Institut Polytechnique de Paris

M1 Internship Report

Quantum Guessing Games

Author:
Ali Almasi

Supervisors:
Peter Brown
Cambyse Rouzé

August 2024

"The more constraints one imposes, the more one frees one's self of the chains that shackle the spirit."

Igor Stravinsky

"All the business of war, and indeed all the business of life, is to endeavour to find out what you don't know by what you do; that's what I called guessing what was at the other side of the hill."

Duke of Wellington

INSTITUT POLYTECHNIQUE DE PARIS

Abstract

Quantum Guessing Games

by Ali ALMASI

In the present work, we study quantum guessing games, a generalization of two well-studied problems in quantum information theory, namely quantum state discrimination and quantum state antidiscrimination. We begin by extending some of the known results about those two special cases to the more general setting of quantum guessing games. In the second part of the thesis, we use semidefinite programming and the theory of Minorization-Maximization (MM) algorithms to introduce iterative algorithms for quantum guessing games and investigate their convergence.

Acknowledgements

First and foremost, I want to express my heartfelt thanks to my supervisors, Peter and Cambyse. I feel incredibly fortunate to have had such wonderful people supervising me—not only for the great mathematical insights they freely shared and their patient help in improving my research abilities and deepening my mathematical understanding, but also for their genuine kindness and unwavering support throughout my internship.

I would also like to thank Titouan Carette for regularly checking on my progress and kindly agreeing to be my internship tutor. Additionally, I am grateful to Thomas Van Himbeeck, Jan Kochanowski, Paul Gondolf, and Nasra Daher Ahmed for the fruitful discussions we had, which are reflected in this work.

A special thanks goes to Niyousha for their kindness and patience in helping me with my writing during about half of my internship when my hand was broken.

Contents

List of Abbreviations

lsc	lower semi-continuous
LP	Linear Programming / Linear Program
PGM	Pretty Good Measurement
POVM	Positive Operator Valued Measure
PSD	Positive Semidefinite
QSD	Quantum State Discrimination
SDP	Semidefinite Programming / Semidefinite Program

Chapter 1

Introduction

Background and Motivation In quantum information theory, information is encoded in the states of quantum systems, and processing this information involves manipulating the state via quantum operations. Once processing is complete, extracting the encoded information requires determining the system's state by performing quantum measurements. However, due to the principles of quantum mechanics, it is not often possible to find a quantum measurement that can perfectly distinguish all possible states of the system after processing. This challenge brings the concept of *quantum state discrimination* (QSD) to the forefront. Quantum state discrimination is the task of finding the optimal measurement to distinguish between a set of quantum states according to a specified figure of merit.

The study of quantum state discrimination began in the late 1960s, with Helstrom's pioneering works [Hel67; Hel69]. Since then, it has attracted considerable attention and found several applications in quantum information theory [BK02; Hay06; Spe14], quantum query complexity [Mon07], quantum cryptography [DJL00; Ber07], and quantum computing [BCD05].

Researchers have considered various figures of merit in these studies, leading to the development of different strategies for quantum state discrimination, including minimum-error discrimination [Hel69], unambiguous discrimination [Iva87], maximum confidence discrimination [Cro+06], quantum hypothesis testing [Hay06], and quantum guesswork discrimination [Che+15]. Among these, minimum-error discrimination is the most well-known, which aims to minimize the probability of error when distinguishing between states. In this thesis, unless otherwise stated, "quantum state discrimination" will refer to minimum-error discrimination.

Finding the optimal success probability and measurement for quantum state discrimination is the central focus of the literature on this topic. To this end, two main approaches have been explored. The first approach seeks closed-form solutions, which are only known for a few special cases, such as the discrimination of two states [Hel67; Hol72], equiprobable qubits [DT10], geometrically uniform states [EMV04], and mirror-symmetric states [And+02]. The second approach involves developing numerical schemes to find solutions with arbitrary precision.

These numerical schemes fall into two categories. The first category is based on *semidefinite programming* (SDP) [JŘF02; Eld03; WGM08; SC23], a powerful tool for solving many quantum information problems. The availability of efficient SDP solvers [Ali91; NN92; NN94] has made it possible to numerically solve the quantum state discrimination problem for reasonable size systems. The second category comprises iterative algorithms specifically designed for quantum state discrimination [JŘF02; Tys09a; NKU15]. These iterative algorithms are often derived from the problem's optimality conditions and can sometimes outperform SDP solvers in terms of efficiency.

While the problem of quantum state discrimination primarily focuses on finding the optimal success probability and the corresponding measurement, in many real-world applications, obtaining a good bound on the success probability or a near-optimal measurement is often sufficient. Several studies have focused on finding upper and lower bounds on the optimal success probability [NS06; HKK06; HB10; QL10; NUK18; Lou22]. Some of these bounds are derived from explicit measurements that are not necessarily optimal but perform pretty well on this task [BK02; Mon07; Tys09b]. Notable examples include the pretty good measurement (also known as the square-root measurement) [Hau+96; PW91] and the Holevo measurement (also known as the quadratically-weighted measurement) [Hol79]. Both of these are special cases within a broader class of measurements known as Belavkin measurements, which also includes the optimal measurement [BM04].

A closely related problem to quantum state discrimination is *quantum state antidiscrimination*, also known as quantum state elimination or exclusion, which has recently attracted the attention of researchers [CFS02]. In the quantum state antidiscrimination problem, the goal is to determine which state the quantum system is not in, excluding one state from the set of possible candidates. This problem has found applications in quantum cryptography [DWA14], quantum communication complexity [HB20], and the foundations of quantum mechanics [PBR12; Lei14]. Similar to the work done on quantum state discrimination, research on state antidiscrimination has explored various aspects, such as establishing conditions for perfect antidiscrimination [HK18; JRS23], formulating the problem as a semidefinite program (SDP) [Ban+14], developing a pretty good measurement for antidiscrimination [MMS24], and studying error exponents in the asymptotic regime [MNW24].

Although quantum state antidiscrimination can be framed as a state discrimination problem [Ban+14; MMS24], and the two share several similarities, antidiscrimination has distinct characteristics that set it apart. For instance, the conditions for perfect antidiscrimination are different and more complicated than those for perfect discrimination. The study of antidiscrimination is still in its early stages, and many questions remain unanswered, potentially revealing fundamental differences or deeper connections between the two problems.

Research Problem and Contributions In this thesis, we explore a generalized formulation of the quantum state discrimination and antidiscrimination problems, known as *quantum guessing games*. In quantum guessing games, two parties, Alice and Bob, engage in the following scenario: Alice selects a quantum state from an ensemble she possesses and sends it to Bob. Bob's task is to guess which state Alice has chosen, using a quantum measurement on the received state. Each possible state Alice selects and each guess Bob makes has an associated reward. Bob's objective is to devise a guessing strategy—a measurement—that maximizes his expected reward. This formulation has appeared in several works related to quantum state discrimination (see e.g. [Hel69]), but the study of this problem in such a general context has only recently begun [CHT22; MSU23].

The first contribution of this thesis is establishing a standard form for quantum guessing games, to which any quantum guessing game can be reduced. In Chapter 2, we introduce this standard form and demonstrate that every quantum guessing game can be transformed into a quantum state discrimination problem, highlighting the significance of studying quantum state discrimination. While this reduction is possible, it does not preserve all properties of the original quantum guessing games, meaning the reduction does not trivialize the study of quantum guessing games. For instance, an ensemble that can be perfectly antidiscriminated may not be reduced to one that can be perfectly discriminated.

The second contribution of this work is the generalization of the pretty good measurement, previously studied in the context of quantum state discrimination, to quantum guessing games. This is done in Chapter 3. In the specific case of quantum state antidiscrimination, this generalization was recently explored by McIrvin et al. in [MMS24] under the term "*pretty bad measurements*." Our definition of pretty good measurements extends both the definitions of pretty good measurements for discrimination and pretty bad measurements for antidiscrimination. Additionally, we derive a lower bound on the success probability of the pretty good measurement and an upper bound on the optimal success probability in terms of the pretty good measurement in the general case of quantum guessing games, extending the bounds obtained in [MMS24; BK02].

Our third contribution is studying a class of iterative algorithms for finding the optimal measurement in quantum guessing games. We begin by formulating quantum guessing games as semidefinite programs and deriving their optimality conditions in Chapter 4. Building on these conditions, in Section 5.1, we generalize the algorithm proposed by Ježek, Řeháček, and Fiurášek [JŘF02] to the context of quantum guessing games and provide numerical evidence supporting the convergence of the algorithm.

Since the convergence of the original algorithm has not been proven for general quantum ensembles yet, we focus on studying its convergence in the case of quantum state discrimination. Using previously established results in [Tys10], we connect the algorithm to a well-known class of iterative optimization methods called minorization-maximization (MM) algorithms in Section 5.3. We then analyze the fixed points of both the algorithm and its MM counterpart in Section 5.4, proposing a potential direction for proving convergence. To the best of our knowledge, this direction has not been explored in the literature.

Outline

- In Chapter 2, we introduce quantum guessing games and establish a standard form for them. Using this standard form, we show that any quantum guessing game can be reduced to a quantum state discrimination problem. Following this, we review key foundational results for both quantum state discrimination and antidiscrimination. The chapter concludes by discussing how quantum guessing games naturally arise in other quantum information problems.
- In Chapter 3, we investigate the pretty good measurement in the context of quantum guessing games. We derive a lower bound on the success probability of the pretty good measurement and an upper bound on the optimal success probability in terms of the pretty good measurement, extending previously known bounds for quantum state discrimination. We then set the stage for our upcoming discussion on iterative algorithms, by reviewing the Belavkin measurements, a more general class of measurements that includes the pretty good measurement.
- In Chapter 4, we analyze quantum guessing games by formulating them as semidefinite programs and derive the optimality conditions for these problems.
- In Chapter 5, we study an iterative algorithm for determining the optimal measurement in quantum guessing games. We begin by generalizing the algorithm proposed by Ježek, Řeháček, and Fiurášek [JŘF02] to quantum guessing games and analyze its convergence numerically. We then examine the convergence of this algorithm specifically for quantum state discrimination, reviewing two modified versions proposed in [Tys10; NKU15]. Our focus then shifts to the one proposed in [Tys10], discussing its interpretation as a minorization-maximization (MM) algorithm. The chapter ends with an analysis of the algorithms' fixed points and a suggestion for a potential approach towards proving their convergence.
- In Chapter 6, we conclude the thesis and discuss some possible directions for future research.
- We also include some basic definitions and results from linear algebra and convex analysis that are needed throughout the thesis in Appendix A.

Chapter 2

Quantum Guessing Games

In this chapter, we introduce the concept of quantum guessing games, a generalization of the quantum state discrimination problem. We also survey some of the most important results known about two special instances of quantum guessing games: the quantum state discrimination problem and the quantum state antidiscrimination problem. Although the notion of quantum guessing games has attracted the attention of researchers quite recently [CHT22; MSU23], one can trace back the roots of this general formulation to the seminal work of Helstrom [Hel69] and since then, it has also appeared in several other works [Che+15; Wat18], not necessarily under the name of quantum guessing games.

2.1 Problem Formulation

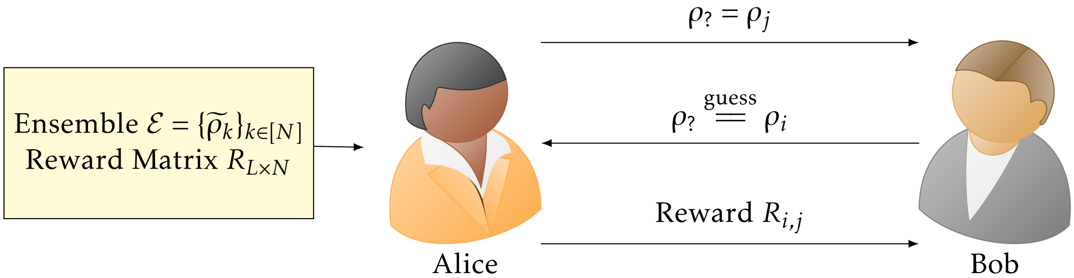
In the quantum state discrimination problem, Alice prepares a quantum state from a known ensemble and sends it to Bob, who performs a quantum measurement to determine the state. The goal is to maximize the probability of Bob correctly identifying the state. In the quantum guessing games, we generalize this problem by assigning rewards to the outcomes of Bob's measurement. The formulation we present here is similar to the one of [MSU23].

Suppose that Alice possesses an ensemble $\mathcal{E} = \{\widetilde{\rho}_k\}_{k\in[N]}$, and she picks a state ρ_j from this ensemble and sends it to Bob, who performs a POVM measurement $M = \{M_s\}_{s\in[L]}$ on the receiving state, obtaining a measurement output $i \in [M]$. For each $i \in [L]$ and $j \in [N]$, if the receiving state is ρ_j and Bob's measurement output is i, he obtains a reward $R_{i,j} \in \mathbb{R}$ (see Figure 2.1). Bob wants to maximize his expected reward obtained by performing a POVM M, which is denoted by $\mathcal{R}_M(G)$ (or simply $\mathcal{R}_M$ when the game is clear from the context), subject to all POVM measurements with L outputs he can perform, meaning that he wants to maximize

$$\mathcal{R}_M \overset{\text{def}}{=} \sum_{i=1}^{L}\sum_{j=1}^{N} R_{i,j}\,\mathrm{tr}(M_i\widetilde{\rho}_j), \tag{2.1}$$

over all POVMs $M = \{M_k\}_{k\in[L]}$.

FIGURE 2.1: A schematic representation of a quantum guessing game. For guessing the unknown state $\rho_?$, Bob can use classical probabilistic strategies, as well as quantum strategies (i.e. performing a quantum measurement on the received state). It is easy to see that the classical probabilistic strategies are a special case of quantum strategies.



Definition 2.1.1 (Quantum Guessing Game)**.** A quantum guessing game is a pair $G = (\mathcal{E}, R)$, where $\mathcal{E} = \{\widetilde{\rho}_k\}_{k\in[N]}$ is a quantum ensemble and $R \in \mathbb{R}^{L\times N}$ is a matrix, with which the following maximization problem is associated:

$$\begin{aligned} \text{maximize}: \ & \mathcal{R}_M = \sum_{i=1}^{L}\sum_{j=1}^{N} R_{i,j}\,\mathrm{tr}(M_i\widetilde{\rho}_j), \\ \text{subject to}: \ & \sum_{k=1}^{N} M_k = \mathbb{1}, \\ & M_k \geq 0 \quad k \in \{1,2,\dots,L\}. \end{aligned} \tag{2.2}$$

The optimal value of this optimization problem is called the *optimal reward*, denoted by $\mathcal{R}_{\text{opt}}(G)$ (or $\mathcal{R}_{\text{opt}}$ when the game is clear from the context), and the optimal solution is called the *optimal measurement*.

Remark 2.1.2. Problem (2.2) is a linear optimization problem over the convex set of POVMs. Thus, every optimal solution for this problem is an extreme point of the feasible set. Notably, characterization theorems for the extremality of POVMs are known [see, e.g., HP20], which can, in principle, be used to find the optimal POVMs for quantum guessing games.

2.2 Normal Forms

The rewards can be arbitrary real numbers in the general form of quantum guessing games introduced in Definition 2.1.1. However, it is easier to analyze the problem when the rewards are restricted to the interval $[0,1]$. In this section, we show that any quantum guessing game can be reduced into a standard form, where the rewards are in $[0,1]$.

Proposition 2.2.1. *Let $G = (\mathcal{E}, R_{L\times N})$ be a quantum guessing game.*

1. *There exists a quantum guessing game $G' = (\mathcal{E}, R'_{L\times N})$, such that for all $(i,j) \in [L]\times[N]$, $R'_{i,j} \geq 0$, and the optimal measurements of G and G' coincide.*
2. *There exists a quantum guessing game $G'' = (\mathcal{E}, R''_{L\times N})$, such that for all $(i,j) \in [L]\times[N]$, $R''_{i,j} \in [0,1]$, and the optimal measurements of G and G'' coincide.*

Proof. 1. Without loss of generality, suppose that there exists $(s,t) \in [L]\times[N]$ such that $R_{s,t} < 0$, and for all $(i,j) \neq (s,t)$, $R_{i,j} \geq 0$. Let $G' = (\mathcal{E}, R'_{L\times N})$ be a quantum guessing game, where R' is defined as follows:

$$R'_{i,j} = \begin{cases} R_{i,j} & j \neq t \\ R_{i,j} - R_{s,t} & j = t \end{cases} \tag{2.3}$$

Note that for all (i,j), $R'_{i,j} \geq 0$. Moreover, for any POVM $M = \{M_k\}_{k\in[L]}$ we have:

$$\mathcal{R}'_M = \sum_{i=1}^{L}\sum_{j=1}^{N} R'_{i,j}\,\mathrm{tr}(M_i\widetilde{\rho}_j) \tag{2.4}$$

$$= \sum_{i=1}^{L}\Big((R_{i,t} - R_{s,t})\,\mathrm{tr}(M_i\widetilde{\rho}_t) + \sum_{j\neq t} R_{i,j}\,\mathrm{tr}(M_i\widetilde{\rho}_j)\Big) \tag{2.5}$$

$$= \sum_{i=1}^{L}\Big(-R_{s,t}\,\mathrm{tr}(M_i\widetilde{\rho}_t) + \sum_{j=1}^{N} R_{i,j}\,\mathrm{tr}(M_i\widetilde{\rho}_j)\Big) \tag{2.6}$$

$$= \sum_{i=1}^{L} -R_{s,t}\,\mathrm{tr}(M_i\widetilde{\rho}_t) + \sum_{i=1}^{L}\sum_{j=1}^{N} R_{i,j}\,\mathrm{tr}(M_i\widetilde{\rho}_j) \tag{2.7}$$

$$= -R_{s,t}\,\mathrm{tr}\Big(\sum_{i=1}^{L} M_i\widetilde{\rho}_t\Big) + \mathcal{R}_M \tag{2.8}$$

$$= -R_{s,t}\,\mathrm{tr}(\mathbb{1}\widetilde{\rho}_t) + \mathcal{R}_M \tag{2.9}$$
$$= -R_{s,t}q_t + \mathcal{R}_M \tag{2.10}$$

Note that $-R_{s,t}q_t$ is a constant that does not depend on M, which implies that the optimal measurements of G and G' are equal.

2. Having an equivalent quantum guessing game $G' = (\mathcal{E}, R'_{L\times N})$ from the previous part, consider the quantum guessing game $G'' = (\mathcal{E}, R''_{L\times N})$, where $R''_{i,j} \overset{\text{def}}{=} \frac{R'_{i,j}}{\max_{i,j} R'_{i,j}}$. The expected reward of G'' when Bob performs a measurement M is

$$\mathcal{R}''_M = \sum_{i=1}^{L}\sum_{j=1}^{N} R''_{i,j}\,\mathrm{tr}(M_i\widetilde{\rho}_j) \tag{2.11}$$
$$= (\max_{i,j} R'_{i,j})^{-1}\sum_{i=1}^{L}\sum_{j=1}^{N} R'_{i,j}\,\mathrm{tr}(M_i\widetilde{\rho}_j) \tag{2.12}$$
$$= (\max_{i,j} R'_{i,j})^{-1}\mathcal{R}'_M. \tag{2.13}$$

As $(\max_{i,j} R'_{i,j})^{-1}$ is a constant, the optimal measurements of G' and G'' remain the same. Moreover, we have $R''_{i,j} \in [0,1]$ for all (i,j) by our construction.

□

Definition 2.2.2. A quantum guessing game $G = (\mathcal{E}, R_{L\times N})$ is called standard, if for all $(i,j) \in [L]\times[N]$, $R_{i,j} \in [0,1]$.

In the rest of this thesis, all quantum guessing games are assumed to be standard unless otherwise stated.

Example 2.2.3 (State Discrimination and Antidiscrimination Problems)**.** A well-studied special case of quantum guessing games is the minimum-error state discrimination problem, which can be formulated as a quantum guessing game with a reward matrix $R_{N\times N}$ of the following form:

$$R_{i,j} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \tag{2.14}$$

In the case of the minimum-error state discrimination problem, we usually refer to the (optimal) reward as the *(optimal) success probability*, and denote it by $(\mathcal{P}_{\text{opt}}(\mathcal{E}))\ \mathcal{P}_M(\mathcal{E})$.
Another example of quantum guessing games is the state antidiscrimination problem, that has a reward matrix $R_{N\times N}$ of the following form:

$$R_{i,j} = \begin{cases} 0 & i = j \\ 1 & i \neq j. \end{cases} \tag{2.15}$$

2.2.1 Reduction to the Quantum State Discrimination Problem

As we have seen, quantum state discrimination is a special case of quantum guessing games. However, as noted by [CHT22; MMS24], quantum guessing games can be reduced to QSD problems over new ensembles constructed from the original ensembles and the reward matrices.

Let $G = (\mathcal{E}, R_{L\times N})$ be a quantum guessing game. Recall that the optimal reward of G is given by

$$\mathcal{R}_{\text{opt}} = \max_M \sum_{i=1}^{L}\sum_{j=1}^{N} R_{i,j}\,\mathrm{tr}(M_i\widetilde{\rho}_j), \tag{2.16}$$

where the maximization is taken over all POVMs $M = \{M_i\}_{i\in[L]}$. Define

$$\Delta_G = \sum_{i=1}^{L}\sum_{j=1}^{N} R_{i,j}\,\mathrm{tr}(\widetilde{\rho}_j), \tag{2.17}$$

and consider the ensemble $\mathcal{E}_G = \{\widetilde{\tau}_i\}_{i\in[L]}$, defined as

$$\widetilde{\tau}_i \stackrel{\text{def}}{=} \frac{1}{\Delta_G}\sum_{j=1}^{N} R_{i,j}\widetilde{\rho}_j. \tag{2.18}$$

Then, for any POVM $M = \{M_i\}_{i\in[L]}$, the success probability of M in a QSD problem over $\mathcal{E}_G$ is given by

$$\mathcal{P}_M(\mathcal{E}_G) = \sum_{i=1}^{L} \operatorname{tr}(M_i\widetilde{\tau}_i) = \frac{1}{\Delta_G}\sum_{i=1}^{L}\sum_{j=1}^{N} R_{i,j}\operatorname{tr}(M_i\widetilde{\rho}_j) = \frac{\mathcal{R}_M(G)}{\Delta_G}. \tag{2.19}$$

The above equality also shows that the optimal measurements for the QSD problem over $\mathcal{E}_G$ are the same as the optimal measurements for the quantum guessing game G, implying that

$$\mathcal{P}_{\text{opt}}(\mathcal{E}_G) = \frac{\mathcal{R}_{\text{opt}}(G)}{\Delta_G}. \tag{2.20}$$

2.3 Quantum State Discrimination and Antidiscrimination

In Example 2.2.3, we saw that the quantum state discrimination and antidiscrimination problems, two well-studied problems in quantum information theory with several applications in quantum foundation [PBR12], quantum communication [Che04; HB20] and quantum cryptography [Ben92; Ber07], are special cases of quantum guessing games. Moreover, in Subsection 2.2.1, we discussed how any quantum guessing game can be reduced to a QSD problem. This section briefly overviews these two problems and presents some essential related results.

2.3.1 Quantum State Discrimination

In the quantum state discrimination problem, Alice chooses a state from a known ensemble $\mathcal{E} = \{\widetilde{\rho}_k\}_{k\in[N]}$ and sends it to Bob, who performs a POVM $M = \{M_i\}_{i\in[N]}$ to determine which state Alice has sent. Ideally, Bob wants to find a POVM $M = \{M_i\}_{i\in[N]}$ such that

$$\operatorname{tr}(M_i\rho_j) = \delta_{i,j}, \quad \forall i,j \in [N], \tag{2.21}$$

where $\delta_{i,j}$ is the Kronecker delta function. However, this is not always possible due to the nature of quantum mechanics. The following proposition governs the possibility of perfect discrimination of quantum states.

Proposition 2.3.1. *Let $\mathcal{E} = \{\widetilde{\rho}_k\}_{k\in[N]}$ be a quantum ensemble. The states in $\mathcal{E}$ can be perfectly discriminated if and only if they are mutually orthogonal, i.e., $\rho_i\rho_j = 0$ for all $i \neq j$.*

Proof. Since the states in $\mathcal{E}$ are mutually orthogonal, for all $i \neq j$, we have $\rho_i\rho_j = \rho_i\rho_j = 0$, which implies that

$$\operatorname{Im}(\rho_i) \subseteq \ker(\rho_j) \quad \text{and} \quad \operatorname{Im}(\rho_j) \subseteq \ker(\rho_i). \tag{2.22}$$

Let $\operatorname{Im}(\mathcal{E}) = \cup_{i=1}^{N}\operatorname{Im}(\rho_i)$, and $M_0 = \mathbb{1} - \Pi_{\operatorname{Im}(\mathcal{E})}$ be the projector onto the orthogonal complement of $\operatorname{Im}(\mathcal{E})$. Let us define the POVM $M = \{M_i\}_{i\in[N]}$ as

$$M_i \stackrel{\text{def}}{=} \Pi_{\operatorname{Im}(\rho_i)} + \frac{1}{N}M_0. \tag{2.23}$$

Then, for all $i,j \in [N]$, we have

$$\operatorname{tr}(M_i\rho_j) = \operatorname{tr}(\Pi_{\operatorname{Im}(\rho_i)}\rho_j) + \frac{1}{N}\operatorname{tr}(M_0\rho_j) = \delta_{i,j} + 0 = \delta_{i,j}. \tag{2.24}$$

For the converse, let us first prove the statement for a pure state ensemble and then extend it to the general case. Suppose that $\mathcal{E} = \{\widetilde{|\psi_k\rangle}\}_{k\in[N]}$ is a pure state ensemble. The states in $\mathcal{E}$ can be perfectly discriminated if and only if they satisfy (2.21). Thus, for any two states $|\psi_i\rangle$ and $|\psi_j\rangle$, we have

$$\langle\psi_i|M_i|\psi_i\rangle = \langle\psi_j|M_j|\psi_j\rangle = 1, \quad \langle\psi_i|M_j|\psi_i\rangle = \langle\psi_j|M_i|\psi_j\rangle = 0. \tag{2.25}$$

Assume that $|\psi_i\rangle$ and $|\psi_j\rangle$ are not orthogonal, i.e., $|\psi_i\rangle = \alpha|\psi_j\rangle + \beta|\psi_j^\perp\rangle$ for some $\alpha, \beta \neq 0$, where $|\psi_j^\perp\rangle$ is a state orthogonal to $|\psi_j\rangle$.

Since $M_i \geq 0$, one can define $\sqrt{M_i} = \sum_{k=1}^N \sqrt{\lambda_k}|\phi_k\rangle\langle\phi_k|$, where $M_i = \sum_{k=1}^N \lambda_k|\phi_k\rangle\langle\phi_k|$ is the spectral decomposition of M_i. Then, from $\langle\psi_j|M_i|\psi_j\rangle = \langle\psi_j|\sqrt{M_i}\sqrt{M_i}|\psi_j\rangle = 0$, we imply that $\sqrt{M_i}|\psi_j\rangle = 0$. This implies that $\sqrt{M_i}|\psi_i\rangle = \sqrt{M_i}(\alpha|\psi_j\rangle + \beta|\psi_j^\perp\rangle) = \beta\sqrt{M_i}|\psi_j^\perp\rangle$. Thus, $\langle\psi_i|M_i|\psi_i\rangle = |\beta|^2 < 1$, which is a contradiction.

For the case of a mixed state ensemble $\mathcal{E} = \{\widetilde{\rho}_k\}_{k\in[N]}$, from the spectral decomposition theorem, for all $k \in [N]$, we can write $\widetilde{\rho}_k = \sum_{i=1}^d p_{k,i}|\psi_{k,i}\rangle\langle\psi_{k,i}|$, where $p_{k,i} \geq 0$, $\sum_{i=1}^d p_{k,i} = 1$, and $\{|\psi_{k,i}\rangle\}_{i\in[d]}$ is an orthonormal set. Now, the conditions $\mathrm{tr}(M_i\widetilde{\rho}_j) = \delta_{i,j}$ for all $i,j \in [N]$ are equivalent to having $\langle\psi_{j,k}|M_i|\psi_{j,k}\rangle = \delta_{i,j}$ for all $i,j \in [N]$ and $k \in [d]$. Therefore, using the result for pure state ensembles, we conclude that $\{|\psi_{j,k}\rangle\}_{k\in[d],j\in[N]}$ must be mutually orthogonal, which implies that $\{\rho_k\}_{k\in[N]}$ must be mutually orthogonal. □

Corollary 2.3.2. *In a QSD problem over an ensemble $\mathcal{E} = \{\widetilde{\rho}_k\}_{k\in[N]}$, if $\mathcal{E}$ contains a full-rank state, then it cannot be perfectly discriminated.*

When an ensemble can not be perfectly discriminated, different figures of merit can be used to quantify the performance of Bob's discrimination strategy. In this thesis, we focus on the minimum-error state discrimination, in which the goal is to minimize the probability of Bob making an error in identifying the state. The optimal success probability of Bob is then formalized as

$$\mathcal{P}_{\mathrm{opt}}(\mathcal{E}) = \max_M \sum_{k=1}^N \mathrm{tr}(M_k\widetilde{\rho}_k), \tag{2.26}$$

where the maximization is taken over all POVMs $M = \{M_k\}_{k\in[N]}$.

Corollary 2.3.2 shows that having full-rank states in the ensemble prevents perfect discrimination. However, in many situations, it is helpful to consider ensembles with only full-rank states. The following proposition shows that for any ensemble, we can always find a full-rank ensemble with an optimal success probability arbitrarily close to the optimal success probability of the original ensemble.

Proposition 2.3.3. *Let $\mathcal{E} = \{\widetilde{\rho}_k\}_{k\in[N]}$ be a quantum ensemble. For any $\epsilon > 0$, there exists a full-rank ensemble $\mathcal{E}' = \{\widetilde{\sigma}_k\}_{k\in[N]}$ such that*

$$|\mathcal{P}_{\mathrm{opt}}(\mathcal{E}) - \mathcal{P}_{\mathrm{opt}}(\mathcal{E}')| \leq \epsilon. \tag{2.27}$$

Proof. First, we prove that the set of full-rank states is dense in the set of all states. Let ρ be a state with a spectral decomposition $\rho = \sum_{i=1}^r p_i|\psi_i\rangle\langle\psi_i|$, where $p_i > 0$, and let $\{|\phi_j\rangle\}_{j=1}^{d-r}$ be an orthonormal basis for the subspace $\mathrm{Im}(\rho)^\perp$. For any $\epsilon > 0$, consider the operator $\sigma = \sum_{i=1}^r (p_i - \frac{\epsilon}{r})|\psi_i\rangle\langle\psi_i| + \sum_{j=1}^{d-r} \frac{\epsilon}{d-r}|\phi_j\rangle\langle\phi_j|$. It is clear that σ is a full-rank quantum state, and $\|\rho - \sigma\|_\infty < \epsilon$.

Now, let $\mathcal{E} = \{\widetilde{\rho}_k\}_{k\in[N]}$ be a quantum ensemble, and let $\epsilon > 0$. For each $k \in [N]$, let σ_k be a full-rank state such that $\|\rho_k - \sigma_k\|_1 < \frac{\epsilon}{N}$. We have proven the existence of such states in the case that the distance on matrices is induced by the operator norm, and since all norms are equivalent in finite dimensions, this result extends to the trace norm. Define the ensemble $\mathcal{E}' = \{\widetilde{\sigma}_k\}_{k\in[N]}$, where $\widetilde{\sigma}_k = \mathrm{tr}(\widetilde{\rho}_k)\sigma_k$. Then, for any POVM $M = \{M_i\}_{i\in[N]}$, we have

$$|\mathcal{P}_M(\mathcal{E}) - \mathcal{P}_M(\mathcal{E}')| = \left|\sum_{k=1}^N \mathrm{tr}(M_k\widetilde{\rho}_k) - \sum_{k=1}^N \mathrm{tr}(M_k\widetilde{\sigma}_k)\right| = \left|\sum_{k=1}^N \mathrm{tr}(M_k(\widetilde{\rho}_k - \widetilde{\sigma}_k))\right| \tag{2.28}$$

$$\leq \sum_{k=1}^N |\mathrm{tr}(M_k(\widetilde{\rho}_k - \widetilde{\sigma}_k))| \tag{2.29}$$

$$\leq \sum_{k=1}^N \|M_k\|_\infty \|\widetilde{\rho}_k - \widetilde{\sigma}_k\|_1 \tag{2.30}$$

$$< \epsilon, \tag{2.31}$$

where (2.30) follows from the Holder's inequality. Having $|\mathcal{P}_M(\mathcal{E}) - \mathcal{P}_M(\mathcal{E}')| < \epsilon$ for all POVMs M, we conclude that $|\mathcal{P}_{\text{opt}}(\mathcal{E}) - \mathcal{P}_{\text{opt}}(\mathcal{E}')| \le \epsilon$. □

Finding closed-form expressions for the optimal success probability of a QSD problem does not seem to be easy, and except for some special cases [Hel67; Hol72; And+02; EMV04; DT10], a closed-form solution is not known. One of these exceptional cases is the binary state discrimination problem, where the ensemble consists of only two states. The following proposition provides a closed-form solution for both the optimal success probability and the optimal POVM in that case.

Proposition 2.3.4 (Holevo-Helstrom Theorem [Hel67; Hol72])**.** *Let $\mathcal{E} = \{\widetilde{\rho}_1, \widetilde{\rho}_2\}$ be a binary ensemble. The optimal success probability of the QSD problem over $\mathcal{E}$ is given by*

$$\mathcal{P}_{\text{opt}} = \frac{1}{2} + \frac{1}{2}\|\widetilde{\rho}_1 - \widetilde{\rho}_2\|_1, \tag{2.32}$$

and the optimal POVM is given by $M = \{M_1, M_2\}$, where M_1 and M_2 are the projectors onto the eigenspaces of $\widetilde{\rho}_1 - \widetilde{\rho}_2$ corresponding to the positive and negative eigenvalues.

Proof. We have

$$\begin{aligned}
\mathcal{P}_M &= \operatorname{tr}(M_1\widetilde{\rho}_1) + \operatorname{tr}(M_2\widetilde{\rho}_2) && (2.33)\\
&= \frac{1}{2}\operatorname{tr}(M_1\widetilde{\rho}_1) + \frac{1}{2}\operatorname{tr}(M_1\widetilde{\rho}_1) + \frac{1}{2}\operatorname{tr}(M_2\widetilde{\rho}_2) + \frac{1}{2}\operatorname{tr}(M_2\widetilde{\rho}_2) && (2.34)\\
&= \frac{1}{2}\operatorname{tr}(M_1\widetilde{\rho}_1) + \frac{1}{2}\operatorname{tr}((\mathbb{1} - M_2)\widetilde{\rho}_1) + \frac{1}{2}\operatorname{tr}(M_2\widetilde{\rho}_2) + \frac{1}{2}\operatorname{tr}((\mathbb{1} - M_1)\widetilde{\rho}_2) && (2.35)\\
&= \frac{1}{2}\operatorname{tr}(\widetilde{\rho}_1) + \frac{1}{2}\operatorname{tr}(\widetilde{\rho}_2) - \frac{1}{2}\operatorname{tr}(M_2\widetilde{\rho}_1) - \frac{1}{2}\operatorname{tr}(M_1\widetilde{\rho}_2) + \frac{1}{2}\operatorname{tr}(M_1\widetilde{\rho}_1) + \frac{1}{2}\operatorname{tr}(M_2\widetilde{\rho}_2) && (2.36)\\
&= \frac{1}{2} + \frac{1}{2}\operatorname{tr}(M_2(\widetilde{\rho}_1 - \widetilde{\rho}_2)) - \frac{1}{2}\operatorname{tr}(M_1(\widetilde{\rho}_1 - \widetilde{\rho}_2)) && (2.37)\\
&= \frac{1}{2} + \frac{1}{2}\operatorname{tr}((M_2 - M_1)(\widetilde{\rho}_1 - \widetilde{\rho}_2)) && (2.38)\\
&\le \frac{1}{2} + \frac{1}{2}\|M_2 - M_1\|_\infty \|\widetilde{\rho}_1 - \widetilde{\rho}_2\|_1 && (2.39)\\
&\le \frac{1}{2} + \frac{1}{2}\|\widetilde{\rho}_1 - \widetilde{\rho}_2\|_1, && (2.40)
\end{aligned}$$

where the first inequality follows from the Holder's inequality, and the second inequality follows from the fact that the operator norm of $M_2 - M_1$ is equal to its largest singular value, which is at most 1.

To see that the equality can be achieved, consider the spectral decomposition of $\widetilde{\rho}_1 - \widetilde{\rho}_2$ as

$$\widetilde{\rho}_1 - \widetilde{\rho}_2 = \sum_{i=1}^{r} \lambda_i |\psi_i\rangle\langle\psi_i| + \sum_{j=1}^{d-r} \mu_j |\phi_j\rangle\langle\phi_j|, \tag{2.41}$$

where $\lambda_i \ge 0$ for $i \in [r]$ and $\mu_j < 0$ for $j \in [d-r]$. Define POVM $M = \{M_1, M_2\}$ as $M_1 = \sum_{i=1}^{r} |\psi_i\rangle\langle\psi_i|$ and $M_2 = \sum_{j=1}^{d-r} |\phi_j\rangle\langle\phi_j|$. Then, we have

$$\begin{aligned}
\operatorname{tr}((M_2 - M_1)(\widetilde{\rho}_1 - \widetilde{\rho}_2)) &= \sum_{i=1}^{r} \lambda_i + \sum_{j=1}^{d-r} |\mu_j| && (2.42)\\
&= \|\widetilde{\rho}_1 - \widetilde{\rho}_2\|_1, && (2.43)
\end{aligned}$$

which completes the proof. □

Remark 2.3.5. As a consequence of the above proposition, the optimal success probability of discriminating between two states, ρ_1 and ρ_2, with equal prior probabilities is $\frac{1}{2} + \frac{1}{4}\|\rho_1 - \rho_2\|_1$. This provides an operational interpretation of the trace distance between two quantum states, relating it to the optimal success probability of discriminating them.

Remark 2.3.6. Putting together Proposition 2.3.4 and our discussion in Subsection 2.2.1, we see that for any quantum guessing game $G = (\mathcal{E}, R_{2\times N})$, the optimal reward of G is given by

$$\mathcal{R}_{\text{opt}} = \frac{\Delta_G}{2} + \frac{1}{2}\|(\sum_{j=1}^{N} R_{1,j}\widetilde{\rho}_j - \sum_{j=1}^{N} R_{2,j}\widetilde{\rho}_j)\|_1, \tag{2.44}$$

and the optimal POVM is given by $M = \{M_1, M_2\}$, where M_1 and M_2 are the projectors onto the eigenspaces of $\sum_{j=1}^{N} R_{1,j}\widetilde{\rho}_j - \sum_{j=1}^{N} R_{2,j}\widetilde{\rho}_j$ corresponding to the positive and negative eigenvalues.

2.3.2 Quantum State Antidiscrimination

The quantum state antidiscrimination problem, or the quantum state exclusion [Ban+14], is a problem closely related to the quantum state discrimination problem but with a less ambitious goal. In the quantum state antidiscrimination, Bob's task is to identify which state has not been sent by Alice. More formally, Bob wants to find a POVM $M = \{M_i\}_{i\in[N]}$ such that

$$\operatorname{tr}(M_i\rho_i) = 0, \quad \forall i \in [N]. \tag{2.45}$$

If there exists a POVM M satisfying the above condition, then the states in the ensemble are said to be *antidistinguishable*, or they can be *perfectly antidiscriminated*. If the states in the ensemble can not be perfectly antidiscrimination, then Bob's goal is to minimize the probability of making an error, i.e., to find a POVM M that minimizes $\sum_{i=1}^{N} \operatorname{tr}(M_i\widetilde{\rho}_i)$.

The quantum state antidiscrimination problem was first introduced in [CFS02]. The first significant application of this problem appeared in [PBR12], where the authors used the problem to show that within the framework of quantum theory, quantum states must be considered ontic, i.e., a state of reality, rather than epistemic, i.e., a state of knowledge [Lei14]. Other applications of the quantum state antidiscrimination have been studied in [Bar+14; DWA14; HB20; LD20].

Proposition 2.3.7. *If a quantum ensemble $\mathcal{E} = \{\widetilde{\rho}_i\}_{i\in[N]}$ can be perfectly discriminated, then it can be perfectly antidiscriminated.*

Proof. Let $M = \{M_i\}_{i\in[N]}$ be a POVM that perfectly discriminates the states in $\mathcal{E}$. Then, for all $i \in [N]$, we have $\operatorname{tr}(M_i\widetilde{\rho}_i) = \delta_{i,j}$. Define $M' = \{\frac{1}{N-1}(\mathbb{1} - M_i)\}_{i\in[N]}$. Then, for all $i \in [N]$, we have

$$\operatorname{tr}(M_i'\rho_i) = \frac{1}{N-1}\operatorname{tr}((\mathbb{1} - M_i)\rho_i) = \frac{1}{N-1}(1 - \operatorname{tr}(M_i\rho_i)) = 0. \tag{2.46}$$

Thus, the states in $\mathcal{E}$ can be perfectly antidiscriminated. □

Proposition 2.3.8. *Let $\mathcal{E} = \{\widetilde{\rho}_1, \widetilde{\rho}_2\}$ be a binary ensemble. $\mathcal{E}$ can be perfectly antidiscriminated if and only it can be perfectly discriminated.*

Proof. One direction of the proof is clear from the previous proposition. For the converse, suppose that $\mathcal{E}$ can be perfectly antidiscriminated. Then, there exists a POVM $M = \{M_1, M_2\}$ such that $\operatorname{tr}(M_i\rho_i) = 0$ for $i = 1, 2$. This implies that $\operatorname{tr}(M_2\rho_1) = \operatorname{tr}(M_1\rho_2) = 1$. Thus, POVM $M' = \{M_2, M_1\}$ perfectly discriminates the states in $\mathcal{E}$. □

Proposition 2.3.9. *If an ensemble $\mathcal{E} = \{\widetilde{\rho}_i\}_{i\in[N]}$ contains at least two orthogonal states, then it can be perfectly antidiscriminated.*

Proof. Without the loss of generality, suppose that ρ_1 and ρ_2 are orthogonal states. Let $M_0 \stackrel{\text{def}}{=} \mathbb{1} - \Pi_{\text{Im}(\rho_1)\cup\text{Im}(\rho_2)}$ be the projector onto the orthogonal complement of the subspace spanned by ρ_1 and ρ_2. Define the POVM $M = \{M_i\}_{i\in[N]}$ as $M_1 = \Pi_{\text{Im}(\rho_2)} + \frac{1}{2}M_0$, $M_2 = \Pi_{\text{Im}(\rho_1)} + \frac{1}{2}M_0$, and $M_i = 0$ for $i > 2$. Then, it can be seen that for all $i \in [N]$, we have $\operatorname{tr}(M_i\rho_i) = 0$. □

Corollary 2.3.10. *For $N > 2$, there exist ensembles with N states that can be perfectly antidiscriminated, but not perfectly discriminated.*

Proof. It is enough to consider an ensemble containing two orthogonal states and two non-orthogonal states. Then, from Proposition 2.3.1, the ensemble can not be perfectly discriminated, but it is implied from Proposition 2.3.9 that the ensemble can be perfectly antidiscriminated. □

Remark 2.3.11. The conditions under which a quantum ensemble can be perfectly antidiscriminated are more complicated than the conditions for perfect discrimination, and has been the subject of study in [HK18; JRS23], where necessary and sufficient conditions for perfect antidiscrimination of pure states have been derived.

2.4 Applications of Quantum Guessing Games

So far, we have seen that quantum guessing games are a general framework that encompasses two well-studied problems in quantum information theory, the quantum state discrimination and antidiscrimination problems. A compelling question, however, is whether quantum guessing games, in their most general form, find applications in quantum information theory. In this section, we present two scenarios where quantum guessing games naturally arise, offering further motivation for their study.

2.4.1 Quantum Random Access Codes

Quantum random access codes are useful tools that have found several applications in quantum information science, including quantum cryptography [PB11] and quantum information theory [Paw+09]. Here, we use the formalism introduced in [PAGK23] to show how quantum guessing games can be used to study quantum random access codes.

Consider a communication scenario between Alice and Bob, in which Alice is given an $n-$character string $\mathbf{x} \in [m]^n$, and she encodes this string in a qudit state via an encoding function $\mathbf{x} \mapsto \rho_{\mathbf{x}} \in \mathcal{L}(\mathbb{C}^d)$, and sends it to Bob. Bob, on the other hand, is given $y \in [n]$ as input and is asked to perform a decoding measurement $M_y = \{M_{y,b}\}_{b\in[m]}$ on the received state to determine the $y-$th character of the string. The goal is to find a pair of encoding preparation and decoding measurement that maximizes the average success probability of Bob, given by

$$\mathcal{F} \stackrel{\text{def}}{=} \sum_{\mathbf{x},\mathbf{y}} \alpha_{\mathbf{x},y} \operatorname{tr}(M_{y,\mathbf{x}_y} \rho_{\mathbf{x}}), \tag{2.47}$$

where $\{\alpha_{\mathbf{x},y}\}_{\mathbf{x}\in[m]^n, y\in[n]}$ is a joint probability distribution over the inputs $\mathbf{x}$ and y, and $\mathbf{x}_y$ is the $y-$th character of the string $\mathbf{x}$. Those pairs of encoding preparations and decoding measurements that maximize the above functional are then called the *optimal quantum encoding-decoding strategies* for the scenario described above, which is called an $n^m \overset{d}{\mapsto} 1$ random access code (RAC).

A well-known algorithm for finding the optimal quantum encoding-decoding strategies for a given RAC is the *see-saw algorithm*, in which one repeatedly optimizes over the encoding and decoding strategies until convergence to a (possibly local) maximum of the success probability is achieved. A more detailed description of the see-saw algorithm is given in Algorithm 1.

Note that in Step (2.49) of this algorithm, the algorithm is actually finding the optimal measurement of a quantum guessing game. To see this, let us define a quantum guessing game $G_y^{(k)} = (\mathcal{E}^{(k)}, R_y)$ for each $y \in [n]$ and $k \in [T]$, where $\mathcal{E}^{(k)} = \{\frac{1}{m^n}\rho_{\mathbf{x}}^{(k)}\}_{\mathbf{x}\in[m]^n}$, and R_y is an $n \times m^n$ matrix, where the $(b,\mathbf{x})$-th entry is given by $\alpha_{\mathbf{x},y}\delta_{b,\mathbf{x}_y}$. Then, we have $\mathcal{R}_{\text{opt}}(G_y^{(k)}) = \frac{1}{m^n} \max_{M_y} \sum_{\mathbf{x},\mathbf{b}} \alpha_{\mathbf{x},y}\delta_{b,\mathbf{x}_y} \operatorname{tr}(M_{y,b}\rho_{\mathbf{x}}^{(k)})$. This provides an example of how quantum guessing games arise in other problems in quantum information theory.

2.4.2 Minimum Guesswork Discrimination

In [Che+15], the authors consider a different figure of merit for the quantum state discrimination problem, called the *guesswork*, and define a new problem, called *minimum guesswork discrimination*. In minimum guesswork discrimination, Alice picks a secret message x_i from her alphabet $\mathcal{X}$, and sends a quantum state ρ_{x_i} to Bob. In minimum-error discrimination, Bob's goal is to identify the message sent by Alice in one-shot, by asking Alice "is the message x_j?", based on his measurement

Algorithm 1 A See-Saw Algorithm for Quantum Random Access Codes

Require: $\mathcal{F} \stackrel{\text{def}}{=} \sum_{\mathbf{x},\mathbf{y}} \alpha_{\mathbf{x},y} \operatorname{tr}(M_{y,\mathbf{x}_y} \rho_{\mathbf{x}})$, to be maximized over all encoding preparations $\mathbf{x} \mapsto \rho_{\mathbf{x}}$ and decoding measurements $\{M_y\}_{y\in[n]}$, where $M_y = \{M_{y,b}\}_{b\in[m]}$.

Initialize: A set of randomly chosen measurements $\{M_y^{(0)}\}_{y\in[n]}$ where $M_y^{(0)} = \{M_{y,b}^{(0)}\}_{b\in[m]}$.

for $k = 1$ **to** $T-1$ **do**

Maximization over the encoding preparations:

$$\rho_{\mathbf{x}}^{(k)} \leftarrow \arg\max_{\rho_{\mathbf{x}}} \operatorname{tr}\left(\rho_{\mathbf{x}} \sum_{y,b} \alpha_{\mathbf{x},y} \delta_{b,\mathbf{x}_y} M_{y,b}^{(k-1)}\right), \qquad \forall \mathbf{x} \in [m]^n, \tag{2.48}$$

The above maximizations are done over the set of all $d \times d$ density matrices.

Maximization over the decoding measurements:

$$M_y^{(k)} \leftarrow \arg\max_{M_y} \sum_{\mathbf{x},\mathbf{b}} \alpha_{\mathbf{x},y} \delta_{b,\mathbf{x}_y} \operatorname{tr}(M_{y,b} \rho_{\mathbf{x}}^{(k)}), \qquad \forall y \in [n], \tag{2.49}$$

The above maximizations are done over the set of all POVMs.

end for

Return: The quantum encoding-decoding strategies $\{\rho_{\mathbf{x}}^{(T)}\}_{\mathbf{x}\in[m]^n}$ and $\{M_y^{(T)}\}_{y\in[n]}$.

outcomes. In minimum guesswork discrimination, Bob can query Alice multiple times, until he hits the correct message. The goal is to minimize the expected number of queries Bob needs to make to identify the message sent by Alice. It is notable that in quantum strategies, Bob can perform multiple measurements on the quantum state, however, as was shown in [Han+21], the set of random variables corresponding to the number of queries when Bob performs multiple measurements is equal to the case where Bob performs a single measurement.

To state the problem formally, following the formalization given in [Che+15], let us start with the classical setting. Consider a random variable X with alphabet $\mathcal{X} = \{x_i\}_{i\in[n]}$ and associated probability distribution $\{p(x_i)\}_{i\in[n]}$. Bob, who knows both the alphabet and the distribution, aims to identify the true value of X by sequentially asking questions of the form "is $X = x_i$?" until he hits the correct value. The expected number of guesses Bob needs to make is known as the guesswork. In [Mas94], it was shown that Bob's optimal strategy is to arrange his queries in non-increasing order of probabilities $p(x_i)$. The optimal guesswork of Bob is then given by

$$\mathcal{G}(X) \stackrel{\text{def}}{=} \sum_{i=1}^{n} \sigma(i) p(x_i), \tag{2.50}$$

where $\sigma \in \mathcal{S}_n$ such that $p(x_i) \geq p(x_j)$ implies $\sigma(i) \leq \sigma(j)$.

This can be extended to the conditional case, where it is assumed that there exists a channel between Alice and Bob with input set $\mathcal{X}$ and output set $\mathcal{Y} = \{y_j\}_{j\in[m]}$. The channel is then specified by the conditional distribution $\{p(y_j \mid x_i)\}_{i\in[n],j\in[m]}$. The distribution of the random variable Y is given by

$$p(y_j) = \sum_{i=1}^{n} p(x_i) p(y_j \mid x_i), \tag{2.51}$$

for $j \in [m]$. In the conditional case, when Bob detects y_j, he applies the optimal strategy discussed above on the posterior distribution $\{p(x_i \mid y_j)\}_{i\in[n]}$ to identify the true value of X. The guesswork is then given by

$$\mathcal{G}(X \mid Y) \stackrel{\text{def}}{=} \sum_{j=1}^{m} p(y_j) \sum_{i=1}^{n} \sigma_j(i) p(x_i \mid y_j), \tag{2.52}$$

where $\sigma_j \in \mathcal{S}_n$ is a permutation such that $p(x_i \mid y_j) \geq p(x_k \mid y_j)$ implies $\sigma_j(i) \leq \sigma_j(k)$.

Now, it is easy to define the guesswork in the quantum setting. Let X be Alice's random variable, which gives rise to an ensemble $\mathcal{E} = \{\widetilde{\rho}_{x_i}\}_{i\in[n]}$. Bob performs a POVM $M = \{M_{y_j}\}_{j\in[m]}$ on the quantum state received from Alice, giving rise to a conditional distribution $\{p(y_j \mid x_i)\}_{i\in[n],j\in[m]}$, where $p(y_j \mid x_i) = \mathrm{tr}(M_{y_j}\rho_{x_i})$. The random variable Y_M is then determined according to (2.51). The minimum guesswork of the ensemble $\mathcal{E}$ is defined as

$$\mathcal{G}_{\mathcal{E}} \stackrel{\text{def}}{=} \min_{M} \mathcal{G}(X \mid Y_M), \tag{2.53}$$

where the minimization is taken over all POVMs M.

To see how the minimum guesswork of an ensemble $\mathcal{E}$ can be related to the quantum guessing games, let us rewrite $\mathcal{G}(X \mid Y_M)$ as

$$\mathcal{G}(X \mid Y_M) = \sum_{j=1}^{m} p(y_j) \sum_{i=1}^{n} \sigma_j(i) p(x_i \mid y_j) \tag{2.54}$$

$$= \sum_{j=1}^{m} \sum_{i=1}^{n} \sigma_j(i) p(x_i, y_j) \tag{2.55}$$

$$= \sum_{j=1}^{m} \sum_{i=1}^{n} \sigma_j(i) p(x_i) p(y_j \mid x_i) \tag{2.56}$$

$$= \sum_{j=1}^{m} \sum_{i=1}^{n} \sigma_j(i) \,\mathrm{tr}(M_{y_j}\widetilde{\rho}_{x_i}). \tag{2.57}$$

(2.57) shows that the optimal measurements for the minimum guesswork discrimination of an ensemble $\mathcal{E}$ are equal to the optimal measurements for the quantum guessing game $G = (\mathcal{E}, R_{m\times n})$, where $R_{i,j} \stackrel{\text{def}}{=} 1 - \sigma_i(j)$. However, it has to be noted that the entries of the reward matrix R are not constant; they depend on Bob's measurement. This motivates extending the definition of quantum guessing games to more general forms of reward matrices, which could be a possible direction for future research.

Chapter 3

Explicit Measurements

In the previous chapter, we addressed the challenges in finding a closed-form solution for the optimal reward and measurement in a quantum guessing game. In Chapter 4, we will discuss numerical methods developed to find these solutions with arbitrary precision. While these methods work well for small-sized instances, they become impractical as the problem size increases. This limitation motivates the investigation of explicit measurements that can be computed efficiently without the need for solving an optimization problem, yet still perform well for the task.

The importance of such explicit measurements is underscored by the fact that in some quantum guessing game scenarios, calculating the exact optimal reward isn't necessary. Often, a lower or upper bound on the optimal reward is sufficient. As we will see, some explicit measurements can provide such bounds.

In this chapter, we first generalize the pretty good measurement, an explicit measurement extensively studied in the context of quantum state discrimination [PW91; Hau+96; EF01; Mon07], to the broader setting of quantum guessing games, and obtain two lower bounds on its expected reward. We then review the Belavkin measurements, a larger family of explicit measurements that includes the pretty good measurement as a special case [BM04]. Our discussion on Belavkin measurements will be focused on the quantum state discrimination problem and will set the stage for our later discussions on iterative algorithms in Chapter 5.

3.1 Pretty Good Measurement

The pretty good measurement (PGM), also known as the square-root measurement, is probably the most well-known suboptimal measurement. It was first introduced by Peres and Wootters in [PW91] for a special instance of QSD problem, and later by Hausladen et al. for a more general setting of QSDs in [Hau+96]. In this section, we will generalize these measurements to the more general setting of quantum guessing games. For the following discussion, unless otherwise stated, we will assume that our ensembles are full-rank, and the reward matrices have no zero rows.

Definition 3.1.1. Let $G = (\mathcal{E}, R_{L\times N})$ be a quantum guessing game. The pretty good measurement $M = \{M_k\}_{k\in[L]}$ is a POVM defined as

$$M_i \stackrel{\text{def}}{=} \Sigma^{-\frac{1}{2}} \left(\sum_{j=1}^{N} R_{i,j} \widetilde{\rho}_j \right) \Sigma^{-\frac{1}{2}}, \tag{3.1}$$

for $i \in [L]$, where

$$\Sigma \stackrel{\text{def}}{=} \sum_{i=1}^{L} \sum_{j=1}^{N} R_{i,j} \widetilde{\rho}_j. \tag{3.2}$$

Notation 3.1.2. We denote the expected reward of the pretty good measurement by $\mathcal{R}_{\text{PGM}}(G)$ (or simply $\mathcal{R}_{\text{PGM}}$ when the game is clear from the context).

In Definition 3.1.1, if we consider the reward matrix R to be the reward matrix corresponding to the QSD problem, the pretty good measurement will be of the form $M = \{M_i\}_{i\in[L]}$ where

$$M_i = \Sigma^{-\frac{1}{2}} \widetilde{\rho}_i \Sigma^{-\frac{1}{2}}, \quad \Sigma = \sum_{i=1}^{L} \widetilde{\rho}_i, \tag{3.3}$$

and the corresponding success probability is denoted by $\mathcal{P}_{\text{PGM}}(\mathcal{E})$. Recall from Section **??** that any quantum guessing game $G = (\mathcal{E}, R_{L\times N})$ can be reduced to a QSD problem over the ensemble $\mathcal{E}_G = \{\widetilde{\tau}_i\}_{i\in[L]}$. The pretty good measurement associated to this QSD problem is the POVM $M = \{M_i\}_{i\in[L]}$ where

$$M_i = \left(\sum_{i=1}^{L} \widetilde{\tau}_i\right)^{-\frac{1}{2}} \widetilde{\tau}_i \left(\sum_{i=1}^{L} \widetilde{\tau}_i\right)^{-\frac{1}{2}} \tag{3.4}$$

$$= \left(\sum_{i=1}^{L}\sum_{j=1}^{N} R_{i,j}\widetilde{\rho}_j\right)^{-\frac{1}{2}} \left(\sum_{j=1}^{N} R_{i,j}\widetilde{\rho}_j\right) \left(\sum_{i=1}^{L}\sum_{j=1}^{N} R_{i,j}\widetilde{\rho}_j\right)^{-\frac{1}{2}}, \tag{3.5}$$

which coincides with the measurement defined in Definition 3.1.1. Moreover, from (2.19), we have

$$\mathcal{P}_{\text{PGM}}(\mathcal{E}_G)\Delta_G = \mathcal{R}_{\text{PGM}}(G), \tag{3.6}$$

where $\Delta_G \stackrel{\text{def}}{=} \sum_{i=1}^{L}\sum_{j=1}^{N} R_{i,j}\,\text{tr}(\widetilde{\rho}_j)$.

Remark 3.1.3. In [MMS24], the authors define a measurement called the "*pretty bad measurement*," which performs pretty well for the task when Bob is trying to minimize his success probability in a QSD problem (or equivalently, the task of quantum state antidiscrimination). Let $G = \{G_i\}_{i\in[N]}$ be a pretty good measurement for a QSD problem. The pretty bad measurement $B = \{B_i\}_{i\in[N]}$ is defined as

$$B_i \stackrel{\text{def}}{=} \frac{1}{N-1}(\mathbb{1} - G_i). \tag{3.7}$$

To see that this measurement coincides with the pretty good measurement defined for the quantum state antidiscrimination problem, note that for the pretty good measurement $M = \{M_i\}_{i\in[N]}$ for the state antidiscrimination, we have

$$M_i = \left(\sum_{i=1}^{N}\sum_{j\neq i} \widetilde{\rho}_j\right)^{-\frac{1}{2}} \left(\sum_{j\neq i} \widetilde{\rho}_j\right) \left(\sum_{i=1}^{N}\sum_{j\neq i} \widetilde{\rho}_j\right)^{-\frac{1}{2}} \tag{3.8}$$

$$= \left(\frac{1}{N-1}\sum_{i=1}^{N} \widetilde{\rho}_i\right)^{-\frac{1}{2}} \left(\sum_{j\neq i} \widetilde{\rho}_j\right) \left(\frac{1}{N-1}\sum_{i=1}^{N} \widetilde{\rho}_i\right)^{-\frac{1}{2}} \tag{3.9}$$

$$= \frac{1}{N-1}\left(\sum_{i=1}^{N} \widetilde{\rho}_i\right)^{-\frac{1}{2}} \left(\sum_{j\neq i} \widetilde{\rho}_j\right) \left(\sum_{i=1}^{N} \widetilde{\rho}_i\right)^{-\frac{1}{2}} \tag{3.10}$$

$$= \frac{1}{N-1}\left(\sum_{i=1}^{N} \widetilde{\rho}_i\right)^{-\frac{1}{2}} \left(\sum_{j=1}^{N} \widetilde{\rho}_j - \widetilde{\rho}_i\right) \left(\sum_{i=1}^{N} \widetilde{\rho}_i\right)^{-\frac{1}{2}} \tag{3.11}$$

$$= \frac{1}{N-1}(\mathbb{1} - G_i). \tag{3.12}$$

This reflects what Hamlet once said to Rosencrantz: "*There is nothing either good or bad, but thinking makes it so.*"

3.1.1 Lower Bounds on the PGM's Reward

The pretty good measurement is called pretty good, as its expected reward provides a decent lower bound on the optimal reward of a quantum guessing game. The following proposition shows that the pretty good measurement is at least as good as the uniform guessing strategy.

Proposition 3.1.4. *For any quantum guessing game $G = (\mathcal{E}, R_{L\times N})$,*

$$\mathcal{R}_{\text{PGM}} \geq \frac{\Delta_G}{L}, \tag{3.13}$$

where $\Delta_G \stackrel{\text{def}}{=} \sum_{i=1}^{L}\sum_{j=1}^{N} R_{i,j}\,\mathrm{tr}(\widetilde{\rho}_j)$*.*

Proof. To prove the above bound, we follow the same steps as in the proof of [MMS24, Theorem 1]. Define $r = \{r_i\}_{i\in[L]}$ as $r_i \stackrel{\text{def}}{=} \|\Sigma^{\frac{-1}{4}}(\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j)\Sigma^{\frac{-1}{4}}\|_F$, where $\Sigma = \sum_{i=1}^{L}\sum_{j=1}^{N} R_{i,j}\widetilde{\rho}_j$, and $\|\cdot\|_F$ denotes the Frobenius norm (2-Schatten norm). Suppose that $\oplus$ denotes the addition modulo L and $M = \{M_i\}_{i\in[L]}$ is the pretty good measurement. We have the following:

$$\Delta_G = \sum_{s=0}^{L-1}\left(\sum_{i=1}^{L}\sum_{j=1}^{N} R_{ij}\,\mathrm{tr}\left(\widetilde{\rho}_j M_{i\oplus s}\right)\right) \tag{3.14}$$

$$= \sum_{s=0}^{L-1}\sum_{i=1}^{L}\mathrm{tr}\left(\Sigma^{\frac{-1}{4}}\left(\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j\right)\Sigma^{\frac{-1}{4}}\Sigma^{\frac{-1}{4}}\left(\sum_{j=1}^{N} R_{(i\oplus s)j}\widetilde{\rho}_j\right)\Sigma^{\frac{-1}{4}}\right) \tag{3.15}$$

$$\leq \sum_{s=0}^{L-1}\sum_{i=1}^{L} r_i r_{i\oplus s} = \|r\|_1^2 \leq L\|r\|_2^2, \tag{3.16}$$

where the inequalities in the last line are obtained using the Cauchy–Schwarz inequality. We also have:

$$\mathcal{R}_{\mathrm{PGM}} = \sum_{i=1}^{L}\sum_{j=1}^{N} R_{ij}\,\mathrm{tr}(\widetilde{\rho}_j M_i) \tag{3.17}$$

$$= \sum_{i=1}^{L}\mathrm{tr}\left(\Sigma^{\frac{-1}{4}}\left(\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j\right)\Sigma^{\frac{-1}{4}}\Sigma^{\frac{-1}{4}}\left(\sum_{k=1}^{N} R_{ik}\widetilde{\rho}_k\right)\Sigma^{\frac{-1}{4}}\right) \tag{3.18}$$

$$= \sum_{i=1}^{L} r_i^2 = \|r\|_2^2 \tag{3.19}$$

Thus,

$$\mathcal{R}_{\mathrm{PGM}} \geq \frac{\Delta_G}{L}. \tag{3.20}$$

□

Remark 3.1.5. Although the above proposition for the special case of the QSD problem is one of the main results in [MMS24], it is notable that this lower bound for the pure state ensembles can be obtained more easily using the previously established lower bounds for PGM proved in [Mon07]. In this paper, the author proves that for a QSD problem over an ensemble $\mathcal{E} = \{\widetilde{\rho}_i\}_{i\in[N]}$, where ρ_i's are pure quantum states, i.e. $\rho_i = |\psi_i\rangle\langle\psi_i|$, the expected success probability of the pretty good measurement admits the following lower bound [Mon07, Section 2.2]:

$$\mathcal{P}_{\mathrm{PGM}} \geq \sum_{i=1}^{N}\frac{q_i^2}{\sum_{j=1}^{N} q_j\left|\left\langle\psi_i \mid \psi_j\right\rangle\right|^2}. \tag{3.21}$$

Note that the right-hand side of (3.21) is also greater than or equal to $\sum_{i=1}^{N} q_i^2$, which is by the AM-QM inequality greater than or equal to $\frac{1}{N}$.

What highlights the importance of the pretty good measurement is that it is not only a good lower bound for the optimal reward, but it also provides an upper bound on the optimal reward of a quantum guessing game. This was shown by Barnum and Knill in [BK02] for the special case of the QSD problem, and it can directly be generalized to the quantum guessing games as follows.

Proposition 3.1.6. *For any quantum guessing game* $G = (\mathcal{E}, R_{L\times N})$*,*

$$\mathcal{R}_{\mathrm{opt}} \leq \sqrt{\mathcal{R}_{\mathrm{PGM}}}\sqrt{\Delta_G}, \tag{3.22}$$

where $\Delta_G \stackrel{\text{def}}{=} \sum_{i=1}^{L}\sum_{j=1}^{N} R_{i,j}\,\mathrm{tr}(\widetilde{\rho}_j)$*.*

Proof. The proof we present here follows the same steps as in the proof of [Wat18, Theorem 3.10]. Defining $r = \{r_i\}_{i\in[L]}$ as in the proof of Proposition 3.1.4, for any measurement $M = \{M_k\}_{k\in[L]}$ we have

$$\mathcal{R}_M = \sum_{i=1}^{L} \operatorname{tr}\left(\sum_{j=1}^{N} R_{i,j}\widetilde{\rho}_j M_i\right) \tag{3.23}$$

$$= \sum_{i=1}^{L} \operatorname{tr}\left(\Sigma^{\frac{-1}{4}}\left(\sum_{j=1}^{N} R_{i,j}\widetilde{\rho}_j\right)\Sigma^{\frac{-1}{4}}\Sigma^{\frac{1}{4}}M_i\Sigma^{\frac{1}{4}}\right) \tag{3.24}$$

$$\leq \sum_{i=1}^{L} r_i \left\|\Sigma^{\frac{1}{4}}M_i\Sigma^{\frac{1}{4}}\right\|_F \tag{3.25}$$

$$\leq \sqrt{\|r\|_2^2 \sum_{i=1}^{L} \left\|\Sigma^{\frac{1}{4}}M_i\Sigma^{\frac{1}{4}}\right\|_F^2} \tag{3.26}$$

$$= \sqrt{\|r\|_2^2 \sum_{i=1}^{L} \operatorname{tr}(M_i\Sigma^{\frac{1}{2}}M_i\Sigma^{\frac{1}{2}})}, \tag{3.27}$$

where (3.25) and (3.26) are Cauchy–Schwarz inequalities. Since

$$\operatorname{tr}(M_i\Sigma^{\frac{1}{2}}M_i\Sigma^{\frac{1}{2}}) \leq \operatorname{tr}(\Sigma^{\frac{1}{2}}M_i\Sigma^{\frac{1}{2}}) = \operatorname{tr}(M_i\Sigma), \tag{3.28}$$

we have

$$\mathcal{R}_M \leq \sqrt{\|r\|_2^2 \sum_{i=1}^{L} \operatorname{tr}(M_i\Sigma^{\frac{1}{2}}M_i\Sigma^{\frac{1}{2}})} \tag{3.29}$$

$$\leq \sqrt{\|r\|_2^2 \sum_{i=1}^{L} \operatorname{tr}(M_i\Sigma)} \tag{3.30}$$

$$= \sqrt{\|r\|_2^2 \operatorname{tr}(\Sigma)}. \tag{3.31}$$

We saw earlier that $\|r\|_2^2 = \mathcal{R}^{\text{PGM}}$, and $\operatorname{tr}(\Sigma) = \Delta_G$. Thus,

$$\mathcal{R}_{\text{opt}} \leq \sqrt{\mathcal{R}_{\text{PGM}}}\sqrt{\Delta_G}. \tag{3.32}$$

□

Remark 3.1.7. Since a quantum guessing game's expected reward is always between 0 and 1, the above bound is meaningful only if $\sqrt{\mathcal{R}_{\text{PGM}}}\sqrt{\Delta_G} \leq 1$.

Remark 3.1.8. The lower and upper bounds derived in Propositions 3.1.4 and 3.1.6 can both be obtained by applying the same bounds for the pretty good measurement in the QSD problem to (3.6).

3.2 Belavkin Measurements

In this section, we limit our discussion to the QSD problem. The Belavkin measurements are a family of (possibly) suboptimal measurements introduced in [Bel75] for this problem. In what follows, assume that $\mathcal{E}$ is an ensemble of quantum states $\{\rho_i\}_{i\in[N]}$ with associated probabilities $\{q_i\}_{i\in[N]}$, and for all $i \in [N]$, $\rho_i = \psi_i\psi_i^\dagger$, where ψ_i is a $d \times \operatorname{rank}(\rho_i)$ matrix.

Definition 3.2.1. For an ensemble $\mathcal{E}$, the Belavkin measurement with associated weights $w = \{w_i \in \mathrm{Pos}(\mathbb{C}^{\mathrm{rank}\,\rho_i})\}_{i\in[N]}$ is defined as

$$M_i \stackrel{\text{def}}{=} (\Sigma^{(w)})^{-\frac{1}{2}} \psi_i w_i {\psi_i}^{\dagger} (\Sigma^{(w)})^{-\frac{1}{2}}, \tag{3.33}$$

for $i \in [N]$, where

$$\Sigma^{(w)} \stackrel{\text{def}}{=} \sum_{i=1}^{N} \psi_i w_i {\psi_i}^{\dagger}. \tag{3.34}$$

Remark 3.2.2. If ρ_i is a pure state, w_i has to be a positive real number.

Notation 3.2.3. For $i \in [N]$, we denote the operator $\psi_i^{\dagger} \Sigma^{(w)^{-\frac{1}{2}}} \psi_i$ by Y_i.

Remark 3.2.4. Let $M = \{M_i\}_{i\in[N]}$ be the Belavkin measurement associated with weights $w = \{w_i\}_{i\in[N]}$. We have

$$\begin{aligned}
\mathrm{tr}(\rho_i M_i) &= \mathrm{tr}(\rho_i (\Sigma^{(w)})^{-\frac{1}{2}} \psi_i w_i {\psi_i}^{\dagger} (\Sigma^{(w)})^{-\frac{1}{2}}) && (3.35)\\
&= \mathrm{tr}(\psi_i {\psi_i}^{\dagger} (\Sigma^{(w)})^{-\frac{1}{2}} \rho_i (\Sigma^{(w)})^{-\frac{1}{2}} w_i) && (3.36)\\
&= \mathrm{tr}({\psi_i}^{\dagger} (\Sigma^{(w)})^{-\frac{1}{2}} \psi_i w_i {\psi_i}^{\dagger} (\Sigma^{(w)})^{-\frac{1}{2}} \psi_i) && (3.37)\\
&= \mathrm{tr}(Y_i w_i Y_i). && (3.38)
\end{aligned}$$

Example 3.2.5. Let $M = \{M_i\}_{i\in[N]}$ be the Belavkin measurement associated with weights $w = \{w_i\}_{i\in[N]}$, where $w_i = q_i$ for all $i \in [N]$. This measurement coincides with the pretty good measurement.

Definition 3.2.6. For an ensemble $\mathcal{E}$, the Belavkin measurement with associated weights $w = \{w_i\}_{i\in[N]}$ where $w_i = q_i \widetilde{\rho_i}$ is called the *Holevo measurement*.

3.2.1 Optimality Condition for Belevkin Measurements

This short section reviews the condition under which a Belavkin measurement is optimal for a QSD problem. We restate the result from [BM04] without proof.

Proposition 3.2.7 (Optimality Condition for Belevkin Measurements [BM04])**.** *A POVM $M = \{M_i\}_{i\in[N]}$ is optimal for a quantum state discrimination problem over an ensemble $\mathcal{E} = \{(\rho_i, q_i)\}_{i\in[N]}$, where $\rho_i = \psi_i \psi_i^{\dagger}$, if and only if it is identical to a Belavkin measurement $M^{(w)} = \{M_i^{(w)}\}_{i\in[N]}$ associated with weights $w = \{w_i\}_{i\in[N]}$ such that there exists $\alpha > 0$ satisfying*

$$q_i Y_i^{(w)} \leq \alpha \mathbb{1}_{\mathrm{rank}(\rho_i)}, \tag{3.39}$$

$$q_i Y_i^{(w)} w_i = \alpha w_i, \tag{3.40}$$

for all $i \in [N]$, where

$$Y_i^{(w)} \stackrel{\text{def}}{=} \psi_i^{\dagger} \left(\Sigma^{(w)}\right)^{-1/2} \psi_i, \quad \textit{and} \quad \Sigma^{(w)} \stackrel{\text{def}}{=} \sum_{i=1}^{N} \psi_i w_i \psi_i^{\dagger}. \tag{3.41}$$

Remark 3.2.8. Proposition 3.2.7 highlights the importance of the Belavkin measurements, as the optimal measurements for a QSD problem are always Belavkin measurements.

Chapter 4

Semidefinite Programming

Semidefinite programming (SDP) is a subfield of convex optimization that extends linear programming to optimize linear objective functions over the intersection of the cone of positive semidefinite matrices and an affine space. This optimization framework has proven to be a powerful tool in various fields, including quantum information theory, control theory, and combinatorial optimization. Its effectiveness is due to two main factors:

- Semidefinite programs can be efficiently solved using interior point methods. These methods were initially developed by Karmarkar [Kar84] for linear programs and later extended to semidefinite programming by Nesterov and Nemirovskii [NN92; NN94], and Alizadeh [Ali91].
- Many problems, particularly in quantum information theory, can be directly formulated as semidefinite programs or approximated by SDPs. This allows SDP solvers to find exact or approximate solutions for these problems.

In this chapter, we study how quantum guessing games can be formulated as SDPs and what understanding this formulation offers about them. We first review a standard form of semidefinite programs in Section 4.1 and the fundamental concept of duality of the SDPs. Then, in Section 4.2 we formulate the quantum guessing games as standard SDPs, and by studying the duality, we obtain necessary and sufficient conditions for the optimality of a measurement for a quantum guessing game. Although the contents of this chapter have all been investigated before, mainly in [MSU23] in the case of quantum guessing games, the derivation of the results presented here has been done independently by the author.

4.1 SDP Primal-Dual Formulation

Semidefinite programs are constrained optimization problems where the objective function is linear in a Hermitian matrix variable, and the constraints are linear matrix inequalities. In this section, we discuss the *standard form* of a semidefinite program, its dual problem, and two important properties known as weak and strong duality. This section is based on the presentation in [SC23].

A general form of a semidefinite program is as follows:

$$\begin{aligned} &\text{Primal SDP} \\ \text{maximize}: \;& \operatorname{tr}(AX), \\ \text{subject to}: \;& \Phi_i(X) = B_i, \quad i \in [m] \\ & \Gamma_j(X) \leq C_j, \quad j \in [n] \end{aligned} \tag{4.1}$$

where X is a Hermitian matrix variable, A, B_i's and C_j's are Hermitian matrix constants, and Φ_i's and Γ_j's are Hermiticity-preserving linear maps. Even though the above form is general, many SDPs encountered in practice are not initially written in this form. We refer to an optimization problem where the objective function is linear in Hermitian matrix variables and the constraints are linear matrix (in)equalities as a general SDP, while those in the above form are called *standard* SDPs.

Having an SDP formulation for a problem, which we call the primal problem, finding a lower bound on the optimal value is straightforward. One can find a feasible solution for the problem, and then the objective value of this feasible solution is a lower bound on the optimal value of the primal. Finding upper bounds, however, is not as simple, at least at first glance.

One can associate a *dual* problem with any SDP, which is an SDP, and its feasible solutions provide upper bounds on the optimal value of the primal problem. This dual problem can be obtained by writing the Lagrangian for the primal problem and then imposing proper conditions on the Lagrange multipliers to ensure that the derived constrained problem has the desired properties.

Associating Lagrange multipliers Y_i and Z_j with the equality and inequality constraints of the primal problem, respectively, the Lagrangian for the primal can be written as:

$$\mathcal{L}(X,\mathbf{Y},\mathbf{Z}) = \mathrm{tr}(AX) + \sum_{i=1}^{m} \mathrm{tr}(Y_i(B_i - \Phi_i(X))) + \sum_{j=1}^{n} \mathrm{tr}\left(Z_j\left(C_j - \Gamma_j(X)\right)\right) \tag{4.2}$$

$$= \mathrm{tr}\left(\left(A - \sum_{i=1}^{m} \Phi_i^*(Y_i) - \sum_{j=1}^{n} \Gamma_j^*\left(Z_j\right)\right)X\right) + \sum_{i=1}^{m} \mathrm{tr}(Y_i B_i) + \sum_{j=1}^{n} \mathrm{tr}\left(Z_j C_j\right). \tag{4.3}$$

We impose the constraints $Z_j \geq 0$ for all $j \in [n]$, to ensure that $\mathcal{L}(X,\mathbf{Y},\mathbf{Z}) \geq \mathrm{tr}(AX)$ for all feasible solutions $(X,\mathbf{Y},\mathbf{Z})$. Next, we impose the constraint $A - \sum_{i=1}^{m} \Phi_i^*(Y_i) - \sum_{j=1}^{n} \Gamma_j^*(Z_j) = 0$ to make the Lagrangian independent of X. This leads us to the dual problem, which is formulated as follows:

$$\begin{aligned} &\text{Dual SDP} \\ \text{minimize}: &\sum_{i=1}^{m} \mathrm{tr}(B_i Y_i) + \sum_{j=1}^{n} \mathrm{tr}(C_j Z_j), \\ \text{subject to}: &A - \sum_{i=1}^{m} \Phi_i^*(Y_i) - \sum_{j=1}^{n} \Gamma_j^*(Z_j) = 0, \\ &Z_j \geq 0, \quad j \in [n]. \end{aligned} \tag{4.4}$$

From our construction, it is evident that the value of every feasible solution of the dual is an upper bound on the optimal value of the primal. Hence, the optimal value of the dual is also an upper bound on the optimal value of the primal. This relationship is called *weak duality* and holds for every SDP.

Under some conditions that are satisfied in many practical examples, including almost all SDPs appearing in quantum information theory, the optimal values of the primal and dual problems are equal. This property is known as *strong duality*. When strong duality holds, the dual problem provides an alternative formulation of the original problem, initially formulated as the primal.

One famous condition for strong duality is given by Slater's theorem. According to this theorem, if the primal (or dual) has a finite optimal value and there exists a feasible solution that satisfies the inequality constraints strictly, then strong duality holds. It is sufficient for the primal and dual to be both strictly feasible for these conditions to be met.

When strong duality holds, from the equality of the objective values of the primal and the dual, one can derive a relationship between the optimal solutions of the primal and the dual, called the *complementary slackness*. Let X^* and $\mathbf{Y}^*, \mathbf{Z}^*$ be the optimal solutions of the primal and the dual, respectively. Then, if strong duality holds, from (4.2) we have

$$\sum_{i=1}^{m} \mathrm{tr}(Z_j^*(\Gamma_j(X^*) - C_j)) = 0, \tag{4.5}$$

implying that for all $j \in [n]$,

$$Z_j^*(\Gamma_j(X^*) - C_j) = 0 \tag{4.6}$$

4.2 Quantum Guessing Games as SDPs

In this section, we formulate quantum guessing games as standard SDPs. For quantum state discrimination this was done in [JŘF02; Eld03; WGM08; SC23], for quantum state antidiscrimination in [Ban+14] and for quantum guessing games in [MSU23]

4.2.1 Standard SDP Formulation

Remember that a quantum guessing game $G = (\mathcal{E}, R_{L\times N})$ is associated with the following optimization problem:

$$\begin{aligned} \text{maximize}: & \sum_{i=1}^{L}\sum_{j=1}^{N} R_{ij}\,\mathrm{tr}(\widetilde{\rho}_j M_i), \\ \text{subject to}: & \sum_{i=1}^{L} M_i = \mathbb{1}_d, \quad M_i \geq 0, \quad i \in [L]. \end{aligned} \tag{4.7}$$

It is clear that the above maximization problem is an informal semidefinite program. In the following, we will show how to formulate this problem as a formal SDP.

For a quantum guessing game $G = (\mathcal{E}, R_{L\times N})$, where $\widetilde{\rho}_1, \ldots, \widetilde{\rho}_N \in \mathcal{L}(\mathbb{C}^d)$, let $\mathcal{X} \stackrel{\text{def}}{=} \mathcal{L}(\mathbb{C}^d)$ and $\mathcal{Y} \stackrel{\text{def}}{=} \mathcal{L}(\mathbb{C}^L)$. We define the following matrices and maps:

$$A \stackrel{\text{def}}{=} \sum_{i=1}^{L} |i\rangle\langle i| \otimes (\sum_{j=1}^{N} R_{i,j}\widetilde{\rho}_j) \in \mathcal{Y}\otimes\mathcal{X}, \quad B \stackrel{\text{def}}{=} \mathbb{1}_d \in \mathcal{X}, \quad C \stackrel{\text{def}}{=} 0 \in \mathcal{X}, \tag{4.8}$$

$$\Gamma \stackrel{\text{def}}{=} -\mathcal{I}_{\mathcal{Y}\otimes\mathcal{X}}, \quad \text{and} \quad \Phi \stackrel{\text{def}}{=} \mathrm{tr}_{\mathcal{Y}}. \tag{4.9}$$

It is straightforward to see that A, B, and C are Hermitian matrices, and Φ and Γ are Hermiticity-preserving linear maps. Now, consider the following formal SDP:

$$\begin{aligned} & \text{Primal (Standard)} \\ \text{maximize}: \; & \mathrm{tr}(AX), \\ \text{subject to}: \; & \Phi(X) = B, \\ & \Gamma(X) \leq C. \end{aligned} \tag{4.10}$$

Proposition 4.2.1. *Problems (4.10) and (4.7) are equivalent.*

Proof. To see why Problem (4.10) is equivalent to the general SDP (4.7), note that for any POVM $M = \{M_i\}_{i\in[L]}$, we can define a matrix $X \in \mathcal{Y}\otimes\mathcal{X}$ as $X \stackrel{\text{def}}{=} \sum_{i=1}^{L} |i\rangle\langle i| \otimes M_i$. Then, we have

$$\Phi(X) = \mathrm{tr}_{\mathcal{Y}}(X) = \mathrm{tr}_{\mathcal{Y}}\left(\sum_{i=1}^{L} |i\rangle\langle i| \otimes M_i\right) = \sum_{i=1}^{L} M_i = \mathbb{1}_d = B, \tag{4.11}$$

and $X \geq 0$. Thus, X is a feasible solution for the standard SDP (4.10). The objective value of X is

$$\mathrm{tr}(AX) = \mathrm{tr}\left(\left(\sum_{i=1}^{L} |i\rangle\langle i| \otimes (\sum_{j=1}^{N} R_{i,j}\widetilde{\rho}_j)\right)\left(\sum_{i=1}^{L} |i\rangle\langle i| \otimes M_i\right)\right) \tag{4.12}$$

$$= \mathrm{tr}\left(\sum_{i=1}^{L} |i\rangle\langle i| \otimes (\sum_{j=1}^{N} R_{i,j}\widetilde{\rho}_j M_i)\right) \tag{4.13}$$

$$= \sum_{i=1}^{L}\sum_{j=1}^{N} R_{ij}\,\mathrm{tr}(\widetilde{\rho}_j M_i), \tag{4.14}$$

which is the objective value of the general SDP (4.7). This shows that the optimal value of the standard SDP (4.10) is greater than or equal to the optimal value of the general SDP (4.7).

On the other hand, for any feasible solution X of the standard SDP (4.10), we can write $X = \sum_{i,j\in[L]} |i\rangle\langle j| \otimes X_{i,j}$. Define $M_i \stackrel{\text{def}}{=} X_{i,i}$ for $i \in [L]$. Each M_i is a diagonal block of a positive definite matrix X, thus $M_i \geq 0$. Moreover,

$$\sum_{i=1}^{L} M_i = \sum_{i=1}^{L} X_{i,i} = \mathrm{tr}_{\mathcal{Y}}(\sum_{i=1}^{L} |i\rangle\langle i| \otimes X_{i,i}) = \mathrm{tr}_{\mathcal{Y}}(X) = B, \tag{4.15}$$

which implies that $M = \{M_i\}_{i\in[L]}$ is a POVM. Finally, we have

$$\operatorname{tr}(AX) = \operatorname{tr}\left(\left(\sum_{i=1}^{L} |i\rangle\langle i| \otimes (\sum_{j=1}^{N} R_{i,j}\widetilde{\rho}_j)\right)\left(\sum_{i,j\in[L]} |i\rangle\langle j| \otimes X_{i,j}\right)\right) \tag{4.16}$$

$$= \operatorname{tr}\left(\sum_{i=1}^{L}\sum_{k=1}^{L} |i\rangle\langle k| \otimes (\sum_{j=1}^{N} R_{i,j}\widetilde{\rho}_j X_{i,k})\right) \tag{4.17}$$

$$= \sum_{i=1}^{L}\sum_{j=1}^{N} R_{ij} \operatorname{tr}(\widetilde{\rho}_j M_i), \tag{4.18}$$

which shows that the optimal value of the standard SDP (4.10) is less than or equal to the optimal value of the general SDP (4.7). Therefore, Problems (4.10) and (4.7) are equivalent. □

Now, using the dual form provided in (4.4), we can write the dual problem of the standard SDP (4.10) as

$$\begin{aligned} &\text{Dual (standard)} \\ \text{minimize}:\ &\operatorname{tr}(Y), \\ \text{subject to}:\ &A - \mathbb{1}_{\mathcal{Y}} \otimes Y + Z = 0, \\ &Z \geq 0. \end{aligned} \tag{4.19}$$

or equivalently,

$$\begin{aligned} &\text{Dual (general)} \\ \text{minimize}:\ &\operatorname{tr}(Y), \\ \text{subject to}:\ &Y \geq \sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j, \quad i \in [L]. \end{aligned} \tag{4.20}$$

Proposition 4.2.2. *Strong duality holds for the SDP formulation of quantum guessing games.*

Proof. The primal has finite optimal value, as the optimal reward is always bounded above by LN. It is also strictly feasible, as $M = \{\frac{1}{L}\mathbb{1}_d\}_{i\in[L]}$ is a feasible solution. Thus, by Slater's theorem, strong duality holds. □

Remark 4.2.3. The dual problem (4.20) admits a simple geometric interpretation, as illustrated in 4.2.1. Consider the cone of positive semidefinite matrices $\operatorname{Pos}(\mathbb{C}^d)$, with zero matrix as the apex. For $i \in [L]$, each $\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j$ is a point in this cone, and we can translate a copy of the cone of PSD matrices to each of these points, which specify the constraints $Y \geq \sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j$. The dual problem (4.20) asks for a point in the intersection of these translated cones that is closest to the apex of the PSD cone with respect to the trace distance. This optimal distance is equal to the optimal reward of the quantum guessing game.

4.2.2 Optimality Criteria

Although finding a closed-form solution for the optimal measurements of a quantum guessing game is not easy, using the semidefinite primal and dual problems, we can obtain a simple-to-check necessary and sufficient optimality criterion for a measurement. The key tool for this is the complementary slackness. Using (4.6) and (4.19), the complementary slackness condition for a quantum guessing game can be written as:

$$(Y^* - \sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j)M_i^* = M_i^*(Y^* - \sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j) = 0, \quad i \in [L]. \tag{4.21}$$

where $M^* = \{M_i^*\}_{i\in[L]}$ and Y^* are the optimal solutions of the primal and the dual, respectively.

The following proposition is a generalization of the Holevo–Yuen–Kennedy–Lax theorem [Wat18]. A similar statement can be found in [Hel82].

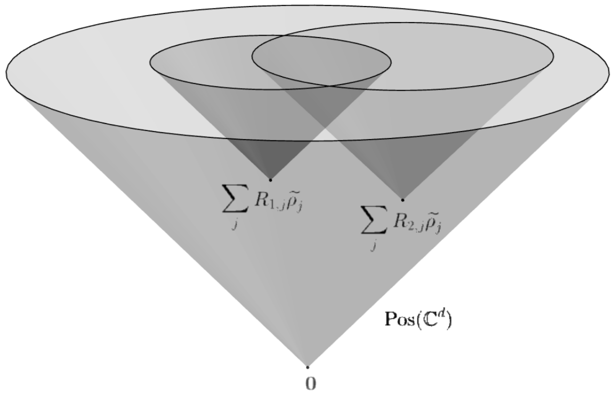


FIGURE 4.1: The cone of positive semidefinite matrices and the cones corresponding to the constraints of the dual problem of a quantum guessing game.

Proposition 4.2.4. *Let $G = (\mathcal{E}, R_{L\times N})$ be a quantum guessing game, and $M = \{M_i\}_{i\in[L]}$ be a measurement. The followings are equivalent:*

1. *$M = \{M_k\}_{k\in[L]}$ is an optimal measurement for G.*
2. *$Z \stackrel{\text{def}}{=} \sum_{i=1}^{L}(\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j)M_i$ is Hermitian and for all $i \in [L]$, $Z \geq \sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j$.*
3. *There exists an operator Y such that for all $i \in [L]$,*

$$Y \geq \sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j,$$

and

$$YM_i = \Big(\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j\Big)M_i.$$

Proof. ($1 \implies 3$): Let $M = \{M_i\}_{i\in[L]}$ be an optimal measurement for G. For any dual-optimal solution Y, by the complementary slackness, we have

$$YM_i = \Big(\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j\Big)M_i. \tag{4.22}$$

Note that Y is in particular a feasible solution of the dual, which implies that for $i \in [L]$, $Y \geq \sum_{j=1}^{N} R_{i,j}\widetilde{\rho}_j$.

($3 \implies 2$): Suppose that Y satisfies $YM_i = \big(\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j\big)M_i$ for all $i \in [L]$. By summimg up both sides over i, we have

$$Y = \sum_{i=1}^{L}\Big(\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j\Big)M_i. \tag{4.23}$$

Moreover, Y is Hermitian, as $Y \geq \sum_{j=1}^{N} R_{i,j}\widetilde{\rho}_j$ for all $i \in [L]$, and $\widetilde{\rho}_j$'s are Hermitian.

($2 \implies 1$): Note that Z is a feasible solution of the dual, and its objective value is $\mathrm{tr}(Z) = \sum_{i=1}^{N}\sum_{j=1}^{L} R_{ij}\,\mathrm{tr}(\widetilde{\rho}_j M_i)$. Since the dual and the primal objectives are equal for Z and M, it is implied that Z and M are optimal solutions of the dual and the primal. □

Chapter 5

Iterative Algorithms for Quantum Guessing Games

In the previous chapter, we discussed how the quantum state discrimination problem, or more generally, quantum guessing games, can be formulated as semidefinite programs. Although finding closed-form solutions for these problems is difficult, interior-point methods offer numerical solutions in polynomial time with respect to the input size. However, these methods become intractable as the input size increases, which has led researchers to explore alternative optimization algorithms that are more efficient and specifically designed for quantum state discrimination [Hel82; Tys09a; JŘF02].

One such alternative approach that has gained attention in recent years involves designing iterative algorithms based on the optimality conditions of an optimization problem. Over the past two decades, various iterative algorithms have been proposed for different quantum information problems, including quantum state discrimination [JŘF02; NKU15], maximum likelihood estimation in quantum process tomography [Hra+04], average fidelity estimation [AKF22], and the matrix projection problem [BRT23].

In this chapter, we focus on an iterative algorithm proposed by Ježek, Řeháček, and Fiurášek [JŘF02], referred to as the JRF algorithm. This algorithm is derived from the complementary slackness condition, and while numerical evaluations suggest it converges to the optimal solution, a formal proof of convergence is still missing in the general case. However, it has been proven to converge for the special case of linearly independent pure ensembles [NKU15].

We begin by generalizing the JRF algorithm to the setting of quantum guessing games in Section 5.1. In Section 5.2, we review a modified version of the JRF algorithm, proposed by Nakahira, Kato, and Usuda [NKU15], which we refer to as the NKU algorithm. In Section 5.3, we connect the JRF algorithm to a class of iterative methods known as Minorization-Maximization (MM) algorithms. The chapter concludes with a discussion on the fixed-point analysis of the JRF algorithm and its MM counterpart in Section 5.4.

5.1 JRF Algorithm

In [JŘF02], the authors proposed an iterative algorithm to solve the QSD problem based on the complementary slackness condition of semidefinite programs. In this section, we extend their algorithm to the more general setting of quantum guessing games. Throughout this section, unless stated otherwise, we assume that the ensemble $\mathcal{E} = \{\widetilde{\rho}_k\}_{k\in[N]}$ is full-rank and that the reward matrix R has no rows identical to zero. This assumption ensures that the operators for which we compute the inverse are indeed invertible. However, if this assumption does not hold, operator inverse can be replaced by the Moore-Penrose pseudoinverse, as was done by the authors in [JŘF02].

From Proposition 4.2.4, we know that a POVM $M = \{M_i\}_{i\in[L]}$ is optimal, if and only if there exists an operator Y such that for all $i \in [L]$,

$$Y \geq \sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j, \tag{5.1}$$

and

$$YM_i = \Big(\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j\Big)M_i, \quad \forall i \in [L]. \tag{5.2}$$

Condition (5.2) is thus a necessary condition for optimality. Note that (5.1) together with our assumption that R has no rows identical to zero and $\mathcal{E}$ is full-rank, it implies that Y is full-rank and, therefore, invertible.

Now assume that an invertible Hermitian operator Y satisfies (5.2). Right-multiplying both sides of (5.2) by Y^{-1}, we obtain

$$M_i = Y^{-1}\Big(\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j\Big)M_i. \tag{5.3}$$

Since Y, M_i's and $\widetilde{\rho}_j$'s are Hermitian, we can write (5.3) as

$$M_i = M_i\Big(\sum_{j=1}^{N} \widetilde{\rho}_j R_{ij}\Big)Y^{-1}. \tag{5.4}$$

Substituting (5.4) into (5.3), we obtain

$$M_i = Y^{-1}\Big(\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j\Big)M_i\Big(\sum_{j=1}^{N} \widetilde{\rho}_j R_{ij}\Big)Y^{-1}. \tag{5.5}$$

The above equation specifies another necessary condition for optimality of a measurement.

An iterative algorithm based on this condition can be constructed as follows. The algorithm starts by setting $M_i^{(0)} = \mathbb{1}_d/L$ for all $i \in [L]$, and at each iteration, it updates the measurements by

$$M_i^{(k+1)} = \Sigma^{(k)^{\frac{-1}{2}}}\Big(\sum_{j=1}^{N} \widetilde{\rho}_j R_{ij}\Big)M_i^{(k)}\Big(\sum_{j=1}^{N} \widetilde{\rho}_j R_{ij}\Big)\Sigma^{(k)^{\frac{-1}{2}}}, \tag{5.6}$$

where

$$\Sigma^{(k)} \stackrel{\text{def}}{=} \sum_{i=1}^{L}\Big(\sum_{j=1}^{N} \widetilde{\rho}_j R_{ij}\Big)M_i^{(k)}\Big(\sum_{j=1}^{N} \widetilde{\rho}_j R_{ij}\Big). \tag{5.7}$$

Note that at each iteration k, $\Sigma^{(k)}$ is a Hermitian operator, and since $M^{(0)}$ is a full-rank POVM, the ensemble is full-rank, and R has no rows identical to zero, $\Sigma^{(k)}$ is invertible for all k's.

One can define a mapping $M = \{M_i\}_{i\in[L]} \stackrel{\mathcal{T}}{\mapsto} M^{(+)} = \{M_i^{(+)}\}_{i\in[L]}$ as

$$\mathcal{T}(M) = \bigoplus_k \left(\Sigma^{\frac{-1}{2}}\left(\sum_{j=1}^{N} \widetilde{\rho}_j R_{ij}\right)M_i\left(\sum_{j=1}^{N} \widetilde{\rho}_j R_{ij}\right)\Sigma^{\frac{-1}{2}}\right), \tag{5.8}$$

on the space of L-tuples of operators acting on $\mathbb{C}^d$, where

$$\Sigma \stackrel{\text{def}}{=} \sum_{i=1}^{L}\left(\sum_{j=1}^{N} \widetilde{\rho}_j R_{ij}\right)M_i\left(\sum_{j=1}^{N} \widetilde{\rho}_j R_{ij}\right). \tag{5.9}$$

A sequence of measurements $\{M^{(k)}\}$ that is obtained by the JRF algorithm can be seen as a sequence obtained by the dynamic $\mathcal{T}$, i.e., $M^{(k+1)} = \mathcal{T}(M^{(k)})$. If this sequence converges to a fixed-point M, then M satisfies Condition (5.5), which is a necessary condition for optimality.

The JRF algorithm for quantum guessing games is summarized in Algorithm 2.

Remark 5.1.1. The map $\mathcal{T}$ defined in (5.8) can be seen as a composition of:

1. a sandwiching map $\mathcal{S}(M) \stackrel{\text{def}}{=} \bigoplus_i \big(\sum_{j=1}^{N} \widetilde{\rho}_j R_{ij}\big)M_i\big(\sum_{j=1}^{N} \widetilde{\rho}_j R_{ij}\big)$, and

Algorithm 2 JRF Algorithm for Quantum Guessing Games

Input: A quantum guessing game $G = \{\mathcal{E}, R_{L\times N}\}$, and the number of iterations T.

Initialize: Initial measurement $M^{(0)} = \{\frac{\mathbb{1}_d}{L}\}_{i\in[L]}$.

for $k = 0, 1, \ldots, T-1$ **do**

$$\Sigma^{(k)} \leftarrow \sum_{i=1}^{L}\Big(\sum_{j=1}^{N}\widetilde{\rho}_j R_{ij}\Big)M_i^{(k)}\Big(\sum_{j=1}^{N}\widetilde{\rho}_j R_{ij}\Big), \tag{5.10a}$$

$$M_i^{(k+1)} \leftarrow \Sigma^{(k)^{\frac{-1}{2}}}\Big(\sum_{j=1}^{N}\widetilde{\rho}_j R_{ij}\Big)M_i^{(k)}\Big(\sum_{j=1}^{N}\widetilde{\rho}_j R_{ij}\Big)\Sigma^{(k)^{\frac{-1}{2}}}. \tag{5.10b}$$

end for

Output: The measurement $M^{(T)} = \{M_i^{(T)}\}_{i\in[L]}$.

2. a normalization map $\mathcal{N}(M) \stackrel{\text{def}}{=\!=} \bigoplus_i (\sum_{i=1}^{L} M_i)^{\frac{-1}{2}} M_i (\sum_{i=1}^{L} M_i)^{\frac{-1}{2}}$.

The sandwiching map $\mathcal{S}$ is a linear map, and thus, it is continuous. However, we do not have a proof for the continuity of the normalization map $\mathcal{N}$, and it might not be true, as the map $M = \{M_i\}_{i\in[L]} \mapsto (\sum_{i=1}^{L} M_i)^{\frac{-1}{2}}$ is not continuous when $\sum_{i=1}^{L} M_i$ is singular.

Similar to the observation of Ježek, Řeháček, and Fiurášek [JŘF02], by running numerical simulations, we observe that our generalized JRF algorithm converges to the optimal solution. This is illustrated in Figure 5.1, where we plot the difference $\mathcal{R}_{\text{opt}} - \mathcal{R}_{M^{(k)}}$ against the number of iterations k for 50 random instances with $N = L = 3$ and $d = 2$. To obtain the optimal reward, we use the python API `PICOS` [SS22] with the solver `MOSEK` [ApS24]. For each instance, we generate a set of Haar-random states $\{\rho_i\}_{i\in[N]}$, using methods provided in the python package `toqito` [Rus21], and the prior probabilities and the reward matrices are generated uniformly at random.

For each iteration, the median of the differences $\mathcal{R}_{\text{opt}} - \mathcal{R}_{M^{(k)}}$ over the 50 instances and the interquartile range are plotted. As can be understood from the plot, the expected reward of the measurement $M^{(k)}$ converges to the optimal reward $\mathcal{R}_{\text{opt}}$ as the number of iterations increases. This convergence is, however, not so fast for some instances, and the algorithm might require a large number of iterations to converge.

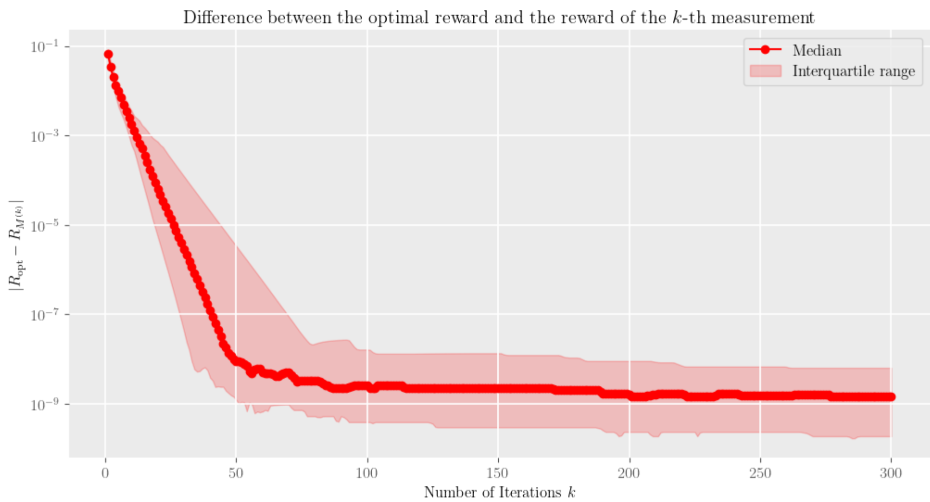


FIGURE 5.1: Convergence of the JRF algorithm for quantum guessing games. The difference $\mathcal{R}_{\text{opt}} - \mathcal{R}_{M^{(k)}}$ is plotted against the number of iterations for 50 random instances, each with $N = L = 3$ and $d = 2$.

To ensure that convergence behavior is observed across all samples, we plotted the difference $|\mathcal{R}_{\text{opt}} - \mathcal{R}_{M^{(k)}}|$ against the number of iterations for each random instance. An example of one such plot is shown in Figure 5.1. While a monotonically decreasing behavior is observed for $|\mathcal{R}_{\text{opt}} - \mathcal{R}_{M^{(k)}}|$ among all instances, as comparing Figures 5.1 and 5.2 suggests, the rate of convergence can vary significantly among different games.

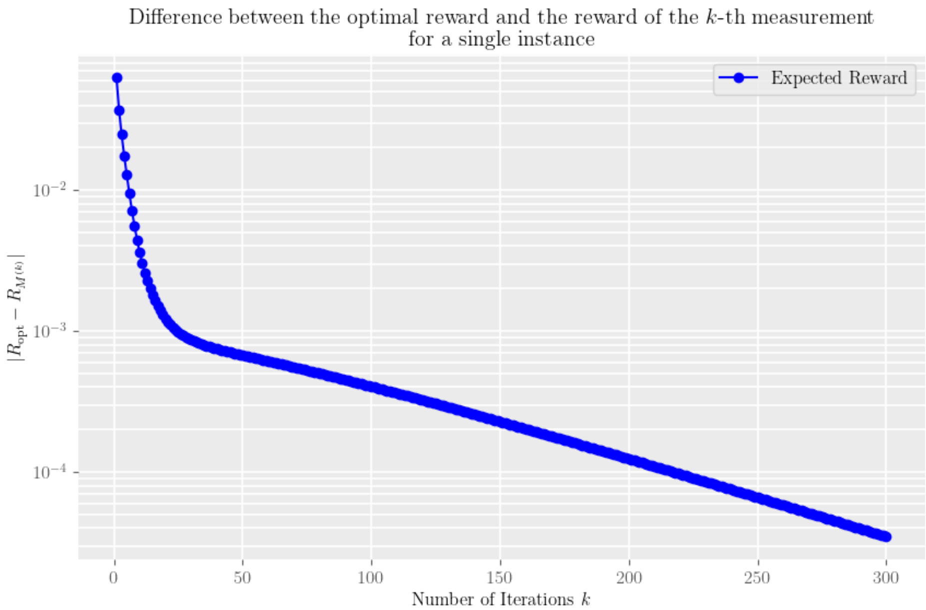


FIGURE 5.2: An example of the convergence behavior of the JRF algorithm for quantum guessing games. Here, the difference $\mathcal{R}_{\text{opt}} - \mathcal{R}_{M^{(k)}}$ is plotted against the number of iterations for one of the random instances used in the numerical results of Figure 5.1.

This low rate of convergence of the JRF algorithm can make it inefficient when a high precision solution is required. In Figure 5.3, we compare the runtime of the JRF algorithm with the runtime of the SDP solver `MOSEK` for different number of states N and different accuracy levels for the JRF measurements. We take our quantum guessing game to be the quantum state discrimination game over qubit ensembles with equal prior probabilities.

We plot the runtime of both algorithms against the number of states. The JRF output is considered for two accuracy levels, 10^{-2} and 10^{-5}, meaning that as soon as the difference between the optimal reward and the reward of the JRF measurements becomes smaller than the accuracy level, the JRF algorithm terminates. The runtimes are evaluated for 100 Haar-random ensembles for each N. The medians and the interquartile ranges are depicted in the plots. As can be seen from the plots, for low accuracy levels, the JRF algorithm is preferable to the SDP solvers, as it performs faster. However, due to the slow convergence of the JRF algorithm, it becomes inefficient for high accuracy levels, and the SDP solver is more preferable.

5.2 NKU Algorithm

A modified version of the JRF algorithm was proposed by Nakahira, Kato, and Usuda [NKU15], which we refer to as the NKU algorithm. Similar to the JRF algorithm, the NKU algorithm is based on a necessary condition for the optimality of a Belavkin measurement. However, it can also be seen as a special case of the JRF algorithm, in which instead of starting from the indifferent measurement $M^0 = \{\mathbb{1}/L\}i \in [L]$, one starts from a Belavkin measurement.

Proposition 5.2.1. *Let* $M^{(w)} = \{M_i^{(w)}\}_{i\in[N]}$ *be a Belavkin measurement with associated weights* $w = \{w_i\}_{i\in[N]}$ *for a QSD problem over the ensemble* $\mathcal{E} = \{(\rho_i, q_i)\}_{i\in[N]}$. *Then the JRF iterate of* $M^{(w)}$ *according to (5.6),* $(M^{(w)})^{(+)} = \{(M_i^{(w)})^{(+)}\}_{i\in[N]}$, *is also a Belavkin measurement with corresponding weights*

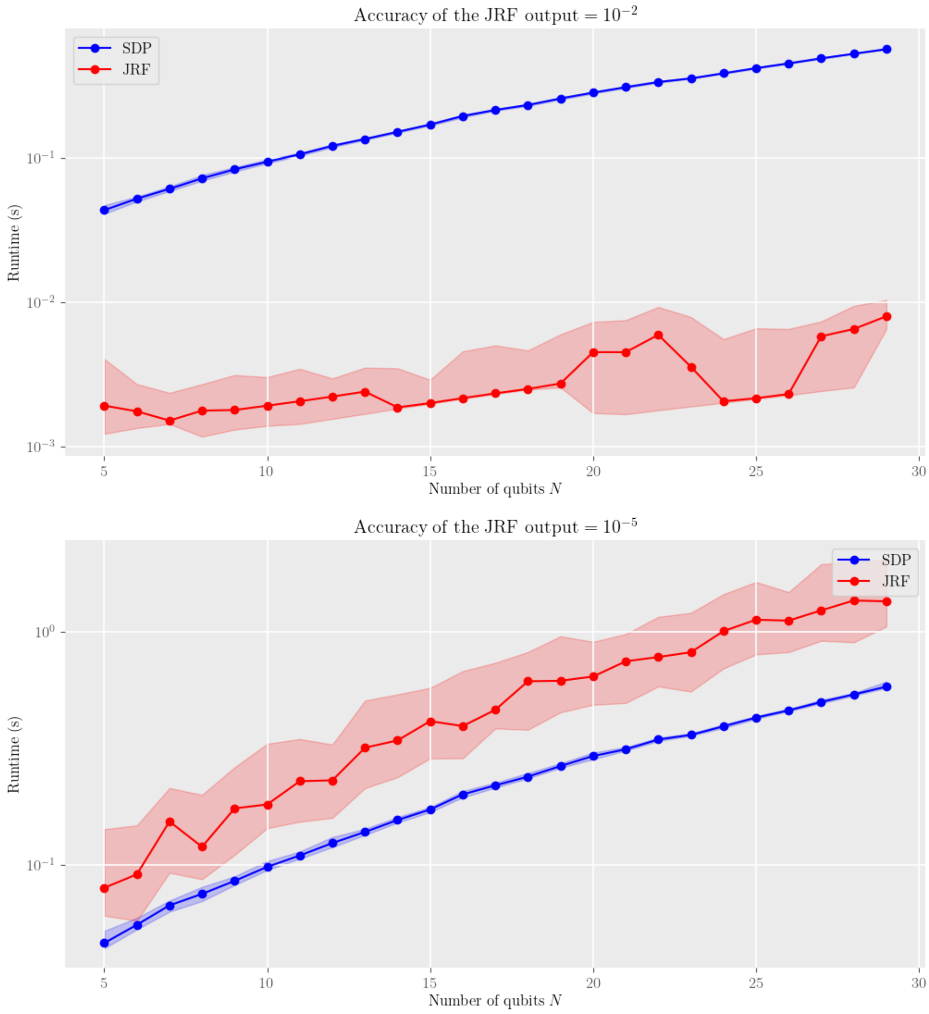


FIGURE 5.3: Runtime comparison of the JRF algorithm and the SDP solver MOSEK for quantum state discrimination over N qubits with equal prior probabilities. The accuracy of the JRF algorithm has been considered for two levels, 10^{-2} and 10^{-5}.

$w^{(+)} = \{w_i^{(+)}\}_{i\in[N]}$, *where*

$$w_i^{(+)} = q_i^2 Y_i^{(w)} w_i Y_i^{(w)}, \tag{5.11}$$

for all $i \in [N]$.

Proof. Recall from Section 3.2 that a Belavkin measurement $M^{(w)}$ is a POVM with elements $M_i^{(w)} = (\Sigma^{(w)})^{-\frac{1}{2}} \psi_i w_i \psi_i^\dagger (\Sigma^{(w)})^{-\frac{1}{2}}$. Applying the JRF update rule (5.6) to $M^{(w)}$, we have

$$(M_i^{(w)})^{(+)} = \left(\sum_{i=1}^{N} \widetilde{\rho}_i M_i^{(w)} \widetilde{\rho}_i\right)^{-\frac{1}{2}} \widetilde{\rho}_i M_i^{(w)} \widetilde{\rho}_i \left(\sum_{i=1}^{N} \widetilde{\rho}_i M_i^{(w)} \widetilde{\rho}_i\right)^{-\frac{1}{2}} \tag{5.12}$$

$$= \left(\sum_{i=1}^{N} q_i^2 \psi_i \psi_i^\dagger M_i^{(w)} \psi_i \psi_i^\dagger\right)^{\frac{-1}{2}} q_i^2 \psi_i \psi_i^\dagger M_i^{(w)} \psi_i \psi_i^\dagger \left(\sum_{i=1}^{N} q_i^2 \psi_i \psi_i^\dagger M_i^{(w)} \psi_i \psi_i^\dagger\right)^{\frac{-1}{2}} \tag{5.13}$$

Note that

$$\psi_i^\dagger M_i^{(w)}\psi_i = \psi_i^\dagger(\Sigma^{(w)})^{-\frac{1}{2}}\psi_i w_i \psi_i{}^\dagger(\Sigma^{(w)})^{-\frac{1}{2}}\psi_i = Y_i^{(w)} w_i Y_i^{(w)}. \tag{5.14}$$

Substituting $\psi_i^\dagger M_i^{(w)}\psi_i = Y_i^{(w)} w_i Y_i^{(w)}$ into (5.13), we obtain

$$(M_i^{(w)})^{(+)} = \left(\sum_{i=1}^N \psi_i(q_i^2 Y_i^{(w)} w_i Y_i^{(w)})\psi_i^\dagger\right)^{\frac{-1}{2}} \psi_i(q_i^2 Y_i^{(w)} w_i Y_i^{(w)})\psi_i^\dagger \left(\sum_{i=1}^N \psi_i(q_i^2 Y_i^{(w)} w_i Y_i^{(w)})\psi_i^\dagger\right)^{\frac{-1}{2}},$$

which shows that $(M^{(w)})^{(+)} = \{(M_i^{(w)})^{(+)}\}_{i\in[N]}$ is a Belavkin measurement with weights $w_i^{(+)} = q_i^2 Y_i^{(w)} w_i Y_i^{(w)}$. □

A detailed description of the NKU algorithm is provided in Algorithm 3.

Algorithm 3 NKU Algorithm for the QSD problem [NKU15]

Input: An ensemble $\mathcal{E} = \{\widetilde{\rho}\}_{i\in[N]}$, where $\rho_i = \psi_i\psi_i^\dagger$

Initialize: Initial weights $w^{(0)} = \{w_i^{(0)}\}_{i\in[N]}$, where $w_i^{(0)} = q_i \mathbb{1}_d$ (corresponding to the PGM), or $w_i^{(0)} = q_i\widetilde{\rho}_i$ (corresponding to the Holevo measurement).

for $k = 0, 1, \ldots, T-1$ **do**

$$Y_i^{(k)} \leftarrow \psi_i^\dagger \left(\sum_{j=1}^N \psi_j w_j^{(k)} \psi_j^\dagger\right)^{\frac{-1}{2}} \psi_i, \tag{5.15a}$$

$$w_i^{(k+1)} \leftarrow q_i^2 Y_i^{(k)} w_i^{(k)} Y_i^{(k)}. \tag{5.15b}$$

end for

Output: The Belavkin measurement $M^{(w^{(T)})}$, associated with the weights $w^{(T)} = \{w_i^{(T)}\}_{i\in[N]}$.

The main difference between the JRF algorithm and the NKU algorithm is that the NKU algorithm starts from a Belavkin measurement, and instead of updating the measurements, it updates the weights of the Belavkin measurement according to (5.11). As discussed in [NKU15], this modification has advantages: the NKU algorithm has lower computational complexity compared to the JRF algorithm, as the size of the weights can be smaller than the size of the measurement operators. Moreover, the speed of convergence can be accelerated by modifying the update rule to

$$w_i^{(+)} = q_i^2 (Y_i^{(w)})^\alpha w_i (Y_i^{(w)})^\alpha, \tag{5.16}$$

with $\alpha > 1$.

As mentioned earlier, the update rule of the NKU algorithm can be derived from a necessary condition for the optimality of a measurement. Recall from Proposition 3.2.7 that a necessary condition for optimality of a Belavkin measurement $M^{(w)}$ is that there exists $\alpha > 0$ such that for all $i \in [N]$,

$$w_i = q_i \alpha^{-1} Y_i^{(w)} w_i. \tag{5.17}$$

Note that in the NKU algorithm, if the weights $w^{(k)}$ converge to a fixed-point $w = \{w_i\}_{i\in[N]}$, that satisfies

$$w_i = (q_i Y_i^{(w)}) w_i (q_i Y_i^{(w)}), \tag{5.18}$$

from the fact that w_i's and $q_i Y_i$'s are PSD matrices, one can conclude [NKU15, Remark 10] that

$$w_i = q_i Y_i^{(w)} w_i, \tag{5.19}$$

which is Condition (5.17) with $\alpha = 1$.

We end this section by mentioning that the convergence of the NKU algorithm has been proven for the case of linearly independent pure ensembles [NKU15, Theorem 5]. The proof extensively uses the fact that when the states are pure, w_i's and Y_i's are positive numbers, and thus they commute with other operators.

5.3 JRF Algorithm as an MM Algorithm

In this section, building on the results of Tyson [Tys10], we relate the JRF algorithm to a well-studied class of iterative optimization methods known as MM (Minorization-Maximization) algorithms. Viewing the JRF algorithm as an MM algorithm not only deepens our understanding of the algorithm itself, but also allows us to explore whether the convergence of the JRF algorithm can be derived from the well-established convergence results of MM algorithms.

MM Algorithms

MM algorithms, which stand for minorization-maximization algorithms in maximization problems (or majorization-minimization in minimization problems), are a class of iterative optimization algorithms constructed based on the *MM principle*. The MM principle for maximizing an objective function $f(x)$ prescribes the following procedure: At each iteration k, first, the objective function $f(x)$ must be *minorized* by a surrogate function $g(x|x^{(k)})$, and then the surrogate function $g(x|x^{(k)})$ must be maximized to obtain a new iterate $x^{(k+1)}$. A schematic representation of the MM principle is shown in Figure 5.3, and the MM algorithm is summarized in Algorithm 4.

Algorithm 4 Minorization-Maximization (MM) Algorithm

Require: Objective function $f(x)$ to be maximized
Initialize: Initial iterate $x^{(0)}$
for $k = 0$ **to** $T-1$ **do**
    Minorize $f(x)$ by a surrogate function $g(x \mid x^{(k)})$ at $x^{(k)}$
    $x^{(k+1)} \leftarrow \arg\max_x g(x \mid x^{(k)})$
end for
Return: $x^{(T)}$

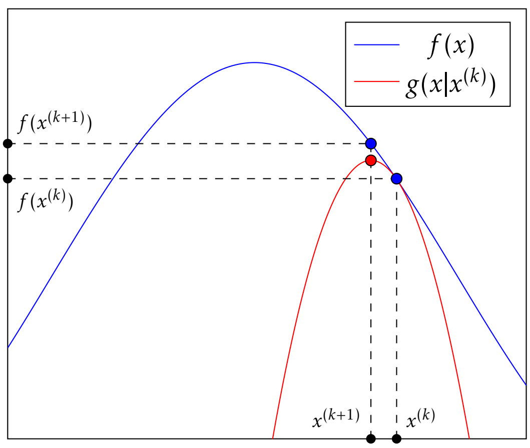


FIGURE 5.4: A schematic representation of the MM principle for maximization of an objective function $f(x)$. At each iteration, f is minorized by a surrogate function $g(x|x^{(k)})$ at the current iterate $x^{(k)}$, and then the surrogate function is maximized to obtain a new iterate $x^{(k+1)}$.

Minorization of f by a surrogate function $g(x|x^{(k)})$ at $x^{(k)}$ means that (1) $g(x|x^{(k)})$ is tangent to $f(x)$ at $x^{(k)}$, i.e., $g(x^{(k)}|x^{(k)}) = f(x^{(k)})$, and (2) $g(x|x^{(k)})$ is a lower bound for $f(x)$, i.e., $g(x|x^{(k)}) \leq f(x)$ for all x [Lan16]. These two conditions ensure that the sequence $\{f(x^{(k)})\}$ is non-decreasing (ascent property), as we have

$$f(x^{(k+1)}) \geq g(x^{(k+1)}|x^{(k)}) \geq g(x^{(k)}|x^{(k)}) = f(x^{(k)}). \tag{5.20}$$

Tyson's Directional Iteration

In [Tys10], Tyson showed that the quantum state discrimination can be seen as an instance of a more general optimization problem, which they called the *maximal seminorm problem*. The maximal seminorm problem is as follows:

Problem 5.3.1 (Maximal Seminorm Problem [Tys10])**.** Let V be a linear (real or complex) space equipped with a semidefinite inner product $\langle\cdot,\cdot\rangle : V \times V \to V$, and consider the seminorm induced by this semidefinite inner product on V. Let $S \subseteq V$. Find an element of S which is maximal with respect to this seminorm.

He suggested an iterative algorithm, called the *directional iteration*, as the consecutive applications of the following update rule:

Definition 5.3.2 (Directional Iteration [Tys10])**.** A directional iterate of $\mathbf{v} \in V$, is an element $\mathbf{v}^{(+)} \in \mathbf{S}$ such that $\mathbf{v}^{(+)} \in \arg\max_{\mathbf{s}\in S} \mathfrak{Re}\langle \mathbf{s}, \mathbf{v}\rangle$.

A directional iterate of an element is not necessarily unique. In the rest of this chapter, we denote the set-valued function mapping $\mathbf{v}$ to the set of its directional iterates by $\mathcal{D} : V \to 2^S$, i.e. $\mathbf{v}^{(+)} \in \mathcal{D}(\mathbf{v})$ for all $\mathbf{v} \in V$.

To better understand the directional iteration, let us compare it to the Frank-Wolfe algorithm [Pok24], which might appear quite similar at first glance. Assume that $\langle\cdot,\cdot\rangle$ is an inner product. Define $f : V \to \mathbb{R}$ as $f(\mathbf{v}) = \|\mathbf{v}\|^2$. Then, we have

$$\langle \nabla f(\mathbf{v}), \mathbf{s}\rangle = \frac{d}{dt}\bigg|_{t=0} f(\mathbf{v} + t\mathbf{s}) = 2\mathfrak{Re}\langle \mathbf{s}, \mathbf{v}\rangle, \tag{5.21}$$

which indicates that the directional iterate of $\mathbf{v}$ is the element $\mathbf{s} \in S$ that maximizes the inner product $\langle \nabla f(\mathbf{v}), \mathbf{s}\rangle$, where $f(\mathbf{v}) \overset{\text{def}}{=\!=} \|\mathbf{v}\|^2$.

This may remind the reader of the Frank-Wolfe algorithm, as the general idea behind both algorithms is indeed similar. However, there are significant differences between Tyson's directional iteration and the Frank-Wolfe algorithm. For instance, the Frank-Wolfe algorithm is typically applied to convex optimization problems, whereas the problem Tyson considered is not necessarily convex. Specifically, the constraint set S in Problem 5.3.1 is not necessarily convex.

JRF Algorithm as a Directional Iteration

In this Section, we show how the JRF Algorithm for quantum guessing games, introduced in Section 5.1 can be seen as a directional iteration. This was shown by Tyson in [Tys10] in the case of quantum state discrimination. For the moment, consider the quantum guessing games in their most general form, without any further assumptions on the ensembles nor reward matrices.

Example 5.3.3 (Quantum Guessing Games as Maximal Seminorm Problems)**.** Let $V = \mathcal{L}(\mathbb{C}^d)^L$ be the space of L-tuples of operators acting on $\mathbb{C}^d$. For two vectors of operators $\mathbf{E} = \{E_k\}_{k\in[L]} \in V$ and $\mathbf{F} = \{F_k\}_{k\in[L]} \in V$ and a quantum guessing game $G = (\mathcal{E}, R_{L\times N})$, we define the semidefinite inner $\langle\cdot,\cdot\rangle_G : V \times V \to V$ as

$$\langle \mathbf{E}, \mathbf{F}\rangle_G \overset{\text{def}}{=\!=} \sum_{i=1}^{L}\sum_{j=1}^{N} R_{ij}\,\mathrm{tr}(E_i^\dagger F_i \widetilde{\rho}_j). \tag{5.22}$$

Define $S_G \overset{\text{def}}{=\!=} \{\mathbf{E} \in V \mid \sum_{i=1}^{L} E_i^\dagger E_i = \mathbb{1}_d\}$. This way, we have $\mathcal{R}_{\text{opt}} = \max_{\mathbf{E}\in S_G} \|\mathbf{E}\|_G^2$. Now it is clear that the task of finding the optimal success probability of discriminating between the states of a quantum ensemble is a special case of the maximal seminorm problem. The elements of the set S_G are known as the *generalised quantum measurements*.

Remark 5.3.4. Under the further assumption that the ensemble is full-rank and the reward matrix R has no rows identical to zero, the semidefinite inner product in Example 5.3.3 turns into an inner product (See Proposition A.1.7), and Problem 5.3.1 will become a norm maximization problem.

Remark 5.3.5. The set of generalized quantum measurements S_G is a compact set, as it is bounded, since for any quantum guessing game $G = (\mathcal{E}, R_{L\times N})$ that $\|\cdot\|_G^2$ is a norm,

$$\sup_{\mathbf{E}\in S_G} \|\mathbf{E}\|_G^2 \leq 1,$$

and it is closed, as S_G is the pre-image of the closed set $\{\mathbb{1}_d\}$ under the continuous map $\mathbf{E} \mapsto \sum_{i=1}^{L} E_i^\dagger E_i$. However, S_G is not convex.

Now, we can show how the JRF algorithm can be related to the directional iteration. As we will see, a sequence generated by the JRF algorithm is a possible sequence of directional iterations. This was first shown in [Tys10, Theorem 24] for the QSD problem, and we present it here for quantum guessing games. Hereafter, we again make the assumption on the games we consider that their ensembles are full-rank and their reward matrices have no rows identical to zero.

Proposition 5.3.6. *Let $G = (\mathcal{E}, R_{L\times N})$ be a quantum guessing game, and consider the maximal norm problem introduced in Example 5.3.3. For $\mathbf{E} \in S_G$, one of the possible directional iterates of $\mathbf{E} = \{E_i\}_{i\in[L]}$, $\mathbf{E}^{(+)} = \{E_i^{(+)}\}_{i\in[L]}$, is given by*

$$E_i^{(+)} = E_i\left(\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j\right)\left(\sum_{i=1}^{L}\left(\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j\right)E_i^\dagger E_i\left(\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j\right)\right)^{-\frac{1}{2}}. \tag{5.23}$$

Proof. By definition, $\mathbf{E}^{(+)} \in \arg\max_{\mathbf{F}\in S_G} \Re\langle \mathbf{F}, \mathbf{E}\rangle_G$. Consider the following operators:

$$V_{\mathbf{E}} \overset{\text{def}}{=} \sum_{i=1}^{L} E_i\left(\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j\right)\otimes|i\rangle, \quad U_{\mathbf{F}} \overset{\text{def}}{=} \sum_{i=1}^{L} F_i\otimes|i\rangle. \tag{5.24}$$

Note that

$$\operatorname{tr}(V_{\mathbf{E}}{}^\dagger U_{\mathbf{F}}) = \operatorname{tr}\left(\left[\sum_{i=1}^{L}\left(\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j\right)E_i{}^\dagger\otimes\langle i|\right]\left[\sum_{i'=1}^{L} F_{i'}\otimes|i'\rangle\right]\right) \tag{5.25}$$

$$= \operatorname{tr}\left(\sum_{i=1}^{L}\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j E_i{}^\dagger F_i\right) = \langle \mathbf{E}, \mathbf{F}\rangle_G, \tag{5.26}$$

which implies that $\mathbf{E}^{(+)} \in \arg\max_{\mathbf{F}\in S_G} \Re\operatorname{tr}(V_{\mathbf{E}}{}^\dagger U_{\mathbf{E}})$.
Note that for any $\mathbf{F}\in S_G$, $\|U_{\mathbf{F}}\|_\infty = 1$. On the other hand, from Hölder's inequality, we know that

$$\max_{\|U\|_\infty\leq 1} \Re\operatorname{tr}(V_{\mathbf{E}}{}^\dagger U) = \|V_{\mathbf{E}}\|_1, \tag{5.27}$$

and

$$V_{\mathbf{E}}\left(V_{\mathbf{E}}{}^\dagger V_{\mathbf{E}}\right)^{-\frac{1}{2}} \in \arg\max_{\|U\|_\infty\leq 1} \Re\operatorname{tr}(V_{\mathbf{E}}{}^\dagger U). \tag{5.28}$$

Now, it is easy to see that for $\mathbf{F}\in S_G$, defined as

$$F_i \overset{\text{def}}{=} E_i\left(\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j\right)\left(\sum_{i=1}^{L}\left(\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j\right)E_i^\dagger E_i\left(\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j\right)\right)^{-\frac{1}{2}}, \tag{5.29}$$

we have

$$U_{\mathbf{F}} = V_{\mathbf{E}}\left(V_{\mathbf{E}}{}^\dagger V_{\mathbf{E}}\right)^{-\frac{1}{2}}, \tag{5.30}$$

which completes the proof. □

Directional Iteration as a Bregman Minorization

It is easy to see that Tyson's directional iteration is a special case of MM algorithms. Let us first convert the constrained optimization in Problem 5.3.1 to an unconstrained optimization, by defining the function

$$f(\mathbf{v}) \stackrel{\text{def}}{=} \begin{cases} \|\mathbf{v}\|_G^2 & \mathbf{v} \in S_G \\ -\infty & \mathbf{v} \in V \setminus S_G \end{cases}. \tag{5.31}$$

The maximum of f is then equal to the optimal value of Problem 5.3.1. Define a surrogate function $g(\mathbf{v}|\mathbf{v}^{(k)})$ as

$$g(\mathbf{v}|\mathbf{v}^{(k)}) \stackrel{\text{def}}{=} \begin{cases} \|\mathbf{v}^{(k)}\|_G^2 + 2\Re\langle \mathbf{v} - \mathbf{v}^{(k)}, \mathbf{v}^{(k)}\rangle_G & \mathbf{v} \in S_G \\ -\infty & \mathbf{v} \in V \setminus S_G \end{cases}. \tag{5.32}$$

We see that g is a minorization of f, as

$$g(\mathbf{v}|\mathbf{v}^{(k)}) = \|\mathbf{v}^{(k)}\|_G^2 + 2\Re\langle \mathbf{v} - \mathbf{v}^{(k)}, \mathbf{v}^{(k)}\rangle_G \leq \|\mathbf{v}\|_G^2 = f(\mathbf{v}), \tag{5.33}$$

and $g(\mathbf{v}^{(k)}|\mathbf{v}^{(k)}) = \|\mathbf{v}^{(k)}\|_G^2 = f(\mathbf{v}^{(k)})$. Note that the maximization of the surrogate function $g(\mathbf{v}|\mathbf{v}^{(k)})$ is equivalent to maximizing $\Re\langle \mathbf{v}, \mathbf{v}^{(k)}\rangle_G$ subject to $\mathbf{v} \in S_G$. This is exactly the definition of the directional iterate. Thus, the directional iteration is a special case of the MM algorithms.

Since for a sequence $\{\mathbf{v}^{(k)}\}$ generated by an MM algorithm, we have $f(\mathbf{v}^{(k)}) \leq f(\mathbf{v}^{(k+1)})$, and we showed that the directional iteration is an MM algorithm, the sequence $\{\|\mathbf{v}^{(k)}\|_G^2\}$ generated by the directional iteration is non-decreasing. It is, however, possible to obtain a stronger result for the directional iteration:

Proposition 5.3.7 ([Tys10, Lemma 7]). *For a sequence $\{\mathbf{v}^{(k)}\}$ generated by the directional iteration, we have*

$$\|\mathbf{v}^{(k+1)}\|_G^2 \geq \|\mathbf{v}^{(k)}\|_G^2 + \|\mathbf{v}^{(k+1)} - \mathbf{v}^{(k)}\|_G^2. \tag{5.34}$$

Proof. We have

$$\|\mathbf{v}^{(k+1)}\|_G^2 = \|\mathbf{v}^{(k)} + (\mathbf{v}^{(k+1)} - \mathbf{v}^{(k)})\|_G^2 \tag{5.35}$$

$$= \|\mathbf{v}^{(k)}\|_G^2 + \|\mathbf{v}^{(k+1)} - \mathbf{v}^{(k)}\|_G^2 + 2\Re\langle \mathbf{v}^{(k)}, \mathbf{v}^{(k+1)} - \mathbf{v}^{(k)}\rangle_G, \tag{5.36}$$

and since from the definition of the directional iterate, we have

$$\Re\langle \mathbf{v}^{(k)}, \mathbf{v}^{(k+1)}\rangle_G \geq \Re\langle \mathbf{v}^{(k)}, \mathbf{v}^{(k)}\rangle_G, \tag{5.37}$$

(5.34) follows. □

Hereafter, we will limit ourselves to guessing games with full-rank ensembles, where the reward matrices have no row identical to zero.

Corollary 5.3.8. *Let $\{M^{(k)}\}$ be a sequence generated by the JRF algorithm, where $M^{(k)} = \{(E_i^{(k)})^\dagger E_i^{(k)}\}_{i\in[L]}$, and*

$$E_i^{(k+1)} = E_i^{(k)} \left(\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j\right) \left(\sum_{i=1}^{L} \left(\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j\right) (E_i^{(k)})^\dagger E_i^{(k)} \left(\sum_{j=1}^{N} R_{ij}\widetilde{\rho}_j\right)\right)^{-\frac{1}{2}}. \tag{5.38}$$

As shown above, $\mathcal{R}_{M^{(k)}} = \|\mathbf{E}^{(k)}\|_G^2$. The reward sequence $\{\mathcal{R}_{M^{(k)}}\}$ is non-decreasing and bounded, thus it converges. From Proposition 5.3.7, and knowing that

$$\lim_{k\to\infty} \|\mathbf{E}^{(k+1)}\|_G^2 - \|\mathbf{E}^{(k)}\|_G^2 = 0, \tag{5.39}$$

we conclude that the sequence $\{\|\mathbf{E}^{(k+1)} - \mathbf{E}^{(k)}\|_G^2\}$ converges to zero. From this, together with the fact that $\{\|\mathbf{E}^{(k)}\|_G^2\}$ is convergent, it is implied that the sequence $\{\mathbf{E}^{(k)}\}$ is a Cauchy sequence, and since S_G is a compact subset of the complete space V, the sequence $\{\mathbf{E}^{(k)}\}$ converges to an element of S_G. This implies that the JRF iterations are also convergent to a POVM.

The surrogate function $g(\mathbf{v}|\mathbf{v}^{(k)})$ in (5.32) is in fact an example of a class of minorization functions known as *Bregman minorizations*. Bregman minorization utilizes the Bregman divergence to define the surrogate function. The Bregman divergence is a measure of the difference between two points in a space, defined with respect to a convex function, as illustrated in Figure 5.3.

Definition 5.3.9 (Bregman Divergence [Teb18])**.** Let V be an inner product space and $\phi : V \to [-\infty, \infty]$ be a Legendre function. The Bregman divergence associated to ϕ, which is denoted by $D_\phi : \operatorname{dom}\phi \times \mathbf{int}\operatorname{dom}\phi \to \mathbb{R}_+$, is defined as

$$D_\phi(\mathbf{x}\|\mathbf{y}) \stackrel{\text{def}}{=} \phi(\mathbf{x}) - \phi(\mathbf{y}) - \langle \nabla\phi(\mathbf{y}), \mathbf{x} - \mathbf{y}\rangle. \tag{5.40}$$

Proposition 5.3.10. *Let D_ϕ be the Bregman divergence associated to a Legendre function ϕ. Then, for any $\mathbf{x}, \mathbf{y} \in V$, we have*

1. $D_\phi(\mathbf{x}\|\mathbf{y}) \geq 0$,
2. $D_\phi(\mathbf{x}\|\mathbf{y}) = 0$ *if and only if* $\mathbf{x} = \mathbf{y}$.

Proof. Note that ϕ is a Legendre function and in particular a strictly convex function differentiable on $\mathbf{int}\operatorname{dom}\phi$. Convexity of the differentiable function ϕ implies that

$$\phi(\mathbf{x}) \geq \phi(\mathbf{y}) + \langle \nabla\phi(\mathbf{y}), \mathbf{x} - \mathbf{y}\rangle, \tag{5.41}$$

and the statement follows from this inequality. □

Using the Bregman divergence, one can define a surrogate function $g(\mathbf{x}|\mathbf{x}^{(k)})$ for maximizing $f(\mathbf{x})$ as follows [Lan16]:

$$g(\mathbf{x}|\mathbf{x}^{(k)}) \stackrel{\text{def}}{=} f(\mathbf{x}) - D_\phi(\mathbf{x}\|\mathbf{x}^{(k)}). \tag{5.42}$$

From Proposition 5.3.10, it is clear that $g(\mathbf{x}|\mathbf{x}^{(k)}) \leq f(\mathbf{x})$ for all $\mathbf{x}$, and $g(\mathbf{x}^{(k)}|\mathbf{x}^{(k)}) = f(\mathbf{x}^{(k)})$. Thus, the function $g(\mathbf{x}|\mathbf{x}^{(k)})$ is a minorization of $f(\mathbf{x})$. Note that for using the Bregman divergence in constructing a surrogate function, we implicitely assume that at each iteration, $\mathbf{x}^{(k)} \in \mathbf{int}\operatorname{dom}\phi$ (as noted in [Byr08]).

Example 5.3.11. For a quantum guessing game $G = (\mathcal{E}, R_{L\times N})$, let $\phi_G : \mathcal{L}(\mathbb{C}^d)^L \to \mathbb{R}$ be the function defined as

$$\phi_G(\mathbf{x}) \stackrel{\text{def}}{=} \|\mathbf{x}\|_G^2, \tag{5.43}$$

where $\|\cdot\|_G^2$ is the norm induced by the inner product defined in Example 5.3.3. Let $\langle\cdot,\cdot\rangle$ be the real-valued inner product on $\mathcal{L}(\mathbb{C}^d)^L$ defined as

$$\langle \mathbf{x}, \mathbf{y}\rangle \stackrel{\text{def}}{=} \mathfrak{Re}\langle \mathbf{x}, \mathbf{y}\rangle_G. \tag{5.44}$$

The function ϕ_G is a Legendre function, and the Bregman divergence associated to ϕ_G is

$$D_{\phi_G}(\mathbf{x}\|\mathbf{y}) = \|\mathbf{x}\|_G^2 - \|\mathbf{y}\|_G^2 - \mathfrak{Re}\langle 2\mathbf{y}, \mathbf{x} - \mathbf{y}\rangle_G. \tag{5.45}$$

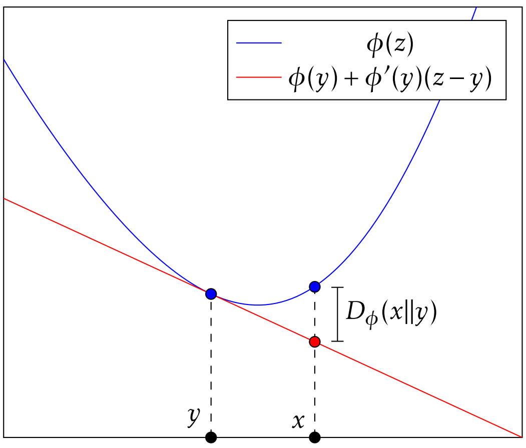


FIGURE 5.5: Illustration of the Bregman divergence $D_\phi(\mathbf{x}\|\mathbf{y})$ associated to a Legendre function $\phi : \mathbb{R} \to \mathbb{R}$.

We can now use the Bregman divergence defined in Example 5.3.11 to obtain a surrogate function for the objective

$$f(\mathbf{v}) = \begin{cases} \|\mathbf{v}\|_G^2 & \mathbf{v} \in S_{\mathcal{E}}, \\ -\infty & \mathbf{v} \in V \setminus S_G. \end{cases} \tag{5.46}$$

Rewriting (5.42) for this function, we obtain

$$g(\mathbf{v}|\mathbf{v}^{(k)}) \stackrel{\text{def}}{=} f(\mathbf{v}) - \|\mathbf{v}\|_G^2 + \|\mathbf{v}^{(k)}\|_G^2 + \Re\langle 2\mathbf{v}^{(k)}, \mathbf{v} - \mathbf{v}^{(k)}\rangle_G \tag{5.47}$$

$$= \begin{cases} \|\mathbf{v}^{(k)}\|_G^2 + 2\Re\langle \mathbf{v} - \mathbf{v}^{(k)}, \mathbf{v}^{(k)}\rangle_G, & \mathbf{v} \in S_G \\ -\infty & \mathbf{v} \in V \setminus S_G. \end{cases} \tag{5.48}$$

This way, we see that the surrogate function defined in (5.32) is indeed a Bregman minorization for the function $f(\mathbf{v})$.

Remark 5.3.12. In [Byr08], the author introduces a condition, called the *SUMMA condition*, under which the global convergence of MM algorithms is guaranteed. The SUMMA condition for a minorization-maximization algorithm can be expressed as

$$g(\mathbf{x}^{(k+1)}|\mathbf{x}^{(k)}) - g(\mathbf{x}|\mathbf{x}^{(k)}) \geq f(\mathbf{x}) - g(\mathbf{x}|\mathbf{x}^{(k+1)}). \tag{5.49}$$

It is proven in [Byr08] that if $p^* = \sup_{\mathbf{x}\in S} f(\mathbf{x})$ is finite and the SUMMA condition (5.49) holds, then $\lim_{k\to\infty} f(\mathbf{x}^{(k)}) = p^*$. Moreover, it is shown in [Byr08] and in a more general setting in [Lan16] that the SUMMA condition holds for Bregman minorizations for maximizing concave functions. The latter result, however, can not be applied to our case, as the function f defined in (5.46) is not concave.

5.4 Fixed-point Analysis

Fixed-point theory plays a crucial role in the development and analysis of iterative algorithms. In particular, as mentioned in Section 5.1, if the JRF iterations converge to a fixed-point of the update rule, the limit will satisfy a necessary optimality condition. Therefore, to analyse the convergence of the JRF algorithm, it is helpful to study the fixed-points of the iterations we have discussed so far.

For simplicity, the following discussion is limited to the QSD problem over full-rank ensembles. We consider fixed-points of three maps:

- the directional iteration map $\mathcal{D}$ for the maximal norm problem

$$\max_{\mathbf{E}\in S_G} \|\mathbf{E}\|_{\mathcal{E}}^2, \tag{5.50}$$

  where $\|\mathbf{E}\|_{\mathcal{E}}^2 \stackrel{\text{def}}{=} \sum_{i=1}^N \operatorname{tr}(E_i^\dagger E_i \widetilde{\rho}_i)$, and S_G is the set of generalized measurements,

- the map $\mathbf{E} = \{E_i\}_{i\in[N]} \stackrel{\mathcal{D}_{\text{JRF}}}{\mapsto} \mathbf{E}^{(+)} = \{E_i^{(+)}\}_{i\in[N]}$ where

$$E_i^{(+)} = E_i \widetilde{\rho}_i \left(\sum_{i=1}^N \widetilde{\rho}_i E_i^\dagger E_i \widetilde{\rho}_i \right)^{-\frac{1}{2}}, \tag{5.51}$$

- the JRF iteration map $\mathcal{T}$ defined as

$$\mathcal{T}(M) = \bigoplus_i \left(\sum_{i=1}^N \widetilde{\rho}_i M_i \widetilde{\rho}_i \right)^{-\frac{1}{2}} \widetilde{\rho}_i M_i \widetilde{\rho}_i \left(\sum_{i=1}^N \widetilde{\rho}_i M_i \widetilde{\rho}_i \right)^{-\frac{1}{2}}. \tag{5.52}$$

We saw earlier that $\mathcal{D}_{\text{JRF}}$ maps $\mathbf{E}$ to a directional iteration of $\mathbf{E}$, meaning that $\mathcal{D}_{\text{JRF}}(\mathbf{E}) \in \mathcal{D}(\mathbf{E})$. Thus, the fixed-points of $\mathcal{D}_{\text{JRF}}$ are also fixed-points of $\mathcal{D}$.

On the Fixed-points of the JRF Iteration

Proposition 5.4.1. *Let $M = \{M_i\}_{i\in[N]}$ be an optimal measurement for the QSD problem over $\mathcal{E}$. Then, M is a fixed-point of the JRF iteration, i.e. $M = \mathcal{T}(M)$.*

Proof. We need to show that for all $i \in [N]$, we have

$$M_i = \left(\sum_{i=1}^{N} \widetilde{\rho}_i M_i \widetilde{\rho}_i\right)^{-\frac{1}{2}} \widetilde{\rho}_i M_i \widetilde{\rho}_i \left(\sum_{i=1}^{N} \widetilde{\rho}_i M_i \widetilde{\rho}_i\right)^{-\frac{1}{2}}. \tag{5.53}$$

Recall from Proposition 4.2.4 that for an optimal measurement M, we have

$$\left(\sum_{i=1}^{N} M_i \widetilde{\rho}_i\right) M_i = \widetilde{\rho}_i M_i, \tag{5.54}$$

for all $i \in [N]$, or equivalently

$$M_i = \left(\sum_{i=1}^{N} M_i \widetilde{\rho}_i\right)^{-1} \widetilde{\rho}_i M_i. \tag{5.55}$$

Left-multiplying both sides of (5.54) by $\widetilde{\rho}_i$, and summing over i, we obtain

$$\left(\sum_{i=1}^{N} \widetilde{\rho}_i M_i\right)^2 = \sum_{i=1}^{N} \widetilde{\rho}_i M_i \widetilde{\rho}_i. \tag{5.56}$$

Thus, we have

$$\left(\sum_{i=1}^{N} \widetilde{\rho}_i M_i \widetilde{\rho}_i\right)^{-\frac{1}{2}} \widetilde{\rho}_i M_i \widetilde{\rho}_i \left(\sum_{i=1}^{N} \widetilde{\rho}_i M_i \widetilde{\rho}_i\right)^{-\frac{1}{2}} = \left(\sum_{i=1}^{N} \widetilde{\rho}_i M_i\right)^{-1} \widetilde{\rho}_i M_i \widetilde{\rho}_i \left(\sum_{i=1}^{N} \widetilde{\rho}_i M_i\right)^{-1} \tag{5.57}$$

$$= M_i \widetilde{\rho}_i \left(\sum_{i=1}^{N} \widetilde{\rho}_i M_i \widetilde{\rho}_i\right)^{-1} \tag{5.58}$$

$$= M_i, \tag{5.59}$$

where the second line follows from (5.55), and the third line follows from taking the adjoint of the both sides of (5.55). ☐

On the Fixed-points of the Directional Iteration

Proposition 5.4.2. *Let V be a linear space equipped with an inner product $\langle\cdot,\cdot\rangle$, and let $S \subseteq V$ be a compact set. Let $\mathcal{D} : V \longrightarrow 2^S$ be the set-valued function mapping $\mathbf{v}$ to the set of its directional iterates. Then, for any sequence $\{\mathbf{v}^{(k)}\}$ generated by the directional iteration, $\{\mathbf{v}^{(k)}\}$ converges to $\mathbf{v}^* \in S$, where $\mathbf{v}^*$ is a fixed-point of the directional iteration, i.e. $\mathbf{v}^* \in \mathcal{D}(\mathbf{v}^*)$.*

Proof. The proof is similar to the proof of [FKP21, Theprem 1]. First, note that $f(\mathbf{v}) = \|\mathbf{v}\|^2$ is a continuous function, and since S is compact, f attains its maximum on S. Let $p^* = \max_{\mathbf{v}\in S} \|\mathbf{v}\|^2$.

For a sequence $\{\mathbf{v}^{(k)}\}$ generated by the directional iteration, the sequence $\{\|\mathbf{v}^{(k)}\|^2\}$ is bounded, and by Proposition 5.3.7, is non-decreasing. Therefore, it is convergent. By a similar argument as what we have done in Corollary 5.3.6, it can be implied that the sequence $\{\mathbf{v}^{(k)}\}$ is convergent. If $\lim_{k\to\infty} \mathbf{v}^{(k)} = \mathbf{v}^*$, then since $\lim_{k\to\infty} \|\mathbf{v}^{(k)}\|^2 = p^*$ and $\|\cdot\|^2$ is continuous, we have $\|\mathbf{v}^*\|^2 = p^*$.

Now, we can show that $\mathbf{v}^*$ is a fixed-point of the directional iteration map. Assume that $\mathbf{v}^* \notin \mathcal{D}(\mathbf{v}^*)$. Then, for all $(\mathbf{v}^*)^{(+)} \in \mathcal{D}(\mathbf{v}^*)$, we have

$$\mathfrak{Re}\langle(\mathbf{v}^*)^{(+)}, \mathbf{v}^*\rangle > \mathfrak{Re}\langle\mathbf{v}^*, \mathbf{v}^*\rangle = p^*. \tag{5.60}$$

Since $\{\mathbf{v}^{(k)}\}$ converges to $\mathbf{v}^*$, and $\Re\langle\cdot,\cdot\rangle$ is continuous in its both arguments, there exists k_0 such that for all $k > k_0$, we have,

$$\Re\langle(\mathbf{v}^*)^{(+)}, \mathbf{v}^{(k)}\rangle > \Re\langle\mathbf{v}^*, \mathbf{v}^*\rangle. \tag{5.61}$$

Moreover, from the definition of the directional iteration, we have

$$\Re\langle\mathbf{v}^{(k+1)}, \mathbf{v}^{(k)}\rangle \geq \Re\langle(\mathbf{v}^{(*)})^{(+)}, \mathbf{v}^{(k)}\rangle, \tag{5.62}$$

and from $\|\mathbf{v}^{(k+1)} - \mathbf{v}^{(k)}\|^2 \geq 0$, we have

$$\Re\langle\mathbf{v}^{(k+1)}, \mathbf{v}^{(k+1)}\rangle \geq \Re\langle\mathbf{v}^{(k+1)}, \mathbf{v}^{(k)}\rangle. \tag{5.63}$$

Inequalities (5.61), (5.62) and (5.63) together imply that for all $k > k_0$,

$$\|\mathbf{v}^{(k+1)}\|^2 = \Re\langle\mathbf{v}^{(k+1)}, \mathbf{v}^{(k+1)}\rangle > p^*, \tag{5.64}$$

which is a contradiction. □

Corollary 5.4.3. *Consider a sequence of JRF iterations* $\{M^{(k)}\}$, *where* $M_i^{(0)} = {E_i^{(0)}}^\dagger E_i^{(0)}$. *For* $k \in \mathbb{N}$, $M_i^{(k)} = {E_i^{(k)}}^\dagger E_i^{(k)}$, *where*

$$E_i^{(k+1)} = E_i^{(k)} \widetilde{\rho}_i \left(\sum_{i=1}^{N} \widetilde{\rho}_i {E_i^{(k)}}^\dagger E_i^{(k)} \widetilde{\rho}_i \right)^{-\frac{1}{2}}. \tag{5.65}$$

As we saw in Proposition 5.3.6, the sequence $\{\mathbf{E}^{(k)}\}$ *is a sequence of directional iterations for Problem (5.50). Since the set of generalized quantum measurements* S_G *is compact, as a consequence of Proposition 5.4.2, the sequence* $\{\mathbf{E}^{(k)}\}$ *converges to a fixed-point of the directional iteration. Moreover, since the mapping* $\mathbf{E} = \{E_i\}_{i\in[N]} \mapsto M = \{E_i^\dagger E_i\}_{i\in[N]}$ *is continuous, it is implied that the sequence* $\{M^{(k)}\}$ *is convergent to a POVM which is the image of* $\mathbf{E}$ *under the mapping* $\mathbf{E} = \{E_i\}_{i\in[N]} \mapsto M = \{E_i^\dagger E_i\}_{i\in[N]}$, *and* $\mathbf{E}$ *is a fixed-point of the directional iteration.*

Proposition 5.4.4. *The set of limit-points of a sequence of directional iterations for Problem (5.50) is connected.*

To prove the above proposition, let us recall a well-known result from basic analysis.

Lemma 5.4.5 ([AA70, Theorem 1])**.** *Let* X *be a compact metric space, and* $\{x_k\}$ *be a sequence of points in* X *such that*

$$\lim_{k\to\infty} d(x_k, x_{k+1}) = 0. \tag{5.66}$$

Then, the set of limit-points of the sequence $\{x_k\}$ *is connected.*

Proof of Proposition 5.4.4. The proof is a direct implication of Lemma 5.4.5, noting that from Remark 5.3.8, we have

$$\lim_{k\to\infty} \|\mathbf{E}^{(k+1)} - \mathbf{E}^{(k)}\|_G^2 = 0. \tag{5.67}$$

□

On the Fixed-points of $\mathcal{D}_{\text{JRF}}$

Let us consider the sequence $\{\mathbf{E}^{(k)}\}$ of directional iterations, discussed in Corollary 5.4.3. We showed that $\lim_{k\to\infty} \mathbf{E}^{(k)} = \mathbf{E}$, where $\mathbf{E} \in \mathcal{D}(\mathbf{E})$.

Conjecture 5.4.6. *The map* $\mathbf{F} = \{F_i\}_{i\in[N]} \mapsto \left(\sum_{i=1}^N \widetilde{\rho}_i F_i^\dagger F_i \widetilde{\rho}_i\right)^{-\frac{1}{2}}$ *is not singular at* $\mathbf{E}$.

Assume that the above conjecture is true. Then, $\mathcal{D}_{\text{JRF}}$ is continuous at $\mathbf{E}$. From $\lim_{k\to\infty} \mathbf{E}^{(k)} = \mathbf{E}$ and the assumption that the conjecture holds, it is implied that $\|\mathcal{D}_{\text{JRF}}(\mathbf{E})\|_{\mathcal{E}}^2 = \|\mathbf{E}\|_{\mathcal{E}}^2$, and then using Proposition 5.3.7, we conclude $\mathcal{D}_{\text{JRF}}(\mathbf{E}) = \mathbf{E}$.

Now, let $M = \{M_i\}_{i\in[N]}$ be the image of $\mathbf{E}$ under $\mathbf{E} = \{E_i\}_{i\in[N]} \mapsto M = \{E_i^\dagger E_i\}_{i\in[N]}$. Recall that M is the limit of the JRF sequence. Since $\mathbf{E}$ is a fixed-point of $\mathcal{D}_{\text{JRF}}$,

$$E_i = E_i\widetilde{\rho}_i\left(\sum_{i=1}^{N}\widetilde{\rho}_i E_i^\dagger E_i\widetilde{\rho}_i\right)^{-\frac{1}{2}}. \tag{5.68}$$

Therefore,

$$M_i = E_i^\dagger E_i\widetilde{\rho}_i\left(\sum_{i=1}^{N}\widetilde{\rho}_i E_i^\dagger E_i\widetilde{\rho}_i\right)^{-\frac{1}{2}} = M_i\widetilde{\rho}_i\left(\sum_{i=1}^{N}\widetilde{\rho}_i M_i\widetilde{\rho}_i\right)^{-\frac{1}{2}}. \tag{5.69}$$

Similarly, we obtain

$$M_i = \left(\sum_{i=1}^{N}\widetilde{\rho}_i M_i\widetilde{\rho}_i\right)^{-\frac{1}{2}}\widetilde{\rho}_i M_i. \tag{5.70}$$

If we have (5.69) and (5.70), we will be only one step away from proving the convergence of the JRF algorithm. Note that (5.69) can be written as

$$M_i\left(\sum_{i=1}^{N}\widetilde{\rho}_i M_i\widetilde{\rho}_i\right)^{\frac{1}{2}} = M_i\widetilde{\rho}_i. \tag{5.71}$$

Recall from Proposition 4.2.4 that M is an optimal measurement if $\left(\sum_{i=1}^{N}\widetilde{\rho}_i M_i\widetilde{\rho}_i\right)^{\frac{1}{2}}$ is Hermitian and for all $i \in [N]$, we have

$$\left(\sum_{i=1}^{N}\widetilde{\rho}_i M_i\widetilde{\rho}_i\right)^{-\frac{1}{2}} \geq \widetilde{\rho}_i. \tag{5.72}$$

Hermiticity of $\left(\sum_{i=1}^{N}\widetilde{\rho}_i M_i\widetilde{\rho}_i\right)^{\frac{1}{2}}$ is clear. If the conjecture holds, and if we prove (5.72) for all $i \in [N]$, then the convergence of the JRF algorithm will immediately be implied.

Chapter 6

Conclusion and Future Works

Conclusion In this thesis, we studied quantum guessing games, a generalisation of two well-studied problems in quantum information theory, namely quantum state discrimination and quantum state antidiscrimination. Both these special cases, and in general quantum guessing games, arise in many quantum information theoretic scenarios, which motivates their study. The present work consists of two main parts. In the first part, which includes Chapters 2, 3 and 4, we investigated properties of quantum guessing games trying to generalize the known results in the special cases and understand the similarities and differences between different quantum guessing games. A fundamental result that was previously known in special cases and was proved for a more general case in this thesis is that quantum guessing games can be reduced to quantum state discrimination problems.

Having this reduction, several known results about quantum state discrimination can be directly extended to quantum guessing games. The existence of closed-form solutions for the optimal reward of guessing games with two possible guesses, the pretty good measurement and lower bounds on its expected reward, and the derivation of optimality conditions from the SDP formulation of the problem are examples of such direct extensions that were investigated in our work. On the other hand, there are properties that can not be directly extended to quantum guessing games and differ across games, for example the notion of perfectness, which we discuss for state discrimination and antidiscrimination.

The second part, which is presented in Chapter 5, is devoted to the study of iterative algorithms for quantum guessing games, particularly for quantum state discrimination. We focused on an algorithm proposed by Ježek, Řeháček, and Fiurášek [JŘF02], for which the authors numerically observed the convergence, but a proof has not yet been found. We investigated how this algorithm can be related to other optimization algorithms for which convergence has been studied in the literature of mathematical optimization. Moreover, by analysing the fixed-points of the algorithm and one of its counterparts, we addressed a possible direction for proving the convergence, which can be a potential direction for future works.

In more detail, the author's contributions presented in this thesis are as follows.

1. In Section 2.2, we derived a standard form for quantum guessing games, such that every quantum guessing game can be reduced to this standard form.
2. Using this standard form, in Section 2.2.1, we showed how quantum guessing games can be reduced to the quantum state discrimination problem.
3. Using the reduction to QSD, in Remark 2.3.6, we obtained a closed-form solution for quantum guessing games with only two possible guesses.
4. In Section 2.4, we discussed two quantum information theoretic scenarios in which quantum guessing games naturally arise.
5. In Section 3.1, we generalized the pretty good measurement, previously introduced and studied for the QSD problem, and the pretty bad measurement, recently introduced by McIrvin, Mohan, and Sikora [MMS24] to the more general setting of quantum guessing games.
6. In Section 3.1.1, we derived two lower bounds for the reward of our generalized pretty good measurements. The first bound, which is given in Proposition 3.1.4, extends the bound proven in [MMS24] and shows that the generalized pretty good measurement is at least as good as uniform guessing. The second bound proven in Proposition 3.1.6 extends the bound proven in

[BK02] and provides an upper bound for the optimal reward in terms of the expected reward of the pretty good measurement.

7. In Remark 3.1.5, we gave an alternative proof for [MMS24, Theorem 1] in the case of pure state ensembles.
8. In Section 5.1, we generalized the iterative algorithm proposed by Ježek, Řeháček, and Fiurášek [JŘF02] to the setting of quantum guessing games. Our derivation is simpler than the one in [JŘF02], as we use the complementary slackness to define the fixed point iteration.
9. In Section 5.1, we also provide numerical results suggesting the convergence of our generalized algorithm. Moreover, we provide numerical comparisons between the runtime of the iterative algorithm and the SDP solver.
10. In Section 5.3, we provide an in-depth study of the directional iteration proposed by Tyson [Tys10], and we show how the JRF algorithm for quantum guessing games can be related to the directional iteration extending the results in [Tys10]. We then provide a Bregman minorization view of the directional iteration through which the JRF algorithm can be related to a well-known class of iterative optimization methods known as MM algorithms. Additionally, we address some hurdles in applying the known results about the global convergence of MM algorithms to the JRF algorithm.
11. In Section 5.4, we study fixed points of three maps discussed throughout Chapter 5. We show that any sequence of directional iterations when the feasible set is compact is convergent to a fixed point of the directional iteration. We then relate the fixed points of these three maps by proposing a conjecture. We study some implications of this conjecture and show that if the conjecture is proved, we will be one step away from proving the convergence of the JRF algorithm.

Future Work Some of the potential directions for future work are as follows:

1. One direction is investigating which quantum guessing games naturally arise in quantum theory, particularly in quantum information. We presented some examples in this work, but finding more examples will deepen our understanding of quantum guessing games and their similarities and differences.
2. Finding upper and lower bounds on the optimal reward of quantum guessing games that are not direct extensions of the known bound for the QSD problem is another possible direction of work. This can also be investigated for a specific game, e.g. quantum state antidiscrimination.
3. A proof of convergence for the JRF algorithm is still missing. In the last part of the thesis, we addressed a potential trajectory of steps that may lead to a proof. It is also interesting to study the connection between the iterative methods discussed in this thesis and those iterative algorithms known as the *first-order optimization methods*, as the behaviour of the two categories seems to have much in common.

Bibliography

[AA70] M. D. Ašić and D. D. Adamović. “Limit Points of Sequences in Metric Spaces”. In: *The American Mathematical Monthly* 77.6 (1970), pp. 613–616. ISSN: 00029890, 19300972. URL: http://www.jstor.org/stable/2316738.

[AKF22] A. Afham, Richard Kueng, and Chris Ferrie. “Quantum mean states are nicer than you think: fast algorithms to compute states maximizing average fidelity”. In: *arXiv preprint arXiv:2206.08183* (2022). URL: https://arxiv.org/abs/2206.08183.

[Ali91] Farid Alizadeh. “Combinatorial optimization with interior point methods and semi-definite matrices”. PhD thesis. University of Minnesota, 1991.

[And+02] Erika Andersson et al. “Minimum-error discrimination between three mirror-symmetric states”. In: *Phys. Rev. A* 65 (5 2002), p. 052308. DOI: 10.1103/PhysRevA.65.052308. URL: https://link.aps.org/doi/10.1103/PhysRevA.65.052308.

[ApS24] MOSEK ApS. *MOSEK Optimizer API for Python 10.2.3¶*. 2024. URL: https://docs.mosek.com/latest/pythonapi/index.html.

[Ban+14] Somshubhro Bandyopadhyay et al. “Conclusive exclusion of quantum states”. In: *Phys. Rev. A* 89 (2 2014), p. 022336. DOI: 10.1103/PhysRevA.89.022336. URL: https://doi.org/10.1103/PhysRevA.89.022336.

[Bar+14] Jonathan Barrett et al. “No ψ-epistemic model can fully explain the indistinguishability of quantum states”. In: *Physical review letters* 112.25 (2014), p. 250403. DOI: 10.1103/PhysRevLett.112.250403. URL: https://doi.org/10.1103/PhysRevLett.112.250403.

[BCD05] Dave Bacon, Andrew M Childs, and Wim van Dam. “From optimal measurement to efficient quantum algorithms for the hidden subgroup problem over semidirect product groups”. In: *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS’05)*. IEEE. 2005, pp. 469–478. DOI: 10.1109/SFCS.2005.38. URL: https://doi.org/10.1109/SFCS.2005.38.

[Bel75] Viacheslav P. Belavkin. “Optimal multiple quantum statistical hypothesis testing”. In: *Stochastics: An International Journal of Probability and Stochastic Processes* 1.1-4 (1975), pp. 315–345. DOI: 10.1080/17442507508833114. URL: https://doi.org/10.1080/17442507508833114.

[Ben92] Charles H. Bennett. “Quantum cryptography using any two nonorthogonal states”. In: *Phys. Rev. Lett.* 68 (21 1992), pp. 3121–3124. DOI: 10.1103/PhysRevLett.68.3121. URL: https://doi.org/10.1103/PhysRevLett.68.3121.

[Ber07] János A. Bergou. “Quantum state discrimination and selected applications”. In: *Journal of Physics: Conference Series*. Vol. 84. 1. IOP Publishing. 2007, p. 012001. DOI: 10.1088/1742-6596/84/1/012001. URL: https://doi.org/10.1088/1742-6596/84/1/012001.

[BK02] H. Barnum and E. Knill. “Reversing quantum dynamics with near-optimal quantum and classical fidelity”. In: *Journal of Mathematical Physics* 43.5 (May 2002), pp. 2097–2106. ISSN: 0022-2488. DOI: 10.1063/1.1459754. URL: https://doi.org/10.1063/1.1459754.

[BM04] V. P. Belavkin and V. P. Maslov. *Design of Optimal Dynamic Analyzers: Mathematical Aspects of Wave Pattern Recognition*. 2004. DOI: 10.48550/arXiv.quant-ph/0412031. URL: https://doi.org/10.48550/arXiv.quant-ph/0412031.

[BRT23] Shrigyan Brahmachari, Roberto Rubboli, and Marco Tomamichel. “A fixed-point algorithm for matrix projections with applications in quantum information”. In: *arXiv preprint arXiv:2312.14615* (2023). URL: https://arxiv.org/abs/2312.14615.

[Byr08] Charles Byrne. “Sequential unconstrained minimization algorithms for constrained optimization”. In: *Inverse Problems* 24.1 (2008), p. 015013. DOI: 10.1088/0266-5611/24/1/015013. URL: https://doi.org/10.1088/0266-5611/24/1/015013.

[CFS02] Carlton M. Caves, Christopher A. Fuchs, and Rüdiger Schack. “Conditions for compatibility of quantum-state assignments”. In: *Phys. Rev. A* 66 (6 2002), p. 062111. DOI: 10.1103/PhysRevA.66.062111. URL: https://doi.org/10.1103/PhysRevA.66.062111.

[Che04] Anthony Chefles. “Quantum States: Discrimination and Classical Information Transmission. A Review of Experimental Progress”. In: *Quantum State Estimation*. Ed. by Matteo Paris and Jaroslav Řeháček. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004. Chap. 12, pp. 467–511. ISBN: 978-3-540-44481-7. DOI: 10.1007/978-3-540-44481-7_12. URL: https://doi.org/10.1007/978-3-540-44481-7_12.

[Che+15] Weien Chen et al. “Minimum guesswork discrimination between quantum states”. In: *Quantum Inf. Comput.* 15.9&10 (2015), pp. 737–758. DOI: 10.26421/QIC15.9-10-2. URL: https://doi.org/10.26421/QIC15.9-10-2.

[CHT22] Claudio Carmeli, Teiko Heinosaari, and Alessandro Toigo. “Quantum guessing games with posterior information”. In: *Reports on Progress in Physics* 85.7 (2022), p. 074001. DOI: 10.1088/1361-6633/ac6f0e. URL: https://doi.org/10.1088/1361-6633/ac6f0e.

[Cro+06] Sarah Croke et al. “Maximum Confidence Quantum Measurements”. In: *Phys. Rev. Lett.* 96 (7 2006), p. 070401. DOI: 10.1103/PhysRevLett.96.070401. URL: https://doi.org/10.1103/PhysRevLett.96.070401.

[DJL00] Miloslav Dušek, Mika Jahma, and Norbert Lütkenhaus. “Unambiguous state discrimination in quantum cryptography with weak coherent states”. In: *Physical Review A* 62.2 (2000), p. 022306. DOI: 10.1103/PhysRevA.62.022306. URL: http://doi.org/10.1103/PhysRevA.62.022306.

[DT10] Matthieu E. Deconinck and Barbara M. Terhal. “Qubit state discrimination”. In: *Phys. Rev. A* 81 (6 2010), p. 062304. DOI: 10.1103/PhysRevA.81.062304. URL: https://doi.org/10.1103/PhysRevA.81.062304.

[DWA14] Vedran Dunjko, Petros Wallden, and Erika Andersson. “Quantum Digital Signatures without Quantum Memory”. In: *Phys. Rev. Lett.* 112 (4 2014), p. 040502. DOI: 10.1103/PhysRevLett.112.040502. URL: https://doi.org/10.1103/PhysRevLett.112.040502.

[EF01] Yonina C. Eldar and G. David Forney. "On quantum detection and the square-root measurement". In: *IEEE Transactions on Information Theory* 47.3 (2001), pp. 858–872. DOI: 10.1109/18.915636.

[Eld03] Yonina C. Eldar. "A semidefinite programming approach to optimal unambiguous discrimination of quantum states". In: *IEEE Transactions on information theory* 49.2 (2003), pp. 446–456. DOI: 10.1109/TIT.2002.807291. URL: https://doi.org/10.1109/TIT.2002.807291.

[EMV04] Yonina C. Eldar, Alexandre Megretski, and George C. Verghese. "Optimal Detection of Symmetric Mixed Quantum States". In: *IEEE Trans. Inf. Theory* 50.6 (2004), pp. 1198–1207. DOI: 10.1109/TIT.2004.828070. URL: https://doi.org/10.1109/TIT.2004.828070.

[FKP21] Pedro Felzenszwalb, Caroline Klivans, and Alice Paul. "Iterated linear optimization". In: *Quarterly of Applied Mathematics* 79.4 (2021), pp. 601–615. DOI: 10.1090/qam/1594. URL: https://doi.org/10.1090/qam/1594.

[Han+21] Eric P. Hanson et al. "Guesswork with quantum side information". In: *IEEE Transactions on Information Theory* 68.1 (2021), pp. 322–338. DOI: 10.1109/TIT.2021.3118878. URL: https://doi.org/10.1109/TIT.2021.3118878.

[Hau+96] Paul Hausladen et al. "Classical information capacity of a quantum channel". In: *Phys. Rev. A* 54 (3 1996), pp. 1869–1876. DOI: 10.1103/PhysRevA.54.1869. URL: https://doi.org/10.1103/PhysRevA.54.1869.

[Hay06] Masahito Hayashi. *Quantum information : an introduction*. Springer, 2006. ISBN: 978-3-540-30265-0.

[HB10] Won-Young Hwang and Joonwoo Bae. "Minimum-error state discrimination constrained by the no-signaling principle". In: *Journal of mathematical physics* 51.2 (2010). DOI: 10.1063/1.3298647. URL: https://doi.org/10.1063/1.3298647.

[HB20] Vojtěch Havlíček and Jonathan Barrett. "Simple communication complexity separation from quantum state antidistinguishability". In: *Physical Review Research* 2.1 (2020), p. 013326. DOI: 10.1103/PhysRevResearch.2.013326. URL: https://doi.org/10.1103/PhysRevResearch.2.013326.

[Hel67] Carl W. Helstrom. "Detection theory and quantum mechanics". In: *Information and Control* 10.3 (1967), pp. 254–291. DOI: 10.1016/S0019-9958(67)90302-6. URL: https://doi.org/10.1016/S0019-9958(67)90302-6.

[Hel69] Carl W. Helstrom. "Quantum detection and estimation theory". In: *Journal of Statistical Physics* 1 (1969), pp. 231–252. DOI: 10.1007/BF01007479. URL: https://doi.org/10.1007/BF01007479.

[Hel82] Carl W. Helstrom. "Bayes-cost reduction algorithm in quantum hypothesis testing". In: *IEEE Trans. Inf. Theory* 28.2 (1982), pp. 359–366. DOI: 10.1109/TIT.1982.1056470. URL: https://doi.org/10.1109/TIT.1982.1056470.

[HK18] Teiko Heinosaari and Oskari Kerppo. "Antidistinguishability of pure quantum states". In: *Journal of Physics A: Mathematical and Theoretical* 51.36 (2018), p. 365303. DOI: 10.1088/1751-8121/aad1fc. URL: https://doi.org/10.1088/1751-8121/aad1fc.

[HKK06] Masahito Hayashi, Akinori Kawachi, and Hirotada Kobayashi. "Quantum measurements for hidden subgroup problems with optimal sample complexity". In: *arXiv preprint quant-ph/0604174* (2006). DOI: 10.48550/arXiv.quant-ph/0604174. URL: https://doi.org/10.48550/arXiv.quant-ph/0604174.

[Hol72] Alexander Semenovich Holevo. "An analogue of statistical decision theory and noncommutative probability theory". In: *Trudy Moskovskogo Matematicheskogo Obshchestva* 26 (1972), pp. 133–149.

[Hol79] Alexander S. Holevo. "On asymptotically optimal hypothesis testing in quantum statistics". In: *Theory of Probability & Its Applications* 23.2 (1979), pp. 411–415. DOI: 10.1137/1123048. URL: https://doi.org/10.1137/1123048.

[HP20] Erkka Haapasalo and JUHA-PEKKA Pellonpää. "Effective methods for constructing extreme quantum observables". In: *Journal of Physics A: Mathematical and Theoretical* 53.24 (2020), p. 245301. DOI: 10.1088/1751-8121/ab8d52. URL: https://doi.org/10.1088/1751-8121/ab8d52.

[Hra+04] Zdeněk Hradil et al. "Maximum-Likelihood Methods in Quantum Mechanics". In: *Quantum State Estimation*. Ed. by Matteo Paris and Jaroslav Řeháček. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004. Chap. 3, pp. 59–112. ISBN: 978-3-540-44481-7. DOI: 10.1007/978-3-540-44481-7_3. URL: https://doi.org/10.1007/978-3-540-44481-7_3.

[Iva87] Igor D. Ivanovic. "How to differentiate between non-orthogonal states". In: *Physics Letters A* 123.6 (1987), pp. 257–259. DOI: 10.1016/0375-9601(87)90222-2. URL: https://doi.org/10.1016/0375-9601(87)90222-2.

[JŘF02] Miroslav Ježek, Jaroslav Řeháček, and Jaromír Fiurášek. "Finding optimal strategies for minimum-error quantum-state discrimination". In: *Physical Review A* 65.6 (2002), p. 060301. DOI: 10.1103/PhysRevA.65.060301. URL: https://doi.org/10.1103/PhysRevA.65.060301.

[JRS23] Nathaniel Johnston, Vincent Russo, and Jamie Sikora. "Tight bounds for antidistinguishability and circulant sets of pure quantum states". In: *arXiv preprint arXiv:2311.17047* (2023). DOI: 10.48550/arXiv.2311.17047. URL: https://doi.org/10.48550/arXiv.2311.17047.

[Kar84] Narendra Karmarkar. "A New Polynomial-Time Algorithm for Linear Programming". In: *Proceedings of the 16th Annual ACM Symposium on Theory of Computing, April 30 - May 2, 1984, Washington, DC, USA*. ACM, 1984, pp. 302–311. DOI: 10.1145/800057.808695. URL: https://doi.org/10.1145/800057.808695.

[Lan16] Kenneth Lange. *MM optimization algorithms*. SIAM, 2016. ISBN: 978-1-611-97439-3.

[LD20] Matthew Leifer and Cristhiano Duarte. "Noncontextuality inequalities from antidistinguishability". In: *Phys. Rev. A* 101 (6 2020), p. 062113. DOI: 10.1103/PhysRevA.101.062113. URL: https://doi.org/10.1103/PhysRevA.101.062113.

[Lei14] Matthew S. Leifer. "Is the Quantum State Real? An Extended Review of ψ-ontology Theorems". In: *Quanta* 3.1 (2014), pp. 67–155. DOI: 10.12743/quanta.v3i1.22. URL: http://doi.org/10.12743/quanta.v3i1.22.

[Lou22] Elena R. Loubenets. "General lower and upper bounds under minimum-error quantum state discrimination". In: *Phys. Rev. A* 105 (3 2022), p. 032410. DOI: 10.1103/PhysRevA.105.032410. URL: https://doi.org/10.1103/PhysRevA.105.032410.

[Mas94] J.L. Massey. "Guessing and entropy". In: *Proceedings of 1994 IEEE International Symposium on Information Theory*. 1994, pp. 204–. DOI: 10.1109/ISIT.1994.394764. URL: https://doi.org/10.1109/ISIT.1994.394764.

[MMS24] Caleb McIrvin, Ankith Mohan, and Jamie Sikora. "The pretty bad measurement". In: *arXiv preprint arXiv:2403.17252* (2024). URL: https://doi.org/10.48550/arXiv.2403.17252.

[MNW24] Hemant K. Mishra, Michael Nussbaum, and Mark M. Wilde. "On the optimal error exponents for classical and quantum antidistinguishability". In: *Letters in Mathematical Physics* 114.3 (2024), p. 76. DOI: 10.48550/arXiv.2309.03723. URL: https://doi.org/10.48550/arXiv.2309.03723.

[Mon07] Ashley Montanaro. "On the distinguishability of random quantum states". In: *Communications in mathematical physics* 273 (2007), pp. 619–636. DOI: 10.1007/s00220-007-0221-7. URL: https://doi.org/10.1007/s00220-007-0221-7.

[MSU23] Ankith Mohan, Jamie Sikora, and Sarvagya Upadhyay. "A generalized framework for quantum state discrimination, hybrid algorithms, and the quantum change point problem". In: *arXiv preprint arXiv:2312.04023* (2023). DOI: 10.48550/arXiv.2312.04023. URL: https://doi.org/10.48550/arXiv.2312.04023.

[NKU15] Kenji Nakahira, Kentaro Kato, and Tsuyoshi Sasaki Usuda. "Iterative methods for finding optimal quantum measurements under minimum-error and minimax criteria". In: *Phys. Rev. A* 91 (1 2015), p. 012318. DOI: 10.1103/PhysRevA.91.012318. URL: https://doi.org/10.1103/PhysRevA.91.012318.

[NN92] Yurii Nesterov and Arkadi Nemirovski. "Conic formulation of a convex programming problem and duality". In: *Optimization Methods and Software* 1.2 (1992), pp. 95–115. DOI: 10.1080/10556789208805510. URL: https://doi.org/10.1080/10556789208805510.

[NN94] Yurii E. Nesterov and Arkadi Nemirovski. *Interior-point polynomial algorithms in convex programming*. Vol. 13. Siam studises in applied mathematics. SIAM, 1994. ISBN: 978-0-89871-319-0. DOI: 10.1137/1.9781611970791. URL: https://doi.org/10.1137/1.9781611970791.

[NS06] Ashwin Nayak and Julia Salzman. "Limits on the ability of quantum states to convey classical messages". In: *J. ACM* 53.1 (Jan. 2006), 184–206. ISSN: 0004-5411. DOI: 10.1145/1120582.1120587. URL: https://doi.org/10.1145/1120582.1120587.

[NUK18] Kenji Nakahira, Tsuyoshi Sasaki Usuda, and Kentaro Kato. "Upper and lower bounds on optimal success probability of quantum state discrimination with and without inconclusive results". In: *Phys. Rev. A* 97 (1 2018), p. 012103. DOI: 10.1103/PhysRevA.97.012103. URL: https://doi.org/10.1103/PhysRevA.97.012103.

[PAGK23] Gabriel Pereira Alves, Nicolas Gigena, and J ędrzej Kaniewski. "Biased random access codes". In: *Phys. Rev. A* 108 (4 2023), p. 042608. DOI: 10.1103/PhysRevA.108.042608. URL: https://link.aps.org/doi/10.1103/PhysRevA.108.042608.

[Paw+09] Marcin Pawłowski et al. "Information causality as a physical principle". In: *Nature* 461.7267 (2009), pp. 1101–1104. DOI: 10.1038/nature08400. URL: https://doi.org/10.1038/nature08400.

[PB11] Marcin Pawłowski and Nicolas Brunner. "Semi-device-independent security of one-way quantum key distribution". In: *Phys. Rev. A* 84 (1 2011), p. 010302. DOI: 10.1103/PhysRevA.84.010302. URL: https://doi.org/10.1103/PhysRevA.84.010302.

[PBR12] Matthew F. Pusey, Jonathan Barrett, and Terry Rudolph. "On the reality of the quantum state". In: *Nature Physics* 8.6 (2012), pp. 475–478. DOI: 10.1038/nphys2309. URL: https://doi.org/10.1038/nphys2309.

[Pok24] Sebastian Pokutta. "The Frank-Wolfe algorithm: a short introduction". In: *Jahresbericht der Deutschen Mathematiker-Vereinigung* 126 (2024), pp. 3–35. DOI: 10.1365/s13291-023-00275-x.

[PW91] Asher Peres and William K. Wootters. "Optimal detection of quantum information". In: *Physical Review Letters* 66.9 (1991), p. 1119. DOI: 10.1103/physrevlett.66.1119. URL: https://doi.org/10.1103/physrevlett.66.1119.

[QL10] Daowen Qiu and Lvjun Li. "Minimum-error discrimination of quantum states: Bounds and comparisons". In: *Phys. Rev. A* 81 (4 2010), p. 042329. DOI: 10.1103/PhysRevA.81.042329. URL: https://doi.org/10.1103/PhysRevA.81.042329.

[Roc70] R. Tyrrell Rockafellar. *Convex Analysis*. Princeton Landmarks in Mathematics and Physics. Princeton University Press, 1970. ISBN: 978-1-4008-7317-3.

[Rus21] Vincent Russo. "toqito–Theory of quantum information toolkit: A Python package for studying quantum information". In: *Journal of Open Source Software* 6.61 (2021), p. 3082.

[SC23] Paul Skrzypczyk and Daniel Cavalcanti. *Semidefinite Programming in Quantum Information Science*. 2053-2563. IOP Publishing, 2023. ISBN: 978-0-7503-3343-6. DOI: 10.1088/978-0-7503-3343-6. URL: https://dx.doi.org/10.1088/978-0-7503-3343-6.

[Spe14] Dominique Spehner. "Quantum correlations and distinguishability of quantum states". In: *Journal of Mathematical Physics* 55.7 (2014). DOI: 10.1063/1.4885832. URL: https://doi.org/10.1063/1.4885832.

[SS22] Guillaume Sagnol and Maximilian Stahlberg. "PICOS: A Python interface to conic optimization solvers". In: *Journal of Open Source Software* 7.70 (Feb. 2022), p. 3915. ISSN: 2475-9066. DOI: 10.21105/joss.03915.

[Teb18] Marc Teboulle. "A simplified view of first order methods for optimization". In: *Mathematical Programming* 170.1 (2018), pp. 67–96. DOI: 10.1007/s10107-018-1284-2. URL: https://doi.org/10.1007/s10107-018-1284-2.

[Tys09a] Jon Tyson. "Estimates of non-optimality of quantum measurements and a simple iterative method for computing optimal measurements". In: *arXiv preprint arXiv:0902.0395* (2009). DOI: 10.48550/arXiv.0902.0395. URL: https://doi.org/10.48550/arXiv.0902.0395.

[Tys09b] Jon Tyson. "Two-sided estimates of minimum-error distinguishability of mixed quantum states via generalized Holevo–Curlander bounds". In: *Journal of mathematical physics* 50.3 (2009). DOI: 10.1063/1.3094322. URL: https://doi.org/10.1063/1.3094322.

[Tys10] Jon Tyson. "Two-sided bounds on minimum-error quantum measurement, on the reversibility of quantum dynamics, and on maximum overlap using directional iterates". In: *Journal of Mathematical Physics* 51.9 (Sept. 2010), p. 092204. ISSN: 0022-2488. DOI: 10.1063/1.3463451. URL: https://doi.org/10.1063/1.3463451.

[Wat18] John Watrous. *The theory of quantum information*. Cambridge university press, Mar. 2018. ISBN: 9781316848142. URL: https://doi.org/10.1017/9781316848142.

[WGM08] Wang Wengang, Dong Guohua, and Li Mingshan. "Minimum-error quantum state discrimination based on semidefinite programming". In: *2008 27th Chinese Control Conference*. 2008, pp. 521–524. DOI: 10.1109/CHICC.2008.4605546. URL: https://doi.org/10.1109/CHICC.2008.4605546.

Appendix A

Preliminaries

A.1 Semidefinite Inner Products and Seminorms

A.1.1 Semidefinite Inner Products

Definition A.1.1. Let V be a vector space over $\mathbb{C}$. A function $\langle\cdot,\cdot\rangle : V\times V \to \mathbb{C}$ is called a semidefinite inner product if it satisfies the following properties for all $\mathbf{v},\mathbf{w},\mathbf{z}\in V$ and $\alpha,\beta\in\mathbb{C}$:

1. $\langle \mathbf{v},\mathbf{w}\rangle = \overline{\langle \mathbf{w},\mathbf{v}\rangle}$,
2. $\langle \mathbf{v},\mathbf{v}\rangle \geq 0$,
3. $\langle \alpha\mathbf{v}+\beta\mathbf{w},\mathbf{z}\rangle = \alpha\langle \mathbf{v},\mathbf{z}\rangle + \beta\langle \mathbf{w},\mathbf{z}\rangle$.

Remark A.1.2. If we replace the second property in Definition A.1.1 with $\langle \mathbf{v},\mathbf{v}\rangle > 0$ for all $\mathbf{v}\neq 0$, then $\langle\cdot,\cdot\rangle$ will be an inner product.

Example A.1.3. Let $V = \mathcal{L}(\mathbb{C}^d)^L$ be the vector space of L-tuples of operators acting on $\mathbb{C}^d$. For two vectors $\mathbf{E} = \{E_k\}_{k\in[L]} \in V$ and $\mathbf{F} = \{F_k\}_{k\in[L]} \in V$, and a quantum guessing game $G = (\mathcal{E}, R_{L\times N})$, the biform $\langle\cdot,\cdot\rangle_G$ is defined as

$$\langle \mathbf{E},\mathbf{F}\rangle_G \stackrel{\text{def}}{=} \sum_{i=1}^{L}\sum_{j=1}^{N} R_{ij}\,\mathrm{tr}(E_i^\dagger F_i\widetilde{\rho}_j). \tag{A.1}$$

This biform is a semidefinite inner product on V, as:

1. $\overline{\langle \mathbf{F},\mathbf{E}\rangle_G} = \overline{\sum_{i=1}^{L}\sum_{j=1}^{N} R_{ij}\,\mathrm{tr}(F_i^\dagger E_i\widetilde{\rho}_j)} = \sum_{i=1}^{L}\sum_{j=1}^{N} R_{ij}\,\mathrm{tr}(E_i^\dagger F_i\widetilde{\rho}_j) = \langle \mathbf{E},\mathbf{F}\rangle_G$,
2. $\langle \mathbf{E},\mathbf{E}\rangle_G = \sum_{i=1}^{L}\sum_{j=1}^{N} R_{ij}\,\mathrm{tr}(E_i^\dagger E_i\widetilde{\rho}_j) \geq 0$, since $E_i^\dagger E_i \geq 0$, $\widetilde{\rho}_j \geq 0$, and $R_{ij}\geq 0$ for all i,j.
3. $\langle \alpha\mathbf{E}+\beta\mathbf{F},\mathbf{H}\rangle_G = \sum_{i=1}^{L}\sum_{j=1}^{N} R_{ij}\,\mathrm{tr}((\alpha E_i+\beta F_i)^\dagger H_i\widetilde{\rho}_j) = \alpha\langle \mathbf{E},\mathbf{H}\rangle_G + \beta\langle \mathbf{F},\mathbf{H}\rangle_G$, by linearity of the trace.

Proposition A.1.4 (Cauchy–Schwarz Inequality)**.** *Let V be a vector space over $\mathbb{C}$ equipped with a semidefinite inner product $\langle\cdot,\cdot\rangle$. Then, for all $\mathbf{v},\mathbf{w}\in V$, we have*

$$|\langle \mathbf{v},\mathbf{w}\rangle|^2 \leq \langle \mathbf{v},\mathbf{v}\rangle\langle \mathbf{w},\mathbf{w}\rangle. \tag{A.2}$$

Proof. For any $t\in\mathbb{C}$, we have $\langle \mathbf{v}+t\mathbf{w},\mathbf{v}+t\mathbf{w}\rangle \geq 0$ by the definition of a semidefinite inner product. Suppose that $\langle \mathbf{w},\mathbf{w}\rangle \neq 0$. Let $t = -\frac{\langle \mathbf{v},\mathbf{w}\rangle}{\langle \mathbf{w},\mathbf{w}\rangle}$. Then, we have

$$0 \leq \langle \mathbf{v} - \frac{\langle \mathbf{v},\mathbf{w}\rangle}{\langle \mathbf{w},\mathbf{w}\rangle}\mathbf{w}, \mathbf{v} - \frac{\langle \mathbf{v},\mathbf{w}\rangle}{\langle \mathbf{w},\mathbf{w}\rangle}\mathbf{w}\rangle \tag{A.3}$$

$$= \langle \mathbf{v},\mathbf{v}\rangle - \frac{|\langle \mathbf{v},\mathbf{w}\rangle|^2}{\langle \mathbf{w},\mathbf{w}\rangle} - \frac{|\langle \mathbf{v},\mathbf{w}\rangle|^2}{\langle \mathbf{w},\mathbf{w}\rangle} + \frac{|\langle \mathbf{v},\mathbf{w}\rangle|^2}{\langle \mathbf{w},\mathbf{w}\rangle} \tag{A.4}$$

$$= \langle \mathbf{v},\mathbf{v}\rangle - \frac{|\langle \mathbf{v},\mathbf{w}\rangle|^2}{\langle \mathbf{w},\mathbf{w}\rangle}. \tag{A.5}$$

Rearranging the above inequality, we obtain

$$|\langle \mathbf{v},\mathbf{w}\rangle|^2 \leq \langle \mathbf{v},\mathbf{v}\rangle\langle \mathbf{w},\mathbf{w}\rangle. \tag{A.6}$$

Now suppose that $\langle \mathbf{w}, \mathbf{w} \rangle = 0$. Let $t = -n\langle \mathbf{v}, \mathbf{w} \rangle$ for some $n \in \mathbb{N}$. Then, we have

$$0 \le \langle \mathbf{v} - n\langle \mathbf{v}, \mathbf{w}\rangle \mathbf{w}, \mathbf{v} - n\langle \mathbf{v}, \mathbf{w}\rangle \mathbf{w}\rangle \tag{A.7}$$
$$= \langle \mathbf{v}, \mathbf{v}\rangle - 2n|\langle \mathbf{v}, \mathbf{w}\rangle|^2. \tag{A.8}$$

Since (A.8) holds for all $n \in \mathbb{N}$, we must have $\langle \mathbf{v}, \mathbf{w} \rangle = 0$. Thus, the inequality $|\langle \mathbf{v}, \mathbf{w}\rangle|^2 \le \langle \mathbf{v}, \mathbf{v}\rangle\langle \mathbf{w}, \mathbf{w}\rangle$ holds in this case as well. □

A.1.2 Seminorms

Definition A.1.5. Let V be a vector space over $\mathbb{C}$. A function $\|\cdot\| : V \to \mathbb{R}$ is called a seminorm if it satisfies the following properties for all $\mathbf{v}, \mathbf{w} \in V$ and $\alpha \in \mathbb{C}$:

1. $\|\alpha \mathbf{v}\| = |\alpha|\|\mathbf{v}\|$, and
2. $\|\mathbf{v} + \mathbf{w}\| \le \|\mathbf{v}\| + \|\mathbf{w}\|$.

Proposition A.1.6. *Let $\langle \cdot, \cdot \rangle$ be a semidefinite inner product on a vector space V. Then, the function $\|\mathbf{v}\| \stackrel{\text{def}}{=} \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}$ is a seminorm on V.*

Proof. To prove that $\|\cdot\|$ is a seminorm, we need to show that it satisfies the above two properties. For all $\mathbf{v}, \mathbf{w} \in V$ and $\alpha \in \mathbb{C}$, we have

1. $\|\alpha \mathbf{v}\| = \sqrt{\langle \alpha\mathbf{v}, \alpha\mathbf{v}\rangle} = \sqrt{|\alpha|^2\langle \mathbf{v}, \mathbf{v}\rangle} = |\alpha|\sqrt{\langle \mathbf{v}, \mathbf{v}\rangle} = |\alpha|\|\mathbf{v}|$,
2. We have

$$\|\mathbf{v} + \mathbf{w}\|^2 = \langle \mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w}\rangle \tag{A.9}$$
$$= \langle \mathbf{v}, \mathbf{v}\rangle + \langle \mathbf{w}, \mathbf{w}\rangle + \langle \mathbf{v}, \mathbf{w}\rangle + \langle \mathbf{w}, \mathbf{v}\rangle \tag{A.10}$$
$$\le \langle \mathbf{v}, \mathbf{v}\rangle + \langle \mathbf{w}, \mathbf{w}\rangle + 2|\langle \mathbf{v}, \mathbf{w}\rangle| \tag{A.11}$$
$$\le \langle \mathbf{v}, \mathbf{v}\rangle + \langle \mathbf{w}, \mathbf{w}\rangle + 2\sqrt{\langle \mathbf{v}, \mathbf{v}\rangle\langle \mathbf{w}, \mathbf{w}\rangle} \tag{A.12}$$
$$= (\sqrt{\langle \mathbf{v}, \mathbf{v}\rangle} + \sqrt{\langle \mathbf{w}, \mathbf{w}\rangle})^2 \tag{A.13}$$
$$= (\|\mathbf{v}\| + \|\mathbf{w}\|)^2, \tag{A.14}$$

   where (A.12) follows from Proposition A.1.4. Therefore, $\|\mathbf{v} + \mathbf{w}\| \le \|\mathbf{v}\| + \|\mathbf{w}\|$.

□

A.1.3 Norms

Proposition A.1.7. *Let $G = (\mathcal{E}, R_{L\times N})$ be a quantum guessing game, such that*

1. *For all $i \in [L]$, there exists $j \in [N]$ such that $R_{i,j} > 0$,*
2. *The ensemble $\mathcal{E}$ is full-rank, i.e. $\widetilde{\rho}_j > 0$ for all $j \in [N]$.*

Then the biform $\langle \cdot, \cdot \rangle_G$, defined in Example A.1.3, is an inner product on V.

Proof. Suppose that $\langle \mathbf{E}, \mathbf{E}\rangle_G = 0$ for some $\mathbf{E} \in V$. Then, we have

$$\sum_{i=1}^{L}\sum_{j=1}^{N} R_{ij}\,\mathrm{tr}(E_i^\dagger E_i \widetilde{\rho}_j) = 0. \tag{A.15}$$

Since $R_{ij} \ge 0$, $E_i^\dagger E_i \ge 0$, and $\widetilde{\rho}_j > 0$ for all i, j, we must have $R_{ij}\,\mathrm{tr}(E_i^\dagger E_i \widetilde{\rho}_j) = 0$ for all i, j. For $i \in [L]$, let $j_i \in [N]$ be such that $R_{i,j_i} > 0$. Then, we have $\mathrm{tr}(E_i^\dagger E_i \widetilde{\rho}_{j_i}) = 0$, implying that $E_i^\dagger E_i \widetilde{\rho}_{j_i} = 0$. Since $\widetilde{\rho}_{j_i} > 0$, we must have $E_i^\dagger E_i = 0$, which implies that $E_i = 0$. Therefore, $\mathbf{E} = 0$, and $\langle \cdot, \cdot \rangle_G$ is an inner product on V. □

Corollary A.1.8. *When the premises of Proposition A.1.7 are satisfied, the seminorm $\|\cdot\|_G$ defined by $\|\mathbf{E}\|_G \stackrel{\text{def}}{=} \sqrt{\langle \mathbf{E}, \mathbf{E}\rangle_G}$ is a norm on V.*

Proposition A.1.9. *For any norm $\|\cdot\|$ on a vector space V, induced by an inner product $\langle\cdot,\cdot\rangle$, the function $\phi(\mathbf{v}) \overset{\text{def}}{=\!=} \|\mathbf{v}\|^2$ is strictly convex, i.e. for all $\mathbf{v} \neq \mathbf{w} \in V$ and $0 < \lambda < 1$, we have*

$$\phi(\lambda\mathbf{v} + (1-\lambda)\mathbf{w}) < \lambda\phi(\mathbf{v}) + (1-\lambda)\phi(\mathbf{w}). \tag{A.16}$$

Proof. For all $\mathbf{v} \neq \mathbf{w} \in V$ and $0 < \lambda < 1$, we have

$$\begin{aligned}
\phi(\lambda\mathbf{v} + (1-\lambda)\mathbf{w}) &= \|\lambda\mathbf{v} + (1-\lambda)\mathbf{w}\|^2 && \text{(A.17)}\\
&= \langle \lambda\mathbf{v} + (1-\lambda)\mathbf{w}, \lambda\mathbf{v} + (1-\lambda)\mathbf{w}\rangle && \text{(A.18)}\\
&= \lambda^2\|\mathbf{v}\|^2 + (1-\lambda)^2\|\mathbf{w}\|^2 + \lambda(1-\lambda)2\,\mathrm{Re}\langle\mathbf{v},\mathbf{w}\rangle && \text{(A.19)}\\
&\leq \lambda^2\|\mathbf{v}\|^2 + (1-\lambda)^2\|\mathbf{w}\|^2 + 2\lambda(1-\lambda)\|\mathbf{v}\|\|\mathbf{w}\| && \text{(A.20)}\\
&= (\lambda\|\mathbf{v}\| + (1-\lambda)\|\mathbf{w}\|)^2 && \text{(A.21)}\\
&< \lambda\|\mathbf{v}\|^2 + (1-\lambda)\|\mathbf{w}\|^2 && \text{(A.22)}\\
&= \lambda\phi(\mathbf{v}) + (1-\lambda)\phi(\mathbf{w}), && \text{(A.23)}
\end{aligned}$$

where (A.20) follows from the Cauchy–Schwarz inequality, and (A.22) follows from the strict convexity of the function x^2. □

A.2 Convex Analysis

The definitions provided here are taken from [Roc70]. Throughout this section, let V be an inner product space with inner product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\|$.

Definition A.2.1 (Epigraph)**.** For a function $f : \mathcal{X} \subseteq V \to [-\infty,\infty]$, the *epigraph* of f is defined as

$$\operatorname{epi} f \overset{\text{def}}{=\!=} \{(x,t) \in V \times \mathbb{R} \mid x \in \mathcal{X}, f(x) \leq t\}. \tag{A.24}$$

Definition A.2.2 (Effective Domain)**.** For a function $f : \mathcal{X} \subseteq V \to [-\infty,\infty]$, the *effective domain* of f is defined as

$$\operatorname{dom} f \overset{\text{def}}{=\!=} \{x \in \mathcal{X} \mid \exists t \in \mathbb{R}, (x,t) \in \operatorname{epi} f\}. \tag{A.25}$$

Definition A.2.3 (Properness)**.** A function $f : \mathcal{X} \subseteq V \to [-\infty,\infty]$ is called *proper* if $\operatorname{dom} f \neq \emptyset$ and for all $x \in \mathcal{X}$, $f(x) > -\infty$.

Definition A.2.4 (Lower Semi-Continuity)**.** A function $f : \mathcal{X} \subseteq V \to [-\infty,\infty]$ is called *lower semi-continuous* if for all $x \in \mathcal{X}$, we have

$$f(x) = \liminf_{y\to x} f(y). \tag{A.26}$$

Definition A.2.5 (Directional Derivative)**.** Let $f : \mathcal{X} \subseteq V \to [-\infty,\infty]$ be a function and $x \in \operatorname{dom} f$. The *one-sided directional derivative* of f at x in the direction $\mathbf{d} \in V$ is defined as

$$f'(x;\mathbf{d}) \overset{\text{def}}{=\!=} \lim_{t\downarrow 0} \frac{f(x+t\mathbf{d}) - f(x)}{t}. \tag{A.27}$$

The one-sided directional derivative $f'(x;\mathbf{d})$ is *two-sided* if $f'(x;\mathbf{d}) = -f'(x;-\mathbf{d})$ for all $\mathbf{d} \in V$.

Definition A.2.6 (Differentiability)**.** Let $f : \mathcal{X} \subseteq V \to [-\infty,\infty]$ be a function and $x \in \operatorname{dom} f$. The function f is called *differentiable* at x if there exists a (necessarily unique) vector $\nabla f(x) \in V$ such that

$$\lim_{y\to x} \frac{f(y) - f(x) - \langle \nabla f(x), y-x\rangle}{\|y-x\|} = 0. \tag{A.28}$$

Proposition A.2.7. *If f is differentiable at x, then for any $\mathbf{d} \neq 0$, the directional derivative $f'(x;\mathbf{d})$ exists and is equal to $\langle \nabla f(x), \mathbf{d}\rangle$.*

Proof. By the Differentiability of f at x, we have

$$\lim_{t\downarrow 0} \frac{f(x+t\mathbf{d}) - f(x) - \langle \nabla f(x), t\mathbf{d}\rangle}{t\|\mathbf{d}\|} = 0. \tag{A.29}$$

Thus, we have

$$\frac{f'(x;\mathbf{d}) - \langle \nabla f(x), \mathbf{d}\rangle}{\|\mathbf{d}\|} = 0, \tag{A.30}$$

which implies that $f'(x;\mathbf{d}) = \langle \nabla f(x), \mathbf{d}\rangle$. □

Example A.2.8. Let V be the vector space introduced in Example 5.3.3, equipped with the real-linear inner product $\langle \mathbf{x}, \mathbf{y}\rangle \stackrel{\text{def}}{=} \text{Re}\langle \mathbf{x}, \mathbf{y}\rangle_{\mathcal{E}}$. Let $f : V \to \mathbb{R}$ be a function defined as $f(\mathbf{x}) \stackrel{\text{def}}{=} \|\mathbf{x}\|_{\mathcal{E}}^2$. We have

$$\lim_{\mathbf{y}\to\mathbf{x}} \frac{\|\mathbf{y}\|_{\mathcal{E}}^2 - \|\mathbf{x}\|_{\mathcal{E}}^2 - \text{Re}\langle 2\mathbf{x}, \mathbf{y}-\mathbf{x}\rangle_{\mathcal{E}}}{\|\mathbf{y}-\mathbf{x}\|_{\mathcal{E}}} = \lim_{\mathbf{y}\to\mathbf{x}} \frac{\|\mathbf{y}\|_{\mathcal{E}}^2 - \|\mathbf{x}\|_{\mathcal{E}}^2 + 2\|\mathbf{x}\|_{\mathcal{E}}^2 - 2\,\text{Re}\langle \mathbf{x}, \mathbf{y}\rangle_{\mathcal{E}}}{\|\mathbf{y}-\mathbf{x}\|_{\mathcal{E}}} \tag{A.31}$$

$$= \lim_{\mathbf{y}\to\mathbf{x}} \frac{\|\mathbf{y}-\mathbf{x}\|_{\mathcal{E}}^2}{\|\mathbf{y}-\mathbf{x}\|_{\mathcal{E}}} = 0, \tag{A.32}$$

where the last equality follows from the continuity of the norm. Thus, f is differentiable at $\mathbf{x}$ with $\nabla f(\mathbf{x}) = 2\mathbf{x}$.

Definition A.2.9 (Essential Smoothness)**.** A proper convex function $f : \mathcal{X} \subseteq V \to [-\infty, \infty]$ is called *essentially smooth* if

1. $\mathbf{int}\,\text{dom}\, f \neq \emptyset$,
2. f is differentiable on $\mathbf{int}\,\text{dom}\, f$,
3. $\lim_{i\to\infty} \|\nabla f(x_i)\| = \infty$ for any sequence $\{x_i\}$ in $\mathbf{int}\,\text{dom}\, f$ converging to the boundary of $\mathbf{int}\,\text{dom}\, f$.

Definition A.2.10 (Legendre Function)**.** A function $f : V \to [-\infty, \infty]$ is called a *Legendre function* if it is proper, lsc, strictly convex, and essentially smooth.