9-2019

Sequential Discrimination Between Non-Orthogonal Quantum States

Dov L. Fields
The Graduate Center, City University of New York

Sequential Discrimination Between Non-Orthogonal Quantum States

by

Dov Fields

A dissertation submitted to the Graduate Faculty in Physics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2019

This manuscript has been read and accepted by the Graduate Faculty in Physics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Professor János Bergou

Date	Chair of Examining Committee

Professor Igor L. Kuskovsky

Date	Executive Officer

Mark Hillery

Neepa Maitra

Edgar Feldman

Christopher Gerry

Larry Liebovitch

Supervisory Committee

Abstract

Sequential Discrimination Between Non-Orthogonal Quantum States

by

Dov Fields

Adviser: János Bergou

The problem of discriminating between non-orthogonal states is one that has generated a lot of interest. This basic formalism is useful in many areas of quantum information. It serves as a fundamental basis for many quantum key distribution schemes, it functions as an integral part of other quantum algorithms, and it is useful in experimental settings where orthogonal states are not always possible to generate. Additionally, the discrimination problem reveals important fundamental properties, and is intrinsically related to entanglement. In this thesis, the focus is on exploring the problem of sequentially discriminating between non-orthogonal states. In the simplest version these schemes, Alice sends one of two known pure states to Bob who performs a non-optimal discrimination procedure such that the post measurement states resulting from his measurement can then be discriminated by a third participant, Charlie. In these schemes, the goal is to optimize the joint probability of both Bob and Charlie succeeding. In devising such a scheme, there are several different criteria that can be prioritized. The most basic scheme, referred to as Minimum Error (ME) discrimination, prioritizes Bob's and Charlie's abilities to successfully determine which state was sent by Alice. In this scheme, Bob and Charlie each set up two detectors and based on the result from the detector they make a guess as to which state was sent. For instance, if Bob registers a click in his first detector, he concludes that Alice sent the first state. As each detector has some probability to produce a result for either incoming result, Bob and Charlie optimize their joint probability of success by optimizing the probability that each detector

will fire when the correlated state is sent by Alice. Another possible scheme, referred to as Unambiguous Discrimination (UD), prioritizes Bob's and Charlie's ability to correctly determine the state sent by Alice. In this scheme, Bob and Charlie each set up three detectors, where if a result is obtained from the first two detectors Bob or Charlie can determine with certainty which state was sent by Alice. One final setup, referred to as Discrimination with a Fixed Rate of Inconclusive Outcome, is a combination of the previous two schemes, where Bob and Charlie maximize their probability of successfully determining the state sent by Alice where they allow some fixed probability that they will not be able to determine which state Alice sent. This fixed inconclusive probability allows Bob and Charlie to control how much they prioritize correctly determining the state that was sent, as in the Unambiguous Discrimination, versus prioritizing successfully determining the state sent by Alice, as in Minimum Error discrimination. One final topic that will be discussed by this thesis is Quantum Retrodiction. Quantum Retrodiction applies an alternate perspective on the communication protocol between Alice and Bob. In the predictive model, Alice calculates the probability that Bob gets a specific measurement result given that she prepares her system in a specific state. In the retrodictive model, Bob calculates the probability that Alice prepared her system in a specific state given the result of his measurement. This alternate perspective on the communication procedure gives new a new understanding and new tools for approaching the problem of state discrimination, as exemplified by applying the retrodictive formalism to unambiguous discrimination.

Acknowledgments

In writing this dissertation, I have been fortunate to have had support from a number of people. First, and foremost, I would like to profusely thank my adviser, János Bergou, for all of the mentorship I have recieved over the past 4 years. Over the past 4 years, Bergou has been a role model in his patience, care, and thoroughness that he applies to every problem that he approaches. Additionally, he constantly supported and encourage me throughout my PhD work. I would also like to appreciate Mark Hillery for acting, sometimes, as a surrogate adviser. There were countless times when I would pop into his office to ask him questions. Any time he was free, he was happy and willing to help listen to any problem I was having. Furthermore, I would like to acknowledge the remaining members of my thesis committee, Ed Feldman, Neepa Maitra, Christopher Gerry, and Larry Liebovitch. I appreciate the time they have taken out of their busy schedules for me.

I would also like to appreciate my friends and family who helped me get through the process. I would like to acknowledge the members of my cohort, Daniel Koch, Steven Muñoz, and David Ascienzo. Their presence at Hunter, their constant encouragement and support was vital to me being able to make it through the program. Finally, and most important, I owe an endless amount of thanks to my spouse, Dana Kline. I would not be who I am or have gotten to where I have without them and their constant support.

The Sequential Unambiguous Discrimination, Sequential Minimum Error, Sequential Discrimination with a Fixed Rate of Inconclusive Outcome, and Quantum Retrodiction sections

of this work (Chapters 3-6), are reproductions of work that is currently in the process of being submitted for publication [1–4].

Contents

List of Figures

Chapter 1

Quantum Measurement Theory

The problem of discriminating between quantum states is intrinsic to the field of Quantum Information. The goal of this chapter is to provide a framework for understanding the concepts behind quantum state discrimination that will be used throughout the rest of this thesis. Before discussing state discrimination directly, it is necessary to first review the basics of quantum measurement theory. Quantum measurement theory relies on a number of fundamental postulates that are based on experimental findings:

1. Any quantum system can be represented by a vector in a complex Hilbert space, i.e. $|\psi\rangle = \sum_i \alpha_i |i\rangle$. The basis vectors $|i\rangle$ represent the possible states the system will be found in after a measurement, and span the Hilbert space. The value $|\alpha_i|^2$ is the probability that the measurement will show that the system is in state $|i\rangle$.

2. Every observable can be represented as a Hermitian operator on this Hilbert space, i.e. $X = \sum_i \lambda_i |i\rangle \langle i|$, where $\lambda_i \in \mathbb{R}\ \forall i$. Measuring an observable will always result in one of the eigenvalues, λ_i, of X. The probability of measuring the eigenvalue λ_i for the state $|\psi\rangle$ can be calculated as $\langle\psi|i\rangle \langle i|\psi\rangle = |\alpha_i|^2$. In order for this measurement to be complete, the eigenvectors $|i\rangle$ of X must span the entire Hilbert space, even if the eigenvalues corresponding to some of these states are 0. The average value of

the observable X can be found as $\sum_i \lambda_i |\alpha_i|^2 = \sum_i \lambda_i \langle\psi| |i\rangle \langle i| |\psi\rangle = \langle\psi|X|\psi\rangle$ for pure quantum states.

3. After a measurement, the final state of the system is the basis vector that corresponds to the measured result. If, after measuring the observable X, one obtains the value λ_i, then the post measurement state of the quantum system is $|i\rangle$.

These basic postulates serve as a foundation for mathematically describing quantum measurements. Before using these postulates to reach a complete description of quantum measurement, it is first helpful to introduce a number of crucial concepts. The first of these relevant concepts are projectors. Projectors are an important instance of an observable. A projector, $P_j = |j\rangle \langle j|$, is an observable that only has a value when the system is in state $|j\rangle$. In other words, a projector is a basic detector that can only detect when a system is in state $|j\rangle$. The most basic type of measurement can be described as having a set of detectors P_j such that they span the entire Hilbert space, i.e. $\sum_j P_j = I$. After this measurement, if the P_j detector clicks, then the post measurement state is $|j\rangle = \frac{P_j|\psi\rangle}{\sqrt{\langle\psi|P_j|\psi\rangle}}$, which is found with probability $\langle\psi|P_j|\psi\rangle = tr\left(P_j |\psi\rangle \langle\psi|\right) = |\alpha_j|^2$.

A second relevant concept is the description of mixed states. In some cases, instead of having a single possible known quantum state, it is useful to be able to describe a system that can be in a number of different possible quantum states. For instance, a mixed state $\rho = \sum_i p_i |\psi_i\rangle \langle\psi_i|$ describes an ensemble of states $\{|\psi_i\rangle\}$ such that the state $|\psi_i\rangle$ is prepared with probability p_i. One can also perform a measurement on a mixed state. To compare with the example above, if the detector P_j clicks, then the post measurement state is $\rho_j = \frac{P_j \rho P_j}{tr(P_j \rho)}$. This post measurement state is found with probability $tr\left(P_j \rho\right)$.

A final necessary concept is partial measurement. If given a quantum state $|\psi\rangle_{ab} = |\theta\rangle_a |\varphi\rangle_b$ in the Hilbert space $H = H_a \otimes H_b$, then one can consider a measurement only on the H_b Hilbert space. For instance, if H_a is spanned by the eigenvectors $\{|i\rangle_a\}$ and H_b is spanned

by the eigenvectors $\{|j\rangle_b\}$, then one can perform a projective measurement described by the projectors $P_{kb} = |k\rangle_b \langle k|$. The resulting final state from a detection by the kth projector gives the post-measurement state $|\psi_f\rangle = \frac{P_{kb}|\psi\rangle_{ab}}{\sqrt{\langle\psi|P_{kb}|\psi\rangle}}$. Any arbitrary unitary operator U can be applied before this partial measurement, giving the final state of $|\psi_k\rangle = \frac{|k\rangle\langle k|U|\psi_{ab}\rangle}{\sqrt{\langle\psi|U^\dagger P_{kb} U|\psi\rangle}}$. Using the substitution $A_k |\theta\rangle_a = \langle k| U |\theta\rangle_a |\varphi\rangle_b$ and $\Pi_k = A_k^\dagger A_k$, the previous result can be rewritten as $|\psi_k\rangle = \frac{A_k|\theta\rangle_a|k\rangle_b}{\sqrt{tr(\Pi|\psi\rangle\langle\psi|)}}$.

These mathematical formalisms can be combined to describe any general quantum measurement that can be performed on some quantum state $|\psi\rangle$ or ρ. All quantum measurements can be represented as a Positive Operator Valued Measurement (POVM), determined by a complete set of positive operators $\{\Pi_i\}$ and the normalization condition, $\sum_i \Pi_i = I$. These positive operators can be expressed in terms of operators A_i such that $\Pi_i = A_i^\dagger A_i$. After a measurement, one of these detectors, Π_i will be detected with probability $p(i|\rho) = tr(\Pi_i \rho)$, resulting in the post measurement state $\rho_i = \frac{A_i \rho A_i^\dagger}{tr(\Pi_i \rho)}$. If the initial state is a pure state then the probability of detection is given as $\langle\psi|\Pi_i|\psi\rangle$ and the post measurement state is $|\psi_i\rangle = \frac{A_i|\psi\rangle}{\sqrt{\langle\psi|\Pi_i|\psi\rangle}}$. The conditions of positivity and normalization on the detectors Π_i ensure that the probabilities $p(i|\rho)$ fulfill all the criteria of probabilities: $p(i|\rho) \in \mathbb{R}$, $1 \geq p(i|\rho) \geq 0$, and $\sum_i p(i|\rho) = 1$. In the case that one wants to describe a system that has been measured but the measurement outcome is unknown, the post measurement state can be described as a mixed state constructed from the ensemble of the possible measurement outcomes multiplied by their probability of occurring: $\tilde{\rho} = \sum_i p(i|p)\, \rho_i = \sum_i A_i \rho A_i^\dagger$.

One important subtlety that we will heavily rely upon the intrinsic relation between the POVM formalism and partial measurements of a system. Neumark showed [5] that every POVM can be formulated in terms of entangling the original state with an ancilla and applying a partial measurement to the ancilla system. As already hinted at early, if there is a POVM represented by $\Pi = \{\Pi_i\}$ with the corresponding states $A_i^\dagger A_i = \Pi_i$, then we can define the transformation $A_k |\psi\rangle = \langle k| U_{ab} |\psi\rangle_a |\phi\rangle_b$, where $|\phi\rangle_b$ is an ancillary state appended

to the system. In general, one can express the POVM acting on the state $|\psi\rangle$ via the unitary as follows:

$$U_{ab} |\psi\rangle_a |\phi\rangle_b = \sum_i A_i |\psi\rangle |i\rangle . \tag{1.1}$$

At this point, the generalized measurement can be recovered by considering a projective measurement on the ancilla states. For more detail on Quantum Measurement Theory, one can refer to [6], [7], and [8].

Chapter 2

Quantum State Discrimination

Now equipped with the tools of Quantum Measurement theory, the problem of state discrimination can be explored. The basic setup is that Alice randomly prepares a system in a state from the set $\{\rho_i\}$ with the corresponding probability $\{\eta_i\}$ and sends that state to Bob. Bob's task is to perform a measurement that maximizes his probability of determining which state Alice prepared. If Bob's measurement is described by the POVM elements Π_i, the probability that Bob's ith POVM element will detect a result given that Alice sent the ρ_j state is: $p(i|j) = tr(\Pi_i \rho_j)$. If upon getting a click in the ith detector, Bob concludes that Alice sent ρ_i, then Bob's average probability of being correct is $\sum_i \eta_i tr(\Pi_i \rho_i)$. Bob's goal in the state discrimination problem is to optimize his average probability of success depending on any criteria he might impose on his results. Three standard formulations of this state discrimination are Minimum Error (ME), Unambiguous Discrimination (UD), and Discrimination with a Fixed Rate of Inconclusive Outcome (FRIO). The setups and optimization of these three discrimination strategies will be discussed for the case where Alice only sends Bob one out of two possible states. While the following overview is sufficient for the purposes of this dissertation, further overviews of state discrimination can be found in [9], [10], [11], and [12]. While these three discrimination strategies are fundamentally important schemes,

there are additional approaches to state discrimination not covered by this thesis, including the Maximum Confidence measurement [13–17] and the Minimum Cost measurement [18]. Additionally, the problem of Quantum State discrimination is intrinsically linked to the no-cloning theorem [19–22] and the no-broadcasting theorem [23]. While this connection is outside the scope of this dissertation, the interested reader can refer to [24] as a starting point. Finally, there are also other approaches to communication of quantum states between two parties that aren't state discrimination, including state exclusion [25], for example.

2.1 Minimum Error Discrimination

In ME, Bob's criteria for success is simply that his probability of making an error in determining which state was sent by Alice is minimized. This criteria was introduced by [26],[27], and [28]. For the case where Alice sends one of two mixed states, Bob's optimization problem can be expressed as:

$$\begin{aligned} &\text{minimize} \\ &\qquad P_e \;=\; \eta_1 tr\left(\Pi_2\rho_1\right) + \eta_2 tr\left(\Pi_1\rho_2\right), \\ &\text{subject to} \\ &\Pi_1 + \Pi_2 \;=\; I \text{ and } \Pi_i \geq 0 \text{ for } i = 1, 2. \end{aligned}$$

In this setup, Bob has two detectors, Π_1 and Π_2. When he gets a result from the Π_i detector, Bob concludes, correctly or erroneously, that Alice sent the state ρ_i. One approach to solving this problem is to substitute $\Lambda = \eta_1\rho_1 - \eta_2\rho_2$ and $\Pi_2 = I - \Pi_1$. After this substitution, one obtains two equivalent terms: $P_e = \eta_1 - tr\left(\Pi_1\Lambda\right) = \eta_2 + tr\left(\Pi_2\Lambda\right)$. Combining these terms, the resulting equation is $2P_e = 1 + tr\left(\Pi_2\Lambda\right) - tr\left(\Pi_1\Lambda\right)$. We can break up Λ into its positive and negative components: $\Lambda = \Lambda_+ + \Lambda_-$, where $\Lambda_+ \geq 0$ and $\Lambda_- < 0$. From this, the P_e term

is minimized when $tr\left(\Pi_2\Lambda_+\right) = tr\left(\Pi_1\Lambda_-\right) = 0$. Thus the solution that gives the minimum error is $P_e = \frac{1}{2}\left(1 - ||\Lambda||\right)$, where $||\Lambda|| = tr\left(\sqrt{\Lambda\Lambda^\dagger}\right)$. If Alice only sends one of two pure states, $|\psi_1\rangle$ or $|\psi_2\rangle$, then this reduces to the $P_e = \frac{1}{2}\left(1 - \sqrt{1 - 4\eta_1\eta_2|\left\langle\psi_1|\psi_2\right\rangle|^2}\right)$.

While this solution is complete, it is useful to see how the same result can be derived from the Neumark formulation of the problem [29]. Specifically focusing on the case where Alice sends pure states, Bob's measurement can be represented as entangling Alice's state with an ancilla state and applying a projective measurement to the ancilla state. Bob's measurement depends on the entangling unitary as follows:

$$U\left|\psi_1\right\rangle\left|i\right\rangle = \sqrt{p_1}\left|\varphi\right\rangle\left|1\right\rangle + \sqrt{r_1}\left|\varphi\right\rangle\left|2\right\rangle, \tag{2.1}$$

$$U\left|\psi_2\right\rangle\left|i\right\rangle = \sqrt{r_2}\left|\varphi\right\rangle\left|1\right\rangle + \sqrt{p_2}\left|\varphi\right\rangle\left|2\right\rangle. \tag{2.2}$$

After measuring the ancilla, Bob uses the result to guess which state Alice sent. If Bob measures the ancilla in state $|i\rangle$, he concludes that Alice sent state $|\psi_i\rangle$. This means that p_i is the probability that Bob succeeds given that Alice sent state $|\psi_i\rangle$, and r_i is the probability that Bob incorrectly identifies the state Alice sent given that Alice sent state $|\psi_i\rangle$. Bob's average probability of erroneously determining the state Alice sent is $P_e = \eta_1 r_1 + \eta_2 r_2$, which is subject to the constraints $\langle\psi_1|\psi_2\rangle = \sqrt{p_1 r_2} + \sqrt{r_1 p_2}$ and $p_i + r_i = 1$. These constraints come from the requirement that Bob's operation U must be unitary. One method to find the minimum probability of error is to make the substitution $\sqrt{r_i} \equiv sin\left(\theta_i\right)$. With this substitution, the constraint becomes $\langle\psi_i|\psi_2\rangle \equiv sin\left(\varphi\right) = sin\left(\theta_1 + \theta_2\right)$. This fixes $\theta_1 + \theta_2 = \varphi$ and leaves only the free variable $\Phi \equiv \theta_1 - \theta_2$. Substituting back into the original expression, one can minimize the average probability of error as follows:

$$\begin{aligned} P_e &= \eta_1 sin\left(\frac{\varphi+\Phi}{2}\right)^2 + \eta_2 sin\left(\frac{\varphi-\Phi}{2}\right)^2, \\ \frac{\partial P_e}{\partial\Phi} = 0 &= 2\eta_1 sin\left(\frac{\varphi+\Phi}{2}\right) cos\left(\frac{\varphi+\Phi}{2}\right) - 2\eta_2 sin\left(\frac{\varphi-\Phi}{2}\right) cos\left(\frac{\varphi-\Phi}{2}\right), \end{aligned}$$

$$
\begin{aligned}
&\Rightarrow \quad \eta_1 sin\left(\varphi+\Phi\right) = \eta_2 sin\left(\varphi-\Phi\right), \\
&\Rightarrow \quad tan\left(\Phi\right) = \left(\eta_2-\eta_1\right) tan\left(\varphi\right), \\
&\Rightarrow \quad r_{1,2} = sin^2\left(\frac{\varphi\pm\Phi}{2}\right) = \frac{1}{2}\left(1-cos\left(\varphi\pm\Phi\right)\right), \\
cos\left(\varphi\pm\Phi\right) &= cos\left(\varphi\right)cos\left(\Phi\right) \mp sin\left(\varphi\right)sin\left(\Phi\right) \\
&= \frac{cos\left(\varphi\right)}{\sqrt{1+\left(\eta_2-\eta_1\right)^2 tan^2\left(\varphi\right)}} \mp \frac{sin\left(\varphi\right)\left(\eta_2-\eta_1\right)tan\left(\varphi\right)}{\sqrt{1+\left(\eta_2-\eta_1\right)^2 tan^2\left(\varphi\right)}} \\
&= \frac{cos^2\left(\varphi\right) \mp sin^2\left(\varphi\right)\left(\eta_2-\eta_1\right)}{\sqrt{cos^2\left(\varphi\right)+\left(\eta_2-\eta_1\right)^2 sin^2\left(\varphi\right)}} = \frac{1-\left(\eta_1+\eta_2\pm\left(\eta_2-\eta_1\right)\right)sin^2\left(\varphi\right)}{\sqrt{1-4\eta_1\eta_2 sin^2\left(\varphi\right)}} \\
r_i &= \frac{1}{2}\left(1-\frac{1-2\left(1-\eta_i\right)\left\langle\psi_1|\psi_2\right\rangle^2}{\sqrt{1-4\eta_1\eta_2\left\langle\psi_1|\psi_2\right\rangle^2}}\right), \\
P_e &= \frac{1}{2}\left(1-\sqrt{1-4\eta_1\eta_2\left|\left\langle\psi_1|\psi_2\right\rangle\right|^2}\right).
\end{aligned} \tag{2.3}
$$

For a calculation of the optimal probability of error for various overlaps see Figure 2.1. Additionally, one helpful visualization of the Minimum Error Discrimination scheme is depicted in Figure 2.2, which is a reproduction of Figure 11.6 in [9]. For other applications and extensions of the minimum error discrimination scheme outside of the scope of this thesis, including minimum error discrimination between more than two states, refer to [30–34].

2.2 Unambiguous State Discrimination

In UD, Bob's measurement criteria is that he never incorrectly identifies the state sent by Alice. This criteria was first introduced and explored by Ivanovic, Dieks, and Peres, in [35], [36], and [37]. It was then later expanded upon by Jaeger and Shimony in [38]. In this setup, Bob has three detectors: Π_1, Π_2, Π_0. When either the Π_1 or Π_2 detectors click, Bob concludes with certainty that Alice sent the corresponding state. When Bob's third detector, Π_0, clicks, Bob's measurement result is inconclusive and he is unable to determine the state sent by

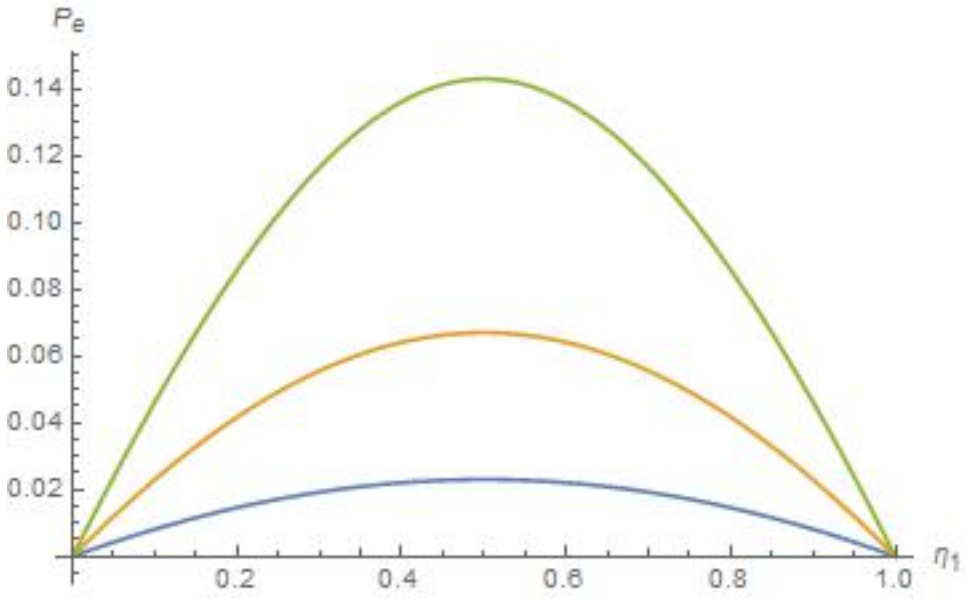

Figure 2.1: Optimal probability of error versus the prior probability for the overlaps $\langle\psi_1|\psi_2\rangle = 0.3, 0.5, 0.7$ in blue, orange, and green respectively.

Alice. While some recent work has explored the possibility of unambiguously discriminating between mixed states [39, 40], this thesis will focus solely on applying UD to pure states. In this case, Bob's optimization problem can be described as follows:

$$\begin{aligned} &\text{maximize} \\ &\qquad P_s &&= \eta_1 \langle\psi_1|\Pi_1|\psi_1\rangle + \eta_2 \langle\psi_2|\Pi_2|\psi_2\rangle , \\ &\text{subject to} \\ \Pi_1 + \Pi_2 + \Pi_0 &= I \text{ and } \Pi_i \geq 0 \text{ for } i = 1, 2, \\ \langle\psi_1|\Pi_2|\psi_1\rangle &= \langle\psi_2|\Pi_1|\psi_2\rangle = 0. \end{aligned}$$

Using the Neumark formulation, Bob's measurement can be written as follows:

$$U|\psi_1\rangle|i\rangle = \sqrt{p_1}|\varphi_1\rangle|1\rangle + \sqrt{q_1}|\varphi_2\rangle|0\rangle , \tag{2.4}$$

$$U|\psi_2\rangle|i\rangle = \sqrt{p_2}|\varphi_3\rangle|2\rangle + \sqrt{q_2}|\varphi_4\rangle|0\rangle . \tag{2.5}$$

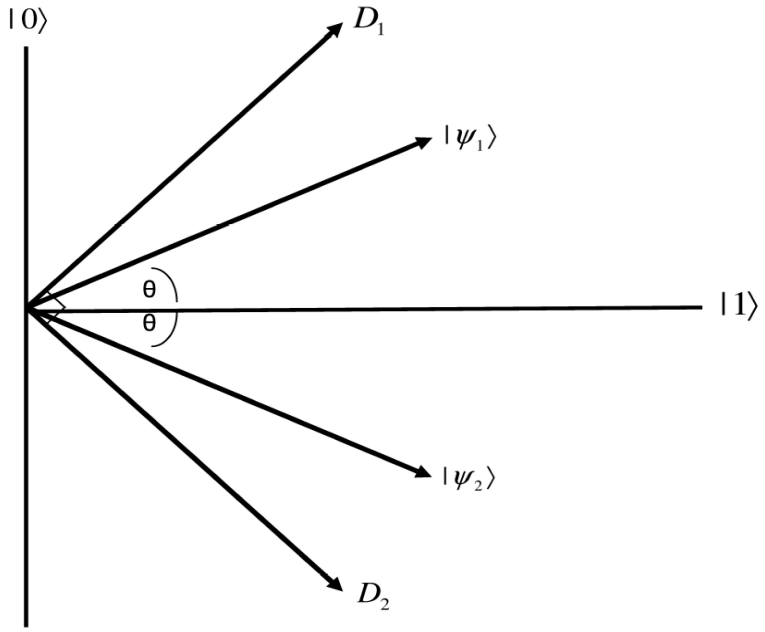

Figure 2.2: Graphical representation of the Minimum Error Discrimination procedure. $|\psi_1\rangle$ and $|\psi_2\rangle$ are projected onto a 2D real Hilbert space and oriented symmetrically around $|1\rangle$. The detection operators Π_i are represented by the projective detection operators D_i in this 2D Hilbert space. The relative angle between D_i and $|\psi_j\rangle$ affects the probability of detecting the state. For equal priors, the optimal solution corresponds to orienting D_1 and D_2 symmetrically around $|\psi_1\rangle$ and $|\psi_2\rangle$.

Here, $p_i \equiv \langle\psi_i|\Pi_i|\psi_i\rangle$ and $q_i \equiv \langle\psi_i|\Pi_0|\psi_i\rangle$. Both of these formulations lead to the constraint: $q_1 q_2 \geq |\langle\psi_1|\psi_2\rangle|^2 \equiv s^2$. In the case of the POVM formulation this constraint comes from the positivity of Π_0, and in the case of the Neumark formulation this constraint comes from the unitarity of Bob's operation. Substituting the saturated constraint $q_1 q_2 = s^2$ into the original problem, the optimal solution can be found:

$$\begin{aligned}
P_s &= 1 - \eta_1 q_1 - \eta_2 q_2 = 1 - \eta_1 q_1 - \eta_2 \frac{s^2}{q_1}, \\
\frac{\partial P_s}{\partial q_1} &= 0 = -\eta_1 + \eta_2 \frac{s^2}{q_1^2}, \\
q_i &= \sqrt{\frac{1-\eta_i}{\eta_i}}\, s, \\
P_s &= 1 - 2\sqrt{\eta_1 \eta_2}\, s.
\end{aligned}$$

One important note is that this solution is only valid for $q_1, q_2 < 1$. This means that the full solution becomes:

$$P_s = \begin{cases} \eta_1 \left(1 - s^2\right) & \eta_1 > \frac{1}{1+s} \\ 1 - 2\sqrt{\eta_1 \eta_2} s & \frac{1}{1+s} \geq \eta_1 \geq \frac{s}{1+s} \\ \eta_2 \left(1 - s^2\right) & \frac{s}{1+s} \geq \eta_1 \end{cases} \tag{2.6}$$

For a calculation of the optimal probability of success, see Figure 2.3. Additionally, a

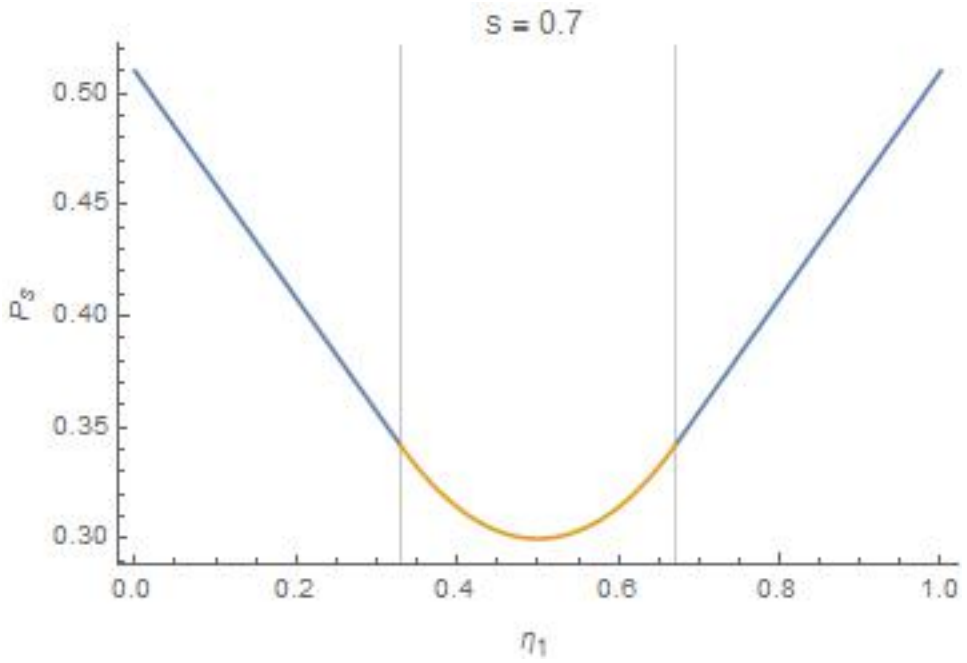

Figure 2.3: Optimal probability of success versus prior probability for $s = 0.7$. The vertical lines indicate the different regions of the solution.

useful graphical depiction of the UD measurement is depicted in Figure 2.4, which is a reproduction of Figure 11.3 in [9]. The Unambiguous Discrimination procedure has been realized experimentally in a number of different physical systems [41–43].

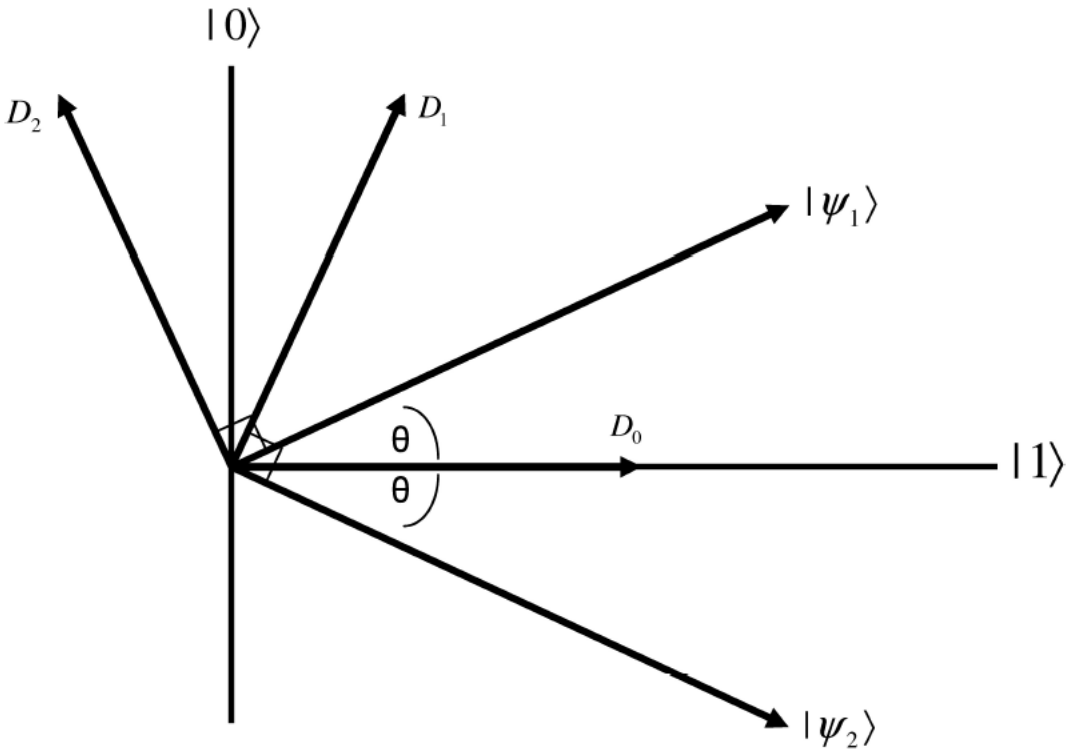

Figure 2.4: Graphical representation of the Unambiguous Discrimination procedure. $|\psi_1\rangle$ and $|\psi_2\rangle$ are projected onto a 2D real Hilbert space and oriented symmetrically around $|1\rangle$. The detection operators Π_i are represented by the detection operators D_i in this 2D Hilbert space. The right angle between D_1 and $|\psi_2\rangle$ and D_2 and $|\psi_1\rangle$ ensure that these detectors are unambiguous. D_0 is symmetrically placed between $|\psi_1\rangle$ and $|\psi_2\rangle$, ensuring that a result from this detector is inconclusive. It is important to note that the actually detection operation does not correspond to a simple projection along the vector representing the detection operators as it would for a projective measurement.

2.3 Discrimination with a Fixed Rate of Inconclusive Outcome

One final discrimination strategy, referred to as Discrimination with a Fixed Rate of Inconclusive Outcome (FRIO), connects the previous two strategies. In this strategy, introduced by [44], Bob again has three detectors, Π_1, Π_2, Π_0. When getting a click in either the Π_1 or Π_2 detectors, Bob guesses that the associated state was sent by Alice. If Bob, instead, gets a

click from the Π_0 detector, Bob treats the measurement as a failure and the state Alice sent remains inconclusive. Instead of requiring the Π_1 and Π_2 to be error free, Bob instead simply optimizes the probability that he succeeds while fixing the probability that the inconclusive channel, Π_0 gives a result. In other words, Bob's optimization problem can be described as follows:

$$\begin{aligned} \text{maximize} & \\ P_s &= \eta_1 \langle\psi_1|\Pi_1|\psi_1\rangle + \eta_2 \langle\psi_2|\Pi_2|\psi_2\rangle , \\ \text{subject to} & \\ \Pi_1 + \Pi_2 + \Pi_0 &= I \text{ and } \Pi_i \geq 0 \text{ for } i = 1, 2, \\ Q &= \eta_1 \langle\psi_1|\Pi_0|\psi_1\rangle + \eta_2 \langle\psi_2|\Pi_0|\psi_2\rangle = \text{fixed}. \end{aligned}$$

Here, Q is the average rate of inconclusive outcomes that is fixed by Bob. This parameter Q allows parameterization between the minimum error and unambiguous discrimination problems. When $Q = 0$ this setup is equivalent to the ME strategy, and when $Q = 2\sqrt{\eta_1\eta_2}s$, this setup is equivalent to the UD strategy. Using the Neumark formulation, Bob's measurement can be written as follows:

$$U|\psi_1\rangle|i\rangle = \sqrt{p_1}|\varphi\rangle|1\rangle + \sqrt{r_1}|\varphi\rangle|2\rangle + \sqrt{q_1}|\varphi\rangle|0\rangle \tag{2.7}$$

$$U|\psi_2\rangle|i\rangle = \sqrt{p_2}|\varphi\rangle|2\rangle + \sqrt{r_2}|\varphi\rangle|1\rangle + \sqrt{q_2}|\varphi\rangle|0\rangle \tag{2.8}$$

Here, p_i is Bob's probability of succeeding given that state i was sent, r_i is Bob's probability of making an error, and q_i is Bob's probability that his measurement will fail given when Alice sends state i. One strategy for optimizing this problem is to rewrite it in terms of a minimum error problem. From the positive operator formulation, one can rewrite the

problem as follows:

$$\begin{aligned}
\Omega &\equiv I - \Pi_0 = \Pi_1 + \Pi_2, \\
\tilde{\Pi}_i &\equiv \Omega^{-\frac{1}{2}} \Pi_i \Omega^{-\frac{1}{2}} \quad \Rightarrow \tilde{\Pi}_1 + \tilde{\Pi}_2 = I, \\
|\tilde{\psi}_i\rangle &\equiv \frac{\Omega^{\frac{1}{2}} |\psi_i\rangle}{\sqrt{\langle\psi_i|\Omega|\psi_i\rangle}}, \quad \tilde{\eta}_i \equiv \frac{\eta_i \langle\psi_i|\Omega|\psi_i\rangle}{\eta_1 \langle\psi_1|\Omega|\psi_1\rangle + \eta_2 \langle\psi_2|\Omega|\psi_2\rangle}, \\
\eta_i \langle\psi_i|\Pi_i|\psi_i\rangle &= (1-Q)\tilde{\eta}_i \langle\tilde{\psi}_i|\tilde{\Pi}_i|\tilde{\psi}_i\rangle , \\
P_s &= (1-Q)\, \tilde{P}_s = (1-Q) \sum_{i=1}^{2} \tilde{\eta}_i \langle\tilde{\psi}_i|\tilde{\Pi}_i|\tilde{\psi}_i\rangle .
\end{aligned}$$

In this final step, $\tilde{P}_s$ is simply the probability of success for ME discrimination between $\left\{|\tilde{\psi}_i\rangle\right\}$, with priors $\{\tilde{\eta}_i\}$, and with detectors $\tilde{\Pi}_i$. As the optimal solution to this problem is already known, this can be plugged in to give the following:

$$\begin{aligned}
P_s &= \frac{(1-Q)}{2}\left(1 + \sqrt{1 - 4\tilde{\eta}_1\tilde{\eta}_2 |\langle\tilde{\psi}_1|\tilde{\psi}_2\rangle|^2}\right), \\
&= \frac{(1-Q)}{2}\left(1 + \sqrt{1 - 4\frac{\eta_1\eta_2}{(1-Q)^2} |\langle\psi_1|\Omega|\psi_2\rangle|^2}\right), \\
&= \frac{(1-Q)}{2}\left(1 + \sqrt{1 - 4\frac{\eta_1\eta_2}{(1-Q)^2} (|\langle\psi_1|\psi_2\rangle| - \sqrt{q_1 q_2})^2}\right), \\
P_s &= \frac{1}{2}\left(\bar{Q} + \sqrt{\bar{Q}^2 - (2\sqrt{\eta_1\eta_2}|\langle\psi_1|\psi_2\rangle| - Q)^2}\right). \qquad (2.9)
\end{aligned}$$

In the second to last line of the derivation, it is assumed, without any loss of generality, that both $|\psi_1\rangle$ and $|\psi_2\rangle$ lie in a real plane. Along with the definition of $q_i \equiv \langle\psi_i|\Pi_0|\psi_i\rangle$, this allows the reduction $\langle\psi_1|\Pi_0|\psi_2\rangle = \sqrt{q_1 q_2}$ used to derive the second to last line of the previous equations. This equation is then optimized with respect to the constraint $Q = \eta_1 q_1 + \eta_2 q_2$, giving $\eta_1 q_1 = \eta_2 q_2 = \frac{Q}{2}$ and resulting in the final equation. For the last line, the variable $\bar{Q} \equiv 1 - Q$ is introduced to simplify some of the notation. This final equation gives the optimal probability of success for a fixed rate of inconclusive results. For a numeric calculation of

this result, see Figure 2.5.

It is worth noting that this solution can equivalently be achieved from the Neumark formalism. The Neumark formulation gives the obvious constraints:

$$\begin{aligned} s &\equiv \langle\psi_1|\psi_2\rangle = \sqrt{p_1 r_2} + \sqrt{p_2 r_1} + \sqrt{q_1 q_2}, \\ 1 &= p_i + r_i + q_i \quad i = 1, 2. \end{aligned}$$

By making the substitutions $\tilde{p}_i \equiv \frac{p_i}{1-q_i}$ and $\tilde{r}_i \equiv \frac{r_i}{1-q_i}$, one can rewrite the above constraints:

$$\begin{aligned} \tilde{s} &\equiv \frac{\langle\psi_1|\psi_2\rangle - \sqrt{q_1 q_2}}{\sqrt{1-q_1}\sqrt{1-q_2}} = \sqrt{\tilde{p}_1 \tilde{r}_2} + \sqrt{\tilde{p}_2 \tilde{r}_1}, \\ 1 &= \tilde{p}_i + \tilde{r}_i \quad i = 1, 2. \end{aligned}$$

At this point, the problem is again effectively rewritten in terms of an ME discrimination problem with states that have an effective overlap of $\tilde{s}$. Using this, one can derive the same optimal result given in 2.3 for maximizing the probability of success.

As in the case of the UD discrimination, this solution is only valid under certain conditions. The boundaries of calculating a valid solution can be more readily derived from the Neumark formulation of the problem. Clearly, $\tilde{s} \geq 0$, which requires $s \geq \sqrt{q_1 q_2}$ and therefore $Q_0 \equiv 2\sqrt{\eta_1 \eta_2} s \geq Q$ in order for the above solution to be valid. An additional restriction on this solution can be derived by revisiting the initial constraint $s = \sqrt{p_1 r_2} + \sqrt{p_2 r_1} + \sqrt{q_1 q_2}$. This constraint can be rewritten by redefining the parameters p_i,r_i, and q_i in terms of the angles θ_i and φ_i as follows:

$$p_i \equiv cos^2(\theta_i)\, cos^2(\varphi_i), \quad r_i \equiv cos^2(\theta_i)\, sin^2(\varphi_i), \quad q_i \equiv sin^2(\theta_i).$$

Using this, the constraint can be rewritten as:

$$s = cos\left(\theta_1\right) cos\left(\theta_2\right) sin\left(\varphi_1 + \varphi_2\right) + sin\left(\theta_1\right) sin\left(\theta_2\right). \tag{2.10}$$

Given that $1 \geq p_i \geq 0$ and $1 \geq r_i \geq 0$ gives that $\frac{\pi}{2} \geq \varphi_i \geq 0$, and thus that $1 \geq sin\left(\varphi_1 + \varphi_2\right) \geq 0$. This gives the constraint that $cos\left(\theta_1 - \theta_2\right) \geq s \geq sin\left(\theta_1\right) sin\left(\theta_2\right)$. The right hand inequality was used above, but, using the substitution $\theta \equiv \theta_1 - \theta_2$, the left hand inequality gives an additional constraint:

$$s \leq cos\left(\theta\right). \tag{2.11}$$

Using the result that $\eta_1 q_1 = \eta_2 q_2$ in the optimal solution, this results in the following expression for Q:

$$\begin{aligned}
Q &= \eta_1 cos^2\left(\theta_1\right) + \eta_2 cos^2\left(\theta_2\right) = 2\sqrt{\eta_1\eta_2} cos\left(\theta_1\right) cos\left(\theta_2\right) \\
&= 2\sqrt{\eta_1\eta_2} cos\left(\theta_1\right) cos\left(\theta_2\right) \frac{\frac{cos(\theta_1) sin^2(\theta_2)}{cos(\theta_2)} + \frac{cos(\theta_2) sin^2(\theta_1)}{cos(\theta_1)} - 2sin\left(\theta_1\right) sin\left(\theta_2\right)}{\frac{cos(\theta_1) sin^2(\theta_2)}{cos(\theta_2)} + \frac{cos(\theta_2) sin^2(\theta_1)}{cos(\theta_1)} - 2sin\left(\theta_1\right) sin\left(\theta_2\right)} \\
&= 2\sqrt{\eta_1\eta_2} \frac{cos^2\left(\theta_1\right) sin^2\left(\theta_2\right) + cos^2\left(\theta_2\right) sin^2\left(\theta_1\right) - 2sin\left(\theta_1\right) sin\left(\theta_2\right) cos\left(\theta_1\right) cos\left(\theta_2\right)}{\frac{\sqrt{\eta_1}}{\sqrt{\eta_2}} sin^2\left(\theta_2\right) + \frac{\sqrt{\eta_2}}{\sqrt{\eta_1}} sin^2\left(\theta_1\right) - 2sin\left(\theta_1\right) sin\left(\theta_2\right)} \\
&= 2\eta_1\eta_2 \frac{\left(sin\left(\theta_1\right) cos\left(\theta_2\right) + cos\left(\theta_1\right) sin\left(\theta_2\right)\right)^2}{\eta_2 sin^2\left(\theta_2\right) + \eta_1 sin^2\left(\theta_1\right) - 2\sqrt{\eta_1\eta_2}\left(cos\left(\theta\right) - cos\left(\theta_1\right) cos\left(\theta_2\right)\right)}, \\
Q &= \frac{2\eta_1\eta_2 sin^2\left(\theta\right)}{1 - 2\sqrt{\eta_1\eta_2} cos\left(\theta\right)}.
\end{aligned}$$

The constraint $s \leq cos\left(\theta\right)$ is saturated when $sin\left(\varphi_i + \varphi_2\right) = 1$. This occurs when one incoming states are ignored, for instance, when $p_1 = 0$ and $r_2 = 0$. In this regime, $Q_{th} = \frac{2\eta_1\eta_2\left(1-s^2\right)}{1-2\sqrt{\eta_1\eta_2}s}$. For $Q \leq Q_{th}$, there exists a solution such that both the condition that $\eta_1 q_1 = \eta_2 q_2$ and the condition that $s \leq cos\left(\theta\right)$ hold. To see this graphically, see Figure 2.6. For $Q > Q_{th}$ this solution is no longer valid.

These two boundaries give two regions for consideration where the optimal solution calculated above is invalid: $2\sqrt{\eta_1\eta_2}s \geq Q$ and $Q_{th} = \frac{2\eta_1\eta_2\left(1-s^2\right)}{1-2\sqrt{\eta_1\eta_2}s} \leq Q$. For a graphical comparison of these two boundaries, Q_{th} and $2\sqrt{\eta_1\eta_2}s$ see Figure 2.7. It is important to note that $Q_{th} \geq 2\sqrt{\eta_1\eta_2}s$ when $\frac{1}{1+s^2} \geq \eta_1 \geq \frac{s^2}{1+s^2}$. Unsurprisingly, this region matches the region for which the UD solution is also valid. At this point we can construct the full solution. For $Q > 2\sqrt{\eta_1\eta_2}s$, the constraint $s \geq \sqrt{\eta_1\eta_2}$ must be saturated, forcing $r_1 = r_2 = 0$. Solving for $Q = \eta_1 q_1 + \eta_2 q_2$ and $s = \sqrt{q_1 q_2}$ gives $q_i = \frac{Q \pm \sqrt{Q - 4\eta_1\eta_2 s}}{2\eta_i}$, and $P_s = (1 - Q)$. For $2\sqrt{\eta_1\eta_2}s \geq Q \geq Q_{th}$, the inequality $s \leq cos\,(\theta)$ is saturated, requiring $p_{1(2)} = r_{2(1)} = 0$ for $\eta_2 > \eta_1 (\eta_1 > \eta_1)$. Assuming $\eta_2 > \eta_1$ and starting from the resulting constraint $s = cos\,(\theta) = cos\,(\theta_1 - \theta_2)$ and using the facts that in this case $P_s = \eta_2 cos\,(\theta_2)$, $P_e = \eta_1 cos\,(\theta_1)$, and $Q = 1 - P_e - P_s$, gives the following form for Q:

$$\begin{aligned} Q &= 1 - P_s - \eta_1 cos^2\,(\theta_2 + \theta) \\ &= 1 - P_s - \eta_1 \left(cos\,(\theta_2)\, cos\,(\theta) - sin\,(\theta_2)\, sin\,(\theta)\right)^2 \\ Q &= 1 - P_s - \eta_1 \left(\sqrt{\frac{P_s}{\eta_2}}s - \sqrt{1 - \frac{P_s}{\eta_2}}\sqrt{1 - s^2}\right)^2 . \end{aligned} \tag{2.12}$$

While this equation can be inverted to obtain an expression for P_s as a function of Q, the critical piece of this solution is that it matches up with the two boundary solutions. At $P_e = 0$, $P_s = \eta_2\,(1 - s^2)$, and in this expression $Q = \eta_1 + \eta_2 s^2$, the expected value for the UD boundary solution. For $Q = Q_{th}$, one can show that P_s is equal to the form given by (2.9).

2.4 Sequential Discrimination Strategies

Having reviewed three fundamental strategies for discriminating between quantum states, the question is raised of how one might include a third party, Charlie. The easiest way to

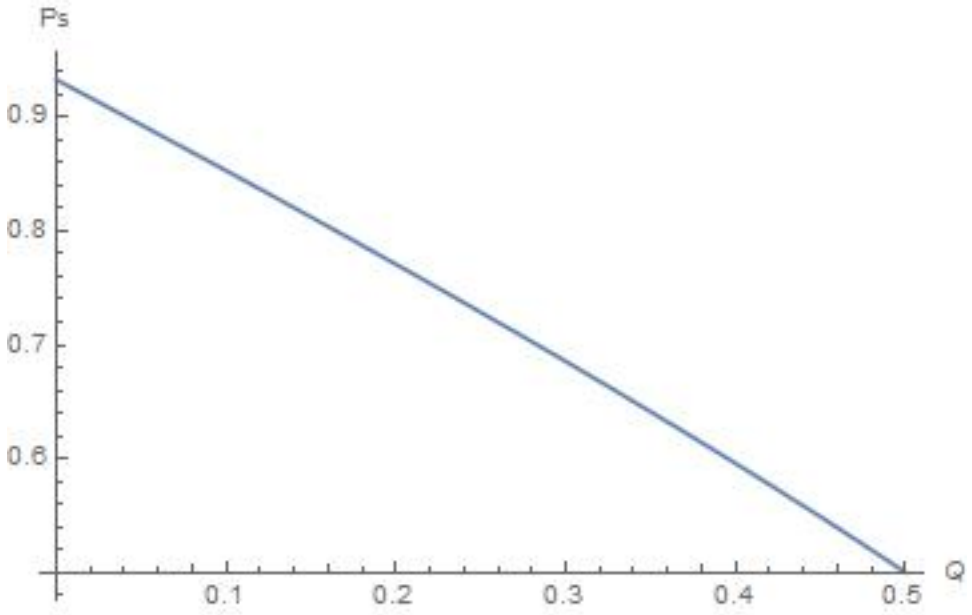

Figure 2.5: P_s versus Q for $\eta_1 = \eta_2 = s = 0.5$, plotted up to $Q = Q_0 \equiv 2\sqrt{\eta_1 \eta_2} s$

include Charlie is for Alice to make 2 copies of her state and send them independently to both Bob and Charlie. However, a more interesting and secure scheme is for Alice to send states to Bob. After performing his own measurement on the state sent by Alice, Bob then sends the resulting state to Charlie. Finally, Charlie then measures the state sent by Bob, and can also come to some conclusion about the state initially sent by Alice. One concern for any sequential scheme is the collapse postulate [45] - that after Bob's measurement the state that Bob measures should collapse leaving nothing useful for Charlie to measure. However, as we shall see, by using the generalized measurements (POVMs) rather than a projective measurement, Bob can avoid this concern. The basic setup for a sequential discrimination procedure is, as before, for Alice to start with random distribution of states $\{\rho_i\}$ that she will choose with the corresponding probability $\{\eta_i\}$. Alice will randomly pick one of the states and send it to Bob, who will apply a measurement described by the POVM elements Π_{bi}. The probability that Bob's ith POVM element will click if Alice sends the state ρ_j is $p_b(i|j) = tr(\Pi_{bi}\rho_j)$. If $\Pi_{bi} \equiv B_i^\dagger B_i$, then after Bob's measurement, the resulting state is

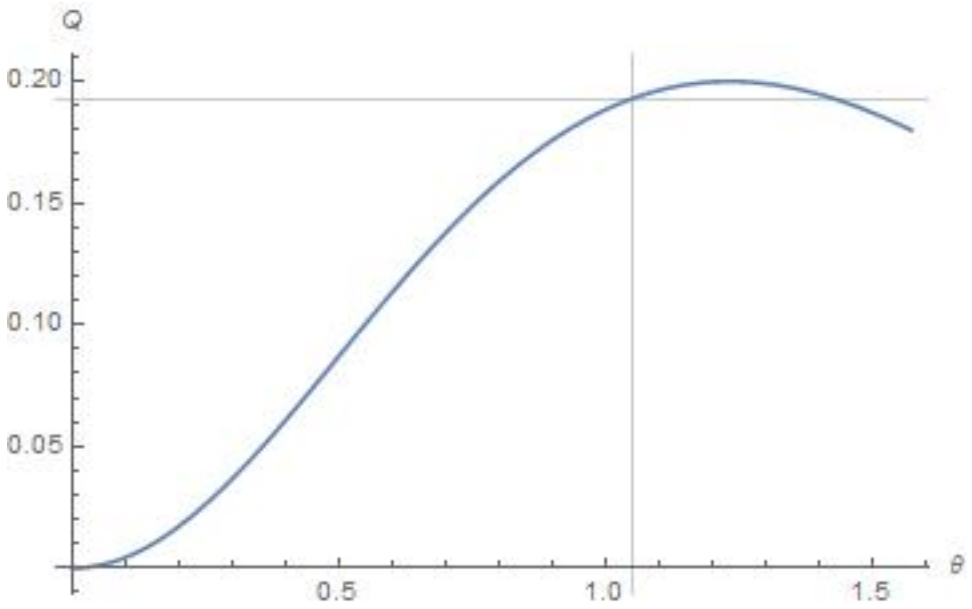

Figure 2.6: $Q = \frac{2\eta_1\eta_2 sin^2(\theta)}{1-2\sqrt{\eta_1\eta_2}cos(\theta)}$ versus θ for $\eta_1 = \eta_2 = s = 0.5$. The grid lines shown intersect at Q_{th}

.

$\theta_{ij} = \frac{B_i\rho_j B_i^\dagger}{tr(\Pi_{bi}\rho_j)}$. Charlie can then perform a measurement described by the POVM elements Π_{ci} on these resulting states. The probability that Charlie's ith POVM element will click if Alice sends the sate ρ_i is $p_c(i|j) = \sum_k tr(\Pi_{ci}\theta_{kj})$. One useful simplification is to note that one can define $\theta_j \equiv \sum_i \theta_{ij}$. From this perspective, one can see θ_j as the post-measurement states of Bob's measurements when Alice sends the corresponding ρ_j, and, therefore, Charlie's goal is to optimally discriminate between the states $\{\theta_j\}$. If upon getting a result in the ith detector, both Bob and Charlie guess that Alice sent ρ_i, then the joint probability that both Bob and Charlie successfully determined the state sent by Alice is $P_{ss} = \sum_i \eta_i p_b(i|i)\, p_c(i|i)$. While this formulation can be considered in general, the focus of this thesis is to explore the extensions to the fundamental strategies discussed earlier. The three sequential strategies that will be discussed are the Sequential Minimum Error (SME), Sequential Unambiguous Discrimination (SUD), and Sequential Discrimination with a Fixed Rate of Inconclusive Outcome (SFRIO) strategies. Additionally, while these strategies can be considered more

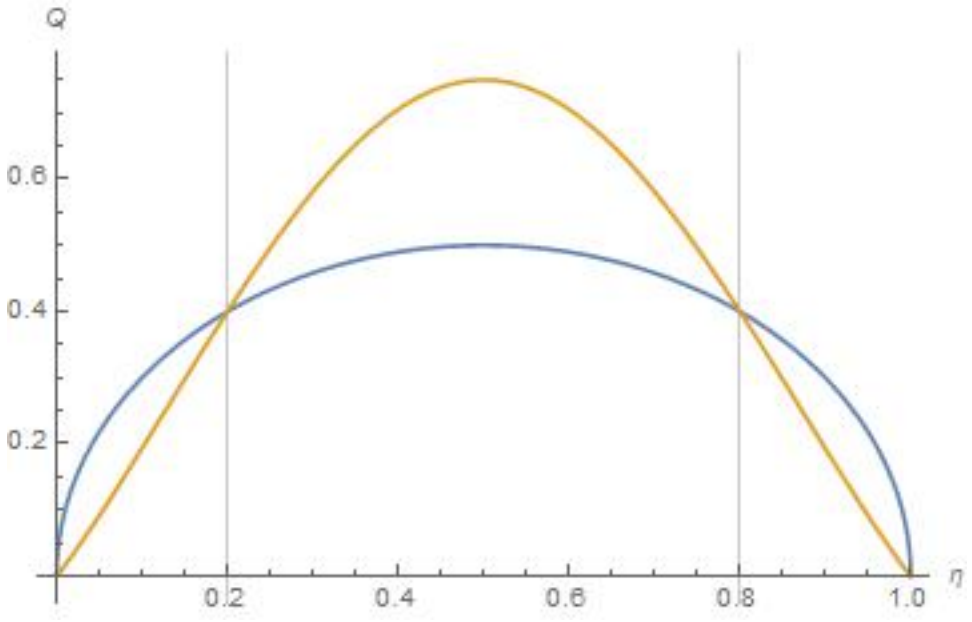

Figure 2.7: Q vs η_1 for $Q = 2\sqrt{\eta_1\eta_2}s$ (Blue) and $Q = \frac{2\eta_1\eta_2\left(1-s^2\right)}{1-2\sqrt{\eta_1\eta_2}s}$ (Yellow), for $s = 0.5$. The grid lines show the boundary values of $\eta_1 = \frac{s^2}{1+s^2}$ and $\eta_1 = \frac{1}{1+s^2}$.

broadly, the focus of this thesis will be restricted to considering only the formulations of these strategies in which both Bob and Charlie are discriminating between two pure states. For more details on how multiple observers can scavenge information from a quantum system, refer to [46].

Chapter 3

Sequential Minimum Error Discrimination

For the problem of SME discrimination between two non-orthogonal pure states, Alice sends the states $|\psi_1\rangle$ or $|\psi_2\rangle$ with probability η_1 and η_2, respectively. These states then get passed sequentially through a number of receivers, each of whom performs their own POVM on the states they receive from their predecessor, and then passes their post measurement states along to the next link in the chain. The goal in this problem is to maximize the average probability that all of them succeed. To see how this sequence can be set up, it helps to start with the case of only two receivers, Bob and Charlie. In this situation, Bob's POVM takes the following form:

$$\sum_{i=1}^{2} \Pi_{bi} = I, \tag{3.1}$$

$$\Pi_{bi} \geq 0 \quad \text{for} \quad i = 1, 2\ . \tag{3.2}$$

Then $\langle\psi_i|\Pi_{bi}|\psi_i\rangle = p_{bi}$ is Bob's probability of correctly determining that state i was sent and $\langle\psi_i|\Pi_j|\psi_i\rangle = r_{bi}$ for $i \neq j$ is Bob's probability of making an erroneous identification. Because

the POVM elements, Π_{bi}, are positive operators, we can write them in the form $\Pi_{bi} = B_i^\dagger B_i$. The detection operators, B_i determine the effect of Bob's POVM on the input states:

$$B_1 = \beta_{11} |v_{11}\rangle \langle \psi_2^\perp| + \beta_{12} |v_{12}\rangle \langle \psi_1^\perp| , \tag{3.3}$$

$$B_2 = \beta_{21} |v_{21}\rangle \langle \psi_2^\perp| + \beta_{22} |v_{22}\rangle \langle \psi_1^\perp| . \tag{3.4}$$

Here $|\psi_i^\perp\rangle$ is the state orthogonal to $|\psi_i\rangle$, $\langle \psi_i^\perp | \psi_i \rangle = 0$. While the problem can be completely determined from these conditions, it is convenient to represent Bob's POVM through the Neumark representation:

$$U_b |\psi_1\rangle |i\rangle = \sqrt{p_{1b}} |v_{11}\rangle |1\rangle + \sqrt{r_{1b}} |v_{12}\rangle |2\rangle \tag{3.5}$$

$$U_b |\psi_2\rangle |i\rangle = \sqrt{r_{2b}} |v_{21}\rangle |1\rangle + \sqrt{p_{2b}} |v_{22}\rangle |1\rangle \tag{3.6}$$

Bob's measurement consists of entangling his state and an ancilla state and then measuring the ancilla state. If he measures the ancilla to be in the $|i\rangle$ state (corresponding to the Π_i detector), Bob concludes that the state sent was ψ_i. Bob's probability of being correct given that states $|i\rangle$ was sent is p_{ib}. It is now straightforward to derive the constraints that Bob's probabilities must satisfy from the conditions of unitarity of the operation performed by Bob:

$$p_{ib} + r_{ib} = 1 \quad \forall i \in 1, 2 \tag{3.7}$$

$$\langle \psi_2 | \psi_1 \rangle = \sqrt{p_{b1} r_{b2}} \langle v_{11} | v_{12} \rangle + \sqrt{p_{b2} r_{b1}} \langle v_{21} | v_{22} \rangle . \tag{3.8}$$

After Bob's measurement the qubit is in one of two mixed states, $\rho_i = p_{bi} |v_{ii}\rangle \langle v_{ii}| + r_{bi} |v_{ji}\rangle \langle v_{ji}|$ $(i = 1, 2,\ i \neq j)$, depending on what state Alice sent. These states can then be

discriminated by Charlie. If Bob chooses his POVM such that $|v_{11}\rangle = |v_{12}\rangle = |v_{21}\rangle = |v_{22}\rangle$ then Bob performs an optimal ME discrimination and leaves no possibility for Charlie to perform any type of discrimination. Alternatively, Bob can choose his POVM such that $|v_{11}\rangle = |v_{21}\rangle$ and $|v_{12}\rangle = |v_{22}\rangle$ ensuring that his output to Charlie is always a pure state. In this case, Bob has denied Charlie any possibility to learn the outcome of his measurement, while still allowing Charlie to have a chance to guess the state initially sent by Alice. While Bob can also choose to send Charlie a set of mixed states for Charlie to discriminate, this paper will focus on the case where Charlie only has to discriminate between pure states.

After Bob's measurement, Charlie can perform a ME discrimination on the resulting pure or mixed states from Bob's POVM. In order for Charlie's measurement to be optimal, he defines Π_c such that there is only one output state to his measurement. The remaining problem lies in choosing Π_b and Π_c such that the joint probability of both Bob and Charlie succeeding in identifying the state sent by Alice is optimized. Formally, one can state the problem as follows. Find the maximum of

$$P_{ss} = \sum_{i=1}^{2} \eta_i \langle \psi_{bi}|\Pi_{bi}|\psi_{bi}\rangle \langle v_i|\Pi_{ci}|v_i\rangle \,, \tag{3.9}$$

subject to Eqs. (3.1) and (3.2) and their analogs for Charlie,

$$\sum_{i}^{N} \Pi_{ci} = I, \tag{3.10}$$

$$\Pi_{ci} \geq 0 \quad \text{for} \quad i = 1, 2 \,. \tag{3.11}$$

It is assumed that Bob picks a POVM such that when Alice sends $|\psi_i\rangle$ the output state is $|v_i\rangle$. This problem can be equivalently formulated in the following way. Find the maximum of

$$P_{ss} = \sum_{i=1}^{2} \eta_i p_{bi} p_{ci}, \tag{3.12}$$

subject to

$$\frac{s}{t} = \sqrt{p_{b1}\left(1-p_{b2}\right)} + \sqrt{p_{b2}\left(1-p_{b1}\right)}, \tag{3.13}$$

$$t = \sqrt{p_{c1}\left(1-p_{c2}\right)} + \sqrt{p_{c2}\left(1-p_{c1}\right)}. \tag{3.14}$$

where p_{bi} and p_{ci} are Bob and Charlie's probabilities of correctly identifying that Alice sent state $|\psi_i\rangle$, and where $s \equiv \langle\psi_1|\psi_2\rangle$ and $t \equiv \langle v_1|v_2\rangle$.

3.1 N receivers

From this example, it is clear how this problem should be extended so that there are N sequential receivers. Instead of choosing the optimal measurement, Charlie also needs to set up his POVM so that the post measurement states from his measurement are not identical. This is true for all $N-1$ receivers, where the final link in the chain performs the optimal measurement on the post measurement states produced by the $N-1$th observer. Assuming that the post measurements states are restricted to being pure states, the problem can then be formulated as follows. Find the maximum of

$$P_{js}^{N} = \eta_1 \prod_{n=1}^{N} p_{1n} + \eta_2 \prod_{n=1}^{N} p_{2n}, \tag{3.15}$$

subject to

$$\begin{aligned} \frac{t_n}{t_{n+1}} &= \sqrt{p_{1n}\left(1-p_{2n}\right)}+\sqrt{p_{2n}\left(1-p_{1n}\right)} \quad n \leq N-1, \\ t_N &= \sqrt{p_{1N}\left(1-p_{2N}\right)}+\sqrt{p_{2N}\left(1-p_{1N}\right)}, \end{aligned}$$

where p_{in} is the nth receiver's probability of succeeding given that Alice sent state $|\psi_i\rangle$, where $t_1 \equiv \langle\psi_1|\psi_2\rangle$, and where t_n is the overlap of the post-measurement states received by the Nth receiver.

3.2 Simplifying the problem

In order to optimize the SME problem for arbitrary priors, it is important to realize that the problem can be simplified significantly. In order to show that this is possible, we need to use induction, starting from the case with only two receivers. Using Lagrange's principle, one can reformulate the SME problem for two receivers as follows:

$$\begin{aligned} F &= \eta_1 p_{1b} p_{1c} + \eta_2 p_{2b} p_{2c} \\ &\quad +\lambda_1\left(\frac{s}{t}-\left(\sqrt{p_{1b}\left(1-p_{2b}\right)}+\sqrt{p_{2b}\left(1-p_{1b}\right)}\right)\right) \\ &\quad +\lambda_2\left(t-\left(\sqrt{\left(p_{1c}\left(1-p_{2c}\right)\right)}+\sqrt{p_{2c}\left(1-p_{1c}\right)}\right)\right). \end{aligned}$$

Optimizing with respect to p_{ib}, p_{ic}, and t gives five Lagrange conditions for the optimal solution:

$$\eta_i p_{ic}\sqrt{p_{ib}\left(1-p_{ib}\right)} = \frac{\lambda_1}{2}\left[\sqrt{\left(1-p_{1b}\right)\left(1-p_{2b}\right)}-\sqrt{p_{1b}p_{2b}}\right], \tag{3.16}$$

$$\eta_i p_{ib}\sqrt{p_{ic}\left(1-p_{ic}\right)} = \frac{\lambda_2}{2}\left[\sqrt{\left(1-p_{1c}\right)\left(1-p_{2c}\right)}-\sqrt{p_{1c}p_{2c}}\right], \tag{3.17}$$

$$\frac{s}{t^2} = \frac{\lambda_2}{\lambda_1}. \tag{3.18}$$

By noting that the right hand sides of the first two sets of equations above do not depend on i, one is able to derive the following equivalences:

$$\eta_1 p_{1c}\sqrt{p_{1b}\left(1-p_{1b}\right)} = \eta_2 p_{2c}\sqrt{p_{2b}\left(1-p_{2b}\right)}, \tag{3.19}$$

$$\eta_1 p_{1b}\sqrt{p_{1c}\left(1-p_{1c}\right)} = \eta_2 p_{2b}\sqrt{p_{2c}\left(1-p_{2c}\right)}. \tag{3.20}$$

If we divide Eq. (3.20) by Eq. (3.19) and rearrange slightly, we get the relation:

$$\frac{\sqrt{p_{2b}\left(1-p_{1b}\right)}}{\sqrt{p_{1b}\left(1-p_{2b}\right)}} = \frac{\sqrt{p_{2c}\left(1-p_{1c}\right)}}{\sqrt{p_{1c}\left(1-p_{2c}\right)}}. \tag{3.21}$$

By taking Eq. (3.13) and dividing by Eq. (3.14), and using Eq. (3.21), we can derive that $\frac{s}{t^2} = \frac{\sqrt{p_{1b}(1-p_{2b})}}{\sqrt{p_{1c}(1-p_{2c})}}$. By equating this with $\frac{\lambda_2}{\lambda_1}$ using Eq. (3.18), and dividing Eq. (3.17) by Eq. (3.16) to get another formula for $\frac{\lambda_2}{\lambda_1}$, we can derive the following:

$$\begin{aligned}
\frac{\sqrt{p_{1b}\left(1-p_{2b}\right)}}{\sqrt{p_{1c}\left(1-p_{2c}\right)}} &= \frac{\sqrt{p_{1b}\left(1-p_{1c}\right)}}{\sqrt{p_{1c}\left(1-p_{1b}\right)}}\frac{\sqrt{\left(1-p_{1b}\right)\left(1-p_{2b}\right)}-\sqrt{p_{1b}p_{2b}}}{\sqrt{\left(1-p_{1c}\right)\left(1-p_{2c}\right)}-\sqrt{p_{1c}p_{2c}}}, \\
\frac{\sqrt{\left(1-p_{1b}\right)\left(1-p_{2b}\right)}}{\sqrt{\left(1-p_{1c}\right)\left(1-p_{2c}\right)}} &= \frac{\sqrt{\left(1-p_{1b}\right)\left(1-p_{2b}\right)}-\sqrt{p_{1b}p_{2b}}}{\sqrt{\left(1-p_{1c}\right)\left(1-p_{2c}\right)}-\sqrt{p_{1c}p_{2c}}}, \\
\frac{\sqrt{p_{1c}\left(1-p_{1b}\right)}}{\sqrt{p_{1b}\left(1-p_{1c}\right)}}\sqrt{p_{2c}\left(1-p_{2b}\right)} &= \sqrt{p_{2b}\left(1-p_{2c}\right)}, \\
p_{2c}\left(1-p_{2b}\right) &= p_{2b}\left(1-p_{2c}\right) \Rightarrow p_{2c}=p_{2b}.
\end{aligned}$$

In the second to last line, we used Eq. (3.21) to derive the final line. Using this method, we can conclude that in the optimal solution $p_{1b} = p_{1c}$, $p_{2b} = p_{2c}$ and $t = \sqrt{s}$. This conclusion means that the problem can be reduced to a simpler form of simply maximizing:

$$P_{ss} = \eta_1 p_1^2 + \eta_2 p_2^2, \tag{3.22}$$

with respect to the constraint:

$$\sqrt{s} = \sqrt{p_1\left(1-p_2\right)} + \sqrt{p_2\left(1-p_1\right)}. \tag{3.23}$$

In order to simplify the problem of N receivers, we can use induction to reduce the problem. Applying the Lagrange formalism the SME problem with N receivers, the following optimization conditions can simply be derived:

$$\begin{aligned}
\eta_1 \prod_{j\neq i}^{N} p_{1j}\sqrt{p_{1i}\left(1-p_{1i}\right)} &= \frac{\lambda_i}{2}\left[\sqrt{\left(1-p_{1i}\right)\left(1-p_{2i}\right)} - \sqrt{p_{1i}p_{2i}}\right] \\
\eta_2 \prod_{j\neq i}^{N} p_{2j}\sqrt{p_{2i}\left(1-p_{2i}\right)} &= \frac{\lambda_i}{2}\left[\sqrt{\left(1-p_{1i}\right)\left(1-p_{2i}\right)} - \sqrt{p_{1i}p_{2i}}\right] \\
\frac{t_n}{t_{n+1}^2} &= \frac{\lambda_{n+1}}{\lambda_n t_{n+2}}\ n \leq N-2 \\
\frac{t_{N-1}}{t_N^2} &= \frac{\lambda_N}{\lambda_{N-1}}.
\end{aligned}$$

Focusing on the constraints for N and N-1, it can be shown that:

$$\begin{aligned}
\eta_1 \prod_{j}^{N-2} p_{1j}p_{1(N)}\sqrt{p_{1(N-1)}\left(1-p_{1(N-1)}\right)} &= \eta_2 \prod_{j}^{N-2} p_{2j}p_{2(N)}\sqrt{p_{2(N-1)}\left(1-p_{2(N-1)}\right)}, \\
\eta_1 \prod_{j}^{N-2} p_{1j}p_{1(N-1)}\sqrt{p_{1(N)}\left(1-p_{1(N)}\right)} &= \eta_2 \prod_{j}^{N-2} p_{2j}p_{2(N-1)}\sqrt{p_{2(N)}\left(1-p_{2(N)}\right)}, \\
\frac{\sqrt{p_{1(N)}\left(1-p_{1(N-1)}\right)}}{\sqrt{p_{1(N-1)}\left(1-p_{1(N)}\right)}} &= \frac{\sqrt{p_{2(N)}\left(1-p_{2(N-1)}\right)}}{\sqrt{p_{2(N-1)}\left(1-p_{2(N)}\right)}}.
\end{aligned}$$

This final condition can be used to show that:

$$\frac{t_{N-1}}{t_N^2} = \frac{\sqrt{p_{1(N-1}\left(1-p_{2(N-1)}\right)}}{\sqrt{p_{1(N)}\left(1-p_{2(N)}\right)}} = \frac{\lambda_N}{\lambda_{N-1}}.$$

In an identical fashion to the 2 receiver case, this can then be used to show that the optimal solution occurs for $p_{i(N-1)} = p_{i(N)}$ and $t_N = \sqrt{t_{N-1}}$. By substituting t_N^2 for t_{N-1}, the same procedure can be used N times to finally derive that for the optimal solution $t_N^N = s$. Using this result, it is trivial to show that this allows the SME problem for N receivers to be simplified in the following form. Maximize

$$P_{js}^n = \eta_1 p_1^n + \eta_2 p_2^n, \tag{3.24}$$

with respect to the constraint:

$$s^{\frac{1}{n}} = \sqrt{p_1 (1 - p_2)} + \sqrt{p_2 (1 - p_1)}. \tag{3.25}$$

3.3 Optimizing for Arbitrary Priors

While there is no simple analytic solution known for optimizing the general SME problem for arbitrary priors, the solution for equal priors, $\eta_1 = \eta_2 = \frac{1}{2}$ is known. The Lagrange constraints for the problem of equal priors can always be satisfied when $p_1 = p_2 = \frac{1}{2}\left(1 + \sqrt{1 - s^{\frac{1}{n}}}\right)$. This gives the following optimal solution:

$$P_{js,eq}^n = \left[\frac{1}{2}\left(1 + \sqrt{1 - s^{\frac{2}{n}}}\right)\right]^n. \tag{3.26}$$

It is important to note that this solution is only optimal within a range of values. This solution must be compared against the boundary solution, where $p_1 = 1$ and $p_2 = 1 - s^{\frac{2}{n}}$. Comparing $\left[\frac{1}{2}\left(1 + \sqrt{1 - s^{\frac{2}{n}}}\right)\right]^n$ with $\frac{1}{2}\left(1 + \left(1 - s^{\frac{2}{n}}\right)^n\right)$, gives the maximum value for s, defining this value as s_{bound}, where the above solution (3.26) is valid. For $s > s_{bound}$, the boundary solution is optimal. As can be seen in the figure above 3.1, as n increases, s_{bound} decays exponentially, asymptotically approaching 0. In other words, for any n greater than

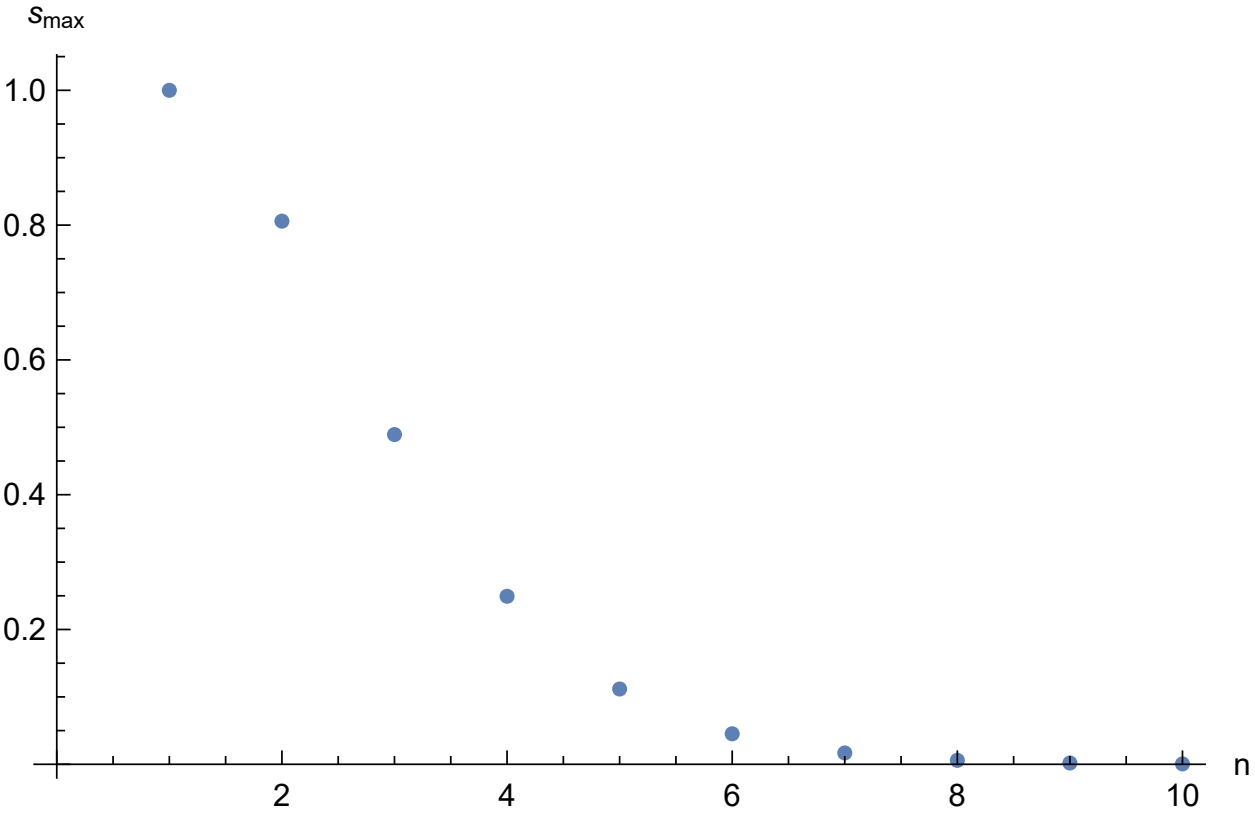

Figure 3.1: Values where $\left[\frac{1}{2}\left(1+\sqrt{1-s^{\frac{2}{n}}}\right)\right]^n = \frac{1}{2}\left(1+\left(1-s^{\frac{2}{n}}\right)^n\right)$ as a function of n

4, the solution given by the boundary solution is the optimal solution in most cases.

Without an analytic solution for the case of arbitrary priors, the optimal probability of success can be computed numerically. This is shown for the case of two receivers, as shown in Figure 3.2. Doing so requires rewriting the constraint in Eq. (3.25) for p_2. This can most easily be done by using the substitution $\sqrt{p_i} = \{sin(\theta_i), cos(\theta_i)\}$ and $s^{\frac{1}{n}} = sin(\phi)$. Doing so rewrites the constraint as $sin(\phi) = sin(\theta_1 + \theta_2)$, or $\phi = \theta_1 + \theta_2$. This gives two possible ways to express p_2:

$$\begin{aligned} p_2 &= sin^2(\phi - \theta_1) = \left(s^{\frac{1}{n}}\sqrt{1-p_1} - \sqrt{p_1}\sqrt{1-s^{\frac{2}{n}}}\right)^2, \\ &= cos^2(\phi - \theta_1) = \left(s^{\frac{1}{n}}\sqrt{1-p_1} + \sqrt{p_1}\sqrt{1-s^{\frac{2}{n}}}\right)^2. \end{aligned} \tag{3.27}$$

One can easily show that the second of these two solutions for p_2 is always greater, and therefore should be the one used. By substituting this expression into the original joint success probability in Eq. (3.24) and optimizing with respect to p_1, one derives the following expression:

$$0 = n\eta_1 p_1 + 2\left(s^{\frac{1}{n}}\sqrt{1-p_1} + \sqrt{p_1}\sqrt{1-s^{\frac{2}{n}}}\right)^{2n-1}\left(\sqrt{1-p_1}s^{\frac{1-n}{n}} - \frac{p_1}{\sqrt{1-s^{\frac{2}{n}}}}s^{\frac{2-n}{n}}\right)$$

By solving this equation for p_1, one can obtain the optimal joint probability of success. The numerical solution for the case of two receivers, or $N = 2$, can be seen in Figure 3.2.

One valid alternative approach is to construct an approximate solution. While not optimal, one can propose using the following solutions for p_1 and p_2.

$$p_i = \frac{1}{2}\left(1 + \frac{1 - 2(1-\eta_i)s^{\frac{2}{n}}}{\sqrt{1 - 4(\eta_1\eta_2 s^{\frac{2}{n}})}}\right).$$

This solution optimizes P_{js}^1 (see Eq. (3.24)) with respect to the constraint given by Eq. (3.25) and reduces to $P_{js,eq}^n$ for equal priors. We can quantify the effectiveness of this approximate solution by evaluating $E \equiv P_{js,opt}^n - P_{js,appx}^n$. For a calculation of this value for $n = 2$ as a function of the prior probability for a variety of state overlaps, see Figure 3.3.

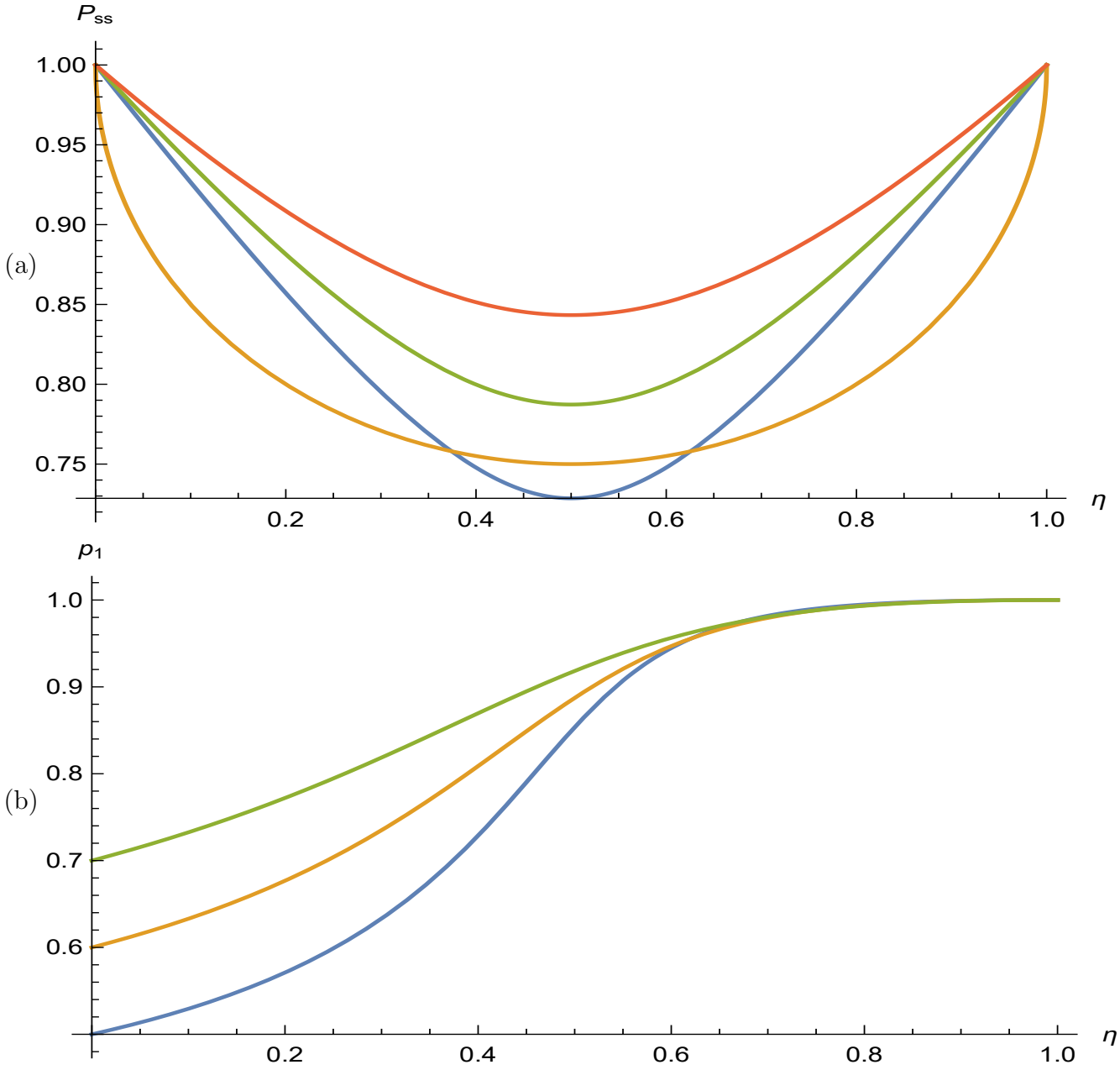

Figure 3.2: (a) Optimal Joint probability of Success and (b) Optimal success for the first state for two receivers versus the prior probability of sending the first state, calculated numerically for the given values of the overlap. One might be surprised that the term for p_1 does not go to zero when $\eta_1 \to 0$. This feature is an artifact of constraint $\sqrt{s} = \sqrt{p_1\,(1-p_2)} + \sqrt{p_2\,(1-p_1)}$ for $p_2 = 1$, which requires $p_1 = 1 - s$. This is closely related the fact that in the case that one measures the states in the $|\psi_2\rangle\,, |\psi_2^{\perp}\rangle$ basis, the probability of getting the result $|\psi_2^{\perp}\rangle$ when measuring the state $|\psi_1\rangle$ is $1 - s^2$.

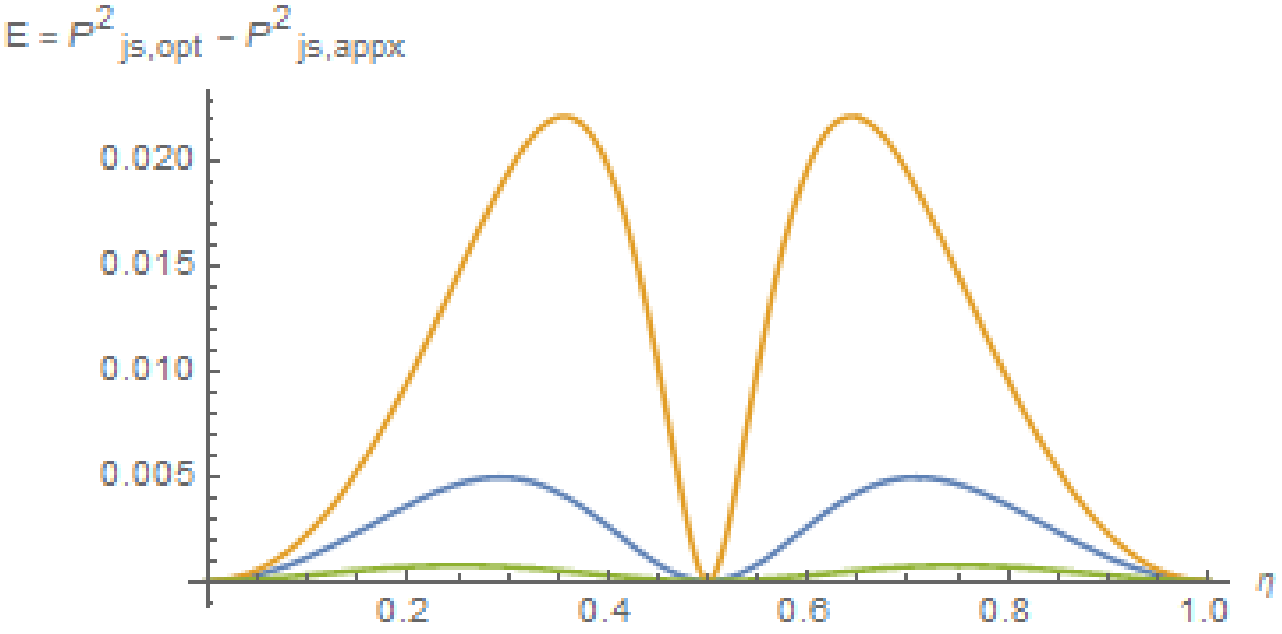

Figure 3.3: A plot of the $E = |P^2_{js,eq} - P^2_{js,appx}|$ as a function of prior probability η for $s = 0.7$ (Yellow), $s = 0.5$(Blue), and $s = 0.3$(Green).

Chapter 4

Sequential Unambiguous Discrimination

In the sequential unambiguous discrimination scheme, Alice prepares a qubit in one of two non-orthogonal states, either $|\psi_1\rangle$ or $|\psi_2\rangle$. The prior probability that $|\psi_i\rangle$ is prepared is η_i $\{i = 1, 2\}$, such that $\eta_1 + \eta_2 = 1$, so one of the states is always prepared. Sequential unambiguous discrimination was introduced in [47], where only the case $\eta_1 = \eta_2 = 1/2$ was considered. Additionally, the sequential unambiguous discrimination scheme for equal priors was experimentally verified [48], extended from discrete to continuous variable states [49], and extended to a sequence of multiple of observers [50]. Here we address the sequential unambiguous discrimination problem with two observers, arbitrary priors, and other generalizations.

Similar to the SME scheme, the states and their priors are also known to Bob and Charlie, they just do not know which state the qubit was actually prepared in. After the preparation Alice sends the qubit to Bob who performs a measurement (POVM) on the qubit, and sends the qubit he just measured to Charlie, who then also performs a measurement (POVM) on the qubit he received. In order for their measurements to be unambiguous, both Bob and

Charlie require that they have detection outcomes that can only result from a specific state sent by Alice. The goal for them is to maximize their joint probability of succeeding subject to this constraint. This goal is compatible with additional optimizations and in what follows we will analyze these options in detail.

In the original presentation of this scheme in [47],the standard POVM formalism was employed. For the purposes of this thesis, we utilize an alternative but equivalent formalism, based on Neumark's extension. In this formalism one first entangles the qubit with an ancilla and then performs a standard projective measurement on the ancillary system. The interaction of the qubit with the ancilla is described by a unitary time evolution operator,

$$U_{\rm b}\left|\psi_1\right\rangle\left|i\right\rangle = \sqrt{p_{1\rm b}}\left|\varphi_1\right\rangle\left|1\right\rangle + \sqrt{q_{1\rm b}}\left|\Phi_1\right\rangle\left|0\right\rangle, \tag{4.1}$$

$$U_{\rm b}\left|\psi_2\right\rangle\left|i\right\rangle = \sqrt{p_{2\rm b}}\left|\varphi_2\right\rangle\left|2\right\rangle + \sqrt{q_{2\rm b}}\left|\Phi_2\right\rangle\left|0\right\rangle. \tag{4.2}$$

Here the subscript b stands for Bob, $|i\rangle$ is the initial state of the ancilla while $|0\rangle$, $|1\rangle$ and $|2\rangle$ are three orthogonal states of the ancillary system. If Bob performs a measurement on the ancilla in the basis formed by these three states and finds either $|1\rangle$ or $|2\rangle$ as the measurement outcome, he will know what state Alice has prepared. If, on the other hand, he finds $|0\rangle$ as the outcome of his measurement, he will not acquire unambiguous information about the input state, hence this result is inconclusive. Therefore, $p_{i\rm b}$ is Bob's success probability of unambiguously identifying the input state $|\psi_i\rangle$ and $q_{i\rm b}$ is Bob's probability of failing to identify the input state. $|\varphi_i\rangle$ and $|\Phi_i\rangle$ $(i = 1, 2)$ are the post-measurement states of the qubit associated with the various outcomes of the measurement performed on the ancilla.

After Bob has performed his state-identifying measurement, he passes the qubit to Charlie, whose task is also to unambiguously identify the initial state of the qubit that Alice prepared. It is known that for unambiguous identification the states to be identified must be linearly independent [24]. For a qubit, this means that two pure states can be unam-

biguously discriminated. This requirement puts serious restrictions on how Bob can design the post-measurement states. There is one additional requirement. We also want that the post-measurement states of the qubit carry no information about the outcome of Bob's measurement, a condition that is central to applications for quantum communication.

The following choice satisfies these requirements and allows for Charlie to unambiguously identify the state initially prepared by Alice. At the same time he does not learn anything about the outcome of Bob's measurement. We set $|\varphi_i\rangle = |\Phi_i\rangle$, following the original proposal for the sequential unambiguous discrimination scheme [47]. This choice ensures that Charlie receives one of two pure states that, when discriminated can be correlated back to Alice's original state. Given this, Charlie's measurement, again employing the Neumark method, can be represented as

$$U_{\rm c}\,|\varphi_1\rangle\,|i\rangle \;=\; \sqrt{p_{1\rm c}}\,|\theta_1\rangle\,|1\rangle + \sqrt{q_{1\rm c}}\,|\Theta_1\rangle\,|0\rangle\,, \tag{4.3}$$

$$U_{\rm c}\,|\varphi_2\rangle\,|i\rangle \;=\; \sqrt{p_{2\rm c}}\,|\theta_2\rangle\,|2\rangle + \sqrt{q_{2\rm c}}\,|\Theta_2\rangle\,|0\rangle\,. \tag{4.4}$$

It has been shown previously that, in order to optimally discriminate between $|\varphi_1\rangle$ and $|\varphi_2\rangle$, Charlie must choose $|\Theta_1\rangle = |\Theta_2\rangle \equiv |\theta_0\rangle$ (see, e.g., [10]). In order to simplify the following discussion, we introduce the notation $\langle\psi_1|\psi_2\rangle = s$ and $\langle\varphi_1|\varphi_2\rangle = t$, We can express the constraints, resulting from the unitarity of U_b and U_c, in terms of these parameters as

$$p_{\rm jb} + q_{jb} = p_{jc} + q_{jc} = 1 \tag{4.5}$$

for $j = 1, 2$ and

$$\frac{s}{t} = \sqrt{q_{1\rm b}q_{2\rm b}},\; t = \sqrt{q_{1\rm c}q_{2\rm c}}. \tag{4.6}$$

The average probability that both Bob and Charlie succeed in unambiguously identifying

the state that Alice sent, the joint success probability, can be written as:

$$P_{\rm ss} = \eta_1 p_{\rm 1b} p_{\rm 1c} + \eta_2 p_{\rm 2b} p_{\rm 2c} \tag{4.7}$$

This is the central quantity for the rest of this work. The main goal is to optimize this expression under the constraints given by Eqs.(4.5) and (4.6).

By making use of the constraints given in Eqs. (4.5) and (4.6), we can write $P_{\rm ss}$ as

$$\begin{aligned} P_{\rm ss} &= \eta_1 \left(1 - q_{\rm 1b}\right)\left(1 - q_{\rm 1c}\right) \\ &\quad + \eta_2 \left(1 - \frac{s^2}{t^2 q_{\rm 1b}} - \frac{t^2}{q_{\rm 1c}} + \frac{s^2}{q_{\rm 1b} q_{\rm 1c}}\right). \end{aligned} \tag{4.8}$$

Equations (4.5)-(4.8) represent the starting point for the various optimization schemes and discussions in the next four sections. In particular, Eq. (4.8) is a function of three independent parameters, t, q_{1b} and q'_{1c}. Their range is given by $s \leq t \leq 1$, $\frac{s^2}{t^2} \leq q_{1b} \leq 1$ and $t^2 \leq q_{1c} \leq 1$. For the optimal $P_{\rm ss}$, the parameters are either internal points in these intervals or lie at the boundary. In the first case the derivatives of $P_{\rm ss}$ with respect to the variables t, q_{1b} and q_{1c} exist and the optimum can be found analytically. The boundary points need to be investigated separately and then compared to the internal optimum points, which may give local optimum only, to find global optimum.

Before we move on to discuss the general case, we deal with the special case of $\eta_1 = \eta_2 = 1/2$, which was the case considered in Refs. [47] and [51]. It was shown in [47] that $t^2 = s$ for optimum joint probability of success. Actually, we will see in the next sections that this remains the optimal choice for general priors, as well. Under this condition, the equations (4.5)-(4.8) are completely symmetric in the indices 1 and 2, and also in b and c. This immediately yields $q_{1b} = q_{1c} = q_{2b} = q_{2c} = \sqrt{s}$ for the internal point solution. Inserting

all these values into Eq. (4.8), gives

$$P_{ss,1}^{opt} = (1 - \sqrt{s})^2 \tag{4.9}$$

for the optimum joint probability of success, which is the result found in [47]. For the boundary solution we can choose either $q_{1b} = 1$ or $q_{2b} = 1$ but not both. For the sake of concreteness, let us make the first choice. From Eq. (4.6) we then find $q_{2b} = s$. Similarly, for the boundary solution, we can choose either $q_{1c} = 1$ or $q_{2c} = 1$ but not both. If we choose $q_{2c} = 1$ then Bob will always fail to identify the first state and sometimes identifies the second state. Charlie, however always fails to identify the second state and sometimes identifies the first. Clearly, their joint probability of success is zero in this case, giving the minimum of P_{ss}. So, we must choose $q_{1c} = 1$ leading to $q_{2b} = s$. Inserting all these values into Eq. (4.8) again, gives

$$P_{ss,2}^{opt} = \frac{1}{2}(1 - s)^2 \tag{4.10}$$

for the optimum joint probability of success, which is the result found in [51]. As it turns out, $P_{ss,1}$ is optimum if $s \leq s_{crit}$ and $P_{ss,2}$ is optimum if $s > s_{crit}$, where $s_{crit} = (\sqrt{2} - 1)^2$ is the critical value of the overlap parameter where the two solutions intersect.

Clearly, a two-state QKD protocol can be based on the sequential scheme. It is very closely related to the B'92 protocol [52], extending it to multiple recipients. Alice encodes the bit value 0 into the first state and 1 into the second state. She prepares a large number of qubits at random in one of these states and sends them to Bob who performs the above described state identifying measurement on them and sends the qubits in their post-measurement states to Charlie who performs an optimal UD measurement on them. They publicly announce the instances when they succeeded but not the result. They keep the results when they succeed and discard the rest. Since Alice knows what she prepared in those instances, she will share a string of 0's and 1's with Bob in those instances when Bob

succeeds and, similarly, a separate string with Charlie in the instances when Charlie succeeds. In addition, in the instances when both Bob and Charlie succeed, they will share a subset of their bit-strings that is common to all three of them. These bit strings serve as the raw key and the rest of the protocol (checking for the presence of eavesdropper(s) and distilling a communication key) follows the same lines as in the original B'92 protocol. The established communication keys can serve as secure keys for a secure three-way communication protocol. It is clear that for this QKD protocol the measurement presented in [47] has to employed. The measurement presented in [51] cannot be used in communication protocols since it generates a string of identical bit values, either all 0's or all 1's, which is clearly not what is needed for a key.

After these preliminaries, we now proceed to the discussion of the general case. At this point, we arrive at a juncture, one can follow one of two ways. One can maximize the joint probability of success and simultaneously minimize the joint probability of failure,

$$P_{\text{ff}} = \eta_1 q_{1b} q_{1c} + \eta_2 q_{2b} q_{2c}. \tag{4.11}$$

Alternatively, one can maximize the joint probability of success only, without minimizing the joint probability of failure. The two methods yield slightly different results. In addition, the first allows for a fully analytical treatment while the second also involves numerics. We present the first approach in the next section and then the second method in Section 4.2.

4.1 Simultaneous optimization of the joint probability of success and the joint probability of failure

The joint probability of failure, Eq. (4.11), can be optimized independently of the rest of the problem, based on the following observation. Taking the product of the constraints in

Eq. (4.6) yields $q_{1b}q_{1c}q_{2b}q_{2c} = s^2$, which is independent of t. So, $q_{2b}q_{2c}$ can be expressed in terms of of the failure probabilities of the first state,

$$q_{2b}q_{2c} = \frac{s^2}{q_{1b}q_{1c}}. \tag{4.12}$$

Inserting this expression into Eq. (4.11), $P_{\rm ff}$ will depend only on the single combination of the parameters, $q_{1b}q_{1c}$. The optimization with respect to this parameter is straightforward, with the result

$$q_{1b}^{opt}q_{1c}^{opt} = \begin{cases} \sqrt{\frac{\eta_2}{\eta_1}}s & \text{if } \frac{s^2}{1+s^2} \leq \eta_1 \leq \frac{1}{1+s^2} \ , \\ 1 & \text{if } \eta_1 < \frac{s^2}{1+s^2} \ , \\ s^2 & \text{if } \frac{1}{1+s^2} < \eta_1 \ . \end{cases} \tag{4.13}$$

Substituting the optimal values into Eq. (4.11) yields

$$P_{\rm ff}^{\rm opt} = \begin{cases} 2\sqrt{\eta_1\eta_2}s & \text{if } \frac{s^2}{1+s^2} \leq \eta_1 \leq \frac{1}{1+s^2} \ , \\ \eta_1 + \eta_2 s^2 & \text{if } \eta_1 < \frac{s^2}{1+s^2} \ , \\ \eta_2 + \eta_1 s^2 & \text{if } \frac{1}{1+s^2} < \eta_1 \ . \end{cases} \tag{4.14}$$

Interestingly, this expression is identical to the one obtained for optimal unambiguous discrimination of the two states by Bob alone [10, 38]. This was to be expected, since Bob can first perform a partial discrimination of the two states and then in a second step a full discrimination of the remaining states, i.e., he can assume the role of Charlie in the sequence. What the above result tells us is that no matter in how many steps the discrimination is performed, its optimal failure probability is always given by the above equation. Thus, quantum mechanics sets a universal bound on the global failure probability.

The individual success probabilities of Bob and Charlie are, however, subject to further optimization. In addition to Eqs. (4.5) and (4.6), we now have Eqs. (4.12) and (4.13) as

constraints for the optimization of $P_{\rm ss}$. With the help of the first line in Eq. (4.13), we can express q_{1c} in terms of q_{1b},

$$q_{1c}^{opt} = \sqrt{\frac{\eta_2}{\eta_1}} \frac{s}{q_{1b}^{opt}}. \tag{4.15}$$

Inserting this expression in Eq (4.8), $P_{\rm ss}$ becomes a function of t and q_{1b}^{opt} only. After some straightforward algebra, it can be written as

$$\begin{aligned} P_{\rm ss} &= 1 + 2\sqrt{\eta_1\eta_2} - \eta_1 q_{1\rm b} - \sqrt{\eta_1\eta_2}\frac{s}{q_{1\rm b}} \\ &\quad -\eta_2 \frac{s^2}{t^2 q_{1\rm b}} - \sqrt{\eta_1\eta_2}\frac{t^2}{s} q_{1\rm b}. \end{aligned} \tag{4.16}$$

The optimization with respect to t and q_{1b}^{opt} is again straightforward, yielding the unique solutions $t^2 = s$ and

$$q_{1b}^{opt} = \begin{cases} (\frac{\eta_2}{\eta_1})^{1/4}\sqrt{s} & \text{if } \frac{s^2}{1+s^2} \le \eta_1 \le \frac{1}{1+s^2} \ , \\ 1 & \text{if } \eta_1 < \frac{s^2}{1+s^2} \ , \\ s & \text{if } \frac{1}{1+s^2} < \eta_1 \ . \end{cases} \tag{4.17}$$

Inserting these expressions in Eq. (4.16), we obtain the optimal joint success probability, under the condition that the joint probability of failure is minimum, as

$$P_{\rm ss}^{\rm opt} = \begin{cases} \left(\sqrt{\eta_1} - (\eta_1\eta_2)^{1/4}\sqrt{s}\right)^2 + \left(\sqrt{\eta_2} - (\eta_1\eta_2)^{1/4}\sqrt{s}\right)^2 & \text{if } \frac{s^2}{1+s^2} \le \eta_1 \le \frac{1}{1+s^2} \ , \\ \eta_2(1-s)^2 & \text{if } \eta_1 < \frac{s^2}{1+s^2} \ , \\ \eta_1(1-s)^2 & \text{if } \frac{1}{1+s^2} < \eta_1 \ . \end{cases} \tag{4.18}$$

This solution is unique and it completely solves the problem of sequential state discrimination under the condition that the joint probability of failure is minimum simultaneously with the condition that the joint probability of success is maximum. However, if we relax the

requirement that the joint probability of failure is at its minimum, the joint probability of success can still be optimized and it can exceed the value given in Eq. (4.18). We will present this case in the next subsection.

4.2 Optimizing the joint probability of success without minimizing the joint probability of failure

First, let us consider the extrema of the joint probability of success with respect to t. They require that t is either on the boundary of the allowed range or the derivative with respect to t is zero, i.e.,

$$\frac{d}{dt}P_{\rm ss} = 2\eta_2 \left(\frac{s^2}{t^3 q_{1b}} - \frac{t}{q_{1c}} \right) = 0 \,, \tag{4.19}$$

which gives $q_{1c} = q_{1b}t^4/s^2$. Together with the two constraints of (4.6), we have $q_{2c} = \frac{t^2}{q_{1c}} = \frac{s^2}{t^2 q_{1b}} = q_{2b}$. In a similar way, we also have $q_{1c} = q_{1b}$ for the optimal solution due to the symmetry of the discrimination scheme for the two signal states. Thus, we have demonstrated that, in order to optimize the joint probability of success, Bob and Charlie must have the same probability for correctly identifying Alice's message, i.e., $q_{1c} = q_{1b}$ and $q_{2c} = q_{2b}$. Inserting these conditions into the constraints, yields immediately $s/t^2 = \sqrt{\frac{q_{1b}q_{2b}}{q_{1c}q_{2c}}} = 1$. Hence, $t = \sqrt{s}$. After the elimination of two of the three independent parameters, there is only one parameter left for the optimization of the joint probability of success

$$P_{\rm ss} = \eta_1(1 - q_{1b})^2 + \eta_2 \left(1 - \frac{s}{q_{1b}} \right)^2 . \tag{4.20}$$

At the optimal value, the derivative with respect to the parameter q_{1b} must also vanish. $\frac{d}{dq_{1b}}P_{\rm ss} = 0$ yields a quartic equation to solve,

$$\frac{\eta_1}{\eta_2}q_{1b}^3(1-q_{1b}) - (q_{1b}-s)s = 0\,. \tag{4.21}$$

This equation has four real or complex solutions. However, the physical solutions must be real and within the range of $s \le q_{1b} \le 1$, depending on the value of η_1/η_2 and s.

The joint probability of success $P_{\rm ss}$ of (4.20) against q_{1b}

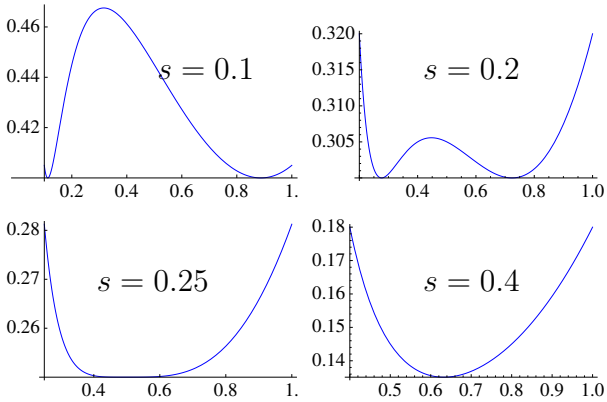

Figure 4.1: For $\eta_1 = \eta_2$, the joint probability of success $P_{\rm ss}$ of (4.20) as a function of its only free parameter q_{1b} for $s = 0.1, 0.2, 0.25$ and 0.4, respectively. For each of the subfigures, the physical range of $s \le q_{1b} \le 1$ is plotted. For $s \le 1/4$, the function has a local maximum at $q_{1b} = \sqrt{s}$ and two global minima at $q_{1b} = 1/2(1 \pm \sqrt{1-4s})$. The optimal value of $P_{\rm ss}$ is given by its local maximum at q_{1b} for very small s, and it is given by its boundary values as the local maximum value gets smaller with increasing s. There is only one extremal value within the physical range for $s \ge 1/4$ and it is a global minimum. This minimum is of second order when $s = 1/4$.

In the following, we will first illustrate as an example the case for equal priors $\eta_1 = \eta_2$ where the quartic equation is solved analytically and then extend our solution to the general case of arbitrary priors. For equal priors, the four solutions of the equation are $\{\pm\sqrt{s}, 1/2(1\pm\sqrt{1-4s})\}$. For $s < 1/4$, there are three physical solutions: $q_{1b} = \{\sqrt{s}, 1/2(1\pm\sqrt{1-4s})\}$. For $s \ge 1/4$, there is only one physical solution at $q_{1b} = \sqrt{s}$. The solutions for $q_{1b} = 1/2(1 \pm \sqrt{1-4s})$ (if exist) always give the location of the minima of $P_{\rm ss}$. Thus, the

maximum of $P_{\rm ss}$ must either be its value at the extremal point of $q_{1b} = \sqrt{s}$, or its value at the boundary solutions $q_{1b} = s$ or $q_{1b} = 1$; see Fig. 4.1. Evaluating the joint probability of success at these values, we have its local extremal given by $P_{\rm ss}(q_{1b}{=}\sqrt{s}) = (1 - \sqrt{s})^2$, and its boundary values given by $P_{\rm ss}(q_{1b}{=}s) = P_{\rm ss}(q_{1b}{=}1) = \frac{1}{2}(1 - s)^2$. The boundary values are larger than the local maximum when $1 - s > \sqrt{2}(1 - \sqrt{s})$, i.e., $s > 3 - 2\sqrt{2} = 0.1716$. Thus, for the global optimum we have

$$(P_{\rm ss})_{\rm max} = \begin{cases} (1 - \sqrt{s})^2 & \text{if} \quad s \leq 3 - 2\sqrt{2} \\ \frac{1}{2}(1 - s)^2 & \text{if} \quad s > 3 - 2\sqrt{2} \end{cases} . \tag{4.22}$$

The dependence of the optimal joint probability of success on the overlap of the states s is illustrated in Fig. 4.2.

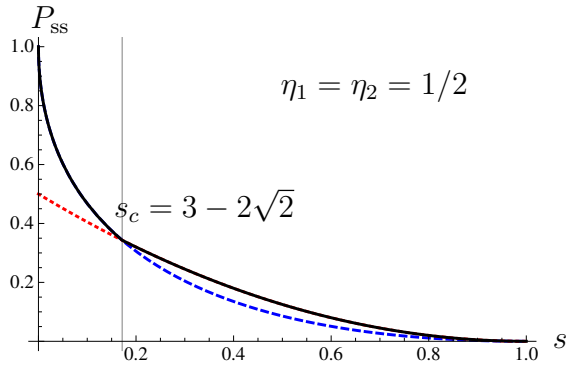

Figure 4.2: The solid black curve gives the optimal joint success probability $P_{\rm ss}$ as a function of $s = \langle\psi_1|\psi_2\rangle$ for equal priors. The dotted red curve shows the boundary value solutions and the dashed blue curve shows the value of the function at $q_{1b} = \sqrt{s}$; the critical value of $s_c = 3 - 2\sqrt{2}$ is the value at which these two curves intersect.

For general priors, $\eta_1 \neq \eta_2$, the optimal value of $P_{\rm ss}$ must also be either given by one of the physical solutions of (4.21) in the interval $s < q_{1b} < 1$, or by its value on the boundary, $q_{1b} = s$ or $q_{1b} = 1$. The two boundary solutions, however, are not the same as in the case of equal priors. If $\eta_1 > \eta_2$, the boundary solution at the lower boundary, $q_{1b} = s$, is larger than the value at the upper boundary $q_{1b} = 1$; and vice versa. The larger boundary value

solution is given by

$$P_{\rm ss}^{b} = \eta_{max}(1-s)^2\,, \tag{4.23}$$

where $\eta_{max} = \max\{\eta_1, \eta_2\}$. For every set of priors, there is a critical value of s for which the boundary value solution of $P_{\rm ss}$ is the same as its value at the local maximum between $s \leq q_{1b} < 1$. This switching of the optimal value between the local maximum and the boundary values can be understood by the relation between $P_{\rm ss}$ and the constraint shown in Fig. 4.3.

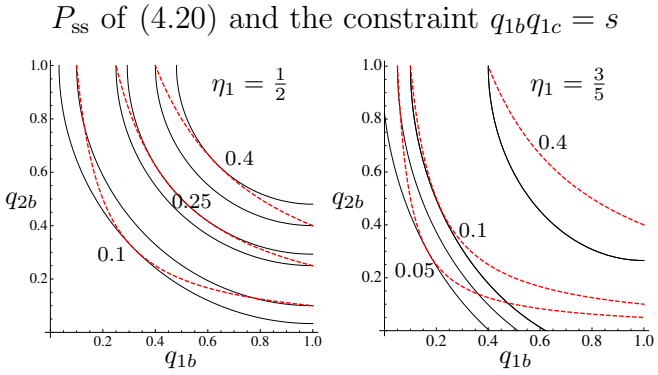

Figure 4.3: Contour plot of the joint probability of success $P_{\rm ss}$ of (4.20) and the constraint $q_{1b}q_{1c} = s$ as a function of q_{1b} and q_{1c}.The solid black curves are the contours of $P_{\rm ss}$ and the dashed red curves are plot of $q_{1b}q_{1c} = s$ where the value of s are labeled next to the curves. For $\eta_1 = \eta_2$ (left), contours of the joint probability of success $P_{\rm ss}$ are quarters of circles and they are symmetric in the variables of both axis; for $\eta_1 \neq \eta_2$ (left), contours of the joint probability of success $P_{\rm ss}$ are sections of ellipses, instead.

We label the critical value of s by s_c, and for $s \geq s_c$, we have $P_{\rm ss} \leq P_{ss}^{b}$. The dependence of s_c on the prior probability η_1 is shown in Fig. 4.4. The critical value $s_c = 3-2\sqrt{2} \approx 0.1716$ for equal prior probability distribution, and s_c decreases as the prior probability distribution becomes more biased. The parameter region where the local maximum of (4.20) is the optimal value for the joint probability of success, shown by the shaded region in Fig. 4.4, is quite small compared to the entire parameter regime of s and $\{\eta_1, \eta_2\}$, which is given by the unit square $0 \leq s, \eta_1 \leq 1$. Thus, for most of the range of s and given priors, $P_{\rm ss}$ is optimized at its boundary solution for $q_1 = q_1' = s$ and $q_2 = q_2' = 1$. In this case, both Bob and Charlie

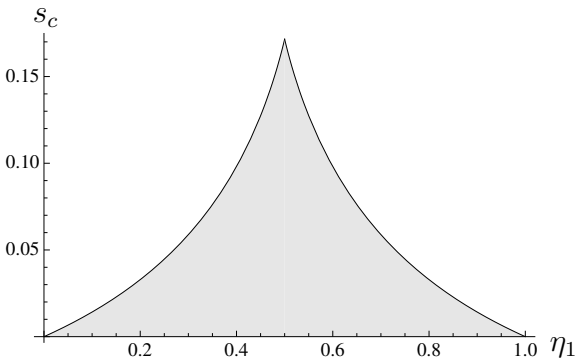

Figure 4.4: The critical value for the overlap of states s_c, as a function of prior probability η_1. The shaded area indicates the parameter regime where the nontrivial solution for the local maximum of P_{ss} is larger than the boundary solutions.

fail to detect state $|\psi_2\rangle$ (or $|\psi_1\rangle$) at all time but they optimize their set up such that state $|\psi_1\rangle$ (or $|\psi_2\rangle$) is successfully identified at a high probability of $1 - s$. Although the joint probability of success can be optimized by the boundary solutions, the information that Bob and Charlie share with each other and with Alice is of no use for communication as they only get a string of identical bits, after discarding the inconclusive outcomes. For example, in the case of $\eta_1 > \eta_2$, they share a string of 0's which carries no useful information. However, in the next section we discuss a measurement scheme that salvages the boundary solution and makes it useful even for communication purposes.

4.3 The Flip-Flop Measurement

In the previous sections, we found that for a large range of the overlap parameter, s, the measurement that optimizes the joint probability of success is the one which unambiguously identifies one of the states and misses the other completely. From the discussion in Sec. 2, we showed that simply performing this measurement cannot transmit information that is useful for quantum communication.

It was noticed, however, already in the case of two-party communication between Alice

and Bob (e.g. in the B92 cryptography protocol [52]) that the von Neumann setup can be used to generate a random key. In this case, Bob randomly choses between the two von Neumann setups, one that projects on $\{|\psi_1\rangle, |\psi_1^\perp\rangle\}$ and the other that projects on $\{|\psi_2\rangle, |\psi_2^\perp\rangle\}$. For the first setup, $P_0^{(1)} = |\psi_1\rangle\langle\psi_1|$ is the inconclusive detector since a click of this detector may originate from either of the input states, and $\mathcal{I} - P_0^{(1)} = |\psi_1^\perp\rangle\langle\psi_1^\perp|$ is the one that unambiguously identifies the input as $|\psi_2\rangle$ since it never clicks for $|\psi_1\rangle$. The action of the second setup can be obtained by interchanging the indices 1 and 2. In the flip-flop measurement Bob randomly chooses between the two setups. With probability c he chooses the setup that succeeds only for the first state and with probability $1-c$ he choses the setup that succeeds only for the second state. What this means is that Bob effectively flip-flops between the two von Neumann setups. The failure probability, averaged over the flipping rate, is

$$q_1 = (1-c)\langle\psi_1|P_0^{(1)}|\psi_1\rangle + c\langle\psi_1|P_0^{(2)}|\psi_1\rangle = (1-c) + cs^2 \tag{4.24}$$

for the first state, and

$$q_2 = (1-c)\langle\psi_2|P_0^{(1)}|\psi_2\rangle + c\langle\psi_2|P_0^{(2)}|\psi_2\rangle = c + (1-c)s^2 \tag{4.25}$$

for the second. Clearly, we have that $q_1 q_2 = s^2 + c(1-c)(1-s^2)^2 \geq s^2$, so this is not the optimal discrimination procedure unless $c = 0$ or $c = 1$.

The success probability averaged over the flipping rate is $p_1 = 1 - q_1 = c(1-s^2)$ for the first state, and $p_2 = 1 - q_2 = (1-c)(1-s^2)$ for the second. Thus, the average success probability for the flip-flop measurement is given by

$$P_{\text{succ}} = \eta_1 p_1 + \eta_2 p_2 = [\eta_1 c + \eta_2(1-c)](1-s^2)\,. \tag{4.26}$$

The average probability of failure, Q, is given by $Q = 1 - P_{\text{succ}}$.

P_{succ} is a linear function of the flipping rate c, so the function is either monotonically increasing, monotonically decreasing or constant. If $\eta_1 = \eta_2 = 1/2$, the function is constant. Otherwise, the maximum is on one of the boundaries of the $0 \leq c \leq 1$ interval. Clearly, P_{succ} reaches its maximum for $c = 1$ when $\eta_1 > \eta_2$ and for $c = 0$ when $\eta_1 < \eta_2$. Thus, the strategy that maximizes the success probability is to always bet on the state with the larger prior probability.

Obviously, the flip-flop measurement with $c \neq 0, 1$ has a lower success probability than the optimal boundary solution. However, it has the capability of generating a bit string that contains both 0's and 1's, not just one of them. In this respect, one particular choice of c stands out. For $c = \eta_2$, the two terms on the r.h.s. of (4.26) become equal, yielding $P_{\text{succ}} = 2\eta_1\eta_2(1 - s^2)$. In this case, the flip-flop measurement generates a random string of 0's and 1's where the occurrence probability of the 0's is equal to that of the 1's, a very desirable feature for QKD application.

After the discussion of the flip-flop measurement on the example of two-party communication, we now extend these considerations to the sequential UD scheme. In the sequential version of the flip-flop measurement both Bob Charlie choose randomly between the two setups. For simplicity, we assume that their flipping rates are equal. Independently, each with probability c chooses the setup that succeeds only for the first state and with probability $1 - c$ chooses the setup that succeeds only for the second state. What this means is that Bob and Charlie independently flip-flop between the corresponding two setups. Their failure probabilities, averaged over the flipping rate, are

$$q_{1\text{b}} = cs + (1 - c)\,, \tag{4.27}$$

$$q_{1\text{c}} = cs + (1 - c)\,, \tag{4.28}$$

$$q_{2\text{b}} = c + (1 - c)s\,, \tag{4.29}$$

$$q_{2c} = c + (1-c)s\,. \tag{4.30}$$

The corresponding success probabilities averaged over the flipping rate are

$$p_{1b} = c(1-s)\,, \tag{4.31}$$

$$p_{1c} = c(1-s)\,, \tag{4.32}$$

$$p_{2b} = (1-c)(1-s)\,, \tag{4.33}$$

$$p_{2c} = (1-c)(1-s)\,. \tag{4.34}$$

Thus, the average joint probability of success for the flip-flop measurement is then given by

$$P_{\rm ss}^{(\rm f)} = c^2\eta_1(1-s)^2 + (1-c)^2\eta_2(1-s)^2. \tag{4.35}$$

This is a simple quadratic function of the flipping rate, c, reaching its maximum at $c = 1$ if $\eta_1 > \eta_2$ and at $c = 0$ if $\eta_2 > \eta_1$. Perhaps more interesting, it is minimum when $c = \eta_2$ with the minimum value,

$$P_{\rm ss,min}^{(\rm f)} = \eta_1\eta_2(1-s)^2\,. \tag{4.36}$$

Clearly, this is the worst strategy for unambiguous identification of the states prepared by Alice. However, what is worst for one thing is best for another. This strategy will generate an unbiased bit string of 0's and 1's, so this best for application in QKD or, in general, quantum communication schemes.

4.4 Mutual Information

4.4.1 Unambiguous communication channel

One negative feature, when the optimal solution is on the boundary, is that it does not lead to any quantum communication protocol between Alice and Bob (and Charlie). In this case, Bob effectively ignores one of the states, setting the probability of successfully detecting that state to 0. If Bob wants to restrict himself to only keeping a result when he is certain about it, then he will end up with a string of identical bits. This is, of course, useless for establishing a secret key with Alice, so there is no way to share information between the two parties. We can quantify the amount of information transmitted by the mutual information. We adopt the common convention of denoting the message of the sender by X and the message the receiver decoded by Y. The mutual information of the communication channel is defined as

$$I\left(A:B\right)=H(X)-H(X|Y). \tag{4.37}$$

$H(X)=H(\eta_1)\equiv=-\eta_1\log_2\eta_1-(1-\eta_1)\log_2(1-\eta_1)$ denotes the Shannon entropy of the sender's binary information and $H(X|Y)$ denotes the conditional Shannon entropy [53]. For a general three-element POVM $\{\Pi_1,\Pi_2,\Pi_0\}$, the mutual information is given by [8]

$$I(A:B)=H(\eta_1)-\sum_{j=0}^{2}P\left(\Pi_j\right)H\left(X|\Pi_j\right), \tag{4.38}$$

where $P(\Pi_j)$ denotes the probability of having measurement outcome Π_j. If Bob gets a click in either the Π_1 or Π_2 detectors, he has no uncertainty as to what state Alice sent, therefore $H(X|\Pi_1)=H(X|\Pi_2)=0$. If $\{q_1,q_2\}$ represent the failure probabilities when Bob attempts to detect states $\{|\psi_1\rangle,|\psi_2\rangle\}$, then $P\left(\Pi_0\right)=\eta_1q_1+\eta_2q_2\equiv Q$ and $H\left(X|\Pi_0\right)=H\left(\frac{\eta_1q_1}{Q}\right)$.

Plugging these values in to (4.38), we have

$$I\left(A:B\right)=H\left(\eta_1\right)-QH\left(\frac{\eta_1 q_1}{Q}\right). \tag{4.39}$$

The mutual information is maximized when $QH\left(\frac{\eta_1 q_1}{Q}\right)$ is minimized. This calculation suggests that information is maximally detected when Bob is only able to detect one of the two incoming states, which is a counterintuitive result. To resolve this quandary, one must realize that this formulation of mutual information has no requirement that Bob determine the state sent by Alice definitively. Instead, this calculation relies on treating all three detection outcomes by Bob, $\{\Pi_1,\Pi_2,\Pi_0\}$ as a source for information. If for the Π_0 detection outcome, Bob guesses which state Alice sent him based on which state was more likely to have failed. In this case, Bob will make some errors, but will still obtain some information. It is clear that Bob succeeds in this strategy the most when the Π_0 channel produces the least uncertainty, which is the result calculated. Here, Bob treats the inconclusive outcomes and the conclusive outcomes in the same way, and he does not share with Alice when his outcome is inclusive to discard those result. Thus, this treatment is in the same spirit of the minimum error state discrimination strategy but not a unambiguous discrimination strategy as errors are permitted.

The mutual information for a truly unambiguous channel, however, has to take into fully consideration that only error-free messages are taken into account. The outcome is conclusive with probability P_s, and the outcome is inconclusive with probability Q. Hence, after discarding the inconclusive outcomes, the mutual information for this unambiguous state discrimination channel is

$$I_{\mathrm{USD}}(A:B)=P_s\Big[H(X_c)-\underbrace{H(X_c|Y_c)}_{=0}\Big]=P_sH(X_c)\,, \tag{4.40}$$

where X_c and Y_c denotes the messages of the sender and the receiver for conclusive outcomes, respectively. $H(X_c|Y_c) = 0$ because there is no uncertainty among conclusive outcomes (i.e., $X_c = Y_c$). The prior probability for Alice's message X_c is given by the confidence probabilities $\{C_{s,1}, 1 - C_{s,1}\}$ corresponding to states $\{|\psi_1\rangle, |\psi_2\rangle\}$, where

$$C_{s,1} = \frac{\eta_1(1 - q_1)}{P_s} \,. \tag{4.41}$$

Hence, the correct expression of mutual information for USD is

$$I_{\mathrm{USD}}(A : B) = P_s H(X_c) = P_s H(C_{s,1}) \,.$$

With this expression, it is clear that if Bob restricts himself to only gaining information from error-free results, the amount of information gained by the boundary solutions, i.e., when either q_1 or q_2 are set to 1, is zero.

The fundamental difference between the mutual information $I(A{:}B)$ of Eq. (4.38) and $I_{\mathrm{USD}}(A{:}B)$ of Eq. (4.40) comes from Bob sharing the classical information of whether his measurement outcome is conclusive. Upon having this classical information, the mutual information of this quantum communication channel is reduced to $I_{\mathrm{USD}}(A{:}B)$ even if we take into account all of the measurement outcomes including the inconclusive ones. Alice's Shannon entropy can be divided into the uncertainty coming from the conclusive outcomes and the uncertainty coming from the inconclusive ones, i.e., $H_{\mathrm{USD}}(X) = P_{\mathrm{s}} H(X_c) + Q H(X_{inc})$. The conditional entropy, $H(X|Y) = P_s \cdot 0 + Q H(X|\Pi_0) = Q H(X_{inc})$. Thus, the mutual information given by the difference is $I_{\mathrm{USD}}(A : B) = P_{\mathrm{s}} H(X_c)$. This shows that, although Bob can obtain information from the inconclusive outcomes, this part of the information is already shared between everybody including Alice or an eavesdropper through classical communication and not through quantum communication.

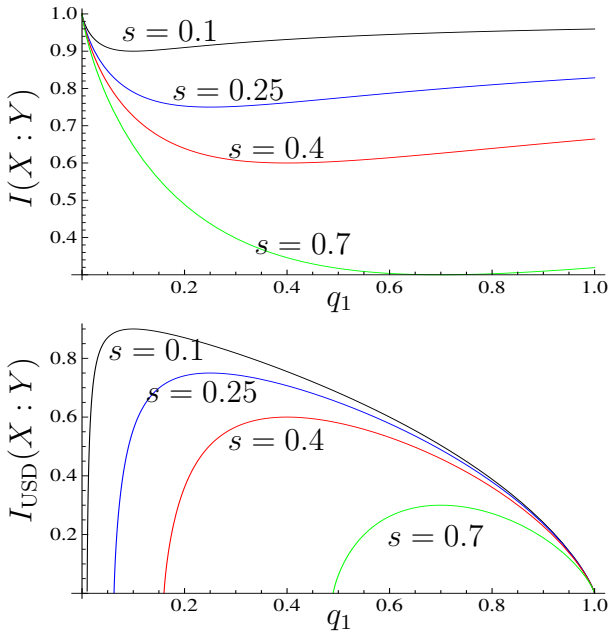

Figure 4.5: For equal prior probability distribution, $\eta_1 = \eta_2 = 1/2$, plots of the mutual information $I(A{:}B)$ of Eq. (4.38) (upper) and $I_{\rm USD}(A{:}B)$ of Eq. (4.40) (lower) as a function of q_1. Each of the quantities are plotted for four different value of $s = \langle\psi_1|\psi_2\rangle$. We choose $q_2 = s^2/q_1$ for the optimal USD that no information is left in the post-measurement states.

For $\eta_1 = \eta_2$, the maximum of $I_{\rm USD}(A{:}B)$ and the minimum of $I(A{:}B)$ occurs at the nontrivial solution for the optimization of the probability of success.

4.4.2 Optimization of the mutual information

Upon using the constraint $q_2 = s^2/q_1$ for the optimal USD, the mutual information $I_{\rm USD}(X:Y)$ can be written as a function of a single parameter q_1. It is optimized when its derivative with respect to q_1 vanishes, i.e.,

$$\frac{d}{dq_1} I_{\rm USD} = \eta_1 \log_2 \frac{\eta_1(1-q_1)}{P_s} - \eta_2 \frac{s^2}{q_1^2} \log_2 \frac{\eta_2(1-q_2)}{P_s} = 0\,. \tag{4.42}$$

The above equation has a simple solution $q_1 = q_2 = s$ for the case of equal priors, which is same as the local maximum solution for the success probability P_s. Since the mutual

information is a concave function within the physical range of the parameter $s^2 \leq q_1 \leq 1$, the solution of (4.42) maximizes the mutual information $I_{\rm USD}(A:B)$. For $\eta_1 = \eta_2 = 1/2$, $H(X_c) = 1$ is maximized and P_s is at its local extrema when $q_1 = s\sqrt{\eta_2/\eta_1} = s$. Thus, $q_1 = q_2 = s$ must be the solution for equal priors and the optimal mutual information is $I_{\rm USD}(A:B) = 1 - s$. For the case of unequal priors $\eta_1 \neq \eta_2$, however, we are not able to solve the equation analytically and have to rely on numerical methods; see Fig. 4.6.

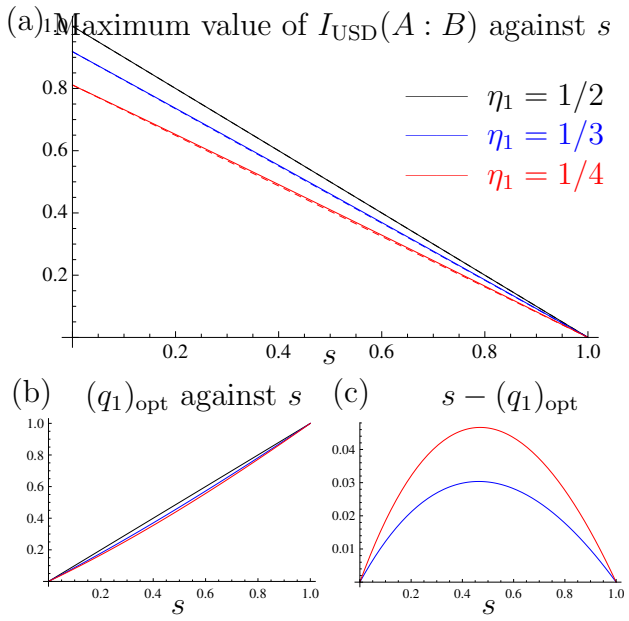

Figure 4.6: (a) The solid curves give the upper bounds of the mutual information $I_{\rm USD}(A:B)$ for the unambiguous communication channel between Alice and Bob as a function of the overlap of states $s = \langle\psi_1|\psi_2\rangle$ for three different sets of prior probability distributions. The dashed lines are plotted for the mutual information given in Eq. (4.43) with $q_1 = s$. (b) The plot of the values of q_1 that optimize $I_{\rm USD}(A:B)$ as a function of s. (c) The difference between the optimal value of q_1 and $q_1 = s$. For all these subfigures, the black, blue and red curves correspond to the cases for $\eta_1 = 1/2$, $\eta_1 = 1/3$ and $\eta_1 = 1/4$, respectively.

Figure 4.6(a) shows the s-dependence of the mutual information $I_{\rm USD}(A:B)$. For equal prior probabilities, we have obtained the upper bound analytically and it is a linear function in s, i.e., $I_{\rm USD}(A:B) \leq 1 - s$. For $\eta_1 \neq \eta_2$, the s-dependence of its upper bound is almost linear but not exactly. Fig. 4.6(b) and (c) show that the value of q_1 that optimizes

$I_{\rm USD}(A:B)$ depends on the prior distributions. This dependence on priors is, however, quite weak. The difference between $(q_1)_{\rm opt}$ for any arbitrary priors and $(q_1)_{\rm opt} = s$ for equal priors is largest at $s = 1/2$ and it is symmetric as s gets larger or smaller. The dashed curve shows the approximated upper bound of the mutual information given by $q_1 = s$, i.e.,

$$I_{\rm USD}(A:B)(q_1 = s) = (1-s)H(\eta_1)\,. \tag{4.43}$$

Thus, we can conclude that for equal priors, the mutual information for unambiguous state discrimination is optimized exact by the local extremal of the probability of success obtained at $q_1 = q_2 = s$. For unequal priors, the mutual information is not given by the local extremal point of the probability of success ($q_1 = s\sqrt{\eta_1/\eta_2}$). Instead, the optimal value of mutual information is extremely close to the value given also by $q_1 = q_2 = s$, that is $I_{\rm USD}(A:B) = (1-s)H(\eta_1)$.

4.4.3 The sequential measurement scheme

For the sequential measurement scheme discussed in the previous section, Bob's probability of success to correctly identify Alice's message is $P_{\rm sb} = \eta_1(1-q_{1b}) + \eta_2(1-q_{2b})$ and the probability for Charlie to correctly identify Alice's message is $P_{\rm sc} = \eta_1(1-q_{1c}) + \eta_2(1-q_{2c})$. Taking into account only conclusive outcomes, the mutual information for the communication channel between Alice and Bob and the channel between Alice and Charlie are, respectively,

$$I_{\rm USD}(A:B) = P_{\rm sb} H\left(\frac{\eta_1(1-q_{1\rm b})}{P_{\rm sb}}\right), \tag{4.44}$$

$$I_{\rm USD}(A:C) = P_{\rm sc} H\left(\frac{\eta_1(1-q_{1\rm c})}{P_{\rm sc}}\right). \tag{4.45}$$

The mutual information between the two receivers Bob and Charlie relies on the case where both of them have successfully identified Alice's message, such that they also share the same

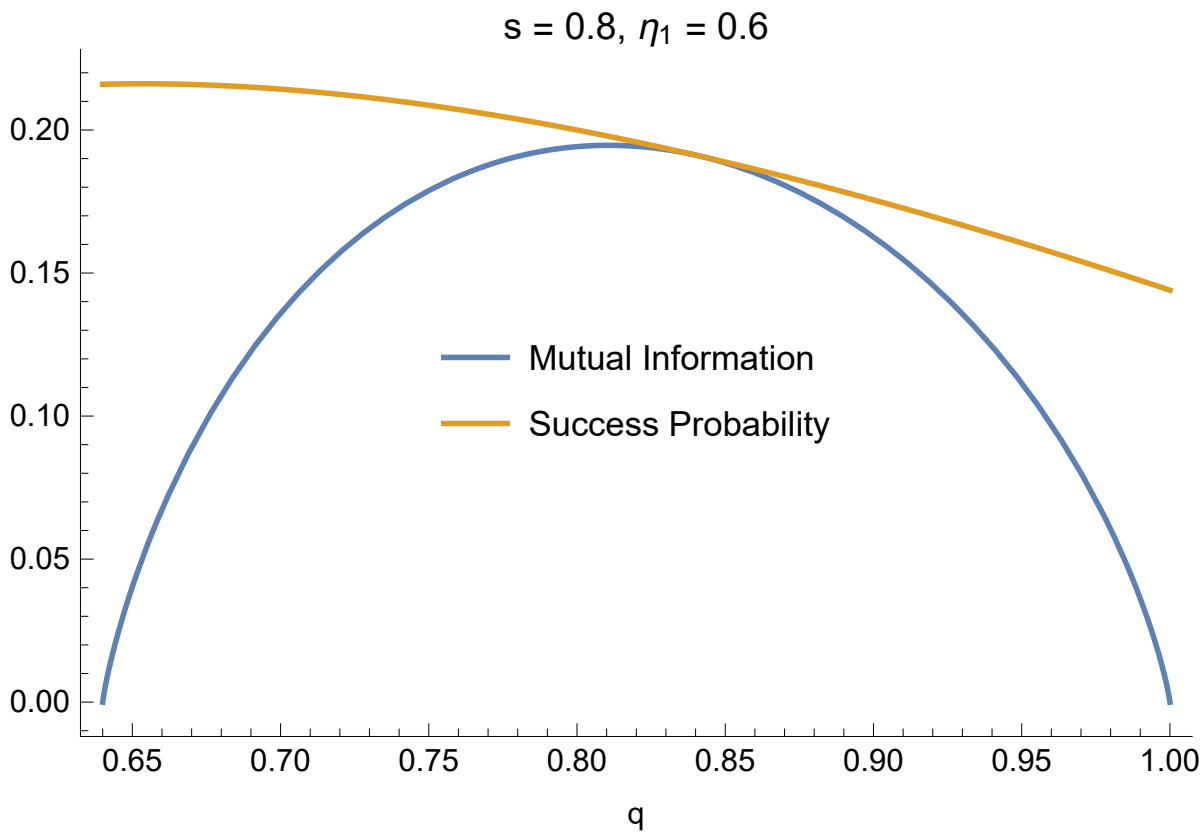

Figure 4.7: Mutual Information and Success Probability are plotted together on one graph for $s = 0.8$ and $\eta_1 = 0.6$. This graph illustrates that while the maximum of the Success probability occurs when $q_1 = s^2$ and $q_2 = 1$, the mutual information for these values is zero.

bit of unambiguous information. This joint probabilities of success $P_{\rm ss} = \eta_1 p_{1\rm b} p_{1\rm c} + \eta_2 p_{2\rm b} p_{2\rm c} = \eta_1(1-q_{1b})(1-q_{1c}) + \eta_2(1-q_{2b})(1-q_{2c})$ given in Eq. (4.8) has been studied in detail in the previous section. Thus, the unambiguous communication channel among all three parties, Alive, Bob and Charlie, can be characterized by the mutual information between Bob and Charlie,

$$I_{\rm USD}(B:C) = P_{\rm ss} H\left(\frac{\eta_1(1-q_{1\rm b})(1-q_{1\rm c})}{P_{\rm ss}}\right). \tag{4.46}$$

This joint mutual information is maximized when the information extracted by Bob and Charlie is symmetric, which requires $p_{1b} = p_{1c}$, $p_{2b} = p_{2c}$ and $t = \sqrt{s}$. This can be shown by setting $\frac{\partial}{\partial t} P_{ss} = 0$, which leads to $s^2 q_{1c} = t^4 q_{1b}$. (One may note that there is another other optimal solution at $P_{ss} = \eta_2 p_{2b} p_{2c}$, but this is a minimum and at this solution, $I_{USD}\,(B:C) = 0$.) Upon inserting these conditions in (4.46), we have

$$I_{\rm USD}(B:C) = P_{\rm ss} H\left(\frac{\eta_1(1-q_{1b})^2}{P_{\rm ss}}\right), \tag{4.47}$$

where $P_{\rm ss}$ is given by Eq. (4.20) and the information is symmetrically distributed between Bob and Charlie.

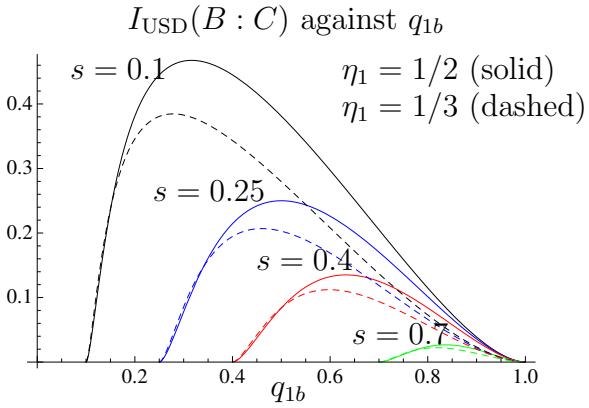

Figure 4.8: The mutual information between Bob and Charlie through their unambiguous sequential state discrimination scheme $I_{\rm USD}(B:C)$ given by Eq. (4.47) as a function of q_{1b} for different values of s.

The solid curves in Fig. 4.8 illustrate how $I_{\rm USD}$ depends on $q_{1b}(=q_{1c})$ for different values of s for the case of equal priors $\eta_1 = \eta_2$. Analogous to the optimization of mutual information between Alice and Bob, $I_{\rm USD}(B:C)$ is optimized by the same solution $p_{1b} = p_{2b} = 1-\sqrt{s}$ that gives a local extremal for the joint probability of success $P_{\rm ss}$. Hence, the upper bound of mutual information for equal priors is

$$I_{\rm USD}(B:C) \leq (1-\sqrt{s})^2 \quad \text{for} \quad \eta_1 = \eta_2 = \frac{1}{2}\,. \tag{4.48}$$

For the optimal solution, the frequencies of having bit 0 and bit 1 are the same among the unambiguous message shared between Bob and Charlie. Obviously, it is also shown that there is no information transmitted through this quantum communication channel at the boundary solutions $q_{1b} = 1$ or $q_{1b} = s^2$, where $P_{\rm ss}$ can be maximized. It is because at the boundary solution, only one type of bit can be sent and no useful information is effectively communicated through the quantum channel.

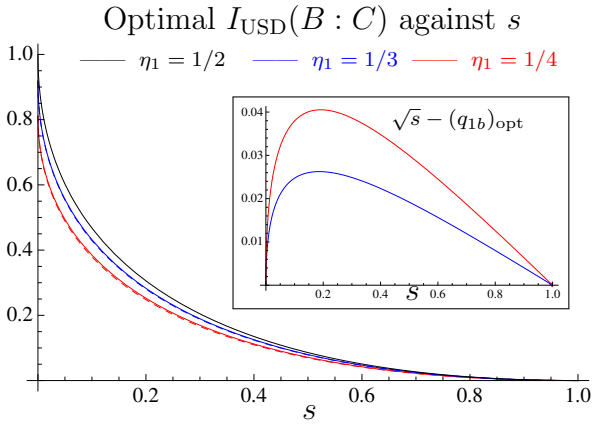

Figure 4.9: The upper bounds of the mutual information between Bob and Charlie $I_{\rm USD}(B:C)$ with respect to the overlap of signal states s are given by the solid curves for different prior distributions. The dashed curves show the mutual information $I_{\rm USD}(B:C) = (1-\sqrt{s})^2 H(\eta_1)$ obtained for $q_{1b} = q_{1c} = \sqrt{s}$. The optimal value of q_{1b} given the upper bound of the mutual information is $q_{1b} = \sqrt{s}$ for equal priors, and it is very close to this value even when the priors are biased. Their difference $\sqrt{s} - (q_{1b})_{\rm opt}$ is shown by the insert plot as a function of s.

For unequal priors, the mutual information depends on q_{1b} in a similar way as it does for equal priors; see Fig. 4.8. For $\eta_1 \neq \eta_2$, however, we no longer have a simple close analytical upper bound of the mutual information, instead, we can obtain it with numerical methods. The optimal value sof the mutual information $I_{\mathrm{USD}}(B : C)$ for different prior distributions are shown in Fig. 4.9. For equal priors, we have $I_{\mathrm{USD}}(B : C) \leq (1-\sqrt{s})H(\frac{1}{2}) = (1-\sqrt{s})^2$ with the optimal obtained at $q_{1b} = \sqrt{s}$. For the case of unequal priors, the optimal $I_{\mathrm{USD}}(B : C)$ is only slightly larger than its value obtained at $q_{1b} = \sqrt{s}$, which is $I_{\mathrm{USD}}(B : C) = (1-\sqrt{s})^2 H(\eta_1)$. Clearly, the mutual information and the joint probability of success are not optimized simultaneously with the same set of parameter values.

Although the flip-flop measurement decreases the success probability, it enables useful and unambiguous information to be transmitted through the communication channel using only von Neumann measurements. Given by Eq. (4.40), the unambiguous mutual information depends on the probability of success and the Shannon entropy of the conclusive outcomes. Thus, the mutual information of the flip-flop measurement is

$$\begin{aligned} I_{\mathrm{USD}}(A : B) &= P_s H\left(\frac{\eta_1 c(1-s^2)}{P_s}\right) \\ &= [\eta_1 c + \eta_2(1-c)](1-s^2) H\left(\frac{\eta_1 c}{\eta_1 c + \eta_2 (1-c)}\right). \end{aligned} \tag{4.49}$$

It is interesting to note that, for $c = \eta_2$, this expression reduces to $2\eta_1\eta_2(1-s^2)$. Additionally, given that the FFM is simply a subspace of the POVM space, no FFM can achieve greater unambiguous mutual information than the calculated maximum mutual information.

4.4.4 Three-party communication

Using the results from the end of the previous section, Eqs. (4.27)-(4.35), the mutual information between Bob and Charlie can be calculated as:

$$\begin{aligned} I(B:C) &= (1-\frac{s^2}{t^2})\left(1-t^2\right)\left(\eta_1 c^2+\eta_2\left(1-c\right)^2\right) \\ &\quad *H\left(\frac{\eta_1 c^2}{\eta_1 c^2+\eta_2\left(1-c\right)^2}\right). \end{aligned} \tag{4.50}$$

This is clearly minimum (=0) when $c = 0$ or $c = 1$, i.e., at the boundaries of the allowed range for the flipping rate. It should be noted that in this expression, the value of t is fixed by $\sqrt{q_1 q_2} = \frac{s}{t}$. This shows that in this case, no unambiguous information is transferred between Bob and Charlie.

It is interesting to note, however, that for $c = \eta_2$ and $t^2 = s$, we obtain

$$I(B:C) = \eta_1\eta_2(1-s)^2 H\left(\eta_2\right) , \tag{4.51}$$

which is the expression we expect for transmitting unambiguous information between Bob and Charlie that us useful for establishing a quantum communication protocol between them.

Chapter 5

Sequential Discrimination with a Fixed Rate of Inconclusive Outcome

5.1 Sequential Discrimination with Fixed Rate Inconclusive Outcome

The natural continuation of the sequential minimum error and the sequential unambiguous discrimination strategies is sequential discrimination with a fixed rate of inconclusive outcome (SFRIO). Similar to the FRIO strategy, the goal is to have Both bob and Charlie optimize their joint probability of success for a fixed parameter of the probability of obtaining an inconclusive outcome. By parameterizing the rate of inconclusive outcome, we can connect the SME strategy, with no inconclusive outcomes, to the SUD outcome, with maximal rate of inconclusive outcomes. More precisely, Alice prepares $\{|\psi_i\rangle\}$ with probabilities $\{\eta_i\}$. Bob performs a POVM Π_b with three outcomes, $\Pi_{0b}, \Pi_{1b}, \Pi_{2b}$. If either of the Π_{1b} or Π_{2b} detectors click, Bob concludes that Alice sent the associated state. If the Π_{0b} detector clicks, with probability $Q_b = tr\left(\Pi_{0b}\left(\eta_1 |\psi_1\rangle \langle\psi_1| + \eta_2 |\psi_2\rangle \langle\psi_2|\right)\right)$, then Bob's result is incon-

clusive. Bob's post measurement states, $\{\rho_{ib}|i=1,2\}$ are sent to Charlie who can apply a similar measurement scheme. If Bob restricts his measurement so that his post measurement states are pure, Bob and Charlie's POVMs can be represented via the Neumark formalism as follows:

$$\begin{aligned}
U_b\left|\psi_1\right\rangle\left|i\right\rangle &= \sqrt{p_{1b}}\left|\theta_1\right\rangle\left|1\right\rangle + \sqrt{r_{1b}}\left|\theta_1\right\rangle\left|2\right\rangle + \sqrt{q_{1b}}\left|\theta_1\right\rangle\left|0\right\rangle, \\
U_b\left|\psi_2\right\rangle\left|i\right\rangle &= \sqrt{r_{2b}}\left|\theta_2\right\rangle\left|1\right\rangle + \sqrt{p_{2b}}\left|\theta_2\right\rangle\left|2\right\rangle + \sqrt{q_{2b}}\left|\theta_2\right\rangle\left|0\right\rangle, \\
U_c\left|\theta_1\right\rangle\left|i\right\rangle &= \sqrt{p_{1c}}\left|\varphi\right\rangle\left|1\right\rangle + \sqrt{r_{1c}}\left|\varphi\right\rangle\left|2\right\rangle + \sqrt{q_{1c}}\left|\varphi\right\rangle\left|0\right\rangle, \\
U_c\left|\theta_2\right\rangle\left|i\right\rangle &= \sqrt{r_{2c}}\left|\varphi\right\rangle\left|1\right\rangle + \sqrt{p_{2c}}\left|\varphi\right\rangle\left|2\right\rangle + \sqrt{q_{2c}}\left|\varphi\right\rangle\left|0\right\rangle.
\end{aligned}$$

Optimizing the joint success means maximizing:

$$P_{ss} = \eta_1 p_{1b} p_{1c} + \eta_2 p_{2b} p_{2c},$$

subject to the constraints:

$$\begin{aligned}
\frac{s}{t} &= \sqrt{p_{1b}r_{2b}} + \sqrt{p_{2b}r_{1b}} + \sqrt{q_{1b}q_{2b}}, \\
t &= \sqrt{p_{1c}r_{2c}} + \sqrt{p_{2c}r_{1c}} + \sqrt{q_{1c}q_{2c}}.
\end{aligned}$$

If we assume that there is a fixed rate of inconclusive errors, Q, for Bob and Charlie, then $Q_b = \eta_1 p_{1b} + \eta_2 p_{2b}$ and similarly for Charlie. As in the case of the FRIO strategy, we can employ a change of variables in order to reframe the problem:

$$\begin{aligned}
\tilde{p}_{ib} &= \frac{p_{ib}}{1-q_{ib}}, \quad \tilde{p}_{ic} = \frac{p_{ic}}{1-q_{ic}}, \\
\tilde{r}_{ib} &= \frac{r_{ib}}{1-q_{ib}}, \quad \tilde{r}_{ic} = \frac{r_{ic}}{1-q_{ic}} \\
\tilde{\eta}_i &= \frac{\eta_i\left(1-q_{ib}\right)\left(1-q_{ic}\right)}{C},
\end{aligned}$$

$$\begin{aligned} C &= \sum_i^2 \eta_i \left(1 - q_{ib}\right)\left(1 - q_{ic}\right). \\ \tilde{s} &= \frac{\frac{s}{t} - \sqrt{q_{1b}q_{2b}}}{\sqrt{(1-q_{1b})(1-q_{2b})}}, \quad \tilde{t} = \frac{t - \sqrt{q_{1c}q_{2c}}}{\sqrt{(1-q_{1c})(1-q_{2c})}}. \end{aligned}$$

Using these variables, the problem becomes maximizing:

$$P_{ss} = C\left(\tilde{\eta}_1\tilde{p}_{1b}\tilde{p}_{1c} + \tilde{\eta}_2\tilde{p}_{2b}\tilde{p}_{2c}\right), \tag{5.1}$$

with respect to the following constraints:

$$\begin{aligned} \tilde{s} &= \sqrt{\tilde{p}_{1b}\tilde{r}_{2b}} + \sqrt{\tilde{p}_{2b}\tilde{r}_{1b}}, \\ \tilde{t} &= \sqrt{\tilde{p}_{1c}\tilde{r}_{2c}} + \sqrt{\tilde{p}_{2c}\tilde{r}_{1c}}, \\ C &= \sum_i^2 \eta_i \left(1 - q_{ib}\right)\left(1 - q_{ic}\right), \\ \tilde{p}_{ib} + \tilde{r}_{ib} &= 1 \; \tilde{p}_{ic} + \tilde{r}_{ic} = 1. \end{aligned}$$

This method of rewriting the problem allows it to be approached through the lens of an already understood optimization - the sequential minimum error problem.

5.2 Optimizing for Equal Priors

For equal priors, where $\eta_1 = \eta_2 = \frac{1}{2}$, an analytic solution can be found. Starting from Eq. (5.1) and the relevant constraints, one can use Lagrange multipliers to optimize as follows:

$$\begin{aligned} F &= P_{ss} + \lambda_1\left(\tilde{s} - \sqrt{\tilde{p}_{1b}\tilde{r}_{2b}} + \sqrt{\tilde{p}_{2b}\tilde{r}_{1b}}\right) + \lambda_2\left(\tilde{t} - \sqrt{\tilde{p}_{1c}\tilde{r}_{2c}} + \sqrt{\tilde{p}_{2c}\tilde{r}_{1c}}\right), \\ \frac{\partial F}{\partial \tilde{p}_{1b}} &= C\tilde{\eta}_1\tilde{p}_{1c} - \frac{\lambda_1}{2}\left(\frac{\sqrt{\tilde{r}_{2b}}}{\sqrt{\tilde{p}_{1b}}} - \frac{\sqrt{\tilde{p}_{2b}}}{\sqrt{\tilde{r}_{1b}}}\right) = 0, \\ \frac{\partial F}{\partial \tilde{p}_{2b}} &= C\tilde{\eta}_2\tilde{p}_{2c} - \frac{\lambda_1}{2}\left(\frac{\sqrt{\tilde{r}_{1b}}}{\sqrt{\tilde{p}_{2b}}} - \frac{\sqrt{\tilde{p}_{1b}}}{\sqrt{\tilde{r}_{2b}}}\right) = 0, \end{aligned}$$

$$\Rightarrow \quad C\tilde{\eta}_1\tilde{p}_{1c}\sqrt{\tilde{p}_{1b}\tilde{r}_{1b}} = C\tilde{\eta}_2\tilde{p}_{2c}\sqrt{\tilde{p}_{2b}\tilde{r}_{2b}},$$

$$\Rightarrow \quad \eta_1 p_{1c}\sqrt{p_{1b}r_{1b}} = \eta_2 p_{2c}\sqrt{p_{2b}r_{2c}}.$$

Applying the same method for Charlie, one obtains the two Langrange constraints of an optimal solution for $\eta_1 = \eta_2 = \frac{1}{2}$:

$$p_{1c}\sqrt{p_{1b}r_{1b}} = p_{2c}\sqrt{p_{2b}r_{2b}}, \tag{5.2}$$

$$p_{1b}\sqrt{p_{1c}r_{1c}} = p_{2b}\sqrt{p_{2c}r_{2c}}. \tag{5.3}$$

These two constraints can be easily satisfied by choosing $p_{1b} = p_{2b}$, $r_{1b} = r_{2b}$, $p_{1c} = p_{2c}$, and $r_{1c} = r_{2c}$. This choice then requires $q_{1b} = q_{2b} = Q_b$ and $q_{1c} = q_{2c} = Q_c$. Furthermore, this result ensures that $\tilde{\eta}_2 = \tilde{\eta}_2 = \frac{1}{2}$, $\tilde{p}_{2b} = \tilde{p}_{2b}$, and $\tilde{p}_{1c} = \tilde{p}_{2c}$. With some basic substitutions into the original constraint equations, one can derive that this result gives:

$$\begin{aligned}
\tilde{p}_{ib} &= \frac{1}{2}\left(1 + \sqrt{1 - \tilde{s}^2}\right), \\
\tilde{p}_{ic} &= \frac{1}{2}\left(1 + \sqrt{1 - \tilde{t}^2}\right), \\
P_{ss} &= \frac{(1 - Q_b)(1 - Q_c)}{4}\left(1 + \sqrt{1 - \left(\frac{\frac{s}{t} - Q_b}{1 - Q_b}\right)^2}\right)\left(1 + \sqrt{1 - \left(\frac{t - Q_c}{1 - Q_c}\right)^2}\right).
\end{aligned}$$

For $Q_b = Q_c$, this is easily optimized with respect to t to give $t = \sqrt{s}$. Furthermore, if, instead of fixing Q_b and Q_c individually, we fix a global parameter $Q_p = Q_b Q_c$, then, assuming $t = \sqrt{s}$, it is easy to show that the above equation is optimized in the case that $Q_b = Q_c$. Finally, if we plot P_{ss} as a function of the two unconstrained variables, t and Q_b, we can clearly see that this function is concave (see Figure 5.1). This concavity implies that this stable optimum of $t = \sqrt{s}$ and $Q_b = Q_c$ is a global optimum for the function, if we

constrain $Q_p = Q_b Q_c$. Defining $Q = Q_b = Q_c$, this gives the final result:

$$P_{ss} = \frac{(1-Q)^2}{4}\left(1+\sqrt{1-\left(\frac{\sqrt{s}-Q}{1-Q}\right)^2}\right)^2. \tag{5.4}$$

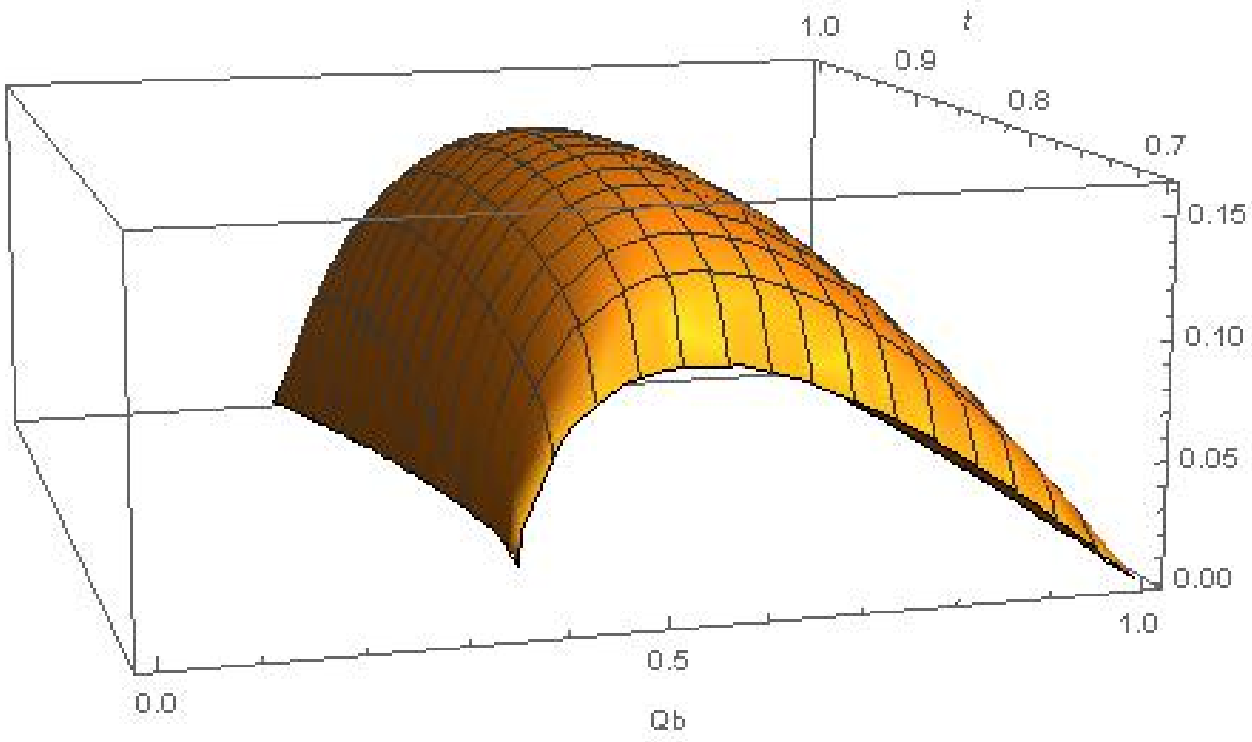

Figure 5.1: P_{ss} as a function of t and Q_b for $s = 0.7$ and $Q = 0.3$

5.3 Boundary Solutions

It is important to realize that the optimal solution derived in the above section is not universally optimal. One such region where the above solution is no longer valid can be derived from the fact that $\tilde{s} > 0$ and $\tilde{t} > 0$. In the above solution, the optimal solution derived relied on $q_{1b} = q_{2b} = Q_b$ and $q_{1c} = q_{2c} = Q_c$, meaning that the above solution

is only valid where both $\frac{s}{t} > Q_b$ and $t > Q_c$, or $s > Q_b Q_c$. If this condition is violated, then either $q_{1b} = q_{2b} = Q_b$ or $q_{1c} = q_{2b} = Q_c$ is no longer a valid solution. For instance, if we assume that $Q_c = q_{1c} = q_{2c}$ no longer holds, then we need to revisit the constraint $t = \sqrt{p_{1c} r_{2c}} + \sqrt{p_{2c} r_{1c}} + \sqrt{q_{1c} q_{2c}}$. In this case, the optimal solution is clearly found by setting $\sqrt{q_{1c} q_{2c}} = t$, minimizing the effective overlap $\tilde{t}$. The new constraint $0 = \sqrt{p_{1c} r_{2c}} + \sqrt{p_{2c} r_{1c}}$ trivially implies that we can choose $r_{1c} = r_{2c} = 0$. Using $t = \sqrt{q_{1c} q_{2c}}$ and $Q_c = \frac{1}{2}(q_{1c} + q_{2c})$, the final solution requires that $p_{1c} = \left(1 - Q_c \mp \sqrt{Q_c^2 - t^2}\right)$ and $p_{2c} = \left(1 - Q_c \pm \sqrt{Q_c^2 - t^2}\right)$. In this case, the optimization problem can be rewritten as maximizing:

$$P_{ss} = (1 - Q_b)(1 - Q_c)\left(\eta_1' \tilde{p}_{1b} + \eta_2' \tilde{p}_{2b}\right), \tag{5.5}$$

subject to:

$$\tilde{s} \equiv \frac{\frac{s}{t} - Q_b}{1 - Q_b} = \sqrt{\tilde{p}_{1b} \tilde{r}_{2b}} + \sqrt{\tilde{p}_{2b} \tilde{r}_{1b}}, \tag{5.6}$$

where $\eta_i' \equiv \frac{1}{2(1-Q_C)}\left(1 - Q_c \pm \sqrt{Q_c^2 - t^2}\right)$. Using the optimal solution for the standard ME problem, the optimal solution for this problem is:

$$P_{ss} = \frac{(1 - Q_b)(1 - Q_c)}{2}\left(1 + \sqrt{1 - \left(\frac{(1 - 2Q_c + t)\left(\frac{s}{t} - Q_b\right)^2}{(1 - Q_b)^2 (1 - Q_c)^2}\right)}\right). \tag{5.7}$$

Optimizing with respect to t gives:

$$\frac{\partial P_{ss}}{\partial t} = 0 = \frac{(1 - Q_b)(1 - Q_c)}{2\left(\sqrt{1 - \left(\frac{(1-2Q_c+t^2)\left(\frac{s}{t} - Q_b\right)^2}{(1-Q_b)^2(1-Q_c)^2}\right)}\right)} \frac{\left(\frac{s}{t} - Q_b\right)\left(2t\left(\frac{s}{t} - Q_b\right)\right) - 2\frac{s}{t^2}(1 - 2Q_c + t^2)}{(1 - Q_b)^2 (1 - Q_c)^2}.$$

This gives an optimum point of $\frac{s}{t} = Q_b$, which makes sense, as this choice of t minimizes the effective overlap $\tilde{s}$. Clearly, for this solution, the optimum probability of success is

simply $P_{ssa} = (1 - Q_b)(1 - Q_c)$. In other words, for $Q_b Q_c > s$, both Bob and Charlie can setup their experiments so that they are guaranteed to succeed when they do not receive an inconclusive outcome.

Another boundary that needs to be considered is given by the solutions for which ignoring one of the incoming states is optimal, i.e. $p_{1b} = p_{1c} = 0$. In this region, the constraints become:

$$\frac{s}{t} = \sqrt{p_{2b} r_{1b}} + \sqrt{q_{1b} q_{2b}}, \tag{5.8}$$

$$t = \sqrt{p_{2c} r_{1c}} + \sqrt{q_{1c} q_{2c}}. \tag{5.9}$$

As there is no longer a constraint on r_{2b} and r_{2c}, it is obvious that these should be set to zero. Through some straightforward substitutions, these constraints can be rewritten as:

$$\frac{s}{t} = \sqrt{p_{2b}(2 - 2Q_b - p_{2b})} + \sqrt{(2Q_b + p_{2b} - 1)(1 - p_{2b})}, \tag{5.10}$$

$$t = \sqrt{p_{2c}(2 - 2Q_c - p_{2c})} + \sqrt{(2Q_c + p_{2c} - 1)(1 - p_{2c})}. \tag{5.11}$$

Through some straight forward algebra, these terms can be solved in terms of p_{2b} and p_{2c}:

$$p_{2b} = (1 - Q_b)\left(1 + \sqrt{1 - \left(\frac{\frac{s^2}{t^2} - 2Q_b + 1}{\left(2\frac{s}{t}\right) 1 - Q_b}\right)^2}\right) \tag{5.12}$$

$$p_{2c} = (1 - Q_c)\left(1 + \sqrt{1 - \left(\frac{t^2 - 2Q_c + 1}{(2t) 1 - Q_c}\right)^2}\right) \tag{5.13}$$

For $Q_b = Q_c \equiv Q$ it is trivial to show that the optimal solution occurs for $t = \sqrt{s}$. This

gives the joint probability of success:

$$P_{ssb} = \frac{(1-Q)^2}{2}\left(1+\sqrt{1-\left(\frac{s-2Q+1}{2\sqrt{s}\,(1-Q)}\right)^2}\right)^2. \tag{5.14}$$

Before analyzing the regime in which this boundary solution is optimal, it is helpful to first consider one final boundary. This final boundary exists in the regime where prioritizing one of the incoming states is optimal, i.e. $p_{1b} = p_{1c} = 1$. In this region, the constraints become:

$$\frac{s}{t} = \sqrt{r_{2b}}, \quad t = \sqrt{r_{2c}}.$$

Since $p_{1c} = p_{1c} = 1$, this requires that $q_{1b} = q_{1c} = 0$, and therefore $q_{2b} = 2Q_b$ and $q_{2c} = 2Q_c$. Putting this all together, it can easily be derived that the probability of success is:

$$P_{ss} = \frac{1}{2}\left\{1+\left[1+\left(1-\frac{s^2}{t^2}-2Q_b\right)\left(1-t^2-2Q_c\right)\right]\right\}. \tag{5.15}$$

For $Q_b = Q_c = Q$, this can be optimized with respect to t, giving $t = \sqrt{s}$:

$$P_{ssc} = \frac{1}{2}\left(1+(1-s-2Q)^2\right) \quad Q < \frac{1-s}{2}. \tag{5.16}$$

The final condition $Q < \frac{1-s}{2}$, comes from the requirement that $p_2 > 0$.

Putting all this together, we arrive at the solution to this problem:

$$P_{ss} = \begin{cases} (1-Q)^2 & Q \geq b_a \\ \frac{(1-Q)^2}{2}\left(1+\sqrt{1-\left(\frac{2-2Q+1}{2\sqrt{s}(1-Q)}\right)^2}\right)^2 & b_a > Q > \frac{1-s}{2} \\ \frac{1}{2}\left(1+(1-s-2Q)^2\right) & \frac{1-s}{2} \geq Q > b_b \\ \frac{(1-Q)^2}{4}\left(1+\sqrt{1-\left(\frac{\sqrt{s}-Q}{1-Q}\right)^2}\right)^2 & b_b \geq Q \end{cases}, \tag{5.17}$$

where b_a and b_b are numerically solved boundaries. It is important to note that this full solution should be used cautiously. For values of $Q \geq b_b$, the solution is valid, but artificial due to the forced constraint of defining a Q. In this regime, Q exceeds the value necessary to achieve unambiguous discrimination, thus the extra probability that the inconclusive detector click is superfluous. The purpose of this full solution is that it helps to connect the minimum error solution at $Q = 0$ to the unambiguous discrimination solution at $Q = \sqrt{s}$.

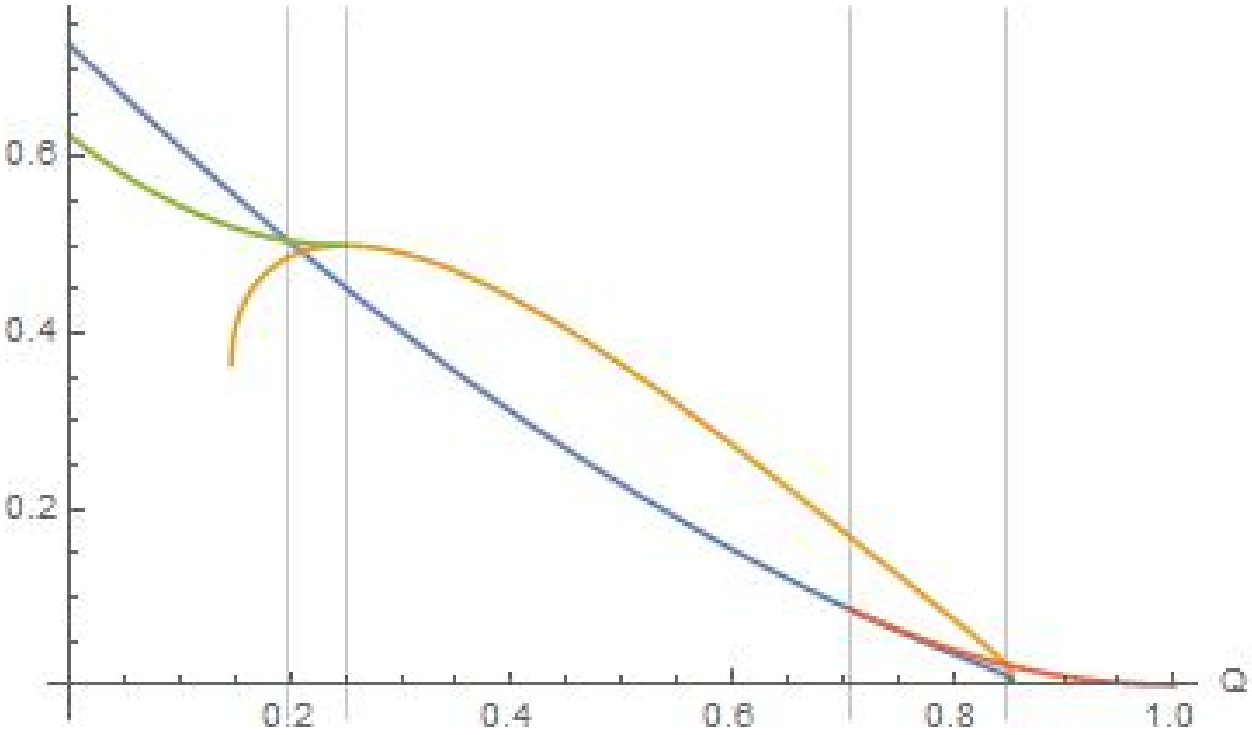

Figure 5.2: A plot of all the possible solutions as a function of Q for $s = 0.5$. The gridlines are at Q = 0.197871 (b_b), 0.25 $\left(\frac{1-s}{2}\right)$, 0.707107 ($\sqrt{s}$), 0.847894($b_a$), respectively. The graphs of $P_{ss}, P_{ssa}, P_{ssb}, P_{ssc}$ are represented by the blue, red, yellow, and green lines respectively.

Chapter 6

Quantum Retrodiction

One final topic that will be discussed by this dissertation is quantum retrodiction. This topic does not explicitly connect to sequential discrimination between non-orthogonal states. However, the result that will be derived will show that the formalism of quantum retrodiction gives a new perspective and method for approaching questions such as state discrimination. Here we focus on applying quantum retrodiction only to the standard unambiguous discrimination problem, but this method suggests that the same approach can be extended to the sequential discrimination problem.

6.1 Introduction to Quantum Retrodiction

In [54], Barnett, Pegg, and Jeffers introduce the concept of quantum retrodiction and derive their formalism using Bayes theorem, building of work done by Aharonov in [55–57]. In order to understand their proposal, it is necessary to review the standard formulation of a measurement scheme. The standard description of a measurement one typically considers two parties, Alice and Bob, where Alice prepares the states $\{\rho_i\}$ with the prior probabilities η_i and Bob measures the states with some detectors $\{\Pi_j\}$. With this setup, Alice can predict

what measurement outcome Bob will get for any given state that she sends:

$$P\left(b_j|a_i\right) = Tr\left(\Pi_j \rho_i\right). \tag{6.1}$$

The essence of the idea proposed by Barnett, et al, is that we can also imagine the measurement from Bob's perspective. From Bob's perspective, he has a set of measurement detectors $\{\Pi_j\}$ that have detected results with some probabilities μ_j caused by some set of states sent by Alice, ρ_i. At this point, Bob can determine, or retrodict, the probability that, given that a detector of his clicked, the odds that Alice sent a particular state. Using Bayes theorem, we can calculate such a probability:

$$\begin{aligned} P\left(a_i|b_j\right) &= \frac{P\left(b_j|a_i\right)P\left(a_i\right)}{P\left(b_j\right)}, \\ &= \frac{Tr\left(\Pi_j\rho_i\right)\eta_i}{\sum_i Tr\left(\Pi_j\rho_i\right)\eta_i}. \end{aligned} \tag{6.2}$$

At this point, Barnett, et al, note that in the case that $\sum_i^D \eta_i\rho_i = \frac{I}{D}$, the formalism of the retrodictive approach can be massaged in order to be defined symmetrically with the predictive probability. By defining $\rho_j^{ret} \equiv \frac{\Pi_j}{Tr(\Pi_j)}$ and $\Pi_i^{ret} \equiv D\eta_i\rho_i$ we can determine the retrodictive probability as:

$$P\left(a_i|b_j\right) = Tr\left(\Pi_i^{ret}\rho_j^{ret}\right). \tag{6.3}$$

The theory proposed by the Barnett paper provides an extremely useful foundation for the retrodictive approach. For more on the fundamentals of quantum retrodiction and its applications, see [58–62]. Up until the following result, the general consensus in the current theory of quantum retrodiciton is that there is no way to find a general definition for the retrodictive states and operators so that the retrodictive probability is always symmetrically defined to the predictive probability. This conclusion is surprising, as the symmetric forms of

Eqs. (6.1) and (6.3), hint that the retrodictive approach can be viewed as a dual problem to the predictive approach. If this is the case, then the formalism should be entirely symmetric, and not only for the specific case of an unbiased source ($\sum_i^D \eta_i \rho_i = \frac{I}{D}$). In fact, with some simple adjustments, we can indeed see that this is the case. First we can use the following definitions:

$$\begin{aligned} \Omega &\equiv \sum_i \eta_i \rho_i, && (6.4) \\ \Pi_i^{ret} &\equiv \Omega^{-\frac{1}{2}} \eta_i \rho_i \Omega^{-\frac{1}{2}}, && (6.5) \\ \mu_j &\equiv \sum_i Tr\left(\Pi_j \rho_i \eta_i\right), && (6.6) \\ \rho_j^{ret} &\equiv \frac{\sqrt{\Omega}\Pi_j\sqrt{\Omega}}{\mu_j}. && (6.7) \end{aligned}$$

First, it is straightforward to see that the previous definitions are a special case of the above definitions. Additionally, using these definitions we can define the dual probabilities of the retrodictive problem without making any assumptions about Alice's source:

$$P\left(a_i|b_j\right) = Tr\left(\Pi_i^{ret}\rho_j^{ret}\right). \quad (6.8)$$

It is worth noting that this definition of the retrodictive operators obey the expected constraints:

$$\begin{aligned} \sum_i \Pi_i^{ret} &= I, \\ tr\left(\rho_j^{ret}\right) &= 1. \end{aligned}$$

Additionally, this formalism also requires the following constraint:

$$\sum_i \eta_i \rho_i = \Omega = \sum_j \mu_j \rho_j^{ret}. \tag{6.9}$$

6.2 Applications of the Retrodictive Formalism

6.2.1 Unambiguous Discrimination in the Retrodictive Formalism

One very useful place to start for looking at understanding and applying the retrodictive formalism is the problem of unambiguous state discrimination. In the standard unambiguous discrimination problem, introduced by Ivanovic [35],Peres [37], and Dieks [36], Alice has a set of pure states $\{|\psi_1\rangle, |\psi_2\rangle\}$ that she sends to Bob with the respective probabilities $\{\eta_1, \eta_2\}$. Bob's task is to discriminate between these states using the Positive Operator Valued Measurement (POVM) defined by $\Pi = \{\Pi_1, \Pi_2, \Pi_0\}$ subject to the standard constraints: $\sum_i \Pi_i = I$, $\Pi_i > 0$. By ensuring that the Π_i detector will only click if the $|\psi_i\rangle$ for $i \in (1, 2)$ or, in other words $\langle\psi_i|\Pi_j|\psi_i\rangle = 0$ for $i \neq j$, then Bob can ensure that, when he gets a detection in either the Π_1 or Π_2 detectors, his measurement is unambiguous. The goal in UD is to maximize the average probability of Bob's success:

$$P_s = \eta_1 tr\left(\Pi_1 |\psi_1\rangle \langle\psi_1|\right) + \eta_2 tr\left(\Pi_2 |\psi_2\rangle \langle\psi_2|\right). \tag{6.10}$$

While the optimal success probability is well known, let us reformulate this problem using the retrodictive formalism. Doing so will allow us to formulate a dual problem, through which we will be able to derive the optimal success probability and highlight some important features of the retrodictive formalism. Using the definitions posed above in Eqs. (6.4) - (6.7), we can define the following dual problem: Given a set of measurements $\{\Pi_1^{ret}, \Pi_2^{ret}\}$ and matrix Ω, what is the set of states $\rho_1^{ret}, \rho_2^{ret}, \rho_0^{ret}$ that satisfies the constraint that $\mu_1\rho_1^{ret} + \mu_2\rho_2^{ret} + \mu_0\rho_0^{ret} =$

Ω for which this set of measurements is optimal. In this formalism the quantity to be optimized is:

$$P_s = \mu_1 tr\left(\Pi_1^{ret}\rho_1^{ret}\right) + \mu_2 tr\left(\Pi_2^{ret}\rho_2^{ret}\right). \tag{6.11}$$

. This is should look familiar to Eq. (6.10) as $P(a,b) = P(b,a)$. For the retrodictive formalism of the UD problem, there are also the additional constraints that $\mu_2 tr\left(\Pi_1^{ret}\rho_2^{ret}\right) = \mu_1 tr\left(\Pi_2^{ret}\rho_1^{ret}\right) = 0$, mirroring the fact that $\eta_1 tr\left(\Pi_2\rho_1\right) = \eta_2 tr\left(\Pi_1\rho_2\right) = 0$.

The first feature that we can exploit to optimize this problem is the normalization condition on the retrodictive detectors: $\Pi_1^{ret} + \Pi_2^{ret} = I$. Without losing any generality, we can assume that we are working within a 2-D Hilbert space. In combination with Eq. (6.5), this gives that $\Pi_i^{ret} = |\phi_i^{ret}\rangle\langle\phi_i^{ret}|$, where $\langle\phi_1^{ret}|\phi_2^{ret}\rangle = 0$. This effectively implies that the states $|\phi_i^{ret}\rangle = \Omega^{-\frac{1}{2}}\sqrt{\eta_i}\,|\psi_i\rangle$ are normalized and orthogonal. We can also work out this result explicitly by starting with the following definitions:

$$\begin{aligned}
|\psi_1\rangle &= cos(\theta)\,|0\rangle + sin(\theta)\,|1\rangle, \\
|\psi_2\rangle &= cos(\theta)\,|0\rangle - sin(\theta)\,|1\rangle, \\
\Omega &= \eta_1\,|\psi_1\rangle\langle\psi_1| + \eta_2\,|\psi_2\rangle\langle\psi_2|, \\
&= w_1\,|\omega_1\rangle\langle\omega_1| + w_2\,|\omega_2\rangle\langle\omega_2|.
\end{aligned}$$

In the final equation above, $\langle\omega_1|\omega_2\rangle = 0$. First we can start by deriving the eigenvalues $\{w_1, w_2\}$ of the matrix Ω as follows:

$$\begin{aligned}
\Omega &= \begin{pmatrix} cos^2(\theta) & (\eta_1-\eta_2)\,sin(\theta)\,cos(\theta) \\ (\eta_1-\eta_2)\,sin(\theta)\,cos(\theta) & sin^2(\theta) \end{pmatrix}, \\
&\Rightarrow \quad w^2 - w + cos^2(\theta)\,sin^2(\theta) - (\eta_1-\eta_2)^2\,cos^2(\theta)\,sin^2(\theta) = 0,
\end{aligned}$$

$$\Rightarrow \quad w_{1,2} = \frac{1}{2}\left(1 \pm \sqrt{1 - 4\eta_1\eta_2 sin^2(2\theta)}\right).$$

If we define the eigenvectors of this matrix as $|\omega_1\rangle = cos(\omega)|0\rangle + sin(\omega)|1\rangle$ and $|\omega_2\rangle = -sin(\omega)|0\rangle + cos(\omega)|1\rangle$, then we can utilize the following transformation:

$$\begin{aligned} |\psi_1\rangle &= cos(\theta-\omega)|\omega_1\rangle + sin(\theta-\omega)|\omega_2\rangle, \\ |\psi_2\rangle &= cos(\theta+\omega)|\omega_1\rangle - sin(\theta+\omega)|\omega_2\rangle. \end{aligned}$$

Using this, we can now express the matrix Ω in the $|\omega_1\rangle, |\omega_2\rangle$ basis:

$$\begin{pmatrix} \eta_1 cos^2(\theta-\omega)\,\eta_2 cos^2(\theta+\omega) & \frac{\eta_1}{2}sin(2(\theta-\omega)) - \frac{\eta_2}{2}sin(2(\theta+\omega)) \\ \frac{\eta_1}{2}sin(2(\theta-\omega)) - \frac{\eta_2}{2}sin(2(\theta+\omega)) & \eta_1 sin^2(\theta-\omega)\,\eta_2 sin^2(\theta+\omega) \end{pmatrix}.$$

In order for this matrix to be diagonal in this basis, we must require that $\frac{\eta_1}{2}sin(2(\theta-\omega)) = \frac{\eta_2}{2}sin(2(\theta+\omega))$, or $tan(2\omega) = (\eta_1-\eta_2)\,tan(2\theta)$. With this we can now explicitly calculate $|\phi_1^{ret}\rangle = \Omega^{-\frac{1}{2}}\sqrt{\eta_1}|\psi_1\rangle$:

$$\begin{aligned} \Omega^{-\frac{1}{2}} &= \frac{1}{\sqrt{w_1}}|\omega_1\rangle\langle\omega_1| + \frac{1}{\sqrt{w_2}}|\omega_2\rangle\langle\omega_2|, \\ \Omega^{-\frac{1}{2}}\sqrt{\eta_1}|\psi_1\rangle &= \sqrt{\eta_1}\left(\frac{cos(\theta-\omega)}{\sqrt{w_1}}|\omega_1\rangle + \frac{sin(\theta-\omega)}{\sqrt{w_2}}|\omega_2\rangle\right). \end{aligned}$$

We can now show that this state is normalized:

$$\begin{aligned} cos^2(\theta-\omega) &= \frac{1}{2}(1 + cos(2(\theta-\omega))) \\ &= \frac{1}{2}(1 + cos(2\theta)\,cos(2\omega) + sin(2\theta)\,sin(2\omega)) \\ &= \frac{1}{2}\left(1 + \frac{cos(2\theta)}{\sqrt{1+(\eta_1-\eta_2)^2 tan^2(2\theta)}} + \frac{sin(2\theta)(\eta_1-\eta_2)tan(2\theta)}{\sqrt{1+(\eta_1-\eta_2)^2 tan^2(2\theta)}}\right) \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2}\left(1+\frac{1-2\eta_2 sin^2(2\theta)}{\sqrt{1-4\eta_1\eta_2 sin^2(2\theta)}}\right)\\
\eta_1\left(\frac{cos^2(\theta-\omega)}{\sqrt{w_1}} + \frac{sin^2(\theta-\omega)}{w_2}\right) &= \eta_1\left(\frac{(w_2-w_1)\,cos^2(\theta-\omega)+w_1}{w_1 w_2}\right)\\
&= \eta_1\left(\frac{\frac{-\sqrt{1-4\eta_1\eta_2 sin^2(2\theta)}}{2}\left(1+\frac{1-2\eta_2 sin^2(2\theta)}{\sqrt{1-4\eta_1\eta_2 sin^2(2\theta)}}\right)+\frac{1}{2}\left(1+\sqrt{1-4\eta_1\eta_2 sin^2(2\theta)}\right)}{\eta_1\eta_2 sin^2(2\theta)}\right)\\
&= 1.
\end{aligned}
$$

We can use this result to represent both $|\phi_1^{ret}\rangle$ and $|\phi_2^{ret}\rangle$:

$$
\begin{aligned}
|\phi_1^{ret}\rangle &= \sqrt{\eta_1}\left(\frac{cos(\theta-\omega)}{\sqrt{w_1}}|\omega_1\rangle + \frac{sin(\theta-\omega)}{\sqrt{w_2}}|\omega_2\rangle\right),\\
|\phi_2^{ret}\rangle &= \sqrt{\eta_2}\left(\frac{cos(\theta+\omega)}{\sqrt{w_1}}|\omega_1\rangle - \frac{sin(\theta+\omega)}{\sqrt{w_2}}|\omega_2\rangle\right).
\end{aligned}
\tag{6.12}
$$

While not immediately obvious, we can show that these states are orthogonal:

$$
\begin{aligned}
\langle\phi_1^{ret}|\phi_2^{ret}\rangle &= \sqrt{\eta_1\eta_2}\left(\frac{cos(\theta-\omega)\,cos(\theta+\omega)}{w_1} - \frac{sin(\theta-\omega)\,sin(\theta+\omega)}{w_2}\right)\\
&= \frac{\sqrt{\eta_1\eta_2}}{2w_1w_2}\left[w_2\left(cos(2\omega)+cos(2\theta)\right) - w_1\left(cos(2\omega)-cos(2\theta)\right)\right]\\
&= \frac{\sqrt{\eta_1\eta_2}}{2w_1w_2}\left[-cos(2\omega)\sqrt{1-4\eta_1\eta_2 sin^2(2\theta)}+cos(2\theta)\right]\\
&= \frac{\sqrt{\eta_1\eta_2}cos(2\theta)}{2w_1w_2}\left(-cos(2\omega)\sqrt{1+(\eta_1-\eta_2)^2 tan^2(2\theta)}+1\right)\\
&= \frac{\sqrt{\eta_1\eta_2}cos(2\theta)}{2w_1w_2}\left(-cos(2\omega)\sqrt{1+tan^2(2\omega)}+1\right) = 0.
\end{aligned}
$$

The second feature that we can exploit to optimize this problem is the UD requirement that $\langle\psi_1|\Pi_2|\psi_1\rangle = \langle\psi_2|\Pi_1|\psi_2\rangle = 0$. We can use the fact that $P(b_j,a_i) = P(a_i,b_j)$ to conclude that $\eta_2 tr\left(\Pi_1|\psi_2\rangle\langle\psi_2|\right) = \mu_1 tr\left(\rho_1^{ret}\Pi_2^{ret}\right) = 0$. From here it is trivial to show that $tr\left(\rho_1^{ret}\Pi_2^{ret}\right) = tr\left(\rho_2^{ret}\Pi_1^{ret}\right) = 0$ and, from the normalization on Π^{ret}, $tr\left(\rho_1^{ret}\Pi_1^{ret}\right) =$

$tr\left(\rho_2^{ret}\Pi_2^{ret}\right)=1$. This implies $\rho_j^{ret}=\left|\phi_i^{ret}\right\rangle\left\langle\phi_i^{ret}\right|$. In order to show this explicitly, we can use the fact that $\Pi_i=c_i\left|\psi_j^{\perp}\right\rangle\left\langle\psi_j^{\perp}\right|$, where $\left\langle\psi_i|\psi_i^{\perp}\right\rangle=0$. First this gives us the result that ρ_j^{ret} is a pure state:

$$tr\left((\rho_j^{ret})^2\right)=c_j^2\frac{tr\left(\sqrt{\Omega}\left|\psi_i^{\perp}\right\rangle\left\langle\psi_i^{\perp}\right|\Omega\left|\psi_i^{\perp}\right\rangle\left\langle\psi_i^{\perp}\right|\sqrt{\Omega}\right)}{c_j^2\left\langle\psi_i^{\perp}|\Omega|\psi_i^{\perp}\right\rangle^2}=1.$$

We can amend the original definition in Eq. (6.7) for pure states:

$$\left|\psi_j^{ret}\right\rangle \;=\; \frac{\sqrt{c_j\Omega}\left|\psi_i^{\perp}\right\rangle}{\sqrt{c_j\left\langle\psi_i^{\perp}|\Omega|\psi_i^{\perp}\right\rangle}}.$$

We can also use this to verify that $\left|\psi_j^{ret}\right\rangle=\left|\phi_j^{ret}\right\rangle$:

$$\begin{aligned}\eta_j\left|\psi_j\right\rangle &= \frac{\Omega\left|\psi_i^{\perp}\right\rangle}{\sqrt{\left\langle\psi_i^{\perp}|\Omega|\psi_i^{\perp}\right\rangle}},\\ \Rightarrow\left|\phi_j^{ret}\right\rangle &= \Omega^{-\frac{1}{2}}\sqrt{\eta_j}\left|\psi_j\right\rangle=\frac{\sqrt{\Omega}\left|\psi_i^{\perp}\right\rangle}{\sqrt{\left\langle\psi_i^{\perp}|\Omega|\psi_i^{\perp}\right\rangle}}=\left|\psi_j^{ret}\right\rangle.\end{aligned}$$

All of these features combined, allow us to have an elegant reformulation of the original problem. The UD problem in the retrodictive perspective is, for a given set of states $\left\{\left|\phi_1^{ret}\right\rangle,\left|\phi_2^{ret}\right\rangle\right\}$ and a matrix Ω, to optimize the function:

$$P_s=\mu_1+\mu_2 \tag{6.13}$$

subject to the constraint:

$$\mu_1\left|\phi_1^{ret}\right\rangle\left\langle\phi_1^{ret}\right|+\mu_2\left|\phi_2\right\rangle\left\langle\phi_2\right|+\mu_0\rho_0^{ret}=\Omega. \tag{6.14}$$

We can solve this problem by expressing the Ω matrix in the $\left\{\left|\phi_1^{ret}\right\rangle,\left|\phi_2^{ret}\right\rangle\right\}$ basis. In order to do this, we can use the following transformation between the eigenvectors of the Ω matrix

and these states:

$$\begin{aligned} |\omega_1\rangle &= \frac{1}{\sqrt{w_1}}\left(\sqrt{\eta_1}cos\left(\theta-\omega\right)|\phi_1^{ret}\rangle + \sqrt{\eta_2}cos\left(\theta+\omega\right)|\phi_2^{ret}\rangle\right), \\ |\omega_2\rangle &= \frac{1}{\sqrt{w_2}}\left(\sqrt{\eta_1}sin\left(\theta-\omega\right)|\phi_1^{ret}\rangle - \sqrt{\eta_2}sin\left(\theta+\omega\right)|\phi_2^{ret}\rangle\right). \end{aligned} \tag{6.15}$$

This gives us the following result:

$$\begin{aligned} \Omega &= w_1|\omega_1\rangle\langle\omega_1| + w_2|\omega_2\rangle\langle\omega_2| \\ &= \left(\eta_1 cos^2\left(\theta-\omega\right)+\eta_1 sin^2\left(\theta-\omega\right)\right)|\phi_1^{ret}\rangle\langle\phi_1^{ret}| + \left(\eta_2 sin^2\left(\theta-\omega\right)+\eta_2 cos^2\left(\theta-\omega\right)\right)|\phi_2^{ret}\rangle\langle\phi_2^{ret}| \\ &\quad +\sqrt{\eta_1\eta_2}\left(cos\left(\theta-\omega\right)cos\left(\theta+\omega\right)-sin\left(\theta-\omega\right)sin\left(\theta+\omega\right)\right)\left(|\phi_1^{ret}\rangle\langle\phi_2^{ret}| + |\phi_2^{ret}\rangle\langle\phi_1^{ret}|\right). \end{aligned}$$

This allows us to rewrite Eq. (6.14) as:

$$\begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix} + \mu_0\rho_0^{ret} = \begin{pmatrix} \eta_1 & \sqrt{\eta_1\eta_2}cos\left(2\theta\right) \\ \sqrt{\eta_1\eta_2}cos\left(2\theta\right) & \eta_2 \end{pmatrix}. \tag{6.16}$$

This same optimization problem appears in a paper deriving optimizing UD discrimination from the no-signaling condition by Barnett and Andersson [63]. In the next section we will flesh out more the connections between the retrodictive approach and the no-signaling condition. For now, we can proceed with the optimization problem by rewriting the constraint Eq. (6.14) above as:

$$\mu_0\rho_0^{ret} = \begin{pmatrix} \eta_1-\mu_1 & \sqrt{\eta_1\eta_2}cos\left(2\theta\right) \\ \sqrt{\eta_1\eta_2}cos\left(2\theta\right) & \eta_2-\mu_2 \end{pmatrix}. \tag{6.17}$$

We can use the normalization $tr\left(\rho_0^{ret}\right)=1$ to derive $\mu_0+\mu_1+\mu_2=1$. The positivity of ρ_0^{ret} gives the conditions that $\eta_1\geq\mu_1$, $\eta_2\geq\mu_2$, and $\left(\eta_1-\mu_1\right)\left(\eta_2-\mu_2\right)-\eta_1\eta_2 cos^2\left(2\theta\right)\geq 0$. This

final constraint can be rewritten as:

$$\eta_2 - \frac{\eta_1\eta_2 cos^2(2\theta)}{(\eta_1 - \mu_1)} \geq \mu_2. \tag{6.18}$$

We can substitute this into the probability of success and optimize:

$$\begin{aligned} P_s &= \mu_1 + \mu_2 \leq \mu_1 + \eta_2 - \frac{\eta_1\eta_2 cos^2(2\theta)}{\eta_1 - \mu_1}, \\ \frac{\partial P}{\partial \mu_1} &= 0 = 1 - \frac{\eta_1\eta_2 cos^2(2\theta)}{(\eta_1 - \mu_1)^2} \\ &\Rightarrow \mu_1 = \eta_1 - \sqrt{\eta_1\eta_2} cos(2\theta) \geq 0, \\ &\Rightarrow \mu_2 = \eta_2 - \sqrt{\eta_1\eta_2} cos(2\theta) \geq 0. \end{aligned}$$

This gives the expected UD optimization result:

$$P_s = \begin{cases} \eta_1(1 - cos^2(2\theta)) & \eta_1 \leq \frac{cos^2(2\theta)}{1+cos^2(2\theta)} \\ 1 - 2\sqrt{\eta_1\eta_1} cos(2\theta) & \frac{cos^2(2\theta)}{1+cos^2(2\theta)} < \eta_1 < \frac{1}{1+cos^2(2\theta)} \\ \eta_2(1 - cos^2(2\theta)) & \frac{1}{1+cos^2(2\theta)} \leq \eta_1 \end{cases}. \tag{6.19}$$

One other result worth noting is that in the optimal solution, $det(\rho_0^{ret}) = 0$, meaning that $\rho_0^{ret} = |\phi_0^{ret}\rangle \langle\phi_0^{ret}|$, where $|\phi_0^{ret}\rangle = \frac{1}{\sqrt{2}}(|\phi_1^{ret}\rangle + |\phi_2^{ret}\rangle)$.

6.2.2 Connecting Retrodiction to the No-Signaling Principle

Earlier we pointed out that the optimization problem we pointed out that the UD optimization condition could be solved using the retrodictive formalism by relying on a constraint in the form of Eq. (6.16), This constraint for optimization is identical to the constraint for optimization derived from the no-signaling condition in [63]. Given the intrinsic connection between the no-signaling condition and state discrimination [63–65], it is essential tease out

the ways in which the retrodictive formulation and the no-signaling condition are related. In order to do so, we need to slightly reformulate the UD discrimination problem. In the standard formulation, Alice sends Bob one of two states, $|\psi_1\rangle, |\psi_2\rangle$ with the respective probabilities η_1, η_2. This can also be accomplished if, instead, Alice and Bob share the state $|\Psi\rangle_{AB} = \sqrt{\eta_1}\,|0\rangle\,|\psi_1\rangle + \sqrt{\eta_1}\,|1\rangle\,|\psi_2\rangle$. If Alice measures her qubit in the $|0\rangle, |1\rangle$ basis, then Bob's state is found in the state $|\psi_1\rangle, |\psi_2\rangle$ with appropriate probability. At this point, we can also introduce the no signaling condition. If Bob performs an optimal UD measurement to distinguish between $|\psi_1\rangle, |\psi_2\rangle$ before Alice performs her own measurement, she should be able to gain no information about the outcome of Bob's measurement. Before Bob's measurement, the possible outcomes of Alice measuring her state can be described using the reduced density matrix $\rho_a = tr_b\left(|\Psi\rangle_{ab}\langle\Psi|\right)$:

$$\rho_a = \begin{pmatrix} \eta_1 & \sqrt{\eta_1\eta_2}cos\,(2\theta) \\ \sqrt{\eta_1\eta_2}cos\,(2\theta) & \sqrt{\eta_2} \end{pmatrix}. \tag{6.20}$$

Here, $\langle\psi_1|\psi_2\rangle = cos\,(2\theta)$. If Bob succeeds in his measurement, then Alice's state is in the corresponding pure state, $|0\rangle, |1\rangle$. This outcome happens with the respective probabilities that Bob's measurement succeeds: p_1, p_2, where p_i is the probability that Bob's measurement succeeds given the state $|\psi_i\rangle$. If Bob's measurement fails, Alice's state can be described by some unknown mixed state ρ_0, which is the outcome with probability p_0. Given this, If Alice does not know the outcome of Bob's measurement, she can describe her state as follows:

$$\tilde{\rho}_a = p_1\,|0\rangle\langle 0| + p_2\,|1\rangle\langle 1| + p_0\rho_0. \tag{6.21}$$

The no-signaling condition requires that Alice gain no information from Bob's measurement, or $\rho_a = \tilde{\rho}_a$, giving the earlier constraint Eq. (6.16).

At first glance, the use of $|\Psi\rangle_{ab} = \sqrt{\eta_1}\,|0\rangle\,|\psi_1\rangle + \sqrt{\eta_1}\,|1\rangle\,|\psi_2\rangle$ as a communication channel

seems asymmetrical, skewed towards Alice perspective. It is easy to see how Alice can use this channel to correlate her information with Bob's results. Additionally this channel analysis shows that Alice's information (reduced density matrix) restricts the efficacy of Bob's measurement. However, with the retrodictive formalism,it is worth questioning whether this apparent asymmetry is intrinsic to the problem, or simply a result of the specific formulation of the problem. Let us consider the opposite formulation for a minute - Bob can also measure his states in an orthonormal basis, leaving Alice with two pure states that are non-orthogonal. We can then ask the question, what is the optimal basis for Bob to measure in so that if Alice performs UD on her resulting states she maximizes her ability to learn the state that Bob measured. As an aside, it is important to realize the subtlety that there is no issue with no-signaling here. Since Bob can't control the outcome of his measurement, only correlations are transfered and not true information. This question of the optimal basis for Bob to measure in is effectively answered by the retrodictive formalism. Consider that instead of Alice using states in the $|0\rangle, |1\rangle$ basis, she uses states in the orthonormal basis $|\phi_1^{ret}, \phi_2^{ret}\rangle$, where $|\phi_i^{ret}\rangle = \Omega^{-\frac{1}{2}}\sqrt{\eta_i}\,|\psi_i\rangle$. In this case, using the positivity of $\Omega^{-\frac{1}{2}}$, the initial state shared by Alice and Bob can be shown to be symmetric. In order to do so, we first need to take advantage of the fact that the definition for $|\phi_i^{ret}\rangle$ implies that $\sqrt{\eta_1}\,|\psi_i\rangle = \sqrt{\Omega}\,|\phi_i^{ret}\rangle = \left(\sqrt{\Omega}\right)_{i1}|\phi_1^{ret}\rangle + \left(\sqrt{\Omega}\right)_{i2}|\phi_2^{ret}\rangle$. Using this and the fact the positivity of $\sqrt{\Omega}$ implies that $\left(\sqrt{\Omega}\right)_{12} = \left(\sqrt{\Omega}\right)_{21}$, gives the following result:

$$\begin{aligned}
|\Psi\rangle_{ab} &= \sqrt{\eta_1}\,|\phi_1^{ret}\rangle\,|\psi_1\rangle + \sqrt{\eta_2}\,|\phi_2^{ret}\rangle\,|\psi_2\rangle \\
&= |\phi_1^{ret}\rangle\left(\left(\sqrt{\Omega}\right)_{11}|\phi_1^{ret}\rangle + \left(\sqrt{\Omega}\right)_{12}|\phi_2^{ret}\rangle\right) + |\phi_2^{ret}\rangle\left(\left(\sqrt{\Omega}\right)_{21}|\phi_1^{ret}\rangle + \left(\sqrt{\Omega}\right)_{22}|\phi_2^{ret}\rangle\right) \\
&= \left(\left(\sqrt{\Omega}\right)_{11}|\phi_1^{ret}\rangle + \left(\sqrt{\Omega}\right)_{12}|\phi_2^{ret}\rangle\right)|\phi_1^{ret}\rangle + \left(\left(\sqrt{\Omega}\right)_{21}|\phi_1^{ret}\rangle + \left(\sqrt{\Omega}\right)_{22}|\phi_2^{ret}\rangle\right)|\phi_2^{ret}\rangle \\
&= \sqrt{\eta_1}\,|\psi_1\rangle\,|\phi_1^{ret}\rangle + \sqrt{\eta_2}\,|\psi_2\rangle\,|\phi_2^{ret}\rangle\,.
\end{aligned}$$

The incredible conclusion of this realization is that the retrodictive basis $|\phi_i^{ret}\rangle$ is the basis

for which the communication channel is symmetric from both Bob and Alice's perspective. Additionally, it is clear that the no-signaling condition must be identical to the retrodictive constraint.

Chapter 7

Conclusion

The problem of state discrimination is essential to Quantum Information Theory. This theory is essential in designing communication protocols, experiments, and quantum algorithms. In chapters 1 and 2 of this dissertation, we review the fundamentals of quantum measurement theory and quantum state discrimination. These chapters serve as a foundation for understanding how one can use the theory of generalized measurements (POVMs) to design optimal communication schemes between two parties, Alice and Bob. In addition, this section explores various criteria for optimization, including minimizing error or requiring unambiguous communication channels.

One important extension of the communication protocols introduced in chapter 2 is extending the communication to multiple parties. The goal of sequential state discrimination is to take advantage of the security provided by the quantum communication channel to allow Alice to communicate with multiple parties in sequence. At the end of chapter 3, we explore the way the standard state discrimination protocols can be modified to allow Bob to make a suboptimal measurement in order to pass along some usable information to a third party, Charlie. Similar to the standard discrimination protocols, there are various possible criteria for optimization, three of which are explored in detail in chapters 3-5.

In chapter 3, we focus on the problem of Sequential Minimum Error. In this scheme, we show how one can design and optimize a communication channel between Alice, Bob, and Charlie, such that Bob and Charlie maximize the probability that they can both correctly identify the state sent by Alice. In chapter 4, we focus on the problem of Sequential Unambiguous Discrimination. In the SUD protocol, Bob and Charlie both require that they have measurements that can identify the state sent by Alice with certainty. In order to achieve this certainty, they also are forced to have inconclusive results. In this chapter, we thoroughly explore the ways they can design and optimize this procedure, as well as considering the mutual information for the communication channel. In chapter 4, we explore an interpolation between these two schemes. In the interpolation, Bob and Charlie control the probability of receiving an inconclusive results, and optimize their sequential discrimination to this fixed rate of inconclusive results. When this rate is zero, they recover the SME protocol and when this rate matches the inconclusive rate for SUD they recover the SUD protocol.

In the final chapter of this dissertation, we explore quantum state discrimination from a new perspective: quantum retrodiction. We review the definition of quantum retrodiction and show how this definition can be modified to form a dual problem to the state discrimination problem. By applying this new technique, we reexamine and explore the standard unambiguous discrimination scheme and derive the expected results. This new perspective on quantum retrodiction promises to be a fruitful avenue for exploration in future analysis of quantum state discrimination protocols.

Ultimately, while there is significant progress made to the understanding of state discrimination protocols made by this thesis, there are still many open problems. One basic extension of the state discrimination protocols is to extend the Sequential Discrimination with a Fixed Rate of Inconclusive Outcome to arbitrary priors. Additionally, the question of how to optimize these schemes for mixed states and for protocols where Bob and Charlie utilize different strategies warrants more exploration. Finally, there is a lot of promise in the

proposed formulation of quantum retrodiction. Extending this formalism to other discrimination problems, including sequential discrimination problems, could yield some interesting results.

Bibliography

[1]D. Fields, H. Rui, M. Hillery, and J. Bergou, "Sequential unambiguous discrimination", In preparation.

[2]D. Fields and J. Bergou, "Sequential minimum error discrimination", in preparation.

[3]D. Fields and J. Bergou, "Sequential discrimination with a fixed rate of inconclusive outcome", in preparation.

[4]D. Fields, A. Sajia, and J. Bergou, "New insights on quantum state discrimination from quantum retrodiction", in preparation.

[5]M. A. Neumark, Compt. Rend. Acad. Sci. USSR **28** (1943).

[6]J. A. Bergou and M. Hillery, *Introduction to the theory of quantum information processing*, Graduate Texts in Physics (Springer, New York,NY, USA, 2013).

[7]M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information: 10th anniversary edition*, 10th (Cambridge University Press, New York, NY, USA, 2011).

[8]A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer Academic Publishers, New York, 2002).

[9]J. A. Bergou, U. Herzog, and M. Hillery, "Discrimination of Quantum States", in *Quantum state estimation*, Vol. 649, Lecture Notes in Physics (Springer, 2004), pp. 417–465.

[10]J. A. Bergou, "Discrimination of quantum states", Journal of Modern Optics **57**, 160–180 (2010).

[11]S. M. Barnett and S. Croke, "Quantum state discrimination", Adv. Opt. Photon. **1**, 238–278 (2009).

[12]J. Bae and L.-C. Kwek, "Quantum state discrimination and its applications", Journal of Physics A: Mathematical and Theoretical **48**, 083001 (2015).

[13]S. Croke, E. Andersson, S. M. Barnett, C. R. Gilson, and J. Jeffers, "Maximum confidence quantum measurements", Phys. Rev. Lett. **96**, 070401 (2006).

[14]P. J. Mosley, S. Croke, I. A. Walmsley, and S. M. Barnett, "Experimental realization of maximum confidence quantum state discrimination for the extraction of quantum information", Phys. Rev. Lett. **97**, 193601 (2006).

[15]S. Croke, P. J. Mosley, S. M. Barnett, and I. A. Walmsley, "Maximum confidence measurements and their optical implementation", The European Physical Journal D **41**, 589–598 (2007).

[16]U. Herzog, "Optimized maximum-confidence discrimination of n mixed quantum states and application to symmetric states", Phys. Rev. A **85**, 032312 (2012).

[17]O. Jiménez, M. A. Solís-Prosser, A. Delgado, and L. Neves, "Maximum-confidence discrimination among symmetric qudit states", Phys. Rev. A **84**, 062315 (2011).

[18]P. Wallden, V. Dunjko, and E. Andersson, "Minimum-cost quantum measurements for quantum information", Journal of Physics A: Mathematical and Theoretical **47**, 125303 (2014).

[19]W. K. Wooters and W. M. Zurek, "A single quantum cannot be cloned", Nature **299**, 802–803 (1982).

[20]D. Dieks, "Communication by epr devices", Physics Letters A **92**, 271–272 (1982).

[21]L.-M. Duan and G.-C. Guo, "Probabilistic cloning and identification of linearly independent quantum states", Phys. Rev. Lett. **80**, 4999–5002 (1998).

[22]V. Yerokhin, A. Shehu, E. Feldman, E. Bagan, and J. A. Bergou, "Probabilistically perfect cloning of two pure states: geometric approach", Phys. Rev. Lett. **116**, 200401 (2016).

[23]H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher, "Noncommuting mixed states cannot be broadcast", Phys. Rev. Lett. **76**, 2818–2821 (1996).

[24]A. Chefles, "Unambiguous discrimination between linearly independent quantum states", Physics Letters A **239**, 339–347 (1998).

[25]S. Bandyopadhyay, R. Jain, J. Oppenheim, and C. Perry, "Conclusive exclusion of quantum states", Phys. Rev. A **89**, 022336 (2014).

[26]A. Holevo, "Statistical decision theory for quantum systems", Journal of Multivariate Analysis **3**, 337–394 (1973).

[27]C. W. Helstrom, "Quantum detection and estimation theory", Journal of Statistical Physics **1**, 231–252 (1969).

[28]H. Yuen, R. Kennedy, and M. Lax, "Optimum testing of multiple hypotheses in quantum detection theory", IEEE Transactions on Information Theory **21**, 125–134 (1975).

[29]R. Han, J. A. Bergou, and G. Leuchs, "Near optimal discrimination of binary coherent signals via atom–light interaction", New Journal of Physics **20**, 043005 (2018).

[30]B.-G. Englert, "Fringe visibility and which-way information: an inequality", Phys. Rev. Lett. **77**, 2154–2157 (1996).

[31]E. Andersson, S. M. Barnett, C. R. Gilson, and K. Hunter, "Minimum-error discrimination between three mirror-symmetric states", Phys. Rev. A **65**, 052308 (2002).

[32]T. Singal and S. Ghosh, "Minimum error discrimination for an ensemble of linearly independent pure states", Journal of Physics A: Mathematical and Theoretical **49**, 165304 (2016).

[33]S. M. Barnett and S. Croke, "On the conditions for discrimination between quantum states with minimum error", Journal of Physics A: Mathematical and Theoretical **42**, 062001 (2009).

[34]J. Bae, "Structure of minimum-error quantum state discrimination", New Journal of Physics **15**, 073037 (2013).

[35]I. Ivanovic, "How to differentiate between non-orthogonal states", Physics Letters A **123**, 257–259 (1987).

[36]D. Dieks, "Overlap and distinguishability of quantum states", Physics Letters A **126**, 303–306 (1988).

[37]A. Peres, "How to differentiate between non-orthogonal states", Physics Letters A **128**, 19 (1988).

[38]G. Jaeger and A. Shimony, "Optimal distinction between two non-orthogonal quantum states", Physics Letters A **197**, 83–87 (1995).

[39]T. Rudolph, R. W. Spekkens, and P. S. Turner, "Unambiguous discrimination of mixed states", Phys. Rev. A **68**, 010301 (2003).

[40]C. Zhang, Y. Feng, and M. Ying, "Unambiguous discrimination of mixed quantum states", Physics Letters A **353**, 300–306 (2006).

[41]B. Huttner, A. Muller, J. D. Gautier, H. Zbinden, and N. Gisin, "Unambiguous quantum measurement of nonorthogonal states", Phys. Rev. A **54**, 3783–3789 (1996).

[42]R. B. M. Clarke, A. Chefles, S. M. Barnett, and E. Riis, "Experimental demonstration of optimal unambiguous state discrimination", Phys. Rev. A **63**, 040305 (2001).

[43]G. Waldherr, A. C. Dada, P. Neumann, F. Jelezko, E. Andersson, and J. Wrachtrup, "Distinguishing between nonorthogonal quantum states of a single nuclear spin", Phys. Rev. Lett. **109**, 180501 (2012).

[44]E. Bagan, R. Muñoz-Tapia, G. A. Olivares-Rentería, and J. A. Bergou, "Optimal discrimination of quantum states with a fixed rate of inconclusive outcomes", Physical Review A **86**, 040303 (2012).

[45]J. von Neumann, *Mathematical foundations of quantum mechanics* (Princeton University Press, Princeton, 1955).

[46]P. Rapčan, J. Calsamiglia, R. Muñoz-Tapia, E. Bagan, and V. Bužek, "Scavenging quantum information: multiple observations of quantum systems", Phys. Rev. A **84**, 032326 (2011).

[47]J. Bergou, E. Feldman, and M. Hillery, "Extracting information from a qubit by multiple observers: Toward a theory of sequential state discrimination", Physical Review Letters **111**, 100501 (2013).

[48]M. A. Solís-Prosser, P. González, J. Fuenzalida, S. Gómez, G. B. Xavier, A. Delgado, and G. Lima, "Experimental multiparty sequential state discrimination", Phys. Rev. A **94**, 042309 (2016).

[49]M. Namkung and Y. Kwon, "Sequential state discrimination of coherent states", Scientific Reports **8**, 16915 (2018).

[50]M. Hillery and J. Mimih, "Sequential discrimination of qudits by multiple observers", (2017).

[51]C.-Q. Pang, F.-L. Zhang, L.-F. Xu, M.-L. Liang, and J.-L. Chen, "Sequential state discrimination and requirement of quantum dissonance", Physical Review A **88**, 052331 (2013).

[52]C. H. Bennett, "Quantum cryptography using any two nonorthogonal states", Phys. Rev. Lett. **68**, 3121–3124 (1992).

[53]T. M. Cover, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing) (Hardcover)* (2006).

[54]S. M. Barnett, D. T. Pegg, and J. Jeffers, "Bayes' theorem and quantum retrodiction", Journal of Modern Optics **47**, 1779–1789 (200).

[55]Y. Aharonov, P. G. Bergmann, and J. L. Lebowitz, "Time symmetry in the quantum process of measurement", Phys. Rev. **134**, B1410–B1416 (1964).

[56]Y. Aharonov and D. Z. Albert, "Is the usual notion of time evolution adequate for quantum-mechanical systems? i", Phys. Rev. D **29**, 223–227 (1984).

[57]Y. Aharonov and L. Vaidman, "Complete description of a quantum system at a given time", Journal of Physics A: Mathematical and General **24**, 2315–2328 (1991).

[58]S. Barnett, *Quantum retrodiction*, edited by E. Andersson and P. Ohberg, Scottish Graduate Series (Springer,Cham, 2014).

[59]D. T. Pegg and S. M. Barnett, "Retrodiction in quantum optics", Journal of Optics B:Quantum and Semiclassical Optics **1**, 442.

[60]S. M. Barnett, D. T. Pegg, J. Jeffers, O. Jedrkiewicz, and R. Loudon, "Retrodiction for quantum optical communications", Phys. Rev. A **62**, 022313 (2000).

[61]D. T. Pegg, S. M. Barnett, and J. Jeffers, "Quantum retrodiction in open systems", Phys. Rev. A **66**, 022106 (2002).

[62]M. Hillery and D. Koch, "Retrodiction of a sequence of measurement results in qubit interferometers", Phys. Rev. A **94**, 032118 (2016).

[63]S. M. Barnett and E. Andersson, "Bound on measurement based on the no-signaling condition", Phys. Rev. A **65**, 044307 (2002).

[64]W.-Y. Hwang, "Helstrom theorem from the no-signaling condition", Phys. Rev. A **71**, 062315 (2005).

[65]S. Croke, E. Andersson, and S. M. Barnett, "No-signaling bound on quantum state discrimination", Phys. Rev. A **77**, 012113 (2008).