

$$0 \leq \delta(t) \leq \rho_i(t), \quad \forall i \in \{j \in \Omega: \gamma(j) = 0\}. \quad (5.54)$$

We have

$$\begin{aligned} EJS(\boldsymbol{\rho}(t), \gamma) &= \sum_{i=1}^M \rho_i(t) D\left(P_{\gamma(i)} \left\| \sum_{j \neq i} \frac{\rho_j(t)}{1 - \rho_i(t)} P_{\gamma(j)}\right.\right) \\ &= \sum_{i \in \Omega: \gamma(i)=0} \rho_i(t) D\left(P_0 \left\| \frac{\pi_0(t) - \rho_i(t)}{1 - \rho_i(t)} P_0 + \frac{\pi_1(t)}{1 - \rho_i(t)} P_1\right.\right) \\ &\quad + \sum_{i \in \Omega: \gamma(i)=1} \rho_i(t) D\left(P_1 \left\| \frac{\pi_0(t)}{1 - \rho_i(t)} P_0 + \frac{\pi_1(t) - \rho_i(t)}{1 - \rho_i(t)} P_1\right.\right) \\ &\stackrel{(a)}{\geq} \sum_{i \in \Omega: \gamma(i)=0} \rho_i(t) D\left(P_0 \left\| \frac{1}{2} P_0 + \frac{1}{2} P_1\right.\right) \\ &\quad + \sum_{i \in \Omega: \gamma(i)=1} \rho_i(t) D\left(P_1 \left\| \frac{1}{2} P_0 + \frac{1}{2} P_1\right.\right) \\ &\stackrel{(b)}{=} C, \end{aligned}$$

where (a) follows from the facts that  $\pi_0(t) - \rho_i(t) \leq \pi_1(t)$  for any  $i$  with  $\gamma(i) = 0$ ,  $\pi_1(t) \leq \pi_0(t)$ , and since for two distributions  $P$  and  $Q$  and  $\alpha \in [0, 1]$ ,  $D(P \parallel \alpha P + (1 - \alpha)Q)$  is decreasing in  $\alpha$  (see Lemma 3.3.1); and (b) follows from Fact 5.5.1 and since the capacity of the channel is achieved by the uniform input distribution.

On the other hand, if  $\rho_i(t) \geq \frac{1}{2}$ , then condition (5.43) is only satisfied by the encoding function  $\hat{\gamma}$  under which  $\hat{\gamma}(\hat{i}) = 0$  and  $\hat{\gamma}(j) = 1$  for all  $j \neq \hat{i}$ . Therefore, if  $\rho_i(t) \geq \tilde{\rho}$  we obtain

$$EJS(\boldsymbol{\rho}(t), \hat{\gamma}) \geq \rho_i(t)D(P_0\|P_1) \geq \tilde{\rho}C_1.$$

### 5.5.5 Proof of Proposition 5.4.3

For any encoding function  $\gamma \in \mathcal{E}$ , let

$$\delta_\gamma(t) = \sum_{i \in \Omega: \gamma(i)=0} \rho_i(t) - \sum_{i \in \Omega: \gamma(i)=1} \rho_i(t). \quad (5.55)$$

Algorithm 1 computes  $\delta_\gamma(t)$  for all  $2^M$  encoding functions  $\gamma \in \mathcal{E}$  and selects  $\gamma^{\text{Alg1}}$  such that

$$\gamma^{\text{Alg1}} := \arg \min_{\gamma \in \mathcal{E}: \delta_\gamma(t) \geq 0} \delta_\gamma(t). \quad (5.56)$$

Next we prove by contradiction that  $\gamma^{\text{Alg1}}$  satisfies (5.43), i.e.,

$$\delta_{\gamma^{\text{Alg1}}}(t) \leq \rho_i(t), \quad \forall i \in \{j \in \Omega: \gamma^{\text{Alg1}}(j) = 0\}. \quad (5.57)$$

Suppose there exists  $k \in \Omega$  such that  $\gamma^{\text{Alg1}}(k) = 0$  and  $\rho_k(t) < \delta_{\gamma^{\text{Alg1}}}(t)$ . We consider two cases:

**Case I.**  $0 < \rho_k(t) \leq \frac{1}{2}\delta_{\gamma^{\text{Alg1}}}(t)$ :

Define the encoding function  $\hat{\gamma}_1$  as follows

$$\hat{\gamma}_1(i) = \begin{cases} 1 & \text{if } i = k \\ \gamma^{\text{Alg1}}(i) & \text{otherwise} \end{cases}. \quad (5.58)$$

We have

$$0 \leq \delta_{\hat{\gamma}_1}(t) = \delta_{\gamma^{\text{Alg1}}}(t) - 2\rho_k(t) < \delta_{\gamma^{\text{Alg1}}}(t),$$

which contradicts (5.56).

**Case II.**  $\frac{1}{2}\delta_{\gamma^{\text{Alg1}}}(t) < \rho_k(t) < \delta_{\gamma^{\text{Alg1}}}(t)$ :

Define the encoding function  $\hat{\gamma}_2$  as follows

$$\hat{\gamma}_2(i) = 1 - \hat{\gamma}_1(i), \quad \forall i \in \Omega. \quad (5.59)$$

We have

$$0 < \delta_{\hat{\gamma}_2}(t) = 2\rho_k(t) - \delta_{\gamma^{\text{Alg1}}}(t) < \delta_{\gamma^{\text{Alg1}}}(t),$$

which again contradicts (5.56).

Algorithm 2 constructs an encoding function that satisfies (5.43). Algorithm 2 terminates in at most  $M(M-1)/2$  rounds of operations, where in each round the main computational burden is to find an element of  $S_0$  with the lowest belief. Note that we do not have to search for the element with the lowest belief in each round if we sort all the beliefs once in the beginning, which has complexity order  $O(M \log M)$ .

Chapter 5, in full, is currently being prepared for submission for publication as M. Naghshvar, T. Javidi, and M. Wigger, “Extrinsic Jensen-Shannon divergence: applications to variable-length coding.” The dissertation author was the primary investigator and author of this material.

## Chapter 6

# Two-Dimensional Visual Search

Consider the problem of sequentially searching for one and only target in an image. Let the image be divided into  $M \times M$  equal sized segments where  $M$  determines the resolution of the search. The goal is to find the segment that contains the target quickly and accurately. In each step, the player can visually inspect an allowable combination of the segments, and the outcome of the inspection is noisy.

The visual search defined above is closely related to the problems of *fault detection* and *whereabouts search*. One possible search strategy for these problems is the maximum likelihood policy which inspects a segment with the highest probability of having the target. However, as the number of segments increases, the scheme becomes inefficient. In such a case, it is more intuitive to initially inspect larger areas and narrow down the search to single segments only after we have collected sufficient information supporting the presence of the target in those segments. Following this intuition, policy  $\pi_{EJS}$  is considered as a candidate search strategy and its performance is compared against that of the maximum likelihood policy via numerical and asymptotic analysis.

## 6.1 Introduction

This chapter considers the problem of sequentially looking for a target in an image, where the image is divided into  $M \times M$  equal sized segments and the goal is to find the segment that contains the target in a speedy manner while accounting for the penalty of wrong declarations.

The visual search defined above is closely related to the problems of *fault detection*, *whereabouts search*, and *group testing*. In fault detection, the objective is to determine the faulty component in a system known to have one failed component [71, 72]. In whereabouts search, the goal is to find an object which is hidden in one of  $N$  boxes; where it is usually assumed that there is no *false alarm* (also known as *false negative*), i.e., the outcome of inspecting box  $i$  is always 0 if no object is present, and is a Bernoulli random variable with a known parameter otherwise [73, 74]. In group testing, the goal is to locate the non-zero element<sup>1</sup> of a vector in  $\mathbb{R}^N$  with a possible noisy linear measurement of the vector [11, 75]. One possible search strategy for these problems is the maximum likelihood policy. In case of fault detection/whereabouts search, this policy is equivalent to one that inspects a segment with the highest probability of having the faulty component/hidden object; while in case of group testing, it is equivalent of measuring the most likely non-zero element of the vector. However, as the number of segments or the dimension of vectors,  $N$ , increases, the scheme becomes inefficient. In such a case, it is more intuitive to initially inspect larger areas and narrow down the search to single segments only after we have collected sufficient information supporting the presence of the target in those segments [76, 77]. Following this intuition and using the results of Chapters 3 and 4, policy  $\pi_{EJS}$  is considered as a candidate search strategy and its performance is compared against that of the maximum likelihood policy via numerical and asymptotic analysis.

The remainder of this chapter is organized as follows. In Section 6.2, we formulate the two-dimensional visual search problem. Section 6.3 provides heuristic policies whose performance will be investigated analytically and numerically

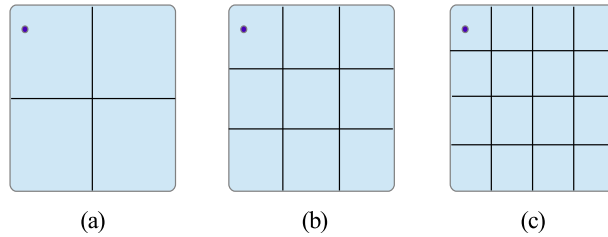
---

<sup>1</sup>Group testing with  $d > 1$  non-zero elements is also a special case of active hypothesis testing with  $\binom{N}{d}$  hypotheses (possible configurations).

in the subsequent sections. In particular, we analyze the performance of the proposed policies in Section 6.4.1, and discuss their complexity in Section 6.4.2. In Section 6.5, we compare the performance of the proposed heuristics in a numerical example.

## 6.2 Problem Formulation

In this section, we formulate the two-dimensional visual search problem. Let the image be divided into  $M \times M$  equal sized segments where  $M$  determines the resolution of the search. In other words, the location of the target is quantized to the index of the segment containing it (hence, we use the phrases “location of the target” and “the segment containing the target” interchangeably). We can think of this segmentation in a matrix form. In particular, let  $G_{ij}$  denote the segment in the  $i^{\text{th}}$  row and  $j^{\text{th}}$  column of the matrix, and  $\Omega := \bigcup_{i=1}^M \bigcup_{j=1}^M \{G_{ij}\}$  be the set of all segments. The goal is to locate the target quickly and accurately via sequentially inspecting the image.



**Figure 6.1:** Two-dimensional visual search under different resolutions.

Let  $a \subset \Omega$  be a subset of locations that can be simultaneously inspected, referred to as the inspection region hereafter, and let  $\mathcal{A} = 2^\Omega$  be the collection of all allowable inspection regions. Generalizing the model in [76], we assume that the outcome of an inspection only depends on the size of the inspection region and the resolution of the search, i.e., the outcome of inspecting region  $a$ , where  $|a| = n$ ,  $1 \leq n \leq M^2$ , is a  $\mathcal{Z}$ -valued random variable with probability density function  $f_{n,M}$  if the target location is included in the inspection region; otherwise,

it is distributed as  $\bar{f}_{n,M}$ . We assume that observation densities are known and the observations are conditionally independent over time. Let  $\theta$  be the location of the target, i.e.,  $\theta$  is the random variable that takes the value  $\theta = G_{ij}$  on the event that  $G_{ij}$  contains the target. Let  $\tau$  be the stopping time at which the player retires and declares its estimated location  $\hat{\theta}$  with the probability of error as

$$\text{Pe} := P(\hat{\theta} \neq \theta).$$

The objective is to find a sequence of inspection regions  $A(0), A(1), \dots, A(\tau - 1)$  and a stopping time  $\tau$  that collectively minimize the expected total cost

$$\mathbb{E}[\tau] + L\text{Pe}, \tag{6.1}$$

where  $L, L > 1$ , denotes the penalty of making a wrong declaration. Here we consider a uniform prior, i.e., initially  $P(\theta = \omega) = 1/M^2$  for all  $\omega \in \Omega$ , and hence the expectation is always taken with respect to this uniform prior on  $\theta$  as well as the distributions of the sequence of inspection regions and outcomes. It is clear that in this setting, the optimum declaration is the maximum a posteriori (MAP) estimate, i.e.,

$$\hat{\theta} := \arg \max_{\omega \in \Omega} P(\{\theta = \omega\} | A^{\tau-1}, Z^{\tau-1}),$$

where  $A^{\tau-1} := [A(0), A(1), \dots, A(\tau - 1)]$  and  $Z^{\tau-1} := [Z(0), Z(1), \dots, Z(\tau - 1)]$  represent respectively the vector of selected inspection regions and observation outcomes up to time  $\tau$ .

**Remark 6.2.1.** It is clear from Fig. 6.1 that a reasonable observation model should account for the equivalence between inspecting one of the segments in (a) and inspecting the 4 corner segments in (c) simultaneously. In other words, a reasonable assumption is one under which the outcome of inspecting region  $a$ ,  $|a| = n$ , depends on the ratio between the size of the inspection region,  $n$ , and the total number of segments, i.e.,  $f_{n,M} = h_{\frac{n}{M^2}}$  and  $\bar{f}_{n,M} = \bar{h}_{\frac{n}{M^2}}$  where  $\{h_r | r \in [0, 1]\}$  and  $\{\bar{h}_r | r \in [0, 1]\}$  are parametric families of distributions.

### 6.3 Heuristics

In this section, we introduce heuristic policies for selecting inspection regions. Note that the two-dimensional visual search defined above is a special case of active hypothesis testing whose observation kernels are of the following form:

$$q_{ij}^a(\cdot) = \begin{cases} f_{n,M}(\cdot) & \text{if } G_{ij} \in a, \text{ and } |a| = n \\ \bar{f}_{n,M}(\cdot) & \text{if } G_{ij} \notin a, \text{ and } |a| = n \end{cases}, \quad 1 \leq i, j \leq M, \forall a \in \mathcal{A}. \quad (6.2)$$

The information state at time  $t$  is the belief vector  $\boldsymbol{\rho}(t) = [\rho_{11}(t), \rho_{12}(t), \dots, \rho_{MM}(t)]$  where  $\rho_{ij}(t) := P(\{\theta = G_{ij}\} | A^{t-1}, Z^{t-1})$ . Accordingly, the information state space is  $\mathbb{P}(\Omega) := \{\boldsymbol{\rho} \in [0, 1]^{M^2} : \sum_{i=1}^M \sum_{j=1}^M \rho_{ij} = 1\}$ .

We consider policy  $\pi_{EJS}$  proposed in Section 3.3.2 as a candidate search strategy. From (3.19) and (6.2), the EJS divergence corresponding to inspection region  $a \in \mathcal{A}$  at belief vector  $\boldsymbol{\rho}$  can be written as:

$$EJS(\boldsymbol{\rho}, a) := \sum_{i=1}^M \sum_{j=1}^M \rho_{ij} \left[ D(f_{|a|,M} \| \frac{\eta^a - \rho_{ij}}{1 - \rho_{ij}} f_{|a|,M} + \frac{\bar{\eta}^a}{1 - \rho_{ij}} \bar{f}_{|a|,M}) \mathbf{1}_{\{G_{ij} \in a\}} \right. \\ \left. + D(\bar{f}_{|a|,M} \| \frac{\bar{\eta}^a - \rho_{ij}}{1 - \rho_{ij}} \bar{f}_{|a|,M} + \frac{\eta^a}{1 - \rho_{ij}} f_{|a|,M}) \mathbf{1}_{\{G_{ij} \notin a\}} \right], \quad (6.3)$$

where  $\eta^a = \sum_{k=1}^M \sum_{l=1}^M \rho_{kl} \mathbf{1}_{\{G_{kl} \in a\}}$ , and  $\bar{\eta}^a = 1 - \eta^a$ .

Policy  $\pi_{EJS}$  is a stationary deterministic Markov policy defined as:

- if  $\rho_{ij}(t) \geq 1 - L^{-1}$ , retire and declare that  $G_{ij}$  contains the target;
- otherwise, select inspection region  $\arg \max_{a \in \mathcal{A}} EJS(\boldsymbol{\rho}(t), a)$ .

Next we define the maximum likelihood policy  $\pi_{ML}$ . This policy, at each step, visually inspects a segment with the highest belief. Policy  $\pi_{ML}$  is a stationary deterministic Markov policy defined as:

- if  $\rho_{ij}(t) \geq 1 - L^{-1}$ , retire and declare  $G_{ij}$  contains the target;
- otherwise, select inspection region  $\{G_{ij}\}$  if  $\rho_{ij}(t) \geq \rho_{kl}(t)$ ,  $1 \leq k, l \leq M$ .

The performance of policies  $\pi_{ML}$  and  $\pi_{EJS}$  will be investigated analytically and numerically in Sections 6.4 and 6.5.

## 6.4 Main Results

In this section, we analyze the performance of the proposed heuristics and provide the main results of the chapter. We have the following assumptions:

**Assumption 6.4.1.** *For all  $M$ ,  $n \leq M^2$ , and  $z \in \mathcal{Z}$ ,*

$$f_{n,M}(z) = \bar{f}_{n,M}(b - z) \text{ for some } b \in \mathbb{R}.$$

Assumption 6.4.1 implies that given a fixed inspection region, the visual samples provide identical information regarding the presence or the absence of the target. This assumption is satisfied if, for instance, observation outcomes are modeled as a signal plus noise, where the signal component appears only if the target is included in the inspection region, and the noise distribution is symmetric with respect to its mean value.

**Assumption 6.4.2.** *For all  $M$ ,  $n < M^2$ ,  $\alpha \in [0, 1]$ , and  $\bar{\alpha} = 1 - \alpha$ ,*

$$D(f_{n,M} \| \alpha f_{n,M} + \bar{\alpha} \bar{f}_{n,M}) \geq D(f_{n+1,M} \| \alpha f_{n+1,M} + \bar{\alpha} \bar{f}_{n+1,M}).$$

Assumption 6.4.2 implies that the visual samples do not become more informative if the size of the inspection region increases.

**Assumption 6.4.3.** *For some  $\gamma > 1$ ,*

$$\sup_M \sup_{n \leq M^2} \int_{\mathcal{Z}} f_{n,M}(z) \left| \log \frac{f_{n,M}(z)}{\bar{f}_{n,M}(z)} \right|^\gamma dz < \infty.$$

Assumption 6.4.3 restricts the amount of information corresponding to each inspection region, and implies that no observation is noise-free. This assumption is important in our asymptotic analysis, and ensures that Assumptions 4.6.3 and 4.5.1 hold.

### 6.4.1 Lower and Upper Bounds

In this section, we specialize the results obtained in Chapter 4 to the problem of two-dimensional visual search, and derive a lower bound on the optimal expected total cost as well as upper bounds on the expected total cost associated with the candidate policies.

**Theorem 6.4.1** (Lower Bound). *Suppose Assumptions 6.4.1, 6.4.2, and 6.4.3 hold. Then the optimal expected total cost,  $V^*$ , is lower bounded as*

$$V^* \geq \left( \frac{\log M^2}{C_{\max}} + \frac{\log L}{E_{\max}} \right) (1 - o(1)), \quad (6.4)$$

where  $o(1) \rightarrow 0$  as  $LM \rightarrow \infty$ , and

$$C_{\max} := \lim_{M \rightarrow \infty} D(f_{1,M} \| \frac{1}{2}f_{1,M} + \frac{1}{2}\bar{f}_{1,M}),$$

$$E_{\max} := \lim_{M \rightarrow \infty} D(f_{1,M} \| \bar{f}_{1,M}).$$

The proof of Theorem 6.4.1 is provided in Section 6.6.1.

**Theorem 6.4.2** (Upper Bound). *Under Assumptions 6.4.1, 6.4.2, and 6.4.3, the expected total cost under policies  $\pi_{ML}$  and  $\pi_{EJS}$  are upper bounded as*

$$V_{\pi_{ML}} \leq \left( \frac{\log M^2}{\frac{1}{M^2}E_{\max}} + \frac{\log L}{E_{\max}} \right) (1 + o(1)), \quad (6.5)$$

$$V_{\pi_{EJS}} \leq \left( \frac{\log M^2}{C_{\min}} + \frac{\log L}{E_{\max}} \right) (1 + o(1)), \quad (6.6)$$

where  $o(1) \rightarrow 0$  as  $LM \rightarrow \infty$ , and

$$C_{\min} := \lim_{M \rightarrow \infty} D(f_{\lfloor M^2/2 \rfloor, M} \| \frac{1}{2}f_{\lfloor M^2/2 \rfloor, M} + \frac{1}{2}\bar{f}_{\lfloor M^2/2 \rfloor, M}).$$

The proof of Theorem 6.4.2 is provided in Section 6.6.2.

Next we provide examples for which Assumptions 6.4.1, 6.4.2, 6.4.3 hold.

**Example 6.4.1.** Consider the case where the observation densities are independent of the resolution and the size of the inspection region, i.e., for all  $M$  and  $n \leq M^2$ ,  $f_{n,M} = f$  and  $\bar{f}_{n,M} = \bar{f}$  where  $f(z) = \bar{f}(b - z)$  for some  $b \in \mathbb{R}$  and  $\int_{\mathcal{Z}} f(z) |\log \frac{f(z)}{\bar{f}(z)}|^\gamma dz < \infty$  for some  $\gamma > 1$ . Then Assumptions 6.4.1, 6.4.2, 6.4.3 hold and  $C_{\min} = C_{\max} = D(f \| \frac{1}{2}f + \frac{1}{2}\bar{f})$  and  $E_{\max} = D(f \| \bar{f})$ . In this case, the upper bound associated with  $\pi_{EJS}$  matches the lower bound obtained asymptotically.

**Example 6.4.2.** Let  $\{p_r\}_{r \in [0,1]}$  be an index family satisfying  $0 < p_0 \leq p_r \leq p_{r'} \leq p_1 < \frac{1}{2}$  for any  $r \leq r'$ . Consider the case of Bernoulli noise where for all  $M$  and  $n \leq M^2$ , we have  $f_{n,M} = \mathcal{B}(1 - p_{\frac{n}{M^2}})$  and  $\bar{f}_{n,M} = \mathcal{B}(p_{\frac{n}{M^2}})$ . Then Assumptions 6.4.1, 6.4.2, 6.4.3 hold and  $C_{\min} = 1 - H([p_{\frac{1}{2}}, 1 - p_{\frac{1}{2}}])$ ,  $C_{\max} = 1 - H([p_0, 1 - p_0])$ , and  $E_{\max} = (1 - 2p_0) \log \frac{1-p_0}{p_0}$ .

**Example 6.4.3.** Let  $\{\sigma_r^2\}_{r \in [0,1]}$  be an index family satisfying  $0 < \sigma_0^2 \leq \sigma_r^2 \leq \sigma_{r'}^2 \leq \sigma_1^2 < \infty$  for any  $r \leq r'$ . Consider the case of Gaussian noise where for all  $M$  and  $n \leq M^2$ ,  $f_{n,M} = N(1, \sigma_{\frac{n}{M^2}}^2)$  and  $\bar{f}_{n,M} = N(0, \sigma_{\frac{n}{M^2}}^2)$ . Then Assumptions 6.4.1, 6.4.2, 6.4.3 hold and  $C_{\min} = D(N(1, \sigma_{\frac{1}{2}}^2) \| \frac{1}{2}N(1, \sigma_{\frac{1}{2}}^2) + \frac{1}{2}N(0, \sigma_{\frac{1}{2}}^2))$ ,  $C_{\max} = D(N(1, \sigma_0^2) \| \frac{1}{2}N(1, \sigma_0^2) + \frac{1}{2}N(0, \sigma_0^2))$ , and  $E_{\max} = \frac{\log e}{2\sigma_0^2}$ .

It is straightforward to show that all assumptions hold in Example 6.4.1. Next we provide the proof of Example 6.4.2. The proof of Example 6.4.3 follows a similar approach and is omitted.

*Proof of Example 6.4.2.* Assumption 6.4.1 holds trivially since  $\bar{f}_{n,M}(z) = p_r \delta(z - 1) + (1 - p_r) \delta(z)$  and  $f_{n,M}(z) = \bar{f}_{n,M}(1 - z)$ , where  $r = \frac{n}{M^2}$  and  $\delta(\cdot)$  denotes the Dirac delta function.

Next we show that Assumption 6.4.2 holds. The proof is divided into two cases. Let  $r = \frac{n}{M^2}$  and  $r' = \frac{n+1}{M^2}$ .

- Case 1:  $\alpha = 0$ . We have

$$\begin{aligned} D(f_{n,M} \| \bar{f}_{n,M}) &= p_r \log \frac{p_r}{1 - p_r} + (1 - p_r) \log \frac{1 - p_r}{p_r} \\ &= (1 - 2p_r) \log \frac{1 - p_r}{p_r} \\ &\stackrel{(a)}{\geq} (1 - 2p_{r'}) \log \frac{1 - p_{r'}}{p_{r'}} \\ &= D(f_{n+1,M} \| \bar{f}_{n+1,M}), \end{aligned}$$

where inequality (a) holds since  $p_r \leq p_{r'}$ .

- Case 2:  $\alpha > 0$ . We have

$$\begin{aligned} &D(f_{n,M} \| \alpha f_{n,M} + \bar{\alpha} \bar{f}_{n,M}) \\ &= p_r \log \frac{p_r}{\alpha p_r + \bar{\alpha}(1 - p_r)} + (1 - p_r) \log \frac{1 - p_r}{\alpha(1 - p_r) + \bar{\alpha} p_r} \\ &= -p_r \log \left( \alpha + \bar{\alpha} \frac{1 - p_r}{p_r} \right) - (1 - p_r) \log \left( \alpha + \bar{\alpha} \frac{p_r}{1 - p_r} \right) \\ &= -\log \alpha - H_\alpha(p_r) \\ &\stackrel{(a)}{\geq} -\log \alpha - H_\alpha(p_{r'}) \\ &= D(f_{n+1,M} \| \alpha f_{n+1,M} + \bar{\alpha} \bar{f}_{n+1,M}), \end{aligned}$$

where

$$H_\alpha(p) := p \log \left( 1 + \frac{\bar{\alpha}}{\alpha} \frac{1-p}{p} \right) + (1-p) \log \left( 1 + \frac{\bar{\alpha}}{\alpha} \frac{p}{1-p} \right),$$

and inequality (a) holds since  $p_r \leq p_{r'}$  and  $H_\alpha(p)$  is concave and symmetric in  $p$  for any  $\alpha \in (0, 1]$ . Note that  $H_{\frac{1}{2}}$  is equal to the entropy function.

Finally, Assumption 6.4.3 holds since for any  $\gamma > 1$ ,

$$\sup_M \max_{n \leq M^2} \int_{\mathcal{Z}} f_{n,M}(z) \left| \log \frac{f_{n,M}(z)}{f_{n,M}(z)} \right|^\gamma dz \leq \left| \log \frac{1-p_0}{p_0} \right|^\gamma.$$

□

## 6.4.2 Inspection Constraints and Computational Complexity

In this section, we discuss the computational complexity of  $\pi_{ML}$  and  $\pi_{EJS}$ . Policy  $\pi_{ML}$  has complexity of order  $O(M^2)$  in each step since its main computational burden is to find a segment of the  $M \times M$  grid with the highest probability of containing the target. The main computational burden in  $\pi_{EJS}$  is the action selection step which requires maximizing the EJS divergence over the entire space of inspection regions  $\mathcal{A}$ . This means that at each step the computational complexity of  $\pi_{EJS}$  is of order  $O(|\mathcal{A}|)$  where  $|\mathcal{A}|$  is the number of allowable inspection regions.

**Table 6.1:** Computational Complexity of  $\pi_{EJS}$

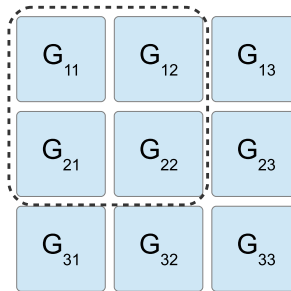
| $\mathcal{A}$    | $ \mathcal{A} $          | Complexity   |
|------------------|--------------------------|--------------|
| all combinations | $2^{M^2}$                | $O(2^{M^2})$ |
| rectangles       | $\frac{M^2(M+1)^2}{4}$   | $O(M^4)$     |
| squares          | $\frac{M(M+1)(2M+1)}{6}$ | $O(M^3)$     |
| thresholds       | $2M^2$                   | $O(M^2)$     |

This provides a recipe for trading off performance for complexity by only considering a smaller number of inspection regions. Table 6.1 shows the complexity of  $\pi_{EJS}$  when the inspection regions are constrained. In particular, if the

inspection regions are always chosen to be contiguous and of certain shape, such as rectangles, etc., the resulting heuristic necessarily has complexity comparable with that of  $\pi_{ML}$ . Constraining the allowable set of inspection regions not only results in a reduction in computational complexity, it also captures realistic limitations in humans' visual abilities. However, a potential drawback of constraining the allowable inspection region is whether similar performance bounds as (6.6) can be obtained. The proof of Theorem 6.4.2 shows that  $\pi_{EJS}$  can achieve upper bound (6.6) so long as  $\mathcal{A}$  contains the single segments and the 2-D threshold-like regions (see Section 6.6.2 for the definition of the latter). It is also conjectured that constraining the inspection region to that of rectangles or squares is of no consequence to upper bound (6.6).

## 6.5 Numerical Example

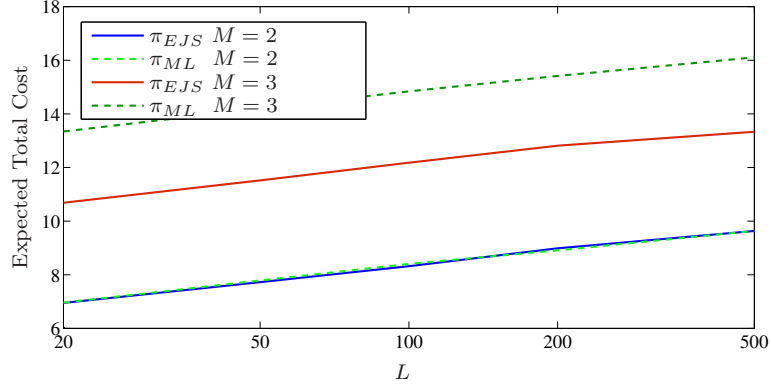
In this section, we compare the performance of  $\pi_{ML}$  and  $\pi_{EJS}$  numerically. We consider the case where the inspection space  $\mathcal{A}$  is constrained to squares in the  $M \times M$  grid. The observation kernels are of the form  $f_{n,M} = N(1, 0.4 + 0.1 \frac{n-1}{M^2})$  and  $\bar{f}_{n,M} = N(0, 0.4 + 0.1 \frac{n-1}{M^2})$ .



**Figure 6.2:**  $3 \times 3$  Grid. The inspection space of all squares in this grid contains 9 single segments, 4 squares of size  $2 \times 2$  (one of which is shown in the figure), and one square of size  $3 \times 3$ .

Figure 6.3 shows that when  $M = 2$ , the performance of the candidate policies is quite similar. This is intuitive since in this case the inspection space contains only the single segments (note that the inspection space contains also

the combination of all 4 segments which does not provide any information and whose EJS is restricted to 0, hence can be ignored). However, for larger  $M$ , the performance gap of the two policies increases which signifies the superiority of  $\pi_{EJS}$  over  $\pi_{ML}$ . Note that the expected total cost corresponding to both policies has the same slope in  $L$ , which verifies the results of Section 6.4.



**Figure 6.3:** Expected total cost for different values of  $L$  and  $M$ .

## 6.6 Proofs

### 6.6.1 Proof of Theorem 6.4.1

In this section, we show that

$$\lim_{M \rightarrow \infty} I_{\max}(M^2) \leq C_{\max} \quad \text{and} \quad \lim_{M \rightarrow \infty} D_{\max}(M^2) = E_{\max}.$$

The proof then simply follows from the above fact and by Proposition 4.6.1.

Remember that  $\eta^a = \sum_{k=1}^M \sum_{l=1}^M \rho_{kl} \mathbf{1}_{\{G_{kl} \in a\}}$ , and  $\bar{\eta}^a = 1 - \eta^a$ . From (6.2) and for any  $a \in \mathcal{A}$  and  $\boldsymbol{\rho} \in \mathbb{P}(\Omega)$ ,

$$\begin{aligned} JS(\boldsymbol{\rho}, a) &= \sum_{i=1}^M \sum_{j=1}^M \rho_{ij} \left[ \mathbf{1}_{\{G_{ij} \in a\}} D(f_{|a|,M} \| \eta^a f_{|a|,M} + \bar{\eta}^a \bar{f}_{|a|,M}) \right. \\ &\quad \left. + \mathbf{1}_{\{G_{ij} \notin a\}} D(\bar{f}_{|a|,M} \| \eta^a f_{|a|,M} + \bar{\eta}^a \bar{f}_{|a|,M}) \right] \\ &= \eta^a D(f_{|a|,M} \| \eta^a f_{|a|,M} + \bar{\eta}^a \bar{f}_{|a|,M}) + \bar{\eta}^a D(\bar{f}_{|a|,M} \| \eta^a f_{|a|,M} + \bar{\eta}^a \bar{f}_{|a|,M}) \end{aligned}$$

$$\begin{aligned}
&\stackrel{(a)}{\leq} \eta^a D(f_{1,M} \| \eta^a f_{1,M} + \bar{\eta}^a \bar{f}_{1,M}) + \bar{\eta}^a D(\bar{f}_{1,M} \| \eta^a f_{1,M} + \bar{\eta}^a \bar{f}_{1,M}) \\
&\stackrel{(b)}{\leq} D(f_{1,M} \| \frac{1}{2} f_{1,M} + \frac{1}{2} \bar{f}_{1,M}),
\end{aligned}$$

where (a) follows from Assumption 6.4.2 and (b) follows from Assumption 6.4.1.

Therefore, by definition (4.19)

$$I_{\max}(M^2) = \max_{a \in \mathcal{A}} \max_{\boldsymbol{\rho} \in \mathbb{P}(\Omega)} JS(\boldsymbol{\rho}, a) \leq D(f_{1,M} \| \frac{1}{2} f_{1,M} + \frac{1}{2} \bar{f}_{1,M}),$$

and hence,

$$\lim_{M \rightarrow \infty} I_{\max}(M^2) \leq C_{\max}.$$

On the other hand, by definition (4.18) and from (6.2)

$$D_{\max}(M^2) = \max_{a \in \mathcal{A}} \max_{1 \leq i,j,k,l \leq M} D(q_{ij}^a \| q_{kl}^a) = D(f_{1,M} \| \bar{f}_{1,M}),$$

which implies that

$$\lim_{M \rightarrow \infty} D_{\max}(M^2) = E_{\max}.$$

### 6.6.2 Proof of Theorem 6.4.2

In this section, we show that

$$EJS(\boldsymbol{\rho}, \pi_{ML}) \geq \begin{cases} \frac{1}{M^2} D(f_{1,M} \| \bar{f}_{1,M}) & \text{if } \max_{i,j} \rho_{ij} < \tilde{\rho} \\ \tilde{\rho} D(f_{1,M} \| \bar{f}_{1,M}) & \text{otherwise} \end{cases}, \quad (6.7)$$

and

$$EJS(\boldsymbol{\rho}, \pi_{EJS}) \geq \begin{cases} D(f_{\lfloor M^2/2 \rfloor, M} \| \frac{1}{2} f_{\lfloor M^2/2 \rfloor, M} + \frac{1}{2} \bar{f}_{\lfloor M^2/2 \rfloor, M}) & \text{if } \max_{i,j} \rho_{ij} < \tilde{\rho} \\ \tilde{\rho} D(f_{1,M} \| \bar{f}_{1,M}) & \text{otherwise} \end{cases}, \quad (6.8)$$

where  $\tilde{\rho} := 1 - \frac{1}{1 + \max\{\log M^2, \log L\}}$ .

The proof then follows from (6.7) and (6.8), and by Theorem 4.2.1 and Proposition 4.6.2.

Under Assumption 6.4.1, we can rewrite (6.3) as

$$EJS(\boldsymbol{\rho}, a) = \sum_{i=1}^M \sum_{j=1}^M \rho_{ij} D(f_{|a|,M} \| \frac{\eta_{ij}^a - \rho_{ij}}{1 - \rho_{ij}} f_{|a|,M} + \frac{\bar{\eta}_{ij}^a}{1 - \rho_{ij}} \bar{f}_{|a|,M}),$$

where

$$\eta_{ij}^a = \sum_{k=1}^M \sum_{l=1}^M \rho_{kl} [\mathbf{1}_{\{G_{kl}, G_{ij} \in a\}} + \mathbf{1}_{\{G_{kl}, G_{ij} \notin a\}}],$$

and  $\bar{\eta}_{ij}^a = 1 - \eta_{ij}^a$ .

Consider a belief vector  $\boldsymbol{\rho}$  at which  $\rho_{ij} \geq \rho_{kl}$  for all  $1 \leq k, l \leq M$ . By definition of the maximum likelihood policy,  $\pi_{ML}(\{G_{ij}\}|\boldsymbol{\rho}) = 1$ . We have,

$$EJS(\boldsymbol{\rho}, \pi_{ML}) = EJS(\boldsymbol{\rho}, \{G_{ij}\}) \geq \rho_{ij} D(f_{1,M} \| \bar{f}_{1,M}),$$

and hence,

$$EJS(\boldsymbol{\rho}, \pi_{ML}) \geq \begin{cases} \tilde{\rho} D(f_{1,M} \| \bar{f}_{1,M}) & \text{if } \rho_{ij} \geq \tilde{\rho} \\ \frac{1}{M^2} D(f_{1,M} \| \bar{f}_{1,M}) & \text{otherwise} \end{cases}. \quad (6.9)$$

Let  $\mathcal{A}_s$  denote the collection of all single segments, and let  $\mathcal{A}_{th}$  be the collection of all 2-D threshold-like regions, i.e., for all  $1 \leq i, j \leq M$ ,  $\bigcup_{k=1}^{i-1} \bigcup_{l=1}^M \{G_{kl}\} \cup \bigcup_{l=1}^{j-1} \{G_{il}\} \in \mathcal{A}_{th}$  and  $\bigcup_{l=j}^M \{G_{il}\} \cup \bigcup_{k=i+1}^M \bigcup_{l=1}^M \{G_{kl}\} \in \mathcal{A}_{th}$ . Let

$$a^* := \arg \min_{a \in \mathcal{A}_{th} : |a| \leq \lfloor M^2/2 \rfloor} |1 - 2\eta^a|.$$

Next we show that

$$EJS(\boldsymbol{\rho}, a^*) \geq D(f_{\lfloor M^2/2 \rfloor, M} \| \frac{1}{2} f_{\lfloor M^2/2 \rfloor, M} + \frac{1}{2} \bar{f}_{\lfloor M^2/2 \rfloor, M}),$$

which together with (6.9) and the fact that

$$EJS(\boldsymbol{\rho}, \pi_{EJS}) = \max_{a \in \mathcal{A}} EJS(\boldsymbol{\rho}, a) \geq \max\{EJS(\boldsymbol{\rho}, \pi_{ML}), EJS(\boldsymbol{\rho}, a^*)\}$$

proves (6.8) and completes the proof (which holds so long as  $\mathcal{A}_s \cup \mathcal{A}_{th} \subseteq \mathcal{A}$ ).

We write the proof for the case  $a^* = \bigcup_{k=1}^{i-1} \bigcup_{l=1}^M \{G_{kl}\} \cup \bigcup_{l=1}^{j-1} \{G_{il}\}$  and  $\eta^{a^*} \leq \frac{1}{2}$  which implies that  $1 - 2\eta^{a^*} \leq \rho_{ij}$ . Other (three) cases can be proved in a similar way.

$$\begin{aligned}
EJS(\boldsymbol{\rho}, a^*) &= \sum_{k=1}^M \sum_{l=1}^M \rho_{kl} D(f_{|a^*|,M} \| \frac{\eta_{kl}^{a^*} - \rho_{kl}}{1 - \rho_{kl}} f_{|a^*|,M} + \frac{\bar{\eta}_{kl}^{a^*}}{1 - \rho_{kl}} \bar{f}_{|a^*|,M}) \\
&\stackrel{(a)}{\geq} \eta^{a^*} D(f_{|a^*|,M} \| \eta^{a^*} f_{|a^*|,M} + \bar{\eta}^{a^*} \bar{f}_{|a^*|,M}) \\
&\quad + \rho_{ij} D(f_{|a^*|,M} \| \frac{1}{2} f_{|a^*|,M} + \frac{1}{2} \bar{f}_{|a^*|,M}) \\
&\quad + (\bar{\eta}^{a^*} - \rho_{ij}) D(f_{|a^*|,M} \| \bar{\eta}^{a^*} f_{|a^*|,M} + \eta^{a^*} \bar{f}_{|a^*|,M}) \\
&\stackrel{(b)}{\geq} D(f_{|a^*|,M} \| \frac{1}{2} f_{|a^*|,M} + \frac{1}{2} \bar{f}_{|a^*|,M}) \\
&\stackrel{(c)}{\geq} D(f_{\lfloor M^2/2 \rfloor, M} \| \frac{1}{2} f_{\lfloor M^2/2 \rfloor, M} + \frac{1}{2} \bar{f}_{\lfloor M^2/2 \rfloor, M}),
\end{aligned}$$

where (a) follows from the fact that  $\eta_{kl}^{a^*} = \eta^{a^*} \mathbf{1}_{\{G_{kl} \in a^*\}} + \bar{\eta}^{a^*} \mathbf{1}_{\{G_{kl} \notin a^*\}}$ ,  $\frac{\eta_{kl}^{a^*} - \rho_{kl}}{1 - \rho_{kl}} \leq \eta_{kl}^{a^*}$ , and  $\frac{\eta_{ij}^{a^*} - \rho_{ij}}{1 - \rho_{ij}} \leq \frac{1}{2}$ ; (b) follows from Jensen's inequality and since  $(\eta^{a^*})^2 + \frac{1}{2}\rho_{ij} + (\bar{\eta}^{a^*} - \rho_{ij})\bar{\eta}^{a^*} \leq \frac{1}{2}$ ; and (c) follows from Assumption 6.4.2.

## Chapter 7

# Noisy Bayesian Active Learning

This chapter considers the problem of noisy Bayesian active learning given a sample space, a finite label set, and a finite set of label generating functions from the sample space to the label set, also known as the function class. The objective is to identify the function in the function class that generates the labels using as few label queries as possible and with low probability of error despite possible corruption by independent noise. The key to achieving this objective relies on the selection of queries in a strategic and adaptive manner. The problem generalizes the problem of noisy generalized binary search [10]. We explore the connection between the above Bayesian active learning problem and the problem of active hypothesis testing. This view of the problem allows for developing a general lower bound on the expected number of queries needed to identify the function that generates the labels in terms of the observation noise statistics, the desired probability of error, and the cardinality of the function class. Furthermore, we compare the performance of  $\pi_{EJS}$ , which in each step queries the label of a sample that maximizes the Extrinsic Jensen–Shannon divergence, with the state of the art strategies for noisy generalized binary search and for different function classes. In the case where the function class is *sample-rich*, it is shown that  $\pi_{EJS}$  is better than previous results in the literature and, in particular, matches the earlier proposed lower bound asymptotically.

## 7.1 Introduction

We consider the problem of Bayesian active learning. We are given a finite set of functions  $\mathcal{H}$ , and a sample space  $\mathcal{X}$ . Each function in  $\mathcal{H}$  assigns a label to a sample in  $\mathcal{X}$ , and the result of a label query on a sample is corrupted by independent noise whose distribution is known. Our goal is to identify the function in  $\mathcal{H}$  that generates the labels with high confidence using as few label queries as possible, by selecting the queries adaptively in a strategic manner.

A special case of the problem, first considered by [78], arises when the label set is binary and the natural sampling strategy for Bayesian active learning becomes closely related to Generalized Binary Search (GBS). In the binary label setting, GBS queries the label of a sample  $x$  for which the size of the subsets of functions that label  $x$  as  $+1$  and  $-1$ , respectively, are as balanced as possible. A variant of GBS is Modified Soft-Decision Generalized Binary Search (MSGBS), which was introduced by [78] to address the case when the observed labels may be noisy. [78] analyzes the performance of MSGBS, under a symmetric and non-persistent noise model which flips the labels randomly, and shows that the (fixed) number of samples required to identify the correct function with probability of error satisfying  $\text{Pe} \leq \epsilon$  is  $O\left(\frac{\log M + \log \frac{1}{\epsilon}}{\lambda}\right)$ , where  $M$  is the number of functions in the class  $\mathcal{H}$ , and  $\lambda$  is a parameter which depends on the structure of the function class and the noise rate. The main contribution of this chapter is to generalize the above problem to the case of general multiclass (non-binary) label set with general (and potentially non-symmetric) non-persistent observation noise.

We observe that Bayesian active learning is a special case of active hypothesis testing, where the hypotheses map to functions, actions map to samples, and the outcomes map to noisy observation of labels. This view of the problem allows for a natural extension of the model of [78] to the non-binary Bayesian active learning setting, where the noise might be label dependent and asymmetric. Relying on this connection and applying the results of Chapter 4, we derive a universal lower bound on the *expected* number of samples required to identify the true hypothesis among  $M$  with error probability  $\text{Pe} \leq \epsilon$  as a function of noise model parameters. This lower bound, when specialized for the noisy generalized binary search

suggests that the proposed schemes of [78], in general, are suboptimal. The next contribution of this work is, thus, to propose and analyze an alternative strategy for sample collection.

To find an alternative strategy, we, again, take advantage of the connection between Bayesian learning and active sequential hypothesis testing. In Chapter 3, we introduced the notion of Extrinsic Jensen–Shannon (EJS) divergence, and proposed  $\pi_{EJS}$ , an active sequential hypothesis test that, at each step, selects the action that maximizes the EJS divergence. In this chapter, we apply the corresponding sampling strategy to Bayesian active learning, and characterize the performance of this strategy. Our bounds show that the number of label queries required by our algorithm is  $O\left(\frac{\log M}{\alpha} + \frac{\log \frac{1}{\epsilon}}{\beta}\right)$ , where  $M$  is the number of functions and  $\alpha$  and  $\beta$  are terms, different from  $\lambda$ , that depend on the structure of the function class, the sample space, and the noise model.

To illustrate our bounds, in Section 7.4, we focus on the generalized binary search studied in [78] and consider the class of *locally identifiable* functions and its three specific subclasses: intervals on the line, thresholds on the line, and a set of sample-rich function classes. We show that the upper bounds on the number of labels required by the EJS policy are superior to those of [78] for all three subclasses for the asymptotic values of  $M$  and  $\epsilon$ . In addition, we show through numerical simulations that our policy has better performance than the algorithms of [78] also in non-asymptotic regimes of practical interest.

The remainder of this chapter is organized as follows. In Section 7.2, we formulate the problem of Bayesian active learning and provide a summary of related works. In Section 7.3, we propose our heuristic policy for selecting samples, and provide the main results of the chapter. As a special case, noisy generalized binary search is discussed in Section 7.4 and a comparison to some of the known results is provided.

## 7.2 Problem Formulation

Let  $\Omega = \{1, 2, \dots, M\}$ . Let  $\mathcal{H} = \{h_1, h_2, \dots, h_M\}$  denote a collection of  $M$  unique functions defined on the *sample space*  $\mathcal{X}$  that take values in a finite *label set*  $\mathcal{L}$ , i.e.,  $h_i : \mathcal{X} \rightarrow \mathcal{L}$ ,  $i \in \Omega$ . Assume that one of the functions in  $\mathcal{H}$  is the true function, i.e., it produces the correct labeling on  $\mathcal{X}$ . However, the observation of the labels is subject to noise, i.e., if the true function is  $h_i$  and the decision maker selects sample  $x$  from  $\mathcal{X}$ , the observation of label  $l = h_i(x)$  is a  $\mathcal{Y}$ -valued random variable with the probability density function  $f_l(\cdot)$ . We assume the observation densities  $\{f_l(\cdot)\}_{l \in \mathcal{L}}$  are fixed and known, and the observations are conditionally independent over time. The goal is to determine the true function with a small number of samples and acceptable accuracy in face of noisy observation by making adaptive queries. More formally, if  $\tau$  is the total number of label queries made by our algorithm, then, we would like to find the true function in  $\mathcal{H}$  with the objective of:

$$\text{minimize } \mathbb{E}[\tau] \text{ subject to } \text{Pe} \leq \epsilon, \quad (7.1)$$

where  $\text{Pe}$  is the probability of error.

Let  $\theta$  be the index of the true function, i.e.,  $\theta$  is the random variable that takes the value  $\theta = i$  on the event that  $h_i$  is the true function for  $i \in \Omega$ . Let  $\tau$  be the stopping time at which the decision maker retires and guesses the true index  $\theta$  as the maximum a posteriori (MAP) declaration<sup>1</sup>, i.e.,

$$\hat{\theta} = \arg \max_{i \in \Omega} P(\theta = i | X^{\tau-1}, Y^{\tau-1}),$$

where  $X^{\tau-1} := [X(0), X(1), \dots, X(\tau-1)]$  and  $Y^{\tau-1} := [Y(0), Y(1), \dots, Y(\tau-1)]$  represent respectively the vector of selected samples and observed labels up to time  $\tau$ . The probability of error is

$$\text{Pe} := P(\hat{\theta} \neq \theta).$$

For a given  $\epsilon > 0$ , the goal is to find a stopping time  $\tau$  and a sequence of samples  $X(0), X(1), \dots, X(\tau-1)$  such that the probability of error satisfies  $\text{Pe} \leq \epsilon$  and

---

<sup>1</sup>The probability of error is minimized under MAP declaration.

the expected number of samples  $\mathbb{E}[\tau]$  is small. Here we consider a uniform prior, i.e., initially  $P(\theta = i) = 1/M$  for all  $i \in \Omega$ , and hence the expectation is always taken with respect to this uniform prior as well as the distribution of the sequence of selected samples and observations.

### 7.2.1 Related Work

Our work naturally builds on and extends on the model proposed in [78]. In the introduction, we detailed the connection between our work and [78].

In recent years, extensive studies have been conducted on the theory of active learning [79–86]. The various studies incorporate a broad range of models and assumptions which, at times, end up in contradictory conclusions. Here we make an attempt to detail the specific attributes of these papers and the connection/disconnect between our work and this literature. The earlier work on active learning [79–81] considered the realizable case where the binary labels are produced by a function in a given function class and are observed noise-free. Here, in contrast to our setting, the function class is either finite, or infinite but equipped with a fixed structure, such as the class of thresholds on a line, or the class of linear classifiers. Another point of difference with our work, the learner here is only allowed to query the labels of samples among an unlabeled set of points which are drawn from the unlabeled data distribution. The goal is to find a function which has low *prediction error with respect to the data distribution*. Here the challenge is to identify a function in the function class where the disagreements with the true labeling function is less than the required accuracy, and the prediction error occurs due to infiniteness of the function class or due to the indistinguishability of the functions with respect to the data distribution as opposed to noisy observations of the labels.

Since the realizability assumption can hardly ever be justified in practice, more recent literature [82–86] has considered active learning in the agnostic setting, where the goal is to find a function in a function class which has low prediction error with respect to the data distribution; however the binary labels are not necessarily generated by a function in the given function class. Perhaps, the most relevant

special case is the bounded rate class noise of [82] in which labels are produced by a member of a given function class but are subjected to an exogenous (and non-persistent) observation noise. In such a setting, repeat queries can be effectively utilized to mitigate the effect of noise [82, 87]. Furthermore, instead of analyzing a specific query strategy, we are interested in obtaining universal lower bounds on the query complexity which are related to the information theoretic analysis of [86].

In summary, our study differs from the above work in three important ways. Firstly, we are interested in a generalized learning setup where labels can be non-binary and observation noise can have a general non-symmetric and non-discrete nature. Secondly, we are interested in a sequential learning setting where the learner is allowed not only to query individual examples (hence, rendering the data distribution irrelevant), but also to determine the number of queries in an online fashion as a function of observations so far. Thirdly, by considering the simpler setup of a finite function class as well as an exogenous and non-persistent observation noise, we provide sharp lower and upper bounds on the query complexity.

## 7.3 Main Results

In this section, we provide the main results of the chapter. Many of the results here are a consequence of a connection between Bayesian active learning and the active sequential hypothesis testing.

### 7.3.1 Preliminaries

The problem of noisy Bayesian active learning defined above is a special case of active sequential hypothesis testing with hypotheses set  $\Omega$ , action space  $\mathcal{X}$ , observation space  $\mathcal{Y}$ , and observation kernels  $\{q_i^x\}_{i \in \Omega, x \in \mathcal{X}}$  where  $q_i^x = f_{h_i(x)}$ .

Let  $\boldsymbol{\rho}(t) = [\rho_1(t), \rho_2(t), \dots, \rho_M(t)]$  denote the information state at time  $t$  where  $\rho_i(t) := P(\{\theta = i\} | X^{t-1}, Y^{t-1})$ . Accordingly, the information state space is  $\mathbb{P}(\Omega) := \{\boldsymbol{\rho} \in [0, 1]^M : \sum_{i=1}^M \rho_i = 1\}$ .

## Heuristic Policy

We consider policy  $\pi_{EJS}$  proposed in Section 3.3.2 as a candidate sampling strategy. From (3.19), we can compute the EJS divergence associated with querying sample  $x$  at belief vector  $\boldsymbol{\rho}$  as

$$EJS(\boldsymbol{\rho}, x) = \sum_{i=1}^M \rho_i D(f_{h_i(x)} \| \sum_{j \neq i} \frac{\rho_j}{1 - \rho_i} f_{h_j(x)}). \quad (7.2)$$

Policy  $\pi_{EJS}$  is a stationary deterministic Markov policy defined as:

- if  $\rho_i(t) \geq 1 - \epsilon$ , retire and declare  $h_i$  as the true function;
- otherwise, select sample  $\arg \max_{x \in \mathcal{X}} EJS(\boldsymbol{\rho}, x)$ .

## Assumptions

We have the following technical assumptions:

**Assumption 7.3.1.**  $C := \min_g \max_{l \in \mathcal{L}} D(f_l \| g) > 0$ , where  $g$  is a convex combination of  $f_l, l \in \mathcal{L}$ .

**Assumption 7.3.2.**  $C_1 := \max_{k, l \in \mathcal{L}} D(f_k \| f_l) < \infty$ .

**Assumption 7.3.3.**  $C_2 := \max_{k, l \in \mathcal{L}} \sup_{y \in \mathcal{Y}} \frac{f_k(y)}{f_l(y)} < \infty$ .

The formulation  $C$  above is equal to the capacity of a communication channel with input alphabet set  $\mathcal{L}$ , output alphabet set  $\mathcal{Y}$ , and transition probabilities  $f_l(\cdot), l \in \mathcal{L}$ . We also denote the capacity of a channel with two inputs  $k, l \in \mathcal{L}$  by  $C_{kl}$ , i.e.,

$$C_{kl} := \min_g \max \{D(f_k \| g), D(f_l \| g)\}, \text{ where } g \text{ is a pdf on } \mathcal{Y}. \quad (7.3)$$

If Assumption 7.3.1 does not hold, that is if  $C = 0$ , the label queries will be completely noisy and no information can be retrieved from the label queries regarding the true function. In this sense, Assumption 7.3.1 is a necessary condition that ensures that the problem of active learning has a meaningful solution. For observation kernels with bounded support, Assumption 7.3.3 is a necessary condition to ensure that no observation is noise-free and is sufficient to ensure Assumption 7.3.2 since  $C_1 \leq \log C_2$ .

### 7.3.2 Lower and Upper Bounds

We have the following lower bound on the minimum expected number of samples required to achieve  $\text{Pe} \leq \epsilon$ .

**Theorem 7.3.1.** *Under Assumptions 7.3.1, 7.3.2, and 7.3.3,*

$$\mathbb{E}[\tau_\epsilon^*] \geq \left[ \frac{\left(1 - \frac{3}{\log \frac{4}{\epsilon}} - \frac{\epsilon}{2} \log \frac{1}{\epsilon}\right) \log M - 2}{C} + \frac{\log \frac{1-\epsilon}{\epsilon} - 2 \log \log \frac{4}{\epsilon} - \log C_2 - 4}{C_1} \right]^+.$$

*Proof.* The proof follows by combining (4.109) with (4.112) and (4.117) for values  $L = \frac{2}{\epsilon \log \frac{4}{\epsilon}}$  and  $\delta = \frac{1}{\log \frac{4}{\epsilon}}$ , and the fact that  $I_{\max}(M) \leq C$ ,  $D_{\max}(M) \leq C_1$ , and  $\xi_M = \log C_2$  for the problem of noisy Bayesian active learning.  $\square$

Next we specialize Theorem 4.2.1 and Corollary 4.2.1 to the problem of noisy Bayesian active learning and we provide an upper bound on the expected number of samples required by  $\pi_{EJS}$  to achieve a desired probability of error  $\text{Pe} \leq \epsilon$ . The upper bounds depend on two quantities  $\alpha$  and  $\beta$ , which depend on the function class  $\mathcal{H}$  and sample space  $\mathcal{X}$ .

**Theorem 7.3.2.** *If there exists  $\alpha > 0$  such that at any given belief vector  $\boldsymbol{\rho} \in \mathbb{P}(\Omega)$ , it is possible to find a sample  $x \in \mathcal{X}$  satisfying  $EJS(\boldsymbol{\rho}, x) \geq \alpha$ , then*

$$\mathbb{E}_{\pi_{EJS}}[\tilde{\tau}_\epsilon] \leq \frac{\log M + \max\{\log \log M, \log \frac{1}{\epsilon}\} + 4C_2}{\alpha}. \quad (7.4)$$

*Furthermore, if there exist positive values  $\alpha$  and  $\beta$  such that for all  $\boldsymbol{\rho} \in \mathbb{P}(\Omega)$ ,*

$$\max_{x \in \mathcal{X}} EJS(\boldsymbol{\rho}, x) \geq \begin{cases} \alpha & \text{if } \max_{i \in \Omega} \rho_i < \tilde{\rho} \\ \beta & \text{otherwise} \end{cases}, \quad (7.5)$$

*where*

$$\tilde{\rho} := 1 - \frac{1}{1 + \max\{\log M, \log \frac{1}{\epsilon}\}}, \quad (7.6)$$

*then the following bound is obtained*

$$\mathbb{E}_{\pi_{EJS}}[\tilde{\tau}_\epsilon] \leq \frac{\log M + \max\{\log \log M, \log \log \frac{1}{\epsilon}\}}{\alpha} + \frac{\log \frac{1}{\epsilon}}{\beta} + \frac{3(4C_2)^2}{\alpha\beta}. \quad (7.7)$$

We now instantiate these upper bounds for some learning problems of interest. For a broad class of learning problems, where the function class in question is a *locally identifiable* class, we can show a general result.

**Definition 7.3.1.** A class of functions  $\mathcal{H}$  is referred to as *locally identifiable* if for any  $h_i \in \mathcal{H}$ , there exist samples  $x, x' \in \mathcal{X}$  and labels  $l, l' \in \mathcal{L}$  such that either

$$[h_i(x), h_i(x')] = [l, l'], \quad \text{and} \quad [h_j(x), h_j(x')] \in \bigcup_{k \in \mathcal{L}} \{[k, k]\} \cup \{[l', l]\}, \quad \forall j \neq i;$$

or

$$[h_i(x), h_i(x')] = [l, l], \quad \text{and} \quad [h_j(x), h_j(x')] \in \{[l, l'], [l', l], [l', l']\}, \quad \forall j \neq i.$$

In essence, the locally identifiable condition implies that for any function  $h_i \in \mathcal{H}$ , there are (at least) two samples in  $\mathcal{X}$  using which  $h_i$  can be distinguished from all other functions.

**Proposition 7.3.1.** When function class  $\mathcal{H}$  is locally identifiable,  $\alpha = \frac{1}{M}C_{\min}$  and  $\beta = \tilde{\rho}C_{\min}$ , where  $C_{\min} = \min_{k, l \in \mathcal{L}, k \neq l} \min \{C_{kl}, D(f_k \| \frac{1}{2}f_k + \frac{1}{2}f_l)\}$ , i.e.,

$$\max_{x \in \mathcal{X}} EJS(\boldsymbol{\rho}, x) \geq \begin{cases} \frac{1}{M}C_{\min} & \text{if } \max_{i \in \Omega} \rho_i < \tilde{\rho} \\ \tilde{\rho}C_{\min} & \text{otherwise} \end{cases}. \quad (7.8)$$

The proof of Proposition 7.3.1 is provided in Section 7.5.1.

For certain subclasses of the locally identifiable functions, one might get substantially tighter bounds. In particular, in Section 7.4, we do this for certain subclasses of binary-valued locally identifiable functions. Here, however, to illustrate how tight these bounds can be, we define an  $\mathcal{L}$ -ary-valued function class for which the upper bound on  $\pi_{EJS}$  matches the lower bound of Theorem 7.3.1 asymptotically.

**Definition 7.3.2.** A class of functions  $\mathcal{H}$  is called *sample-rich* if for any vector  $\mathbf{v} \in \mathcal{L}^M$ , there exists a sample  $x \in \mathcal{X}$  such that  $\mathbf{h}(x) = \mathbf{v}$ , where  $\mathbf{h}(x) := [h_1(x), h_2(x), \dots, h_M(x)]$ . In other words,  $\cup_{x \in \mathcal{X}} \{\mathbf{h}(x)\} = \mathcal{L}^M$ .

**Proposition 7.3.2.** *When function class  $\mathcal{H}$  is sample-rich,  $\alpha = C$  and  $\beta = \tilde{\rho}C_1$ , i.e.,*

$$\max_{x \in \mathcal{X}} EJS(\boldsymbol{\rho}, x) \geq \begin{cases} C & \text{if } \max_{i \in \Omega} \rho_i < \tilde{\rho} \\ \tilde{\rho}C_1 & \text{otherwise} \end{cases} \quad (7.9)$$

The proof of Proposition 7.3.2 is provided in Section 7.5.2.

In the following section, we characterize the performance of  $\pi_{EJS}$  for a special case and compare its performance with the state of the art strategies.

## 7.4 Special Case: Noisy Generalized Binary Search

Noisy generalized binary search, first introduced in [78], is a special case of our problem where functions are binary-valued, i.e.,  $\mathcal{L} = \{-1, +1\}$ , and observations have probability densities of the following form:

$$f_l(y) = \begin{cases} 1 - p & \text{if } y = l \\ p & \text{if } y = -l \end{cases},$$

for some  $p \in (0, 1/2)$ . In other words, label  $l$  is observed through a binary symmetric channel with crossover probability  $p$ .

**Lemma 7.4.1.** *For the case of noisy generalized binary search,  $C$ ,  $C_1$ , and  $C_2$  defined in Section 7.3 can be further simplified to*

$$\begin{aligned} C &:= 1 + p \log p + (1 - p) \log(1 - p), \\ C_1 &:= p \log \frac{p}{1 - p} + (1 - p) \log \frac{1 - p}{p}, \\ C_2 &:= \frac{1 - p}{p}. \end{aligned}$$

In order to emphasize the dependence of  $C$ ,  $C_1$ , and  $C_2$  on the Bernoulli parameter  $p$  (corresponding to the observation noise), we denote them by  $C(p)$ ,  $C_1(p)$ , and  $C_2(p)$ , respectively. Note that from Jensen's inequality,  $C_1(p) \geq 2C(p)$ .

**Corollary 7.4.1.** *For the special case of noisy generalized binary search, when function class  $\mathcal{H}$  is locally identifiable, we have  $\alpha = \frac{C(p)}{M}$  and  $\beta = \tilde{\rho}C(p)$ .*

Next we define three important subclasses of binary-valued locally identifiable functions: 1) class of disjoint interval functions  $\mathcal{H}_D$ ; 2) class of threshold functions  $\mathcal{H}_T$ ; and 3) class of sample-rich functions  $\mathcal{H}_R$ . We analyze the performance of  $\pi_{EJS}$  on these three classes by using Theorem 7.3.2 and finding the corresponding  $\alpha$  and  $\beta$ .

**Definition 7.4.1.** Let  $\mathbf{e}_i$ ,  $i \in \Omega$ , represent a vector of size  $M$  whose  $i^{\text{th}}$  element is  $+1$  and all other elements are  $-1$ . A collection of functions  $\mathcal{H}$  is referred to as *disjoint interval functions* if  $\cup_{x \in \mathcal{X}} \{\mathbf{h}(x)\} = \cup_{i \in \Omega} \{\mathbf{e}_i\} \subset \{-1, +1\}^M$ , where  $\mathbf{h}(x) := [h_1(x), h_2(x), \dots, h_M(x)]$ . In other words, for any sample  $x \in \mathcal{X}$ , only one function in  $\mathcal{H}$  takes value  $+1$  and all other functions take value  $-1$ .

**Definition 7.4.2.** Let  $\mathbf{u}_i$ ,  $i \in \Omega$ , represent a vector of size  $M$  whose first  $i$  elements are  $-1$  and all other elements are  $+1$ . A collection of functions  $\mathcal{H}$  is referred to as *threshold functions* if  $\cup_{x \in \mathcal{X}} \{\mathbf{h}(x)\} = \cup_{i \in \Omega} \{\mathbf{u}_i\} \subset \{-1, +1\}^M$ .

**Definition 7.4.3.** A collection of functions  $\mathcal{H}$  is referred to as *sample-rich functions* if for any vector  $\mathbf{v} \in \{-1, +1\}^M$ , there exists a sample  $x \in \mathcal{X}$  such that  $\mathbf{h}(x) = \mathbf{v}$ . In other words,  $\cup_{x \in \mathcal{X}} \{\mathbf{h}(x)\} = \{-1, +1\}^M$ .

Table 7.1 shows the values of  $\alpha$  and  $\beta$  for the above classes where recall that  $\tilde{\rho} = 1 - \frac{1}{1 + \max\{\log M, \log \frac{1}{\epsilon}\}}$  (see Section 7.5.3 for the proofs).

**Table 7.1:**  $\alpha$  and  $\beta$  for different function classes.

| Function class                     | $\alpha$           | $\beta$              |
|------------------------------------|--------------------|----------------------|
| Disjoint intervals $\mathcal{H}_D$ | $\frac{C_1(p)}{M}$ | $\tilde{\rho}C_1(p)$ |
| Thresholds $\mathcal{H}_T$         | $C(p)$             | $C(p)$               |
| Sample-rich $\mathcal{H}_R$        | $C(p)$             | $\tilde{\rho}C_1(p)$ |

For class of disjoint interval functions  $\mathcal{H}_D$  and from Theorem 7.3.2 and Table 7.1,

$$\begin{aligned}
\mathbb{E}_{\pi_{EJS}}[\tilde{\tau}_\epsilon] &\leq \frac{\log M + \max\{\log \log M, \log \log \frac{1}{\epsilon}\}}{\frac{1}{M}C_1(p)} + \frac{\log \frac{1}{\epsilon}}{\tilde{\rho}C_1(p)} + \frac{3(4C_2(p))^2}{\frac{1}{M}C_1(p)\tilde{\rho}C_1(p)} \\
&\stackrel{(a)}{\leq} \frac{M \log M + M \log \log \frac{M}{\epsilon}}{C_1(p)} + \frac{\log \frac{1}{\epsilon} + 1}{C_1(p)} + \frac{6M(4C_2(p))^2}{(C_1(p))^2} \\
&\leq \left( \frac{M \log M}{C_1(p)} + \frac{\log \frac{1}{\epsilon}}{C_1(p)} \right) \left( 1 + \frac{M \log \log \frac{M}{\epsilon} + 1 + 6M(4C_2(p))^2/C_1(p)}{M \log M + \log \frac{1}{\epsilon}} \right) \\
&= \left( \frac{M \log M}{C_1(p)} + \frac{\log \frac{1}{\epsilon}}{C_1(p)} \right) (1 + o(1)), \tag{7.10}
\end{aligned}$$

where  $o(1) \rightarrow 0$  as  $\epsilon \rightarrow 0$  or  $M \rightarrow \infty$  and (a) holds since  $\frac{1}{\tilde{\rho}} = 1 + \frac{1}{\max\{\log M, \log \frac{1}{\epsilon}\}} \leq 2$ .

For class of threshold functions  $\mathcal{H}_T$  and from Theorem 7.3.2 and Table 7.1,

$$\begin{aligned}
\mathbb{E}_{\pi_{EJS}}[\tilde{\tau}_\epsilon] &\leq \frac{\log M + \max\{\log \log M, \log \log \frac{1}{\epsilon}\} + 4C_2(p)}{C(p)} \\
&\leq \left( \frac{\log M}{C(p)} + \frac{\log \frac{1}{\epsilon}}{C(p)} \right) \left( 1 + \frac{\log \log M + 4C_2(p)}{\log \frac{M}{\epsilon}} \right) \\
&= \left( \frac{\log M}{C(p)} + \frac{\log \frac{1}{\epsilon}}{C(p)} \right) (1 + o(1)), \tag{7.11}
\end{aligned}$$

where  $o(1) \rightarrow 0$  as  $\epsilon \rightarrow 0$  or  $M \rightarrow \infty$ .

For class of sample-rich functions  $\mathcal{H}_R$  and from Theorem 7.3.2 and Table 7.1,

$$\begin{aligned}
\mathbb{E}_{\pi_{EJS}}[\tilde{\tau}_\epsilon] &\leq \frac{\log M + \max\{\log \log M, \log \log \frac{1}{\epsilon}\}}{C(p)} + \frac{\log \frac{1}{\epsilon}}{\tilde{\rho}C_1(p)} + \frac{3(4C_2(p))^2}{C(p)\tilde{\rho}C_1(p)} \\
&\stackrel{(a)}{\leq} \frac{\log M + \log \log \frac{M}{\epsilon}}{C(p)} + \frac{\log \frac{1}{\epsilon} + 1}{C_1(p)} + \frac{6(4C_2(p))^2}{C(p)C_1(p)} \\
&\leq \left( \frac{\log M}{C(p)} + \frac{\log \frac{1}{\epsilon}}{C_1(p)} \right) \left( 1 + \frac{C_1(p) \log \log \frac{M}{\epsilon} + C(p) + 6(4C_2(p))^2}{C(p) \log \frac{M}{\epsilon}} \right) \\
&= \left( \frac{\log M}{C(p)} + \frac{\log \frac{1}{\epsilon}}{C_1(p)} \right) (1 + o(1)), \tag{7.12}
\end{aligned}$$

where  $o(1) \rightarrow 0$  as  $\epsilon \rightarrow 0$  or  $M \rightarrow \infty$  and (a) holds since  $\frac{1}{\tilde{\rho}} = 1 + \frac{1}{\max\{\log M, \log \frac{1}{\epsilon}\}} \leq 2$ .

Furthermore, it follows from Theorem 7.3.1 that

$$\mathbb{E}[\tau_\epsilon^*] \geq \left( \frac{\log M}{C(p)} + \frac{\log \frac{1}{\epsilon}}{C_1(p)} \right) (1 - o(1)), \tag{7.13}$$

where  $o(1) \rightarrow 0$  as  $\epsilon \rightarrow 0$ .

### 7.4.1 Comparison to Known Results

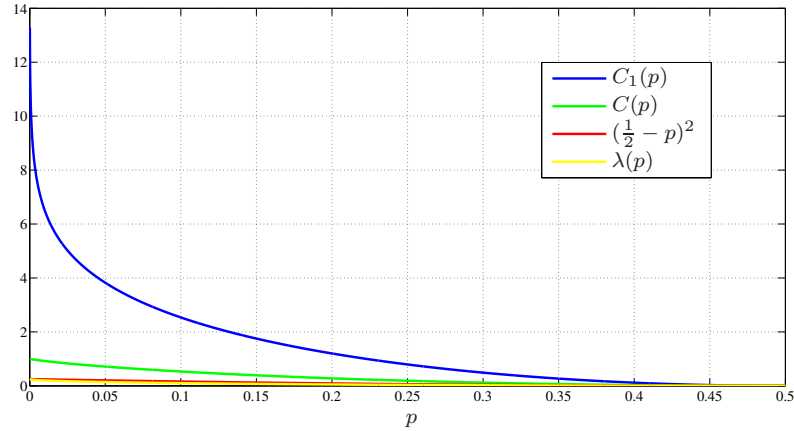
In this section, we compare the performance of  $\pi_{EJS}$  to that of NGBS and MSGBS policies proposed in [10]. The performance of NGBS and MSGBS was analyzed in [10] for the *1-neighborly* function class and its subclasses. It can be shown that the 1-neighborly class is a subclass of the binary-valued locally identifiable functions. Furthermore, function classes  $\mathcal{H}_D$ ,  $\mathcal{H}_T$ , and  $\mathcal{H}_R$  are also 1-neighborly and hence, we can compare the performance of  $\pi_{EJS}$  to that of NGBS and MSGBS over these function classes.

Table 7.2 shows the number of samples required by the policies NGBS, MSGBS, and  $\pi_{EJS}$  to attain  $\text{Pe} \leq \epsilon$ , where  $\lambda(p) = \max_{p' \in (p, 1/2)} \frac{1}{4} \left( 1 - \frac{p'(1-p)}{1-p'} - \frac{(1-p')p}{p'} \right)$ . Policies NGBS and MSGBS are non-sequential and the numbers shown in Table 7.2 are the actual number of samples that these policies require to achieve  $\text{Pe} \leq \epsilon$ ; while policy  $\pi_{EJS}$  is sequential and Table 7.2 shows the expected number of samples required by this policy to achieve  $\text{Pe} \leq \epsilon$ .

**Table 7.2:** Performance comparison of NGBS, MSGBS, and  $\pi_{EJS}$  for different function classes.

| Function class  | NGBS                                                                      | MSGBS                                                     | $\pi_{EJS}$                                                                                  |
|-----------------|---------------------------------------------------------------------------|-----------------------------------------------------------|----------------------------------------------------------------------------------------------|
| $\mathcal{H}_D$ | $\frac{M(\log M + \log \frac{1}{\epsilon})}{(\frac{1}{2}-p)^2}$           | $\frac{M(\log M + \log \frac{1}{\epsilon})}{4\lambda(p)}$ | $\left( \frac{M \log M}{C_1(p)} + \frac{\log \frac{1}{\epsilon}}{C_1(p)} \right) (1 + o(1))$ |
| $\mathcal{H}_T$ | $\frac{\log M(\log \log M + \log \frac{1}{\epsilon})}{(\frac{1}{2}-p)^2}$ | $\frac{\log M + \log \frac{1}{\epsilon}}{\lambda(p)}$     | $\left( \frac{\log M}{C(p)} + \frac{\log \frac{1}{\epsilon}}{C(p)} \right) (1 + o(1))$       |
| $\mathcal{H}_R$ | $\frac{\log M(\log \log M + \log \frac{1}{\epsilon})}{(\frac{1}{2}-p)^2}$ | $\frac{\log M + \log \frac{1}{\epsilon}}{\lambda(p)}$     | $\left( \frac{\log M}{C(p)} + \frac{\log \frac{1}{\epsilon}}{C_1(p)} \right) (1 + o(1))$     |

Figure 7.1 compares the denominators of the upper bounds given in Table 7.2. Note that our upper bound provides improvement over those corresponding to NGBS and MSGBS. Particularly, the gap between the bounds is very significant for small values of the Bernoulli parameter  $p$  and for large values of  $\frac{1}{\epsilon}$  and  $M$ . However, with no tight lower bound on the performance of NGBS and MSGBS, the analysis is inconclusive as to the actual performance of  $\pi_{EJS}$  in comparison with NGBS and MSGBS.



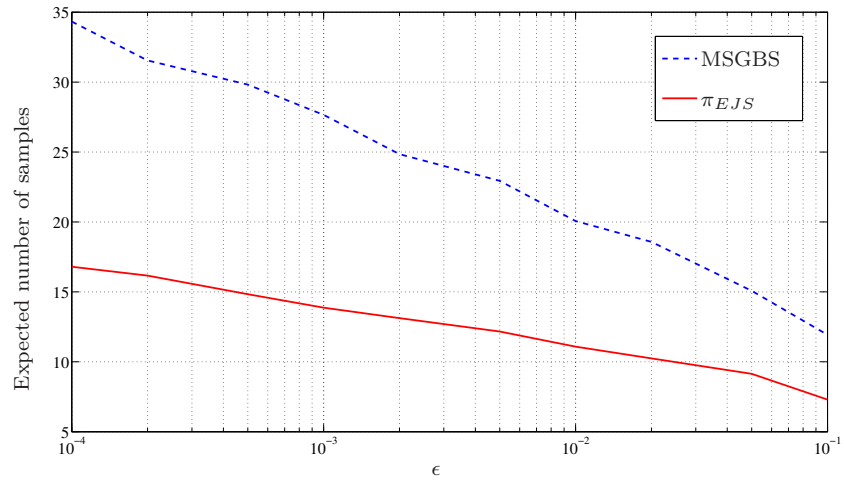
**Figure 7.1:** Comparison of  $C(p)$ ,  $C_1(p)$ ,  $(\frac{1}{2} - p)^2$ , and  $\lambda(p)$ , for  $p \in (0, 1/2)$ .

In the next section, we provide a numerical example and compare the performance of  $\pi_{EJS}$  and MSGBS (we do not consider NGBS since it is outperformed by MSGBS).

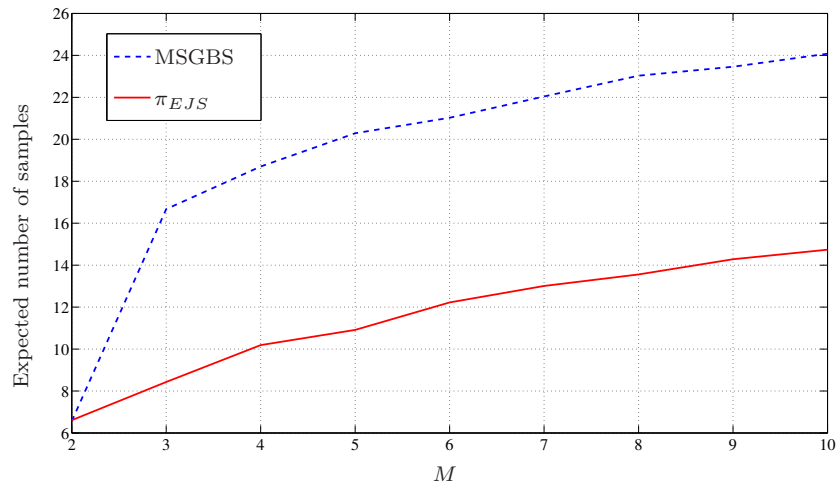
### 7.4.2 Numerical Example

In this section, policies  $\pi_{EJS}$  and MSGBS are compared numerically for the problem of noisy generalized binary search with parameter  $p$  and a sample-rich function class of size  $M$ . In order to have a fair comparison, the candidate policies are compared in both sequential and non-sequential scenarios. In the sequential scenario, the policies stop as soon as the belief about one of the functions passes a threshold  $1 - \epsilon$ , and the expected number of queries is considered as a measure of performance; while in the non-sequential scenario, the policies are compared based on their average probability of making a wrong declaration after  $N$  number of label queries.

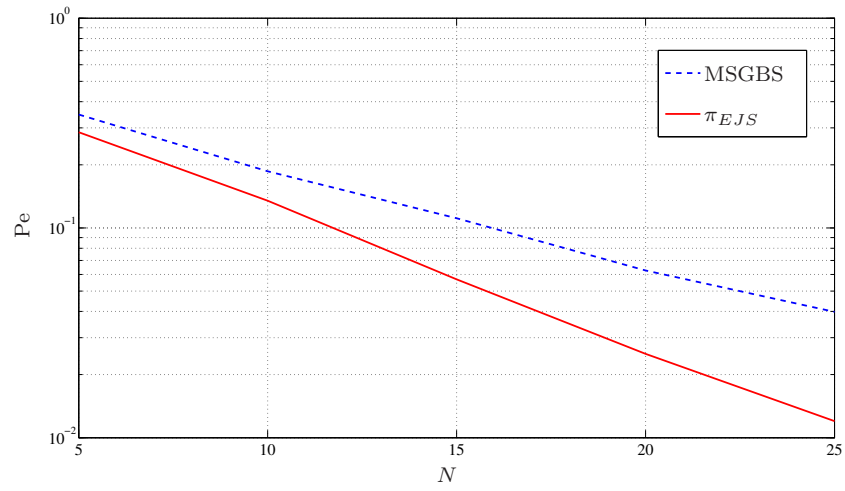
Figures 7.2 and 7.3 show the performance of  $\pi_{EJS}$  and MSGBS for the sequential scenario while Figs. 7.4 and 7.5 compare their performance for the non-sequential scenario. The figures show the superior performance of  $\pi_{EJS}$  over MSGBS in both scenarios and for different values of  $\epsilon$ ,  $N$ , and  $M$ .



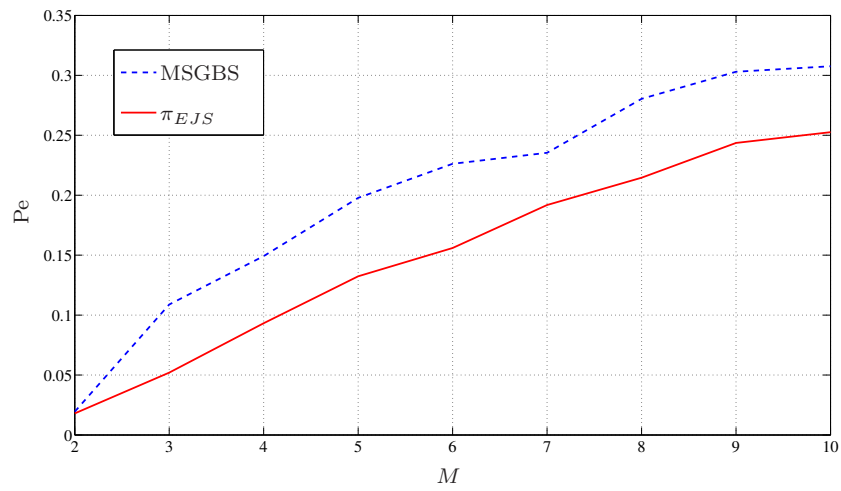
**Figure 7.2:** Expected number of samples as  $\epsilon$  varies for  $M = 5$  and  $p = 0.2$ .



**Figure 7.3:** Expected number of samples as  $M$  varies for  $\epsilon = 0.01$  and  $p = 0.2$ .



**Figure 7.4:** Average probability of error as  $N$  varies for  $M = 5$  and  $p = 0.2$ .



**Figure 7.5:** Average probability of error as  $M$  varies for  $N = 10$  and  $p = 0.2$ .

## 7.5 Proofs

### 7.5.1 Proof of Proposition 7.3.1

To prove Proposition 7.3.1, it suffices to show that

$$\max_{x \in \mathcal{X}} EJS(\boldsymbol{\rho}, x) \geq \max_{i \in \Omega} \rho_i C_{\min}.$$

Let  $\hat{i} = \arg \max_{i \in \Omega} \rho_i$ . By definition of the locally identifiable class, there exist  $x_{\hat{i}}, x'_{\hat{i}} \in \mathcal{X}$  and  $l, l' \in \mathcal{L}$  such that one of the following conditions holds

$$[h_{\hat{i}}(x_{\hat{i}}), h_{\hat{i}}(x'_{\hat{i}})] = [l, l'] \quad \text{and} \quad [h_j(x_{\hat{i}}), h_j(x'_{\hat{i}})] \in \bigcup_{k \in \mathcal{L}} \{[k, k]\} \cup \{[l', l]\} \quad \forall j \neq \hat{i}, \quad (7.14)$$

$$[h_{\hat{i}}(x_{\hat{i}}), h_{\hat{i}}(x'_{\hat{i}})] = [l, l] \quad \text{and} \quad [h_j(x_{\hat{i}}), h_j(x'_{\hat{i}})] \in \{[l, l'], [l', l], [l', l']\} \quad \forall j \neq \hat{i}. \quad (7.15)$$

For any  $k, k' \in \mathcal{L}$ , let

$$\pi_{kk'} := \sum_{j \in \Omega: [h_j(x_{\hat{i}}), h_j(x'_{\hat{i}})] = [k, k']} \frac{\rho_j}{1 - \rho_{\hat{i}}}.$$

Suppose (7.14) holds. Then

$$\begin{aligned} & \max_{x \in \mathcal{X}} EJS(\boldsymbol{\rho}, x) \\ & \geq \max \{EJS(\boldsymbol{\rho}, x_{\hat{i}}), EJS(\boldsymbol{\rho}, x'_{\hat{i}})\} \\ & \geq \rho_{\hat{i}} \max \left\{ D \left( f_{h_{\hat{i}}(x_{\hat{i}})} \parallel \sum_{j \neq \hat{i}} \frac{\rho_j}{1 - \rho_{\hat{i}}} f_{h_j(x_{\hat{i}})} \right), D \left( f_{h_{\hat{i}}(x'_{\hat{i}})} \parallel \sum_{j \neq \hat{i}} \frac{\rho_j}{1 - \rho_{\hat{i}}} f_{h_j(x'_{\hat{i}})} \right) \right\} \\ & = \rho_{\hat{i}} \max \left\{ D \left( f_l \parallel \sum_{k \in \mathcal{L}} \pi_{kk} f_k + \pi_{l'l} f_{l'} \right), D \left( f_{l'} \parallel \sum_{k \in \mathcal{L}} \pi_{kk} f_k + \pi_{l'l} f_l \right) \right\} \\ & \stackrel{(a)}{\geq} \rho_{\hat{i}} \max \left\{ D \left( f_l \parallel \frac{\sum_{k \in \mathcal{L}} \pi_{kk} f_k + \pi_{l'l} f_{l'} + \pi_{l'l} f_l}{1 + \pi_{l'l}} \right), D \left( f_{l'} \parallel \frac{\sum_{k \in \mathcal{L}} \pi_{kk} f_k + \pi_{l'l} f_l + \pi_{l'l} f_{l'}}{1 + \pi_{l'l}} \right) \right\} \\ & \geq \rho_{\hat{i}} \min_g \max \{D(f_l \parallel g), D(f_{l'} \parallel g)\} \\ & = \rho_{\hat{i}} C_{ll'} \\ & \geq \max_{i \in \Omega} \rho_i C_{\min}, \end{aligned} \quad (7.16)$$

where (a) follows by Lemma 3.3.1.

On the other hand, if (7.15) holds, then

$$\begin{aligned}
& \max_{x \in \mathcal{X}} EJS(\boldsymbol{\rho}, x) \\
& \geq \rho_i \max \left\{ D(f_l \| \pi_{ll'} f_l + (\pi_{l'l} + \pi_{l'l'}) f_{l'}) , D(f_l \| \pi_{l'l} f_l + (\pi_{ll'} + \pi_{l'l'}) f_{l'}) \right\} \\
& \stackrel{(a)}{\geq} \rho_i D\left(f_l \| \frac{1}{2} f_l + \frac{1}{2} f_{l'}\right) \\
& \geq \max_{i \in \Omega} \rho_i C_{\min},
\end{aligned} \tag{7.17}$$

where (a) follows by Lemma 3.3.1 and the fact that  $\min\{\pi_{ll'}, \pi_{l'l}\} \leq \frac{1}{2}$ .

Combining (7.16) and (7.17), we have the assertion of the proposition.

### 7.5.2 Proof of Proposition 7.3.2

To prove Proposition 7.3.2, we will show that

$$\max_{x \in \mathcal{X}} EJS(\boldsymbol{\rho}, x) \geq C,$$

and furthermore,

$$\max_{x \in \mathcal{X}} EJS(\boldsymbol{\rho}, x) \geq \max_{i \in \Omega} \rho_i C_1.$$

Recall from Section 7.3.1 that

$$C = \min_g \max_{l \in \mathcal{L}} D(f_l \| g) \tag{7.18}$$

where  $g$  is a convex combination of  $f_l, l \in \mathcal{L}$ . Let  $g^* = \sum_{l \in \mathcal{L}} \pi_l^* f_l$  be the distribution that achieves the minimum in (7.18). From Fact 5.5.1,

$$D\left(f_k \| \sum_{l \in \mathcal{L}} \pi_l^* f_l\right) = C \quad \text{for any } k \in \mathcal{L} \text{ such that } \pi_k^* > 0. \tag{7.19}$$

By definition of the sample-rich function class, for each  $\mathbf{v} \in \mathcal{L}^M$ , there exists a sample in  $\mathcal{X}$ , say  $x_{\mathbf{v}}$ , that satisfies  $\mathbf{h}(x_{\mathbf{v}}) = \mathbf{v}$ . Let

$$\lambda_{\mathbf{v}} = \prod_{i=1}^M \pi_{v_i}^*.$$

Note that  $\sum_{\mathbf{v} \in \mathcal{L}^M} \lambda_{\mathbf{v}} = 1$ . Moreover, for any  $i, j \in \Omega, i \neq j$ ,

$$\sum_{\mathbf{v} \in \mathcal{L}^M : v_i = k} \lambda_{\mathbf{v}} = \pi_k^*, \quad \sum_{\mathbf{v} \in \mathcal{L}^M : v_i = k, v_j = l} \lambda_{\mathbf{v}} = \pi_k^* \pi_l^*.$$

Using weights  $\{\lambda_{\mathbf{v}}\}_{\mathbf{v} \in \mathcal{L}^M}$  and taking average over all  $\mathbf{v} \in \mathcal{L}^M$ , we obtain

$$\begin{aligned}
\max_{x \in \mathcal{X}} EJS(\boldsymbol{\rho}, x) &\geq \sum_{\mathbf{v}} \lambda_{\mathbf{v}} EJS(\boldsymbol{\rho}, x_{\mathbf{v}}) \\
&= \sum_{\mathbf{v}} \lambda_{\mathbf{v}} \sum_{i=1}^M \rho_i D\left(f_{h_i(x_{\mathbf{v}})} \parallel \sum_{j \neq i} \frac{\rho_j}{1 - \rho_i} f_{h_j(x_{\mathbf{v}})}\right) \\
&= \sum_{i=1}^M \rho_i \sum_{k \in \mathcal{L}} \pi_k^* \sum_{\mathbf{v}: v_i=k} \frac{\lambda_{\mathbf{v}}}{\pi_k^*} D\left(f_k \parallel \sum_{j \neq i} \frac{\rho_j}{1 - \rho_i} f_{v_j}\right) \\
&\stackrel{(a)}{\geq} \sum_{i=1}^M \rho_i \sum_{k \in \mathcal{L}} \pi_k^* D\left(f_k \parallel \sum_{j \neq i} \frac{\rho_j}{1 - \rho_i} \sum_{\mathbf{v}: v_i=k} \frac{\lambda_{\mathbf{v}}}{\pi_k^*} f_{v_j}\right) \\
&= \sum_{i=1}^M \rho_i \sum_{k \in \mathcal{L}} \pi_k^* D\left(f_k \parallel \sum_{j \neq i} \frac{\rho_j}{1 - \rho_i} \sum_{l \in \mathcal{L}} \sum_{\mathbf{v}: v_i=k, v_j=l} \frac{\lambda_{\mathbf{v}}}{\pi_k^*} f_l\right) \\
&= \sum_{i=1}^M \rho_i \sum_{k \in \mathcal{L}} \pi_k^* D\left(f_k \parallel \sum_{l \in \mathcal{L}} \pi_l^* f_l\right) \\
&\stackrel{(b)}{=} \sum_{i=1}^M \rho_i C \\
&= C,
\end{aligned}$$

where (a) follows from Jensen's inequality and (b) follows from (7.19).

Let  $\hat{i} = \arg \max_{i \in \Omega} \rho_i$ . Let  $k, l \in \mathcal{L}$  be the labels satisfying  $D(f_k \parallel f_l) = C_1$ . By definition of the sample-rich function class, there exists a sample  $x_{\hat{i}} \in \mathcal{X}$  that satisfies  $h_{\hat{i}}(x_{\hat{i}}) = k$  and  $h_j(x_{\hat{i}}) = l$  for all  $j \neq \hat{i}$ . We have

$$\max_{x \in \mathcal{X}} EJS(\boldsymbol{\rho}, x) \geq EJS(\boldsymbol{\rho}, x_{\hat{i}}) \geq \rho_{\hat{i}} D\left(f_{h_{\hat{i}}(x_{\hat{i}})} \parallel \sum_{j \neq \hat{i}} \frac{\rho_j}{1 - \rho_{\hat{i}}} f_{h_j(x_{\hat{i}})}\right) = \max_{i \in \Omega} \rho_i C_1.$$

### 7.5.3 Finding $\alpha$ and $\beta$ for Different Function Classes

Let  $g_p(\cdot)$  and  $\bar{g}_p(\cdot)$  be probability density functions on  $\mathcal{Y}$  defined as follows:

$$g_p(y) = \begin{cases} p & \text{if } y = -1 \\ 1 - p & \text{if } y = +1 \end{cases}, \bar{g}_p(y) = g_p(-y). \quad (7.20)$$

It can be easily shown that:

$$C(p) = D(g_p \parallel \frac{g_p + \bar{g}_p}{2}) = D(\bar{g}_p \parallel \frac{g_p + \bar{g}_p}{2}) \quad \text{and} \quad C_1(p) = D(g_p \parallel \bar{g}_p) = D(\bar{g}_p \parallel g_p).$$

The result for the sample-rich class follows from Proposition 7.3.2. Next we provide the proof for the class of disjoint interval functions and threshold functions.

### Class of disjoint interval functions:

To prove this case, we will show that

$$\max_{x \in \mathcal{X}} EJS(\boldsymbol{\rho}, x) \geq \max_{i \in \Omega} \rho_i C_1(p).$$

Let  $\hat{i} = \arg \max_{i \in \Omega} \rho_i$ . By definition of the class of disjoint interval functions, there exists a sample  $x_{\hat{i}} \in \mathcal{X}$  that satisfies  $\mathbf{h}(x_{\hat{i}}) = \mathbf{e}_{\hat{i}}$ . We have

$$EJS(\boldsymbol{\rho}, x_{\hat{i}}) \geq \rho_{\hat{i}} D\left(f_{h_{\hat{i}}(x_{\hat{i}})} \parallel \sum_{j \neq \hat{i}} \frac{\rho_j}{1 - \rho_{\hat{i}}} f_{h_j(x_{\hat{i}})}\right) = \rho_{\hat{i}} D(g_p \parallel \bar{g}_p) = \rho_{\hat{i}} C_1(p).$$

### Class of threshold functions:

We will prove that

$$\max_{x \in \mathcal{X}} EJS(\boldsymbol{\rho}, x) \geq C(p).$$

At any belief vector  $\boldsymbol{\rho} \in \mathbb{P}(\Omega)$ , there exists  $k, k \in \Omega$ , such that  $\sum_{j=1}^k \rho_j \leq \frac{1}{2}$  and  $\sum_{j=1}^{k+1} \rho_j > \frac{1}{2}$ . Let  $x_k$  and  $x_{k+1}$  be samples in  $\mathcal{X}$  that satisfy  $\mathbf{h}(x_k) = \mathbf{u}_k$  and  $\mathbf{h}(x_{k+1}) = \mathbf{u}_{k+1}$  respectively. Let  $\delta_1 = \frac{1}{2} - \sum_{j=1}^k \rho_j$  and  $\delta_2 = \sum_{j=1}^{k+1} \rho_j - \frac{1}{2}$ . Notice that  $\rho_{k+1} = \delta_1 + \delta_2$ . There are two cases:

- Case 1:  $\delta_1 \leq \delta_2$ . We have

$$\begin{aligned} EJS(\boldsymbol{\rho}, x_k) &= \sum_{i=1}^M \rho_i D\left(f_{h_i(x_k)} \parallel \sum_{j \neq i} \frac{\rho_j}{1 - \rho_i} f_{h_j(x_k)}\right) \\ &= \sum_{i=1}^k \rho_i D\left(\bar{g}_p \parallel \frac{1/2 - \delta_1 - \rho_i}{1 - \rho_i} \bar{g}_p + \frac{1/2 + \delta_1}{1 - \rho_i} g_p\right) \\ &\quad + \rho_{k+1} D\left(g_p \parallel \frac{1/2 - \delta_1}{1 - \rho_{k+1}} \bar{g}_p + \frac{1/2 - \delta_2}{1 - \rho_{k+1}} g_p\right) \\ &\quad + \sum_{i=k+2}^M \rho_i D\left(g_p \parallel \frac{1/2 - \delta_1}{1 - \rho_i} \bar{g}_p + \frac{1/2 + \delta_1 - \rho_i}{1 - \rho_i} g_p\right) \end{aligned}$$

$$\begin{aligned}
&\stackrel{(a)}{\geq} (1/2 - \delta_1)D\left(g_p\|(1/2 + \delta_1)\bar{g}_p + (1/2 - \delta_1)g_p\right) \\
&\quad + (\delta_1 + \delta_2)D\left(g_p\|\frac{1}{2}\bar{g}_p + \frac{1}{2}g_p\right) \\
&\quad + (1/2 - \delta_2)D\left(g_p\|(1/2 - \delta_1)\bar{g}_p + (1/2 + \delta_1)g_p\right) \\
&\stackrel{(b)}{\geq} D\left(g_p\|(1 - \gamma)\bar{g}_p + \gamma g_p\right) \\
&\stackrel{(c)}{\geq} D\left(g_p\|\frac{1}{2}\bar{g}_p + \frac{1}{2}g_p\right) \\
&= C(p),
\end{aligned}$$

where

$$\gamma = (1/2 - \delta_1)^2 + \frac{1}{2}(\delta_1 + \delta_2) + (1/2 - \delta_2)(1/2 + \delta_1),$$

inequality (a) follows from Lemma 3.3.1 and (7.20), (b) follows since KL divergence is convex, and (c) follows from the fact that  $\gamma = \frac{1}{2} + \delta_1(\delta_1 - \delta_2) \leq \frac{1}{2}$  and by Lemma 3.3.1.

- Case 2:  $\delta_1 > \delta_2$ . We have

$$\begin{aligned}
EJS(\boldsymbol{\rho}, x_{k+1}) &= \sum_{i=1}^k \rho_i D\left(\bar{g}_p\|\frac{1/2 + \delta_2 - \rho_i}{1 - \rho_i}\bar{g}_p + \frac{1/2 - \delta_2}{1 - \rho_i}g_p\right) \\
&\quad + \rho_{k+1} D\left(\bar{g}_p\|\frac{1/2 - \delta_1}{1 - \rho_{k+1}}\bar{g}_p + \frac{1/2 - \delta_2}{1 - \rho_{k+1}}g_p\right) \\
&\quad + \sum_{i=k+2}^M \rho_i D\left(g_p\|\frac{1/2 + \delta_2}{1 - \rho_i}\bar{g}_p + \frac{1/2 - \delta_2 - \rho_i}{1 - \rho_i}g_p\right) \\
&\stackrel{(a)}{\geq} (1/2 - \delta_1)D\left(g_p\|(1/2 - \delta_2)\bar{g}_p + (1/2 + \delta_2)g_p\right) \\
&\quad + (\delta_1 + \delta_2)D\left(g_p\|\frac{1}{2}\bar{g}_p + \frac{1}{2}g_p\right) \\
&\quad + (1/2 - \delta_2)D\left(g_p\|(1/2 + \delta_2)\bar{g}_p + (1/2 - \delta_2)g_p\right) \\
&\stackrel{(b)}{\geq} D\left(g_p\|(1 - \gamma')\bar{g}_p + \gamma'g_p\right) \\
&\stackrel{(c)}{\geq} D\left(g_p\|\frac{1}{2}\bar{g}_p + \frac{1}{2}g_p\right) \\
&= C(p),
\end{aligned}$$

where

$$\gamma' = (1/2 - \delta_1)(1/2 + \delta_2) + \frac{1}{2}(\delta_1 + \delta_2) + (1/2 - \delta_2)^2,$$

inequality (a) follows from Lemma 3.3.1 and (7.20), (b) follows since KL divergence is convex, and (c) follows from the fact that  $\gamma' = \frac{1}{2} + \delta_2(\delta_2 - \delta_1) < \frac{1}{2}$  and by Lemma 3.3.1.

Therefore,

$$\max_{x \in \mathcal{X}} EJS(\boldsymbol{\rho}, x) \geq \max \{EJS(\boldsymbol{\rho}, x_k), EJS(\boldsymbol{\rho}, x_{k+1})\} \geq C(p).$$

Chapter 7, in full, is currently being prepared for submission for publication as M. Naghshvar, T. Javidi, and K. Chaudhuri, “Noisy Bayesian active learning.” The dissertation author was the primary investigator and author of this paper.

# Bibliography

- [1] A. Wald and J. Wolfowitz. Optimum character of the sequential probability ratio tests. *The Annals of Mathematical Statistics*, 19(3):326–339, 1948.
- [2] P. Armitage. Sequential analysis with more than two alternative hypotheses, and its relation to discriminant function analysis. *Journal of the Royal Statistical Society, Series B*, 12(1):137–144, 1950.
- [3] G. Lorden. Nearly-optimal sequential tests for finitely many parameter values. *The Annals of Statistics*, 5(1):1–21, 1977.
- [4] V. P. Dragalin, A. G. Tartakovsky, and V. V. Veeravalli. Multihypothesis sequential probability ratio tests. I. Asymptotic optimality. *IEEE Transactions on Information Theory*, 45(7):2448–2461, November 1999.
- [5] H. Chernoff. Sequential design of experiments. *The Annals of Mathematical Statistics*, 30:755–770, 1959.
- [6] S. M. Berry, B. P. Carlin, J. J. Lee, and P. Müller. *Bayesian adaptive methods for clinical trials*, volume 38 of *Chapman & Hall/CRC Biostatistics Series*. CRC Press, Boca Raton, FL, 2011.
- [7] P. Shenoy and A. J. Yu. Rational decision-making in inhibitory control. *Frontiers in Human Neuroscience*, 5(48), 2011.
- [8] A. O. Hero and D. Cochran. Sensor management: past, present, and future. *IEEE Sensors Journal*, 11(12):3064–3075, December 2011.
- [9] G. A. Hollinger, U. Mitra, and G. S. Sukhatme. Active classification: Theory and application to underwater inspection. In *International Symposium on Robotics Research*, 2011.
- [10] R. D. Nowak. The geometry of generalized binary search. *IEEE Transactions on Information Theory*, 57(12):7893–7906, December 2011.
- [11] C. L. Chan, P. H. Che, S. Jaggi, and V. Saligrama. Non-adaptive probabilistic group testing with noisy measurements: near-optimal bounds with efficient

- algorithms. In *49th Annual Allerton Conference on Communication, Control, and Computing*, pages 1832–1839, 2011.
- [12] M. V. Burnashev. Data transmission over a discrete channel with feedback. Random transmission time. *Problemy Peredachi Informatsii*, 12(4):10–30, 1975.
  - [13] D. Blackwell. Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, 24:265–272, 1953.
  - [14] D. V. Lindley. On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, 27(4):986–1005, 1956.
  - [15] L. Le. Cam. Sufficiency and approximate sufficiency. *The Annals of Mathematical Statistics*, 35(4):1419–1455, 1964.
  - [16] M. H. DeGroot. *Optimal statistical decisions*. McGraw-Hill Book Co., New York, NY, 1970.
  - [17] K. Goel and M. H. DeGroot. Comparison of experiments and information measures. *The Annals of Statistics*, 7(5):1066–1077, 1979.
  - [18] E. L. Lehmann. Comparing location experiments. *The Annals of Statistics*, 16(2):521–533, 1988.
  - [19] E. Torgersen. *Stochastic orders and comparison of experiments*, volume 19 of *IMS Lecture Notes - Monograph Series*, pages 334–371. Institute of Mathematical Statistics, Hayward, CA, 1991.
  - [20] M. H. DeGroot. Uncertainty, information, and sequential experiments. *The Annals of Mathematical Statistics*, 33(2):404–419, 1962.
  - [21] M. Horstein. Sequential transmission using noiseless feedback. *IEEE Transactions on Information Theory*, 9(3):136–143, July 1963.
  - [22] M. V. Burnashev and K. Sh. Zigangirov. An interval estimation problem for controlled observations. *Problemy Peredachi Informatsii*, 10(3):51–61, 1974.
  - [23] M. Malloy and R. Nowak. Sequential analysis in high-dimensional multiple testing and sparse recovery. In *IEEE International Symposium on Information Theory (ISIT)*, pages 2661–2665, 2011.
  - [24] M. A. Iwen. Group testing strategies for recovery of sparse signals in noise. In *43rd Asilomar Conference on Signals, Systems and Computers*.
  - [25] S. A. Bessler. Theory and applications of the sequential design of experiments,  $K$ -Actions and infinitely many experiments: Part I — Theory. *Technical Report no. 55, Department of Statistics, Stanford University*, 1960.

- [26] R. E. Blahut. Hypothesis testing and information theory. *IEEE Transactions on Information Theory*, 20(4):405–417, July 1974.
- [27] E. Tuncel. On error exponents in hypothesis testing. *IEEE Transactions on Information Theory*, 51(8):2945–2950, August 2005.
- [28] M. Hayashi. Discrimination of two channels by adaptive methods and its application to quantum system. *IEEE Transactions on Information Theory*, 55(8):3807–3820, August 2009.
- [29] Y. Polyanskiy and S. Verdú. Hypothesis testing with feedback. In *Information Theory and Applications Workshop (ITA)*, 2011.
- [30] S. Nitinawarat, G. Atia, and V. V. Veeravalli. Controlled sensing for multi-hypothesis testing. 2012. available on arXiv:1205.0858v4.
- [31] P. R. Kumar and P. Varaiya. *Stochastic systems: estimation, identification, and adaptive control*. Prentice-Hall, Inc., Upper Saddle River, NJ, 1986.
- [32] A. E. Albert. The sequential design of experiments for infinitely many states of nature. *The Annals of Mathematical Statistics*, 32:774–799, 1961.
- [33] J. Kiefer and J. Sacks. Asymptotically optimum sequential inference and design. *The Annals of Mathematical Statistics*, 34(3):705–750, 1963.
- [34] W. J. Blot and D. A. Meeter. Sequential experimental design procedures. *Journal of the American Statistical Association*, 68(343):586–593, 1973.
- [35] R. Keener. Second order efficiency in the sequential design of experiments. *The Annals of Statistics*, 12(2):510–532, 1984.
- [36] S. P. Lalley and G. Lorden. A control problem arising in the sequential design of experiments. *The Annals of Probability*, 14(1):136–172, 1986.
- [37] M. Naghshvar and T. Javidi. Information utility in active sequential hypothesis testing. In *48th Annual Allerton Conference on Communication, Control, and Computing*, 2010.
- [38] M. Naghshvar and T. Javidi. Active sequential hypothesis testing. 2012. available on arXiv:1203.4626.
- [39] E. A. Haroutunian, M. E. Haroutunian, and A. N. Harutyunyan. Reliability criteria in information theory and in statistical hypothesis testing. *Foundations and Trends in Communications and Information Theory*, 4(2):97–263, January 2007.
- [40] O. Shayevitz. On Rényi measures and hypothesis testing. In *IEEE International Symposium on Information Theory (ISIT)*, pages 894–898, 2011.

- [41] T. van Erven and P. Harremoës. Rényi divergence and Kullback–Leibler divergence. 2012. available on arXiv:1206.2459v1.
- [42] M. Sakaguchi. Information theory and decision making. 1964. Unpublished lecture notes, Statistic Department, The George Washington University, Washington, D.C.
- [43] I. Csiszár and P. C. Shields. Information theory and statistics: a tutorial. *Foundations and Trends in Communications and Information Theory*, 1(4):417–528, December 2004.
- [44] E. R. Berlekamp. Block coding with noiseless feedback. 1964. PhD Thesis, MIT, Cambridge, MA.
- [45] B. Nakiboglu. Exponential bounds on error probability with feedback. 2011. PhD Thesis, MIT, Cambridge, MA.
- [46] F. Liese and I. Vajda. *Convex statistical distances*, volume 95 of *Teubner-Texte zur Mathematik*. Teubner, Leipzig, 1987.
- [47] P. Billingsley. *Probability and measure*. John Wiley & Sons Inc., New York, NY, third edition, 1995.
- [48] C. McDiarmid. On the method of bounded differences. *Surveys in Combinatorics, London Mathematical Society Lecture Note Series 141*, Cambridge University Press, pages 148–188, 1989.
- [49] D. P. Bertsekas and S. E. Shreve. *Stochastic optimal control: The discrete-time case*. Athena Scientific, Belmont, MA, 2007.
- [50] H. Jeffreys. An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society. London. Series A.*, 186:453–461, 1946.
- [51] J. Lin. Divergence measures based on the Shannon entropy. *IEEE Transactions on Information Theory*, 37(1):145–151, January 1991.
- [52] J. Burbea and C. R. Rao. On the convexity of some divergence measures based on entropy functions. *IEEE Transactions on Information Theory*, 28(3):489–495, May 1982.
- [53] G. T. Toussaint. Some functional lower bounds on the expected divergence for multihypothesis pattern recognition, communication, and radar systems. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-1:384–385, 1971.
- [54] M. C. Gastpar, P. R. Gill, A. G. Huth, and F. E. Theunissen. Anthropic correction of information estimates and its application to neural coding. *IEEE Transactions on Information Theory*, 56(2):890–900, February 2010.

- [55] D. P. Bertsekas. *Dynamic Programming and Optimal Control, Volume II*. Athena Scientific, Belmont, MA, 1995.
- [56] Y. Polyanskiy, H. V. Poor, and S. Verdu. Feedback in the non-asymptotic regime. *IEEE Transactions on Information Theory*, 57(8):4903–4925, August 2011.
- [57] G. Lorden. On excess over the boundary. *Annals of Mathematical Statistics*, 41:520–527, 1970.
- [58] M. V. Burnashev and K. Sh. Zigangirov. On one problem of observation control. *Problemy Peredachi Informatsii*, 11(3):44–52, 1975.
- [59] M. V. Burnashev. Sequential discrimination of hypotheses with control of observations. *Mathematics of the USSR-Izvestiya*, 15(3):419–440, 1980.
- [60] H. Yamamoto and K. Itoh. Asymptotic performance of a modified Schalkwijk–Barron scheme for channels with noiseless feedback. *IEEE Transactions on Information Theory*, 25:729–733, 1979.
- [61] J. M. Ooi and G. W. Wornell. Fast iterative coding techniques for feedback channels. *IEEE Transactions on Information Theory*, 44(7):2960–2976, November 1998.
- [62] G. Caire, S. Shamai, and S. Verdu. Propagation, feedback and belief. *4th International Symposium on Turbo Codes & Related Topics; 6th International ITG-Conference on Source and Channel Coding (TURBOCODING)*, pages 1–6, April 2006.
- [63] A. Tchamkerten and E. Telatar. Variable length coding over an unknown channel. *IEEE Transactions on Information Theory*, 52:2126–2145, 2006.
- [64] O. Shayevitz and M. Feder. Optimal feedback communication via posterior matching. *IEEE Transactions on Information Theory*, 57(3):1186–1222, March 2011.
- [65] T. M. Cover and J. A. Thomas. *Elements of information theory*. John Wiley & Sons, Inc., Hoboken, NJ, second edition, 2006.
- [66] P. Berlin, B. Nakiboglu, B. Rimoldi, and E. Telatar. A simple converse of Burnashev’s reliability function. *IEEE Transactions on Information Theory*, 55:3074–3080, 2009.
- [67] H. S. Witsenhausen. A counterexample in stochastic optimum control. *SIAM Journal on Control*, 6(1):131–147, 1968.

- [68] A. Mahajan, A. Nayyar, and D. Teneketzis. Identifying tractable decentralized problems on the basis of information structures. In *46th Annual Allerton Conference on Communication, Control, and Computing*, pages 1440–1449, 2008.
- [69] D. P. Bertsekas and J. N. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, Belmont, MA, 1996.
- [70] R. G. Gallager. *Information theory and reliable communication*. John Wiley & Sons, Inc., New York, NY, 1968.
- [71] J. A. Nachlas, S. R. Loney, and B. A. Binney. Diagnostic-strategy selection for series systems. *IEEE Transactions on Reliability*, 39(3):273–280, 1990.
- [72] D. A. Castanon. Optimal search strategies in dynamic hypothesis testing. *IEEE Transactions on Systems, Man and Cybernetics*, 25(7):1130–1138, July 1995.
- [73] K. P. Tognetti. An optimal strategy for a whereabouts search. *Operations Research*, 16(1):209–211, 1968.
- [74] J. B. Kadane. Optimal whereabouts search. *Operations Research*, 19(4):894–904, 1971.
- [75] D. Sejdinovic and O. Johnson. Note on noisy group testing: Asymptotic bounds and belief propagation reconstruction. In *48th Annual Allerton Conference on Communication, Control, and Computing*, pages 998–1003, 2010.
- [76] E. Posner. Optimal search procedures. *IEEE Transactions on Information Theory*, 9(3):157–160, July 1963.
- [77] L. D. Stone. *Theory of optimal search*. Academic Press, New York, NY, 1975.
- [78] R. D. Nowak. Noisy generalized binary search. In *Neural Information Processing Systems (NIPS)*, 2009.
- [79] S. Dasgupta. Coarse sample complexity bounds for active learning. In *Neural Information Processing Systems (NIPS)*, 2005.
- [80] S. Hanneke. A bound on the label complexity of agnostic active learning. In *24th Annual International Conference on Machine Learning (ICML)*, pages 353–360, 2007.
- [81] S. Dasgupta. Two faces of active learning. *Theoretical Computer Science*, 412(19):1767–1781, April 2011.
- [82] M. Kääriäinen. Active learning in the non-realizable case. In *17th International Conference on Algorithmic Learning Theory (ALT)*, pages 63–77, 2006.

- [83] M. F. Balcan, A. Beygelzimer, and J. Langford. Agnostic active learning. *Journal of Computer and System Sciences*, 75(1):78–89, 2009.
- [84] A. Beygelzimer, S. Dasgupta, and J. Langford. Importance weighted active learning. In *26th Annual International Conference on Machine Learning (ICML)*, pages 49–56, 2009.
- [85] A. Beygelzimer, D. Hsu, J. Langford, and T. Zhang. Agnostic active learning without constraints. In *Advances in Neural Information Processing Systems (NIPS)*, 2010.
- [86] M. Raginsky and A. Rakhlin. Lower bounds for passive and active learning. In *Neural Information Processing Systems (NIPS)*, 2011.
- [87] Y. Sakakibara. On learning from queries and counterexamples in the presence of noise. *Information Processing Letters*, 37(5):279–284, March 1991.