Locally Adaptive Protocols for Quantum State Discrimination

by

Sarah Brandsen

Department of Physics
Duke University

Date: _____

Approved:

_____
Henry Pfister, Advisor

_____
Robert Calderbank

_____
Ayana Arce

_____
Iman Marvian

_____
Thomas Barthel

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Physics
in the Graduate School of Duke University
2021

Abstract

Locally Adaptive Protocols for Quantum State Discrimination

by

Sarah Brandsen

Department of Physics
Duke University

Date: _____

Approved:

_____
Henry Pfister, Advisor

_____
Robert Calderbank

_____
Ayana Arce

_____
Iman Marvian

_____
Thomas Barthel

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Physics in the Graduate School of Duke University
2021

Abstract

This dissertation makes contributions to two rapidly developing fields: quantum information theory and machine learning. It has recently been demonstrated that reinforcement learning is an effective tool for a wide variety of tasks in quantum information theory, ranging from quantum error correction to quantum control to preparation of entangled states. In this work, we demonstrate that reinforcement learning is additionally highly effective for the task of multiple quantum hypothesis testing.

Quantum hypothesis testing consists of finding the quantum measurement which allows one to discriminate with minimal error between m possible states $\{\rho_k\}|_{k=1}^m$ of a quantum system with corresponding prior probabilities $p_k = \Pr[\rho = \rho_k]$. In the general case, although semi-definite programming offers a way to numerically approximate the optimal solution [EMV03], a closed-form analytical solution for the optimal measurement is not known.

Additionally, when the quantum system is large and consists of many subsystems, the optimal measurement may be experimentally difficult to implement. In this work, we provide a comprehensive study of locally adaptive approaches to quantum hypothesis testing where only a single subsystem is measured at a time and the order and types of measurements implemented may depend on previous measurement results. Thus, these locally adaptive protocols present an experimentally feasible approach to quantum state discrimination.

We begin with the case of binary hypothesis testing (where $m = 2$), and generalize previous work by Acin et al. (Phys. Rev. A 71, 032338) to show that a simple Bayesian-updating scheme can optimally distinguish between any pair of arbitrary pure, tensor product quantum states. We then demonstrate that this same Bayesian-updating scheme has poor asymptotic behaviour when the candidate states are not pure, and based on this we introduce a modified scheme with strictly better performance. Finally, a dynamic programming (DP) approach is used to find the optimal local protocol for binary state discrimination and numerical simulations are run for both qubit and qutrit subsystems.

Based on these results, we then turn to the more general case of multiple hypothesis testing where there may be several candidate states. Given that the dynamic-programming approach has a high complexity when there are a large number of subsystems, we turn to reinforcement learning methods to learn adaptive protocols for even larger systems. Our numerical results support the claim that reinforcement learning with neural networks (RLNN) is able to successfully find the optimal locally adaptive approach for up to 20 subsystems. We additionally find the optimal collective measurement through semidefinite programming techniques, and demonstrate that the RLNN approach meets or comes close to the optimal collective measurement in every random trial.

Next, we focus on quantum information theory and provide an operational interpretation for the entropy of a channel. This task is motivated by the central role of entropy across several areas of physics and science. We use games of chance as a more systematic and unifying approach to define entropy, as a system's performance in any game of chance depends solely on the uncertainty of the output. We construct families of games which result in a pre-order on channels and provide an operational interpretation for all pre-orders (corresponding to majorization, conditional majorization, and channel majorization respectively), and this defines the

unique asymptotically continuous entropy function for classical channels.

Contents

List of Figures

Acknowledgements

One of the best parts of completing a PhD thesis is the opportunity to finally express gratitude towards the many people who have been so helpful along the way.
I am extremely grateful for the opportunity to work in Dr. Henry Pfister's research group. I don't think it would have been possible to ask for a better academic advisor. I am also grateful to Dr. Robert Calderbank for his support and for agreeing to co-supervise my research.

Likewise, I am grateful to the rest of the Pfister group and those in Gross Hall, especially Kevin Stubbs, Narayanan Rengaswamy, and Mengke Lian. Mengke's extensive knowledge of computer programming and willingness to share that knowledge was invaluable for starting the work done in this thesis. I am grateful to Narayanan for his contribution to our research projects, as well as for being such an excellent teacher in general. I appreciated the opportunity to work with Kevin, and am grateful for his contributions to both works on quantum hypothesis testing. Finally, I am grateful to James Leung for helpful physics discussions and for teaching me more about the experimental side of quantum information.

I also am grateful to the Eisenstein group at Caltech for helping me find a start in physics, and to the Vedral group at the National University of Singapore (NUS) for first introducing me to quantum information theory. As a beginning researcher, it was Dr. Johannes Pollanen, Dr. Michele Dall'Arno, and Dr. Francesco Buscemi who first taught me how to approach research. Over the past year, it has also

been a privilege to collaborate with the University of Calgary. I appreciate the time investment that Dr. Gilad Gour and Dr. Carlo Maria Scandolo have made in explaining their research projects, and have greatly enjoyed collaborating with Isabelle Jianing Geng on the channel entropy work.

Finally, I am grateful to Grants No. 1908730 and 1910571 from the National Science Foundation (NSF) which financially supported this work.

Aside from physics, I am grateful to have become acquainted with many other wonderful people- Sehaj, Xiaoqing, Zehui, Meimei, Erin, Faye, Dave, and many others. Above all, I am deeply grateful to my family, my spouse Ruijie, and my toddler Johann. Watching Johann grow from a frail preemie to a happy and healthy toddler has been the greatest source of joy during my time at Duke.

1

Introduction

Quantum information theory and reinforcement learning are both rapidly advancing fields with enormous potential. The development of a working quantum computer would revolutionize computing and offer speedups in areas ranging from cryptography to quantum search and solving linear equations. Likewise, the ability of machine learning algorithms to independently learn and adapt from any dataset has applications ranging from self-driving cars to fraud detection. At the intersection of these two fields, machine learning can be applied to problems in quantum information theory to find flexible and effective quantum protocols.

Quantum information theory has progressed at a dizzying rate over the past few decades. In the 1970s, Richard Feynman famously suggested that quantum physics would be best simulated with quantum computers [Tra12]. Since Feynman's original suggestion, it has become clear that quantum properties such as superposition and entanglement offer unique computational advantages. Perhaps the most famous example of a likely quantum advantage lies in the field of cryptography. Most modern cryptographic systems are based on integer factorization, as there is no currently

known protocol which would allow ordinary computers to factor the product of two sufficiently large prime numbers. On the other hand, Shor's algorithm [Sho94] would allow a quantum computer to efficiently find the factors and thus break most modern cryptographic systems. As a more recent example, in 2019 Google used 54 qubits to complete a series of operations in the span of 200 seconds. At the time Google believed that it would take a classical supercomputer 10,000 years to complete the same operations, though this estimate was subsequently reduced to 2.5 days by IBM. A general introduction to quantum information theory as well as a summary of multiple key results in the field can be found in [NC11a].

Current state-of-the-art experiments include superconducting quantum computers and trapped ion quantum computers. In trapped ion quantum computers, ions are suspended via electromagnetic fields and the electronic state of each ion corresponds to a qubit (quantum bit), and quantum operations are completed through coupling induced by lasers. Such trapped ion quantum computers have demonstrated significant success, including the ability to successfully entangle up to 20 qubits at a time. One of the largest quantum simulations performed thus far utilized 53 trapped ion qubits to simulate the properties of many body magnetic interactions. [ZPH$^+$17]

Despite these remarkable advances, some limitations still exist. For example, IonQ, a state-of-the-art quantum computer, can perform single-qubit gates on 79 qubits and two qubit gates on arbitrary pairs of qubits for chains of up to 11 qubits. Thus, one key barrier to implementing arbitrary quantum algorithms is the difficulty in implementing gates involving more than two qubits. Instead, algorithms involving multi-qubit operations become experimentally feasible when they can be implemented as a short sequence of one- and two-qubit quantum operations.

Quantum hypothesis testing is a key task in quantum information which is affected by limitations on multi-qubit gates. In quantum hypothesis testing, a system is in a state represented by the density matrix ρ. This must belong to one of m candidate states $\{\rho_j\}|_{j=1}^m$ with probabilities $\{p_j = \Pr(\rho = \rho_j)\}$. The state of a quantum system cannot be observed directly– instead, quantum measurements are used to extract information about a system. Thus, the task of quantum hypothesis testing amounts to finding the quantum measurement which allows one to guess the state of the system with the smallest probability of error. A literature review of key results in quantum hypothesis testing can be found in [BC09].

Quantum hypothesis testing has multiple applications, including discriminating between coherent quantum states and decoding codewords that have been sent through a known, noisy quantum channel. Despite the central nature of this task, the optimal measurement generally requires simultaneous operations on all qubits in the system. Thus, if the system is sufficiently large, implementing the optimal measurement will generally be experimentally infeasible. The goal of this thesis is then to implement optimal or almost-optimal quantum hypothesis testing algorithms in a *locally adaptive* way. In the locally adaptive algorithms that we consider, each step consists of a single-qubit measurement, with subsequent steps depending on measurement results from previous rounds. Thus, such a locally adaptive measurement scheme is experimentally feasible.

We begin by investigating several locally adaptive schemes in the case of binary state discrimination, including various locally greedy measurement schemes and a more general dynamic-programming based locally adaptive protocol. The latter protocol finds the optimal locally adaptive measurement sequence (and order) for state discrimination. Despite its optimality, dynamic programming computationally ex-

pensive. Hence, we turn to *reinforcement learning* to approach general quantum hypothesis testing on larger systems.

Reinforcement learning algorithms consist of training an agent to make a sequence of decisions to optimize a future reward [SB18]. More specifically, a reinforcement learning agent controls a policy π where $\pi(a, s)$ is the probability of choosing action a given that the system is in state s. In each round, the agent starts in state s and chooses action a. The agent then receives a reward from the environment, as well as information about the updated state s' of the system. Through repeated training, the agent aims to learn the best choice of a for each state s, thus reaching the optimal policy π^* which maximizes expected reward.

The use of reinforcement learning is further motivated by its remarkable success across multiple different applications. Perhaps the most famous example occurred in 2016, when the reinforcement-based computer program AlphaGo beat 9-dan professional Go player Lee Sedol in a five game match [BBC16]. The ability of an computer program to beat world-class professionals was a landmark for artificial intelligence. Reinforcement learning has additionally been utilised in a diverse range of tasks, from creating preliminary models of self-driving cars [KHJ$^+$18] to developing novel approaches to medical imaging [MKHV18]. Finally, reinforcement learning has been successfully applied to certain tasks in quantum error correction.

The final portion of this thesis focuses on characterising entropy in physical systems. Entropy is a crucial quantity in multiple different areas of science, from physics to statistics to chemistry. This is also evident in the multiple different definitions of entropy: von Neumann entropy [BZ], Renyi entropy [Ren60], Shannon entropy [Sha48], Boltzmann entropy [Jay65], entropy of mixing [Pri67], and so on.

Despite the very different interpretation of each entropy, each function shares one key trait in common– namely, all entropies are measurements of uncertainty. Thus, the key task finding the entropy of physical systems amounts to characterising their uncertainty.

The second law of thermodynamics motivates the importance of thermodynamic entropy, stating that the total entropy of any isolated thermodynamic system is nondecreasing over time. Thus, entropy provides a way to rule out impossible thermodynamic physical processes (namely, processes that would *decrease* disorder), and provides understanding of the direction of spontaneous change in a physical system [Mac92]. Similar interpretations exist for other forms of entropy.

We introduce a unifying approach for characterising the entropy of a given physical system. In this approach, games of chance are used to quantify the uncertainty of a given system. Such games of chance– also known as gambling games– are ideal for measuring uncertainty, as the probability of winning a gambling game with a given system depends only on the certainty of the system's output. Therefore, if system A outperforms system B for all reasonable games of chance, then system B has at least as much entropy as system A.

Although the entropy of classical and quantum states is known to be the Shannon and Von Neumann entropy respectively, the entropy of classical and quantum channels remain unknown. The entropy of a channel additionally has a special importance given the key role of channels in information theory– channels are one of the most general physical processes, as any physical mapping of an input state to an output state can be represented as a channel. We therefore apply our framework to find the unique, asymptotically continuous entropy of a channel and provide an

operational interpretation for this result.

Structure and Contributions

The remainder of this thesis is structured as follows:

1. Chapter 2: Introduces notation and key concepts in quantum mechanics and information theory

2. Chapter 3: Introduces quantum hypothesis testing and discusses locally greedy algorithms for binary state discrimination. One key result is that locally greedy algorithms are fully optimal when both candidate states are pure.

3. Chapter 4: Extends Chapter 3 and discusses dynamic programming-based locally adaptive protocols for binary state discrimination.

4. Chapter 5: Introduces a reinforcement-learning based algorithm for locally adaptive state discrimination in the general case (that is, when there may be multiple different candidate states.) Numerical simulations are provided comparing the reinforcement based algorithm with an optimal (non-local) measurement.

5. Chapter 6: Introduces families games of chance and finds pre-orders of channels based on games of chance. Operational interpretations are provided for each pre-order.

6. Chapter 7: Based on the pre-orders found in chapter 6, the minimum output entropy [Sho04] is found to be the unique asymptotically continuous entropy function. Additionally, connections to quantum dynamical resource theory are discussed.

7. Chapter 8: Conclusions

Collaboration Statement

The work presented in Chapters 3 and 4 is done in collaboration with Mengke Lian, Kevin D Stubbs, Narayanan Rengaswamy, and Dr. Henry Pfister. I found the main analytical results and wrote the final manuscript. Mengke Lian improved the computer algorithms used and generated several of the numerical results. Kevin Stubbs and Narayanan Rengaswamy provided useful feedback for many ideas on this project, reviewed and edited the final manuscript, and contributed to analytical results for the modified locally greedy algorithm. Dr. Pfister oversaw the project.

The work presented in Chapter 5 is done in collaboration with Kevin D. Stubbs and Dr. Henry Pfister. I developed the reinforcement learning algorithm, found the numerical results, and wrote the final manuscript. Kevin Stubbs showed the RLNN algorithm is robust under noise and contributed to writing the final manuscript. Dr. Pfister oversaw the project.

The work presented in Chapters 6 and 7 is done in collaboration with Isabelle Jianing Geng and Dr. Gilad Gour. Dr. Gour oversaw the project, found results for conditional majorization, and developed the project idea. I found the remaining analytical results, and wrote the final manuscript. Isabelle Jianing Geng made substantial contributions to the final manuscript and verified all analytical results.

2

Brief Overview of Quantum Information Theory

2.1 Dirac Notation and Wavefunctions

In order to represent the state of a quantum mechanical system, we first introduce Dirac notation. In this notation, vectors are represented as *kets* belonging to a Hilbert space $\mathcal{H}$ and are denoted as $|\psi\rangle \in \mathcal{H}$. Physically, kets represent a pure state of some quantum system. Likewise, bras are denoted as $\langle\psi|$ and belong to the dual space of kets such that:

$$|\psi\rangle \triangleq \begin{pmatrix} \psi_1 \\ \psi_2 \\ \cdots \\ \cdots \\ \psi_n \end{pmatrix} \Longrightarrow \langle\psi| = \begin{pmatrix} \psi_1^* & \psi_2^* & .. & .. & \psi_n^* \end{pmatrix}$$

which can be equivalently represented as:

$$(|\psi\rangle)^\dagger = \langle\psi| \quad \text{and} \quad (\langle\psi|)^\dagger = |\psi\rangle$$

where † represents the conjugate transpose operation.

One can then define the inner product of ket $|\psi\rangle$ with bra $\langle\phi|$ as:

$$\langle\phi\,|\psi\rangle = \sum_j \phi_j^* \psi_j$$

and the outer product $|\psi\rangle\langle\phi|$ is defined via standard matrix multiplication rules. A wavefunction is defined to be normalised if the inner product of the wavefunction with itself is one- that is, $|\psi\rangle$ is normalised if $\langle\psi\,|\psi\rangle = 1$.

A complete orthonormal basis for a d-dimensional system is then a set of vectors $\{|i\rangle\}|_{i=1}^{d}$ such that

$$\langle i\,|j\rangle = \delta_{i,j} \quad \forall i, j \in \{1, 2, ..., d\}$$

From the above, it follows that any normalised wavefunction in d-dimensions can be expressed as:

$$|\psi\rangle \triangleq \sum_{j=1}^{d} \alpha_j\,|j\rangle \quad \text{s.t.} \quad \sum_{j=1}^{d} |\alpha_j|^2 = 1$$

Two-dimensional systems have a special place in quantum mechanics and quantum information theory, and are denoted as qubits (quantum bits).

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$$

where $\beta = \sqrt{1-\alpha^2}$.

2.2 Density Matrices and Measurements

We can now introduce the concepts of mixed states. A quantum system with a mixed state cannot be described by a single pure state, but rather by a probabilistic ensemble of pure states. Given that the ensemble of quantum states is $\{|\phi_x\rangle\}_x$ and the corresponding probabilities are $\{p_x\}$, the mixed state may then be represented as a *density matrix* ρ:

$$\rho \triangleq \sum_{j=1}^{d} p_j\,|\phi_j\rangle\langle\phi_j|$$

Given that the probabilities must sum up to 1, it follows that $\mathrm{Tr}[\rho] = 1$. Likewise, ρ is positive semidefinite and Hermitian such that $\rho = \rho^\dagger$.

We denote the set of density matrices compatible with system A as $\mathcal{D}(\mathrm{A})$ or equivalently as $\mathcal{D}(\mathcal{H}_\mathrm{A})$ where $\mathcal{H}_\mathrm{A}$ represents the Hilbert space corresponding to system A.

Information about a quantum state can be extracted through quantum measurement. Quantum measurements are given the name of POVM (Positive Operator Valued Measurements), such that a POVM $\hat{\Pi}$ is a collection of measurement elements $\{\Pi_j\}|_{j=1}^n$. Additionally, such measurement elements must satisfy the constraints:

$$\Pi_j \geqslant 0 \quad \forall j$$

$$\sum_{j=1}^{n} \Pi_j = \mathbb{I}$$

where the last condition ensures that the probability of achieving some measurement result is 1. Given that measurement $\hat{\Pi}$ is implemented on a system in state ρ, the probability of obtaining outcome j is given by:

$$p_j = \mathrm{Tr}[\Pi_j \rho]$$

One unique feature of quantum mechanics is that measurement disturbs the system. In other words, the state after measurement will depend on the measurement result. In the case of a projective measurement where $\Pi_j^2 = \Pi_j$ for all j, then the postmeasurement state given outcome j is observed is defined as:

$$\rho^{(j)} \triangleq \frac{\Pi_j \rho \Pi_j}{\mathrm{Tr}[\Pi_j \rho]}$$

In this thesis, we assume that all measurements are projective unless explicitly stated otherwise.

2.3 Tensor Product States and Local Operations

Suppose that a quantum system contains n subsystems.

One special class of quantum states is tensor product states. Consider a simple bipartite system containing system A and system B. If system A has density matrix ρ_{A} and system B has density matrix ρ_{B}, then the the joint system AB has the state

$$\rho_{AB} \triangleq \rho_A \otimes \rho_B$$

The above state belongs to the class of tensor product quantum states (TPQS). In general, the combined state of a n-subsystems is a tensor product if it can be represented as

$$\rho = \bigotimes_{j=1}^{n} \rho_j$$

where for all j, then ρ_j is the state of the j^{th} subsystem.

Note that not all density matrices can be written as tensor product states. For example, entangled states involve interactions between subsystems and therefore do not belong to the set of TPQS. One simple entangled state for a two-qubit system would be:

$$\rho_{\mathrm{ent}} = \frac{1}{2}\Big(|00\rangle + |11\rangle \Big)\Big(\langle 00| + \langle 11| \Big)$$

and there are no qubit density matrices, ρ_A, ρ_B, such that $\rho_{\mathrm{ent}} = \rho_A \otimes \rho_B$.

Finally, we discuss the implementation of local measurements on a tensor product quantum state. In a local measurement, only one subsystem is measured while the remaining subsystems are acted on trivially. Let $\hat{\Pi} = \{\Pi_1, ..., \Pi_{d_1}\}$ be a quantum measurement acting on system j. This measurement can be extended to the full system by transforming each measurement element as:

$$\Pi \rightarrow \mathbb{I}^{(1)} \otimes ... \otimes \mathbb{I}^{(j-1)} \otimes \Pi^{(j)} \otimes \mathbb{I}^{(j+1)} \otimes ... \otimes \mathbb{I}^{(n)}$$

where $\mathbb{I}^{(k)}$ is the identity matrix acting on the k^{th} subsystem. Equivalently, a local operation on subsystem j is equivalent to a joint operation where all other subsystems are acted on trivially with the identity.

Then the probability of obtaining outcome k when implementing a local projective measurement $\hat{\Pi}$ on subsystem j of a TPQS ρ is:

$$\begin{aligned}\Pr(\text{out} = k) &= \text{Tr}\left[\left(\mathbb{I}^{(1)} \otimes ... \otimes \mathbb{I}^{(j-1)} \otimes \Pi^{(j)} \otimes \mathbb{I}^{(j+1)} \otimes ... \otimes \mathbb{I}^{(n)}\right)\left(\rho_1 \otimes ...\rho_n\right)\right] \\ &= \Big(\prod_{i=1}^{j-1} \text{Tr}[\rho_i]\Big)\text{Tr}[\Pi_k \rho_j]\Big(\prod_{i=j+1}^{n} \text{Tr}[\rho_i]\Big) \\ &= \text{Tr}[\Pi_k \rho_j]\end{aligned}$$

and the post-measurement state given measurement outcome k is then:

$$\rho^{(k)} \triangleq \bigotimes_{\ell=1}^{j-1} \rho_\ell \otimes \frac{\Pi_j \rho \Pi_j}{\text{Tr}[\Pi_j \rho]} \otimes \bigotimes_{\ell=j+1}^{n} \rho_j$$

Thus, a local measurement on a TPQS only changes one subsystem and leaves all others unaffected.

2.4 Quantum Channels and Superchannels

A quantum channel is a completely positive, trace preserving map that takes a quantum state as an input and outputs another quantum state. Thus, if quantum channel $\mathcal{N}$ takes as input density matrices in system A and outputs density matrices in system B, then $\mathcal{N}$ belongs to the set of completely positive trace preserving maps from A to B which we denote as CPTP(A $\rightarrow$ B).

We now introduce a few important types of channels: the family of depolarising channels, classical channels, the identity channel, unitary channels, and the family of replacement channels.

The identity channel which acts on system A, denoted as $\mathcal{I}_\text{A} \in$ CPTP(A $\rightarrow$ A),

preserves the input state and so is defined as:

$$\mathcal{I}_{\mathrm{A}}(\rho) \triangleq \rho \qquad \forall \rho \in \mathcal{D}(\mathrm{A})$$

The family of depolarising channels is determined by a noise parameter γ which corresponds to a uniform noise distribution. Then $\mathcal{D}_\gamma \in \mathrm{CPTP}(\mathrm{A} \to \mathrm{A})$ is defined as:

$$\mathcal{D}_\gamma(\rho) \triangleq (1-\gamma)\rho + \gamma \frac{\mathbb{I}_{|\mathrm{A}|}}{|\mathrm{A}|} \qquad \forall \rho \in \mathcal{D}(\mathrm{A})$$

Unitary channels are determined by a unitary operator U which satisfies $UU^\dagger = U^\dagger U = \mathbb{I}$ where $\dagger$ denotes the Hermitian conjugate. Such channels are typically represented as $\mathcal{U} \in \mathrm{CPTP}(\mathrm{A} \to \mathrm{A})$ where

$$\mathcal{U}(\rho) \triangleq U\rho U^\dagger \quad \forall \rho \in \mathcal{D}(\mathrm{A})$$

Classical channels are those which take diagonal inputs to diagonal outputs. That is, if $\mathcal{N} \in \mathrm{CPTP}(\mathrm{A} \to \mathrm{B})$ is a classical channel, then there exists an orthonormal basis $\{|x\rangle\}$ of system A and $\{|y\rangle\}$ for system B such that

$$\mathcal{N}(|x\rangle\langle x|) \triangleq \sum p_{y|x} |y\rangle\langle y| \quad \forall x$$

Replacement channels, also known as discarding channels, throw away the input state and prepare a fixed output state. A replacement channel $\mathcal{R}_\sigma \in \mathrm{CPTP}(\mathrm{A} \to \mathrm{B})$ is thus defined as:

$$\mathcal{R}_\sigma(\rho) \triangleq \sigma \quad \forall \rho \in \mathcal{D}(\mathrm{A})$$

evidently when A is trivial, this is equivalent to a state preparation device.

Finally, we introduce superchannels. While channels map quantum states to quantum states, superchannels map quantum channels to quantum channels. The set of superchannels which takes as an input a channel in $\mathrm{CPTP}(\mathrm{A} \to \mathrm{B})$ and outputs a

channel in CPTP(A$'$ $\rightarrow$ B$'$) may be denoted as CPTP$_s$(AB $\rightarrow$ A'B'). Such superchannels are typically denoted as $\Theta^{\text{AB}\rightarrow\text{A'B'}}$. Additionally, in the remainder of this work we will interchangeably write $\mathcal{N} \in \text{CPTP}(\text{A} \rightarrow \text{B})$ and $\mathcal{N}^{\text{A}\rightarrow\text{B}}$.

Additionally, a superchannel $\Theta^{\text{AB}\rightarrow\text{A'B'}}$ is called uniformity preserving if:

$$\Theta^{\text{AB}\rightarrow\text{A'B'}}(\mathcal{R}^{\text{A}\rightarrow\text{B}}) = \mathcal{R}^{\text{A}'\rightarrow\text{B}'}$$

where $\mathcal{R} = \mathcal{D}_{\gamma=1}$ is the completely randomising (equivalently, completely depolarising) channel.

2.5 Entropy for Classical and Quantum States

In this section we briefly overview key entropy measures. The entropy (or uncertainty) of a density matrix can be found via the von Neumann entropy, which is defined as

$$H_{\text{VN}}(\rho) \triangleq -\text{Tr}[\rho \log(\rho)]$$

In the case where ρ is a classical state, such that it is diagonal and its diagonal entries represent a probability distribution $\{p_x\}|_x$, the above entropy reduces to the Shannon entropy:

$$H_{\text{S}}(\mathbf{p}) \triangleq -\sum_x p_x \log_2(p_x)$$

Finally, the relative entropy of a quantum state ρ and a positive semidefinite operator σ is defined as:

$$D(\rho\big\|\sigma) \triangleq \text{Tr}[\rho\big(\log(\rho) - \log(\sigma)\big)]$$

In the case where both ρ and σ are classical states which can be represented by probability distributions $\mathbf{p}$ and $\mathbf{q}$ respectively, the above reduces to the Kullback-Leibler divergence:

$$D_{\mathrm{KL}}(\mathbf{p}||\mathbf{q}) \triangleq \sum_x p_x \log(\frac{p_x}{q_x})$$

3

Locally Greedy Protocols for Binary Quantum State Discrimination

Portions of this chapter are adapted from [BLS+19]. Statement on collaborative work: This work is done in collaboration with Mengke Lian, Kevin D Stubbs, Narayanan Rengaswamy, and Dr. Henry Pfister. I found the main analytical results and wrote the final manuscript. Mengke Lian improved the computer algorithms used and generated several of the numerical results. Kevin Stubbs and Narayanan Rengaswamy provided useful feedback for many ideas on this project, reviewed and edited the final manuscript, and contributed to analytical results for the modified locally greedy algorithm. Dr. Pfister oversaw the project.

3.1 Quantum State Discrimination

Measurement lies at the heart of quantum mechanics. Since the exact state of a quantum system cannot be directly observed, measurements are the means of extracting information from quantum systems. However, due to the inherent uncertainty in quantum systems, it is impossible to design a quantum measurement capable of

perfectly discriminating between two non-orthogonal quantum states [NC11b].

Quantum state discrimination is the task of finding the optimal measurement for distinguishing between a set of quantum states. Consider a quantum system whose actual state is ρ, where ρ can take on values from the set of candidate states $\{\rho_1, ..., \rho_m\}$ with corresponding prior probability vector $\mathbf{q} = \{q_1, ..., q_m\}$. This prior probability vector is defined such that $\Pr(\rho = \rho_j) = q_j$ for every $j \in \{1, ..., m\}$. Then the measurement $\hat{\Pi} = \{\Pi_j\}|_{j=1}^{m}$ will yield the success probability

$$\mathrm{P}_{\mathrm{succ}}\big(\{\rho_j\}, \mathbf{q}, \hat{\Pi}\big) = \sum_j q_j \mathrm{Tr}[\Pi_j \rho_j]$$

The aim of minimal-error quantum state discrimination is then to find the measurement $\hat{\Pi}$ which maximizes the probability of success. In the special case of *binary* quantum state discrimination, there are exactly two candidate states and the state set can be represented as $\{\rho_+, \rho_-\}$ and with a scalar prior $\Pr(\rho = \rho_+) = q$ (since $\Pr(\rho = \rho_-) = 1 - q$). In this case, the optimal (Helstrom) measurement [Hel69] has the simple description

$$\Pi \triangleq \sum_{|v\rangle \in \mathcal{V}} |v\rangle\langle v|\,, \text{ where } \mathcal{V} \triangleq \Big\{ |v\rangle \;\Big|\; \exists \lambda \geqslant 0,\ ((1-q)\rho_- - q\rho_+)\,|v\rangle = \lambda\,|v\rangle \Big\}, \quad (3.1)$$

However, for composite quantum systems composed of several subsystems, the Helstrom measurement is often impractical to implement experimentally because it requires simultaneously measuring all subsystems. This creates a need for experimentally feasible state discrimination protocols. We thus aim to achieve or approximate the Helstrom success probability via locally adaptive protocols where only a single subsystem is measured in each round and where each measurement may depend on previous measurement results.

The simplest strategy, a naïve "majority vote", has been shown to have probability of error which approaches zero exponentially fast in N [ABB$^+$05, HDB$^+$11]. Previous algebraic results [VSPM01] demonstrated that locally adaptive protocols are optimal

for pure binary state discrimination. However, a simple description of how to implement the optimal protocol was not discovered until a subsequent work [ABB$^+$05] demonstrated that a greedy adaptive strategy involving Bayesian updates of the prior after each measurement result is optimal in the special case where all subsystems are pure states and identical copies.

In this approach, after each measurement one updates the prior probability using Bayes' theorem. Then in the j^{th} round, a greedy measurement is implemented on subsystem j based on the updated prior (where a greedy measurement is a measurement that would be optimal if j were the last subsystem.)

In this chapter, we generalize previous works and consider the problem of discrimination between two general tensor product states with subsystems of arbitrary dimension and where the subsystems are in general distinct. More specifically, we suppose that we are given either ρ_+ or ρ_- with prior probability q and $1-q$ respectively, where $\rho_\pm = \hat{\rho}_\pm^{(1)} \otimes \cdots \otimes \hat{\rho}_\pm^{(N)}$ and $\hat{\rho}_\pm^{(j)}$ is potentially different for each $j \in \{1, \cdots, N\}$.

We extend the result [ABB$^+$05] and prove that if all of the systems are pure states then the order of measurement does not matter and a Bayesian update-based fixed measurement strategy is optimal. In the case where the states are not pure, we prove that the performance is not only suboptimal, but additionally plateaus as the number of subsystems increases. A similar phenomena was observed and discussed in [FBC19], where it was demonstrated that for the case of noisy qubit copies the error rate of the standard locally greedy strategy is nonzero even in the limit of infinite copies.[1] We then introduce a new modified locally greedy algorithm that exhibits asymptotically optimal performance as the number of subsystems approaches infinity.

[1] Our results were found independently of those in [FBC19]. We presented our results at TQC on 5 June, 2019; and their arXiv paper was released around the same time on 26 June, 2019. Our arXiv paper was released in December, 2019 after collecting additional results related to dyanmical programming.

We additionally conjecture that this algorithm is optimal or near-optimal for the case where there are a finite number of subsystems corresponding to identical copies of depolarized states.

3.2 Notation and Structure

Following the same notation as above, ρ is the random variable representing the given state, so that either $\rho = \rho_+ = \rho_+^{(1)} \otimes \cdots \otimes \rho_+^{(N)}$ or $\rho = \rho_- = \rho_-^{(1)} \otimes \cdots \otimes \rho_-^{(N)}$, and we refer to N as the number of subsystems. Each $\rho_\pm^{(j)}$ is in general any density matrix of finite dimensions, but we will only consider qubits and qutrits in this paper. The prior probability that the given state is ρ_+ is denoted by $q \triangleq \mathbb{P}[\rho = \rho_+]$. The number of subsystems measured jointly in each round is denoted by $m \in \{1, \ldots, N\}$, where m divides N, and $m = 1$ unless otherwise mentioned. The permutation $\sigma \in \mathcal{S}_N$, where $\mathcal{S}_N$ is the symmetric group on N elements, is unknown at the beginning of the protocol, and is defined progressively in each round (index by index) when the algorithm determines the next subsystem to measure (assuming no grouping of subsystems, i.e., $m = 1$). At round $j \in \{1, \ldots, N\}$, we determine the next subsystem $\sigma(j)$ and the action $\mathbf{a}_{\sigma(j)} \in \mathcal{A}$ on it by optimizing a cost function, then execute the action and obtain a result $d_{\sigma(j)} \in \mathcal{D}$. Here $\mathcal{A}$ is a generic action set which is specified by the type of measurements in any specific scheme, and $\mathcal{D}$ is the space containing possible outcomes for the chosen action set. For example, if $\mathcal{A}$ contains projective measurements on qubits, then $\mathcal{D} = \{\pm 1\}$. For a natural number n, define $[n] \triangleq \{1, \ldots, n\}$. Then at round j, the past actions and results are recorded into the vectors $\mathbf{a}_{[j-1]}^\sigma = (\mathbf{a}_{\sigma(1)}, \ldots, \mathbf{a}_{\sigma(j-1)})$ and $\mathbf{d}_{[j-1]}^\sigma = (d_{\sigma(1)}, \ldots, d_{\sigma(j-1)})$ respectively.

3.3 Locally Greedy Algorithm

We first describe a simple locally greedy algorithm, which was labeled the "locally optimal locally adaptive" algorithm in [HDB+11]. For $m = 1$, at each round $j \in [N]$, the algorithm updates the probability that the given state is ρ_+ based on results of past measurements. The algorithm does not consider any non-trivial ordering of the N subsystems, so $\sigma(j) = j$ for all $j \in [N]$. Once the prior is updated at round j, it performs the Helstrom measurement on the subsystem j according to the given $\rho_\pm^{(j)}$ and this updated prior. In order to formally describe this process and later generalize it to the dynamic programming-based algorithm in the next section, we begin by defining the credulity at round j for a non-trivial permutation σ on the N subsystems.

Definition 1. *The* credulity $C_j^\sigma(q, \mathbf{a}_{[j]}^\sigma, \mathbf{d}_{[j]}^\sigma)$ *is defined as the probability that* $\rho = \rho_+$ *given that the starting prior is* q*, that the first* j *rounds of measurement were executed with ordering* σ *and actions* $\mathbf{a}_{[j]}^\sigma$*, and that the results are* $\mathbf{d}_{[j]}^\sigma$*. Therefore,*

$$C_j^\sigma(q, \mathbf{a}_{[j]}^\sigma, \mathbf{d}_{[j]}^\sigma) \triangleq \mathbb{P}\left(\rho = \rho_+ \,\middle|\, \textit{prior} = q,\ \textit{actions} = \mathbf{a}_{[j]}^\sigma,\ \textit{results} = \mathbf{d}_{[j]}^\sigma\right). \tag{3.2}$$

We define the prior as the credulity when $j = 0$*, namely* $C_0^\sigma(q) \triangleq q$*.*

Then the credulity can be computed using past actions and results as

$$C_j^\sigma(q, \mathbf{a}_{[j]}^\sigma, \mathbf{d}_{[j]}^\sigma) = \frac{\mathbb{P}\left(\rho_+, d_{\sigma(j)} \middle| q, \mathbf{a}_{[j]}^\sigma, \mathbf{d}_{[j-1]}^\sigma\right)}{\mathbb{P}\left(d_{\sigma(j)} \middle| q, \mathbf{a}_{[j]}^\sigma, \mathbf{d}_{[j-1]}^\sigma\right)} \tag{3.3}$$

$$= \frac{\mathbb{P}\left(d_{\sigma(j)} \middle| \rho_+, q, \mathbf{a}_{[j]}^\sigma, \mathbf{d}_{[j-1]}^\sigma\right) \mathbb{P}\left(\rho_+ \middle| q, \mathbf{a}_{[j]}^\sigma, \mathbf{d}_{[j-1]}^\sigma\right)}{\mathbb{P}\left(d_{\sigma(j)} \middle| q, \mathbf{a}_{[j]}^\sigma, \mathbf{d}_{[j-1]}^\sigma\right)}, \tag{3.4}$$

where in the second equality we have marginalized over the possible values of ρ in the denominator.

For simplicity, we drop the prior q from $C_j^\sigma(q, \mathbf{a}_{[j]}^\sigma, \mathbf{d}_{[j]}^\sigma)$ in the following. Now we may write $C_j^\sigma(q, \mathbf{a}_{[j]}^\sigma, \mathbf{d}_{[j]}^\sigma)$ recursively as

$$C_j^\sigma(\mathbf{a}_{[j]}^\sigma, d_{[j]}^\sigma) = \frac{\mathbb{P}\left(d_{\sigma(j)} \middle| \rho_+, \mathbf{a}_{\sigma(j)}\right) C_{j-1}(\mathbf{a}_{[j-1]}^\sigma, \mathbf{d}_{[j-1]}^\sigma)}{\mathbb{P}\left(d_{\sigma(j)} \middle| q, \mathbf{a}_{[j]}^\sigma, \mathbf{d}_{[j-1]}^\sigma\right)}. \tag{3.5}$$

Now we observe that in the recursion, only the term $C_{j-1}^\sigma(\mathbf{a}_{[j-1]}^\sigma, \mathbf{d}_{[j-1]}^\sigma)$ involves the variables $\mathbf{a}_{[j-1]}^\sigma, \mathbf{d}_{[j-1]}^\sigma$ and implicitly q. Thus, we can simplify the notation by defining two quantities:

$$\mathrm{L}(q, \mathbf{a}, d) \triangleq \mathbb{P}(d \mid \rho_+, \mathbf{a}) \cdot q + \mathbb{P}(d \mid \rho_-, \mathbf{a}) \cdot (1-q), \qquad \mathrm{P}(q, \mathbf{a}, d) \triangleq \frac{\mathbb{P}(d \mid \rho_+, \mathbf{a}) \cdot q}{\mathrm{Lkhd}(q, \mathbf{a}, d)}. \tag{3.6}$$

Note that the naming follows from observing that they represent a likelihood and posterior, respectively. Thus we can write

$$\mathbb{P}\left(d_{\sigma(j)} \middle| q, \mathbf{a}_{[j]}^\sigma, \mathbf{d}_{[j-1]}^\sigma\right) = \mathrm{L}(C_{j-1}^\sigma(q, \mathbf{a}_{[j-1]}^\sigma, \mathbf{d}_{[j-1]}^\sigma), \mathbf{a}_{\sigma(j)}, d_{\sigma(j)}), \tag{3.7}$$

$$C_j^\sigma(q, \mathbf{a}_{[j]}^\sigma, \mathbf{d}_{[j]}^\sigma) = \mathrm{P}(C_{j-1}^\sigma(q, \mathbf{a}_{[j-1]}^\sigma, \mathbf{d}_{[j-1]}^\sigma), \mathbf{a}_{\sigma(j)}, d_{\sigma(j)}). \tag{3.8}$$

This completes our description of the Bayesian update for the locally greedy algorithm.

Next, we discuss the performance of this algorithm when the N subsystems are identical copies of qubits. In this case, the ordering of the subsystems is clearly immaterial. Thus, at round j, the locally greedy algorithm uses the updated credulity $p_{j-1} = C_{j-1}^\sigma(q, \mathbf{a}_{[j-1]}^\sigma, \mathbf{d}_{[j-1]}^\sigma)$ and applies the (optimal) Helstrom measurement on the j-th subsystem. For qubits, this measurement is defined by the projector

$$\Pi(p_{j-1}, j) \triangleq \sum_{|v\rangle \in \mathcal{V}(p_{j-1}, j)} |v\rangle\langle v| \tag{3.9}$$

where $\mathcal{V}(p_{j-1}, j)$ is taken to be any orthonormal basis for the space spanned by all non-negative eigenvectors of $\left((1-p_{j-1})\rho_-^{(j)} - p_{j-1}\rho_+^{(j)}\right)$.

Since $\rho_{\pm}^{(i)} = \rho_{\pm}^{(j)}$ for all $i, j \in [N]$, $\Pi(p_{j-1}, j)$ changes at every round only because of the changing p_{j-1}. The outcome probabilities for this measurement are given by

$$\mathbb{P}\left(d \mid \rho_{\pm}^{(j)}, \Pi(p_{j-1}, j)\right) = \begin{cases} 1 - \mathrm{Tr}\big[\Pi(p_{j-1}, j)\rho_{\pm}^{(j)}\big] & \text{if } d = +1, \\ \mathrm{Tr}\big[\Pi(p_{j-1}, j)\rho_{\pm}^{(j)}\big] & \text{if } d = -1, \end{cases} \tag{3.10}$$

and the overall probability of error (at round j) is given by

$$P_{\mathrm{err},j} = (1 - \mathrm{Tr}\big[\Pi(p, j)\rho_{-}^{(j)}\big]) \cdot (1 - p_{j-1}) + \mathrm{Tr}\big[\Pi(p, j)\rho_{+}^{(j)}\big] \cdot p_{j-1}. \tag{3.11}$$

Hence, we can define the probability of successfully distinguishing between states ρ_+ and ρ_- under the locally greedy algorithm as

$$P_{\mathrm{s,lg}}(q, \rho_{\pm}) \triangleq 1 - P_{\mathrm{err},N}. \tag{3.12}$$

In the special case where $\rho_{\pm}$ is a tensor product of arbitrary pure states, we now prove analytically that the locally-greedy algorithm achieves the same success probability as the optimal Helstrom measurement.

Theorem 2. *Let $P_{\mathrm{s,h}}(q, \rho_{\pm})$ and $P_{\mathrm{s,lg}}(q, \rho_{\pm})$ denote the probabilities of successful state discrimination, given initial prior $\mathbb{P}(\rho = \rho_+) = q$, using the joint N-system Helstrom measurement and the locally greedy measurement technique, respectively. If ρ_+ and ρ_- are pure states, i.e., $\rho_{\pm}^{(j)} = |{\pm}\theta_j\rangle\langle{\pm}\theta_j|$ where $|\theta\rangle \triangleq \cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}|1\rangle$, for some $\theta_j \in (0, 2\pi)$ for every $j \in [N]$, then*

$$P_{\mathrm{s,h}}(q, \rho_{\pm}) = P_{\mathrm{s,lg}}(q, \rho_{\pm}) \tag{3.13}$$

$$= \frac{1}{2}\left(1 + \sqrt{1 - 4q(1-q)\Pi_{j=1}^{N}\cos^2(\theta_j)}\right). \tag{3.14}$$

Sketch of Proof: The strategy is to prove the result for $N = 2$ and then extend via induction for arbitrary N. A complete proof is provided in Appendix 1. ■

Plateau with locally-greedy algorithm

We observe the plateau in performance using the following experimental setup (dropping for now the prior q as we assume $q = \frac{1}{2}$ in all cases unless specified otherwise):

1. Choose a set of allowed depolarizing parameters and number of trials. In this case, we choose $\mathcal{S}_{\text{dep}} = \{0.01, 0.05, 0.1, 0.3\}$ and $n_{\text{trial}} = 1000$.

2. Generate $\theta_{\pm}^{(t)} \in (0, 2\pi)$ uniformly, where $t \in [n_{\text{trial}}]$ denotes the trial index.

3. For each $\gamma \in \mathcal{S}_{\text{dep}}$, define the corresponding qubit quantum states $\rho_{\pm}(\gamma, t) \triangleq (1-\gamma)\left|\theta_{\pm}^{(t)}\right\rangle\!\left\langle\theta_{\pm}^{(t)}\right| + \frac{\gamma}{2} I$, where

$$|\theta\rangle \triangleq \cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}|1\rangle. \tag{3.15}$$

   Note that the subscript $\pm$ in $\theta_{\pm}^{(t)}$ is used to represent that the angles are chosen independently for the ρ_+ and ρ_- states.

4. For all $\gamma \in \mathcal{S}_{\text{dep}}$ and all $N = 1, 2, \ldots, 12$, we define the candidate TPQS generated by the random sampling as denoted by

$$P_{\text{succ}}(N, \gamma) = \frac{1}{n_{\text{trial}}} \sum_{t=1}^{n_{\text{trial}}} P_{\text{s,lg}}\left(\rho_{\pm}(\gamma, t)^{\otimes N}\right), \tag{3.16}$$

   where $P_{\text{s,lg}}(\rho_{\pm})$ is the success probability for the locally greedy algorithm and candidate states $\rho_{\pm}(\gamma, t)^{\otimes N}$.

   In the above, we randomly sample a set of pure states $\left\{ \left|\theta_{\pm}^{(t)}\right\rangle\!\left\langle\theta_{\pm}^{(t)}\right| \right\}|_{t=1}^{n_{trial}}$ and generate the corresponding set of candidate states $\{\rho_{\pm}(\gamma, t)^{\otimes N}\}|_{t=1}^{n_{trial}}$ for each N and γ. Thus, $P_{\text{succ}}(N, \gamma)$ represents the Monte Carlo average of performance for fixed N and γ.

We plot the results of this computational experiment in Fig. 3.1. We observe that the average probability of success (asymptotically) approaches a value strictly less than 1 when the depolarizing parameter is sufficiently high. In the limit as $\gamma = 0$, the probability of success must approach 1 with increasing N because the locally-greedy approach recovers the optimal Helstrom performance (see Theorem 2). Next, we prove a result that explains the performance plateau in Fig. 3.1 and then define

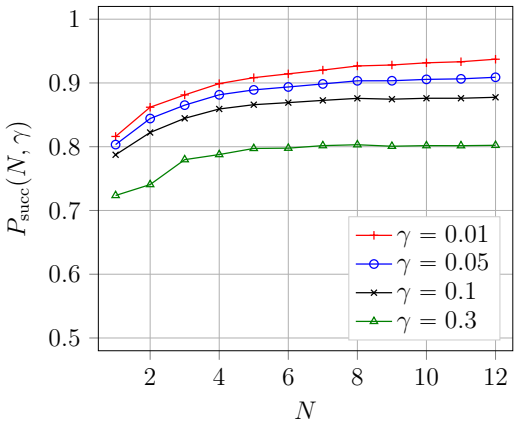

FIGURE 3.1: Comparison of probability of success for varying γ in the case of identical copies, as a function of the number of available systems. Based on the computational results, we observe that as the depolarizing parameter increases, the probability of success levels off for large N.

a modified locally-greedy approach that overcomes this suboptimality. Note that an arbitrary qubit state (density matrix) can always be expressed as a pure state passed through a depolarizing channel, because this procedure can define any state in the Bloch sphere [NC11b].

Lemma 3. *Consider two d-dimensional qudit states ρ_+ and ρ_-. Suppose that we are given ρ_+ with probability q and ρ_- with probability $1-q$ where $q \leqslant \frac{1}{2}$. The depolarized versions of $\rho_\pm$ are defined as follows:*

$$\rho_\pm^{\text{dep}} := (1-\gamma)\rho_\pm + \frac{\gamma}{d} I. \tag{3.17}$$

Consider sufficiently small γ, such that $\frac{\gamma}{1-\gamma}\frac{1-2q}{d}$ is less than the magnitude of the largest negative eigenvalue of $(1-q)\rho_- - q\rho_+$. Then if the probability of distinguishing ρ_+ and ρ_- is P_{succ}, the probability of distinguishing ρ_+^{dep} and ρ_-^{dep} is given by

$$P_{\text{succ}}^{\text{dep}} = \gamma q + \frac{\gamma(1-2q)k}{d} + (1-\gamma)P_{\text{succ}}, \tag{3.18}$$

where k is the rank of the Helstrom projector distinguishing ρ_+ and ρ_-.

Proof. The Helstrom measurement is given by the orthogonal projector onto the positive eigenspace of the operator $\left[(1-q)\rho_- - q\rho_+\right]$. More explicitly, it is given by

the orthogonal projector onto the vector space spanned by all eigenstates $|v\rangle$ such that

$$\langle v| \Big[(1-q)\rho_- - q\rho_+ \Big] |v\rangle \geqslant 0.$$

Let us denote this projector as Π_{Hel}. Using this orthogonal projector, the probability of success is given by:

$$\begin{aligned} P_{\text{succ}} &= q\text{Tr}(I - \Pi_{\text{Hel}})\rho_+ + (1-q)\text{Tr}\Pi_{\text{Hel}}\rho_- \\ &= q + \text{Tr}\Pi_{\text{Hel}}\Big[(1-q)\rho_- - q\rho_+\Big]. \end{aligned}$$

Now let us consider the optimal measurement for distinguishing ρ_+^{dep} and ρ_-^{dep}. Calculating the analogous operator for ρ_+^{dep} and ρ_-^{dep} gives

$$\begin{aligned} (1-q)\rho_-^{\text{dep}} - q\rho_+^{\text{dep}} &= (1-q)\Big[(1-\gamma)\rho_- + \frac{\gamma}{d}I\Big] - q\Big[(1-\gamma)\rho_+ + \frac{\gamma}{d}I\Big] \\ &= \frac{\gamma(1-2q)}{d}I + (1-\gamma)\Big[(1-q)\rho_- - q\rho_+\Big]. \end{aligned}$$

Since $\gamma < 1$, we can divide by $1-\gamma$ without changing the positive eigenspace. Therefore, the Helstrom optimal measurement projects onto the space of eigenstates $|v\rangle$ such that the following is positive:

$$\langle v| \left(\frac{\gamma}{1-\gamma} \frac{1-2q}{d} I + \Big[(1-q)\rho_- - q\rho_+\Big] \right) |v\rangle = \langle v| \Big[(1-q)\rho_- - q\rho_+\Big] |v\rangle \tag{3.19}$$

$$+ \frac{\gamma}{1-\gamma} \frac{1-2q}{d}. \tag{3.20}$$

Therefore, if γ is sufficiently small, then the optimal projector distinguishing ρ_+^{dep} and ρ_-^{dep} is Π_{Hel}. Hence, for γ sufficiently small, we have:

$$\begin{aligned} P_{\text{succ}}^{\text{dep}} &= q + \text{Tr}\Pi_{\text{Hel}}\Big[(1-q)\rho_-^{\text{dep}} - q\rho_+^{\text{dep}}\Big] \\ &= q + \frac{\gamma(1-2q)}{d}\text{Tr}\Pi_{\text{Hel}} + (1-\gamma)\text{Tr}\Pi_{\text{Hel}}\Big[(1-q)\rho_- - q\rho_+\Big] \\ &= \gamma q + \frac{\gamma(1-2q)k}{d} + (1-\gamma)P_{\text{succ}}. \end{aligned}$$

∎

In the case of qubits, this lemma implies the following corollary.

Corollary 4. *Consider the problem of distinguishing between two distinct single qubit states* ρ_+^{dep} *and* ρ_-^{dep} *with prior probabilities* q *and* $1-q$ *respectively. Assume that* ρ_+^{dep} *and* ρ_-^{dep} *are depolarized such that there exist pure states* $|\psi_+\rangle\langle\psi_+|$, $|\psi_-\rangle\langle\psi_-|$ *such that*

$$\gamma_\pm \in [0,1] \quad \textit{and} \quad \rho_\pm^{\mathrm{dep}} \triangleq (1-\gamma_\pm)\,|\psi_\pm\rangle\langle\psi_\pm| + \frac{\gamma_\pm}{2} I.$$

For any choice of $\gamma_\pm$, $q \in [0,1]$ *the probability of correctly distinguishing* ρ_+^{dep} *and* ρ_-^{dep}*, is denoted by* $P_{\mathrm{succ}}^{\mathrm{dep}}$ *and satisfies*

$$P_{\mathrm{succ}}^{\mathrm{dep}} \leqslant \max\left\{1-q, q, 1-\frac{\gamma_{\min}}{2}\right\} \tag{3.21}$$

where $\gamma_{\min} \triangleq \min(\gamma_+, \gamma_-)$.

Proof. Let us denote the Helstrom measurement for $\{|\psi_+\rangle\langle\psi_+|, |\psi_-\rangle\langle\psi_-|\}$ by $\Pi_{\mathrm{Hel},|\psi_\pm\rangle\langle\psi_\pm|}$ and the Helstrom measurement for $\{\rho_+^{\mathrm{dep}}, \rho_-^{\mathrm{dep}}\}$ by $\Pi_{\mathrm{Hel},\rho_\pm}$. Since $\rho_\pm^{\mathrm{dep}}$ are qubit states, $\mathrm{rank}(\Pi_{\mathrm{Hel},\rho_\pm})$ is 0, 1, or 2.

If $\mathrm{rank}(\Pi_{\mathrm{Hel},\rho_\pm}) = 0$, then $\Pi_{\mathrm{Hel},\rho_\pm} = 0$ and

$$P_{\mathrm{succ}}^{\mathrm{dep}} = q + \mathrm{Tr}\left[\Pi_{\mathrm{Hel},\rho_\pm}\Big((1-q)\rho_-^{\mathrm{dep}} - q\rho_+^{\mathrm{dep}}\Big)\right] = q.$$

If $\mathrm{rank}(\Pi_{\mathrm{Hel},\rho_\pm}) = 2$, then $\Pi_{\mathrm{Hel},\rho_\pm} = I$ and

$$P_{\mathrm{succ}}^{\mathrm{dep}} = q + \mathrm{Tr}\left[\Pi_{\mathrm{Hel},\rho_\pm}\Big((1-q)\rho_-^{\mathrm{dep}} - q\rho_+^{\mathrm{dep}}\Big)\right] = 1-q.$$

Finally, consider the case where $\mathrm{rank}(\Pi_{\mathrm{Hel},\rho_\pm}) = 1$. The state discrimination problem between $\{\rho_+^{\mathrm{dep}}, \rho_-^{\mathrm{dep}}\}$ is physically equivalent to a black box which outputs one of the following four separate discrimination problems:

$$\Big\{\big\{\,|\psi_+\rangle\langle\psi_+|\,, |\psi_-\rangle\langle\psi_-|\,\big\},\ \big\{\,|\psi_+\rangle\langle\psi_+|\,, \frac{\mathbb{I}}{2}\big\},\ \big\{\frac{\mathbb{I}}{2}, |\psi_-\rangle\langle\psi_-|\,\big\},\ \big\{\frac{\mathbb{I}}{2}, \frac{\mathbb{I}}{2}\big\}\Big\},$$

with probabilities

$$\Big\{p_1, p_2, p_3, p_4\Big\} \triangleq \Big\{(1-\gamma_+)(1-\gamma_-), (1-\gamma_+)\gamma_-, \gamma_+(1-\gamma_-), \gamma_+\gamma_-\Big\}$$

respectively. (This follows from viewing $\rho_\pm^{\text{dep}}$ as corresponding to a quantum system prepared in state $|\psi_\pm\rangle\langle\psi_\pm|$ with probability $1-\gamma_\pm$ and prepared in state $\frac{\mathbb{I}}{2}$ with probability $\gamma_\pm$.)

We denote by $P_{\text{succ}}(\rho_+, \rho_-, \Pi)$ the probability of successfully discriminating between $\{\rho_+, \rho_-\}$ given measurement $\{\Pi, \mathbb{I}-\Pi\}$ where the prior is implicitly defined as q. Then we can upper bound the success probability as

$$\begin{aligned}
P_{\text{succ}}^{\text{dep}} \leqslant & p_1 \max_{|\psi_+\rangle,|\psi_-\rangle,\Pi} P_{\text{succ}}\Big(|\psi_+\rangle\langle\psi_+|,\ |\psi_-\rangle\langle\psi_-|,\ \Pi\Big) + p_2 \max_{|\psi_+\rangle,\Pi} P_{\text{succ}}\Big(|\psi_+\rangle\langle\psi_+|,\ \frac{\mathbb{I}}{2},\ \Pi\Big) \\
& + p_3 \max_{|\psi_-\rangle,\Pi} P_{\text{succ}}\Big(\frac{\mathbb{I}}{2},\ |\psi_-\rangle\langle\psi_-|,\ \Pi\Big) + p_4 \times \frac{1}{2} \\
= & p_1 P_{\text{succ}}\Big(|0\rangle\langle 0|,\ |1\rangle\langle 1|,\ \Pi_{\text{Hel},\{|0\rangle\langle 0|,|1\rangle\langle 1|\}}\Big) + p_2 P_{\text{succ}}\Big(|0\rangle\langle 0|,\ \frac{\mathbb{I}}{2},\ \Pi_{\text{Hel},\{|0\rangle\langle 0|,|1\rangle\langle 1|\}}\Big) \\
& + p_3 P_{\text{succ}}\Big(\frac{\mathbb{I}}{2},\ |1\rangle\langle 1|,\ \Pi_{\text{Hel},\{|0\rangle\langle 0|,|1\rangle\langle 1|\}}\Big) + \frac{p_4}{2} \\
= & P_{\text{succ}}\Big((1-\gamma_+)|0\rangle\langle 0| + \frac{\gamma_+}{2}\mathbb{I},\ (1-\gamma_-)|1\rangle\langle 1| + \frac{\gamma_-}{2}\mathbb{I},\ \Pi_{\text{Hel},\{|0\rangle\langle 0|,|1\rangle\langle 1|\}}\Big).
\end{aligned}$$

Thus, the success probability for $\rho_\pm^{\text{dep}}$ is upper bounded by the success probability when $|\psi_+\rangle$ and $|\psi_-\rangle$ are orthogonal (w.l.o.g. we have set $|\psi_+\rangle = |0\rangle$ and $|\psi_-\rangle = |1\rangle$). Upon solving for $P_{\text{succ}}\Big(|0\rangle\langle 0|^{\text{dep}},\ |1\rangle\langle 1|^{\text{dep}}, \Pi_{\text{Hel},\{|0\rangle\langle 0|,|1\rangle\langle 1|\}}\Big)$, it immediately follows that:

$$P_{\text{succ}}^{\text{dep}} \leqslant (1-\frac{\gamma_+}{2})q + (1-\frac{\gamma_-}{2})(1-q) \leqslant 1 - \frac{\gamma_{\min}}{2}.$$

∎

Assuming w.l.o.g. that $q \leqslant \frac{1}{2}$, observe that $1-q \geqslant 1-\frac{\gamma}{2}$ implies $\gamma \geqslant 2q$ and therefore $\frac{\gamma}{1-\gamma}\frac{(1-2q)}{qd} \geqslant 1$ ($d=2$). In the notation of Lemma 3, set $\rho_\pm = |\psi_\pm\rangle\langle\psi_\pm|$.

Since the spectrum of $\left[(1-q)\left|\psi_-\right\rangle\left\langle\psi_-\right| - q\left|\psi_+\right\rangle\left\langle\psi_+\right|\right]$ lies in the interval $[-1,1]$, eq. (3.19) implies that the smallest eigenvalue of $\left[(1-q)\rho_-^{\text{dep}} - q\rho_+^{\text{dep}}\right]$ will now be non-negative and hence $\Pi_{\text{Hel},\rho_\pm} = I$ will be trivial. Hence, in this scenario, the Helstrom measurement is equivalent to guessing according to the prior.

In summary, this corollary implies that, for equally depolarized states, once the prior is updated so that either q or $1-q$ is greater than $1-\frac{\gamma}{2}$, the locally greedy algorithm will be stuck making trivial measurements for all subsequent subsystems and therefore the error will not approach 0 as $N \to \infty$. In Appendix 2 we show that the locally greedy method also exhibits plateaus in more general scenarios.

This result provides motivation for us to modify the conventional locally greedy method discussed above (first introduced by [ABB$^+$05] and [HDB$^+$11]). In particular, a "modified Helstrom" measurement is implemented whenever the Helstrom measurement is trivial (namely, $\Pi_{\text{Hel},\rho_\pm} \in \{I, 0\}$). In the next section, we introduce this modified locally greedy method (MLG method) and show that for arbitrary, qubit-subsystem $\rho_\pm$, we have $P_{\text{s,mlg}}(\rho_\pm) \geqslant P_{\text{s,lg}}(\rho_\pm)$ where $P_{\text{s,mlg}}(\rho_\pm)$ is the probability of successful discrimination under the MLG method. We further show that for any $\rho_\pm$, $P_{\text{s,mlg}}(\rho_\pm) \to 1$ as the number of subsystems j such that $\rho_+^{(j)} \neq \rho_-^{(j)}$ approaches infinity.

3.3.1 Modified Locally Greedy (MLG) algorithm

Like the locally greedy algorithm, the MLG algorithm updates the prior after each measurement round. Thus, it performs the modified Helstrom measurement according to the new prior. Before defining the modified Helstrom measurement, we introduce the quantities $\lambda_{\min}$ and $\lambda_{\max}$. For a given ρ_-, ρ_+, and p, these quantities are defined as:

$$\lambda_{\min} \triangleq \min_\lambda \left\{ \lambda \;\middle|\; \left((1-p)\rho_-^{(j)} - p\rho_+^{(j)}\right)\left|v_\lambda\right\rangle = \lambda\left|v_\lambda\right\rangle \right\}$$

$$\lambda_{\max} \triangleq \max_\lambda \left\{ \lambda \;\middle|\; \left((1-p)\rho_-^{(j)} - p\rho_+^{(j)}\right)\left|v_\lambda\right\rangle = \lambda\left|v_\lambda\right\rangle \right\}$$

We now define the modified Helstrom measurement as:

$$\Pi^*(p,j) \triangleq \begin{cases} \Pi(p,j) & \text{if } \Pi(p,j) \notin \{\mathbb{I}, 0\} \\ |v_{\lambda_{\max}}\rangle\langle v_{\lambda_{\max}}| & \text{if } \Pi(p,j) = 0, \\ \mathbb{I} - |v_{\lambda_{\min}}\rangle\langle v_{\lambda_{\min}}| & \text{if } \Pi(p,j) = \mathbb{I}, \end{cases}$$

where the final state is decoded as $\hat{\rho} = \rho_+$ if $C_N^{\sigma}(q, \mathbf{a}_{[N]}^{\sigma}, \mathbf{d}_{[N]}^{\sigma}) \geqslant \frac{1}{2}$ and as $\hat{\rho} = \rho_-$ otherwise.

Whenever the Helstrom measurement is nontrivial, the modified Helstrom measurement is equivalent and thus locally-optimal by definition. In the case where the Helstrom measurement is trivial, then any other measurement and outcome would lead to identical posterior-based decoding (i.e. any measurement is locally optimal). The modified Helstrom measurement takes advantage of this degeneracy to replace the trivial Helstrom measurement with a more informative measurement.

Consider the measurement given by the set of projectors $\left\{ |v_\lambda\rangle\langle v_\lambda| \;\middle|\; \left((1-p)\rho_-^{(j)} - p\rho_+^{(j)}\right)|v_\lambda\rangle = \lambda\,|v_\lambda\rangle \right\}$ and w.l.o.g. let $p \geqslant \frac{1}{2}$. Given a measurement outcome corresponding to projector $|v_\lambda\rangle\langle v_\lambda|$, the posterior-based decoding is uniquely determined by the sign of λ, with larger values of λ being a stronger predictor that $\rho = \rho_-$.

The Helstrom measurement then partitions these projectors by the sign of their eigenvalues and groups all projectors together into a trivial measurement when all eigenvalues have the same sign. In the case where the Helstrom measurement is trivial, the modified Helstrom measurement instead partitions the projectors based on the ordering of their eigenvalues. Thus, it separates out the projector that is most strongly predictive of the less-likely candidate state.

Lemma 5. *Let $\rho_\pm^{(j)}$ and p be such that $\Pi(p,j) = \mathbb{I}$ or $\Pi(p,j) = 0$.*

Then $\max_\Pi \left(\frac{p\mathrm{Tr}[\rho_+^{(j)}\Pi]}{\mathrm{Tr}[\Pi((1-p)\rho_-^{(j)} + p\rho_+^{(j)}]}\right) < \frac{1}{2}$ *or* $\min_\Pi \left(\frac{p\mathrm{Tr}[\rho_+^{(j)}\Pi]}{\mathrm{Tr}[\Pi((1-p)\rho_-^{(j)} + p\rho_+^{(j)}]}\right) \geqslant \frac{1}{2}$ *respectively. Namely, any local measurement is optimal given posterior-based decoding.*

Proof. Define $M \triangleq (1-p)\rho_- - p\rho_+$ and let the resulting projector be $\Pi_h(p, \rho_\pm) = \mathbb{I}$. Then the eigenvalues of M satisfy $\lambda_j > 0 \;\; \forall j$, and as M is Hermitian the eigenvectors

$\{|v_j\rangle\}$ are orthogonal and form a basis. Any projector diagonal in this basis may be defined $\Pi_S \equiv \sum_{j\in S} |v_j\rangle\langle v_j|$ for some set of indices S. It follows that $\mathrm{Tr}[M\Pi_s] = \sum_{j\in S}\lambda_j > 0$, so $\mathrm{Tr}[\Pi_S p\rho_+] < \mathrm{Tr}[\Pi_S(1-p)\rho_-]$. Then, the updated prior upon obtaining measurement corresponding to Π_S is:

$$p' = \frac{\mathrm{Tr}[\Pi_S\rho_+ p]}{\mathrm{Tr}[\Pi_S p\rho_+] + \mathrm{Tr}[\Pi_S(1-p)\rho_-]} < \frac{1}{2} \qquad \forall S$$

Now suppose the projector is diagonal in an arbitrary basis $\{|w_k\rangle\}$ s.t. $|w_k\rangle = \sum_j \alpha_{k,j}|v_j\rangle$ where $\{\alpha_{k,j}\}$ form the entries of some unitary operator. Then it is sufficient to show that $\mathrm{Tr}[M|w_k\rangle\langle w_k|] > 0$ for all k, since then $\mathrm{Tr}[M\sum_{k\in S}|w_k\rangle\langle w_k|] > 0$ for all S. We observe:

$$\begin{aligned}\mathrm{Tr}[M|w_k\rangle\langle w_k|] &= \sum_{j,j'}\alpha_{k,j}\alpha^*_{k,j'}\lambda_j \mathrm{Tr}[|v_j\rangle\langle v_{j'}|] \\ &= \sum_j |\alpha_{k,j}|^2\lambda_j > 0\end{aligned}$$

Similarly, for any basis $\{|w_k\rangle\}$, then $\mathrm{Tr}[M|w_k\rangle\langle w_k|] \leqslant 0$ if $\Pi_h(p,\rho_\pm) = 0$. ∎

Denote by $P_{\mathrm{s,mlg}}(q,\rho_\pm)$ the success probability of distinguishing $\{\rho_+,\rho_-\}$ with initial prior q using the MLG algorithm. We now show that the MLG method exhibits the desired asymptotic behaviour in the limit of large N. Additionally, we show $P_{\mathrm{s,mlg}}(\rho_\pm) \geqslant P_{\mathrm{s,lg}}(\rho_\pm)$ for all $\rho_\pm$ so the MLG algorithm always performs at least as well as the LG algorithm.

Corollary 6. *For any $\rho_\pm$ where $\rho_+^{(j)} \neq \rho_-^{(j)}$ for all subsystems j, then in the limit $N\to\infty$, $P_{\mathrm{s,mlg}}(q,\rho_\pm) = 1$.*

Proof. It is sufficient to show that for all $j \in \{0,1,...,N\}$ and for all $p_j \in (0,1)$ we have

$$f_+(p_j) \triangleq \mathbb{E}[p_{j+1}|\rho = \rho_+^{(j)}, \Pi^*(p_j,j)] > p_j \text{ and } f_-(p_j) \triangleq \mathbb{E}[p_{j+1}|\rho = \rho_-^{(j)}, \Pi^*(p_j,j)] < p_j,$$

and that $f_{\pm}(p_j)$ is continuous with no fixed points other than $p_j = 0$ or 1. For simplicity, we drop the superscript on $\rho_{\pm}^{(j)}$ in the following whenever the subsystem index is unambiguous. We denote the modified Helstrom measurement as $\Pi = \Pi^*(p_j, j)$, such that by definition $\mathrm{Tr}[\Pi\rho_-] > \mathrm{Tr}[\Pi\rho_+]$.

Let $x \triangleq \mathrm{Tr}[\Pi\rho_-] - \mathrm{Tr}[\Pi\rho_+] = \mathrm{Tr}[\Pi^\perp\rho_+] - \mathrm{Tr}[\Pi^\perp\rho_-]$ s.t. $x \in (0, 1]$. Then, there exists $y \in [\frac{x}{2}, 1 - \frac{x}{2}]$ such that the conditional measurement probabilities may be represented as follows:

$$\mathrm{Tr}[\Pi^\perp\rho_\pm] = y \pm \frac{x}{2}, \quad \mathrm{Tr}[\Pi\rho_\pm] = 1 - y \mp \frac{x}{2}$$

Finally, we calculate $f_+(p_j)$ as follows:

$$\begin{aligned}
f_+(p_j) &= \mathrm{Tr}[\Pi\rho_+]\Big(\frac{p_j\mathrm{Tr}[\Pi\rho_+]}{\mathrm{Tr}[\Pi(p_j\rho_+ + (1-p_j)\rho_-)]}\Big) \\
&+ \mathrm{Tr}[\Pi^\perp\rho_+]\Big(\frac{p_j\mathrm{Tr}[\Pi^\perp\rho_+]}{\mathrm{Tr}[\Pi^\perp(p_j\rho_+ + (1-p_j)\rho_-)]}\Big) \\
&= p_j\left(\frac{(1-y-\frac{x}{2})^2}{p_j(1-y-\frac{x}{2}) + (1-p_j)(1-y+\frac{x}{2})} + \frac{(y+\frac{x}{2})^2}{p_j(y+\frac{x}{2}) + (1-p_j)(y-\frac{x}{2})}\right) \\
&> p_j
\end{aligned}$$

where the final line follows from solving symbolically for the range $p_j \in (0, 1);\ x \in (0, 1);\ y \in [\frac{x}{2}, 1 - \frac{x}{2}]$. We then check for any fixed points $p_j^* = f_+(p_j)$. This results in the condition $p_j^*(-x^2 + 2p_j^*x^2 - (p_j^*x)^2) = 0$ so the only fixed points are $p_j^* = 0$ or 1. Additionally, $f_+(p_j)$ is continuous in p_j as it is differentiable for all p_j.

A similar argument holds for the case $\rho = \rho_-$. Thus, the probability of success converges to 1 under the MLG algorithm. ∎

From the above, we can conclude that $P_{s,mlg}(\rho_\pm) \geqslant P_{s,lg}(\rho_\pm)$ as the MLG and LG methods are equivalent whenever the Helstrom measurement is nontrivial. When the Helstrom measurement is trivial, it follows that the MLG method does strictly better. The improved asypmptotic behaviour of the MLG algorithm is depicted in

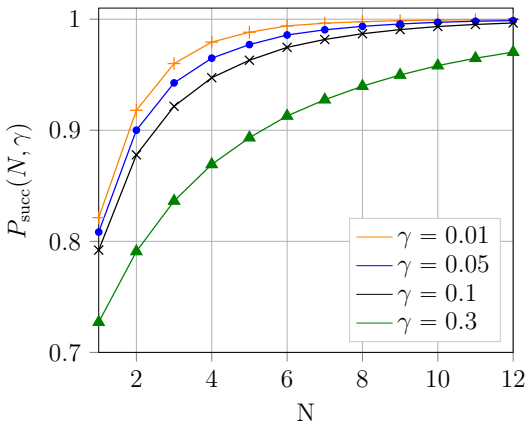

FIGURE 3.2: Comparison of probability of success for varying γ in the case of identical copies, as a function of the number of available systems using the MLG algorithm. We observe that as the depolarizing parameter increases, the probability of success no longer levels off for large N.

Fig. 3.2, where we repeat the previous experimental setup with the MLG algorithm, and plot the resulting $P_{succ}(N,\gamma) = \frac{1}{n_{trial}} \sum_{t=1}^{n_{trial}} P_{s,lg}(\rho_{\pm}(\gamma,t)^{\otimes N})$.

In the next chapter, we generalize to a dynamic programming based algorithm capable of optimizing over the order of subsystem measurement as well as the measurement performed on each subsystem.

4

Dynamic Programming-Based Locally Adaptive Protocols

Portions of this chapter are adapted from [**?**]. Statement on collaborative work: this work is done in collaboration with Mengke Lian, Kevin Stubbs, Narayanan Rengaswamy, and Dr. Henry Pfister. I found the main analytical results and wrote the final manuscript. Mengke Lian improved the computer algorithms used and generated several of the numerical results. Kevin Stubbs and Narayanan Rengaswamy provided useful feedback for many ideas on this project, reviewed and edited the final manuscript, and contributed to analytical results for the modified locally greedy algorithm. Dr. Pfister oversaw the project.

4.1 Introduction

In this chapter, we utilise dynamic programming to recursively minimize the expected future error probability over all allowed measurements, and thus yield the optimal adaptive strategy for any given family of allowable measurements [HDB+11]. This dynamic programming strategy includes the locally greedy strategies introduced in the previous chapter as a special case. Additionally, we extend the dynamic pro-

graming strategy described in [HDB+11] to optimize over not only the adaptive measurement sequence but also over the order in which subsystems are measured.

The dynamic programming (DP) algorithm performance is also tested on qutrit states, and the results demonstrate that, in general, the number of measurement outcomes affects performance (namely, ternary local measurements outperform binary local measurements).

4.2 Order-Optimized Locally Greedy Algorithm

Before introducing the most general DP algorithm, we first discuss a variant on the locally greedy algorithm where we now also choose $\sigma(j)$ carefully at each round j. We recursively compute an *expected future risk* function $R_S \colon [0,1] \to [0,1]$, where S denotes the set of subsystem indices that are yet to be measured and the domain corresponds to the current updated prior (credulity). Formally, at round j,

$$S \triangleq [N] \backslash \sigma([j-1]). \tag{4.1}$$

- For the base case $S = \varnothing$, one can make a hard decision on $C_N^\sigma(q, \mathbf{a}_{[N]}^\sigma, \mathbf{d}_{[N]}^\sigma)$, i.e., by comparing it to 0.5. Hence

$$R_\varnothing \left(C_N^\sigma \left(q, \mathbf{a}_{[N]}^\sigma, \mathbf{d}_{[N]}^\sigma \right) \right) = \min \left(C_N^\sigma \left(q, \mathbf{a}_{[N]}^\sigma, \mathbf{d}_{[N]}^\sigma \right),\ 1 - C_N^\sigma \left(q, \mathbf{a}_{[N]}^\sigma, \mathbf{d}_{[N]}^\sigma \right) \right), \tag{4.2}$$

  which can be written as function of $p \in [0,1]$ as $R_\varnothing(p) = \min(p, 1-p)$.

- For the general case $S \neq \varnothing$ and $j = N - |S| + 1$, consider $N - |S|$ measurements are performed, the goal is to choose the best subsystem to be measured next in order to minimize the expected error probability over the remaining

measurements, i.e.,

$$R_S\left(C^{\sigma}_{N-|S|}\left(q, \mathbf{a}^{\sigma}_{[N-|S|]}, \mathbf{d}^{\sigma}_{[N-|S|]}\right)\right)$$

$$= \min_{k\in S} \sum_{d_k\in\mathcal{D}} \mathbb{P}\left(d_k \middle| q, (\mathbf{a}^{\sigma}_{[N-|S|]}, \mathbf{a}_k), \mathbf{d}^{\sigma}_{[N-|S|]}\right) \tag{4.3}$$

$$\times R_{S\setminus\{k\}}\left(C^{\sigma}_{N-|S|+1}\left(q, (\mathbf{a}^{\sigma}_{[N-|S|]}, \mathbf{a}_k), (\mathbf{d}^{\sigma}_{[N-|S|]}, d_k)\right)\right), \tag{4.4}$$

where $\mathbf{a}_k = \Pi(p,k)$ is defined by (3.9), with

$$p = C^{\sigma}_{N-|S|}(q, \mathbf{a}^{\sigma}_{[N-|S|]}, \mathbf{d}^{\sigma}_{[N-|S|]}) = C^{\sigma}_{j-1}(q, \mathbf{a}^{\sigma}_{[j-1]}, \mathbf{d}^{\sigma}_{[j-1]})$$

We now allow for subsystems $\rho^{(k)}_{\pm}$ to be distinct.

This expression can be written as a function of $p \in [0,1]$ as

$$R_S(p) = \min_{k\in S} \sum_{d_k\in\mathcal{D}} \mathrm{Lkhd}(p, \Pi(p,k), d_k) \cdot R_{S\setminus\{k\}}\left(\mathrm{Post}(p, \Pi(p,k), d_k)\right). \tag{4.5}$$

Hence, during the execution of the algorithm, the next mapping for σ at round j can be defined as

$$\sigma(j) \triangleq \operatorname*{argmin}_{k\in S} \sum_{d_k\in\mathcal{D}} \mathrm{Lkhd}(p, \Pi(p,k), d_k) \cdot R_{S\setminus\{k\}}\left(\mathrm{Post}(p, \Pi(p,k), d_k)\right), \tag{4.6}$$

As for the case of identical copies, the probabilities of measurement comes are given by (3.10), the probability of error at round j is given by (3.11), and the overall probability of success of the order-optimized locally greedy algorithm is given by (3.12) with N replaced by $\sigma(N)$.

The performance plateau remains in the order-optimized algorithm

We generalize the experiment in the identical copies scenario to the case where the subsystems are distinct. The primary change is that we now sample states parameterized by $\theta^{(t,j)}_{\pm}$ so that each subsystem in both ρ_+ and ρ_- can have (potentially) distinct copies. Also, the vector of success probabilities is altered accordingly.

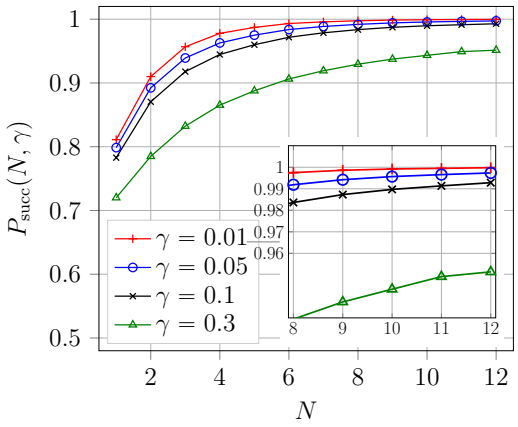

FIGURE 4.1: Comparison of probability of success for varying γ in the distinct subsystems scenario, as a function of the number of available systems, N. Results are average over 1000 trials.

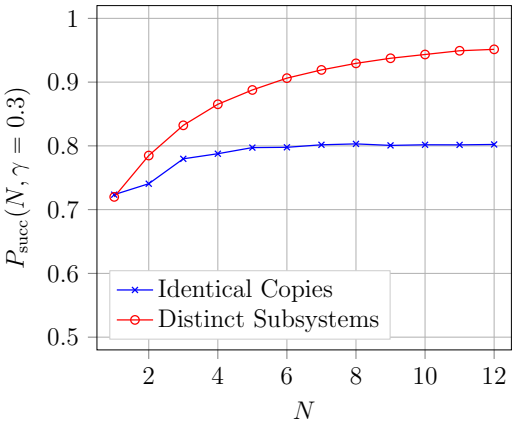

FIGURE 4.2: Comparison of probability of success as a function of the number of available systems, N, for depolarizing parameter $\gamma = 0.3$. Results are averaged over 1000 trials.

1. Choose a set of depolarizing parameters and number of trials. Again we set $\mathcal{S}_{\text{dep}} = \{0.01, 0.05, 0.1, 0.3\}$ and $n_{\text{trial}} = 1000$.

2. Set $N = 12$ and generate $\theta_{\pm}^{(t,j)} \in (0, 2\pi)$ uniformly, where $t \in [n_{\text{trial}}]$ denotes the trial index, and $j \in [N]$ denotes the subsystem index.

3. For each $\gamma \in \mathcal{S}_{\text{dep}}$ and $N \in [12]$, define the corresponding qubit quantum states

$$\rho_{\pm}(\gamma, t, N) \triangleq \bigotimes_{j=1}^{N} \left((1-\gamma) \left|\theta_{\pm}^{(t,j)}\right\rangle\!\left\langle\theta_{\pm}^{(t,j)}\right| + \frac{\gamma}{2} I \right). \tag{4.7}$$

4. For all $\gamma \in \mathcal{S}_{\text{dep}}$ and all $N \in [12]$, denote

$$P_{\text{succ}}(N,\gamma) = \frac{1}{n_{\text{trial}}} \sum_{t=1}^{n_{\text{trial}}} P_{\text{s,lg}}\left(\rho_{\pm}(\gamma, t, N)\right), \tag{4.8}$$

where $P_{\text{s,lg}}(\rho_{\pm})$ indicates that we perform the locally greedy algorithm on states $\rho_{\pm}$.

We plot the results of this experiment in Fig. 4.1.

When the subsystems are not identical copies, we notice that the plateau is much higher when compared to the case of identical copies. At first glance this appears to violate the bound obtained in Corollary 4, since here the pair of states in each subsystem are pure states depolarized with the same parameter γ, as required in the hypothesis of Corollary 4. However, the reason for the higher plateau is as follows. The algorithm orders the subsystems in such a way that the credulity can be updated to be as close to $1-\frac{\gamma}{2}$ as possible (where we assume the states $\rho_{\pm}$ may be relabeled at any step to ensure the credulity is always greater than $\frac{1}{2}$). In the next round, it is still possible to obtain one more non-trivial measurement, after which either the updated credulity exceeds $1-\frac{\gamma}{2}$ and all subsequent rounds are trivial, or the credulity is lowered below the threshold and another measurement is permitted until the updated credulity again exceeds $1-\frac{\gamma}{2}$. This permitted "jump" in credulity due to the final measurement explains why the value appearing as the plateau in Fig. 4.1 can be larger than $1-\frac{\gamma}{2}$.

The best "jump" beyond $1-\frac{\gamma}{2}$ is obtained when the states in that subsystem are an orthogonal pair of pure states subjected to the depolarizing channel and a measurement result which increases the credulity is attained, as formalized in the following lemma.

Lemma 7. *Suppose that we are given one of two quantum states* $\rho_+ = \bigotimes_{j=1}^{N}\left((1-\gamma)\left|\theta_{j,+}\right\rangle\left\langle\theta_{j,+}\right| + \frac{\gamma}{2}\mathbb{I}\right)$ *and* $\rho_- = \bigotimes_{j=1}^{N}\left((1-\gamma)\left|\theta_{j,-}\right\rangle\left\langle\theta_{j,-}\right| + \frac{\gamma}{2}\mathbb{I}\right)$ *where* γ *is a fixed depolarizing parameter. Then an upper bound on the probability of success using the locally greedy method is given by*

$$P_{\mathrm{s,lg}}(q=\frac{1}{2},\rho_{\pm}) \leqslant P_{bound}(\gamma) \equiv \frac{(1-\frac{\gamma}{2})^2}{(1-\frac{\gamma}{2})^2+(\frac{\gamma}{2})^2}.$$

Proof. We show that the probability of success is upper bounded by the maximal attainable credulity after N measurements, and proceed by induction. For details,

see Appendix 2. ∎

To illustrate the predictive value of this bound, we list the observed numerical asymptotic values found when $N = 12$ for non-identical subsystems ($P_{\text{obs}}(\gamma)$) and the predicted upper bound for $\gamma = 0.1, 0.3, 0.4, 0.5$ respectively, as depicted in the following table:

γ	P_{obs}	P_{bound}
0.1	0.9943	0.9972
0.3	0.9549	0.9698
0.4	0.9198	0.9412
0.5	0.8732	0.900

Finally, we compare the two scenarios for the specific value of the depolarizing parameter $\gamma = 0.3$ in Fig. 4.2. This plot shows the non-trivial advantage obtained from subsystems being distinct rather than copies of each other, which is the case most considered in the literature. For the special case of $\gamma = 0$, we have shown in Theorem 2 that the order of subsystems does not matter and that the simple locally greedy algorithm itself achieves the optimal performance obtained with the joint N-system Helstrom measurement.

4.2.1 Ordering and Grouping

The restriction to measurement of a single subsystem is largely artificial, so we now generalize to the case where m subsystems may be jointly measured at each round. We plot the corresponding empirical results for the scenario of identical copies and distinct subsystems in Fig. 4.3 and Fig. 4.4 respectively. When all subsystems are copies, the probability of success strictly increases as a function of m. However, when the subsystems are not copies, this no longer holds. In particular, for γ sufficiently large, the probability of success is (on average) lower for $m = 2$ and $m = 3$ than for $m = 1$ as seen for the case of $\gamma = 0.3$ in Fig. 4.5. This demonstrates the non-trivial nature of the result of measuring multiple subsystems at once. Intuitively, we might

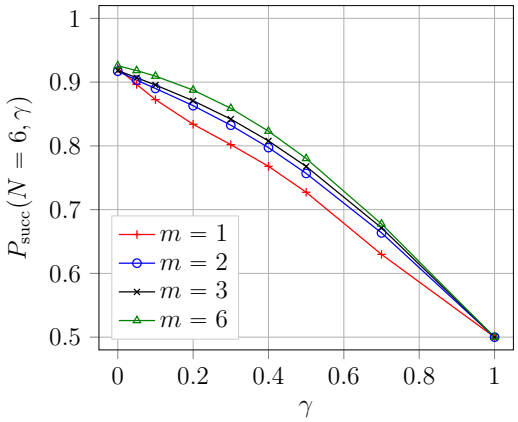

FIGURE 4.3: Probability of success for identical copies as a function of γ and number of subsystems measured simultaneously, m. Here $N = 6$. Results are averaged over 1000 trials.

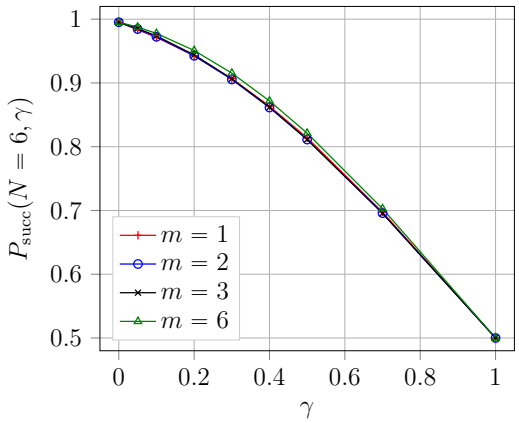

FIGURE 4.4: Probability of success for distinct subsystems as a function of γ and the number of subsystems measured simultaneously, m. Here $N = 6$. Results are averaged over 1000 trials.

expect that when $m > 1$, every measurement biases the credulity by a larger amount thereby increasing the chances of a trivial measurement in the next round. However, the difference in behavior for $m = 2$ and $m = 3$ is not clear. Hence, the complete behavior for $m > 1$ is not yet understood.

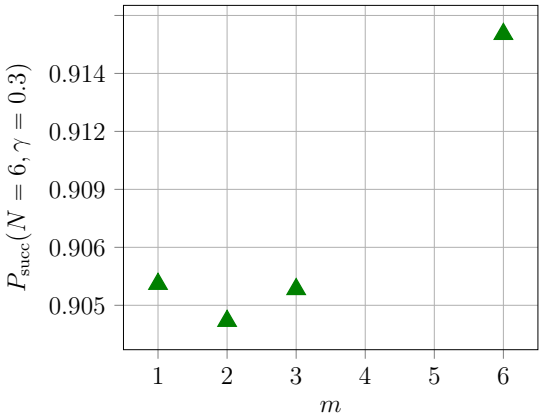

FIGURE 4.5: Probability of success for the special case $\gamma = 0.3$ when the subsystems are not necessarily copies. Here $N = 6$. Results are averaged over 1000 trials. This clearly illustrates the initial dip in probability of success with increasing m.

4.3 Measurement- and Order-Optimized DYnamic (MOODY) Algorithm

The MOODY algorithm is a generalization of the order-optimized locally greedy algorithm described above for the distinct subsystems scenario. At each round j during execution, it optimizes over all choices of $\sigma(j)$ as well as the measurement actions that could be performed over the chosen subsystem $\sigma(j)$. Hence, the expected future risk function is given by

$$R_S(p) = \min_{(k,\mathbf{a}_k)\in S\times\mathcal{A}} \sum_{d_k\in\mathcal{D}} \mathrm{L}(p,\mathbf{a}_k,d_k)\cdot R_{S\backslash\{k\}}\left(\mathrm{Post}(p,\mathbf{a}_k,d_k)\right). \tag{4.9}$$

The optimal choice for the next subsystem, $k \in S$, and the optimal action to be performed on that subsystem, $\mathbf{a}_k \in \mathcal{A}$, form the minimizer $A_S(p) = (k,\mathbf{a}_k)$ of the above function:

$$A_S(p) \triangleq \operatorname*{argmin}_{(k,\mathbf{a}_k)\in S\times\mathcal{A}} \sum_{d_k\in\mathcal{D}} \mathrm{L}(p,\mathbf{a}_k,d_k)\cdot R_{S\backslash\{k\}}\left(\mathrm{Post}(p,\mathbf{a}_k,d_k)\right). \tag{4.10}$$

Therefore, during the execution of the algorithm, at round j, we have $j = N-|S|+1$ and we set

$$(\sigma(j),\mathbf{a}_{\sigma(j)}) = A_S\left(C^{\sigma}_{j-1}\left(q,\mathbf{a}^{\sigma}_{[j-1]},\mathbf{d}^{\sigma}_{[j-1]}\right)\right). \tag{4.11}$$

The MOODY algorithm can be summarized as below. Once the DP subroutine is completed, we have a set of expected future error functions $\{R_S|S\subseteq[N]\}$ and a set of best measurement action functions $\{A_S|S\subseteq[N]\}$. Setting $p_0 = \mathbb{P}(\hat{\rho}=\hat{\rho}_+) = q$ and $S_0 = \{1,\ldots,N\}$, we then have for $i = 0,\ldots,N-1$:

$$\begin{aligned}
P^{\mathrm{err}}_i &= R_{S_i}(p_i),\\
(\sigma(i+1),\mathbf{a}_{\sigma(i+1)}) &= A_{S_i}(p_i),\\
\mathbb{P}(d_{\sigma(i+1)} = d \mid p_i,\mathbf{a}_{\sigma(i+1)}) &= \mathrm{L}(p_i,\mathbf{a}_{\sigma(i+1)},d),\\
p_{i+1} &= \mathrm{P}(p_i,\mathbf{a}_{\sigma(i+1)},d_{\sigma(i+1)}),\\
S_{i+1} &= S_i\backslash\{\sigma(i+1)\}.
\end{aligned}$$

Finally, after N rounds of measurements we can conclude that

$$\text{State Estimate} = \begin{cases} \rho_+, & \text{if } p_N > 0.5, \\ \rho_-, & \text{if } p_N \leqslant 0.5, \end{cases} \quad \text{and} \quad P_{\text{succ}} = \max(p_N, 1 - p_N). \tag{4.13}$$

Implementation and Complexity

We compute functions R_S and A_S using dynamic programming (DP), which requires collapsing the states $(q, \mathbf{a}^\sigma_{[N-|S|]}, \mathbf{d}^\sigma_{[N-|S|]}) \in [0,1] \times \mathcal{A}^{N-|S|} \times \mathcal{D}^{N-|S|}$ to $C^\sigma_{N-|S|}(q, \mathbf{a}^\sigma_{[N-|S|]}, \mathbf{d}^\sigma_{[N-|S|]}) \in [0,1]$. Moreover, these functions can be stored for later use in the following scenarios:

- problems with the same states but different initial priors;
- problems on larger systems where the current system is a subsystem of it.

Since the interval $[0,1]$ and action space $\mathcal{A}$ are continuous, they have to be quantized properly for a tractable implementation. The computation complexity and memory complexity of DP is highly dependent on this quantization. For programming implementation, we again apply a quantized version of the above DP where the input p is quantized into Q_p equi-spaced points over $[0,1]$ and the measurement action space $\mathcal{A}$ is quantized into a size-Q_a set to make the minimization over $\mathcal{A}$ tractable. To store expected future error functions $\{R_S | S \subseteq [N]\}$ and a set of best measurement action functions $\{A_S | S \subseteq [N]\}$, the memory complexity is $\mathrm{O}(2^N Q_p)$. Besides, each value is obtained from a minimization over $S \times \mathcal{A}$, so the total computation is of complexity $\mathrm{O}(2^N Q_p N Q_a)$. The number of DP functions A_S and R_S is 2^N as the order of measurement matters in general. However, DP still represents a speedup over the case of naive exhaustion for all possible orders, which has complexity $N!$.

If all the qubits are identical copies, then the ordering is immaterial, and the different subsets S with same size correspond to the same case. Therefore, in this scenario the memory complexity is $\mathrm{O}(NQ_p)$ and the computation complexity is $\mathrm{O}(NQ_pQ_a)$.

Since there are only two possible states ρ_+ and ρ_-, we can also use log-likelihood ratios (LLR) to describe the probabilities. This parameterization can greatly simplify

the computation of credulities. For $j \in [N]$, define

$$\ell_0^\sigma(q) \triangleq \ln\left(\frac{C_0^\sigma(q)}{1-C_0^\sigma(q)}\right),$$

$$\ell_j^\sigma(q, \mathbf{a}_{[j]}^\sigma, \mathbf{d}_{[j]}^\sigma) \triangleq \ln\left(\frac{C_j^\sigma(q, \mathbf{a}_{[j]}^\sigma, \mathbf{d}_{[j]}^\sigma)}{1-C_j^\sigma(q, \mathbf{a}_{[j]}^\sigma, \mathbf{d}_{[j]}^\sigma)}\right),$$

$$\tilde{\ell}_{\sigma(j)}(\mathbf{a}_{\sigma(j)}, d_{\sigma(j)}) \triangleq \ln\left(\frac{\mathbb{P}(d_{\sigma(j)}|\rho_+, \mathbf{a}_{\sigma(j)})}{\mathbb{P}(d_{\sigma(j)}|\rho_-, \mathbf{a}_{\sigma(j)})}\right).$$

It is easy to check that

$$\begin{aligned}
\exp(\ell_j^\sigma(q, \mathbf{a}_{[j]}^\sigma, \mathbf{d}_{[j]}^\sigma)) &= \frac{C_j^\sigma(q, \mathbf{a}_{[j]}^\sigma, \mathbf{d}_{[j]}^\sigma)}{1-C_j^\sigma(q, \mathbf{a}_{[j]}^\sigma, \mathbf{d}_{[j]}^\sigma)} \\
&= \frac{\mathbb{P}(d_{\sigma(j)}|\rho_+, \mathbf{a}_{\sigma(j)})\, C_{j-1}^\sigma(q, \mathbf{a}_{[j-1]}^\sigma, \mathbf{d}_{[j-1]}^\sigma)}{\mathbb{P}(d_{\sigma(j)}|\rho_-, \mathbf{a}_{\sigma(j)})\,(1-C_{j-1}^\sigma(q, \mathbf{a}_{[j-1]}^\sigma, \mathbf{d}_{[j-1]}^\sigma))} \\
&= \exp(\tilde{\ell}_{\sigma(j)}(\mathbf{a}_{\sigma(j)}, d_{\sigma(j)})) \cdot \exp(\ell_{j-1}^\sigma(q, \mathbf{a}_{[j-1]}^\sigma, \mathbf{d}_{[j-1]}^\sigma)).
\end{aligned}$$

This yields the simpler recursive equation

$$\ell_0^\sigma(q) = \ln\left(\frac{q}{1-q}\right),\ \ell_j^\sigma(q, \mathbf{a}_{[j]}^\sigma, d_{[j]}^\sigma) = \tilde{\ell}_{\sigma(j)}(\mathbf{a}_{\sigma(j)}, d_{\sigma(j)}) + \ell_{j-1}^\sigma(q, \mathbf{a}_{[j-1]}^\sigma, \mathbf{d}_{[j-1]}^\sigma). \tag{4.14}$$

4.3.1 Results for Qubits and Qutrits

The exact properties of the most general adaptive algorithm remain unknown. For the special case where $\hat{\rho}_\pm$ are both pure, we have proven in Theorem 2 that the optimal adaptive strategy consists of binary projective measurements and that its performance is unaffected by subsystem ordering. Therefore, a natural claim to test is whether adaptive binary projective measurements are always sufficient for general states. Additionally, we will address the question of whether subsystem ordering affects the probability of success when optimization is done over all "reasonable" adaptive protocols.

Qubit Results

We first address the question of ordering when $\hat{\rho}_{\pm}^{(j)}$ are all real qubit states, setting

$$\mathcal{A}_{\text{qubit}} \triangleq \left\{ \{ |\phi\rangle\langle\phi| , |\phi^{\perp}\rangle\langle\phi^{\perp}| \} \colon \phi \in \left[0, \frac{\pi}{2}\right] \right\} \tag{4.15}$$

to be the standard action space of real orthogonal projectors [HDB+11] where we quantize ϕ into $Q_{\phi} = 128$ equally spaced points. We perform the following experiments.

1. Choose a set of depolarizing parameters

   $$\mathcal{S}_{\text{dep}} = \{0, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$$

   and number of trials $n_{\text{trial}} = 1000$.

2. Generate $\theta_{\pm}^{(t,j)} \in (0, 2\pi)$ uniformly, where $t \in [n_{\text{trial}}]$ denotes the trial index, and $j = 1, 2, \ldots, 7$ denotes the subsystem index.

3. For each $\gamma \in \mathcal{S}_{\text{dep}}$ and $N \in \{3, 4, 5, 6, 7\}$, define the corresponding qubit quantum states:

   $$\rho_{\pm}(\gamma, t, N) \triangleq \bigotimes_{j=1}^{N} \left((1-\gamma) \, |\theta_{\pm}^{(t,j)}\rangle\langle\theta_{\pm}^{(t,j)}| + \frac{\gamma}{2} I \right). \tag{4.16}$$

4. For each $\hat{\rho}_{\pm}(\gamma, t, N)$ perform two separate optimizations corresponding to the best and worst ordering respectively, where the corresponding future risk functions are

   $$R_{S,\text{best}}\Big(p, \{\hat{\rho}_{\pm}(\gamma, t, N)\}\Big) = \min_{(k, \mathbf{a}_k) \in S \times \mathcal{A}} \sum_{d_k \in \{+,-\}} \mathrm{L}(p, \mathbf{a}_k, d_k) \cdot R_{S \setminus \{k\}}(\mathrm{P}(p, \mathbf{a}_k, d_k)), \tag{4.17}$$

   $$R_{S,\text{worst}}\Big(p, \{\hat{\rho}_{\pm}(\gamma, t, N)\}\Big) = \max_{k \in S} \min_{\mathbf{a}_k \in \mathcal{A}} \sum_{d_k \in \{+,-\}} \mathrm{L}(p, \mathbf{a}_k, d_k) \cdot R_{S \setminus \{k\}}(\mathrm{P}(p, \mathbf{a}_k, d_k)). \tag{4.18}$$

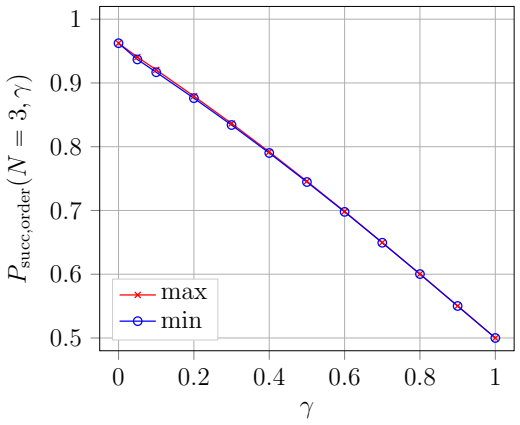

FIGURE 4.6: Comparison of probability of $P_{\text{succ,best}}(N = 3, \gamma)$ and $P_{\text{succ,worst}}(N = 3, \gamma)$ as a function of the depolarising parameter γ over 1000 trials. Although $P_{\text{succ,best}}(N = 3, \gamma) \neq P_{\text{succ,worst}}(N = 3, \gamma)$, the relative difference is small.

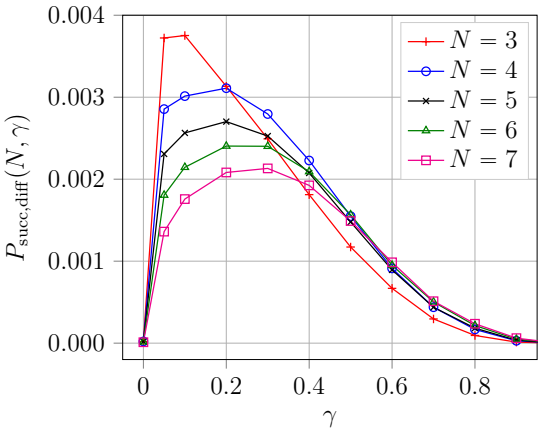

FIGURE 4.7: Comparison of difference in maximum and minimum probability of success, $P_{\text{succ,diff}}(N, \gamma)$, as a function of the depolarizing parameter γ over 1000 trials for $N = 3, 4, 5, 6, 7$.

5. For $\gamma \in \mathcal{S}_{\text{dep}}$ and $N \in \{3, 4, 5, 6, 7\}$, given an order of "best" or "worst", denote:

$$P_{\text{succ,order}}(N, \gamma) = \frac{1}{n_{\text{trial}}} \sum_{t=1}^{n_{\text{trial}}} P_{\text{s,order}}\left(\rho_{\pm}(\gamma, t, N)\right) \tag{4.19}$$

where $P_{\text{s,order}}(\rho_{\pm})$ indicates that we perform the MOODY algorithm with the specified ordering on states $\rho_{\pm}$.

We plot $P_{\text{succ,order}}(N = 3, \gamma)$ as a function of γ in Fig. 4.6 and we also compare the difference $P_{\text{succ,diff}}(N, \gamma) \triangleq P_{\text{succ,best}}(N, \gamma) - P_{\text{succ,worst}}(N, \gamma)$ for $N \in \{3, 4, 5, 6, 7\}$ in Fig. 4.7. From these results, we observe that the difference in probability of success with respect to ordering persists even using the MOODY algorithm.

Qutrit Results

Finally, we investigate whether restricting the action space to binary projectors is sufficient for non-qubit states, and in particular for qutrit states.

Definition 8. *We say that action space $\mathcal{A}$ is sufficient for state space $\mathcal{H}$ if and only if for all $\hat{\rho}_{\pm} \in \mathcal{H}$ and $q \in [0, 1]$,*

$$P_{\text{succ},\mathcal{A}}(q, \hat{\rho}_{\pm}) = P_{\text{succ},\mathcal{A}_{\text{all}}}(q, \hat{\rho}_{\pm}),$$

where $\mathcal{A}_{\text{all}}$ *is the set of all quantum measurements of appropriate dimension, i.e.,* $dim(\rho_{\pm})$.

For pure states Theorem 2 confirms that binary projectors are sufficient, and by the definition of the Helstrom measurement, binary projectors are additionally sufficient whenever $N = 1$.

We show that binary projective measurements are not sufficient for general state spaces. To this aim, we define $\mathcal{H}_{\text{qutrit}}$ to be the space of depolarized, real qutrit states and define the action space of real binary (ternary) measurements as $\mathcal{A}_{\text{b}}$ ($\mathcal{A}_{\text{t}}$).

$$\mathcal{A}_{\text{b}} \triangleq \left\{ \{\Pi_j^{\text{b}}\}_{j=1}^{2} \;\middle|\; \Pi_j^{\text{b}}\Pi_{j'}^{\text{b}} = \delta_{j,j'}\Pi_j^{\text{b}} \;\forall\; j,j' \in \{1,2\},\; \text{rank}(\Pi_1^{\text{b}}) = 2,\; \text{rank}(\Pi_2^{\text{b}}) = 1 \right\}, \tag{4.20}$$

$$\mathcal{A}_{\text{t}} \triangleq \left\{ \{\Pi_j^{\text{t}}\}_{j=1}^{3} \;\middle|\; \Pi_j^{\text{t}}\Pi_{j'}^{\text{t}} = \delta_{j,j'}\Pi_j^{\text{t}},\; \text{rank}(\Pi_j^{\text{t}}) = 1 \;\forall\; j,j' \in \{1,2,3\} \right\}. \tag{4.21}$$

Note that it is sufficient to consider real quantum measurements, as $\text{Tr}[\rho\,\Pi] = \text{Tr}[\rho\,\text{Re}(\Pi)]$ for any Hermitian projector Π. Additionally, any ternary set of orthogonal projectors may be chosen to have all elements rank 1, as any rank 2 or 3 element can be viewed as grouping the corresponding rank 1 projectors post-measurement.

Then parameterizing the action spaces is equivalent to generating (with some quantization) all orthonormal basis $\{|u_1\rangle, |u_2\rangle, |u_3\rangle\}$, and for each choice of basis defining the corresponding ternary POVM $\Pi^{\text{t}}\Big(\{|u_1\rangle, |u_2\rangle, |u_3\rangle\}\Big)$ and three corresponding binary POVMs $\Pi^{\text{b},k}\Big(\{|u_1\rangle, |u_2\rangle, |u_3\rangle\}\Big)$, for $k \in \{1,2,3\}$ as follows:

$$\Pi^{\text{t}}\Big(\{|u_1\rangle, |u_2\rangle, |u_3\rangle\}\Big) \triangleq \{|u_1\rangle\langle u_1|,\; |u_2\rangle\langle u_2|,\; |u_3\rangle\langle u_3|\}, \tag{4.22}$$

$$\Pi^{\text{b},k}\Big(\{|u_1\rangle, |u_2\rangle, |u_3\rangle\}\Big) \triangleq \left\{ \sum_{l\neq k} |u_l\rangle\langle u_l|,\; |u_k\rangle\langle u_k| \right\}, \tag{4.23}$$

We implement this quantization through the following steps:

1. We quantize the unit sphere by subdividing an icosahedron for T steps according to vector $\vec{r} = [r_1, ..., r_T]$ such that at the j^{th} step we subdivide each

segment by r_j. Then according to the Euler characteristic of convex polyhedrons [Eul58], the number of vertices after all subdivisions are complete is given by $10\prod_{i=1}^{T} r_i^2 + 2$, and we denote this set of vertices as $\mathrm{Sub}(\vec{r})$.

2. For each point (x, y, z) in $\mathrm{Sub}([2, 2, 2])$, convert to polar coordinates (ϕ, θ) according to

$$x = \sin(\theta)\cos(\phi), \quad y = \sin(\theta)\sin(\phi), \quad z = \cos(\theta).$$

3. For each pair (ϕ, θ), define the rotation matrix $R(\phi, \theta)$ as

$$R(\phi, \theta) \triangleq \begin{bmatrix} -\sin(\phi) & \cos(\phi)\cos(\theta) & \cos(\phi)\sin(\theta) \\ \cos(\phi) & \sin(\phi)\cos(\theta) & \sin(\phi)\sin(\theta) \\ 0 & -\sin(\theta) & \cos(\theta) \end{bmatrix}.$$

4. Choose Q as the resolution on the equatorial plane, let $\omega \in \{\frac{\pi q}{2Q}\}_{q=0}^{Q-1}$ and define

$$\vec{u}_1(\phi, \theta, \omega) = R(\phi, \theta)\begin{bmatrix} \cos(\omega) \\ \sin(\omega) \\ 0 \end{bmatrix}, \quad \vec{u}_2(\phi, \theta, \omega) = R(\phi, \theta)\begin{bmatrix} -\sin(\omega) \\ \cos(\omega) \\ 0 \end{bmatrix}, \quad \vec{u}_3(\phi, \theta, \omega) = R(\phi, \theta)\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

The created action spaces $\mathcal{A}_{\mathrm{t}}$ and $\mathcal{A}_{\mathrm{b}}$ are then used to compare the probability of successful discrimination in the ternary and binary cases, respectively. Using the following procedure, we demonstrate that for general real depolarized qutrit states $\{\hat{\rho}_+, \hat{\rho}_-\}$, $P_{\mathrm{succ},\mathcal{A}_{\mathrm{b}}}(\frac{1}{2}, \hat{\rho}_\pm) < P_{\mathrm{succ},\mathcal{A}_{\mathrm{t}}}(\frac{1}{2}, \hat{\rho}_\pm)$, and hence binary projective measurements are not sufficient.

1. Fix $N = 3$, and choose a set of allowed depolarizing parameters $\mathcal{S}_{\mathrm{dep}} = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6\}$ and number of trials $n_{\mathrm{trial}} = 1000$.

2. Generate $\alpha_\pm^{(t,j)}, \beta_\pm^{(t,j)} \in (0, 1)$ uniformly, where $t \in [n_{\mathrm{trial}}]$ denotes the trial index, and $j = 1, 2, \ldots, N$ denotes the subsystem index. Set $\phi_\pm^{(t,j)} = 2\pi\alpha_\pm^{(t,j)}$ and $\theta = \arccos(1 - 2\beta_\pm^{(t,j)})$, such that

$$|v(\phi, \theta)\rangle = [\sin(\theta)\cos(\phi),\ \sin(\theta)\sin(\phi),\ \cos(\theta)]$$

is uniformly distributed over the set of all unit vectors.

3. For each $\gamma \in \mathcal{S}_{\text{dep}}$, define the corresponding qutrit quantum states

$$\rho_{\pm}(\gamma, t, N) \triangleq \bigotimes_{j=1}^{N} \left((1-\gamma) \left| v(\phi_{\pm}^{(t,j)}, \theta_{\pm}^{(t,j)}) \right\rangle \left\langle v(\phi_{\pm}^{(t,j)}, \theta_{\pm}^{(t,j)}) \right| + \frac{\gamma}{3} I \right). \tag{4.24}$$

4. For each $\hat{\rho}_{\pm}(\gamma, t, N)$ perform the MOODY algorithm for $\mathcal{A}_{\text{b}}$ and $\mathcal{A}_{\text{t}}$ for both best ordering and worst ordering.

5. For $\gamma \in \mathcal{S}_{\text{dep}}$, given an order of "best" or "worst", and given an action space $\mathcal{A} \in \{\mathcal{A}_{\text{b}}, \mathcal{A}_{\text{t}}\}$, denote

$$P_{\text{succ,order}}(\gamma, \mathcal{A}) = \frac{1}{n_{\text{trial}}} \sum_{t=1}^{n_{\text{trial}}} P_{\text{s,order}} \left(\rho_{\pm}(\gamma, t, N), \mathcal{A} \right) \tag{4.25}$$

   where $P_{\text{s,order}}(\rho_{\pm}, \mathcal{A})$ indicates that we perform the MOODY algorithm over action space $\mathcal{A}$ with corresponding ordering on states $\rho_{\pm}$.

We plot the results for all four methods in Fig. 4.8, and compare the difference of the remaining three methods to the ternary, best ordering method ($P_{\text{diff,order}}(\gamma, \mathcal{A}) = P_{\text{succ,best}}(\gamma, \mathcal{A}_t) - P_{\text{succ,order}}(\gamma, \mathcal{A})$) in Fig. 4.9. We observe that the best ternary ordering is optimal compared to the best binary ordering, and therefore ordering still affects performance even in MOODY algorithm. It remains an open question whether it is sufficient to consider d rank 1 orthogonal projectors for a state space $\mathcal{H}_d$ containing d-dimensional real quantum states.

4.4 Summary

In chapters 3 and 4, we investigate simple locally-greedy and modified locally-greedy algorithms as well as more general dynamic programming algorithms for quantum state discrimination when the given states are tensor products of N arbitrary qubit or qutrit states. We prove analytically that, when the individual subsystems are pure states, the simple locally-greedy algorithm achieves the optimal performance of the joint N-system Helstrom measurement.

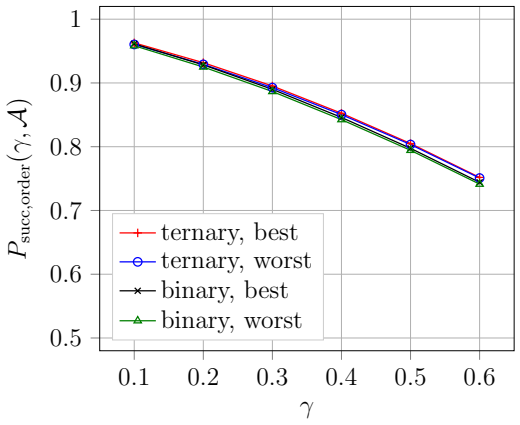

FIGURE 4.8: The average probability of success for the best and worst ordering using both ternary and binary projective measurements for qutrit product states when $N = 3$. Results are averaged over 1000 trials.

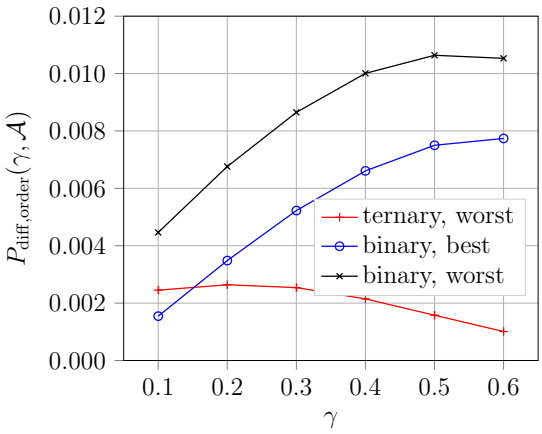

FIGURE 4.9: Difference in average success probability for the various methods, namely, $P_{\text{diff,order}}(\gamma, \mathcal{A})$ as a function of γ when $N = 3$. Results are averaged over 1000 trials.

For the scenario where each subsystem contains arbitrary qubit states, we demonstrate a plateau in the probability of success attained by the locally-greedy algorithm with increasing N. The reason for this plateau is found and an explicit bound is derived for the success probability as a function of the channel depolarizing parameter and initial prior. Based on these results, a modified locally-greedy algorithm is introduced with strictly better performance and its state discrimination becomes perfect in the large N limit.

For the general DP algorithm, we show that ordering of subsystems affects the performance when the individual subsystems have distinct states. For qutrit states, we show that binary projective measurements are inadequate to achieve optimal performance.

In the following chapter, we extend our results developing reinforcement learning algorithms to adaptively distinguish between more than two candidate states.

5

Reinforcement Learning for Quantum Hypothesis Testing

Portions of this chapter are adapted from [BSP20]. This work was done in collaboration with Kevin D. Stubbs and Dr. Henry Pfister. I developed the reinforcement learning algorithm, found the numerical results, and wrote the final manuscript. Kevin Stubbs showed the RLNN algorithm is robust under noise and contributed to writing the final manuscript. Dr. Pfister oversaw the project.

5.1 Introduction

In this chapter, we move from binary state discrimination to arbitrary state discrimination where in general one might have multiple candidate states in a state set. Unlike the Helstrom measurement for binary state discrimination, for more general state discrimination problems there is no known closed-form optimal measurement. Instead, the optimal measurement can be written as the solution of a semidefinite programming problem [Hol73, YKL75]. Techniques for solving semidefinite programming then can be used to find the minimal-error measurement and compute the

optimal success probability [KM18, KRS09].

As demonstrated in the prior chapter, dynamic programming can be used to find an optimal local approach [Bel54]. However, even in the simplest case where $m = 2$, the complexity grows like $O(2^n nQ)$, where n is the number of qubit subsystems and Q is the number of different local measurements considered [BLS$^+$19]. A powerful alternative tool for developing optimal adaptive protocols is reinforcement learning with neural networks (RLNN), where an agent learns an optimized protocol through repeated interaction with an environment. While RLNN was introduced more than 20 years ago [Tes92, Gor95], interest in these methods was recently rekindled by its remarkable success for Atari games [MKS$^+$13, MKS$^+$15]. RLNN and other machine-learning approaches have been successfully applied to a variety of problems in quantum information theory: generating error-correcting sequences [FTWM18, PDM$^+$14], preparation of special quantum states [Buk18, MPNK$^+$18, MDW20], setting up experimental Bell tests [MSS20], quantum communication [WMDB20], fault-tolerant quantum computation [SKvNE21a], quantum control [SKBL20, ZWA$^+$19, SKvNE21b, XLL$^+$19], and nonequilibrium quantum thermodynamics [SPP21]. Additionally, RLNN has been applied in the closely related topic of adaptive quantum metrology [PWS16, HS10, PWZ$^+$17, PS19]. Motivated by these successes, in this work we use RLNN to find optimal locally-adaptive measurement protocols.

We demonstrate that our approach using reinforcement learning with neural networks (RLNN) achieves excellent performance when the total number of subsystems is 10 or fewer. This performance holds even when the candidate states are subjected to small perturbations. For cases where the exact locally optimal protocol is not known, we compare the RLNN performance to an SDP upper bound and find that for each trial the RLNN comes close to the upper bound. Additionally, we introduce a min-entropy based locally adaptive approach which reduces to the optimal local

approach for binary pure states, and show that the RLNN meets or exceeds this approach in every trial.

Finally, we note that for pure states the existence of a gap between local and nonlocal strategies depends *solely* on the number of candidate states. In the previous chapters, we have extended results by [ABB$^+$05] to provide a constructive locally greedy approach which is fully optimal for any tensor product pure state discrimination problem when $m = 2$. However, prior work [CBW17, WHBC18, BDF$^+$99] has demonstrated that the simplest possible nontrivial system of two qubits with three equally likely candidate states is sufficient to find a gap between local and nonlocal measurement strategies.

5.2 Reinforcement Learning

Each round of the reinforcement learning process involves an agent choosing one action from an allowed action space, implementing the action, and receiving a reward from the environment. For a Markov decision process, the agent can eventually learn to choose actions according to an optimal policy that maximizes the expected future reward. For the problem at hand, the agent is trained to learn the optimal adaptive measurement strategy as well as the optimal adaptive order in which subsystems should be measured.

In the context of state discrimination, the environment is a parameterized measurement protocol for the quantum system of interest. The action space (denoted by $\mathcal{A}$) is the set of allowed quantum measurements. Denote by s_t the state of the environment just before round t and let n be the total number of rounds. The agent's policy, $\pi_\theta(a_t|s_t)$, is parameterized by θ and equals the probability of selecting action $a_t \in \mathcal{A}$ in round t conditioned on the state s_t of the environment. The goal of training is for the agent to learn the optimal policy π_θ^* which maximizes a given reward function.

We consider the task of deriving the minimum-error adaptive measurement protocol

to distinguish between m tensor-product quantum states $\{\rho_j\}|_{j=1}^m$ with prior probability vector $\boldsymbol{q}$ where $q_j = \Pr(\rho = \rho_j)$. To reduce the number of measurement parameters and thus the size of the action space, we restrict to the case where each candidate state is real-valued. Since each candidate state is assumed to be a tensor product of n subsystems, it can be written as

$$\rho_j = \bigotimes_{k=1}^{n} \rho_j^{(k)},$$

where $\rho_j^{(k)}$ is a qubit density matrix for all $j \in \{1, ..., m\}$ and all $k \in \{1, ..., n\}$. Thus, the quantum system ρ is composed of n unentangled qubit subsystems.

We build an OpenAI gym environment [BCP+16] capable of simulating local measurement protocols. In each round, the algorithm chooses the next subsystem j to measure as well as which measurement to implement.

The action space $\mathcal{A}$ consists of elements (ℓ, k) where $\ell \in \{1, ..., 20\}$ selects which measurement in the allowed measurement set is to be implemented and $k \in \{1, ..., n\}$ is the subsystem to be measured. More specifically, ℓ corresponds to implementing the binary real qubit POVM

$$\hat{\Pi}_Q(\ell) \triangleq \left\{ \begin{pmatrix} \sin^2(\frac{\ell\pi}{2Q}) & \frac{1}{2}\sin(\frac{\ell\pi}{Q}) \\ \frac{1}{2}\sin(\frac{\ell\pi}{Q}) & \cos^2(\frac{\ell\pi}{2Q}) \end{pmatrix}, \begin{pmatrix} \cos^2(\frac{\ell\pi}{2Q}) & -\frac{1}{2}\sin(\frac{\ell\pi}{Q}) \\ -\frac{1}{2}\sin(\frac{\ell\pi}{Q}) & \sin^2(\frac{\ell\pi}{2Q}) \end{pmatrix} \right\}$$

and $Q = 20$ (unless otherwise specified). The set of allowed measurement is then $\{\hat{\Pi}_Q(\ell)\}|_{\ell=1}^Q$ which corresponds to binary real qubit POVMs spaced evenly on the Bloch sphere. For a given Q, this action set minimizes the worst case quantization error. Increasing the quantization beyond $Q = 20$ (or allowing continuous choice of measurement) slowed the training time and did not offer observable gain in performance, so $Q = 20$ was chosen as the smallest quantization which yields near optimal results.

For a given set of candidate states, the state of the environment consists of the updated probabilities for each candidate state as well as a list of which subsystems have already been measured. Given starting prior $\mathbf{q}$ and measurement results $\mathbf{d}$, the updated prior is denoted by $p(\mathbf{q}, \mathbf{d})$. The list of subsystems which have been measured is given by the length-n vector $\boldsymbol{v}$ where $v_k = 1$ if subsystem k has already been measured and 0 else. Thus, the state of environment before each round can be represented as $s \triangleq (\boldsymbol{v}, p(\boldsymbol{q}, \boldsymbol{d}))$. The episode is terminated when all subsystems except one have been measured, or equivalently when $\sum_i \boldsymbol{v}_i = n - 1$.

When only one subsystem remains unmeasured the optimal final measurement is automatically determined through semidefinite programming. The reward is given by the probability of successfully decoding the actual state ($\rho = \rho_{j*}$) after the final local measurement, where a successful decoding occurs if

$$j^* = \underset{j \in \{1,\ldots,m\}}{\operatorname{argmax}} (p_j(\boldsymbol{q}, \boldsymbol{d})),$$

where $\boldsymbol{d}$ is the vector containing all previous measurement results and $\boldsymbol{p}(\boldsymbol{q}, \boldsymbol{d})$ is the updated probability given initial prior $\mathbf{q}$ and previous measurement results $\boldsymbol{d}$. Additionally, a penalty of -0.3 is given if at any round the agent attempts to re-measure an already measured subsystem, as for qubit subsystems re-measuring an already measured subsystem is non-informative.

5.3 Details of Implementation

We train the agent using the proximal policy optimization (PPO) algorithm [SWD$^+$17]. Results are then generated using the default PPO algorithm from the RLlib package included in Ray version 0.7.3 [LLN$^+$18, LLM$^+$17]. After hyperparameter tuning of the learning rate, we set the learning rate to be $\eta = 5 \times 10^{-5}$. For the remaining hyperparameters, we find the default parameter settings to be optimal, including the clipping parameter $\epsilon = 0.3$ and the discount factor $\gamma = 0.99$. Comparison of the

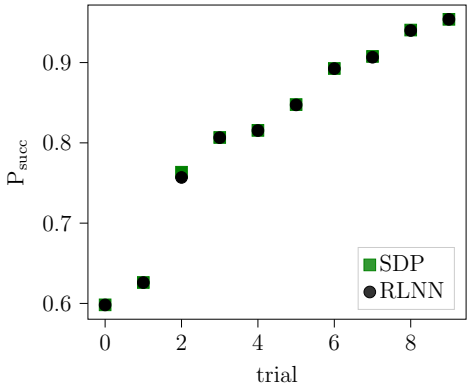

FIGURE 5.1: Performance of tuned network versus the collective Helstrom (SDP) measurement.

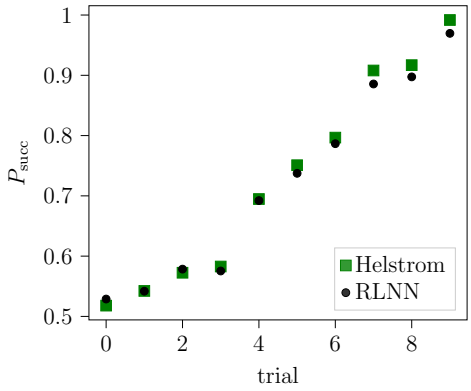

FIGURE 5.2: Performance of network before tuning versus the collective Helstrom measurement.

tuned versus untuned network is depicted in Figures 1 and 2.

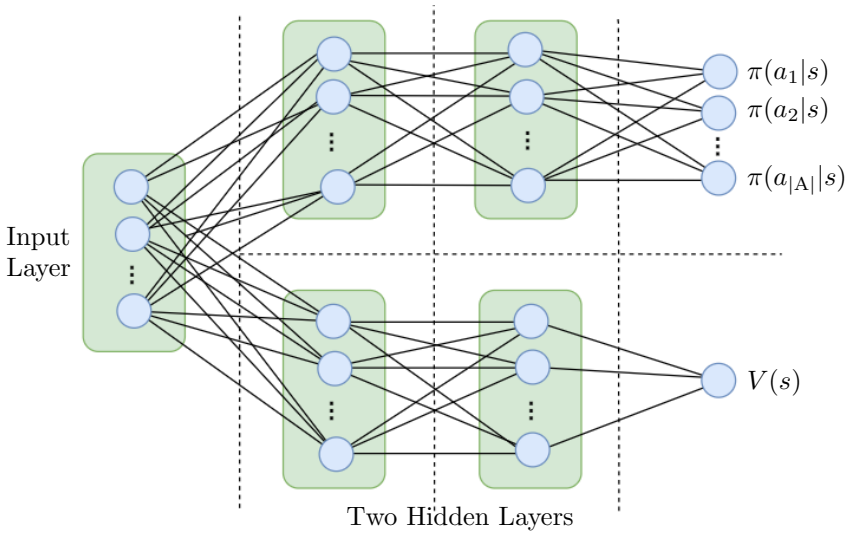

FIGURE 5.3: Neural Network configuration, consisting of one input layer, two parallel subnetworks. One subnetwork outputs an estimate of the value $V(s)$ for state s, one outputs an estimate of the policy $\pi(a|s)$.

5.4 Numerical Results for RLNN Performance

As an initial benchmark of the RLNN performance, we compare it to known optimal results in several special cases.

In the case of binary discrimination (i.e. $m = 2$) between tensor products of pure

states such that $\rho_j^{(k)} = |\psi_j^{(k)}\rangle\langle\psi_j^{(k)}|$ for all $k \in \{1, ..., n\}$ and $j \in \{1, 2\}$, it has been shown that the optimal collective success probability, $\mathrm{P}_{\mathrm{SDP}}$ can be achieved through locally-adaptive strategies [ABB+05, BLS+19]. The collective success probability, $\mathrm{P}_{\mathrm{SDP}}$, is found using semidefinite programming techniques introduced by [EMV03]. We randomly generate ten trials with $n = 3$ and order the trials according to increasing distinguishability measured by $\mathrm{P}_{\mathrm{SDP}}$. For each trial, we compare this success probability with the RLNN success probability, $\mathrm{P}_{\mathrm{RLNN}}$, as shown in Fig. 5.4. The neural network attains the correct (optimal) success probability in each case, with a very small gap that is likely due to action space quantization.

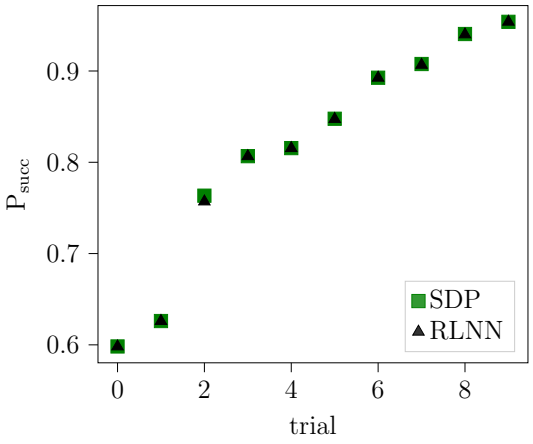

FIGURE 5.4: Probability of success for the optimal RLNN policy after 1000 training iterations vs. the optimal collective measurement for tensor-products of pure states when $m = 2$, $n = 3$. The neural network closely approximates the optimal success probability in each trial.

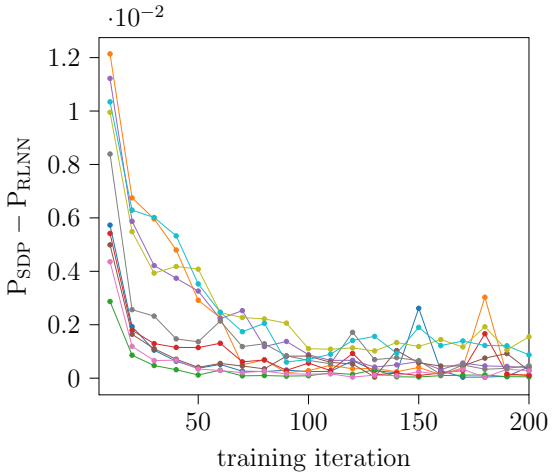

FIGURE 5.5: Difference between RLNN reward and Helstrom (SDP) success probability as a function of training iteration. We observe that the RLNN success probability stabilizes after 100 training iterations, with occasional fluctuations.

An additional case where locally adaptive protocols are strictly optimal has been found by Sasaki et. al in [SKIH98]. Consider a set of states $\mathcal{S}_1 \triangleq \{\rho_j\}|_{j=1}^m$ and associated probabilities $\{q_j\}|_{j=1}^m$. Suppose the known optimal POVM for these is

$\{\Pi_j\}|_{j=1}^m$. The set of n-subsystem product states generated by $\mathcal{S}$ can be written as

$$\mathcal{S}_n \triangleq \left\{ \bigotimes_{j=1}^{n} \rho_{i_j} \;\middle|\; i \in \{1, ..., m\}^n \right\},$$

with corresponding probabilities defined as $q_{i_1 \ldots i_n} \triangleq q_{i_1} \times ... \times q_{i_n}$. Then, the optimal POVM candidate state set $\mathcal{S}_n$ has elements that can be written in tensor product form as:

$$\Pi_{i_1 \ldots i_n} = \bigotimes_{j=1}^{n} \Pi_{i_j}.$$

This provides a useful test of the neural network performance. We take the initial state set to be $\mathcal{S}_1 = \{\rho_1, \rho_2\}$, where $\rho_1 = \left(\begin{smallmatrix} 0.85 & 0 \\ 0 & 0.15 \end{smallmatrix}\right)$ and $\rho_2 = \left(\begin{smallmatrix} 0.15 & 0 \\ 0 & 0.85 \end{smallmatrix}\right)$. Since the optimal local POVM belongs to the allowed action set, there should be no quantization loss. We train the neural network for 1000 iterations (using a custom learning rate schedule where the learning rate starts at 5.5×10^{-5} and decays by 0.95 every 10 iterations), and compare the neural network performance after training to the optimal success probability. For $1 \leqslant n \leqslant 8$, the neural network attains or approximately attains the exact success probability, as depicted in Fig. 5.6.

5.5 Comparison to SDP-based Locally Adaptive Strategies

Just as the collective SDP measurement provides an upper bound for the optimal locally adaptive success probability, simple locally adaptive algorithms such as locally greedy algorithms provide a lower bound. In this section, we introduce a local SDP-based approach, and demonstrate that the RLNN always meets or exceeds the success probability of the local SDP-based approach.

In the case of binary state discrimination, locally greedy algorithms are optimal for pure states and close-to-optimal for mixed states. Our choice of the local SDP-based algorithm as a "good" simple strategy is then motivated by the fact that

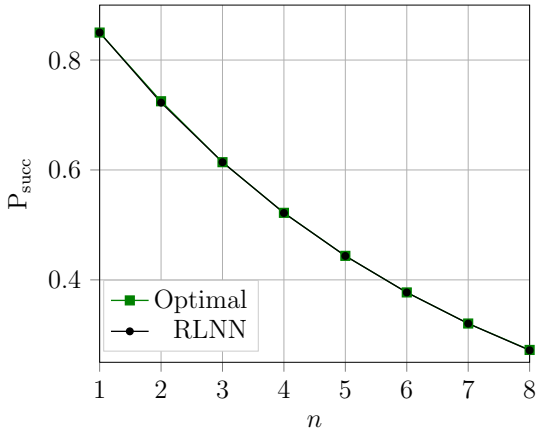

FIGURE 5.6: Performance of the RLNN policy after 150 training iterations vs. the optimal success probability as a function of the number of subsystems n. The RLNN approach converges to the optimal local approach in this example for $1 \leqslant n \leqslant 8$.

it reduces to a locally greedy protocol when $m = 2$. Additionally, for $m > 2$, we found through numerical simulation that the local SDP-based based approach generally performs better than locally greedy protocols. Finally, we compare the RLNN algorithm to the local SDP-based algorithm via simulations with $n = 3$ and $n = 4$ and demonstrate that the RLNN always meets or significantly exceeds the local SDP-based success probability.

The SDP-based local algorithm selects the local measurement which maximizes the expected (collective) success probability of future rounds. Let $\mathcal{S}$ be the set of remaining unmeasured subsystems. For each round, the algorithm chooses to measure the subsystem $l \in \mathcal{S}$ and implement the measurement a such that:

$$(a, l) = \underset{(a,\ell)\in\mathcal{A}\times\mathcal{S}}{\operatorname{argmax}} \sum_{d'=0}^{|a|} \Pr(d_{n-|\mathcal{S}\setminus\ell|} = d') \times \mathrm{P}_{\text{succ, coll}}\left(\{\rho_j^{\mathcal{S}\setminus l}\} \mid q, d_{n-|\mathcal{S}\setminus\ell|} = d'\right)$$

where $\mathrm{P}_{\text{succ, coll}}\left(\{\rho_j^{\mathcal{S}\setminus l}\} \mid q, d_{n-|\mathcal{S}\setminus\ell|} = d'\right)$ equals the success probability of implementing an optimal collective measurement on the remaining subsystems (with indices

belonging to set $\mathcal{S}\backslash l$), given the prior for round j was q and the outcome of action a was d'.

In the special case where $m = n = 3$, all candidate states are pure states, and all subsystems identical copies, the performance of the SDP-based local algorithm and the RLNN algorithm appear to be identical for each of 5 random trials, as depicted in Fig. 5.7. However, we demonstrate that simpler locally adaptive strategies such as the min-entropy approach are not sufficient to find the optimal locally adaptive strategy, as when $n = 4$ and the candidate states are mixed, a significant gap appears between the RLNN results and the SDP-based local algorithm, as shown in Fig. 5.8.

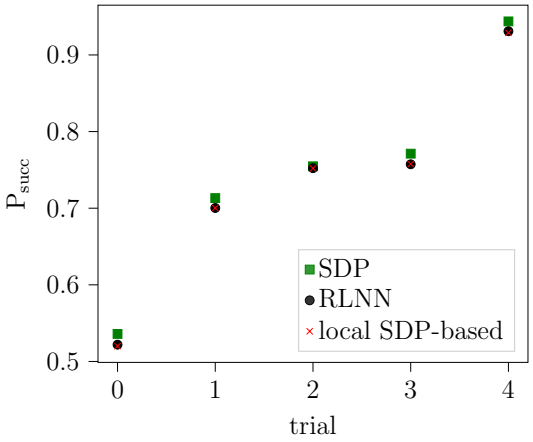

FIGURE 5.7: Plot of success probability for RLNN after 250 training iterations, the collective optimal (SDP) measurement, and the SDP-based local algorithm, over 5 trials with $m = 3$, $n = 3$ and pure states.

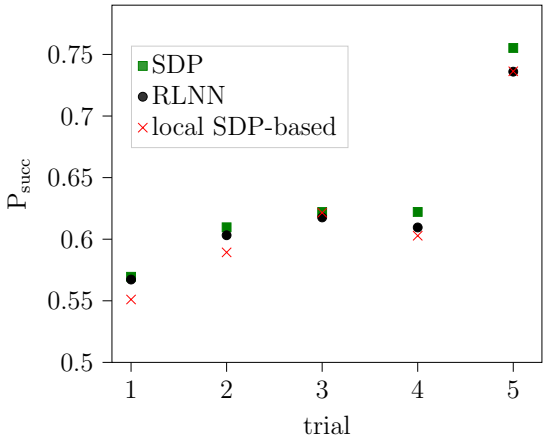

FIGURE 5.8: Plot of success probability for RLNN after 250 training iterations, the collective optimal (SDP) measurement, and the SDP-based local algorithm, over 5 trials with $m = 3$, $n = 5$.

5.6 Pure State Discrimination

In the special case of binary state discrimination ($m = 2$), it has been shown [ABB$^+$05, BLS$^+$19] that locally-greedy algorithms are optimal for distinguishing between pure tensor product states. Thus, for pure binary state discrimination, the success probability of the optimal collective measurement can be achieved through a simple locally-

greedy algorithm.

This optimality result does not extend to $m \geqslant 3$ even in the simplest case of a two-qubit system, as the trine state ensemble provides a counterexample. In particular, the unique locally-greedy algorithm is no longer the optimal local algorithm. Additionally, for the trine state ensemble, the optimal locally adaptive algorithm performs worse than the optimal collective measurement. This raises the question, "what is the maximum possible gap in success probability between the optimal locally adaptive measurement strategy and optimal collective measurement as a function of m and n?"

We now demonstrate that a gap exists even for the case where $n = 2$ and $m = 3$. Original results for the double trine were found in [CBW17]. We present an alternate proof which utilizes the results of subsequent work [WHBC18]. The trine ensemble is defined to be symmetric with

$$\begin{aligned}\rho_j &\triangleq \left(U^j \left|0\right\rangle\!\left\langle 0\right| \left(U^j\right)^\dagger\right)^{\otimes 2} \\ &= (U \otimes U)^j \left|00\right\rangle\!\left\langle 00\right| \left((U \otimes U)^j\right)^\dagger,\end{aligned}$$

where $U \triangleq \begin{pmatrix} \cos(\frac{2\pi}{3}) & -\sin(\frac{2\pi}{3}) \\ \sin(\frac{2\pi}{3}) & \cos(\frac{2\pi}{3}) \end{pmatrix}$ and $\boldsymbol{q} = [1/3, 1/3, 1/3]$.

Since $(U \otimes U)^m = \mathbb{I}$, and the starting prior is balanced, the PGM is optimal [Ban97], with a corresponding collective success probability of $P_{\text{coll}} = \frac{1}{6}(3 + 2\sqrt{2}) \approx 0.971$. Hence to prove the theorem, it is enough to show that any choice of local measurements will give a probability of success which is strictly less than $\frac{1}{6}(3 + 2\sqrt{2})$.

We now show that the optimal locally adaptive method is to measure the first subsystem with an "anti-trine" measurement. The anti-trine is defined as $\hat{\Pi}_{\text{AT}} = \{\frac{2}{3} U^{\frac{1}{2}} \rho_j (U^{\frac{1}{2}})^\dagger\}|_{j=1}^3$, and is depicted in Figure 5.9. After obtaining the measurement outcome for the first subsystem, the updated prior is a permutation of $\mathbf{q} = [\frac{1}{2}, \frac{1}{2}, 0]$,

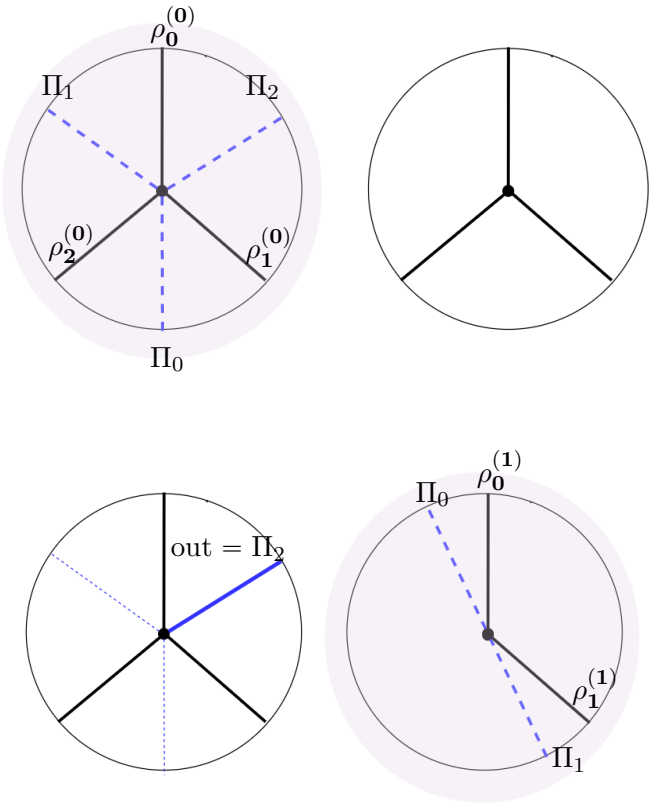

FIGURE 5.9: The top picture illustrates the optimal first (anti-trine) measurement with blue dotted lines and trine states with black lines. The bottom picture illustrates the optimal second measurement conditioned on the first measurement result, namely, a Helstrom measurement on the two remaining states.

and the second subsystem is measured according to the optimal measurement for the remaining two candidate states (see Figure 5.9).

The most general local approach is to implement measurement $\hat{\Pi}_1 = \left\{\Pi_{1,j}\right\}|_{j=1}^{m}$ on the first subsystem. We may label the result of the first subsystem out_1. Then the second and last measurement can be chosen as $\hat{\Pi}_2(\text{out}_1) = \left\{\Pi_{2,j}(\text{out}_1)\right\}|_{j=1}^{3}$ and is in general allowed to depend on the outcome out_1. It is conventional to label the elements of the second measurement such that state ρ is decoded as ρ_j if measurement element j is obtained in the final round, and thus it is sufficient to consider 3-element measurement outcomes in the second and final round given

that there are three candidate states. (Note that a binary measurement may be represented as a three-element measurement with one element set to be zero.)

The success probability of this approach is then:

$$\mathrm{P}_{\mathrm{succ}}\Big(\hat{\Pi}_1, \{\hat{\Pi}_2(\mathrm{out}_1)\}\Big)$$

$$= \sum_{j=0}^{2} \Pr(\rho = \rho_j) \sum_{d_1} \Pr[\mathrm{out}_1 = \Pi_{d_1}^{(1)}, \mathrm{out}_2 = \Pi_j^{(2)}(d_1) \;\Big|\; \rho = \rho_j]$$

$$= \sum_{j=0}^{2} \sum_{d_1} \Pr\Big[\rho = \rho_j \;\Big|\; \mathrm{out}_1 = \Pi_{d_1}^{(1)}, \mathrm{out}_2 = \Pi_j^{(2)}(d_1)\Big] \Pr\Big[\mathrm{out}_1 = \Pi_{d_1}^{(1)}, \mathrm{out}_2 = \Pi_j^{(2)}(d_1)\Big]$$

Suppose the first outcome is d_1 for measurement $\hat{\Pi}_1$. Then the conditional success probability evidently becomes:

$$\mathrm{P}_{\mathrm{succ}}\Big(\hat{\Pi}^{(2)}(d_1)\Big|\Pi_{d_1}^{(1)}\Big)$$

$$= \sum_{j=0}^{2} \Pr\Big[\rho = \rho_j \big| \mathrm{out}_2 = \Pi_j^{(2)}(d_1), \mathrm{out}_1 = \Pi_{d_1}^{(1)}\Big] \Pr\Big[\mathrm{out}_2 = \Pi_j^{(2)}(d_1) \big| \mathrm{out}_1 = \Pi_{d_1}^{(1)}\Big]$$

$$= \sum_{j=0}^{2} \Pr\Big[\rho = \rho_j \big| \mathrm{out}_1 = \Pi_{d_1}^{(1)}\Big] \Pr\Big[\mathrm{out}_2 = \Pi_j^{(2)}(d_1) \big| \mathrm{out}_1 = \Pi_{d_1}^{(1)}, \rho = \rho_j\Big]$$

Thus, in the most general locally adaptive approach, the probability of successful decoding given that the first system was measured with $\hat{\Pi}_1$ and the outcome obtained is out_1 is simply the probability of successfully distinguishing between the states of the second qubit given the updated priors $\mathbf{q}_{d_1} = \Big\{\Pr(\rho = \rho_j | \mathrm{out}_1 = d_1)\Big\}|_{j=0}^{2}$.
It follows that the second measurement is immediately determined to be the optimal measurement for $\{U^j |0\rangle\langle 0| (U^j)^\dagger\}|_{j=0}^{2}$ given the updated prior from the first

measurement. What remains is to optimize over the first measurement.

$$\begin{aligned}\mathrm{P}_{\mathrm{succ}}\Big(\hat{\Pi}_1, \big\{\hat{\Pi}_2(d_1)\big\}\Big) &= \sum_{d_1=1}^{m} \Pr[\mathrm{out}_1 = \Pi_{d_1}^{(1)}]\mathrm{P}_{\mathrm{succ}}\Big(\hat{\Pi}^{(2)}(d_1)\Big|\mathrm{out}_1 = \Pi_{d_1}^{(1)}\Big) \\ &\leqslant \max_{d_1} \mathrm{P}_{\mathrm{succ}}\Big(\hat{\Pi}^{(2)}(d_1)\Big|\mathrm{out}_1 = \Pi_{d_1}^{(1)}\Big)\end{aligned}$$

Evidently, the second line presents an upper bound on the success probability of *any* locally adaptive strategy, as it gives the success probability of the best possible two outcomes.

Let the first step be the anti-trine measurement $\hat{\Pi}_{\mathrm{AT}}$ and denote with $\{\hat{\Pi}_2^*(d_1)\}$ our conjectured optimal second step. It then follows that a sufficient condition for the optimality of this method is

$$\begin{aligned}\mathrm{P}_{\mathrm{succ}}\Big(\hat{\Pi}_{\mathrm{AT}}, \big\{\hat{\Pi}_2^*(d_1)\big\}\Big) &= \max_{\{\hat{\Pi}^{(2)}(d_1)\},\Pi_{d_1}^{(1)}} \left(\mathrm{P}_{\mathrm{succ}}\Big(\hat{\Pi}^{(2)}(d_1)\Big|\mathrm{out}_1 = \Pi_{d_1}^{(1)}\Big)\right) \\ &\geqslant \max_{\{\hat{\Pi}^{(2)}(d_1)\},\hat{\Pi}^{(1)}} \left(\mathrm{P}_{\mathrm{succ}}\Big(\hat{\Pi}_1, \big\{\hat{\Pi}_2(d_1)\big\}\Big)\right)\end{aligned}$$

By symmetry, we know the expected success probability for the anti-trine is equivalent regardless of which outcome is obtained, and can immediately compute $\mathrm{P}_{\mathrm{succ}}(\hat{\Pi}_{\mathrm{AT}}) = 0.933$. From [WHBC18], when the outcome for the first subsystem is $\Pi(\theta) = \begin{pmatrix} \sin^2(\theta) & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \cos^2(\theta) \end{pmatrix}$, the expected success probability given optimal choice of subsequent measurement is given by:

$$\begin{aligned}\max_{\{\hat{\Pi}^{(2)}(d_1)\}} \left(\mathrm{P}_{\mathrm{succ}}\Big(\hat{\Pi}^{(2)}(\theta)\Big|\mathrm{out}_1 = \Pi(\theta)\Big)\right) &= \frac{1}{3} - \frac{1}{12}\cos(2\theta) - 0.288675\cos(\theta)\sin(\theta) \\ &\quad + \frac{1}{2}\sqrt{\frac{5}{12} - \frac{1}{6}\cos(2\theta) - 0.57735\cos(\theta)\sin(\theta)}.\end{aligned}$$

when the optimal measurement on the second subsystem has two outcomes. We can directly verify from [WHBC18] that for the limited range of θ such that the optimal

measurement on the second subsystem has three outcomes, then

$$\max_{\{\hat{\Pi}^{(2)}(d_1)\}} \left(\mathrm{P}_{\mathrm{succ}} \Big(\hat{\Pi}^{(2)}(\theta) \Big| \mathrm{out}_1 = \Pi(\theta) \Big) \right) \leqslant 0.85$$

Finally, it follows from the above that:

$$\begin{aligned} \max_{\{\hat{\Pi}^{(2)}(d_1)\},\Pi^{(1)}_{d_1}} \left(\mathrm{P}_{\mathrm{succ}} \Big(\hat{\Pi}^{(2)}(d_1) \Big| \mathrm{out}_1 = \Pi^{(1)}_{d_1} \Big) \right) &= \max_{\theta} \Big(\frac{1}{3} - \frac{1}{12}\cos(2\theta) - 0.288675\cos(\theta)\sin(\theta) \\ &\quad + \frac{1}{2}\sqrt{\frac{5}{12} - \frac{1}{6}\cos(2\theta) - 0.57735\cos(\theta)\sin(\theta)} \Big) \\ &= \mathrm{P}_{\mathrm{succ}} \Big(\hat{\Pi}_{\mathrm{AT}}, \big\{ \hat{\Pi}^{*}_{2}(d_1) \big\} \Big) \end{aligned}$$

thus proving optimality of the conjectured local method so that $P_{\mathrm{loc}}(\{\rho_j\}, \mathbf{q}) = 0.933$. Clearly, $\mathrm{P}_{\mathrm{loc}}(\{\rho_j\}, \mathbf{q}) < \mathrm{P}_{\mathrm{coll}}(\{\rho_j\}, \mathbf{q})$. ■

We note that the locally greedy method in this case consists of measuring the first subsystem with a local trine measurement, and then measuring the second subsystem with the optimal local measurement for the updated prior. This locally greedy method has a success probability of $\mathrm{P}_{\mathrm{greedy}} = 0.8$, substantially lower than both the collective and optimal locally adaptive success probability.

5.7 Gap between Locally Optimal Algorithm and Collective Measurement

Finally, we use RLNN to estimate the gap between the best locally adaptive algorithm and the optimal collective (non-local) measurement in more general cases where the best locally adaptive algorithm is not otherwise known.

The simulation setup for a given m and n is as follows: for each trial, we randomly generate pure tensor product candidate states and then apply depolarizing noise with a randomly chosen noise parameter. The RLNN algorithm is independently trained 5 times over 2000 iterations, and the average final success probability is compared

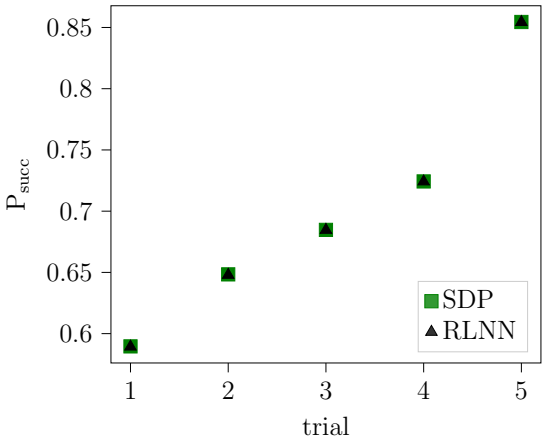

FIGURE 5.10: Probability of success for SDP and RLNN when $m = 2$ and $n = 3$. For each trial, the RLNN success probability is computed by separately training the neural network five times with 2000 iterations each. The error bars, were they visible, would represent the standard deviation in the final success probability over the five independent trainings. But in all trials, the error bars have collapsed to nothing, and the gap between local and non-local measurements is very small.

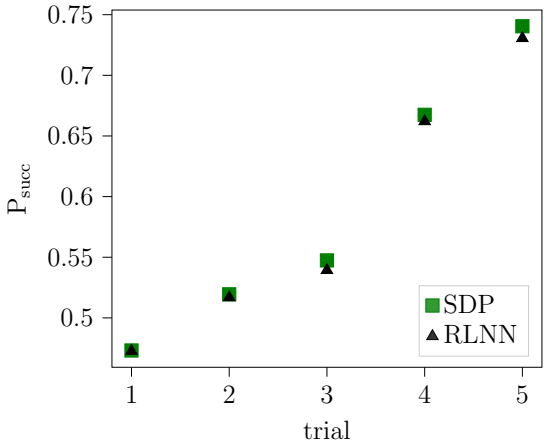

FIGURE 5.11: Probability of success for SDP and RLNN after 2000 training iterations when $m = 3$, $n = 3$. For each trial, the RLNN success probability is computed by separately training the neural network five times with 2000 iterations each. In all trials, the error bars have collapsed to nothing and the gap between local and non-local measurements is very small. Compared to the case where $m = 2$, there is a slightly larger gap between local and non-local measurements.

(with error bars) to the optimal collective success probability found via SDP. Results are plotted for $m = 2$, $n = 3$ in Figure 5.10 and for $m = 3$, $n = 3$ in Figure 5.11, and indicate that the gap between local and collective measurements increases with m.

5.8 Performance for a Large Number of Subsystems

We examine how the RLNN performance varies as a function of n, and demonstrate good performance for up to $n = 10$ subsystems. However, for $n \geqslant 20$, the RLNN begins to have suboptimal performance.

First, we consider the case of binary pure state discrimination, where a locally-greedy (LG) technique is known to be optimal. We restrict the LG algorithm to the

same action space as the RLNN to remove any gap due to action space quantization, and compare the resulting success probabilities. Results are depicted in Fig. 5.13, and indicate that the RLNN matches or almost matches the LG algorithm when $n = 10$ but develops a performance gap for $n = 20$.

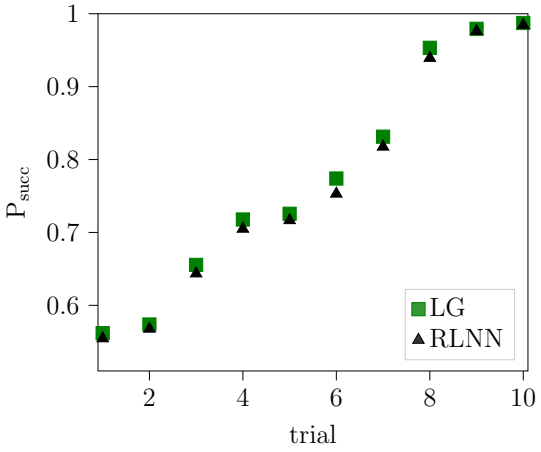

FIGURE 5.12: Success probability for $m = 2$, $n = 10$ where all candidate states are pure. Success probability for the RLNN is based on 750 training iterations for each round.

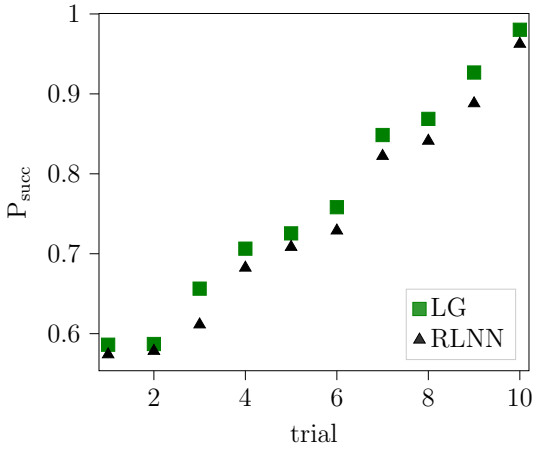

FIGURE 5.13: Success probability for $m = 2$, $n = 20$ where all candidate states are pure. Success probability for the RLNN is based on 1000 training iterations for each round.

Next, we consider the performance of multiple state discrimination where $m = 3$. For $n < 8$ we can compare directly to the collective SDP. For $n \geqslant 8$, computing P_{succ} via SDP techniques is infeasibly slow, so we instead look at the RLNN training curve shape to determine whether the neural network converges to a steady solution.

As a first test of RLNN performance for multiple state discrimination when $m = 3$, we consider the case of pure states with $n = 5$ and plot the success probability vs SDP as well as the training curves in Fig. 5.14.The RLNN success probability plateaus after approximately 100 training iterations, and the RLNN success probability comes close to the collective SDP in each case.

We then examine the training curves for general state discrimination when $m = 3$ as a function of n, with results shown in Fig. 5.16. Although stable plateaus are

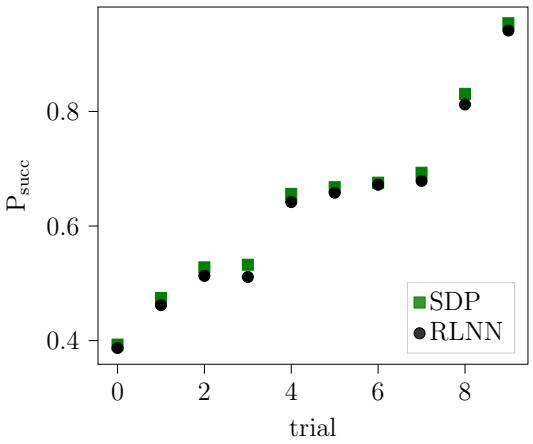

FIGURE 5.14: 10 trials with $m = 3$, $n = 5$. Success probability for RLNN after 300 training iterations compared to success probability of SDP

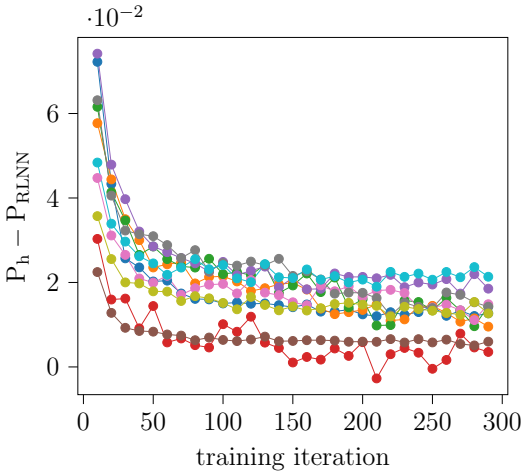

FIGURE 5.15: Difference between RLNN reward and SDP success probability as a function of training iteration. The RLNN success probability plateaus after 100 training iterations

reached for both $n = 10$ and $n = 20$, as n increases the RLNN spends more training iterations learning not to re-measure subsystems. In the case where $n = 50$, approximately 500 training iterations are spent learning not to re-measure subsystems, leading to a negative initial reward. This leads to the question of whether it is possible to extend the RLNN performance to a larger number of subsystems by predetermining a close-to-optimal ordering,

5.9 Robustness under noise

We demonstrate that the success probability is stable when the candidate states are subject to a small perturbation. Consider an over-rotation noise model where the noise is parametrised by rotation angle θ, s.t.

$$\tilde{\rho}_j^{(k)}(\theta) = U(\theta)\rho_j^{(k)}U^\dagger(\theta) \quad \forall \quad j,k$$

where $\tilde{\rho}_j^{(k)}(\theta)$ is the noisy state and we set the rotation matrix as $U(\theta) = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$. We prove that the error due to noise is negligible for sufficiently small enough per-

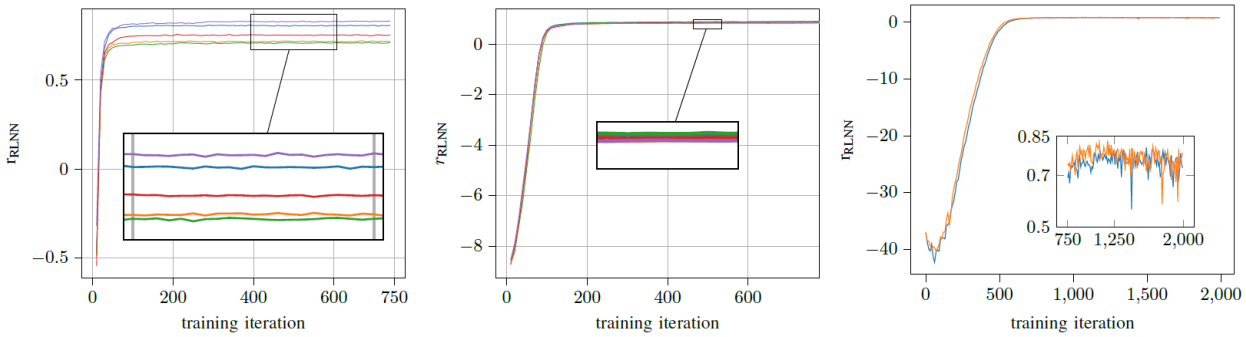

FIGURE 5.16: Training curves for five independent trials where $n = 10$ (right), $n = 20$ (centre), and one trial of $n = 50$ (left). As n increases, the training curve becomes less stable, and the number of iterations required for the reward to reach its maximal value increases. For $n = 50$, the shape of the training curve also changes to include an initial dip in reward before convergence to the ideal, and we observe less stability.

turbations, indicating that the RLNN adaptive method is still close-to-optimal when the candidate states are subjected to sufficiently small amounts of unitary noise.

Theorem 9. *Consider candidate set* $\{\rho_j\}|_{j=1}^m$ *with prior* $\mathbf{q}$. *Denote by* $P_{succ}\Big(\{\rho_j\}\Big)$ *the probability of success using the optimal locally adaptive method on the original state set. Likewise, let* $P_{succ}\Big(\{\tilde{\rho}_j(\theta)\}\Big)$ *be the success probability for the noisy state set. Then for all* θ,

$$|P_{succ}\Big(\{\rho_j\}\Big) - P_{succ}\Big(\{\tilde{\rho}_j(\theta)\}\Big)| \leqslant n|\theta|$$

where n *is the number of subsystems.*

Proof. See Appendix 3 for a complete proof. ∎

Finally, we generate five candidate state sets with $m = 3$, $d = 2$. For each candidate state set $\{\rho_j\}$, we train the neural network to find the optimal locally adaptive method. The original adaptive measurement scheme is then applied to the rotated state set $\{\tilde{\rho}_j(\theta)\}$, and we plot the gap in success probabilities $\text{diff}(\theta) \triangleq \text{P}_{\text{succ}}\Big(\{\rho_j\}\Big) - \text{P}_{\text{succ}}\Big(\{\tilde{\rho}_j(\theta)\}\Big)$.

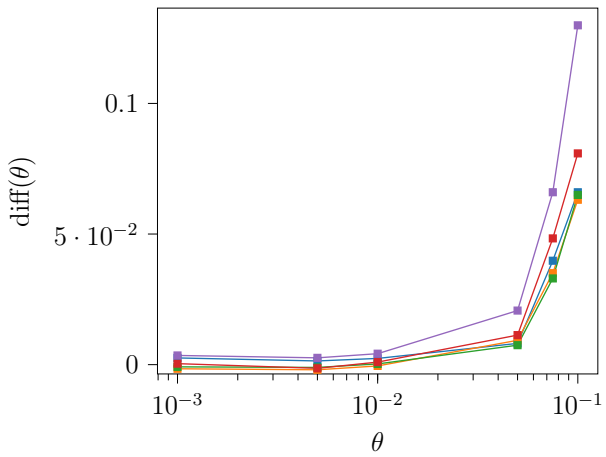

FIGURE 5.17: Gap in success probability as a function of rotation parameter θ for five trials where $m = n = 3$. We observe that the gap is negligible for sufficiently small θ.

5.10 Summary

We apply RLNN to calculate a near-optimal locally-adaptive measurement scheme for multiple state discrimination. We provide preliminary results for the neural network performance in cases where the locally-adaptive probability of success is known, and show that the network can achieve good performance when the total number of subsystems is 10 or fewer. This performance holds even when the candidate states are subjected to small perturbations. For cases where the exact locally optimal protocol is not known, we compare the RLNN performance to an SDP upper bound and find that for each trial the RLNN comes close to the upper bound. Additionally, we introduce a min-entropy based locally adaptive approach which reduces to the optimal local approach for binary pure states, and show that the RLNN meets or exceeds this approach in every trial.

Finally, we discuss pure state discrimination and note that the existence of a gap between adaptive locally greedy algorithms and collective measurements depends solely on the number of candidate states. Future work aims to extend the RLNN performance to a larger number of subsystems, as well as to characterize the maximal

gap between the optimal local and optimal collective success probability as a function of the number of candidate states (m) and subsystems (n).

6

Quantifying Uncertainty Through Games of Chance

Portions of this chapter are adapted from [BGG21]. This work is done in collaboration with Isabelle Jianing Geng and Dr. Gilad Gour. Dr. Gour oversaw the project, found results for conditional majorization, and developed the project idea. I found the remaining analytical results. Isabelle Jianing Geng and I finalized the manuscript.

6.1 Introduction

Entropy plays a central role in many scientific areas including statistical mechanics, thermodynamics, information theory, black hole physics, cosmology, chemistry, and even economics [Jak20, Bek72, Bek73]. Consequently, there are multiple approaches to understanding entropy. In thermodynamics, it can be understood as a measure of energy dispersal at a given temperature. In information theory, entropy can be viewed as a compression rate. Other properties related to entropy such as disorder, chaos, randomness of a system, and the arrow of time [Mac92], have also been studied extensively in literature. Despite the many different definitions of entropy, one

unifying theme is the idea of uncertainty.

The diverse roles of entropy would benefit from a more systematic and unifying approach, where entropy is defined rigorously and in a way that is independent of the physical context. In this work, we introduce games of chance as a means of characterizing the uncertainty of a physical system. Games of chance are ideal candidates for studying uncertainty, as the performance in such games depends solely on the certainty about the outcome of the game such that "more uncertain" systems will have a lower expected reward. Thus, for any physical object such as a system, measurement, or channel, we define its degree of uncertainty in terms of the probability of winning a game of chance. Since we construct a family of games, uncertainty cannot be quantified with one function that is induced by a single game, but rather is characterized with a partial order where system A is said to be "less uncertain" than system B if system A performs at least as well as system B for *all* games of chance. Equivalently, we say that A majorizes B.

We then apply our framework to the task of characterizing the entropy of a channel. The classical channel is a fundamental concept in classical information theory, as classical states (which can be viewed as a probability distribution), random relabelings, and stochastic evolutions can all be viewed as a certain type of classical channels. Likewise, quantum states, unitary evolutions, quantum measurements, state preparations and marginalization over any subsystem of a larger quantum system can be regarded as special cases of quantum channels [GMN$^+$15]. Over the past few years, several works have aimed to characterise and order channels based on their performance for certain tasks [GW18, Yua19, FFRS20, DKJR06, DHBS20, DBB17]. The seminal work by John Kelly [Kel56] in 1956 developed the Kelly criterion which yields the optimal protocol for a player to maximize the growth rate of their gam-

bling profit via N uses of a communication channel in the limit as $N \to \infty$.

The idea of using majorization to study entropy could be found in [Par11], while the notion of conditional majorization was first introduced in [FGG13], and extended in [GGH$^+$18]. However, channel majorization had not been considered in previous literature, nor had a unifying operational interpretation across different types of majorization. In our work, we construct gambling games that give rise to three types of partial orders: majorization, conditional majorization, and channel majorization.The first two partial orders characterize the degree of uncertainty and conditional uncertainty in (possibly composite) physical systems, while the last characterizes the uncertainty associated with a channel. Critically, we provide operational interpretations for each type of majorization and demonstrate that the definition of conditional majorization coincides with the definition provided in [GGH$^+$18]. In the following chapter, we define channel entropy based on the definition of channel majorization given in this chapter.

6.1.1 Notation

We denote the conditional probability distribution of random variable Y on random variable X as $\{p_{y|x}\}$ where $p_{y|x} \triangleq \Pr(X = x \big| Y = y)$. Likewise, the joint distribution for X and Y is denoted as $\{p_{yx}\}$ such that $p_{xy} \triangleq \Pr(X = x, Y = y)$. We denote the set of all $m \times n$ stochastic matrices as $\text{Stoch}(m, n)$, such that $P \in \text{Stoch}(m, n)$ if $P_{X,Y} = p_{y|x}$ (ie, P is a transition probability matrix).

We use calligraphic letters such as $\mathcal{P}$, $\mathcal{Q}$, $\mathcal{R}$ to represent classical channels, and text capital letters such as P, Q, R, S for the corresponding transition probability matrices. At times, we will use superscripts of the form $\mathcal{N}^{m\to n}$ to indicate that the channel $\mathcal{N}$ takes inputs from an alphabet of size m and produces outputs in an alphabet of size n.

6.2 Dice Games and Majorisation

Gambling games are games in which a player is provided with partial information and uses statistical inference to make their best guess in the face of incomplete information. We first consider a gambling game in which the host rolls a biased dice, and the player has to guess its outcome. Denote by $\mathbf{p} = (p_1, ..., p_n)^T$ the probability vector corresponding to the n possible outcomes, and denote by $\mathbf{p}^{\downarrow} = (p_1^{\downarrow}, ..., p_n^{\downarrow})^T$ the vector obtained from $\mathbf{p}$ by rearranging its components in non-increasing order. For simplicity, we will always assume that the components of any probability vectors $\mathbf{p}$ are arranged in non-increasing order so that $\mathbf{p} = \mathbf{p}^{\downarrow}$.

A *w-gambling game* occurs when the player is allowed to provide a set with w-numbers as guesses prior to rolling the dice. The player then wins if the outcome from the dice roll belongs to the set of guesses. For example, if $w = 2$, then the player will choose to provide numbers $\{1, 2\}$ (as these have the highest probability of occurring), and will win the game with probability $p_1^{\downarrow} + p_2^{\downarrow}$. In general, the probability of winning a w-game with dice $\mathbf{p}$ can be denoted as:

$$\text{Prob}_w(\mathbf{p}) = \|\mathbf{p}\|_{(w)} := \sum_{x=1}^{w} p_x^{\downarrow}$$

where $\|\cdot\|_{(w)}$ denotes the Ky-Fan norm.

Suppose that at the beginning of each game, the player is allowed to choose between two dice with corresponding probabilities $\mathbf{p}$ and $\mathbf{q}$. Clearly, the player will choose the dice which gives better odds of winning the game, and so will choose the $\mathbf{p}$-dice if $\|\mathbf{p}\|_{(w)} \geqslant \|\mathbf{q}\|_{(w)}$. In general, the player's choice will depend on the value of w- for example, if $\mathbf{p} = (\frac{1}{2}, \frac{1}{2}, 0)$ and $\mathbf{q} = (\frac{2}{3}, \frac{1}{6}, \frac{1}{6})$ then the player will choose $\mathbf{q}$ when $w = 1$ and $\mathbf{p}$ when $w = 2$.

If the parameter w is chosen randomly from the distribution $\{t_w\}_{w=1}^m$, then the

probability that the optimal play wins is given by

$$\mathrm{Prob}_{\mathbf{t}}(\mathbf{p}) = \sum_{k=1}^{m} t_k \|\mathbf{p}\|_{(k)} = \sum_{x=1}^{m} \sum_{k=x}^{m} t_k p_x \equiv \sum_{x=1}^{m} r_x p_x = \mathbf{r} \cdot \mathbf{p} \tag{6.1}$$

where the vector $\mathbf{r}$ with components $\{r_x\}$ is given by $\mathbf{r} \equiv U\mathbf{t}$ and U is the $m \times m$ invertible upper triangular matrix

$$U \equiv \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \tag{6.2}$$

Note that the vector $\mathbf{r}$ satisfies $\mathbf{r} = \mathbf{r}^{\downarrow}$ and that $\mathbf{r}$ has this property if and only if the vector $\mathbf{t} = U^{-1}\mathbf{r}$ has non-negative components. In general, we allow $t \equiv \sum_w t_w$ to be strictly smaller than one, in which case there is a non-zero probability that the player loses the game irrespective of the dice outcome. We likewise set $p_x^{\downarrow} \equiv 0$ if $x > n$.

Finally, we say that $\mathbf{p}$ majorizes $\mathbf{q}$ and write $\mathbf{q} \precsim \mathbf{p}$ if and only if

$$\mathrm{Prob}_{\mathbf{t}}(\mathbf{q}) \leqslant \mathrm{Prob}_{\mathbf{t}}(\mathbf{p}) \; . \tag{6.3}$$

for all (possibly incomplete) distributions $\{t_w\}$. This is equivalent to saying that $\mathbf{q} \precsim \mathbf{p}$ if the player will *always* choose the $\mathbf{p}$-dice over the $\mathbf{q}$-dice for any gambling game.

One can additionally consider games with resources, where a player may be provided with an additional $\mathbf{s}$-dice. The probability to win such a w-gambling game with resources is given by $\|\mathbf{p} \otimes \mathbf{s}\|_{(w)}$, such that a $\mathbf{p}$-dice has less uncertainty than a $\mathbf{q}$-dice if $\|\mathbf{p} \otimes \mathbf{s}\|_{(w)} \geqslant \|\mathbf{q} \otimes \mathbf{s}\|_{(w)}$ for all w and for all $\mathbf{s}$. However, this condition is equivalent to $\mathbf{q} \precsim \mathbf{p}$, so the class of w-games is sufficient to induce the pre-order generated by the larger class of games with resources.

6.3 Conditional Majorization: Games with a correlated source.

Here we consider a game in which the host rolls a die where each side is labeled with two outcomes, x and y. The host sends the value of x to the player, and the value y is kept hidden from the player. The player knows the distribution $\{p_{xy}\}$ from which x and y are sampled, and the player's goal is to guess the value of y. Our goal is to construct all possible gambling games that incorporate a correlated source, so we allow the player to choose a value z and then have the host select w from a conditional distribution $\{t_{wz}\}$ after receiving the value z from the player. We denote the player's choice of z with a function $z = f(x)$. In general, the player will choose z based on their knowledge of x, as well as the fixed distributions $\{p_{xy}\}$ and $\{t_{w|z}\}$.

In Fig. 6.1 we depict such a $\mathcal{T}$-gambling game.

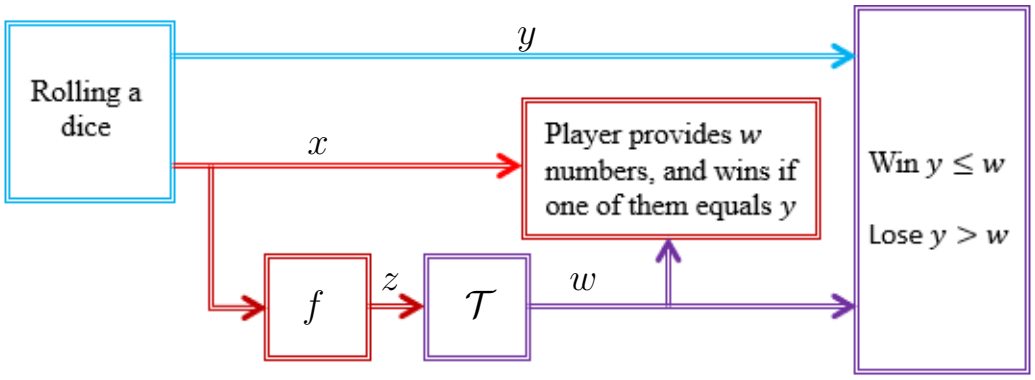

FIGURE 6.1: Classical gambling game with a correlated source. The player is provided with the value x. Based on this value, the player choose z (or the function f) give and send it to the host. The host then chooses the w game based on a (possibly incomplete) distribution matrix $\mathcal{T} = (t_{w|z})$. The player will then guess $\{1, ..., w\}$ and will win the game if $y \leqslant w$.

Let $P = (p_{xy})$ be the $m \times n$ probability matrix, and w.l.o.g. suppose $P = P^{\downarrow}$ such that

$$p_{x1} \geqslant p_{x2} \geqslant \cdots \geqslant p_{xn} \quad \forall\, x = 1, ..., m. \tag{6.4}$$

For a given x and z, the probability to win the game can be expressed as $\mathbf{r}_z \cdot \mathbf{p}_x$,

where $\{\mathbf{p}_x\}$ are the rows of P and $\mathbf{r}_z \equiv U\mathbf{t}_z$, where $\{\mathbf{t}_z\}$ are the columns of the $\ell \times q$ matrix $\mathcal{T} = (t_{w|z})$. As before, we only require that $\sum_{w=1}^{\ell} t_{w|z} \leqslant 1$ for all $z = 1, ..., q$.

Therefore, the optimal probability to win a $\mathcal{T}$-game is given by

$$\mathrm{Prob}_{\mathcal{T}}(P) = \sum_{x=1}^{m} \max_z \mathbf{r}_z \cdot \mathbf{p}_x \ . \tag{6.5}$$

We are now ready to compare between two dice, a P-dice and a Q-dice, and call this comparison conditional majorization.

Definition 10. *Let $P = (p_{xy})$ be an $m \times n$ probability matrix, and $Q = (q_{x'y'})$ be an $m' \times n'$ probability matrix. We say that P conditionally majorizes Q and write*

$$Q \precsim_c P \quad \textit{if and only if} \quad \mathrm{Prob}_{\mathcal{T}}(Q) \leqslant \mathrm{Prob}_{\mathcal{T}}(P) \tag{6.6}$$

for all (column) sub-stochastic matrices $\mathcal{T}$.

We note that the term 'conditional majorization' was first introduced in [GGH$^+$18], however it was not defined in an operational way. We now demonstrate conditional majorization as defined by games of chance is equivalent to the original definition in [GGH$^+$18].

Theorem 11. *Let P and Q be two $m \times n$ column stochastic matrices. Then,*

$$Q \precsim_c P \quad \Longleftrightarrow \quad Q = \sum_z S_z P V_z \tag{6.7}$$

where each S_z is a sub-stochastic matrix such that $\sum_z S_z$ is a column stochastic matrix (i.e. a classical channel), and each V_z is a permutation matrix.

This theorem provides an operational characterization for conditional majorization.[1] It states that conditional majorization is equivalent to a relation induced by a conditional random relabeling map; see Fig 6.3.

[1] For convenience we assume that the two matrices have the same dimensions, as if this is not the case one can add zero rows and columns to make them the same dimension.

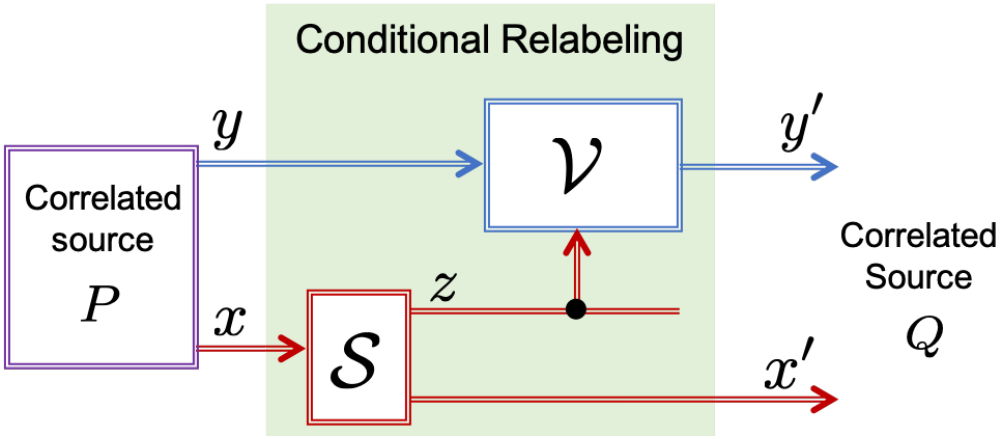

FIGURE 6.2: The action of conditional random relabeling map on a correlated source $P = (p_{xy})$ yields the correlated source $Q = (q_{x'y'}) = \sum_z S_z P V_z$.

Proof. The proof follows from the following two lemmas.

Lemma 12. *[GGH+18] The matrices P and Q are related as in the RHS of (16) if and only if there exists a column stochastic matrix S such that*

$$QU \leqslant SPU \ , \tag{6.8}$$

where the inequality is entry-wise and U is the upper triangular matrix

Lemma 13. *[GGH+18] There exists a column stochastic matrix S that satisfies (6.8) if and only if for any set of m vectors $\mathbf{r}_1, ..., \mathbf{r}_m \in \mathbb{R}^n_+$ whose components are arranged in non-decreasing order,*

$$\sum_{x=1}^{m} \max_z \mathbf{p}_x \cdot \mathbf{r}_z \geqslant \sum_{x=1}^{m} \mathbf{q}_x \cdot \mathbf{r}_x \ . \tag{6.9}$$

We now prove the theorem. If $Q \precsim_c P$ then from (6.5) and Definition 10 we get

$$\sum_{x=1}^{m} \max_z \mathbf{p}_x \cdot \mathbf{r}_z \geqslant \sum_{x=1}^{m} \max_z \mathbf{q}_x \cdot \mathbf{r}_z \geqslant \sum_{x=1}^{m} \mathbf{q}_x \cdot \mathbf{r}_x \ . \tag{6.10}$$

Hence, from the above two lemmas it follows that P and Q are related as in the RHS of (16). ∎

6.4 Channel Majorization: Games with a classical channel.

We now identify the class of gambling games corresponding to the uncertainty of a channel. Roughly speaking, our aim is to identify games in which the player has a lower probability of winning the game with a noisier channel. We will consider a classical channel $\mathcal{M}^{\mathrm{m}\to\mathrm{n}}$ which takes inputs from the alphabet $\mathcal{X} = \{1, ..., m\}$ and gives outputs in the alphabet $\mathcal{Y} = \{1, ..., n\}$. Each channel has a corresponding transition matrix $P = (p_{y|x})$. The goal of the game is for the player to correctly guess the value y at the output of the channel.

In the most general settings, the host does not provide the player the full information about w at the early stage of the game. Instead, the player receives a number z that is sampled from a joint distribution $\{t_{wz}\}$ with $w = 1, ..., m$ and $z = 1, ..., \ell$. Denote by $\{\mathbf{t}_z\}$ the columns of T. Then $\Pr(Z = z) = |\mathbf{t}_z|$, and one can write the conditional probability for w given z as $t_{w|z} = \frac{t_{wz}}{|\mathbf{t}_z|}$ for all w, z.

The player knows the $m \times \ell$ joint probability matrix $T = (t_{wz})$. While $\mathcal{T}$ in the previous section represented a classical channel which takes z as an input and outputs w, here T represents a correlated bipartite source which generates w and z. Based on this partial information about the value of w, the player will choose the optimal value of x to send through the channel. Finally, the host draws the value of w and the player wins if $y \leqslant w$.

Such a T-gambling game with channel $\mathcal{M}$ is then depicted in Fig. 6.1.

Let $P = (p_{y|x}) = P^{\downarrow}$ be the ordered transition matrix of the channel $\mathcal{M}$, in which the columns of P are arranged in non-increasing order. For a given choice of x and z the probability that the player wins the game is given by

$$\sum_{w=1}^{m} t_{w|z} \sum_{y=1}^{w} p_{y|x} = \sum_{y=1}^{m} \sum_{w=y}^{m} t_{w|z} p_{y|x} \equiv \frac{\mathbf{r}_z \cdot \mathbf{p}_x}{|\mathbf{t}_z|} \tag{6.11}$$

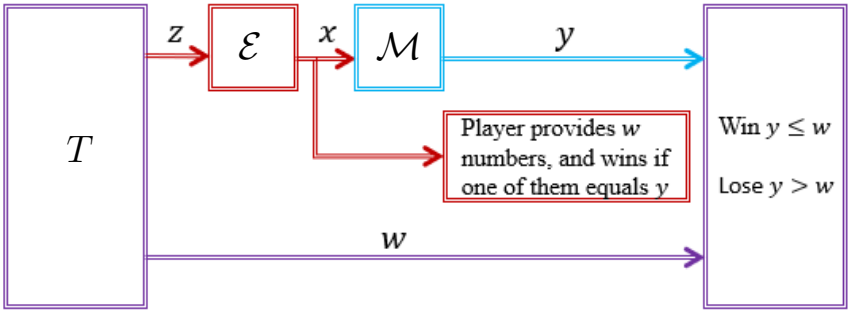

FIGURE 6.3: Classical gambling game with a channel. The host provides the player with a value z that is drawn from a probability distribution $T = (t_{wz})$. The player chooses an input to the channel $x = f(z)$ that may depends of the value of z. The host then selects w, and the player wins if $y \leqslant w$ since they will always provide the first w numbers with the highest corresponding probabilities.

where $\mathbf{r}_z \equiv U\mathbf{t}_z$ and $\{\mathbf{p}_x\}$ are the columns of the transition matrix P; and in particular, each $\mathbf{p}_x$ is a probability vector. For each value of z the player will choose x (i.e. $f(z)$) such that $\mathbf{r}_z \cdot \mathbf{p}_x = \max_{x'} \mathbf{r}_z \cdot \mathbf{p}_{x'}$. We therefore conclude that the optimal probability to win a T-gambling game with a classical channel $\mathcal{M} = P$ is given by

$$\text{Prob}_T(\mathcal{M}) = \sum_{z=1}^{\ell} \max_x \ \mathbf{r}_z \cdot \mathbf{p}_x \tag{6.12}$$

Note that the above quantity is the dual of (6.5) in the sense that the role of x and z are flipped. Unlike (6.5), here $\mathbf{p}_x$ is a probability vector for each x.

Definition 14. *Let $\mathcal{M}$ and $\mathcal{N}$ be two classical channels with corresponding transition matrices $P = (p_{y|x})$ and $Q = (q_{y'|x'})$. We say that the channel $\mathcal{N}$ majorizes $\mathcal{M}$ and write*

$$\mathcal{M} \precsim \mathcal{N} \quad \textit{if and only if} \quad \textit{Prob}_T(\mathcal{M}) \leqslant \textit{Prob}_T(\mathcal{N}) \tag{6.13}$$

for all (column) sub-stochastic matrices T.

We now provide the following characterization of channel majorization, which demonstrates that $\mathcal{M} \precsim \mathcal{N}$ if and only $\mathcal{M}$ can be simulated using $\mathcal{N}$, an arbitrary

pre-processing channel, and random isometry post-processing channels which may in general be correlated with the input to the pre-processing channel.

Theorem 15. *Let $\mathcal{M}^{m\to n}$ and $\mathcal{N}^{m'\to n'}$ be two classical channels with probability transition matrices Q and P respectively. Then, $\mathcal{M} \precsim \mathcal{N}$ if and only if there exists a preprocessing channel $\mathcal{S}^{m\to m'\times \ell}$ with transition probability matrix $S = \{s_{x',w'|x}\}$, and postprocessing channels $\{\mathcal{V}_{w'}^{n'\to n}\}_{w'}$ whose corresponding transition matrices $\{V_{w'}\}_{w'}$ are permutation matrices, such that*

$$Q = \sum_{w'} V_{w'} P S_{w'} \tag{6.14}$$

where $S_{w'} = \{s_{x',w'|x}\}|_{x',x}$ for all w' (see Fig. 7.1).

Proof. Note that w.l.o.g. we can assume $m' = m$, so that P is an $m \times n'$ matrix and Q is an $m \times n$ matrix. We will need the following two lemmas to prove the theorem statement:

Lemma 16. *The matrices P and Q are related as in* (6.14) *if and only if there exists a column stochastic matrix $S = \sum_{w'} S_{w'}$ such that*

$$U^T Q \leqslant U^T P S \ , \tag{6.15}$$

where the inequality is entry-wise and U is the upper triangular matrix as defined in (6.2)*.*

To prove the lemma, denote by $\{\mathbf{p}_{x'}\}_{x'=1}^{n'}$ and $\{\mathbf{q}_x\}_{x=1}^{n}$ the n' and n columns of the matrices P and Q, and by $s_{x'w'|x}$ and $s_{x'|x}$ the components of the matrices $S_{w'}$ and $S \equiv \sum_{w'=1}^{\ell} S_{w'}$. Then, the relation (6.14) can be expressed as:

$$\mathbf{q}_x = \sum_{x'=1}^{n'} s_{x'|x} D_{x'x} \mathbf{p}_{x'} \quad \text{where} \quad D_{x'x} \equiv \sum_{w'=1}^{\ell} \frac{s_{x'w'|x}}{s_{x'|x}} V_{w'} \ . \tag{6.16}$$

Since for each x' and x the matrix $D_{x'x}$ is an $m \times m$ doubly-stochastic matrix, we have that for each x and x', $D_{x'x}\mathbf{p}_{x'} \prec \mathbf{p}_{x'}$. Thus, if the matrices Q and P are related as in (6.14) then there must exists a column stochastic matrix $S = (s_{x'|x})$ such that

$$\mathbf{q}_x \prec \sum_{x'=1}^{n'} s_{x'|x}\mathbf{p}_{x'} \quad \forall x = 1, ..., n. \tag{6.17}$$

The above relation is equivalent to (6.15) since for each x and x', $\mathbf{q}_x = \mathbf{q}_x^{\downarrow}$ and $\mathbf{p}_{x'} = \mathbf{p}_{x'}^{\downarrow}$.

Conversely, suppose (6.15) holds. Therefore, also (6.17) holds, so that there exists an $m \times m$ doubly-stochastic matrix D such that

$$\mathbf{q}_x = \sum_{x'=1}^{n} s_{x'|x} D\mathbf{p}_{x'} \quad \forall x = 1, ..., n. \tag{6.18}$$

Expressing $D = \sum_{w'=1}^{\ell} c_{w'} V_{w'}$ as a convex combination of permutation matrices we obtain

$$\mathbf{q}_x = \sum_{x'=1}^{n'} s_{x'|x} \sum_{w'=1}^{\ell} c_{w'} V_{w'} \mathbf{p}_{x'} = \sum_{x',w'} V_{w'} \mathbf{p}_{x'} s_{x'|x} c_{w'} \quad \forall x = 1, ..., n. \tag{6.19}$$

Finally, denote $s_{x'w'|x} \equiv c_{w'} s_{x'|x}$ and note that with this notation the above equation is equivalent to (6.14), where for each w' the the components of the matrix $S_{w'}$ are $s_{x'w'|x}$. This completes the proof of the lemma.

The question of whether there exists such a column stochastic matrix S that satisfies (6.15) can be solved with linear programming.

Lemma 17. *There exists a $n' \times n$ column stochastic matrix S that satisfies* (6.15) *if and only if for any set of n vectors $\mathbf{r}_1, ..., \mathbf{r}_n \in \mathbb{R}_+^m$ whose components are arranged in non-decreasing order,*

$$\sum_{x=1}^{n} \max_{x'} \ \mathbf{r}_x \cdot \mathbf{p}_{x'} \geqslant \sum_{x=1}^{n} \mathbf{r}_x \cdot \mathbf{q}_x \ . \tag{6.20}$$

To prove the lemma, denote by $\{\mathbf{s}_x\}_{x=1}^n$ the columns of S, so that (6.15) can be expressed as

$$U^T P \mathbf{s}_x \geqslant U^T \mathbf{q}_x \qquad \forall\, x = 1, ..., n\ . \tag{6.21}$$

and the condition that S is column stochastic is equivalent to $\mathbf{e} \cdot \mathbf{s}_x = 1$ for each $x = 1, ..., n$, where $\mathbf{e} \equiv (1, ..., 1)^T \in \mathbb{R}^n$. Note however that it is sufficient to require that $\mathbf{e} \cdot \mathbf{s}_x \leqslant 1$ since by adding to $\mathbf{s}_x$ a non-negative vector, the relation (6.21) is preserved. Denote further by $\mathbf{e}' \equiv (1, ..., 1)^T \in \mathbb{R}^{n'}$,

$$A \equiv \begin{pmatrix} U^T P & 0 & \cdots & 0 \\ 0 & U^T P & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & U^T P \\ -\mathbf{e}^T & 0 & \cdots & 0 \\ 0 & -\mathbf{e}^T & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -\mathbf{e}^T \end{pmatrix} \in \mathbb{R}^{(m+1)n \times n'n} \tag{6.22}$$

$$\mathbf{b} \equiv \begin{pmatrix} U^T \mathbf{q}_1 \\ \vdots \\ U^T \mathbf{q}_n \\ -\mathbf{e}' \end{pmatrix} \in \mathbb{R}^{(m+1)n}, \mathbf{s} \equiv \begin{pmatrix} \mathbf{s}_1 \\ \vdots \\ \mathbf{s}_n \end{pmatrix} \in \mathbb{R}_+^{n'n} \tag{6.23}$$

Using this notation, (6.21) is equivalent to $A\mathbf{s} \geqslant \mathbf{b}$. The question of whether there exists a $\mathbf{s} \geqslant 0$ that satisfies $A\mathbf{s} \geqslant \mathbf{b}$ is a feasibility problem in linear programming. From Farkas lemma, such a row vector $\mathbf{s} \in \mathbb{R}_+^{n'n}$ exists if and only if for any vector $\mathbf{t} \in \mathbb{R}_+^{(m+1)n}$ that satisfies $\mathbf{t}^T A \leqslant 0$ we also have $\mathbf{b} \cdot \mathbf{t} \leqslant 0$. Denote $\mathbf{t} = (\mathbf{t}_1, ..., \mathbf{t}_n, \mathbf{v})^T$, where for each $x' = 1, ..., n$, $\mathbf{t}_x \in \mathbb{R}_+^m$ and $\mathbf{v} = (v_1, ..., v_n)^T \in \mathbb{R}^n$. With this notation, the dual problem has the form

$$v_x \mathbf{e} \geqslant \mathbf{t}_x^T U^T P \quad \forall x = 1, ..., n \quad \Rightarrow \quad \mathbf{e}' \cdot \mathbf{v} \geqslant \sum_{x=1}^{n} \mathbf{t}_x^T U^T \mathbf{q}_x \tag{6.24}$$

Note that the condition $v_x \mathbf{e} \geqslant \mathbf{t}_x^T U^T P$ implies that $v_x \geqslant \mathbf{t}_x^T U^T \mathbf{p}_{x'}$ for all $x' = 1, ..., n'$ and $x = 1, ..., n$. Therefore, taking the optimal value $v_x = \max_{x'} \mathbf{t}_x^T U^T \mathbf{p}_{x'}$ we get

that (6.24) holds if and only if for any n vectors $\mathbf{t}_1, ..., \mathbf{t}_n \in \mathbb{R}_+^m$

$$\sum_{x=1}^{n'} \max_{x'} \mathbf{t}_x^T U^T \mathbf{p}_{x'} \geqslant \sum_{x=1}^{n'} \mathbf{t}_x^T U^T \mathbf{q}_x \ . \tag{6.25}$$

Denoting by $\mathbf{r}_x \equiv U\mathbf{t}_x$ we obtain the the above equation is equivalent to (6.20). This completes the proof of the lemma.

We are now ready to prove Theorem 15. We therefore assume now that $\mathcal{M} \precsim \mathcal{N}$. Then, by definition we have for all ℓ and any $\mathbf{r}_1, ..., \mathbf{r}_\ell \in \mathbb{R}_+^m$ whose components are arranged in a non-decreasing order

$$\sum_{w'=1}^{\ell} \max_{x'} \ \mathbf{r}_{w'} \cdot \mathbf{p}_{x'} \geqslant \sum_{w'=1}^{\ell} \max_{x} \ \mathbf{r}_{w'} \cdot \mathbf{q}_x. \tag{6.26}$$

Taking $\ell = n$ we get that

$$\sum_{x=1}^{n} \max_{x'} \ \mathbf{r}_x \cdot \mathbf{p}_{x'} \geqslant \sum_{w'=1}^{n} \max_{x} \ \mathbf{r}_{w'} \cdot \mathbf{q}_x \geqslant \sum_{x=1}^{n} \ \mathbf{r}_x \cdot \mathbf{q}_x \ . \tag{6.27}$$

Therefore, from the two lemmas above it follows that P and Q are related as in (6.14). This completes the "only if" direction in the theorem statement.

The "if" direction of the theorem follows quickly:

$$\begin{aligned}
&\mathrm{Prob}_T(\mathcal{M}) = \mathrm{Prob}_T(\mathcal{N}) \\
&= \sum_z \max_x \left(\sum_{w=1}^{m} t_{w|z} \Big|\Big| \sum_{w'} V_{w'} P S_{w'} \hat{e}_x \Big|\Big|_{(w)} \right) \\
&\leqslant \sum_z \max_x \left(\sum_{w=1}^{m} t_{w|z} \max_{w'} \Big|\Big| V_{w'} P \hat{e}_x \Big|\Big|_{(w)} \right) \\
&= \sum_z \max_x \left(\sum_{w=1}^{m} t_{w|z} \Big|\Big| P \hat{e}_x \Big|\Big|_{(w)} \right) \\
&= \mathrm{Prob}_T(\mathcal{N})
\end{aligned}$$

where $\hat{e}_x$ is a vector with a one at position x and zeros else. ∎

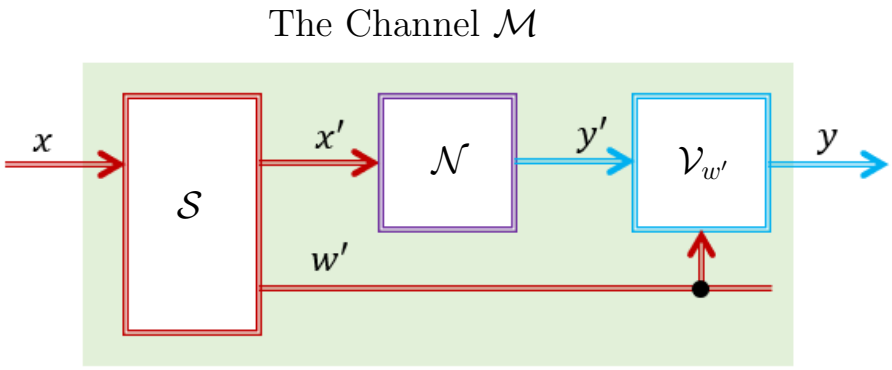

FIGURE 6.4: The simulation of $\mathcal{M}$ with $\mathcal{N}$ in the case that $\mathcal{M} \lesssim \mathcal{N}$.

We now provide a lemma showing that it is sufficient to consider only games with a fixed value of z.

Lemma 18. *The pre-order induced by the set of classical gambling games is unchanged if we restrict to the set of gambling games such that* $\mathcal{T} = \mathbf{p}$ *where* $\mathbf{p}$ *is a vector probability distribution. Equivalently,* $\mathcal{M} \lesssim \mathcal{N}$ *if and only if* $Prob_{\mathbf{p}}(\mathcal{M}) \leqslant Prob_{\mathbf{p}}(\mathcal{N})$ *for all* $\mathbf{p}$.

Proof. Referring to the definition of $\text{Prob}_T(\mathcal{N})$ from the previous section,

$$\begin{aligned}
\text{Prob}_{\mathcal{T}}(\mathcal{N}) &= \sum_{z=1}^{\ell} \max_x \Big(\sum_{y=1}^{m} \sum_{w=y}^{m} t_{wz} p_{y|x} \Big) \\
&= \sum_{z=1}^{\ell} t_z \max_x \Big(\sum_{y=1}^{m} \sum_{w=y}^{m} t_{w|z} p_{y|x} \Big) \\
&= \sum_{z=1}^{\ell} t_z \text{Prob}_{\{\mathbf{t}_z\}}(\mathcal{N})
\end{aligned}$$

where $t_{kz} = t_z t_{w|z}$ such that $\{\mathbf{t}_z\} = \{t_{1|z}, ..., t_{m|z}\}$. The statement then follows. ∎

6.5 Summary

In this chapter, we used games of chance as a method for characterizing the uncertainty of classical states and channels. This gives rise to majorization, conditional majorization, and channel majorization. We prove that the resulting conditional majorization is equivalent to [GGH$^+$18] and thus provide the first operational interpretation of conditional majorization.

In the following chapter, we use the channel majorization found in this chapter to define the entropy of a channel (ie a function is a channel entropy if it is both monotonic under channel majorisation and satisfies additivity under tensor products) and we then find the unique asymptotically continuous entropy. Thus, games of chance are a step towards addressing the open question how to characterise channel entropy.

7

Entropy of a Channel

Portions of this chapter are adapted from [BGG21]. This work is done in collaboration with Isabelle Jianing Geng and Dr. Gilad Gour. Dr. Gour oversaw the project, found results for conditional majorization, and developed the project idea. I found the remaining analytical results, and wrote the final manuscript. Isabelle Jianing Geng made substantial editorial contributions to the final manuscript and verified all analytical results.

7.1 Introduction

The previous chapter provided an ordering of channels based on uncertainty as quantified through games of chance. Previous works have investigated necessary conditions for entropy functions [Abe00, IS14, Tem16]. Based on these results as well as the channel ordering induced by games of chance, we define the set of allowed channel entropy functions. We then find that the minimum output entropy [Sho04] is the unique asymptotically continuous entropy of a channel. For a classical channel, the

minimum output entropy is defined as

$$H(\mathcal{N}) \triangleq \min_{x} \Big(H_S(\mathcal{N}(x)) \Big)$$

where H_S is the Shannon entropy.

Finally, this chapter connects games of chance to quantum dynamical resource theory and discusses additional potential physical applications of channel entropy. More specifically, we prove that the expected reward functions from games of chance provide a first operational interpretation of monotones in quantum dynamical resource theory [GS20, CG19, CFS16].

7.2 Channel Entropy Definition and Properties

Equipped with the pre-order introduced in the previous chapter, we now provide an operational definition for the family of entropy functions.

Definition 19. *(cf. [GW18, Gou19]) Let $L(m \to n)$ be the set of all classical channels which take inputs from $\{1, ..., m\}$ and yield outputs in $\{1, ..., n\}$. A non-zero function*

$$H : \bigcup_{m,n \,\in\, \mathbb{Z}^+} L(m \to n) \to \mathbb{R}$$

is a channel entropy if it satisfies the following two conditions:

1. *It is monotonic under the channel majorization; i.e.*

$$H(\mathcal{M}) \geqslant H(\mathcal{N}) \quad \textit{if} \quad \mathcal{M} \precsim \mathcal{N}$$

2. *It is additive under tensor products; i.e.*

$$H(\mathcal{N} \otimes \mathcal{M}) = H(\mathcal{N}) + H(\mathcal{N}) \tag{7.1}$$

for all channels $\mathcal{N}$ and $\mathcal{M}$.

The above definition is equivalent to the definition of entropy for classical channels previously outlined in [GW18, Gou19]. In these works, entropy was defined based on additivity and monotonicity under uniformity-preserving superchannel, or channels which map the completely randomizing channel $\mathcal{R}$ to itself. That is, if $\mathcal{R}^{m\to n}(\rho) = \frac{\mathbb{I}}{n}$ for all ρ, then $\Theta^{(n,m)\to(n',m')}$ is a completely randomizing superchannel if and only if:

$$\Theta^{(m,n)\to(m',n')}[\mathcal{R}^{m\to n}] = \mathcal{R}^{m'\to n'}$$

The following lemma demonstrates the equivalence of previous entropy definitions with our operational definition:

Lemma 20. *For any classical channels $\mathcal{M}^{m\to n}$ and $\mathcal{N}^{m'\to n'}$, then*

$$\mathcal{M} \preceq \mathcal{N} \iff \mathcal{M} = \Theta^{(m',n')\to(m,n)}(\mathcal{N})$$

where $\Theta^{(m',n')\to(m,n)}$ is a uniformity-preserving superchannel taking an input channel in $L(m' \to n')$ and outputting a channel in $L(m \to n)$.

Proof. In Theorem 15, we proved that $\mathcal{M} \preceq \mathcal{N}$ if and only if there exists a superchannel Θ satisfying $\Theta[\mathcal{N}] = \mathcal{M}$ and where Θ is of the form

$$\Theta^{(m',n')\to(m,n)}[\mathcal{P}] = \sum_z \mathcal{V}_z \circ \mathcal{N} \circ \mathcal{S}_z \quad \forall \mathcal{P} \in \mathrm{L}(m' \to n')$$

where $\mathcal{S}_z$ are sub-stochastic pre-processing processes (given bipartite pre-processing channel $\mathcal{S}$, then $\mathcal{S}_z = \left\{s_{x',z|x}\right\}|_{x',x}$ represents the action of S on the first subsystem when the output of S on the second subsystem is fixed to be z), and $\mathcal{V}$ implements a conditioned post-processing isometry $\mathcal{V}_z$ on the first subsystem (see Figure 6.4). This is equivalent to requiring Θ to be a classical, random unitary superchannel.

Note that this implies for any $\mathcal{P} \in \mathrm{L}(m' \to n')$, then:

$$\Theta^{(m',n')\to(m,n)}[\mathcal{P}](x) = \sum_z p(z|x)\mathcal{V}_z \circ \mathcal{P} \circ \mathcal{S'}_z(x) \quad \forall x \in [1, ..., m]$$

where $S'_z = \left\{s_{x'|x,z}\right\}|_{x,x'}$ is a pre-processing channel.

In the remainder of this proof, we demonstrate that the set of random unitary classical superchannels are equivalent to the set of uniformity-preserving classical superchannels. First, we show that classical random unitary channels are uniformity preserving:

$$\begin{aligned}
\Theta[\mathcal{R}^{m'\to n'}](x) &= \sum_z \mathcal{V}_z\Big(\mathcal{R}(\mathcal{S'}_z(x))\Big) \quad \forall x \\
&= \sum_z p(z|x)\mathcal{V}_z\Big(\frac{\mathbb{I}_n}{n}\Big) \quad \forall x \\
&= \sum_z p(z|x)\frac{\mathbb{I}_n}{n} \quad \forall x \\
&= \frac{\mathbb{I}_n}{n} \quad \forall x
\end{aligned}$$

Next, we show that any classical uniformity preserving superchannel Θ can be written in the above form. First, we note that the most general classical superchannel can be written as

$$\Theta[\mathcal{N}](|x\rangle\langle x|) = \sum_z p(z|x)\mathcal{P}\Big(\mathcal{N}\Big(\mathcal{S}_z(|x\rangle\langle x|)\Big) \otimes |z\rangle\langle z|\Big) \quad \forall\ x$$

where $\mathcal{P}$ is an arbitrary post-processing channel and $\mathcal{S}_z$ is the effect of $\mathcal{S}$ on subsystem 1 when the output of $\mathcal{S}$ on subsystem 2 is z. (See Figure 7.1). We use Dirac notation to denote $x \to |x\rangle\langle x|$

Upon imposing the restraint that Θ be uniformity-preserving, we have

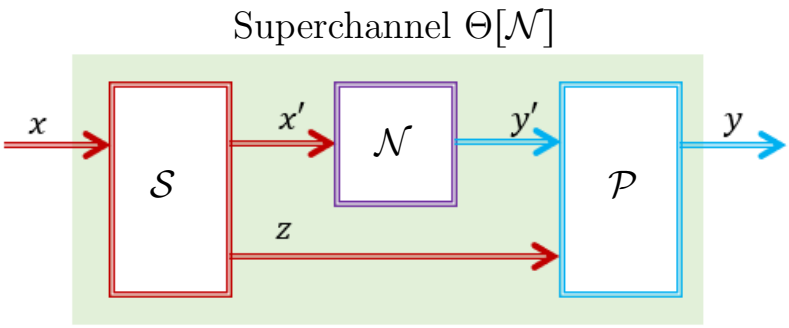

FIGURE 7.1: General classical superchannel.

$$\begin{aligned}
\frac{\mathbb{I}}{n} &= \Theta[\mathcal{R}]\big(\,|x\rangle\langle x|\,\big) \quad \forall x \\
&= \sum_z p(z|x)\mathcal{P}\Big(\mathcal{R}\Big(\mathcal{S}_z(|x\rangle\langle x|)\Big)\otimes|z\rangle\langle z|\,\Big) \quad \forall\ x \\
&= \sum_z p(z|x)\mathcal{P}\Big(\frac{\mathbb{I}}{n}\otimes|z\rangle\langle z|\,\Big) \quad \forall x \\
&= \sum_z p(z|x)\mathcal{P}_z\Big(\frac{\mathbb{I}}{n}\Big) \quad \forall x
\end{aligned}$$

where $\mathcal{P}_z(\rho) \triangleq \mathcal{P}(\rho\otimes|z\rangle\langle z|)$. Finally, denote by P^x the transition probability matrix corresponding to the channel $\sum_z p(z|x)\mathcal{P}_z$. Clearly, $\sum_z p(z|x)\mathcal{P}_z$ must be unital for all x s.t.

$$P^x\mathbf{u}_B = \mathbf{u}_B$$

where $\mathbf{u}_B$ is the uniform distribution for system B. It follows from the above that P^x is doubly stochastic for all x, and therefore (from Birkhoff's lemma), can be written as a convex combination of permutation matrices.

Then there exists a set of permutation matrices $\{V_{w,x}\}$ corresponding to isometry channels $\{\mathcal{V}_{w,x}\}$ and probability distribution $\sum_w q_{w|x} = 1$ for all x s.t.

$$\sum_z p(z|x)\mathcal{P}_z \triangleq \sum_w q_{w|x}\mathcal{V}_{w,x} \quad \forall x$$

Finally, we can then rewrite our superchannel Θ as:

$$\Theta[\mathcal{N}](|x\rangle\langle x|) = \sum_z p(z|x)\mathcal{P}_z\Big(\mathcal{N}\Big(\mathcal{S}_z(|x\rangle\langle x|)\Big)\Big) \quad \forall\ x$$

$$= \sum_t q_t \mathcal{V}_t\Big(\mathcal{N}\Big(\tilde{\mathcal{S}}_z(|x\rangle\langle x|)\Big)\Big)$$

where $t = (w, z)$ and where $\tilde{S}_z$ represents the action of $\tilde{S}$ on the first subsystem when the second output is z. Then $\tilde{S}$ is related to the original pre-processing channel as shown in Fig. 7.2.

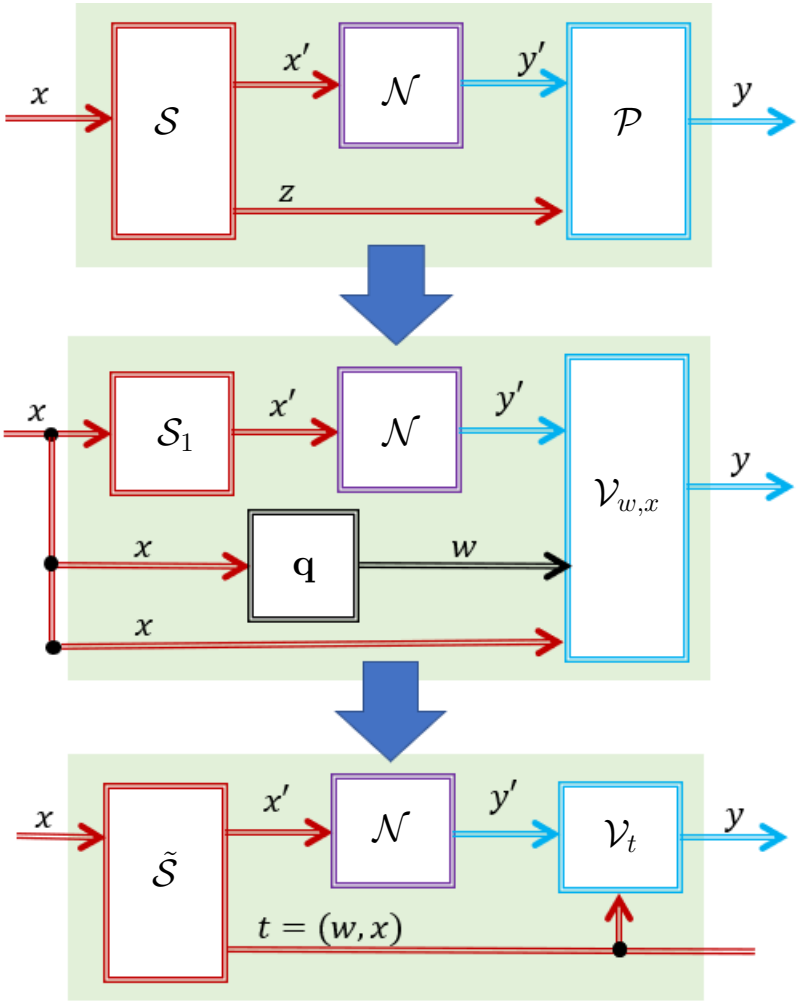

FIGURE 7.2: Restructuring a general uniformity preserving superchannel.

7.3 Uniqueness of Entropy Function

We now prove that there is only one channel entropy function that reduces to the Shannon entropy on classical states. An entropy function H reduces to to the Shannon entropy on states if and only if for *all* state preparation channels $\mathcal{N}_{\mathbf{p}}$ which take a trivial input and prepare a fixed output distribution $\mathbf{p}$, then:

$$H(\mathcal{N}_{\mathbf{p}}) = H_S(\mathbf{p})$$

where H_S is the Shannon entropy. Given that states are special cases of channels, we expect any reasonable channel entropy function to reduce to the Shannon entropy on states.

Theorem 21. *Let H be a classical entropy that reduces to the Shannon entropy, H_S, on classical states. Then for all $\mathcal{N} \in \bigcup_{m,n\in\mathbb{Z}^+} L(m \to n)$,*

$$H(\mathcal{N}) = \min_x H_S\Big(\mathcal{N}(x)\Big)$$

and H is asymptotically continuous.

Proof. We first show that there is a unique function H which reduces to the Shannon entropy on classical states. To this aim, we define the minimal extension $\underline{H}$ and maximal extension $\overline{H}$ of H as:

$$\underline{H}(\mathcal{N}) \triangleq \log|B| - \lim_{k\to\infty}\left(\frac{1}{k}\inf_{\mathbf{p},\mathbf{q}}\Big(D(\mathbf{p}\big\|\mathbf{q})\Big)\right)$$

$$\overline{H}(\mathcal{N}) \triangleq \log|B| - \max_x\Big(D\Big(\mathcal{N}(x)\Big\|\mathcal{R}(x)\Big)\Big)$$

where D is the Kullback-Leibler divergence and the infimum in the second line is over all $\mathbf{p}$, $\mathbf{q}$ satisfying $(\mathbf{p},\mathbf{q}) \prec \big(\mathcal{N}^{\otimes k}(\vec{x}), \mathcal{R}^{\otimes k}(\vec{x})\big)$ for all sequences $\vec{x}$. Since H reduces

to the Shannon entropy on classical states, then $\log|B| - H(\mathcal{N})$ reduces to $D(\mathcal{N}\big|\big|\mathcal{R})$. From [Gou21], it then follows that

$$\underline{H}(\mathcal{N}) \leqslant H(\mathcal{N}) \leqslant \overline{H}(\mathcal{N})$$

It was additionally shown in [Gou21] that

$$\max_x \Big(D\Big(\mathcal{N}(x)\big|\big|\mathcal{R}(x)\Big)\Big) = \lim_{k\to\infty}\left(\frac{1}{k}\inf_{\mathbf{p},\mathbf{q}}\Big(D(\mathbf{p}||\mathbf{q})\Big)\right)$$

where the infimum is over all $\mathbf{p}$, $\mathbf{q}$ satisfying $(\mathbf{p},\mathbf{q}) \prec \big(\mathcal{N}^{\otimes k}(\vec{x}), \mathcal{R}^{\otimes k}(\vec{x})\big)$ for all legnth-k sequences $\vec{x}$. Thus,

$$\underline{H}(\mathcal{N}) \leqslant H(\mathcal{N}) \leqslant \overline{H}(\mathcal{N}) = \underline{H}(\mathcal{N})$$

and so

$$\begin{aligned} H(\mathcal{N}) &= \log|B| - \max_x \Big(D\Big(\mathcal{N}(x)\big|\big|\mathcal{R}(x)\Big)\Big) \\ &= \min_x H_S(\mathcal{N}(x)) \end{aligned}$$

where H_S is the Shannon entropy.

Finally, we demonstrate that $H(\mathcal{N}) = \min_x H_S\big(\mathcal{N}(x)\big)$ satisfies definition 20 and thus is an entropy. It has been shown in [GW18] that H is additive and non-decreasing under the action of a uniformity-preserving superchannel. It then follows from Appendix D that H is non-decreasing under superchannels of the form

$$\Theta[\mathcal{N}](x) = \sum_z p(z|x)\mathcal{V}_z \circ \mathcal{N} \circ \mathcal{S'}_z(x) \quad \forall x$$

and therefore $H(\mathcal{M}) \geqslant H(\mathcal{N})$ if $\mathcal{M} \preceq \mathcal{N}$. ∎

Finally, we note that one natural extension of these results is to characterise the uncertainty of *quantum* channels. Unlike the classical case, we expect the ordering induced by quantum gambling games to also take into account resources such as entanglement.

7.4 Operational Interpretation of Dynamical Monotones

In this section, we demonstrate that the payoff function $\text{Prob}_T(\mathcal{N})$ provides a first operational interpretation of the dynamical monotones introduced in [GW18].

We begin by defining a quantum channel $\mathcal{N}$ as a completely-positive trace preserving map from system A to system B, which can be written as $\mathcal{N} \in \text{CPTP}(\text{A} \to \text{B})$. A superchannel Θ is then defined as a map from quantum channels to quantum channels s.t. $\Theta : \text{CPTP}(\text{A}_0 \to \text{A}_1) \to \text{CPTP}(\text{B}_0 \to \text{B}_1)$.

The Choi matrix $J_{\mathcal{N}}$ of channel $\mathcal{N}$ is defined as

$$J_{\mathcal{N}} \triangleq \sum_{i,j} |i\rangle\langle j| \otimes \mathcal{N}(|i\rangle\langle j|)$$

We first restate the lemma first proved in [GW18], which introduces dynamical monotones $f_{\mathcal{P}}$.

Lemma 22. *Let FREE($A \to B$) be a convex and topologically closed set where $A = (A_0, A_1)$ and $B = (B_0, B_1)$. Denote the Choi matrix for any channel $\mathcal{N}$ as $J_{\mathcal{N}}$ For any quantum channel $\mathcal{P}_B \in CPTP(B_0 \to B_1)$, define*

$$\begin{aligned} f_{\mathcal{P}}(\mathcal{N}_A) &\triangleq \max_{\Theta \in FREE(A \to B)} \left\langle \mathcal{P}, \Theta[\mathcal{N}_A] \right\rangle \\ &= \max_{\Theta \in FREE(A \to B)} Tr\Big(J_{\mathcal{P}}^{\dagger} J_{\Theta[\mathcal{N}]} \Big) \end{aligned}$$

for every $\mathcal{N}_A \in CPTP(A_0 \to A_1)$. Let $\mathcal{N}_A \in CPTP(A_0 \to A_1)$ and $\mathcal{M}_B \in CPTP(B_0 \to B_1)$ be two quantum channels. Then, $\mathcal{M}_B = \Theta_{A \to B}[\mathcal{N}_A]$, for some super-channel $\Theta \in FREE(A \to B)$ if and only if

$$f_{\mathcal{P}}(\mathcal{N}_A) \geqslant f_{\mathcal{P}}(\mathcal{M}_B) \qquad \forall P \in CPTP(B_0 \to B_1)$$

Proof. See [GW18] for a complete proof. ∎

We now demonstrate that our work provides an operational interpretation of Lemma 9. Namely, the reward functions for expected channel payoff, $\text{Prob}_{\mathcal{T}}$ correspond to dynamical monotones in quantum resource theory:

Theorem 23. *Define the set of free classical superchannels $FREE(A \to B)$ to be all classical, uniformity preserving superchannels Θ and define*

$$f_{\mathcal{P}}(\mathcal{N}) \triangleq \max_{\Theta \in FREE(A \to B)} \langle \mathcal{P}, \Theta[\mathcal{N}] \rangle$$

for any quantum channel $\mathcal{P}$. Then, $\mathcal{M} \preceq \mathcal{N}$ if and only if

$$f_{\mathcal{P}}(\mathcal{N}_A) \geqslant f_{\mathcal{P}}(\mathcal{M}_B) \qquad \forall P \in CPTP(B_0 \to B_1)$$

Proof. See Appendix 5 for details. ∎

7.5 Summary

In Chapters 6 and 7, we introduce a new method for characterizing channel uncertainty via payoff functions from games of chance. From this, we introduce families of games of chance which give rise to three different partial orders: majorization, conditional majorization, and channel majorization. For each order we provide an operational interpretation, and for the case of channel majorization we find the unique corresponding asymptotically continuous channel entropy. Finally, we provide a connection between our results and dynamical monotones in quantum dynamical resource theory.

One natural extension of this work is to characterise the uncertainty of quantum channels. Unlike the classical case, we expect the ordering induced by quantum gambling games to also take into account resources such as entanglement. Preliminary work in this direction suggests that their also exists a unique asymptotically continuous channel entropy in the quantum case.

Finally, previous works on majorization have led to a variety of applications such as

finding the capacity of bosonic Gaussian channels, computing quantum discord, characterizing allowed thermodynamic transformations, and developing entanglement detection protocols [MGH14, GHM15, HO15, GJB+18, WWS20]. Likewise, previous entropy definitions such as the Shannon entropy have had wide-reaching applications in thermodynamics and data compression [CC05, BFL11]. Thus, given the central role of channels in multiple areas of physics and our results for characterizing channel entropy via majorization, we expect broad potential applications in thermodynamics and information theory.

8

Conclusions

The first part of this thesis discusses locally adaptive procedures for quantum hypothesis testing. Adaptive protocols offer a local, experimentally feasible alternative to collective measurements on a large quantum system. First, this thesis discusses variations of locally greedy algorithms for binary state discrimination. In the first locally greedy protocol, subsystems are measured in order with each measurement consisting of the locally optimal Helstrom measurement for the updated prior. Thus, all information about previous measurements is encoded via the prior which is updated using Bayes' theorem after each measurement result.

As an extension of previous results by Acin et al [ABB$^+$05], this thesis demonstrates that the locally greedy strategy is completely optimal for distinguishing *any* two pure TPQS. Despite the optimality for pure states, the locally greedy approach is not even asymptotically optimal when the candidate states are mixed. To remedy this, we introduce a modified locally greedy algorithm which is asymptotically optimal for mixed states and retains optimality for pure states. Finally, a general dynamic programming algorithm is introduced which is capable of finding the opti-

mal locally adaptive measurement *and* order of measurement for any pair of TPQS. We provide results of numerical simulations on up to 10 subsystems, and note that beyond 10 subsystems the computational complexity of the DP algorithm makes it difficult to implement.

We then demonstrate that for larger dimensional subsystems (ie qutrits) the dynamic programming algorithm performs better when allowed to implement measurements with multiple outcomes. It remains an open question whether one can upper bound the number of measurement outcomes needed to obtain an optimal adaptive measurement scheme as a function of the subsystem dimension.

The remainder of the work on state discrimination discusses implementation of an effective reinforcement learning (RLNN) algorithm for multiple state discrimination. Multiple numerical simulations are run for a variety of system sizes and numbers of candidate states, and results for the RLNN performance are compared to the optimal (collective) performance as found via SDP programming. The results demonstrate that RLNN is optimal or close-to-optimal in every trial if the system has less than 20 qubits. Additionally, the RLNN algorithm is provably robust under small rotation errors.

This work paves the way for developing even more sophisticated RLNN protocols capable of finding locally adaptive protocols for even larger systems. Other key open questions are whether locally adaptive protocols are equally effective for entangled states, and what the "worst case" gap is between the best locally adaptive procedure and the best collective procedure.

The second part of this thesis discusses using games of chance as a way to measure

uncertainty in physical systems. Such families of games of chance allow for pre-orders on classical channels corresponding to channel entropy. Additionally, this thesis provides an operational interpretation for the preorder– namely, channel $\mathcal{M} \preceq \mathcal{N}$ if and only if $\mathcal{M}$ can be simulated with $\mathcal{N}$ using classical communication and shared randomness.

Finally, we demonstrate that the payoff functions from the games of chance provide a first operational interpretation of monotones in quantum dynamical resource theory. We additionally find the unique asymptotically continuous entropy of a quantum channel. Future work aims to extend this approach to *quantum* gambling games and thus to characterise the entropy of a quantum channel.

In summary, this dissertation contributes to the topics of quantum state discrimination and characterisation of quantum devices. More broadly, this work adds to the building evidence [FTWM18, MKHV18] that reinforcement learning is a powerful tool capable of improving current protocols in quantum information theory.

Appendix 1

Theorem *Let $P_{\mathrm{s,h}}(q,\rho_\pm)$ and $P_{\mathrm{s,lg}}(q,\rho_\pm)$ denote the probabilities of successful state discrimination, given initial prior $\mathbb{P}(\rho=\rho_+)=q$, using the joint N-system Helstrom measurement and the locally greedy measurement technique, respectively. If ρ_+ and ρ_- are pure states, i.e., $\rho_\pm^{(j)}=|\pm\theta_j\rangle\langle\pm\theta_j|$ where $|\theta\rangle\triangleq\cos\frac{\theta}{2}|0\rangle+\sin\frac{\theta}{2}|1\rangle$, for some $\theta_j\in(0,2\pi)$ for every $j\in[N]$, then*

$$P_{\mathrm{s,h}}(q,\rho_\pm)=P_{\mathrm{s,lg}}(q,\rho_\pm) \tag{1}$$

$$=\frac{1}{2}\left(1+\sqrt{1-4q(1-q)\Pi_{j=1}^{N}\cos^2(\theta_j)}\right). \tag{2}$$

We prove the statement by induction, first considering the base case where $N=2$ and additionally specifying $q=\frac{1}{2}$. Then the probability of success can be written as

$$P_{\mathrm{s,lg}}\left(\frac{1}{2},\rho_\pm\right)=\sum_{d_1\in\{+,-\}}\mathbb{P}(\rho=\rho_+)\mathbb{P}\left(d_1\middle|\rho_+,\frac{1}{2}\right)\mathbb{P}\left(d_2=+\middle|\rho_+,d_1,\frac{1}{2}\right) \tag{3}$$

$$+\sum_{d_1\in\{+,-\}}\mathbb{P}(\rho=\rho_-)\mathbb{P}\left(d_1\middle|\rho_-,\frac{1}{2}\right)\mathbb{P}\left(d_2=-\middle|\rho_-,d_1,\frac{1}{2}\right)$$

$$=\frac{1}{2}\sum_{d_1\in\{+,-\}}\left[\mathbb{P}\left(d_1\middle|\rho_+,\frac{1}{2}\right)\mathbb{P}\left(d_2=+\middle|\rho_+,d_1,\frac{1}{2}\right)+\mathbb{P}\left(d_1\middle|\rho_-,\frac{1}{2}\right)\mathbb{P}\left(d_2=-\middle|\rho_-,d_1,\frac{1}{2}\right)\right]$$

$$=\mathbb{P}\left(d_1=+\middle|\rho_+,\frac{1}{2}\right)\mathbb{P}\left(d_2=+\middle|\rho_+,d_1=+,\frac{1}{2}\right)+\mathbb{P}\left(d_1=+\middle|\rho_-,\frac{1}{2}\right)\mathbb{P}\left(d_2=-\middle|\rho_-,d_1=+,\frac{1}{2}\right), \tag{4}$$

where the final equality follows from symmetry of probability outcomes for pure states, i.e.,

$$\mathbb{P}\left(d_j=+\middle|\rho_+,P_{j-1}=\frac{1}{2}\pm\alpha\right)=\mathbb{P}\left(d_j=-\middle|\rho_-,P_{j-1}=\frac{1}{2}\mp\alpha\right).$$

From the definition of the Helstrom measurement, we observe that

$$\mathbb{P}\left(d_1=\pm\middle|\rho_\pm,\frac{1}{2}\right)=\frac{1}{2}(1+\sin(\theta_1)),\ \mathbb{P}\left(d_1=\pm\middle|\rho_\mp,\frac{1}{2}\right)=\frac{1}{2}(1-\sin(\theta_1)). \tag{5}$$

Hence the updated prior is $P_1(q = \frac{1}{2}, d_1) = \frac{1}{2}(1 + d_1 \sin(\theta_1))$. Then according to the locally greedy algorithm,

$$\mathbb{P}\left(d_2 \middle| \rho_\pm, d_1, \frac{1}{2}\right) = \begin{cases} 1 - \mathrm{Tr}\left[\Pi\left(P_1(q = \frac{1}{2}, d_1), j = 1\right)\rho_\pm\right] & \text{if } d_2 = +, \\ \mathrm{Tr}\left[\Pi\left(P_1(q = \frac{1}{2}, d_1), j = 1\right)\rho_\pm\right] & \text{if } d_2 = -. \end{cases} \tag{6}$$

After some simplifications and using a result from [Hel69], we observe

$$\mathbb{P}\left(d_2 = \pm \middle| \rho_+, d_1 = \pm, \frac{1}{2}\right) = \frac{1}{2}\left(1 \pm \frac{\sin^2\theta_2 \pm \cos\theta_2 \sin\theta_1}{2\sqrt{\cos^2(\theta_2)\sin^2(\theta_1) + \sin^2(\theta_2)}}\right). \tag{7}$$

Again using the symmetry property we have

$$\mathbb{P}\left(d_2 = \pm \middle| \rho_-, d_1 = \pm, \frac{1}{2}\right) = \frac{1}{2}\left(1 \mp \frac{\sin^2\theta_2 \pm \cos\theta_2 \sin\theta_1}{2\sqrt{\cos^2(\theta_2)\sin^2(\theta_1) + \sin^2(\theta_2)}}\right). \tag{8}$$

Upon substitution we obtain

$$\begin{aligned} P_{\mathrm{s,lg}}\left(\frac{1}{2}, \rho_\pm\right) &= \mathbb{P}\left(d_1 = + \middle| \rho_+, \frac{1}{2}\right)\mathbb{P}\left(d_2 = + \middle| \rho_+, d_1 = +, \frac{1}{2}\right) \qquad (9) \\ &\quad + \mathbb{P}\left(d_1 = + \middle| \rho_-, \frac{1}{2}\right)\mathbb{P}\left(d_2 = - \middle| \rho_-, d_1 = +, \frac{1}{2}\right) \\ &= \frac{1}{2}\left(1 + \sqrt{\cos^2(\theta_2)\sin^2(\theta_1) + \sin^2(\theta_2)}\right) \\ &= \frac{1}{2}\left(1 + \sqrt{1 - \cos^2(\theta_1)\cos^2(\theta_2)}\right). \qquad (10) \end{aligned}$$

For the inductive step, we define a new variable $\tilde{\theta} \in [0, \pi]$ such that $\cos^2(\tilde{\theta}) \triangleq \Pi_{i=1}^{N-1}\cos^2(\theta_i)$. Then by assumption

$$P_{\mathrm{s,lg}}\left(\frac{1}{2}, \rho_\pm^{(1,\ldots,N-1)}\right) = \frac{1}{2}\left(1 + \sqrt{1 - \Pi_{i=1}^{N-1}\cos^2(\theta_i)}\right) = \frac{1}{2}\left(1 + \sqrt{1 - \cos^2(\tilde{\theta})}\right). \tag{11}$$

We then apply the previously shown statement for $N = 2$, letting the first subsystem

now be the combined subsystems $1, 2, \ldots, N-1$, i.e., $\rho_{\pm}^{(1,\ldots,N-1)}$.

$$P_{\text{s,lg}}\left(\frac{1}{2}, \rho_{\pm}\right) = \frac{1}{2}\left(1 + \sqrt{1 - \cos^2(\tilde{\theta})\cos^2(\theta_N)}\right)$$

$$= \frac{1}{2}\left(1 + \sqrt{1 - \Pi_{i=1}^{N-1}\cos^2(\theta_i)\cos^2(\theta_N)}\right) \tag{12}$$

$$= \frac{1}{2}\left(1 + \sqrt{1 - \Pi_{i=1}^{N}\cos^2(\theta_i)}\right). \tag{13}$$

Now we consider the case of general priors. We can artificially rearrange this problem so that it is mathematically equivalent to a new quantum state discrimination problem between two transformed states $\hat{\rho'}_{\pm}$. We start by defining θ_0 such that $q = \frac{1}{2}(1 + \sin(\theta_0))$. For pure states, we have $P_{\text{s,h}}(q, \rho_{\pm}) = P_{\text{s,h}}(1-q, \rho_{\pm})$, so

$$P_{\text{s,h}}(q, \rho_{\pm}) = \frac{1}{2}\left[P_{\text{s,h}}\left(\frac{1+\sin\theta_0}{2}, \rho_{\pm}\right) + P_{\text{s,h}}\left(\frac{1-\sin\theta_0}{2}, \rho_{\pm}\right)\right]$$

$$= P_{\text{s,lg}}\left(\frac{1}{2}, \rho_{\pm}^{(0)} \otimes \rho_{\pm}\right)$$

$$= \frac{1}{2}\Bigg[\mathbb{P}\left(d_0 = + \middle| q_0 = \frac{1}{2}, \rho^{(0)} = \rho_{+}^{(0)}\right) \cdot \mathbb{P}\left(d_N = + \middle| q = \frac{1+\sin\theta_1}{2}, \rho = \rho_{+}\right)$$

$$+ \mathbb{P}\left(d_0 = + \middle| q_0 = \frac{1}{2}, \rho^{(0)} = \rho_{-}^{(0)}\right) \cdot \mathbb{P}\left(d_N = - \middle| q = \frac{1+\sin\theta_1}{2}, \rho = \rho_{-}\right)$$

$$+ \mathbb{P}\left(d_0 = - \middle| q_0 = \frac{1}{2}, \rho^{(0)} = \rho_{+}^{(0)}\right) \cdot \mathbb{P}\left(d_N = + \middle| q = \frac{1-\sin\theta_1}{2}, \rho = \rho_{+}\right)$$

$$+ \mathbb{P}\left(d_0 = - \middle| q_0 = \frac{1}{2}, \rho^{(0)} = \rho_{-}^{(0)}\right) \cdot \mathbb{P}\left(d_N = - \middle| q = \frac{1-\sin\theta_1}{2}, \rho = \rho_{+}\right)\Bigg],$$

where we define the newly appended states $\rho_{\pm}^{(0)} \triangleq |\theta_{0,\pm}\rangle\langle\theta_{0,\pm}|$ such that $|\langle\theta_{0,+}\rangle\theta_{0,-}|^2 = \cos^2(\theta_0)$. Here $\mathbb{P}(d_N \mid P_1(q_0 = \frac{1}{2}, d_0), \text{state})$ denotes the probability of obtaining d_N as the measurement result on the N^{th} subsystem given the specified (updated) prior and state, with all local measurements determined by the locally greedy algorithm.

Since we have restricted all quantum subsystems to be in pure states, we can simplify through symmetry as follows.

$$\mathbb{P}\left(d_N = + \middle| q = \frac{1-\sin\theta_1}{2}, \rho = \rho_+\right) = \mathbb{P}\left(d_N = - \middle| q = \frac{1+\sin\theta_1}{2}, \rho = \rho_-\right)$$

$$\mathbb{P}\left(d_N = - \middle| q = \frac{1-\sin\theta_1}{2}, \rho = \rho_-\right) = \mathbb{P}\left(d_N = + \middle| q = \frac{1+\sin\theta_1}{2}, \rho = \rho_+\right).$$

Substituting these properties we obtain

$$\begin{aligned} P_{\mathrm{s,h}}(q, \rho_\pm) &= \mathbb{P}\left(d_0 = + \middle| q_0 = \frac{1}{2}, \rho^{(0)} = \rho_+^{(0)}\right) \cdot \mathbb{P}\left(d_N = + \middle| q = \frac{1+\sin\theta_1}{2}, \rho = \rho_+\right) \\ &\quad + \mathbb{P}\left(d_0 = + \middle| q_0 = \frac{1}{2}, \rho^{(0)} = \rho_-^{(0)}\right) \cdot \mathbb{P}\left(d_N = - \middle| q = \frac{1+\sin\theta_1}{2}, \rho = \rho_-\right) \\ &= P_{\mathrm{s,lg}}(q, \rho_\pm). \end{aligned}$$

Thus, the probability of success is equivalent under both the joint N-system Helstrom measurement and the locally greedy method for pure states.

Appendix 2

We show the plateau remains in the order-optimized locally greedy algorithm by generalizing the experiment from identical copies to the case where the subsystems are distinct. The primary change is that we now sample states parameterized by $\theta_{\pm}^{(t,j)}$ so that each subsystem in both ρ_+ and ρ_- can have (potentially) distinct copies. Also, the vector of success probabilities is altered accordingly.

1. Choose a set of depolarizing parameters and number of trials. Again we set $\mathcal{S}_{\text{dep}} = \{0.01, 0.05, 0.1, 0.3\}$ and $n_{\text{trial}} = 1000$.

2. Generate $\theta_{\pm}^{(t,j)} \in (0, 2\pi)$ uniformly, where $t \in [n_{\text{trial}}]$ denotes the trial index, and $j = 1, 2, \ldots, 12$ denotes the subsystem index.

3. For each $\gamma \in \mathcal{S}_{\text{dep}}$ and $N = 1, 2, \ldots, 12$, define the corresponding qubit quantum states

$$\rho_{\pm}(\gamma, t, N) \triangleq \bigotimes_{j=1}^{N} \left((1-\gamma) \left| \theta_{\pm}^{(t,j)} \right\rangle \left\langle \theta_{\pm}^{(t,j)} \right| + \frac{\gamma}{2} I \right). \tag{14}$$

4. For all $\gamma \in \mathcal{S}_{\text{dep}}$ and all $N = 1, 2, \ldots, 12$, denote

$$P_{\text{succ}}(N, \gamma) = \frac{1}{n_{\text{trial}}} \sum_{t=1}^{n_{\text{trial}}} P_{\text{s,lg}} \left(\rho_{\pm}(\gamma, t, N) \right), \tag{15}$$

   where $P_{\text{s,lg}}(\rho_{\pm})$ indicates that we perform the locally greedy algorithm on states $\rho_{\pm}$.

We plot the results of this experiment in Fig. 4.1. When the subsystems are not identical copies, we notice that the plateau is much higher when compared to the case of identical copies. At first glance this appears to violate the bound obtained in Corollary 4, since here the pair of states in each subsystem are pure states depolarized with the same parameter γ, as required in the hypothesis of Corollary 4. However, the reason for the higher plateau is as follows. The algorithm orders the subsystems in such a way that the credulity can be updated to be as close to $1 - \frac{\gamma}{2}$ as possible

(where we assume the states $\rho_\pm$ may be relabeled at any step to ensure the credulity is always greater than $\frac{1}{2}$). In the next round, it is still possible to obtain one more non-trivial measurement, after which either the updated credulity exceeds $1-\frac{\gamma}{2}$ and all subsequent rounds are trivial, or the credulity is lowered below the threshold and another measurement is permitted until the updated credulity again exceeds $1-\frac{\gamma}{2}$. This permitted "jump" in credulity due to the final measurement explains why the value appearing as the plateau in Fig. 1 can be larger than $1-\frac{\gamma}{2}$.

The best "jump" beyond $1-\frac{\gamma}{2}$ is obtained when the states in that subsystem are an orthogonal pair of pure states subjected to the depolarizing channel and a measurement result which increases the credulity is attained, as formalized in the following lemma.

Lemma 24. *Suppose that we are given one of two quantum states* $\rho_+ = \bigotimes_{j=1}^{N}\Big((1-\gamma)\left|\theta_{j,+}\right\rangle\!\left\langle\theta_{j,+}\right| + \frac{\gamma}{2}\mathbb{I}\Big)$ *and* $\rho_- = \bigotimes_{j=1}^{N}\Big((1-\gamma)\left|\theta_{j,-}\right\rangle\!\left\langle\theta_{j,-}\right| + \frac{\gamma}{2}\mathbb{I}\Big)$ *where* γ *is a fixed depolarizing parameter. Then, an upper bound on the probability of success using the locally greedy method is given by*

$$P_{\mathrm{s,lg}}\big(\frac{1}{2},\rho_\pm\big) \leqslant P_{bound}(\gamma) \equiv \frac{(1-\frac{\gamma}{2})^2}{(1-\frac{\gamma}{2})^2+(\frac{\gamma}{2})^2}.$$

Proof. Given the sequence of measurement results $d_{[N]}$, the probability of success is $\max(p_N, 1-p_N)$ where $p_N = C_N^\sigma(q, \mathbf{a}_{[N]}^\sigma, \mathbf{d}_{[N]}^\sigma)$ (and where $\mathbf{a}_{\sigma(j)} = \Pi(j, p_j)$.) We suppose that at any step we may swap the labels of the composite states to enforce $p_j \geqslant \frac{1}{2} \;\; \forall\; j$. Then, the probability of success is upper bounded by the maximal attainable probability, namely $P_{s,lg}(\frac{1}{2},\rho_\pm) \leqslant \max_{\mathbf{d}_{[N]}^\sigma \in \mathcal{D}^N}\big(C_N^\sigma(\frac{1}{2},\mathbf{d}_{[N]}^\sigma)\big)$.

We then show by induction that $p_j \leqslant \frac{(1-\frac{\gamma}{2})^2}{(1-\frac{\gamma}{2})^2+(\frac{\gamma}{2})^2} \;\; \forall\; j \in \{0,1,...,N\}$. For the base case, $j=0$ and the statement is trivially true as $p_0 = q = \frac{1}{2}$. For the inductive step, we assume that the statement holds for $j \in \{0,1,...,N-1\}$ and show that it then also holds for $j+1$. First, consider the case where the updated prior at the j-th step

exceeds the critical value $1 - \frac{\gamma}{2} \leqslant p_j$. Then, all future measurements are trivial and $p_k = p_j$ for all $k \geqslant j$. Thus, the inductive hypothesis again holds.

Next, consider the case where $p_j \in [\frac{1}{2}, 1-\frac{\gamma}{2}]$. For simplicity, we define $p_{j+1}(p_j, \mathbf{d}^\sigma_{j+1}) \triangleq C^\sigma_{j+1}(q, \mathbf{a}^\sigma_{[j+1]}, \mathbf{d}^\sigma_{[j+1]})$ with $a_{\sigma(j)} = \Pi(j, p_j)$. If the Helstrom measurement $\Pi_{hel} \equiv \Pi(j+1, p_j)$ is trivial, then $p_j = p_{j+1}$ and the inductive hypothesis holds. In the case where Π_{hel} is nontrivial, the prior will increase when $d_{j+1} = +$ and the new maximal credulity p^*_{j+1} is defined as follows:

$$
\begin{aligned}
p^*_{j+1} &= \max\Big(p_{j+1}(p_j,\ \mathbf{d}^\sigma_{j+1} = +), p_{j+1}(p_j,\ \mathbf{d}^\sigma_{j+1} = -)\Big) \\
&= p_{j+1}(p_j,\ \mathbf{d}^\sigma_{j+1} = +) \\
&= \frac{p_j \mathrm{Tr}[\big(\mathbb{I} - \Pi_h\big)\big((1-\gamma)\left|\theta_{j,+}\right\rangle\!\left\langle\theta_{j,+}\right| + \frac{\gamma}{2}\mathbb{I})]}{\mathrm{Tr}[\big(\mathbb{I} - \Pi_h\big)\big(p_j(1-\gamma)\left|\theta_{j,+}\right\rangle\!\left\langle\theta_{j,+}\right| + (1-p_j)(1-\gamma)\left|\theta_{j,-}\right\rangle\!\left\langle\theta_{j,-}\right| + \frac{\gamma}{2}\mathbb{I})]} \\
&= \frac{(1-\gamma)p_j \mathrm{Tr}[\big(\mathbb{I} - \Pi_h\big)\left|\theta_{j,+}\right\rangle\!\left\langle\theta_{j,+}\right|] + \frac{p_j\gamma}{2}}{(1-\gamma)\mathrm{Tr}[\big(\mathbb{I} - \Pi_h\big)\big(p_j\left|\theta_{j,+}\right\rangle\!\left\langle\theta_{j,+}\right| + (1-p_j)\left|\theta_{j,-}\right\rangle\!\left\langle\theta_{j,-}\right|] + \frac{\gamma}{2}} \\
&= \frac{(1-\gamma)p_j x_+ + \frac{p_j\gamma}{2}}{(1-\gamma)(p_j x_+ + (1-p_j)x_-) + \frac{\gamma}{2}}
\end{aligned}
$$

for $x_\pm \equiv \mathrm{Tr}[(\mathbb{I} - \Pi_h)\left|\theta_{j,\pm}\right\rangle\!\left\langle\theta_{j,\pm}\right|] \in [0,1]$. The third line follows from substituting into Bayes' law and simplifying. In the following, we derive an upper bound on p^*_{j+1} and thus the success probability by optimizing over x_+, x_-, p_j without placing any restrictions on whether the optimal set $\{x^*_+, x^*_-, p^*_j\}$ is actually physically realizable.

$$
\begin{aligned}
p^*_{j+1} &\leqslant \max_{x_\pm\in[0,1]} \max_{p_j\in[\frac{\gamma}{2},1-\frac{\gamma}{2}]} \Big(\frac{(1-\gamma)p_j x_+ + \frac{p_j\gamma}{2}}{(1-\gamma)(p_j x_+ + (1-p_j)x_-) + \frac{\gamma}{2}} \Big) \\
&= \max_{x_+\in[0,1]} \max_{p_j\in[\frac{\gamma}{2},1-\frac{\gamma}{2}]} \Big(\frac{(1-\gamma)p_j x_+ + \frac{p_j\gamma}{2}}{(1-\gamma)p_j x_+ + \frac{\gamma}{2}} \Big) \\
&= \max_{p_j\in[\frac{\gamma}{2},1-\frac{\gamma}{2}]} \Big(\frac{(1-\gamma)p_j + \frac{p_j\gamma}{2}}{(1-\gamma)p_j + \frac{\gamma}{2}} \Big) \\
&= \frac{(1-\gamma)(1-\frac{\gamma}{2}) + \frac{\gamma}{2}(1-\frac{\gamma}{2})}{(1-\gamma)(1-\frac{\gamma}{2}) + \frac{\gamma}{2}} \\
&= \frac{(1-\frac{\gamma}{2})^2}{(1-\frac{\gamma}{2})^2 + (\frac{\gamma}{2})^2}
\end{aligned}
$$

Thus, $P_{s,lg}\big(\frac{1}{2},\rho_\pm\big) \leqslant p^*_{j+1} \leqslant \frac{(1-\frac{\gamma}{2})^2}{(1-\frac{\gamma}{2})^2+(\frac{\gamma}{2})^2}$. Finally, we note that the bound on p^*_{j+1} is tight and is achieved when $p_j = 1-\frac{\gamma}{2}$ and $\rho_\pm^{(j)} = (1-\gamma)\,|\pm\frac{\pi}{4}\rangle\langle\pm\frac{\pi}{4}| + \frac{\gamma}{2}\mathbb{I}$.

∎

To illustrate the predictive value of this bound, we list the observed numerical asymptotic values found when $N = 12$ for non-identical subsystems ($P_{\text{obs}}(\gamma)$) and the predicted upper bound for $\gamma = 0.1, 0.3, 0.4, 0.5$ respectively:

$$
\Big\{ \big(P_{\text{obs}}(0.1) = 0.9943, P_{\text{bound}}(0.1) = 0.9972\big), \big(P_{\text{obs}}(0.3) = 0.9549, P_{\text{bound}}(0.3) = 0.9698\big),
$$

$$
\big(P_{\text{obs}}(0.4) = 0.9198, P_{\text{bound}}(0.4) = 0.9412\big), \big(P_{\text{obs}}(0.5) = 0.8732, P_{\text{bound}}(0.5) = 0.9000\big) \Big\}.
$$

Finally, we compare the two scenarios for the specific value of the depolarizing parameter $\gamma = 0.3$ in Fig. 2. This plot shows the non-trivial advantage obtained from subsystems being distinct rather than copies of each other, which is the case most considered in the literature. For the special case of $\gamma = 0$, we have shown in Theorem 2 that the order of subsystems does not matter and that the simple locally greedy algorithm itself achieves the optimal performance obtained with the joint N-system Helstrom measurement.

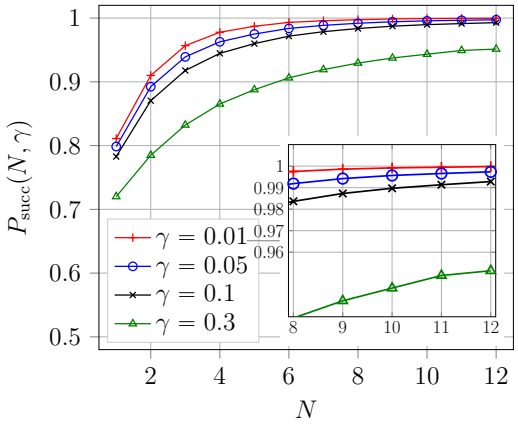

FIGURE 1: Comparison of success probability for varying γ with distinct subsystems, as a function of the number of available systems, N.

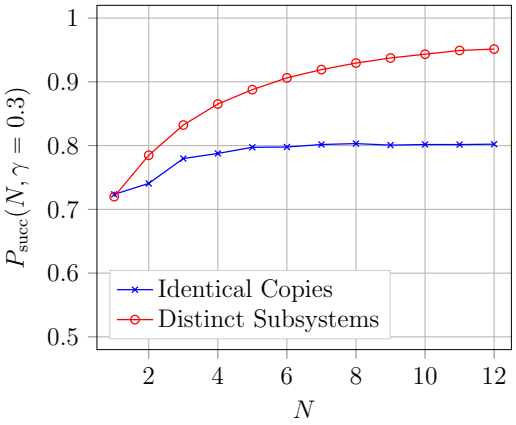

FIGURE 2: Comparison of success probability as a function of the number of available systems, N, for depolarizing parameter $\gamma = 0.3$.

Appendix 3

Theorem: *Consider candidate set* $\{\rho_j\}|_{j=1}^m$ *with prior* $\mathbf{q}$. *Denote by* $P_{succ}\Big(\{\rho_j\}\Big)$ *the probability of success using the optimal locally adaptive method on the original state set. Likewise, let* $P_{succ}\Big(\{\tilde{\rho}_j(\theta)\}\Big)$ *be the success probability for the noisy state set. Then for all* θ,

$$|P_{succ}\Big(\{\rho_j\}\Big) - P_{succ}\Big(\{\tilde{\rho}_j(\theta)\}\Big)| \leqslant n|\theta|$$

where n *is the number of subsystems.* Any adaptive measurement protocol consists of a series of measurements, $\{\Pi_1, \Pi_2(d_1), ...\Pi_n(\mathbf{d}_{[n-1]})\}$ where all measurements after the first depend on previous measurement results. Then any individual measurement sequence can be written as a tensor product:

$$\Pi_{\mathbf{d}_{[n]}} = \bigotimes_{k=1}^{n} \Pi_k(\mathbf{d}_{[k-1]})$$

Let Π_j be the sum of all such measurement sequences s.t. the post-measurement decoding is $\hat{\rho} = \rho_j$. Then the difference between the two success probabilities can be bounded as:

$$\begin{aligned}
\mathrm{P_{succ}}\Big(\{\rho_j\}\Big) - \mathrm{P_{succ}}\Big(\{\tilde{\rho}_j(\theta)\}\Big) &= \sum q_j \mathrm{Tr}[\Pi_j(\rho_j - \tilde{\rho}_j(\theta))] \\
&\leqslant \max_j \left(\mathrm{Tr}\big[\Pi_j\big(\rho_j - \tilde{\rho}_j(\theta)\big)\big] \right) \\
&\leqslant \max_{j,\ \mathbf{d}_{[n]}} \left(\mathrm{Tr}[\Pi_{\mathbf{d}_{[n]}}(\rho_j - \tilde{\rho}_j(\theta))] \right) \\
&= \max_{j,\ \mathbf{d}_{[n]}} \left(\mathrm{Tr}[(\Pi_{\mathbf{d}_{[n]}} - \tilde{\Pi}_{\mathbf{d}_{[n]}}(\theta))\rho_j] \right)
\end{aligned}$$

where

$$\tilde{\Pi}_{\mathbf{d}_{[n]}}(\theta) = \bigotimes_{k=1}^{n} U(\theta)\Pi_k(\mathbf{d}_{[k-1]})U^\dagger(\theta)$$

For simplicity, we here drop notation indicating previous measurement outcomes and represent $\Pi_{\mathbf{d}_{[n]}} = \bigotimes_{k=1}^{n} \Pi_k$. Likewise, we drop the dependence on θ. Then:

$$
\begin{aligned}
\mathrm{Tr}\Big[\bigotimes_k \Pi_k \rho_j^{(k)}\Big] - \mathrm{Tr}\Big[\bigotimes_k \tilde{\Pi}_k \rho_j^{(k)}\Big] &= \mathrm{Tr}\left[\Big(\bigotimes_k \Pi_k - \bigotimes_k \tilde{\Pi}_k\Big) \bigotimes_k \rho_j^{(k)}\right] \\
&\leqslant \Big\|\bigotimes_k \Pi_k - \bigotimes_k \tilde{\Pi}_k\Big\|_\infty \Big\|\bigotimes_k \rho_j^{(k)}\Big\|_1 \\
&\leqslant \Big\|\bigotimes_k \Pi_k - \bigotimes_k \tilde{\Pi}_k\Big\|_\infty \Big\|\Pi_k \otimes \Pi_n - \bigotimes_{k=1}^{n} \tilde{\Pi}_k\Big\|_\infty \\
&\leqslant \Big\|\bigotimes_k^{n} \Pi_k - \bigotimes_{k=1}^{n-1} \tilde{\Pi}_k \otimes \Pi_n\Big\|_\infty + \Big\|\bigotimes_{k=1}^{n} \tilde{\Pi}_k \otimes \Pi_n - \bigotimes_{k=1}^{n} \tilde{\Pi}_k\Big\|_\infty \\
&\leqslant \sum_{\ell=0}^{n-1} \Big\|\bigotimes_{k=1}^{\ell} \tilde{\Pi}_k \otimes \bigotimes_{k=\ell+1}^{n} \Pi_k - \bigotimes_{k=1}^{\ell+1} \tilde{\Pi}_k \otimes \bigotimes_{k=\ell+2}^{n} \Pi_k\Big\|_\infty \\
&= \sum_{\ell=1}^{n} \Big\|\tilde{\Pi}_\ell - \Pi_\ell\Big\|_\infty \\
&\leqslant n \max_\ell \left(\Big\|\tilde{\Pi}_\ell - \Pi_\ell\Big\|_\infty\right).
\end{aligned}
$$

Finally, upon substituting $\tilde{\Pi}_\ell = U\Pi_\ell U^\dagger$, this becomes

$$\begin{aligned} n\max_\ell \left(\left|\left| \tilde{\Pi}_\ell - \Pi_\ell \right|\right|_\infty \right) &= n\max_\ell \left(\left|\left| \Pi_\ell - U\Pi_\ell U^\dagger \right|\right|_\infty \right) \\ &= n\max_\ell \left(\left|\left| \Pi_\ell - ([U,\Pi_\ell] + \Pi_\ell U)U^\dagger \right|\right|_\infty \right) \\ &= n\max_\ell \left(\left|\left| [U,\Pi_\ell]U^\dagger \right|\right|_\infty \right) \\ &= n\max_\ell \left(\left|\left| [(U-\mathbb{I}),(\Pi_\ell - \frac{1}{2}\mathbb{I})] \right|\right|_\infty U^\dagger \right) \\ &= n\max_\ell \left(\left|\left| [(U-\mathbb{I}),(\Pi_\ell - \frac{1}{2}\mathbb{I})] \right|\right|_\infty \right) \\ &\leqslant n\left|\left| U-\mathbb{I} \right|\right|_\infty . \end{aligned}$$

Since $U = U(\theta) = \left(\begin{smallmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{smallmatrix}\right)$ it's easy to check that

$$\begin{aligned} \left\| U - I \right\|_\infty &= \left\| \begin{pmatrix} 1-\cos(\theta) & \sin(\theta) \\ -\sin(\theta) & 1-\cos(\theta) \end{pmatrix} \right\|_\infty \\ &= \sqrt{2}\sqrt{1-\cos(\theta)} \\ &\leqslant |\theta| \end{aligned}$$

Combining everything, we find:

$$|\mathrm{P}_{\mathrm{succ}}\Big(\{\rho_j\}\Big) - \mathrm{P}_{\mathrm{succ}}\Big(\{\tilde{\rho}_j(\theta)\}\Big)| \leqslant \sqrt{2}n|\theta|.$$

Appendix 4

Theorem: *Let P and Q be two $m \times n$ column stochastic matrices. Then,*

$$Q \precsim_c P \quad \iff \quad Q = \sum_z S_z P V_z \tag{16}$$

where each S_z is a sub-stochastic matrix such that $\sum_z S_z$ is a column stochastic matrix (i.e. a classical channel), and each V_z is a permutation matrix.

Lemma 25. *[GGH+18] The matrices P and Q are related as in the RHS of (16) if and only if there exists a column stochastic matrix S such that*

$$QU \leqslant SPU \tag{17}$$

where the inequality is entry-wise and U is the upper triangular matrix

To prove the lemma, denote by $\{\mathbf{p}_x\}_{x=1}^n$ and $\{\mathbf{q}_w\}_{w=1}^n$ the n rows of the matrices P and Q, and by $s_{wz|x}$ and $s_{w|x}$ the components of the matrices S_z and $\sum_z S_z$. Then, the RHS of (16) can be expressed as:

$$\mathbf{q}_w = \sum_{x=1}^n s_{w|x} \mathbf{p}_x D_{wx} \quad \text{where} \quad D_{wx} \equiv \sum_z \frac{s_{wz|x}}{s_{w|x}} V_z \ . \tag{18}$$

Since D_{wx} is a doubly-stochastic matrix we have that for each x and w, $\mathbf{p}_x D_{wx} \prec \mathbf{p}_x$. We therefore conclude that if the matrices Q and P are related as in (18) then there must exists a column stochastic matrix $S = (s_{w|x})$ such that

$$\mathbf{q}_w \precsim \sum_{x=1}^n s_{w|x} \mathbf{p}_x \quad \forall w = 1, ..., n. \tag{19}$$

Conversely, if the relation above holds, then there exists n doubly-stochastic matrices D_w such that

$$\mathbf{q}_w = \sum_{x=1}^n s_{w|x} \mathbf{p}_x D_w \quad \forall w = 1, ..., n. \tag{20}$$

Expressing each $D_w = \sum_z c_{z|w} V_z$ as a convex combination of all permutation matrices we obtain the form (16). Finally, recall that for each w the row vectors $\mathbf{p}_x = \mathbf{p}_x^{\downarrow}$ and $\mathbf{q}_w = \mathbf{q}_w^{\downarrow}$ so that (19) is equivalent to

$$\mathbf{q}_w U \leqslant \sum_{x=1}^{n} s_{w|x} \mathbf{p}_x U \quad \forall w = 1, ..., n, \tag{21}$$

where the inequality is entry-wise. This completes the proof of the lemma.

The question of whether there exists such a column stochastic matrix S that satisfies 17 can be solved with linear programming. In the following lemma, which was proved in [GGH$^+$18], we show that the dual problem can be expressed in terms of sub-linear functionals.

Lemma 26. *[GGH$^+$18] There exists a column stochastic matrix N that satisfies (**??**) if and only if for any set of n vectors $\mathbf{v}_1, ..., \mathbf{v}_n \in \mathbb{R}_+^m$ whose components are arranged in non-decreasing order,*

$$\sum_{w=1}^{n} \max_x \mathbf{p}_w \cdot \mathbf{v}_x \geqslant \sum_{w=1}^{n} \mathbf{q}_w \cdot \mathbf{v}_w \ . \tag{22}$$

To prove the lemma, denote by $\mathbf{s}_x$ the rows of S, so that (**??**) can be expressed as

$$\mathbf{s}_x P U \geqslant \mathbf{q}_x U \tag{23}$$

and the condition that N is column stochastic is equivalent to $\sum_x \mathbf{s}_x = \mathbf{e} \equiv (1, ..., 1)$. Denote further by $\mathbf{s} \equiv (\mathbf{s}_1, ..., \mathbf{s}_n)$ an n^2-dimensional vector, and $\mathbf{b} \equiv (\mathbf{q}_1 U, ..., \mathbf{q}_n U, -\mathbf{e})$ an $(nm + n)$-dimensional vector, and by A the $n^2 \times n(m + 1)$ matrix

$$A \equiv \begin{pmatrix} PU & 0 & \cdots & 0 & -I_n \\ 0 & PU & \cdots & 0 & -I_n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & PU & -I_n \end{pmatrix} \tag{24}$$

With these notations (23) is equivalent to $\mathbf{s}A \geqslant \mathbf{b}$. Now, the question whether there exists $\mathbf{s} \geqslant 0$ that satisfies $\mathbf{s}A \geqslant \mathbf{b}$ is a feasibility problem in linear programming.

From the Farkas lemma, such a row vector $\mathbf{s} \in \mathbb{R}_+^{n^2}$ exists if and only if for any column vector $\mathbf{t} \in \mathbb{R}_+^{n(m+1)}$ that satisfies $A\mathbf{t} \leqslant 0$ we also have $\mathbf{b} \cdot \mathbf{t} \leqslant 0$. Denote $\mathbf{t} = (\mathbf{t}_1, ..., \mathbf{t}_n, \mathbf{r})^T$, where for each $x = 1, ..., n$, $\mathbf{t}_x \in \mathbb{R}_+^m$ and $\mathbf{r} \in \mathbb{R}^n$. Then, with this notation the dual problem has the form

$$\mathbf{r} \geqslant PU\mathbf{t}_x \quad \forall x = 1, ..., n \quad \Rightarrow \quad \mathbf{e} \cdot \mathbf{r} \geqslant \sum_{x=1}^{n} \mathbf{q}_x \cdot (U\mathbf{t}_x) \tag{25}$$

Note that the condition that $\mathbf{r} \geqslant PU\mathbf{t}_x$ for all x can be expressed as $r_w \geqslant \max_x \mathbf{p}_w \cdot (U\mathbf{t}_x)$. Therefore, taking $r_w = \max_x \mathbf{p}_w U\mathbf{t}_x$ we conclude that (25) holds if and only if for any n vectors $\mathbf{t}_1, ..., \mathbf{t}_n \in \mathbb{R}_+^m$

$$\sum_{w=1}^{n} \max_x \mathbf{p}_w \cdot (U\mathbf{t}_x) \geqslant \sum_{x=1}^{n} \mathbf{q}_x \cdot (U\mathbf{t}_x) \ . \tag{26}$$

Denote by $\mathbf{v}_x \equiv U\mathbf{t}_x$, and note that $\mathbf{v}_x = \mathbf{v}_x^{\downarrow}$. Moreover, the inverse of U is given by

$$U^{-1} = \begin{pmatrix} 1 & -1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & -1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \tag{27}$$

Hence, $\mathbf{t}_x \geqslant 0$ if and only if $\mathbf{v}_x = \mathbf{v}_x^{\downarrow}$. We therefore conclude that (25) holds if and only if for any n vectors $\mathbf{v}_1, ..., \mathbf{v}_n \in \mathbb{R}_+^m$ whose components are arranged in non-decreasing order,

$$\sum_{w=1}^{n} \max_x \mathbf{p}_w \cdot \mathbf{v}_x \geqslant \sum_{x=1}^{n} \mathbf{q}_x \cdot \mathbf{v}_x \ . \tag{28}$$

This completes the proof of the lemma.

Appendix 5

Theorem *Define the set of free classical superchannels FREE($A \to B$) to be all classical, random unitary superchannels Θ and define*

$$f_{\mathcal{P}}(\mathcal{N}) \triangleq \max_{\Theta \in FREE(A \to B)} \langle \mathcal{P}, \Theta[\mathcal{N}] \rangle$$

for any quantum channel $\mathcal{P}$. Then, $\mathcal{M} \preceq \mathcal{N}$ if and only if

$$f_{\mathcal{P}}(\mathcal{N}_A) \geqslant f_{\mathcal{P}}(\mathcal{M}_B) \qquad \forall P \in CPTP(B_0 \to B_1)$$

From Theorem 2 of this work, $\text{Prob}_T(\mathcal{M}) \leqslant \text{Prob}_T(\mathcal{N})$ for all T if and only if there exists a free classical superchannel s.t. $\Theta[\mathcal{N}] = \mathcal{M}$. Thus, the application above will hold if we can demonstrate the following:

- For every T with fixed z, there exists a classical channel $\mathcal{P}$ and positive constant α_T such that $\text{Prob}_T(..) = \alpha_T f_{\mathcal{P}}(..)$
- For every classical channel $\mathcal{P}$, there exists a distribution T and positive constant $\alpha_{\mathcal{P}}$ such that $f_{\mathcal{P}}(...) = \alpha_{\mathcal{P}} \text{Prob}_T(...)$

We first prove the second statement.

$$f_{\mathcal{P}}(\mathcal{N}) \triangleq \max_{\Theta \in \text{FREE(A} \to \text{B)}} \text{Tr}\Big(J_{\mathcal{P}}^{\dagger} J_{\Theta[\mathcal{N}]}\Big)$$

$$= \max_{\Theta \in \text{FREE(A} \to \text{B)}} \text{Tr}\left(\Big(\sum_{ij} \mathcal{P}(|i\rangle\langle j|) \otimes |i\rangle\langle j|\Big)\Big(\sum_{i'j'} \Theta[\mathcal{N}](|i\rangle\langle j|) \otimes |i\rangle\langle j|\Big)\right)$$

$$= \max_{\Theta \in \text{FREE(A} \to \text{B)}} \text{Tr}\left(\sum_{j} \mathcal{P}(|j\rangle\langle j|) \otimes |j\rangle\langle j| \Big(\sum_{j'} \Theta[\mathcal{N}](|j'\rangle\langle j'|) \otimes |j'\rangle\langle j'|\Big)\right)$$

$$= \max_{\Theta \in \text{FREE(A} \to \text{B)}} \text{Tr}\left(\sum_{j} \mathcal{P}(|j\rangle\langle j|) \Theta[\mathcal{N}](|j\rangle\langle j|)\right)$$

where the simplification of the Choi Matrix follows from noting that both $\mathcal{P}$ and $\mathcal{N}$

are classical channels. Recalling the definition of free operations, the above becomes

$$\begin{aligned}
f_{\mathcal{P}}(\mathcal{N}) &= \max_{\mathcal{S},\{\mathcal{V}_z\},\{p(k|z)\}} \sum_k \mathrm{Tr}\left[\mathcal{P}(|k\rangle\langle k|)\sum_z p(z|k)\mathcal{V}_z\left(\mathcal{N}\Big(\mathcal{S}(|k\rangle\langle k|)\Big)\right)\right] \\
&= \max_{\mathcal{S},\{\mathcal{V}_z\},\{p(k|z)\}} \sum_k \sum_z p(z|k) \\
&\times \mathrm{Tr}\left[\mathcal{P}(|k\rangle\langle k|)\mathcal{V}_z\left(\mathcal{N}\Big(\mathcal{S}(|k\rangle\langle k|)\Big)\right)\right] \\
&= \max_{\mathcal{S},\{V_k\}} \sum_k \mathrm{Tr}\left[\mathcal{P}(|k\rangle\langle k|)\mathcal{V}_k\left(\mathcal{N}\Big(\mathcal{S}(|k\rangle\langle k|)\Big)\right)\right]
\end{aligned}$$

The third line follows by setting $p(z|k) = \delta_{z,k}$ and noting that this will always be optimal, as

$$\begin{aligned}
&\sum_z p(z|k)\mathrm{Tr}\left[\mathcal{P}(|k\rangle\langle k|)\mathcal{V}_z\left(\mathcal{N}\Big(\mathcal{S}(|k\rangle\langle k|)\Big)\right)\right] \\
&\leqslant \max_z \left(\mathrm{Tr}\left[\mathcal{P}(|k\rangle\langle k|)\mathcal{V}_z\left(\mathcal{N}\Big(\mathcal{S}(|k\rangle\langle k|)\Big)\right)\right]\right)
\end{aligned}$$

We continue simplifying this expression:

$$\begin{aligned}
f_{\mathcal{P}}(\mathcal{N}) &= \max_{\mathcal{S},\{V_k\}} \sum_{k=1}^{d} \mathrm{Tr}\left[\mathcal{P}(|k\rangle\langle k|)\mathcal{V}_k\left(\mathcal{N}\Big(\mathcal{S}(|k\rangle\langle k|)\Big)\right)\right] \\
&= \max_{\mathcal{S},\{V_k\}} \sum_{k=1}^{d}\sum_{y=1}^{d} q_{y|k}\langle y|\,\mathcal{V}_k\left(\mathcal{N}\Big(\mathcal{S}(|k\rangle\langle k|)\Big)\right)|y\rangle \\
&= \max_{\mathcal{S}} \sum_{k=1}^{d}\sum_{y=1}^{d} q^{\downarrow}_{y|k}\langle y|\,\mathcal{N}^{\downarrow}\Big(\mathcal{S}(|k\rangle\langle k|)\Big)|y\rangle \\
&= \sum_{k=1}^{d} \max_x \sum_{y=1}^{d} q^{\downarrow}_{y|k}\langle y|\,\mathcal{N}^{\downarrow}(|x\rangle\langle x|)\,|y\rangle \\
&= \sum_{k=1}^{d} \max_x \sum_{y=1}^{d} q^{\downarrow}_{y|k} p^{\downarrow}_{y|x}
\end{aligned}$$

where $q_{y|k} \triangleq \langle y| \mathcal{P}(|k\rangle\langle k|) |y\rangle$ and $p^{\downarrow}_{y|x} \triangleq \langle y| \mathcal{N}^{\downarrow}(|x\rangle\langle x|) |y\rangle$. Finally, we manipulate the above expression to rewrite it in the form of $\text{Prob}_T(\mathcal{N})$:

$$\begin{aligned}
f_{\mathcal{P}}(\mathcal{N}) &= \sum_{k=1}^{d} \max_x \left(q^{\downarrow}_{1|k} p^{\downarrow}_{1|x} + \sum_{y=2}^{d} q^{\downarrow}_{y|k} (\sum_{j=1}^{y} p^{\downarrow}_{j|x} - \sum_{j=1}^{y-1} p^{\downarrow}_{j|x}) \right) \\
&= \sum_{k=1}^{d} \max_x \left(q^{\downarrow}_{1|k} p^{\downarrow}_{1|x} + \sum_{y=2}^{d} q^{\downarrow}_{y|k} \sum_{j=1}^{y} p^{\downarrow}_{j|x} - \sum_{y=1}^{d} q^{\downarrow}_{y+1|k} \sum_{j=1}^{y} p^{\downarrow}_{j|x}) \right) \\
&= \sum_{k=1}^{d} \max_x \sum_{y=1}^{d} (q^{\downarrow}_{y|k} - q^{\downarrow}_{y+1|k}) \sum_{j=1}^{y} p^{\downarrow}_{j|x} \\
&= \sum_{k=1}^{d} |\tilde{\mathbf{t}}_k| \max_x \sum_{y=1}^{d} \frac{\tilde{t}_{y|k}}{|\tilde{\mathbf{t}}_k|} \sum_{j=1}^{y} p^{\downarrow}_{j|x} \\
&= \sum_k |\tilde{\mathbf{t}}_k| \text{Prob}_{\frac{\tilde{\mathbf{t}}(k)}{|\tilde{\mathbf{t}}(k)|}}(\mathcal{N})
\end{aligned}$$

where $\tilde{\mathbf{t}}_k = \{(q^{\downarrow}_{1|k} - q^{\downarrow}_{2|k}), (q^{\downarrow}_{2|k} - q^{\downarrow}_{3|k}), ..., (q^{\downarrow}_{d-1|k} - q^{\downarrow}_{d|k}), (q^{\downarrow}_{d|k} - 0)\}$ and in line two we use the definition $q^{\downarrow}_{d+1|k} = 0$ (as $q^{\downarrow}_{y|k}$ can always be padded with extra zeros).

Finally, we relabel the dummy variable k to z and simplify:

$$\begin{aligned}
f_{\mathcal{P}}(\mathcal{N}) &= \sum_z |\tilde{\mathbf{t}}_z| \text{Prob}_{\frac{\tilde{\mathbf{t}}_z}{|\tilde{\mathbf{t}}_z|}}(\mathcal{N}) \\
&= (\sum_{z'} |\tilde{\mathbf{t}}_{z'}|) \sum_z \frac{|\tilde{\mathbf{t}}_z|}{\left(\sum_{z'} |\tilde{\mathbf{t}}_{z'}|\right)} \text{Prob}_{\frac{\tilde{\mathbf{t}}_z}{|\tilde{\mathbf{t}}_z|}}(\mathcal{N}) \\
&= (\sum_{z'} |\tilde{\mathbf{t}}_{z'}|) \text{Prob}_T(\mathcal{N})
\end{aligned}$$

where T is defined by its elements:

$$t_{k,z} = \frac{|\tilde{\mathbf{t}}_z|}{\sum_{z'} |\tilde{\mathbf{t}}_{z'}|} \times \frac{\tilde{t}^{\downarrow}_{k|z}}{|\tilde{\mathbf{t}}_z|} = \frac{\tilde{t}^{\downarrow}_{k|z}}{\sum_{z'} |\tilde{\mathbf{t}}_{z'}|}$$

Finally, we need to show that for every T with fixed $z = 1$ (i.e. T is a vector), there exists a classical channel $\mathcal{P}$ and positive constant α_T such that $\text{Prob}_T(\mathcal{N}) =$

$\alpha_T f_{\mathcal{P}}(\mathcal{N})$. However, from the above expression it immediately follows that we can select $\tilde{\mathbf{t}}_{z=1}$ s.t. $T = \frac{\tilde{\mathbf{t}}_{z=1}}{|\tilde{\mathbf{t}}_{z=1}|}$. Then the channel $\mathcal{P}$ corresponding to $\tilde{\mathbf{t}}$ will satisfy $\text{Prob}_T(\mathcal{N}) = \alpha_T f_{\mathcal{P}}(\mathcal{N})$.

Bibliography

[ABB+05] A. Acin, E. Bagan, M. Baig, Ll. Masanes, and R. Munoz-Tapia. Multiple-copy two-state discrimination with individual measurements. *Phys. Rev. A*, 71:032338, 2005.

[Abe00] Sumiyoshi Abe. Axioms and uniqueness theorem for tsallis entropy. *Physics Letters A*, 271(1):74–79, 2000.

[Ban97] M. Ban. Optimum measurements for discrimination among symmetric quantum states and parameter estimation. *International Journal of Theoretical Physics*, 36(6):1269–1288, 1997.

[BBC16] Artificial intelligence: Google's alphago beats go master lee se-dol. *BBC*, 2016.

[BC09] Stephen M. Barnett and Sarah Croke. Quantum state discrimination. *Adv. Opt. Photon.*, 1(2):238–278, 2009.

[BCP+16] Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. OpenAI gym. *arXiv:1606.01540*, 2016.

[BDF+99] Charles H. Bennett, David P. DiVincenzo, Christopher A. Fuchs, Tal Mor, Eric Rains, Peter W. Shor, John A. Smolin, and William K. Wootters. Quantum nonlocality without entanglement. *Physical Review A*, 59(2):1070–1091, 1999.

[Bek72] J. D. Bekenstein. Black holes and the second law. *Lettere al Nuovo Cimento, 4, 737*, 1972.

[Bek73] J. D. Bekenstein. Black holes and entropy. *Physical Review D, 7, 2333*, 1973.

[Bel54] Richard Bellman. The theory of dynamic programming. *Bull. Amer. Math. Soc.*, 60(6):503–515, 1954.

[BFL11] J.C. Baez, T. Fritz, and T. Leinster. A characterization of entropy in terms of information loss. *Entropy , 13(11), 1945-1957*, 2011.

[BGG21] Sarah Brandsen, Isabelle Jianing Geng, and Gilad Gour. What is entropy? a new perspective from games of chance. *arXiv:2103.0861*, 2021.

[BLS+19] Sarah Brandsen, Mengke Lian, Kevin D. Stubbs, Narayanan Rengaswamy, and Henry D. Pfister. Adaptive procedures for discrimination between arbitrary tensor-product quantum states. *arXiv: 1912.05087*, 2019.

[BSP20] Sarah Brandsen, Kevin D. Stubbs, and Henry D. Pfister. Reinforcement learning with neural networks for quantum multiple hypothesis testing. *arXiv:2010.08588*, 2020.

[Buk18] Marin Bukov. Reinforcement learning for autonomous preparation of floquet-engineered states: Inverting the quantum kapitza oscillator. *Phys. Rev. B*, 98:224305, 2018.

[BZ] I. Bengtsson and K. Zyczkowski. Geometry of Quantum States: An Introduction to Quantum Entanglement. *Cambridge University Press.*

[CBW17] Sarah Croke, Stephen M. Barnett, and Graeme Weir. Optimal sequential measurements for bipartite state discrimination. *Physical Review A*, 95(5), 2017.

[CC05] C.G. Chakrabarti and I. Chakrabarty. Shannon entropy: Axiomatic characterization and application. *International Journal of Mathematics and Mathematical Sciences 2005:17, 2847-2854*, 2005.

[CFS16] Bob Coecke, Tobias Fritz, and Robert W. Spekkens. A mathematical theory of resources. *Information and Computation*, 250:59–86, 2016. Quantum Physics and Logic.

[CG19] Eric Chitambar and Gilad Gour. Quantum resource theories. *Rev. Mod. Phys.*, 91:025001, 2019.

[DBB17] M. Dall'Arno, S. Brandsen, and F. Buscemi. Device-independent tests of quantum channels. *Proc. R. Soc. A, 473, 20160721*, 2017.

[DHBS20] M. Dall'Arno, A. Ho, F. Buscemi, and V. Scarani. Data-driven inference and observational completeness of quantum devices. *Phys. Rev. A 102, 062407*, 2020.

[DKJR06] I. Devetak, C. King, M. Junge, and M. B. Ruskai. Multiplicativity of completely bounded p-norms implies a new additivity result. *Communications in Mathematical Physics*, 2006.

[EMV03] Y.C. Eldar, A. Megretski, and G.C. Verghese. Designing optimal quantum detectors via semidefinite programming. *IEEE Transactions on Information Theory*, 49(4):1007–1012, 2003.

[Eul58] Leonhard Euler. Elementa doctrinae solidorum. *Novi Commentarii academiae scientiarum Petropolitanae: 109–140.*, 1758.

[FBC19] Kieran Flatt, Stephen M. Barnett, and Sarah Croke. Multiple-copy state discrimination of noisy qubits. *Physical Review A*, 100(3), 2019.

[FFRS20] K. Fang, O. Fawzi, R. Renner, and D. Sutter. A chain rule for the quantum relative entropy. *Phys. Rev. Lett. 124, 100501*, 2020.

[FGG13] S. Friedland, V. Gheorghiu, and G. Gour. Universal uncertainty relations. Phys. Rev. Lett. 111, 230401, 2013.

[FTWM18] Thomas Fösel, Petru Tighineanu, Talitha Weiss, and Florian Marquardt. Reinforcement learning with neural networks for quantum feedback. *Phys. Rev. X*, 8:031084, 2018.

[GGH+18] G. Gour, A. Grudka, M. Horodecki, W. Klobus, J. Lodyga, and V. Narasimhachar. The Conditional Uncertainty Principle. *Phys. Rev. A 97, 042130*, 2018.

[GHM15] V. Giovannetti, A.S. Holevo, and A. Mari. Majorization and additivity for multimode bosonic gaussian channels. *Theor Math Phys 182, 284–293*, 2015.

[GJB+18] G. Gour, D. Jennings, F. Buscemi, R. Duan, and I. Marvian. Quantum majorization and a complete set of entropic conditions for quantum thermodynamics. *Nat Commun 9, 5352*, 2018.

[GMN+15] G. Gour, M.P. Muller, V. Narasimhachar, R.W. Spekkens, and N.Y. Halpern. The resource theory of informational nonequilibrium in thermodynamics. Physics Reports 583, 1-58, 2015.

[Gor95] Geoffrey J. Gordon. Stable fitted reinforcement learning. In *Proceedings of the 8th International Conference on Neural Information Processing*

Systems, NIPS'95, page 1052–1058, Cambridge, MA, USA, 1995. MIT Press.

[Gou19] G. Gour. Comparison of quantum channels by superchannels. *IEEE Transactions on Information Theory*, 65(9):5880–5904, 2019.

[Gou21] G. Gour. Uniqueness and Optimality of Dynamical Extensions of Divergences. *PRX Quantum 2, 010313*, 2021.

[GS20] G. Gour and C. M. Scandolo. Dynamical Resources. *arXiv:2101.01552v1*, 2020.

[GW18] G. Gour and M. M. Wilde. Entropy of a quantum channel. *arXiv:1808.06980v2*, 2018.

[HDB+11] B. L. Higgins, A. C. Doherty, S. D. Bartlett, G. J. Pryde, and H. M. Wiseman. Multiple-copy state discrimination: Thinking globally, acting locally. *Phys. Rev. A*, 83:052314, 2011.

[Hel69] Carl W Helstrom. Quantum detection and estimation theory. *Journal of Statistical Physics*, 1(2):231–252, 1969.

[HO15] M. Horodecki and J. Oppenheim. Fundamental limitations for quantum and nanoscale thermodynamics. *Nat Communications 4, 2059*, 2015.

[Hol73] Alexander Semenovich Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problemy Peredachi Informatsii*, 9(3):3–11, 1973.

[HS10] Alexander Hentschel and Barry C. Sanders. Machine learning for precise quantum measurement. *Physical Review Letters*, 104(6), 2010.

[IS14] Velimir M. Ilić and Miomir S. Stanković. Generalized shannon–khinchin axioms and uniqueness theorem for pseudo-additive entropies. *Physica A: Statistical Mechanics and its Applications*, 411:138–145, 2014.

[Jak20] A. Jakimowicz. The Role of Entropy in the Development of Economics. *Entropy, 22, 452*, 2020.

[Jay65] E.T. Jaynes. Gibbs vs Boltzmann Entropies. *American Journal of Physics, Vol. 33, No. 5, 391-398*, 1965.

[Kel56] J. L. Kelly. A new interpretation of information rate. *Bell System Technical Journal. 35 (4): 917–926*, 1956.

[KHJ+18] Alex Kendall, Jeffrey Hawke, David Janz, Przemyslaw Mazur, Daniele Reda, John-Mark Allen, Vinh-Dieu Lam, Alex Bewley, and Amar Shah. Learning to drive in a day. *arXiv*, 2018.

[KM18] Alexander Holm Kiilerich and Klaus Mølmer. Multistate and multihypothesis discrimination with open quantum systems. *Physical Review A*, 97(5), 2018.

[KRS09] Robert Koenig, Renato Renner, and Christian Schaffner. The operational meaning of min- and max-entropy. *IEEE Transactions on Information Theory*, 55(9):4337–4347, 2009.

[LLM+17] Eric Liang, Richard Liaw, Philipp Moritz, Robert Nishihara, Roy Fox, Ken Goldberg, Joseph E. Gonzalez, Michael I. Jordan, and Ion Stoica. Rllib: Abstractions for distributed reinforcement learning. *arXiv: 1712.09381*, 2017.

[LLN+18] Richard Liaw, Eric Liang, Robert Nishihara, Philipp Moritz, Joseph E Gonzalez, and Ion Stoica. Tune: A research platform for distributed model selection and training. *arXiv:1807.05118*, 2018.

[Mac92] M. Mackey. Time's Arrow: The Origins of Thermodynamic Behavior. *Berlin Heidelberg New York: Springer*, 1992.

[MDW20] J. Mackeprang, D.B.R. Dasari, and J. Wrachtrup. A reinforcement learning approach for quantum state engineering. *Quantum Mach. Intell. 2, 5*, 2020.

[MGH14] A. Mari, V. Giovannetti, and A. Holevo. Quantum state majorization at the output of bosonic gaussian channels. *Nat Commun 5, 3826*, 2014.

[MKHV18] Mufti Mahmud, Mohammed Shamim Kaiser, Amir Hussain, and Stefano Vassanelli. Applications of deep learning and reinforcement learning to biological data. *IEEE Transactions on Neural Networks and Learning Systems*, 29(6):2063–2079, 2018.

[MKS+13] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Playing Atari with deep reinforcement learning. *arXiv: 1312.5602*, 2013.

[MKS+15] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei Rusu, Joel Veness, Marc Bellemare, Alex Graves, Martin Riedmiller, Andreas Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg, and Demis Hassabis. Human-level control through deep reinforcement learning. *Nature*, 518:529–33, 2015.

[MPNK+18] Alexey A. Melnikov, Hendrik Poulsen Nautrup, Mario Krenn, Vedran Dunjko, Markus Tiersch, Anton Zeilinger, and Hans J. Briegel. Active learning machine learns to create new quantum experiments. *Proceedings of the National Academy of Sciences*, 115(6):1221–1226, 2018.

[MSS20] Alexey A. Melnikov, Pavel Sekatski, and Nicolas Sangouard. Setting up experimental bell tests with reinforcement learning. *Phys. Rev. Lett.*, 125:160401, 2020.

[NC11a] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, USA, 10th edition, 2011.

[NC11b] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, New York, NY, USA, 10th edition, 2011.

[Par11] M. H. Partovi. Majorization formulation of uncertainty in quantum mechanics. Phys. Rev. A 84, 052117, 2011.

[PDM+14] Giuseppe Davide Paparo, Vedran Dunjko, Adi Makmal, Miguel Angel Martin-Delgado, and Hans J. Briegel. Quantum speedup for active learning agents. *Phys. Rev. X*, 4(9), 2014.

[Pri67] I. Prigogine. Introduction to Thermodynamics of Irreversible Processes, third edition. *Interscience Publishers, New York, p. 12.*, 1967.

[PS19] Pantita Palittapongarnpim and Barry C. Sanders. Robustness of quantum-enhanced adaptive phase estimation. *Physical Review A*, 100(1), 2019.

[PWS16] Pantita Palittpongarnpim, Peter Wittek, and Barry C. Sanders. Single-shot adaptive measurement for quantum-enhanced metrology. *Quantum Communications and Quantum Imaging XIV*, 2016.

[PWZ+17] Pantita Palittapongarnpim, Peter Wittek, Ehsan Zahedinejad, Shakib Vedaie, and Barry C. Sanders. Learning in quantum control: High-dimensional global optimization for noisy quantum dynamics. *Neurocomputing*, 268:116–126, 2017.

[Ren60] A. Renyi. On measures of information and entropy. *Proceedings of the fourth Berkeley Symposium on Mathematics, Statistics and Probability*, 1960.

[SB18] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. A Bradford Book, Cambridge, MA, USA, 2018.

[Sha48] C. E. Shannon. A Mathematical Theory of Communication. *The Bell System Technical Journal, Vol. 27, pp. 379–423, 623–656*, 1948.

[Sho94] P.W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. *Proceedings 35th Annual Symposium on Foundations of Computer Science. IEEE Comput. Soc. Press.*, 1994.

[Sho04] Peter W. Shor. Equivalence of additivity questions in quantum information theory. *Communications in Mathematical Physics*, 246(3):473–473, 2004.

[SKBL20] Frank Schäfer, Michal Kloc, Christoph Bruder, and Niels Lörch. A differentiable programming method for quantum control. *Machine Learning: Science and Technology*, 1(3):035009, 2020.

[SKIH98] Masahide Sasaki, Kentaro Kato, Masayuki Izutsu, and Osamu Hirota. Quantum channels showing superadditivity in classical capacity. *Phys. Rev. A*, 58:146–158, 1998.

[SKvNE21a] Ryan Sweke, Markus S Kesselring, Evert P L van Nieuwenburg, and Jens Eisert. Reinforcement learning decoders for fault-tolerant quantum computation. *Machine Learning: Science and Technology*, 2(2):025005, 2021.

[SKvNE21b] Ryan Sweke, Markus S Kesselring, Evert P L van Nieuwenburg, and Jens Eisert. Reinforcement learning decoders for fault-tolerant quantum computation. *Machine Learning: Science and Technology*, 2(2):025005, 2021.

[SPP21] Pierpaolo Sgroi, G. Massimo Palma, and Mauro Paternostro. Reinforcement learning approach to nonequilibrium quantum thermodynamics. *Phys. Rev. Lett.*, 126:020601, 2021.

[SWD+17] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv: 1707.06347*, 2017.

[Tem16] Piergiulio Tempesta. Beyond the shannon-khinchin formulation: The composability axiom and the universal group entropy. *Annals of Physics*, 365:180–197, 01 2016.

[Tes92] Gerald Tesauro. Practical issues in temporal difference learning. *Mach. Learn.*, 8(3–4):257–277, 1992.

[Tra12] A Trabesinger. Quantum simulation. *Nature Phys 8, 263*, 2012.

[VSPM01] S Virmani, M.F Sacchi, M.B Plenio, and D Markham. Optimal local discrimination of two multipartite pure states. *Physics Letters A*, 288(2):62–68, 2001.

[WHBC18] Graeme Weir, Catherine Hughes, Stephen M. Barnett, and Sarah Croke. Optimal measurement strategies for the trine states with arbitrary prior probabilities. *arXiv: 1803.03590*, 2018.

[WMDB20] Julius Wallnöfer, Alexey A. Melnikov, Wolfgang Dür, and Hans J. Briegel. Machine learning for long-distance quantum communication. *PRX Quantum*, 1:010301, 2020.

[WWS20] K. Wang, N. Wu, and F. Song. Entanglement detection via direct-sum majorization. *Sci Rep 10, 452*, 2020.

[XLL+19] Han Xu, Junning Li, Liqiang Liu, Yu Wang, Haidong Yuan, and Xin Wang. Generalizable control for quantum parameter estimation through reinforcement learning. *npj Quantum Inf 5, 82*, 2019.

[YKL75] H. Yuen, R. Kennedy, and M. Lax. Optimum testing of multiple hypotheses in quantum detection theory. *IEEE Transactions on Information Theory*, 21(2):125–134, 1975.

[Yua19] X. Yuan. Relative entropies of quantum channels with applications in resource theory. *Phys. Rev. A 99, 032317*, 2019.

[ZPH+17] J. Zhang, G. Pagano, P. W. Hess, A. Kyprianidis, P. Becker, H. B. Kaplan, A. V. Gorshkov, Z.-X. Gong, and C. Monroe. Observation of a many-body dynamical phase transition in a 53-qubit quantum simulator. *Nature 551, 601*, 2017.

[ZWA+19] Xiao-Ming Zhang, Zezhu Wei, Raza Asad, Xu-Chen Yang, and Xin Wang. When does reinforcement learning stand out in quantum control? a comparative study on state preparation. *npj Quantum Inf 5, 85*, 2019.