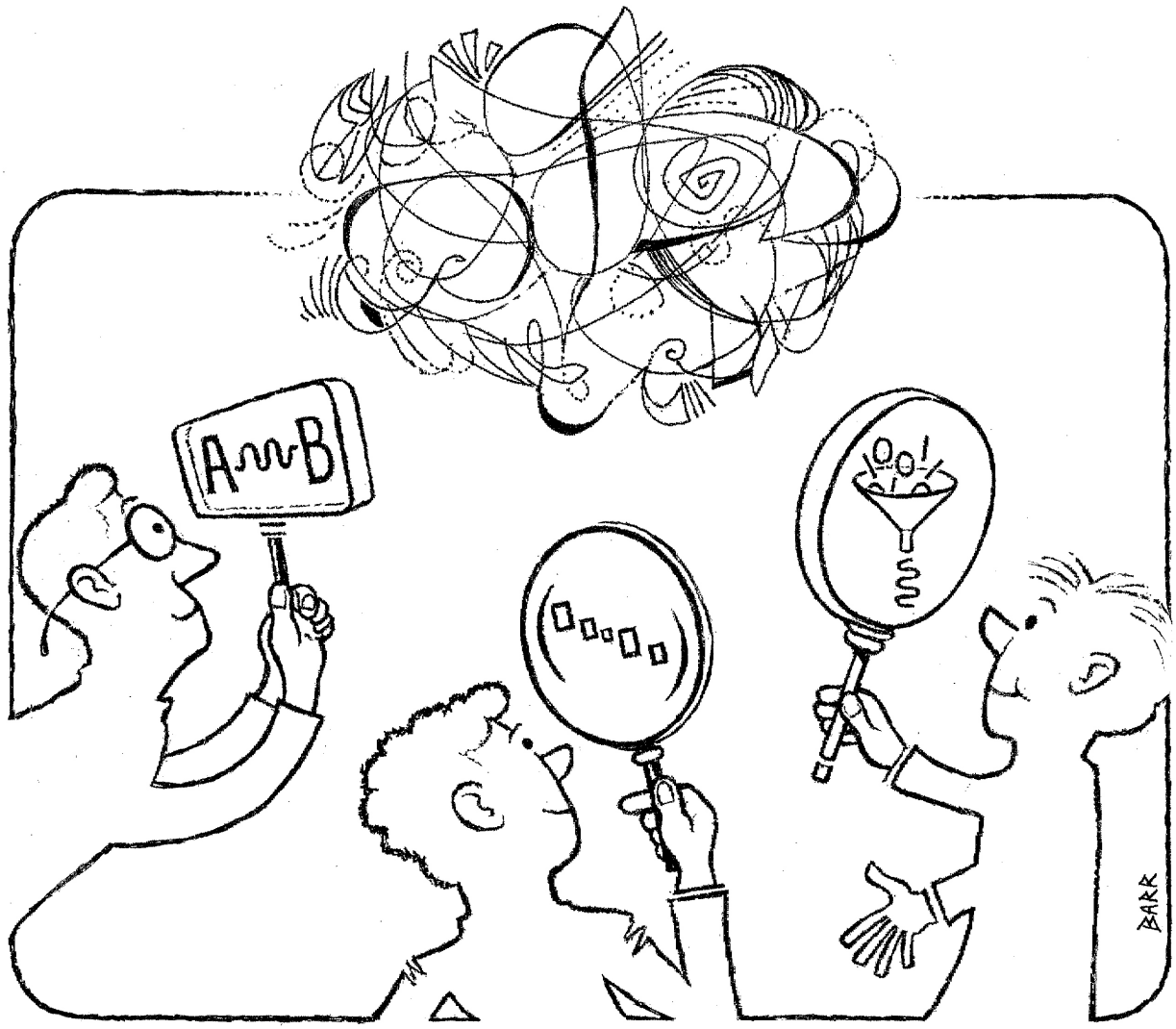
Cryptography in a Quantum World



Stephanie Wehner

Cryptography in a Quantum World

ILLC Dissertation Series DS-2008-01

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

For further information about ILLC-publications, please contact

Institute for Logic, Language and Computation
Universiteit van Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam
phone: +31-20-525 6051
fax: +31-20-525 5206
e-mail: illc@science.uva.nl
homepage: http://www.illc.uva.nl/

Cryptography in a Quantum World

Academisch Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof.dr. D.C. van den Boom
ten overstaan van een door het college voor
promoties ingestelde commissie, in het openbaar
te verdedigen in de Agnietenkapel
op woensdag 27 februari 2008, te 14.00 uur

door

Stephanie Dorothea Christine Wehner

geboren te Würzburg, Duitsland.

Promotiecommissie:

Promotor:	prof.dr. H.M. Buhrman
Overige leden:	prof.dr.ir. F.A. Bais
	prof.dr. R.J.F. Cramer
	prof.dr. R.H. Dijkgraaf
	prof.dr. A.J. Winter
	dr. R.M. de Wolf

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

The investigations were supported by EU projects RESQ IST-2001-37559, QAP IST 015848 and the NWO vici project 2004-2009.

Cover design by Frans Bartels.
Printed and bound by PrintPartners Ipskamp.

ISBN: 90-6196-544-6

Parts of this thesis are based on material contained in the following papers:

- **Cryptography from noisy storage**
  S. Wehner, C. Schaffner, and B. Terhal
  Submitted
  (Chapter 11)

- **Higher entropic uncertainty relations for anti-commuting observables**
  S. Wehner and A. Winter
  Submitted
  (Chapter 4)

- **Security of Quantum Bit String Commitment depends on the information measure**
  H. Buhrman, M. Christandl, P. Hayden, H.K. Lo and S. Wehner
  In Physical Review Letters, 97, 250501 (2006)
  (long version submitted to Physical Review A)
  (Chapter 10)

- **State Discrimination with Post-Measurement Information**
  M. Ballester, S. Wehner and A. Winter
  To appear in IEEE Transactions on Information Theory
  (Chapter 3)

- **Entropic uncertainty relations and locking: tight bounds for mutually unbiased bases**
  M. Ballester and S. Wehner
  In Physical Review A, 75, 022319 (2007)
  (Chapters 4 and 5)

- **Tsirelson bounds for generalized CHSH inequalities**
  S. Wehner
  In Physical Review A, 73, 022110 (2006)
  (Chapter 7)

- **Entanglement in Interactive Proof Systems with Binary Answers**
  S. Wehner
  In Proceedings of STACS 2006, LNCS 3884, pages 162-171 (2006)
  (Chapter 9)

Other papers to which the author contributed during her time as a PhD student:

- **The quantum moment problem**
  A. Doherty, Y. Liang, B. Toner and S. Wehner
  Submitted

- **Security in the Bounded Quantum Storage Model**
  S. Wehner and J. Wullschleger
  Submitted

- **A simple family of non-additive codes**
  J.A. Smolin, G. Smith and S. Wehner
  In Physical Review Letters, 99, 130505 (2007)

- **Analyzing Worms and Network Traffic using Compression**
  S. Wehner
  Journal of Computer Security, Vol 15, Number 3, 303-320 (2007)

- **Implications of Superstrong Nonlocality for Cryptography**
  H. Buhrman, M. Christandl, F. Unger, S. Wehner and A. Winter
  In Proceedings of the Royal Society A, vol. 462 (2071), pages 1919-1932 (2006)

- **Quantum Anonymous Transmissions**
  M. Christandl and S. Wehner
  In Proceedings of ASIACRYPT 2005, LNCS 3788, pages 217-235 (2005)

C'est véritablement utile puisque c'est joli.
Le Petit Prince, Antoine de Saint-Exupéry

Contents

Acknowledgments **xv**

I Introduction 1

1 Quantum cryptography **3**
1.1 Introduction . . . 3
1.2 Setting the state . . . 5
1.2.1 Terminology . . . 5
1.2.2 Assumptions . . . 6
1.2.3 Quantum properties . . . 7
1.3 Primitives . . . 9
1.3.1 Bit commitment . . . 9
1.3.2 Secure function evaluation . . . 11
1.3.3 Secret sharing . . . 17
1.3.4 Anonymous transmissions . . . 18
1.3.5 Other protocols . . . 19
1.4 Challenges . . . 19
1.5 Conclusion . . . 20

II Information in quantum states 23

2 Introduction **25**
2.1 Quantum mechanics . . . 25
2.1.1 Quantum states . . . 25
2.1.2 Multipartite systems . . . 27
2.1.3 Quantum operations . . . 29
2.2 Distinguishability . . . 32

2.3 Information measures . . . 36
2.3.1 Classical . . . 36
2.3.2 Quantum . . . 37
2.4 Mutually unbiased bases . . . 39
2.4.1 Latin squares . . . 39
2.4.2 Generalized Pauli matrices . . . 41
2.5 Conclusion . . . 42

3 State discrimination with post-measurement information 43
3.1 Introduction . . . 43
3.1.1 Outline . . . 45
3.1.2 Related work . . . 46
3.2 Preliminaries . . . 47
3.2.1 Notation and tools . . . 47
3.2.2 Definitions . . . 47
3.2.3 A trivial bound: guessing the basis . . . 48
3.3 No post-measurement information . . . 49
3.3.1 Two simple examples . . . 49
3.3.2 An upper bound for all Boolean functions . . . 50
3.3.3 AND function . . . 50
3.3.4 XOR function . . . 51
3.4 Using post-measurement information . . . 54
3.4.1 A lower bound for balanced functions . . . 54
3.4.2 Optimal bounds for the AND and XOR function . . . 57
3.5 Using post-measurement information and quantum memory . . . 63
3.5.1 An algebraic framework for perfect prediction . . . 63
3.5.2 Using two bases . . . 66
3.5.3 Using three bases . . . 70
3.6 Conclusion . . . 72

4 Uncertainty relations 75
4.1 Introduction . . . 75
4.2 Limitations of mutually unbiased bases . . . 78
4.2.1 MUBs in square dimensions . . . 79
4.2.2 MUBs based on Latin squares . . . 80
4.2.3 Using a full set of MUBs . . . 80
4.3 Good uncertainty relations . . . 83
4.3.1 Preliminaries . . . 84
4.3.2 A meta-uncertainty relation . . . 89
4.3.3 Entropic uncertainty relations . . . 89
4.4 Conclusion . . . 91

5 Locking classical information **93**
5.1 Introduction . . . 93
5.1.1 A locking protocol . . . 94
5.1.2 Locking and uncertainty relations . . . 95
5.2 Locking using mutually unbiased bases . . . 96
5.2.1 An example . . . 96
5.2.2 MUBs from generalized Pauli matrices . . . 99
5.2.3 MUBs from Latin squares . . . 101
5.3 Conclusion . . . 101

III Entanglement **103**

6 Introduction **105**
6.1 Introduction . . . 105
6.1.1 Bell's inequality . . . 106
6.1.2 Tsirelson's bound . . . 108
6.2 Setting the stage . . . 109
6.2.1 Entangled states . . . 109
6.2.2 Other Bell inequalities . . . 110
6.2.3 Non-local games . . . 110
6.3 Observations . . . 113
6.3.1 Simple structural observations . . . 113
6.3.2 Vectorizing measurements . . . 115
6.4 The use of post-measurement information . . . 116
6.5 Conclusion . . . 119

7 Finding optimal quantum strategies **121**
7.1 Introduction . . . 121
7.2 A simple example: Tsirelson's bound . . . 123
7.3 The generalized CHSH inequality . . . 125
7.4 General approach and its applications . . . 128
7.4.1 General approach . . . 128
7.4.2 Applications . . . 129
7.5 Conclusion . . . 130

8 Bounding entanglement in NL-games **131**
8.1 Introduction . . . 131
8.2 Preliminaries . . . 132
8.2.1 Random access codes . . . 132
8.2.2 Non-local games and state discrimination . . . 134
8.3 A lower bound . . . 134
8.4 Upper bounds . . . 136

8.5 Conclusion . . . 138

9 Interactive Proof Systems **139**
9.1 Introduction . . . 139
9.1.1 Classical interactive proof systems . . . 139
9.1.2 Quantum multi-prover interactive proof systems . . . 140
9.2 Proof systems and non-local games . . . 142
9.2.1 Non-local games . . . 142
9.2.2 Multiple classical provers . . . 143
9.2.3 A single quantum prover . . . 145
9.3 Simulating two classical provers with one quantum prover . . . 145
9.4 Conclusion . . . 148

IV Consequences for Crytography **149**

10 Limitations **151**
10.1 Introduction . . . 151
10.2 Preliminaries . . . 152
10.2.1 Definitions . . . 152
10.2.2 Model . . . 153
10.2.3 Tools . . . 154
10.3 Impossibility of quantum string commitments . . . 156
10.4 Possibility . . . 159
10.5 Conclusion . . . 161

11 Possibilities: Exploiting storage errors **163**
11.1 Introduction . . . 163
11.1.1 Related work . . . 165
11.2 Preliminaries . . . 165
11.2.1 Definitions . . . 165
11.3 Protocol and analysis . . . 170
11.3.1 Protocol . . . 170
11.3.2 Analysis . . . 170
11.4 Practical oblivious transfer . . . 171
11.5 Example: depolarizing noise . . . 174
11.5.1 Optimal cheating strategy . . . 175
11.5.2 Noise tradeoff . . . 183
11.6 Conclusion . . . 185

Appendix

A Linear algebra and semidefinite programming **187**
A.1 Linear algebra prerequisites 187
A.2 Definitions 189
A.3 Semidefinite programming 190
A.4 Applications 191

B C^*-Algebra **193**
B.1 Introduction 193
B.2 Some terminology 194
B.3 Observables, states and representations 195
B.3.1 Observables and states 195
B.3.2 Representations 196
B.4 Commuting operators 198
B.4.1 Decompositions 199
B.4.2 Bipartite structure 200
B.4.3 Invariant observables and states 202
B.5 Conclusion 203

C Clifford Algebra **205**
C.1 Introduction 205
C.2 Geometrical interpretation 206
C.2.1 Inner and outer product 206
C.2.2 Reflections 207
C.2.3 Rotations 208
C.3 Application 212
C.4 Conclusion 215

Bibliography **217**

Index **241**

Symbols **249**

Samenvatting **251**

Summary **255**

Acknowledgments

Research has been an extremely enjoyable experience for me, and I had the opportunity to learn many exciting new things. However, none of this would have been possible without the help and support of many people.

First, I would like to thank my supervisor Harry Buhrman for our interesting discussions and for giving me the opportunity to be at CWI which is a truly great place to work. For the freedom to pursue my own interests, I am deeply grateful. My time as a PhD student would have been very different without Andreas Winter, and I would especially like to thank him for our many enjoyable discussions and conversations. I have learned about many interesting things from him, ranging from the beautiful topic of algebras, that I discovered way too late, to his way of taking notes which I have shamelessly adopted. I would also like to thank him for much encouragement, without which I may not have dared to pursue my ideas about uncertainty relations much further. Much of Chapter 4.3 is owed to him. I would also like thank him, as well as Sander Bais, Ronald Cramer, Robbert Dijkgraaf, and Ronald de Wolf for taking part in my PhD committee.

Thanks also to Ronald de Wolf for supervising my Master's thesis, which was of tremendous help to me during my time as a PhD student. Furthermore, I would like to thank Matthias Christandl for our fun collaborations, a great trip to Copenhagen, and the many enjoyable visits to Cambridge. Thanks also to Artur Ekert for making these visits possible, and for the very nice visit to Singapore. I am very grateful for his persistent encouragement, and his advice on giving talks is still extremely helpful to me. For many interestings dicussions and insights I would furthermore like to thank Serge Fehr, Julia Kempe, Iordanis Kerenidis, Oded Regev, Renato Renner and Pranab Sen, as well as my collaborators Manuel Ballester, Harry Buhrman, Matthias Christandl, Andrew Doherty, Patrick Hayden, Hoi-Kwong Lo, Christian Schaffner, Graeme Smith, John Smolin, Barbara Terhal, Ben Toner, Falk Unger, Andreas Winter, Ronald de Wolf, and Jürg Wullschleger. Thanks also to Nebošja Gvozdenović, Dennis Hofheinz, Monique Laurent, Serge Massar and Frank Vallentin for useful point-

ers, and to Boris Tsirelson for supplying me with copies of [Tsi80] and [Tsi87]. Many thanks also to Tim van Erven, Peter Grünwald, Peter Harremoes, Steven de Rooij, and Nitin Saxena for the enjoyable time at CWI, and to Paul Vitányi who let me keep his comfy armchair on which many problems were solved.

Fortunately, I was able to visit many other places during my time as a PhD student. I am grateful to Dorit Aharanov, Claude Crépeau, Artur Ekert, Julia Kempe, Iordanis Kerenidis, Michele Mosca, Michael Nielsen, David Poulin, John Preskill, Barbara Terhal, Oded Regev, Andreas Winter and Andrew Yao for their generous invitations. For making my visits to England and Australia so enjoyable, I would furthermore like to thank Almut Beige, Agata Branczyk, Matthias Christandl, Andrew Doherty, Marie Ericsson, Alistair Kay, Jiannis Pachos, Peter Rohde and Andreas Winter.

Thanks to Manuel Ballester, Cor Bosman, Serge Fehr, Sandor Héman, Oded Regev, Peter Rohde, and especially Christian Schaffner for many helpful comments on this thesis; any remaining errors are of course my own responsibility. Thanks also to Frans Bartels for drawing the thesis cover and the illustrations of Alice and Bob. I am still grateful to Torsten Grust and Peter Honeyman who encouraged me to go to university in the first place.

Finally, many thanks to my family and friends for being who they are.

Amsterdam — Stephanie Wehner
February, 2008.

Part I

Introduction

Chapter 1

Quantum cryptography

Cryptography is the art of secrecy. Nearly as old as the art of writing itself, it concerns itself with one of the most fundamental problems faced by any society whose success crucially depends on knowledge and information: With whom do we want to share information, and when, and how much?

1.1 Introduction

Starting with the first known encrypted texts from 1900 BC in Egypt [Wik], cryptography has a fascinating history [Kah96]. Its goal is simple: to protect secrets as best as is physically possible. Following our increased understanding of physical processes with the advent of quantum mechanics, Wiesner [Wie83] proposed using quantum techniques for cryptography in the early 1970's. Unfortunately, his groundbreaking work, which contained the seed for quantum key distribution, oblivious transfer (as described below), and a form of quantum money, was initially met with rejection [Bra05]. In 1982, Bennett, Brassard, Breitbart and Wiesner joined forces to publish "Quantum cryptography, or unforgeable subway tokens" which luckily found acceptance [BBBW82], leading to the by now vast field of research in quantum key distribution (QKD). Quantum key distribution allows two remote parties who are only connected via a quantum channel to generate an arbitrarily long secret key that they can then use to perfectly shield their messages from prying eyes. The idea is beautiful in its simplicity: unlike with classical data, quantum mechanics prevents us from copying an unknown quantum state. What's more is that any attempt to extract information from such a state can be detected! That is, we can now determine whether an eavesdropper has been trying to intercept our secrets. Possibly the most famous QKD protocol known to date was proposed in 1983 by Bennett and Brassard [BB83], and is more commonly known as BB84 from its 1984 full publication [BB84]. Indeed, many quantum cryptographic protocols to date are inspired in some fashion by BB84. It saw its first experimental implementation in 1989, when Bennett, Bessette,

Brassard, Salvail and Smolin built the first QKD setup covering a staggering distance of 32.5 cm [BB89, BBB$^+$92]! In 1991, Ekert proposed a beautiful alternative view of QKD based on quantum entanglement and the violation of Bell's theorem, leading to the protocol now known as E91 [Eke91]. His work paved the way to establishing the security of QKD protocols, and led to many other interesting tasks such as entanglement distillation. Since then, many other protocols such as B92 [Ben92] have been suggested. Today, QKD and its related problems form a well-established part of quantum information, with countless proposals and experimental implementations. It especially saw increased interest after the discovery of Shor's quantum factoring algorithm in 1994 [Sho97] that renders almost all known classical encryption systems insecure, once a quantum computer is built. Some of the first security proofs were provided by Mayers [May96a], Lo and Chau [LC99], and Shor and Preskill [SP00], finally culminating in the wonderful work of Renner [Ren05] who supplied the most general framework for proving the security of any known QKD protocol. QKD systems are already available commercially today [Qua, Tec]. The best known experimental implementations now cover distances of up to 148.7 km in optical fiber [HRP$^+$06], and 144 km in free space [UTSM$^+$] in an experiment conducted between two Canary islands.

Figure 1.1: Encrypted pottery glaze formula, Mesopotamia 1500 BC

Figure 1.2: QKD today

Traditional cryptography is concerned with the secure and reliable transmission of messages. With the advent of widespread electronic communication, however, new cryptographic tasks have become increasingly important. We would like to construct secure protocols for electronic voting, online auctions, contract signing and many other applications where the protocol participants themselves do not trust each other. Two primitives that can be used to construct all such protocols are bit commitment and oblivious transfer. We will introduce both primitives in detail below. Interestingly, it turns out that despite many initially suggested protocols [BBBW82, Cré94], both primitives are impossible to achieve when we ask for unconditional security. Luckily, as we will see in Chapter 11

we can still implement both building blocks if we assume that our quantum operations are affected by noise. Here, the very problem that prevents us from implementing a full-scale quantum computer can be turned to our advantage.

In this chapter, we give an informal introduction to cryptography in the quantum setting. We first introduce necessary terminology, before giving an overview over the most well-known cryptographic primitives. Since our goal is to give an overview, we will restrict ourselves to informal definitions. Surprisingly, even definitions themselves turn out to be a tricky undertaking, especially when entering the quantum realm. Finally, we discuss what makes the quantum setting so different from the classical one, and identify a range of open problems.

1.2 Setting the state

1.2.1 Terminology

In this text, we consider protocols among multiple participants $P_1, \ldots, P_n$, also called *players*. When considering only two players, we generally identify them with the protagonists Alice and Bob. Each player may hold a *private input*, that is classical and quantum data unknown to the other players. In addition, the players may have access to a shared resource such as classical shared randomness or quantum entanglement that has been distributed before the start of the protocol. We will refer to any information that is available to all players as *public*. A subset of players may also have access to shared information that is known only to them, but not to the remaining players. Such an input is called *private shared input*. In the case of shared randomness, this is also known as *private shared randomness*. The players can be connected by classical as well as quantum channels, and use them to exchange messages during the course of the protocol. A given protocol consists of a set of messages as well as a specification of actions to be undertaken by the players. At the end of the protocol, each player may have a classical as well as a quantum output.

A player is called *honest*, if he follows the protocol exactly as dictated. He is called *honest-but-curious*, if he follows the protocol, but nevertheless tries to gain additional information by processing the information supplied by the protocol in a way which is not intended by the protocol. An honest player, for example, will simply ignore parts of the information he is given, as he will do exactly as he is told. However, a player that is honest-but-curious will take advantage of all information he is given, i.e., he may read and copy all messages as desired, and never forgets any information he is given.[1] Yet, the execution of the protocol itself is unaffected as the player does not change any information used in the protocol,

[1]Note that since an honest-but-curious player never forgets any information, he effectively makes a copy of all messages. He will erase his memory needed for the execution of the protocol if dictated by the protocol: his copy lies outside this memory.

he merely reads it. But what does this mean in a quantum setting? Indeed, this question appears to be a frequent point of debate. We will see in Chapter 2 that he cannot copy arbitrary quantum information, and extracting non-classical information from a quantum state will necessarily lead to disturbance. Evidently, disturbance alters the quantum states during the protocol. Hence, the player actually took actions to alter the execution of the protocol, and we can no longer regard him as honest. After examining quantum operations in Chapter 2 we will return to the definition of an honest-but-curious player in the quantum setting. Finally, a player can also be *dishonest*: he will do anything in his power to break the protocol. Evidently, this is the most realistic setting, and we will always consider it here.

An *adversary* is someone who is trying to break the protocol. An adversary is generally modeled as an entity outside of the protocol that can either be an eavesdropper, or take part in the protocol by taking control of specific players. This makes it easier to model protocols among multiple players, where we assume that all dishonest players collaborate to form a single adversary.

1.2.2 Assumptions

In an ideal world, we could implement any cryptographic protocol described below. Interestingly though, even in the quantum world we encounter physical limits which prevent us from doing so with unconditional security. *Unconditional security* most closely corresponds to the intuitive notion of "secure". A protocol that is unconditionally secure fulfills its purpose and is secure even if an attacker is granted unlimited resources. We happily provide him with the most powerful computer we could imagine and as much memory space as he wants. The main question of unconditional security is thus whether the attacker obtains enough information to defeat the security of the system. Unconditional security is also called *perfect secrecy* in the context of encryption systems, and forms part of *information-theoretic security*.

Most often, however, unconditional security can never be achieved. We must therefore resign ourselves to introducing additional limitations on the adversary: the protocol will only be secure if certain assumptions hold. In practise, these assumptions can be divided into two big categories: In the first, we assume that the players have access to a common resource with special properties. This includes models such as a trusted initializer [Riv99], or another source that provides the players with shared randomness drawn from a fixed distribution. An example of this is also a noisy channel [CK88]: Curiously, a noisy channel that neither player can influence too much turns out to be an incredibly powerful resource. The second category consists of clear limitations on the ability of the adversary. For example, the adversary may have limited storage space available [Mau92, DFSS05], or experience noise when trying to store qubits as we will see in Chapter 11. In multi-player protocols we can also demand that dishonest players cannot commu-

nicate during the course of the protocol, that messages between different players take a certain time to be transmitted, or that only a minority of the players is dishonest. In the quantum case, other known assumptions include limiting the adversary to measure not more than a certain number of qubits at a time [Sal98], or introducing superselection rules [KMP04], where the adversary can only make a limited set of quantum measurements. When introducing such assumptions, we still speak of *information-theoretic security*: Except for these limitations, the adversary remains all-powerful. In particular, he has unlimited computational resources.

Classically, most forms of practical cryptography are shown to be *computationally secure*. In this security model, we do not grant an adversary unlimited computational resources. Instead, we are concerned with the amount of computation required to break the security of a system. We say that a system is *computationally secure*, if the believed level of computation necessary to defeat it exceeds the computational resources of any hypothetical adversary by a comfortable margin. The adversary is thereby allowed to use the best possible attacks against the system. Generally, the adversary is modeled as having only polynomial computational power. This means that any attacks are restricted to time and space polynomial in the size of the underlying security parameters of the system. In this setting the difficulty of defeating the system's security is often proven to be as difficult as solving a well-known problem which is believed to be hard. The most popular problems are often number-theoretic problems such as factoring. Note that for example in the case of factoring, it is not known whether these problems are truly difficult to solve classically. Many such problems, such as factoring, fold with the advent of a quantum computer [Sho97]. It is an interesting open problem to find classical hardness assumptions, which are still secure given a quantum computer. Several proposals are known [Reg03], but so far none of them have been proven secure.

In the realm of quantum cryptography, we are so far only interested in information-theoretic security: we may introduce limitations on the adversary, but we do not resort to computational hardness assumptions.

1.2.3 Quantum properties

Quantum mechanics introduces several exciting aspects to the realm of cryptography, which we can exploit to our benefit, but which also introduce additional complications even in existing classical primitives whose security does not depend on computational hardness assumptions. Here, we give a brief introduction to some of the most striking aspects, which we will explain in detail later on.

1. **Quantum states cannot be copied:** In classical protocols, an adversary can always copy any messages and his classical data at will. Quantum states, however, differ: We will see in Chapter 2 that we cannot copy an

arbitrary qubit. This property led to the construction of the unforgeable subway tokens [BBBW82] mentioned earlier.

2. **Information gain can be detected:** Classically there is no way for an honest player to determine whether messages have been read maliciously outside the scope of the protocol. However, in a quantum setting we can detect whether an adversary tried to extract information from a transmitted message. This property forms the heart of quantum key distribution described below. It also allows us to construct *cheat-sensitive* protocols, a concept which is foreign to classical cryptography: even though we cannot prevent an adversary from gaining information if he intends to do so, we will be able to detect such cheating and take appropriate action. We will return to this aspect in Chapter 2.

3. **Uncertainty relations exist:** Unlike in the classical world, quantum states allow us to encode multiple bits into a single state in such a way that we cannot extract all of them simultaneously. This property is closely related to cheat-sensitivity, and is a consequence of the existence of uncertainty relations we will encounter in Chapter 4. It is also closely related to what is known as quantum random access codes, which will we employ in Chapter 8.

4. **Information can be "locked":** Another aspect we need to take into account when considering quantum protocols is an effect known as locking classical information in quantum states. Surprisingly, the amount of correlation between two parties can increase by much more than the data transmitted. We will examine this effect for a specific measure of correlation in more detail in Chapter 5.

5. **Entanglement allows for stronger correlations:** Entanglement is another concept absent from the classical realm. Whereas entanglement has many useful applications such as quantum teleportation and can also be used to analyze the security of quantum key distribution, it also requires us to be more cautious: In Chapter 9, we will see that the parameters of classical protocols can change dramatically if dishonest players share entanglement, even if they do not have access to a full quantum computer. In Chapter 10, entanglement will enable an adversary to break any quantum string commitment protocol.

6. **Measurements can be delayed:** Finally, we encounter an additional obstacle, which is also entirely missing from classical protocols: Players may delay quantum measurements. In any classical protocol, we can be assured that any input and output is fixed once the protocol ends. In the quantum case, however, players may alter their protocol input retroactively by

delaying quantum measurements that depend on their respective inputs. Essentially, in a classical protocol the players will automatically be "committed" to the run of the protocol, whereas in the quantum setting this property is entirely missing. This can make an important difference in reductions among several protocols as we will see in Section 1.3.2 below.

1.3 Primitives

We now present an overview of the most common multi-party protocol primitives, and what is known about them in the quantum setting. We already encountered quantum key distribution (QKD) in the introduction. In this thesis, our focus lies on cryptographic protocols *other* than QKD.

1.3.1 Bit commitment

Possibly the most active area of quantum cryptography in the early stages next to QKD was quantum bit commitment: Imagine two mutually distrustful parties Alice and Bob at distant locations. They can only communicate over a channel, but want to play the following game: Alice secretly chooses a bit c. Bob wants to be sure that Alice indeed has made her choice. Yet, Alice wants to keep c hidden from Bob until she decides to reveal c. To convince Bob that she made up her mind, Alice sends Bob a commitment. From the commitment alone, Bob cannot deduce c. At a later time, Alice reveals c and enables Bob to open the commitment. Bob can now check if Alice is telling the truth. This scenario is known as *bit commitment*. Commitments play a central role in modern-day cryptography.

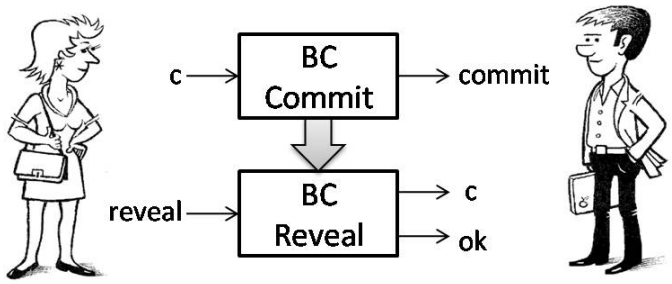


Figure 1.3: Schematic run of a BC protocol when Alice and Bob are honest.

They form an important building block in the construction of larger protocols in, for example, gambling and electronic voting, and other instances of secure two-party computation. In the realm of quantum mechanics, it has been shown that oblivious transfer [BBCS92b] (defined in Section 1.3.2) can be achieved provided there exists a secure bit commitment scheme [Yao95, Cré94]. In turn,

classical oblivious transfer can be used to perform any secure two-party computation defined below [CvdGT95]. Commitments are also useful for constructing zero-knowledge proofs [Gol01] and lead to coin tossing [Blu83]. Informally, bit commitment can be defined as follows:

1.3.1. Definition. *Bit commitment* (BC) is a two-party protocol between Alice (the committer) and Bob (the verifier), which consists of three stages, the committing and the revealing stage, and a final declaration stage in which Bob declares "accept" or "reject". The following requirements should hold:

- (Correctness) If both Alice and Bob are honest, then before the committing stage Alice picks a bit c. Alice's protocol depends on c and any randomness used. At the revealing stage, Alice reveals to Bob the committed bit c. Bob accepts.
- (Binding) If Alice wants to reveal a bit c', then

$$\Pr[\text{Bob accepts } |c' = 0] + \Pr[\text{Bob accepts } |c' = 1] \leq 1.$$

- (Concealing) If Alice is honest, Bob does not learn anything about c before the revealing stage.

Classically, unconditionally secure bit commitment is known to be impossible. Indeed, this is very intuitive if we consider the implications of the concealing condition: This condition implies that exactly the same information exchange must have occurred if Alice committed herself to $c = 0$ or $c = 1$, otherwise Bob would be able to gain information about c. But this means that even if Alice initially made a commitment to $c = 0$, she can later reconstruct the run of the protocol as if she had committed herself to $c = 1$ and thus send the right message to Bob to reveal $c = 1$ instead. Unfortunately, even quantum communication cannot help us to implement unconditionally secure bit commitment without further assumptions: After several quantum schemes were suggested [BB84, BC90a, BCJL93], quantum bit commitment was shown to be impossible, too [May96b, LC97, May97, LC96, BCMS97, CL98, DKSW06], even in the presence of superselection rules [KMP04], where the adversary can only perform a certain restricted set of measurements. In the face of the negative results, what can we still hope to achieve?

Evidently, we need to assume that the adversary is limited in certain ways. In the classical case, bit commitment is possible if the adversary is *computationally bounded* [Gol01], if *one-way functions exist* [Nao91, HR07], if Alice and Bob are connected via a *noisy channel* that neither player can influence too much [CK88, DKS99, DFMS04], or if the adversary is bounded in *space* instead of time, i.e., he is only allowed to use a certain amount of storage space [Mau92]. Unfortunately, the security of the bounded classical storage model [Mau92, CCM98] is somewhat

unsatisfactory: First, a dishonest player needs only quadratically more memory than the honest one to break the security. Second, as classical memory is very cheap, most of these protocols require huge amounts of communication in order to achieve reasonable bounds on the adversaries memory.

Do we gain anything by using quantum communication? Interestingly, even without any further assumptions, quantum cryptography at least allows us to implement imperfect forms of bit commitment, where Alice and Bob both have a limited ability to cheat. That is, we allow Alice to change her mind, and Bob to learn the committed bit with a small probability. These protocols are based on the fact that quantum protocols can exhibit a form of cheat sensitivity unavailable to classical communication [HK04, ATSVY00]. Exact tradeoffs on how well we can implement bit commitment in the quantum world can be found in [SR02a]. Protocols that make use of this tradeoff are cheat-sensitive, as described in Section 1.2.2. Examples of such protocols have been used to implement coin tossing [Amb01] as described in Section 1.3.2. In Chapter 10, we will consider commitments to an entire string of bits at once. Whereas this task turns out to be impossible as well for a strong security definition, we will see that non-trivial quantum protocols do exist for a very weak security definition. Bit commitment can also be implemented under the assumption that faster than light communication is impossible, provided that Alice and Bob are located very far apart [Ken99], or if Alice and Bob are given access to non-local boxes [BCU$^+$06] which provide superstrong non-local correlations.

But even a perfect commitment can be implemented, if we make quantum specific assumptions. For example, it is possible to securely implement BC provided that an adversary cannot measure more than a fixed number of qubits simultaneously [Sal98]. With current-day technology, it is *very* difficult to store states even for a very short period of time. This leads to the protocol presented in [BBCS92a, Cré94], which shows how to implement BC and OT (defined below) if the adversary is not able to store *any* qubits at all. In [DFSS05, DFR$^+$07], these ideas have been generalized in a very nice way to the *bounded-quantum-storage model*, where the adversary is computationally unbounded and is allowed to have an unlimited amount of *classical* memory. However, he is only allowed a limited amount of *quantum* memory. The advantages over the classical bounded-storage model are two-fold: First, given current day technology it is indeed very hard to store quantum states. Secondly, the honest players do not require any quantum storage at all, making the protocol much more efficient. It has been shown that such protocols remain secure when executed many times in a row [WW07].

1.3.2 Secure function evaluation

An important aspect of modern day cryptography is the primitive known as secure function evaluation, and its multi-player analogue, secure multi-party computation, first suggested by Yao [Yao82]. Imagine that Alice and Bob are trying to

decide whether to attend an unpopular administrative event. If Alice attends, Bob feels forced to attend as well and vice versa. However, neither of them wants to announce publicly whether they are planning to attend or whether they would rather make up an excuse to remain at home, as this may have dire consequences. How can Alice and Bob solve their dilemma? Note that their problem can be phrased in the following form: Let x be Alice's private input bit, where $x = 1$ if Alice is planning to attend and $x = 0$ if Alice skips the event. Similarly, let y be Bob's private input bit. Alice and Bob now want to compute $\mathrm{OR}(x, y)$ in such a way that both of them learn the result, but neither of them learns anything more about the input of the other player than can be inferred from the result. In our example, if $\mathrm{OR}(x, y) = 1$, at least one of the players is planning to attend the event. Both Alice and Bob now attend the event, and both of them can safely claim that they really did plan to do so in the first place. If $\mathrm{OR}(x, y) = 0$, Alice and Bob learn that they both agree, and do not need to fear any political consequences.

Secure function evaluation enables Alice and Bob to solve any such task. Protocols for secure function evaluation enable us to construct protocols for electronic voting and secure auctions. Informally, we define:

1.3.2. Definition. *Secure function evaluation* (SFE) is a two-party protocol between Alice and Bob, where Alice holds a private input x and Bob holds a private input y such that

- (Correctness) If both Alice and Bob are honest, then they both output the same value $v = f(x, y)$.
- (Security) If Alice (Bob) is dishonest, then Alice (Bob) does not learn more about x (y) then can be inferred from $f(x, y)$.

A common variant of SFE is so-called *one-sided* SFE: Here, only one of the two players receives the result of the computation, $f(x, y)$. Sadly, we cannot implement SFE for an arbitrary function f classically without additional assumptions, akin to bit commitment. Even in the quantum world, the situation is equally bleak: SFE remains impossible in the quantum setting [Lo97]! Fortunately, the situations improves when we consider multi-party protocols as mentioned below.

Oblivious transfer

A special case of secure function evaluation is the problem of oblivious transfer, which was first introduced by Rabin [Rab81]. The variant of 1-2 OT appeared in a paper by Even, Goldreich and Lempel [EGL85] and also, under a different name, in the well-known paper by Wiesner [Wie83]. 1-2 OT allows Alice and Bob to solve a seemingly uninteresting problem: The sender (Alice) secretly chooses two bits s_0 and s_1, the receiver (Bob) secretly chooses a bit c. The primitive of oblivious transfer allows Bob to retrieve s_c in such a way, that Alice cannot gain any

information about c. At the same time, Alice is ensured that Bob only retrieves s_c and gets no information about the other input bit $s_{\bar{c}}$. Oblivious transfer can be used to perform any secure two-party computation [Kil88, CvdGT95], and is therefore a very important primitive.

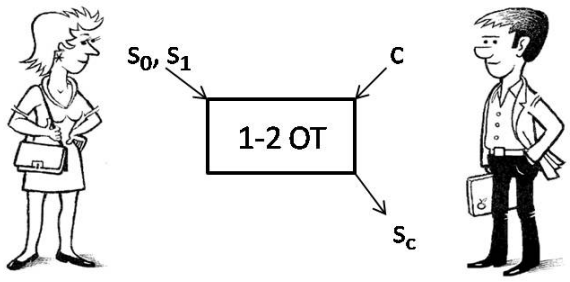


Figure 1.4: Schematic run of a 1-2 OT protocol.

Unlike in the classical setting, oblivious transfer in the quantum world requires additional caution: We want that after the protocol ends, both of Alice's inputs bits s_0, s_1 and Bob's choice bit c have been determined. That is, they are fixed and the players can no longer change their mind. In particular, we do not want Bob to delay his choice of c indefinitely, possibly by delaying a quantum measurement. Similarly, Alice should not be able to change her mind about, for example, the parity of $s_0 \oplus s_1$ after the end of the protocol by delaying a measurement. Informally, we define

1.3.3. Definition. $\binom{2}{1}$-*oblivious transfer* (1-2 OT$(s_0, s_1)(c)$) is a two-party protocol between Alice (the sender) and Bob (the receiver), such that

- (Correctness) If both Alice and Bob are honest, the protocol depends on Alice's two input bits $s_0, s_1 \in \{0, 1\}$ and Bob's input bit $c \in \{0, 1\}$. At the end of the protocol Bob knows s_c.
- (Security against Alice) If Bob is honest, Alice does not learn c.
- (Security against Bob) If Alice is honest, Bob does not learn anything about $s_{\bar{c}}$.

After the protocol ends, s_0, s_1 and c have been chosen.

Classically, 1-2 OT can be obtained from the following simpler primitive, also known as Rabin-OT [Rab81] or erasure channel. Conversely, OT can be obtained from 1-2 OT.

1.3.4. Definition. Rabin *Oblivious transfer* (Rabin-OT) is a two-party protocol between Alice (the sender) and Bob (the receiver), such that

- (Correctness) If both Alice and Bob are honest, the protocol depends on Alice's input bit $b \in \{0,1\}$. At the end of the protocol, Bob obtains b with probability $1/2$ and knows whether he obtained b or not.
- (Security against Alice) If Bob is honest, Alice does not learn whether Bob obtained b.
- (Security against Bob) If Alice is honest, Bob's probability of learning bit b does not exceed $1/2$.

After the protocol ends, b has been chosen.

The fact that Alice and Bob may delay their measurements makes an important difference, as the following simple example shows: Consider the standard reduction of Rabin-OT to 1-2 OT: Alice uses inputs $s_k = b$ and $s_{\bar{k}} = 0$ with $k \in_R \{0,1\}$. Bob uses input $c \in_R \{0,1\}$, for a randomly chosen c. The players now perform 1-2 OT$(s_0, s_1)(c)$ after which the receiver holds s_c. Subsequently, Alice announces k. If $k = c$, Bob succeeded in retrieving b and otherwise he learns nothing. This happens with probability $p = 1/2$ and thus we have constructed Rabin-OT from one instance of 1-2 OT. Clearly, this reduction fails if we use an 1-2 OT protocol in which Bob can defer his choice of c, possibly by delaying a quantum measurement that depends on c. He simply waits until Alice announces k, to retrieve s_k with certainty. This simple example makes it clear that implementing 1-2 OT is far from a trivial task in the quantum setting. Even the classical definitions need to be revised carefully. In this brief overview, we restricted ourselves to the informal definition given above, and refer to [Wul07] for an extensive treatment of the definition of oblivious transfer.

Note that oblivious transfer forms an instance of secure function evaluation with $f : \{0,1\}^2 \times \{0,1\} \to \{0,1\}$ satisfying $f(s_0, s_1, c) = s_c$, where only one player (Bob) learns the output. Hence by Lo's impossibility result for SFE discussed earlier, oblivious transfer is not possible in the quantum setting either without introducing additional assumptions. Indeed, note that there exists a classical reduction of bit commitment to oblivious transfer (up to a vanishing probability), where we reverse the roles of Alice and Bob for bit commitment: Alice simply chooses two n-bit strings $x_0 \in_R \{0,1\}^n$, and $x_1 \in_R \{0,1\}^n$. Alice and Bob now use n rounds of 1-2 OT, where Bob retrieves x_c when he wants to commit to a bit c. To reveal, he then sends c and x_c to Alice. Intuitively, one can thus hope to use the impossibility proof of bit commitment to show that oblivious transfer is impossible as well, without resorting to [Lo97]. However, note that we would first have to show the security of this reduction with respect to a quantum adversary. Fortunately, oblivious transfer becomes possible if we make the same assumptions as for bit commitment described in Section 1.3.1. We will consider how to implement oblivious transfer if the adversary's quantum storage is subject to noise in Chapter 11.

Coin tossing

Another example of SFE is the well-known primitive of coin tossing [Blu83], which can be viewed as an instance of randomized secure function evaluation defined in [Gol01]. Imagine that Alice and Bob want to toss a coin, solely by communicating over a classical and a quantum channel. We thereby want to ensure that neither party can influence the outcome of the coin toss by too much. Unfortunately, we cannot implement this primitive classically without relying on additional assumptions.

What assumptions do we need to implement coin tossing? It is easy to see that we can implement one form of coin tossing, if we could perform bit commitment: Alice chooses a random bit $b \in_R \{0,1\}$ and commits herself to b. Subsequently, Bob chooses a random bit $b' \in_R \{0,1\}$ and sends it to Alice. After receiving b', Alice opens her commitment and reveals b. Both parties now output $c = b \oplus b'$ as their outcome. Thus, any assumptions that enable us to implement bit commitment also lead to coin tossing. Some assumptions even allow for very simple protocols: If we assume that Alice and Bob are located far apart and faster-than-light communication is impossible, they can simply both flip a coin themselves and send it over the channel. They then take the xor of the two bits as the outcome of the coin flip. If Alice and Bob do not receive the other's bit within a certain time frame they reject this execution of the protocol and restart. Since it takes the bit a specific time to travel over the channel, both parties can be sure that it must have been sent before a certain time, i.e., before receiving the other's bit.

Many definitions of coin tossing are known in the literature, which exhibit subtle differences especially whether aborts are allowed during the protocol. In the quantum literature, *strong coin tossing*[2] has been informally defined as follows:

1.3.5. Definition. A quantum *strong coin tossing* protocol with bias ε is a two-party protocol, where Alice and Bob communicate and finally decide on a value $c \in \{0, 1, \perp\}$ such that

- If both parties are honest, then $\Pr[c = 0] = \Pr[c = 1] = 1/2$.
- If one party is honest, then for any strategy of the dishonest player $\Pr[c = 0] \leq 1/2 + \varepsilon$ and $\Pr[c = 1] \leq 1/2 + \varepsilon$.

Sadly, strong coin tossing cannot be implemented perfectly with bias $\varepsilon = 0$ [LC98]. However, one might hope that one could still achieve an arbitrarily small bias $\varepsilon > 0$. Many protocols have been proposed for quantum strong coin tossing and subsequently been broken [MSC99, ZLG00]. Sadly, it was shown that strong coin tossing cannot be implemented with an arbitrarily small bias, and $\varepsilon = 1/\sqrt{2} - 1/2 \approx 0.207$ is the best we could hope to achieve [Kit02]. So far,

[2]Unfortunately, these names carry a slightly different meaning in the classical literature.

quantum protocols for strong coin tossing with a bias of $\varepsilon \approx 0.42$ [ATSVY00] and finally $\varepsilon = 1/4$ [Amb01, SR02a, KN04, Col07] are known. No formal definition of strong coin tossing in the quantum setting is known to date, that specifies how to deal with an abort in the case when the protocol is executed multiple times.

To circumvent this problem, a slightly weaker primitive has been proposed, which carries the name *weak coin tossing* in the quantum literature. Here, we explicitly allow the dishonest party to bias the coin entirely in one direction, but limit his ability to bias the coin the other way. This scenario corresponds to a setting where, for example, Alice wins if the outcome is $c = 0$ and Bob if $c = 1$. However, we do allow each player to give in and loose at will. Intuitively, this setting makes more sense in all common practical examples when considering a standalone run of such a protocol, where each player has a preferred outcome. Informally, we define

1.3.6. Definition. A quantum *weak coin tossing* protocol with bias ε is a two-party protocol, where Alice and Bob communicate and finally decide on a value $c \in \{0, 1, \perp\}$ such that

- If both parties are honest, then $\Pr[c = 0] = \Pr[c = 1] = 1/2$.
- If Alice is honest, then for any strategy of Bob

$$\Pr[c = 1] \leq 1/2 + \varepsilon.$$

- If Bob is honest, then for any strategy of Alice

$$\Pr[c = 0] \leq 1/2 + \varepsilon.$$

Weakening the definition in this way indeed helps us! It has been shown that we can construct a quantum protocol for weak coin tossing that achieves a bias of $\varepsilon \approx 0.239$ [KN04], $\varepsilon \approx 0.207$ [SR02b], $\varepsilon \approx 0.192$ [Moc04], and $\varepsilon \approx 0.167$ [Moc05]. Very recently, however, a protocol with an arbitrarily small bias has been suggested [Moc07b]! To date, there is also no formal definition of weak coin tossing in the quantum setting.

Multiple players

Secure multi-party computation (SMP) concerns an analogous task to SFE, involving n players $P_1, \ldots, P_n$, where P_j has a private input x_j. Their goal is to compute $f(x_1, \ldots, x_n)$, such that none of them can learn more about the input of any other player than they can infer from $f(x_1, \ldots, x_n)$. Fortunately, the situation changes dramatically when extending the protocol to multiple players. SMP can be implemented with unconditional security even classically, provided that $t < n/3$ of the players are dishonest [Gol01]. If the adversary is not dishonest,

but merely honest-but-curious, it is possible to increase t up to $t < n/2$ [Gol01]. We refer to [Cra99] for an overview of classical secure multi-party computation.

Quantumly, one can generalize secure multi-party computation to the following setting. Each player P_j holds an input state $\rho_i \in \mathcal{H}$ (see Chapter 2 for details). Let $\rho \in \mathcal{H}_1 \otimes \ldots \otimes \mathcal{H}_n$ denote the joint state of players $P_1, \ldots, P_n$. Then quantum secure multi-party computation (QSMP) allows the players to compute any quantum transformation U to obtain $U\rho U^\dagger$, where player P_j receives the quantum state on $\mathcal{H}_j$ as his output. QSMP can be implemented securely if $t < n/2$ of the players are dishonest [CGS02, CGS05].

Coin tossing has also been studied in the multi-party setting. Classically, multi-party coin tossing forms part of secure multi-party computation [Gol01], and can thus be implemented under the same assumptions. Quantumly, multi-party coin tossing has been studied in [ABRD04].

1.3.3 Secret sharing

Another interesting problem concerns the sharing of a classical or quantum secret. Imagine Alice holding an important piece of information, for example the launch code to her personal missile silo. Alice would like to enable members of her community to gain access, but wants to prevent a single individual from launching a missile on his own. Secret sharing enables Alice to distribute some secret data d among a set of n players, such that at least $t > 1$ players need to combine their individual shares to reconstruct the original secret d. A trivial secret sharing scheme for a bit $d \in \{0, 1\}$ involving just two players is as follows: Alice picks $r \in_R \{0, 1\}$ and hands $s_1 = d \oplus r$ to the first player, and $s_2 = r$ to the second player. Clearly, if r is chosen uniformly at random from $\{0, 1\}$, none of the individual players can gain any information about d. Yet, when combining their individual shares they can compute $s_1 \oplus s_2 = d$.

General secret-sharing schemes were introduced by Shamir [Sha79] and Blakey [Bla79]. They have found a wide range of applications, most notably to construct protocols for secure multi-party computation as described in Section 1.3.2. Many classical secret sharing schemes are known today [MvOV97]. Quantum secret sharing was first introduced in [HBB99] and shortly after in [CGL99], which also formed a link between quantum secret sharing schemes and error correcting codes. Quantumly, we can distinguish two types of secret sharing schemes: The first allows to share a *quantum* secret, i.e., Alice holds a quantum state ρ and wants to construct n quantum shares $\sigma_1, \ldots, \sigma_n$ such that when t such shares are combined ρ can be reconstructed [HBB99, CGL99, Got00]. The second allows us to share *classical* secrets using quantum states that have very nice *data-hiding* properties [DLT02, DHT03, EW02, HLS05]: it is not sufficient for n parties to perform local measurements and communicate classically in order to reconstruct the secret. To reconstruct the secret data they must communicate quantumly to perform a coherent measurement on their states. It is an exciting open question

whether such schemes can be used to implement *quantum* protocols for secure multi-party computations with *classical* inputs that remain secure as long as the dishonest players can only communicate classically, but not quantumly.

1.3.4 Anonymous transmissions

In all applications we considered so far, we were concerned with two aspects: either, we wanted to protect protocol participants from being cheated by the other players, or, we wanted to protect the secrecy of data from a third party as in the setting of key distribution described in Section 1.1. In the problem of key distribution, sender and receiver know each other, but are trying to protect their data exchange from prying eyes. Anonymity, however, is the secrecy of identity. Primitives to hide the sender and receiver of a transmission have received considerable attention in classical computing. Such primitives allow any member of a group to send and receive data anonymously, even if all transmissions can be monitored. They play an important role in protocols for electronic auctions [SA99], voting protocols and sending anonymous email [Cha81]. An anonymous channel which is completely immune to any active attacks, would be a powerful primitive. It has been shown how two parties can use such a channel to perform key-exchange [AS83].

A considerable number of classical schemes have been suggested for anonymous transmissions. An unconditionally secure classical protocol was introduced by Chaum [Cha88] in the context of the Dining Cryptographers Problem. Such a protocol can also be considered an instance of secure multi-party computation considered above.

Boykin [Boy02] considered a quantum protocol to send classical information anonymously where the players distribute and test pairwise shared EPR pairs, which they then use to obtain key bits. His protocol is secure in the presence of noise or attacks on the quantum channel. In [CW05a], we presented a protocol for anonymous transmissions of classical data that achieves a novel property that cannot be achieved classically: it is completely *traceless*. This property is related, but stronger than the notion of incoercibility in secure multi-party protocols [CG96]. Informally, a protocol is traceless, if a player cannot be forced to reveal his true input at the end of the protocol. Even when forced to hand out his input, output and randomness used during the course of the protocol, a player is able to generate fake input that is consistent with all other data gathered from the run of the protocol. The protocols suggested in [Boy02] are not traceless, but can be modified to exhibit this property. It would be interesting to see whether it is possible to make general protocols for secure multi-party computation similarly traceless.

The first protocol for the anonymous transmission of qubits was constructed in [CW05a]. Whereas the anonymous transmissions of classical bits can be implemented via secure multi-party computation, the scenario is different when we

wish to transmit qubits: as we will see in Chapter 2, qubits cannot be copied. Thus we cannot expect each player to obtain a copy of the output. New protocols for creating anonymous entanglement and anonymously transmitting qubits have since been suggested in [BS05, BBF$^+$].

1.3.5 Other protocols

Besides the protocols above, a variety of other primitives making use of particular quantum effects have been proposed. One of the oldest suggested applications is the one of quantum money that is resistant to copying [Wie83], also proposed as unforgeable subway tokens [BBBW82]. Quantum seals [BP03, Cha03, SS05] employ the notion of cheat sensitivity in order to provide data with a seal that is "broken" once the data is extracted. That is, we can detect whether the data has been read. Perfect quantum seals that allow us to detect tampering with certainty have been shown to be impossible [BPDM05]. Nevertheless, non-trivial constructions are can be implemented.

Furthermore, quantum signature schemes [GC01] have been proposed which exhibit unconditional security: here Bob can verify Alice's signature using a public key given to him ahead of time. Sadly, such a scheme slowly consumes the necessary public key. Finally, protocols have been suggested for the encryption of quantum data which allow n qubits to be encoded using a $2n$ bit key achieving perfect secrecy [BR03, AMTdW00]. Much smaller keys are possible, if we allow for small imperfections [DN06, AS04]. Such encryption schemes have also been used to allow for private circuit evaluation [Chi05]: Here, Alice encrypts her quantum state before handing it to Bob who is capable of running a certain quantum operation that Alice would like to apply. This allows Alice to let her quantum operations be performed by Bob without revealing her quantum input.

1.4 Challenges

As we saw in Section 1.2.3, introducing quantum elements into cryptography leads to interesting new effects. Much progress has been made to exploit these quantum effects, although many open questions remain. In particular, not much is known about how well quantum protocols compose. That is, when we use one protocol as a building block inside a larger application, does the protocol still remain secure as expected? Recall from Section 1.2.3 that especially our ability to delay quantum measurements has a great influence on composition. Fortunately, quantum key distribution has been shown composable [BOHL$^+$05, Ren05, RK05]. However, composability remains a particularly tricky question in protocols where we are not faced with an external eavesdropper, but where the players themselves are dishonest. Composability of quantum protocols was first considered in [vdG98], followed by [CGS02] who addressed the composability of QSMP, and the general

composability frameworks of [Unr04, BOM04] applied to QKD [BOHL$^+$05]. Great care must also be taken when composing quantum protocols in the bounded quantum storage model [WW07]. Even though these composability frameworks exist, very few protocols have been proven secure when composed.

Secondly, we need to consider what happens if an adversary is allowed to store even small amounts of quantum information. There are many examples known where quantum memory can prove much more useful to an adversary than classical memory [GKK$^+$06], and we will encounter such examples in Chapters 3 and 5.

In addition, it is often assumed that the downfall of computational assumptions such as factoring is the only consequence that quantum computing has on the security of classical protocols. Sadly, this is by no means the only problem. Classical protocols where the security depends on the fact that different players cannot communicate during the course of the protocol may be broken when the players can share quantum entanglement and perform even a very limited set of quantum operations, well within the reach of current day technology. We will encounter such an example in Chapter 9.

Furthermore, we may conceive new primitives, unknown to the classical setting. One such primitive is the distribution of shared quantum states in the presence of dishonest players. Here, our goal is to create a protocol among n players such that at the end of the protocol $m \leq n$ players share a specified state ρ, where the dishonest players may apply any measurement to their share. It is conceivable to extend the QSMP protocol of [CGS02] to address this problem, yet, much more efficient protocols may be possible. Such a primitive would also enable us to build up the resources needed by other protocols such as [CW05a].

Finally, it is an interesting question by itself, what cryptographic primitives are possible in a quantum mechanical world. Conversely, it has even been shown that the axioms governing quantum mechanics can in part be obtained from the premise that perfect bit commitment is impossible [CBH03]. Perhaps such connections may lead to novel insights.

1.5 Conclusion

Quantum cryptography beyond quantum key distribution is an exciting subject. In this thesis, we will investigate several aspects that play an important role in nearly all cryptographic applications in the quantum setting.

In part I, we will examine how to extract information from quantum states. We first consider the problem of state discrimination. Here, our goal is to determine the identity of a state ρ within a finite set of possible states $\{\rho_1, \ldots, \rho_n\}$. In Chapter 3, we will examine a special case of this problem that is of particular relevance to quantum cryptography in the bounded quantum storage model: How

well can we perform state discrimination if we are given additional information after an initial quantum measurement, i.e., after a quantum memory bound is applied? In Chapter 4, we address uncertainty relations, which play an important role in nearly all cryptographic applications. We will prove tight bounds for uncertainty relations for certain mutually unbiased measurements. We will also present optimal uncertainty relations for anti-commuting measurements. Finally, in Chapter 5, we then examine a peculiar quantum effect known as locking classical information in quantum states. Such effects are important in the security of QKD, and also play a role in quantum string commitments which we will encounter in part III. In particular, we address the following question: Can we always obtain good locking effects for mutually unbiased measurements?

In part II, we turn to investigate quantum entanglement. In Chapter 7, we show how to find optimal quantum strategies for two parties who cannot communicate, but share quantum entanglement. Understanding such strategies plays an important part in understanding the effect of entanglement in otherwise classical protocols. In Chapter 8, we then present some initial weak result on the amount of entanglement such strategies require. Finally, in Chapter 9, we show how the security of classical protocols can be affected considerably in the presence of entanglement.

In part III, we investigate two cryptographic problems directly. In Chapter 10, we first consider commitments: Quantumly, one may hope that committing to an entire string of bits at once, and allowing Alice and Bob a limited ability to cheat, may still be within the realm of possibilities. This does not contradict that bit commitment itself is impossible. Unfortunately, we will see that for any reasonable security measure, string commitments are also impossible. However, non-trivial protocols do become possible for very weak notions of security.

In Chapter 11, we then introduce the model of noisy-quantum storage that in spirit is very similar to the setting of bounded-quantum storage: Here we assume that the adversary's quantum operations and storage are subject to noise. We show that oblivious transfer can be implemented securely in this model. We give an explicit tradeoff between the amount of noise and the security of our protocol.

Part II

Information in quantum states

Chapter 2

Introduction

To investigate the limitations and possibilities of cryptographic protocols in a physical world, we must familiarize ourselves with its physical theory: quantum mechanics. What are quantum states and what sets them apart from the classical scenario? Here, we briefly recount the most elementary facts that will be necessary for the remainder of this text. We refer to [Per93] for a more gentle introduction to quantum mechanics, to Appendix A for linear algebra prerequisites, and to the symbol index on page 249 for unfamiliar notation. In later chapters, we examine some of the most striking aspects of quantum mechanics, such as uncertainty relations and entanglement in more detail.

2.1 Quantum mechanics

2.1.1 Quantum states

A d-dimensional quantum state is a positive semidefinite operator ρ of norm 1 (i.e., ρ has no negative eigenvalues and $\mathrm{Tr}(\rho) = 1$) living in a d-dimensional Hilbert space[1] $\mathcal{H}$. We commonly refer to ρ as a *density operator* or *density matrix*. A special case of a quantum state is a *pure state*, which has the property that $\mathrm{rank}(\rho) = 1$. That is, there exists some vector $|\Psi\rangle \in \mathcal{H}$ such that we can write $\rho = |\Psi\rangle\langle\Psi|$, where $|\Psi\rangle\langle\Psi|$ is a projector onto the vector $|\Psi\rangle$. If $\{|0\rangle, \ldots, |d-1\rangle\}$ is a basis for $\mathcal{H}$, we can thus write $|\Psi\rangle = \sum_{j=0}^{d-1} \alpha_j |j\rangle$ for some coefficients $\alpha_j \in \mathbb{C}$. Note that our normalization constraint implies that $\mathrm{Tr}(\rho) = \sum_j |\alpha_j|^2 = 1$. We also say that $|\Psi\rangle$ is in a *superposition* of vectors $|0\rangle, \ldots, |d-1\rangle$. Clearly, for a pure state we have that $\rho^2 = \rho$ and thus $\mathrm{Tr}(\rho^2) = 1$.

Let's first look at an example of pure states. Suppose we consider a $d = 2$ dimensional quantum system $\mathcal{H}$, also called a *qubit*. We call $\{|0\rangle, |1\rangle\}$ the

[1] A complete vector space with an inner product. Here, we always consider a vector space over the complex numbers.

computational basis, where

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Any pure qubit state can then be written as $|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle$ for some $\alpha, \beta \in \mathbb{C}$ with $|\alpha|^2 + |\beta|^2 = 1$. We take an *encoding of '0' or '1'* in the computational basis to be $|0\rangle$ or $|1\rangle$ respectively, and use the subscript '+' to refer to an encoding in the computational basis. An alternative choice of basis would be the *Hadamard basis*, given by vectors $\{|+\rangle, |-\rangle\}$, where

$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \text{ and } |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle).$$

We use '$\times$' to refer to an encoding in the Hadamard basis. We will often consider systems consisting of n qubits. If $\mathcal{H}$ is a 2-dimensional Hilbert space corresponding to a single qubit, the system of n qubits is given by the n-fold tensor product $\mathcal{H}^{\otimes n}$ with dimension $d = 2^n$. A basis for this larger Hilbert space can easily be found by forming the tensor products of the basis vectors of a single qubit. For example, the computational basis for an n-qubit system is given by the basis vectors $\{|x_1\rangle \otimes \ldots \otimes |x_n\rangle \mid x_j \in \{0,1\}, j \in [n]\}$ where $[n] = \{1, \ldots, n\}$. We will often omit the tensor product and use the shorthand $|x_1 \ldots x_n\rangle = |x_1\rangle \otimes \ldots \otimes |x_n\rangle$.

If ρ is not pure, then ρ is a *mixed state* and can be written as a *mixture* of pure states. That is, for any state ρ there exist $\lambda_j \geq 0$ with $\sum_j \lambda_j = 1$ and vectors $|\Psi_j\rangle$ such that

$$\rho = \sum_j \lambda_j |\Psi_j\rangle\langle\Psi_j|.$$

Since ρ is Hermitian, we can take λ_j and $|\Psi_j\rangle$ to be the eigenvalues and eigenvectors of ρ respectively. We thus have for any quantum state that $\mathrm{Tr}(\rho^2) \leq 1$, where equality holds if and only if ρ is a pure state. We can also consider a mixture of quantum states, pure or mixed. Suppose we have a physical system whose state ρ_x depends on some value $x \in \mathcal{X}$ of a classical random variable X drawn from $\mathcal{X}$ according to a probability distribution P_X. For anyone who does not know the value of X (but does know the distribution P_X), the state of the system is given as

$$\rho = \sum_x P_X(x)\rho_x.$$

We also call the set $\mathcal{E} = \{(P_X(x), \rho_x) \mid x \in \mathcal{X}\}$ an *ensemble*, that *gives rise* to the density matrix ρ. We generally use the common shorthand $\mathcal{E} = \{P_X(x), \rho_x\}$. Clearly, for any state ρ we can take its eigendecomposition as above to find one possible ensemble that gives rise to ρ. With this interpretation in mind, it is now intuitive why we wanted $\rho \geq 0$ and $\mathrm{Tr}(\rho) = 1$: the first condition ensures that ρ has no negative eigenvalues and hence all probabilities λ_j are non-negative. The

second condition ensures that the resulting distribution in indeed normalized. We will use $\mathcal{S}(\mathcal{H})$ and $\mathbb{B}(\mathcal{H})$ to denote the set of all density matrices and the set of all bounded operators on a system $\mathcal{H}$ respectively.

Let's look at a small example illustrating the concept of mixed quantum states. The density matrices corresponding to $|0\rangle$ and $|1\rangle$ are $\rho_{0+} = |0\rangle\langle 0|$ and $\rho_{1+} = |1\rangle\langle 1|$, and the density matrices corresponding to $|+\rangle$ and $|-\rangle$ are given by $\rho_{0\times} = |+\rangle\langle +|$ and $\rho_{1\times} = |-\rangle\langle -|$. Let's suppose we are now told that we are given a '0' but encoded in either the computational or Hadamard basis, each with probability 1/2. Our quantum state corresponding to this encoding of '0' is now

$$\rho_0 = \frac{1}{2}(\rho_{0+} + \rho_{0\times}).$$

The state corresponding to an encoding of '1' is similarly given by

$$\rho_1 = \frac{1}{2}(\rho_{1+} + \rho_{1\times}).$$

It is important to note that the same density matrix can be generated by two different ensembles. As a simple example, consider the matrix $\rho = (2/3)|0\rangle\langle 0| + (1/3)|1\rangle\langle 1|$. Clearly, $\rho \geq 0$ and $\mathrm{Tr}(\rho) = 1$ and thus ρ forms a valid one qubit quantum state. However, $\mathcal{E}_1 = \{(2/3, |0\rangle), (1/3, |1\rangle)\}$ and $\mathcal{E}_2 = \{(1/2, |\phi_0\rangle), (1/2, |\phi_1\rangle)\}$ with $|\phi_0\rangle = \sqrt{2/3}|0\rangle + \sqrt{1/3}|1\rangle$ and $|\phi_1\rangle = \sqrt{2/3}|0\rangle - \sqrt{1/3}|1\rangle$ both give rise to ρ:

$$\rho = \frac{2}{3}|0\rangle\langle 0| + \frac{1}{3}|1\rangle\langle 1| = \frac{1}{2}|\phi_0\rangle\langle\phi_0| + \frac{1}{2}|\phi_1\rangle\langle\phi_1|.$$

Classical vs. Quantum

Quantum states exhibit an important property known as "no-cloning": very much unlike classical states, we cannot create a copy of an arbitrary quantum state! This is only possible with a small probability. We refer to [SIGA05] for an excellent overview of known results.

In the following, we call an ensemble *classical* if all states ρ_x commute. This is an interesting special case, we discuss in more detail below.

2.1.2 Multipartite systems

We frequently need to talk about a quantum state shared by multiple players in a protocol. Let $\mathcal{H}_1, \ldots, \mathcal{H}_n$ denote the Hilbert spaces corresponding to the quantum systems of players 1 up to n. As outlined in the case of multiple qubits above, the joint system $\mathcal{H}_1 \otimes \ldots \otimes \mathcal{H}_n$ of all players is formed by taking the tensor product. For example, suppose that we have only two players, Alice and Bob. Let $\mathcal{H}^A$ and $\mathcal{H}^B$ be the Hilbert spaces corresponding to Alice's and Bob's quantum systems respectively. Any *bipartite state* ρ^{AB} shared by Alice and Bob is a state living in the joint system $\mathcal{H}^A \otimes \mathcal{H}^B$. Bipartite states can exhibit an interesting

property called entanglement, which we investigate in Chapter 6. In short, if $|\Psi\rangle \in \mathcal{H}^A \otimes \mathcal{H}^B$ is a pure state, we say that $|\Psi\rangle$ is *separable* if and only if there exist states $|\Psi^A\rangle \in \mathcal{H}^A$ and $|\Psi^B\rangle \in \mathcal{H}^B$ such that $|\Psi\rangle = |\Psi^A\rangle \otimes |\Psi^B\rangle$. A separable pure state is also called a *product state*. A state that is not separable is called *entangled*. An example of an entangled pure state is the so-called EPR-pair

$$\frac{1}{2}(|00\rangle + |11\rangle).$$

For mixed states the definition is slightly more subtle. Let $\rho \in \mathcal{S}(\mathcal{H}^A \otimes \mathcal{H}^B)$ be a mixed state. Then ρ is called a *product state* if there exist $\rho^A \in \mathcal{H}^A$ and $\rho^B \in \mathcal{H}^B$ such that $\rho = \rho^A \otimes \rho^B$. The state ρ is called *separable*, if there exists an ensemble $\mathcal{E} = \{p_j, |\Psi_j\rangle\}$ such that $|\Psi_j\rangle = |\Psi_j^A\rangle \otimes |\Psi_j^B\rangle$ with $|\Psi_j^A\rangle \in \mathcal{H}^A$ and $|\Psi_j^B\rangle \in \mathcal{H}^B$ for all j, such that

$$\rho = \sum_j p_j |\Psi_j\rangle\langle\Psi_j| = \sum_j p_j |\Psi_j^A\rangle\langle\Psi_j^A| \otimes |\Psi_j^B\rangle\langle\Psi_j^B|.$$

Intuitively, if ρ is separable then ρ corresponds to a mixture of separable pure states according to a classical joint probability distribution $\{p_j\}$. We return to such differences in Chapter 6. From a cryptographic perspective, it is for now merely important to note that if the state ρ^{AB} shared between Alice and Bob is a pure state, then ρ^{AB} is not entangled with any third system $\mathcal{H}^C$ held by Charlie. That is, ρ^{AB} does not depend on any classical random variable X held by Charlie whose value is unknown to Alice and Bob. An important consequence is that the outcomes of any measurement (see below) that Alice and Bob may perform on ρ^{AB} are therefore independent of X, and hence secret with respect to Charlie.

Given a quantum state in a combined, larger, system, what can we say about the state of the individual systems? For example, given a state ρ^{AB} shared between Alice and Bob, the *reduced state* of Alice's system alone is given by $\rho^A = \mathrm{Tr}_B(\rho^{AB})$, where Tr_B is the partial trace over Bob's system. The partial trace operation $\mathrm{Tr}_B : \mathbb{B}(\mathcal{H}^A \otimes \mathcal{H}^B) \to \mathbb{B}(\mathcal{H}^A)$ is thereby defined as the unique linear operator that for all $A \in \mathbb{B}(\mathcal{H}^A)$ and all $B \in \mathbb{B}(\mathcal{H}^B)$ maps $\mathrm{Tr}_B(A \otimes B) = A\mathrm{Tr}(B)$. We also say that we *trace out* Bob's system from ρ^{AB} to obtain ρ^A. Furthermore, given any state $\rho^A \in \mathcal{H}^A$, we can always find a second system $\mathcal{H}^B$ and a pure state $|\Psi\rangle \in \mathcal{H}^A \otimes \mathcal{H}^B$ such that $\rho^A = \mathrm{Tr}_B(|\Psi\rangle\langle\Psi|)$. We call $|\Psi\rangle$ a *purification* of ρ^A.

Classical vs. Quantum

In the quantum world, we encounter a particular effect known as entanglement. Intuitively, entanglement leads to very strong correlations among Alice and Bob's system, which we will examine in detail in Chapter 6.

2.1.3 Quantum operations

Unitary evolution

The evolution of any closed quantum system is described by a *unitary evolution* U that maps

$$\rho \to U\rho U^\dagger.$$

It is important to note that unitary operations are reversible: We can always apply an additional unitary $V = U^\dagger$ to retrieve the original state since $V(U\rho U^\dagger)V^\dagger = U^\dagger U\rho U^\dagger U = \rho$. In particular, we often make use of the following single qubit unitaries known as the Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Note that $\sigma_y = i\sigma_x\sigma_z$. Furthermore, we also use the Hadamard, and the K-transform given by

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \text{ and } K = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}.$$

Note that $K = (\mathbb{I} + i\sigma_x)/\sqrt{2}$.

Measurements

Besides unitary operations, we can also perform measurements on the quantum state. A *quantum measurement* of a state $\rho \in \mathcal{S}(\mathcal{H})$ is a set of operators $\{M_m\}$ acting on $\mathcal{S}(\mathcal{H})$, satisfying $\sum_m M_m^\dagger M_m = \mathbb{I}$. We will call operators M_m *measurement operators*. The probability of obtaining outcome m when measuring the state ρ is given by

$$\Pr[m] = \mathrm{Tr}(M_m^\dagger M_m \rho).$$

Conditioned on the event that we obtained outcome m, the *post-measurement state* of the system is now

$$\rho_m = \frac{M_m \rho M_m^\dagger}{\mathrm{Tr}(M_m^\dagger M_m \rho)}.$$

Most measurements *disturb* the quantum state and hence ρ_m generally differs from ρ. We will discuss this effect in more detail below. Note that we have $\sum_m \Pr[m] = \mathrm{Tr}\left(\left(\sum_m M_m^\dagger M_m\right)\rho\right) = 1$, and hence the distribution over outcomes $\{m\}$ is appropriately normalized.

A special case of a quantum measurement is a *projective measurement*, where all measurement operators M_m are orthogonal projectors which we write as $P_m = M_m = M_m^\dagger M_m$. Projective measurements are also described via an *observable* $A = \sum_m mP_m$, where $m \in \mathbb{R}$. Note that A is a Hermitian matrix with eigenvalues

$\{m\}$. For any given basis $\mathcal{B} = \{|x_1\rangle, \ldots, |x_d\rangle\}$ we speak of *measuring in the basis* $\mathcal{B}$ to indicate that we perform a projective measurement given by operators $P_k = |x_k\rangle\langle x_k|$ with $k \in [d]$.

If we are only interested in the measurement outcome, but do not care about the post-measurement state, it is often simpler to make use of the POVM (positive operator valued measure) formalism. A POVM is a set of Hermitian operators $\{E_m\}$ such that $\sum_m E_m = \mathbb{I}$ and for all m we have $E_m \geq 0$. Evidently, from a general measurement we can obtain a POVM by letting $E_m = M_m^\dagger M_m$. We now have

$$\Pr[m] = \mathrm{Tr}(E_m \rho).$$

The advantage of this approach is that we can easily solve optimization problems involving probabilities $\Pr[m]$ over the operators E_m, instead of considering the individual operators M_m. Since $E_m \geq 0$ such problems can be solved using semidefinite programming, which we describe in Appendix A. Finally, it is important to note that quantum measurements do not always commute: it matters crucially in which order we execute them. Indeed, as we will see later it is this property that leads to all the interesting quantum effects we will consider.

Let's consider a small example. Suppose we are given a pure quantum state $|\Psi\rangle = \sqrt{2/3}|0\rangle + \sqrt{1/3}|1\rangle$. When measuring $|\Psi\rangle$ in the computational basis, we perform a measurement determined by operators $P_0 = |0\rangle\langle 0|$ and $P_1 = |1\rangle\langle 1|$. Evidently, we have

$$\Pr[0] = \mathrm{Tr}(P_0|\Psi\rangle\langle\Psi|) = \langle\Psi|P_0|\Psi\rangle = \frac{2}{3},$$

and

$$\Pr[1] = \mathrm{Tr}(P_1|\Psi\rangle\langle\Psi|) = \langle\Psi|P_1|\Psi\rangle = \frac{1}{3}.$$

If we obtained outcome '0', the post-measurement state is given by

$$\rho_0 = \frac{P_0|\Psi\rangle\langle\Psi|P_0}{\Pr[0]} = |0\rangle\langle 0|.$$

Similarly, if we obtained outcome '1', the post-measurement state is

$$\rho_1 = \frac{P_1|\Psi\rangle\langle\Psi|P_1}{\Pr[1]} = |1\rangle\langle 1|.$$

Quantum channel

The most general way to describe an operation is by means of a CP (completely positive) map $\Lambda : \mathcal{H}^A \to \mathcal{H}^B$, where $\mathcal{H}^A$ and $\mathcal{H}^B$ denote the in and output systems respectively. We also call Λ a *channel*. Any channel Λ can be written as $\Lambda(\rho) = \sum_m V_m \rho V_m^\dagger$ where V_m is a linear operator from $\mathcal{H}^A$ to $\mathcal{H}^B$, and $\sum_m V_m^\dagger V_m \leq \mathbb{I}$. V_m is also referred to as a *Kraus operator*. Λ is *trace preserving* if $\sum_m V_m^\dagger V_m = \mathbb{I}$. Any

quantum operation can be expressed by means of a CPTP (completely positive trace preserving) map . We sometimes also refer to such a map as a *superoperator*, a *quantum channel*, or a (measurement) *instrument*, if we think of a POVM with elements $\{V_m\}$. A channel is called *unital*, if in addition $\sum_m V_m V_m^\dagger = \mathbb{I}$: we then have $\Lambda(\mathbb{I}) = \mathbb{I}$.

We give two simple examples. Consider the unitary evolution U of a state ρ: here we have $\Lambda(\rho) = U\rho U^\dagger$. When we perform our single qubit measurement in the computational basis described above, and ignore the measurement outcome, we implement the channel $\Lambda(\rho) = P_0\rho P_0 + P_1\rho P_1$. Since P_0 and P_1 form a measurement and are projectors we also have that $P_0P_0^\dagger + P_1P_1^\dagger = \mathbb{I}$ and hence the channel is unital.

Any quantum channel can be described by a unitary transformation on the original and an ancilla system, where the ancilla system is traced out to recover the original operation. More precisely, given a channel $\Lambda : \mathcal{H}^A \to \mathcal{H}^B$ we can choose a Hilbert space $\mathcal{H}^C$ identical to $\mathcal{H}^B$, a pure state $\hat{\rho} \in \mathcal{S}(\mathcal{H}^B \otimes \mathcal{H}^C)$ and a unitary matrix U_Λ acting on $\mathcal{H}^A \otimes \mathcal{H}^B \otimes \mathcal{H}^C$ such that for any $\rho \in \mathcal{S}(\mathcal{H}^A)$ $\Lambda(\rho) = \mathrm{Tr}_{A,C} U_\Lambda(\rho \otimes \hat{\rho})U_\Lambda^\dagger$. This is all that we need here, and we refer to [Hay06] for detailed information.

Of particular interest, especially with regard to constructing cheat-sensitive protocols, is the following statement which specifies which operations leave a given set of states invariant. Clearly, any cheating party may always perform such operations without being detected. It has been shown that

2.1.1. Lemma. *(HKL) [HKL03] Let $\Lambda : \mathcal{H} \to \mathcal{H}$ be a unital quantum channel with $\Lambda(\rho) = \sum_m V_m \rho V_m^\dagger$, and let $\mathcal{S}$ be a set of quantum states. Then*

$$\forall \rho \in \mathcal{S}, \Lambda(\rho) = \rho \text{ if and only if } \forall m \forall \rho \in \mathcal{S}, [V_m, \rho] = 0.$$

Indeed, the converse direction is easy to see. If we have that for all m and for all $\rho \in \mathcal{S}$ $[V_m, \rho] = 0$, then $\Lambda(\rho) = \sum_m V_m \rho V_m^\dagger = \sum_m V_m V_m^\dagger \rho = \rho$, since Λ is unital. If a quantum channel is not of this form, i.e. it does not leave the state invariant, we also say that it *disturbs* the state. The statement above has interesting consequences: consider an ensemble of states $\mathcal{E} = \{p_x, \rho_x\}$ with $\rho_x \in \mathcal{H}$, and suppose that there exists a decomposition $\mathcal{H} = \bigoplus_j \mathcal{H}_j$ such that for all x we have $\rho_x = \sum_j \Pi_j \rho_x \Pi_j$ where Π_j is a projector onto $\mathcal{H}_j$. If we perform the measurement given by operators $\{\Pi_j\}$ then (ignoring the outcome) the states ρ_x are invariant under such a measurement, since clearly $[\Pi_j, \rho_x] = 0$ for all j and x. The outcome of the measurement tells us which $\mathcal{H}_j$ we reside in. However, Lemma 2.1.1 tells us a lot more: We will see in Chapter 3.5.1 that if the measurement operators from a projective measurement commute with all the states ρ_x, they are in fact of this very form (see also Appendix B). In the following, we call the information about which $\mathcal{H}_j$ we reside in the *classical information of the ensemble* $\mathcal{E}$. Any attempt to gain more information, i.e. by

performing measurements which do not satisfy these commutation properties, necessarily leads to disturbance and can be detected.

An adversary can thus always extract this classical information without affecting the quantum state. Looking back at Chapter 1, we can now see that for unital adversary channels we can define an honest-but-curious player to be honest-but-curious with regard to the classical information, and honest with regard to the quantum information: he may extract, copy and memorize the classical information as desired. However, if he wants to leave the protocol execution itself unaltered, he cannot perform any other measurements and must thus be honest on the remaining quantum part of the ensemble.

Classical vs. Quantum

Clearly, Lemma 2.1.1 also tells us that if all the states ρ_x in our ensemble commute, i.e. the ensemble is classical as defined above, then we can always perform a measurement in their common eigenbasis "for free". Furthermore, if our ensemble is classical we have $\dim(\mathcal{H}_j) = 1$, i.e. $\mathcal{H}_j$ itself is also classical: it is just a scalar. We thus see that such an ensemble has no quantum properties: we can extract and copy information at will. Informally, we may think of the different states within the ensemble as different classical probability distributions over their common eigenstates. We will return to this idea shortly.

Furthermore, we can look at measurements or observables themselves. Note again from the above that since a quantum measurement may disturb a state, it matters in which order measurements are executed. That is, quantum operations do not commute. It is this fact that leads to all the interesting effects we observe: uncertainty relations, locking and Bell inequality violations using quantum entanglement are all consequences of the existence of non-commuting measurements in the quantum world. This lies in stark contrast to the classical world, where all our measurement do commute, and we therefore do not encounter such effects.

2.2 Distinguishability

How can we distinguish several quantum states? Suppose we are given states ρ_X where X is a random variable drawn according to a probability distribution P_X over some finite set $\mathcal{X}$. Our goal is now to determine the value of X given an unknown state $\rho \in \{\rho_x \mid x \in \mathcal{X}\}$. Cryptographically, this gives an intuitive measure on how well we can guess the value of X. The problem of finding the optimal distinguishing measurement is called *state discrimination*, where optimal refers to finding the measurement that maximizes the probability of successfully guessing X. For two states, the optimal guessing probability is particularly simple to evaluate. To this end, we first need to introduce the trace distance, and the trace norm:

2.2.1. Definition. The *trace distance* of two states ρ_0 and ρ_1 is given by

$$D(\rho_0, \rho_1) = \frac{1}{2}||\rho_0 - \rho_1||_1,$$

where $||A||_1 = \mathrm{Tr}(\sqrt{A^\dagger A})$ is the *trace norm* of A.

Alternatively, the trace distance may also be expressed as [Hay06]

$$D(\rho_0, \rho_1) = \max_M \mathrm{Tr}(M(\rho_0 - \rho_1)),$$

where the maximization is taken over all $M \geq 0$. Indeed, D is really a "distance" measure, as it is clearly a metric on the space of density matrices: We have $D(\rho_0, \rho_1) = 0$ if and only if $\rho_0 = \rho_1$, and evidently $D(\rho_0, \rho_1) = D(\rho_1, \rho_0)$. Finally, the triangle inequality holds:

$$\begin{aligned} D(\rho_0, \rho_1) &= \max_M \mathrm{Tr}(M(\rho_0 - \rho_1)) = \max_M \left(\mathrm{Tr}(M(\rho_0 - \sigma)) + \mathrm{Tr}(M(\sigma - \rho_1))\right) \\ &\leq D(\rho_0, \sigma) + D(\sigma, \rho_1). \end{aligned}$$

When considering single qubits (such as for example in Chapter 11) it is often intuitive to note that for a single qubit, the trace distance has a particularly simple form. Note that $\mathbb{I}$, σ_x, σ_y and σ_z form a basis for the space of 2×2 complex matrices. Since we have $\mathrm{Tr}(\rho) = 1$ for any quantum state, we can thus write any single qubit state as

$$\rho = \frac{\mathbb{I} + \vec{r} \cdot \vec{\sigma}}{2} = \frac{\mathbb{I} + r_x\sigma_x + r_y\sigma_y + r_z\sigma_z}{2}$$

where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ and $\vec{r} = (r_x, r_y, r_z)$ is the *Bloch vector* as given in Figure 2.1. For $\tau = (\mathbb{I} + \vec{t} \cdot \vec{\sigma})/2$ with $\vec{t} = (t_x, t_y, t_z)$ we then have

$$D(\rho, \tau) = \frac{1}{2}||\rho - \tau||_1 = \frac{1}{2}\left|\left|\sum_{j\in\{x,y,z\}} (r_j - t_j)\sigma_j\right|\right|_1 = \frac{1}{2}\sqrt{\sum_{j\in\{x,y,z\}} (r_j - t_j)^2},$$

where we used the fact that all Pauli matrices anti-commute. Thus, the trace distance between ρ and τ is exactly half the Euclidean distance of the corresponding Bloch vectors.

Using the trace distance, we can address the problem of distinguishing *two* quantum states:

2.2.2. Theorem (Helstrom [Hel67]). *Suppose we are given states ρ_0 with probability q, and ρ_1 with probability $1 - q$. Then the probability to determine whether the state was ρ_0 and ρ_1 is at most*

$$p = \frac{1}{2}\left[1 + ||q\rho_0 - (1-q)\rho_1||_1\right].$$

The measurement that achieves p is given by M_0, and $M_1 = \mathbb{I} - M_0$, where M_0 is the projector onto the positive eigenspace of $q\rho_0 - (1-q)\rho_1$.

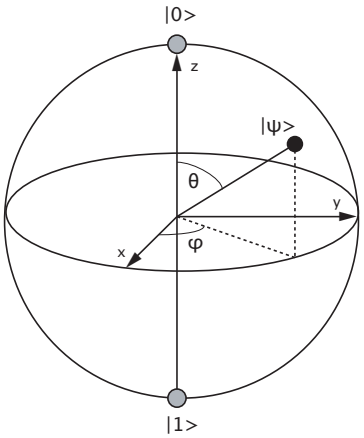


Figure 2.1: Bloch vector $(r_x, r_y, r_z) = (\cos\psi \sin\theta, \sin\psi \sin\theta, \cos\theta)$

For $q = 1/2$, this gives us $p = 1/2 + D(\rho_0, \rho_1)/2$. Indeed, it is easy to see why such M_0 and M_1 form the optimal measurement. Note that here we are only interested in finding a POVM. To find the optimal POVM we must solve the following optimization problem for variables M_0 and M_1:

$$\begin{array}{ll} \text{maximize} & q\mathrm{Tr}(M_0\rho_0) + (1-q)\mathrm{Tr}(M_1\rho_1) \\ \text{subject to} & M_0, M_1 \geq 0, \\ & M_0 + M_1 = \mathbb{I}. \end{array}$$

We can rewrite our target function as

$$\begin{aligned} q\mathrm{Tr}(M_0\rho_0) + (1-q)\mathrm{Tr}(M_1\rho_1) &= q\mathrm{Tr}(M_0\rho_0) + (1-q)\mathrm{Tr}((\mathbb{I} - M_0)\rho_1) \\ &= \mathrm{Tr}(M_0(q\rho_0 - (1-q)\rho_1)) + 1 - q \\ &= \mathrm{Tr}\left(M_0\left(\sum_{\lambda_j \geq 0} \lambda_j |u_j\rangle\langle u_j|\right)\right) \\ &\quad + \mathrm{Tr}\left(M_0\left(\sum_{\lambda_j < 0} \lambda_j |u_j\rangle\langle u_j|\right)\right) + 1 - q, \end{aligned}$$

where $q\rho_0 - (1-q)\rho_1 = \sum_j \lambda_j |u_j\rangle\langle u_j|$. Hence, to maximize the above expression, we need to choose $M_0 = \sum_{\lambda_j \geq 0} |u_j\rangle\langle u_j|$.

Unfortunately, computing the optimal measurement to distinguish more than two states is generally not so easy. Yuen, Kennedy and Lax [YKL75] first showed that this problem can be solved using semidefinite programming, a technique we describe in Appendix A. This technique has since been refined to address other variants such as unambiguous state discrimination where we can output

"don't know", but are never allowed to make a mistake [Eld03]. Evidently, we can express the optimization problem for any state discrimination problem as

$$\begin{array}{ll} \text{maximize} & \sum_x P_X(x)\mathrm{Tr}(M_x\rho_x) \\ \text{subject to} & \forall x \in \mathcal{X}, M_x \geq 0, \\ & \sum_{x\in\mathcal{X}} M_x = \mathbb{I}. \end{array}$$

In Chapter 3, we will use the above formulation. We also show how to address a variant of this problem, where we receive additional classical information after performing the measurement.

Closely related to the trace distance is the notion of fidelity.

2.2.3. Definition. The *fidelity* of states ρ and σ is given by

$$F(\rho,\sigma) = \mathrm{Tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}}.$$

Note that if $\rho = |\Psi\rangle\langle\Psi|$ is a pure state, this becomes

$$F(|\Psi\rangle,\sigma) = \sqrt{\langle\Psi|\sigma|\Psi\rangle}.$$

The fidelity is closely related to the trace distance. In particular, we have that for any states ρ and σ

$$1 - F(\rho,\sigma) \leq D(\rho,\sigma) \leq \sqrt{1-F(\rho,\sigma)^2}.$$

A proof can be found in [NC00, Section 9.2.3]. If $\rho = |\Psi\rangle\langle\Psi|$ is a pure state, the lower bound can be improved to

$$1 - F(|\Psi\rangle,\sigma)^2 \leq D(|\Psi\rangle,\sigma).$$

Many other distance measures of quantum states are known, which may be a more convenient choice for particular problems. We refer to [Fuc95, Hay06] for an overview.

Classical vs. Quantum

Suppose again we are given a classical ensemble of states ρ and σ. That is, both operators commute and hence have a common eigenbasis $\{|u_1\rangle,\ldots,|u_d\rangle\}$. We can thus write $\rho = \sum_j \lambda_j|u_j\rangle\langle u_j|$ and $\sigma = \sum_j \gamma_j|u_j\rangle\langle u_j|$, which allows us to write the trace distance of ρ and σ as

$$D(\rho,\sigma) = \frac{||\sum_j(\lambda_j-\gamma_j)|u_j\rangle\langle u_j|||_1}{2} = \frac{1}{2}\sum_j|\lambda_j-\gamma_j| = D(\lambda_j,\gamma_j),$$

where $D(\lambda_j,\gamma_j)$ is the *classical variational distance* between the distributions $\{\lambda_j\}$ and $\{\gamma_j\}$. Again, we see that there is nothing quantum in this setting. We can view ρ and σ as two different probability distributions over the set $\{|u_j\rangle\}$. Similarly, it is easy to see that

$$F(\rho,\sigma) = \mathrm{Tr}\sqrt{\sum_j \lambda_j\gamma_j|u_j\rangle\langle u_j|} = \sum_j\sqrt{\lambda_j\gamma_j} = F(\lambda_j,\gamma_j),$$

where $F(\lambda_j,\gamma_j)$ is the *classical fidelity* of the distributions $\{\lambda_j\}$ and $\{\gamma_j\}$.

2.3 Information measures

2.3.1 Classical

We also need the following ways of measuring information. Let X be a random variable distributed over a finite set $\mathcal{X}$ according to probability distribution P_X. The *Shannon entropy* of X is then given by

$$H(X) = -\sum_{x\in\mathcal{X}} P_X(x)\log P_X(x).$$

Intuitively, the Shannon entropy measures how much information we gain *on average* by learning X. A complementary view point is that $H(X)$ quantifies the amount of uncertainty we have about X before the fact. We will also use $H(P_X)$, if our discussion emphasizes a certain distribution P_X. If $|\mathcal{X}| = 2$, we also use the term *binary entropy* and use the shorthand

$$h(p) = -p\log p - (1-p)\log(1-p).$$

Let Y be a second random variable distributed over a finite set $\mathcal{Y}$ according to distribution P_Y. The *joint entropy* of X and Y can now be expressed as

$$H(X,Y) = -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} P_{XY}(x,y)\log P_{XY}(x,y),$$

where P_{XY} is the joint distribution over $\mathcal{X}\times\mathcal{Y}$. Furthermore, we can quantify the uncertainty about X given Y by means of the *conditional entropy*

$$H(X|Y) = H(X,Y) - H(Y).$$

To quantify the amount of information X and Y may have in common we use the *mutual information*

$$\mathcal{I}(X,Y) = H(X) + H(Y) - H(X,Y) = H(X) - H(X|Y).$$

Intuitively the mutual information captures the amount of information we gain about X by learning Y. The Shannon entropy has many interesting properties, summarized, for example, in [NC00, Theorem 11.3], but we do not require them here. In Chapter 5, we only need the *classical mutual information* of a bipartite quantum state ρ^{AB}, which is the maximum classical mutual information that can be obtained by local measurements $M^A\otimes M^B$ on the state ρ^{AB} [THLD02]:

$$\mathcal{I}_c(\rho^{AB}) = \max_{M^A\otimes M^B} \mathcal{I}(A,B), \tag{2.1}$$

where A and B are the random variables corresponding to Alice's and Bob's measurement outcomes respectively.

In a cryptographic setting, the Shannon entropy is not always a desirable measure as it merely captures our uncertainty about X *on average*. Often, the Rényi entropy allows us to make stronger statements The *Rényi entropy* [Rén60] of order α is defined as

$$H_\alpha(X) = \frac{1}{1-\alpha} \log \left[\sum_{x \in \mathcal{X}} P_X(x)^\alpha \right].$$

Indeed, the Shannon entropy forms a special case of the Rényi entropy by taking the limit $\alpha \to 1$, i.e., $H_1(\cdot) = H(\cdot)$, where we omit the subscript. Of particular importance is the *min-entropy*, for $\alpha \to \infty$:

$$H_\infty(X) = -\log \left(\max_{x \in \mathcal{X}} P_X(x) \right),$$

and the *collision entropy*

$$H_2(X) = -\log \sum_{x \in \mathcal{X}} P_X(x)^2.$$

We have

$$\log |\mathcal{X}| \geq H(X) \geq H_2(X) \geq H_\infty(X).$$

Intuitively, the min-entropy is determined by the highest peak in the distribution and most closely captures the notion of "guessing" x. Consider the following example: Let $\mathcal{X} = \{0,1\}^n$ and let $x_0 = 0, \ldots, 0$ be the all 0 string. Suppose that $P_X(x_0) = 1/2 + 1/(2^{n+1})$ and $P_X(x) = 1/(2^{n+1})$ for $x \neq x_0$, i.e., with probability $1/2$ we choose x_0 and with probability $1/2$ we choose one string uniformly at random. Then $H(X) \approx n/2$, whereas $H_\infty(X) = 1$! If x would correspond to an encryption key used to encrypt an n bit message, we would certainly not talk about security if we can guess the key with probability $1/2$! Yet, the Shannon entropy is quite high. We refer to [Cac97] for an in-depth discussion of security measures in classical cryptography.

2.3.2 Quantum

Similar to the Shannon entropy, the *von Neumann entropy* of a quantum states ρ is given by

$$S(\rho) = -\mathrm{Tr}(\rho \log \rho).$$

Taking the eigendecomposition of $\rho = \sum_x \lambda_x |x\rangle\langle x|$ we can also write

$$S(\rho) = -\sum_x \lambda_x \log \lambda_x,$$

which corresponds to the Shannon entropy arising from measuring ρ in the basis given by $\{|x\rangle\langle x|\}$. We refer to [NC00, Section 11.3] for the properties of the von Neumann entropy.

Here, we will only be concerned with the *accessible information* [Per93, Eq. (9.75)] of an ensemble $\mathcal{E} = \{p_x, \rho_x\}$ which we encounter again in Chapter 5.

$$\mathcal{I}_{acc}(\mathcal{E}) = \max_M \left(-\sum_x p_x \log p_x + \sum_j \sum_x p_x \alpha_j \mathrm{Tr}(M_j \rho_x) \log \frac{p_x \mathrm{Tr}(M_j \rho_x)}{\mathrm{Tr}(M_j \rho)} \right),$$

where $\rho = \sum_x p_x \rho_x$ and the maximization is taken over all POVMs $M = \{M_j\}$. It has been shown that we can take all POVM elements to be of rank 1 [Dav78]. However, maximizing this quantity still remains a hard task [Per93]. Some upper and lower bounds are known [Fuc95], but sadly none of them are generally very strong. The most well-known upper bound is given by the *Holevo quantity*, which is given by

$$\chi(\rho) = S(\rho) - \sum_x p_x S(\rho_x).$$

Holevo's theorem [NC00] states that

$$\mathcal{I}_{acc}(\mathcal{E}) \leq \chi(\rho). \tag{2.2}$$

Classical vs. Quantum

Equality in Eq. (2.2) is achieved if all states ρ_x have a common eigenbasis (i.e., all ρ_x commute). Hence, for classical ensembles we do not have a gap between these two quantities. The fact that quantumly we can obtain such a gap leads to a peculiar effect known as locking classical information in quantum states in Chapter 5. However, even if the states ρ_x do not commute, we can still extract the "classical information" of the ensemble: Suppose for all $\rho_x \in \mathcal{H}$ from our ensemble there exists a decomposition $\mathcal{H} = \bigoplus_j \mathcal{H}_j$ such that for all x, $\rho_x = \sum_j \Pi_j \rho_x \Pi_j$, where Π_j is a projector onto $\mathcal{H}_j$. That is, there exists a way to simultaneously block-diagonalize all states. Note that for any measurement maximizing the accessible information above, we can find an equivalent measurement with measurement operators $\hat{M} = \sum_j \Pi_j M \Pi_j$, since evidently, $\mathrm{Tr}(\hat{M}\rho_x) = \sum_j \mathrm{Tr}(\Pi_j M \Pi_j \rho_x) = \mathrm{Tr}(M\rho_x)$. Intuitively, this means that we can always first determine which block we are in "for free", followed by our original measurement constrained to this block. Note that $[\Pi_j, \rho_x] = 0$ for all Π_j and ρ_x. Hence, looking back at Section 2.1.3 this is not so surprising: the measurement leaves our states invariant. In general, such commutation relations lead to interesting structural consequences which we examine in more detail in Appendix B and also exploit in Chapter 3. Finally, it will be useful in Chapter 10 that the accessible information is additive [Hol73, DLT02]: For m independent draws of an ensemble $\mathcal{E}$ of separable states (see Chapter 6), i.e., we choose m states from m identical ensembles independently, we have $\mathcal{I}_{acc}(\mathcal{E}^{\otimes m}) = m\mathcal{I}_{acc}(\mathcal{E})$.

2.4 Mutually unbiased bases

In the following chapters, we will be particularly concerned with measurements in *mutually unbiased bases* (MUBs). MUBs were initially introduced in the context of state estimation [WF89], but feature in many other problems in quantum information. The following definition closely follows the one given in [BBRV02].

2.4.1. Definition. [MUBs] Let $\mathcal{B}_1 = \{|b_1^1\rangle, \ldots, |b_d^1\rangle\}$ and $\mathcal{B}_2 = \{|b_1^2\rangle, \ldots, |b_d^2\rangle\}$ be two orthonormal bases in $\mathbb{C}^d$. They are said to be *mutually unbiased* if $|\langle b_k^1|b_l^2\rangle| = 1/\sqrt{d}$, for every $k, l \in [d]$. A set $\{\mathcal{B}_1, \ldots, \mathcal{B}_m\}$ of orthonormal bases in $\mathbb{C}^d$ is called a *set of mutually unbiased bases* if each pair of bases is mutually unbiased.

As an example, consider the computational and the Hadamard basis defined above, and note that we can write $|+\rangle = H|0\rangle$ and $|-\rangle = H|1\rangle$. We then have for $x \in \{0,1\}^n$ that

$$|\langle x|H^{\otimes n}|x\rangle|^2 = \frac{1}{2^n}.$$

Hence, the computational and the Hadamard basis are mutually unbiased in dimension $d = 2^n$.

We use $N(d)$ to denote the maximal number of MUBs in dimension d. In any dimension d, we have that $\mathrm{N}(d) \leq d + 1$ [BBRV02]. If $d = p^k$ is a prime power, we have that $\mathrm{N}(d) = d+1$ and explicit constructions are known [BBRV02, WF89]. If $d = s^2$ is a square, $\mathrm{N}(d) \geq \mathrm{MOLS}(s)$ where $\mathrm{MOLS}(s)$ denotes the number of mutually orthogonal $s \times s$ Latin squares [WB05]. In general, we have $\mathrm{N}(nm) \geq \min\{\mathrm{N}(n), \mathrm{N}(m)\}$ for all $n, m \in \mathbb{N}$ [Zau99, KR03]. It is also known that in any dimension, there exists an explicit construction for 3 MUBs [Gra04]. Unfortunately, not much else is known. For example, it is still an open problem whether there exists a set of 7 MUBs in dimension $d = 6$. We say that a unitary U_t *transforms the computational basis into the t-th MUB* $\mathcal{B}_t = \{|b_1^t\rangle, \ldots, |b_d^t\rangle\}$ if for all $k \in [d]$ we have $|b_k^t\rangle = U_t|k\rangle$. In the next two chapters, we will be particularly concerned with two specific constructions of mutually unbiased bases. There exists a third construction based on Galois rings [KR04], which we do not consider here.

2.4.1 Latin squares

First, we consider MUBs based on mutually orthogonal Latin squares [WB05]. Informally, an $s \times s$ Latin square over the symbol set $[s]$ is an arrangement of elements of $[s]$ into an $s \times s$ square such that in each row and each column every element occurs exactly once. Let L_{ij} denote the entry in a Latin square in row i and column j. Two Latin squares L and L' are called mutually orthogonal if and only if $\{(L_{i,j}, L'_{i,j})|i, j \in [s]\} = \{(u, v)|u, v \in [s]\}$. Intuitively, this means that if we place one square on top of the other, and look at all pairs generated by the

overlaying elements, all possible pairs occur. An example is given in Figures 2.2 and 2.3 below. From any $s \times s$ Latin square we can obtain a basis for $\mathbb{C}^s \otimes \mathbb{C}^s$. First, we construct s of the basis vectors from the entries of the Latin square itself. Let

$$|v_{1,\ell}\rangle = \frac{1}{\sqrt{s}} \sum_{i,j \in [s]} E^L_{i,j}(\ell)|i,j\rangle,$$

where E^L is a predicate such that $E^L_{i,j}(\ell) = 1$ if and only if $L_{i,j} = \ell$. Note that for each ℓ we have exactly s pairs i, j such that $E_{i,j}(\ell) = 1$, because each element of $[s]$ occurs exactly s times in the Latin square. Secondly, from each such vector we obtain $s - 1$ additional vectors by adding successive rows of an $s \times s$ complex Hadamard matrix $H = (h_{ij})$ as coefficients to obtain the remaining $|v_{t,j}\rangle$ for $t \in [s]$, where $h_{ij} = \omega^{ij}$ with $i, j \in \{0, \ldots, s-1\}$ and $\omega = e^{2\pi i/s}$. Two additional MUBs can then be obtained in the same way from the two non-Latin squares where each element occurs for an entire row or column respectively. From each mutually orthogonal Latin square and these two extra squares which also satisfy the above orthogonality condition, we obtain one basis. This construction therefore gives MOLS$(s) + 2$ many MUBs. It is known that if $s = p^k$ is a prime power itself, we obtain $p^k + 1 \approx \sqrt{d}$ MUBs from this construction. Note, however, that there do exist many more MUBs in prime power dimensions, namely $d + 1$. If s is not a prime power, it is merely known that MOLS$(s) \geq s^{1/14.8}$ [WB05].

1	2	3
2	3	1
3	1	2

Figure 2.2: Latin Square (LS)

1	2	3
3	1	2
2	3	1

Figure 2.3: Mutually Orthogonal LS

As an example, consider the 3×3 Latin square depicted in Figure 2.2 and the 3×3 complex Hadamard matrix

$$H = \begin{pmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{pmatrix},$$

where $\omega = e^{2\pi i/3}$. First, we obtain vectors

$$\begin{aligned} |v_{1,1}\rangle &= \frac{1}{\sqrt{3}}(|1,1\rangle + |2,3\rangle + |3,2\rangle) \\ |v_{1,2}\rangle &= \frac{1}{\sqrt{3}}(|1,2\rangle + |2,1\rangle + |3,3\rangle) \\ |v_{1,3}\rangle &= \frac{1}{\sqrt{3}}(|1,3\rangle + |2,2\rangle + |3,1\rangle). \end{aligned}$$

With the help of H we obtain 3 additional vectors from the ones above. From the vector $|v_{1,1}\rangle$, for example, we obtain

$$\begin{aligned} |v_{1,1}\rangle &= \frac{1}{\sqrt{3}}(|1,1\rangle + |2,3\rangle + |3,2\rangle) \\ |v_{2,1}\rangle &= \frac{1}{\sqrt{3}}(|1,1\rangle + \omega|2,3\rangle + \omega^2|3,2\rangle) \\ |v_{3,1}\rangle &= \frac{1}{\sqrt{3}}(|1,1\rangle + \omega^2|2,3\rangle + \omega|3,2\rangle). \end{aligned}$$

This gives us basis $\mathcal{B} = \{|v_{t,\ell}\rangle | t, \ell \in [s]\}$ for $s = 3$. The construction of another basis follows in exactly the same way from a mutually orthogonal Latin square. The fact that two such squares L and L' are mutually orthogonal ensures that the resulting bases will be mutually unbiased. Indeed, suppose we are given another such basis, $\mathcal{B}' = \{|u_{t,\ell}\rangle | t, \ell \in [s]\}$ belonging to L'. We then have for any $\ell, \ell' \in [s]$ that $|\langle u_{1,\ell'}|v_{1,\ell}\rangle|^2 = |(1/s)\sum_{i,j\in[s]} E^{L'}_{i,j}(\ell')E^{L}_{i,j}(\ell)|^2 = 1/s^2$, as there exists exactly only one pair $\ell, \ell' \in [s]$ such that $E^{L'}_{i,j}(\ell')E^{L}_{i,j}(\ell) = 1$. Clearly, the same argument holds for the additional vectors derived from the complex Hadamard matrix.

2.4.2 Generalized Pauli matrices

The second construction we consider is based on the generalized Pauli matrices X_d and Z_d [BBRV02], defined by their actions on the computational basis $C = \{|0\rangle, \ldots, |d-1\rangle\}$ as follows:

$$\begin{aligned} X_d|k\rangle &= |k+1 \mod d\rangle \\ Z_d|k\rangle &= \omega^k|k\rangle, \ \forall |k\rangle \in C, \end{aligned}$$

where $\omega = e^{2\pi i/d}$. We say that $(X_d)^{a_1}(Z_d)^{b_1} \otimes \cdots \otimes (X_d)^{a_N}(Z_d)^{b_N}$ for $a_k, b_k \in \{0, \ldots, d-1\}$ and $k \in [N]$ is a *string of Pauli matrices*.

If d is a prime, it is known that the $d+1$ MUBs constructed first by Wootters and Fields [WF89] can also be obtained as the eigenvectors of the matrices $Z_d, X_d, X_dZ_d, X_dZ_d^2, \ldots, X_dZ_d^{d-1}$ [BBRV02]. If $d = p^k$ is a prime power, consider all d^2-1 possible strings of Pauli matrices excluding the identity and group them into sets $C_1, \ldots, C_{d+1}$ such that $|C_i| = d-1$ and $C_i \cap C_j = \{\mathbb{I}\}$ for $i \neq j$ and

all elements of C_i commute. Let B_i be the common eigenbasis of all elements of C_i. Then $B_1, \ldots, B_{d+1}$ are MUBs [BBRV02]. A similar result for $d = 2^k$ has also been shown in [LBZ02]. A special case of this construction are the three mutually unbiased bases in dimension $d = 2^k$ given by the unitaries $\mathbb{I}^{\otimes k}$,$H^{\otimes k}$ and $K^{\otimes k}$ (as defined on page 29) applied to the computational basis.

2.5 Conclusion

We summarized the most important elements of quantum theory that we need here. We refer to [Per93, NC00, Hay06] for more information about each topic. In Chapters 4 and 6 we investigate the two most striking aspects of quantum theory in detail: uncertainty relations and entanglement. But first, let's examine the case of state discrimination *with* additional post-measurement information.

Chapter 3

State discrimination with post-measurement information

In this chapter, we investigate an extension of the traditional state discrimination problem we encountered in Chapter 2.2: what if we are given some additional information *after* the measurement? Imagine that you are given a string x encoded in an unknown basis chosen from a known set of bases. You may perform any measurement, but you can only store at most q qubits of quantum information afterwards. Later on, you are told which basis was used. How well can you compute a function f of x, given the initial measurement outcome, the q qubits and the additional basis information?

3.1 Introduction

This question is of central importance for protocols in the bounded quantum storage model [DFSS05], which we encountered in Chapter 1. The security of such protocols rests on the realistic assumption that a dishonest player cannot store more than q qubits for long periods of time. In this model, even bit commitment and oblivious transfer can be implemented securely which is otherwise known to be impossible as we saw in Chapter 1. We formalize this general setting as a state discrimination problem: Here, we are given additional information about the state after the measurement or, more generally, after a quantum memory bound is applied. We prove general bounds on the success probability for any balanced function. We also show that storing just a *single* qubit allows you to compute any Boolean function perfectly when two bases are used. However, we also construct three bases for which you need to keep *all* qubits.

In general, we consider the following problem: Take an ensemble of quantum states, $\mathcal{E} = \{p_{yb}, \rho_{yb}\}$, with double indices $yb \in \mathcal{Y} \times \mathcal{B}$, and an integer $q \geq 0$. Suppose Alice sends Bob the state ρ_{yb}, where she alone knows indices y and b. Bob can perform any measurement on his system, but afterwards store at most

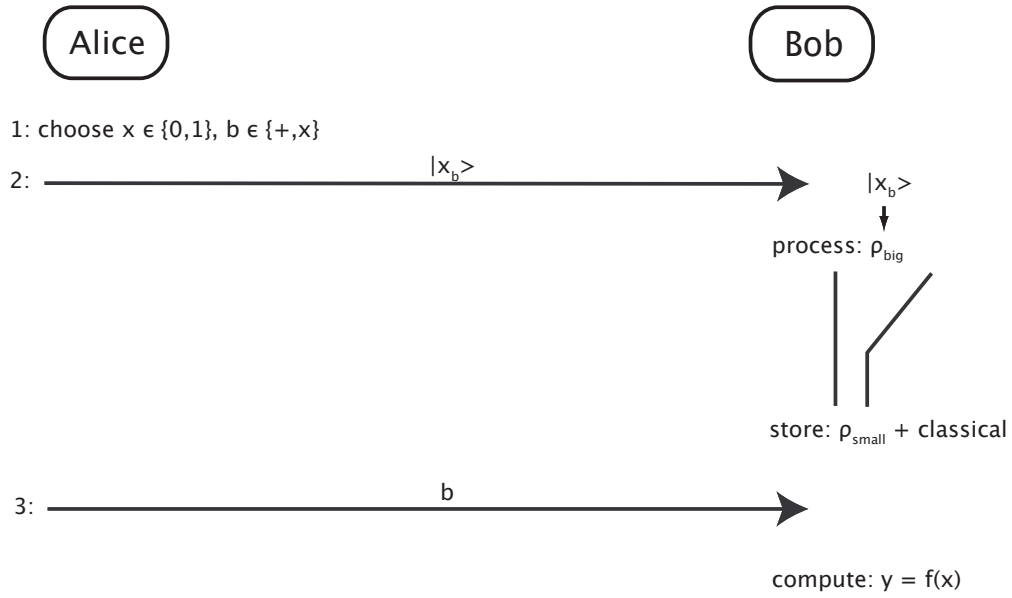


Figure 3.1: Using post-measurement information.

q qubits, and an unlimited amount of classical information. Afterwards, Alice tells him b. Bob's goal is now to approximate y as accurately as possible, which means that he has to make a guess $\hat{Y}$ that maximizes the success probability $P_{\text{succ}} = \sum_{yb} p_{yb} \Pr[\hat{Y} = y|\text{state } \rho_{yb}]$. For $|\mathcal{B}| = 1$, i.e., no post-measurement information is available, q is irrelevant and Bob's task is to discriminate among states ρ_y. This is the well-known state discrimination problem, which we encountered in Chapter 2.2, a problem studied since the early days of quantum information science. A particular case that isolates the aspect of the timing between measurements and side-information is one where for each fixed b, the states ρ_{yb} are mutually orthogonal: if Bob knew b, he could actually compute y perfectly. A special case of this problem is depicted in Figure 3.1. Here, Alice picks a string $x \in_R \{0,1\}^n$, and a basis $b \in \{+,\times\}$. She then encodes the string in the chosen basis and sends the resulting state to Bob. Bob's goal is now to determine $y = f(x)$ for a fixed function f. The states in this particular problem are thus of the form $\rho_{yb} = \sum_{x \in f^{-1}(y)} P_{X|B=b}(x) U_b |x\rangle\langle x| U_b^\dagger$, for a function $f : \mathcal{X} \to \mathcal{Y}$, and a set of mutually unbiased bases (MUBs) $\mathcal{B}$, given by the unitaries $U_0 = \mathbb{I}, U_1, \ldots, U_{|\mathcal{B}|-1}$ on a Hilbert space with basis $\{|x\rangle : x \in \mathcal{X}\}$, where the string x and a basis b are drawn from the distribution $P_{X,B}$. We mostly focus on this special case.

This problem also has an interpretation in terms of communication complexity. Suppose Alice is given b, and Bob is given the state ρ_{yb}. If classical communication is free, what is the minimum number of qubits Bob needs to send to Alice such that Alice learns y? Note that Bob needs to send exactly q qubits if and only if there exists a strategy for Bob to compute y in our task, while storing only q qubits.

3.1.1 Outline

In the following, we will close in on our problem in several stages. First, we briefly recall the case of state discrimination *without* any post-measurement information in Section 3.3. This enables us to draw comparisons later.

Second, in Section 3.4 we assume that Bob does receive post-measurement information, but has no quantum memory at all, i.e. $q = 0$. His goal then is to compute $f(x)$ given the classical outcome obtained by measuring $U_b|x\rangle$ and the later announcement of b. Clearly, a trivial strategy for Bob is to simply guess the basis, measure to obtain some string $\hat{x}$ and take $\hat{y} = f(\hat{x})$ as his answer. We thus want to find a better strategy. In particular, we will see that for any number of MUBs, any number of function outcomes, and any balanced f, Bob has a systematic advantage over guessing the basis, independent of $|\mathcal{X}|$. Furthermore, we show that for *any* Boolean f, Bob can succeed with probability at least $P_{succ} \geq 1/2 + 1/(2\sqrt{2})$ even if he cannot store any qubits at all. The latter result is relevant to the question of whether deterministic privacy amplification is possible in the protocols of [DFSS05]. Here, Alice uses two MUBs, and secretly chooses a function from a set of predetermined functions. She later tells Bob which function he should evaluate, together with the basis information b. Is it possible to use a fixed Boolean function instead? Our result shows that this is not possible.

It is interesting to consider when post-measurement information is useful for Bob, and how large his advantage is compared to the case where he does not receive any post-measurement information. To this end, we show how to phrase our problem as a semidefinite program (SDP), in the case where Bob has no quantum memory. In Section 3.4.2, we examine in detail the specific functions XOR and AND, for which we prove optimal bounds on Bob's success probability. In particular, the XOR on uniformly distributed strings of length n with two or three MUBs provides an extreme example of the usefulness of post-measurement information: We show that for the XOR function with n odd, $P_{\text{succ}} = 1/2 + 1/(2\sqrt{2})$. This is the same as Bob can achieve *without* the extra basis information. For even n, $P_{\text{succ}} = 1$ with the additional basis information. Here, P_{succ} jumps from 3/4 (without) to certainty (with basis information). The advantage that Bob gains can thus be maximal: *without* the post-measurement information, he can do no better than guessing the basis. However, *with* it, he can compute $y = f(x)$ perfectly. For even n, this was also observed in [DFSS05]. However, our analysis for odd n shows that the strategy for even n does *not* work for

any linear function as claimed in [DFSS05]. It remains an interesting question to find general conditions on the ensemble of states that determine how useful post-measurement information can be. We return to this question in Chapter 6.4.

Finally, we address the case where Bob does have quantum memory available. The question we are then interested in is: How large does this memory have to be so that Bob can compute y perfectly? In Section 3.5.1, we derive general conditions that determine when q qubits are sufficient. Our conditions impose a restriction on the rank of Bob's measurement operators and require that all such operators commute with the projector onto the support of ρ_{yb}, for all y and b. In particular, we give a general algebraic framework that allows us to determine q for any number of bases, functions and outcomes, in combination with an algorithm given in [KI02]. In Sections 3.5.2 and 3.5.3, we then consider two specific examples: First, we show that for *any* Boolean f and *any* two bases, storing just a *single* qubit is sufficient for Bob to compute $f(x)$ perfectly. The latter result again has implications to protocols in the bounded quantum storage model: for all existing protocols, deterministic privacy amplification is indeed hopeless. It turns out that part of this specific example also follows from known results derived for non-local games as we will discuss below. Surprisingly, things change dramatically when we are allowed to use three bases: We show how to construct three bases, such that for *any* balanced f Bob needs to keep *all* qubits in order to compute $f(x)$ perfectly!

3.1.2 Related work

In Chapter 2.2, we already examined the traditional setting of state discrimination *without* post-measurement information. Some of the tools we need below have found use in this setting as well. Many convex optimization problems can be solved using semidefinite programming. We refer to Appendix A for a introduction. Eldar [Eld03] and Eldar, Megretski and Verghese [EMV03] used semidefinite programming to solve state discrimination problems, which is one of the techniques we also use here. The square-root measurement [HW94] (also called pretty good measurement) is an easily constructed measurement to distinguish quantum states, however, it is only optimal for very specific sets of states [EF01, EMV04]. Mochon constructed specific pure state discrimination problems for which the square-root measurement is optimal [Moc07a]. We use a variant of the square-root measurement as well. Furthermore, our problem is related to the task of state filtering [BHH03, BHH05, BH05] and state classification [WY06]. Here, Bob's goal is to determine whether a given state is either one specific state or one of several other possible states, or, more generally, which subset of states a given state belongs to. Our scenario differs, because we deal with mixed states and Bob is allowed to use post-measurement information. Much more is known about pure state discrimination problems and the case of unambiguous state discrimination where we are not allowed to make an error. Since

we concentrate on mixed states, we refer to [BHH04] for an excellent survey on the extended field of state discrimination.

Regarding state discrimination with post-measurement information, special instances of the general problem have occurred in the literature under the heading "mean king's problem" [AE01, KR05], where the stress was on the usefulness of entanglement. Furthermore, it should be noted that prepare-and-measure quantum key distribution schemes of the BB84 type also lead to special cases of this problem: When considering optimal individual attacks, the eavesdropper is faced with the task of extracting maximal information about the raw key bits, encoded in an unknown basis, that she learns later during basis reconciliation.

Our result that one qubit of storage suffices for any Boolean function f demonstrates that storing quantum information can give an adversary a great advantage over storing merely classical information. It has also been shown in the context of randomness extraction with respect to a quantum adversary that storing quantum information can sometimes convey much more power to the adversary [GKK$^+$06].

3.2 Preliminaries

3.2.1 Notation and tools

We need the following notions. The Bell basis is given by the vectors $|\Phi^{\pm}\rangle = (|00\rangle \pm |11\rangle)/\sqrt{2}$ and $|\Psi^{\pm}\rangle = (|01\rangle \pm |10\rangle)/\sqrt{2}$. Furthermore, let $f^{-1}(y) = \{x \in \mathcal{X} | f(x) = y\}$. We say that a function f is balanced if and only if any element in the image of f is generated by equally many elements in the pre-image of f, i.e. there exists a $k \in \mathbb{N}$ such that $\forall y \in \mathcal{Y} : |f^{-1}(y)| = k$.

3.2.2 Definitions

We now give a more formal description of our problem. Let $\mathcal{Y}$ and $\mathcal{B}$ be finite sets and let $P_{YB} = \{p_{yb}\}$ be a probability distribution over $\mathcal{Y} \times \mathcal{B}$. Consider an ensemble of quantum states $\mathcal{E} = \{p_{yb}, \rho_{yb}\}$. We assume that $\mathcal{Y}$, $\mathcal{B}$, $\mathcal{E}$ and P_{YB} are known to both Alice and Bob. Suppose now that Alice chooses $yb \in \mathcal{Y} \times \mathcal{B}$ according to probability distribution P_{YB}, and sends ρ_{yb} to Bob. We can then define the tasks:

3.2.1. Definition. *State discRimination* (*STAR*($\mathcal{E}$)) is the following task for Bob. Given ρ_{yb}, determine y. He can perform any measurement on ρ_{yb} immediately upon receipt.

3.2.2. Definition. *State discRimination with Post-measurement Information* (PI_q-*STAR*($\mathcal{E}$)) is the following task for Bob. Given ρ_{yb}, determine y, where Bob can use the following sources of information in succession:

1. First, he can perform any measurement on ρ_{yb} immediately upon reception. Afterwards, he can store at most q qubits of quantum information about ρ_{yb}, and an unlimited amount of classical information.

2. After Bob's measurement, Alice announces b.

3. Then, he may perform any measurement on the remaining q qubits depending on b and the measurement outcome obtained in step 1.

We also say that *Bob succeeds at* $\mathrm{STAR}(\mathcal{E})$ *or* PI_q*-STAR*$(\mathcal{E})$ *with probability* p if and only if p is the average success probability $p = \sum_{yb} p_{yb} \Pr[\hat{Y} = y|\text{state } \rho_{yb}]$, where $\Pr[\hat{Y} = y|\text{state } \rho_{yb}]$ is the probability that Bob correctly determines y given ρ_{yb} in the case of STAR, and in addition using information sources 1, 2 and 3 in the case of PI-STAR.

Here, we are interested in the following special case: Consider a function $f : \mathcal{X} \to \mathcal{Y}$ between finite sets, and a set of mutually unbiased bases $\mathcal{B}$ as defined in Chapter 2, generated by a set of unitaries $U_0, U_1, \ldots, U_{|\mathcal{B}|-1}$ acting on a Hilbert space with basis $\{|x\rangle \mid x \in \mathcal{X}\}$. Take $|\Phi_b^x\rangle = U_b|x\rangle$. Let P_X and P_B be probability distributions over $\mathcal{X}$ and $\mathcal{B}$ respectively. We assume that f, $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{B}$, P_X, P_B, and the set of unitaries $\{U_b|b \in \mathcal{B}\}$ are known to both Alice and Bob. Suppose now that Alice chooses $x \in \mathcal{X}$ and $b \in \mathcal{B}$ independently according to probability distributions P_X and P_B respectively, and sends $|\Phi_b^x\rangle$ to Bob. Bob's goal is now to compute $y = f(x)$. We thus obtain an instance of our problem with states $\rho_{yb} = \sum_{x \in f^{-1}(y)} P_X(x)|\Phi_b^x\rangle\langle\Phi_b^x|$. We write $\mathrm{STAR}(f)$ and PI_q-$\mathrm{STAR}(f)$ to denote both problems in this special case. We concentrate on the case of mutually unbiased bases, as this case is most relevant to our initial goal of analyzing protocols for quantum cryptography in the bounded storage model [DFSS05].

Here, we make use of the basis set $\mathcal{B} = \{+, \times, \odot\}$, where $\mathcal{B}_+ = \{|0\rangle, |1\rangle\}$ is the computational basis, $\mathcal{B}_\times = \{\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\}$ is the Hadamard basis, and $\mathcal{B}_\odot = \{\frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle)\}$ is what we call the K-basis. The unitaries that give rise to these bases are $U_+ = \mathbb{I}$, $U_\times = H$ and $U_\odot = K$ with $K = (\mathbb{I} + i\sigma_x)/\sqrt{2}$ respectively. Recall from Chapter 2 that the Hadamard matrix is given by $H = \frac{1}{\sqrt{2}}(\sigma_x + \sigma_z)$, and that σ_x, σ_z and σ_y are the well-known Pauli matrices. We generally assume that Bob has no a priori knowledge about the outcome of the function and about the value of b. This means that b is chosen uniformly at random from $\mathcal{B}$, and, in the case of balanced functions, that Alice chooses x uniformly at random from $\mathcal{X}$. More generally, the distribution is uniform on all $f^{-1}(y)$ and such that each value $y \in \mathcal{Y}$ is equally likely.

3.2.3 A trivial bound: guessing the basis

Note that a simple strategy for Bob is to guess the basis, and then measure. This approach leads to a lower bound on the success probability for both STAR and

PI-STAR. In short:

3.2.3. Lemma. *Let $P_X(x) = \frac{1}{2^n}$ for all $x \in \{0,1\}^n$. Let $\mathcal{B}$ denote the set of bases. Then for any balanced function $f : \mathcal{X} \to \mathcal{Y}$ Bob succeeds at* STAR(f) *and* PI_0-STAR(f) *with probability at least*

$$p_{guess} = \frac{1}{|\mathcal{B}|} + \left(1 - \frac{1}{|\mathcal{B}|}\right)\frac{1}{|\mathcal{Y}|}.$$

Our goal is to beat this bound. We show that for PI-STAR, Bob can indeed do much better.

3.3 No post-measurement information

We first consider the standard case of state discrimination. Here, Alice does not supply Bob with any additional post-measurement information. Instead, Bob's goal is to compute $y = f(x)$ immediately. This analysis enables us to gain interesting insights into the usefulness of post-measurement information later.

3.3.1 Two simple examples

We now examine two simple one-qubit examples of a state discrimination problem, which we make use of later on. Here, Bob's goal is to learn the value of a bit which has been encoded in two or three mutually unbiased bases while he does not know which basis has been used.

3.3.1. Lemma. *Let $x \in \{0,1\}$, $P_X(x) = \frac{1}{2}$ and $f(x) = x$. Let $\mathcal{B} = \{+, \times\}$ with $U_+ = \mathbb{I}$ and $U_\times = H$. Then Bob succeeds at* STAR(f) *with probability at most*

$$p = \frac{1}{2} + \frac{1}{2\sqrt{2}}.$$

There exists a strategy for Bob that achieves p.

Proof. The probability of success follows from Theorem 2.2.2 with $\rho_0 = \frac{1}{2}(|0\rangle\langle 0| + H|0\rangle\langle 0|H)$, $\rho_1 = \frac{1}{2}(|1\rangle\langle 1| + H|1\rangle\langle 1|H)$ and $q = 1/2$. □

3.3.2. Lemma. *Let $x \in \{0,1\}$, $P_X(x) = \frac{1}{2}$ and $f(x) = x$. Let $\mathcal{B} = \{+, \times, \odot\}$ with $U_+ = \mathbb{I}$, $U_\times = H$ and $U_\odot = K$. Then Bob succeeds at* STAR(f) *with probability at most*

$$p = \frac{1}{2} + \frac{1}{2\sqrt{3}}.$$

There exists a strategy for Bob that achieves p.

Proof. The proof is identical to that of Lemma 3.3.1 using $\rho_0 = \frac{1}{3}(|0\rangle\langle 0| + H|0\rangle\langle 0|H + K|0\rangle\langle 0|K^\dagger)$, $\rho_1 = \frac{1}{3}(|1\rangle\langle 1| + H|1\rangle\langle 1|H + K|1\rangle\langle 1|K^\dagger)$, and $q = 1/2$. □

3.3.2 An upper bound for all Boolean functions

We now show that for any Boolean function f and any number of mutually unbiased bases, the probability that Bob succeeds at STAR(f) is very limited.

3.3.3. THEOREM. *Let $|\mathcal{Y}| = 2$ and let f be a balanced function. Let $\mathcal{B}$ be a set of mutually unbiased bases. Then Bob succeeds at* STAR(f) *with probability at most*

$$p = \frac{1}{2} + \frac{1}{2\sqrt{|\mathcal{B}|}}.$$

In particular, for $|\mathcal{B}| = 2$ we obtain $(1 + 1/\sqrt{2})/2 \approx 0.853$; for $|\mathcal{B}| = 3$, we obtain $(1 + 1/\sqrt{3})/2 \approx 0.789$.

Proof. The probability of success is given by Theorem 2.2.2 where for $y \in \{0, 1\}$

$$\rho_y = \frac{1}{2^{n-1}|\mathcal{B}|} \sum_{b=0}^{|\mathcal{B}|-1} P_{yb},$$

with $P_{yb} = \sum_{x \in f^{-1}(y)} U_b |x\rangle\langle x| U_b^\dagger$. Using the Cauchy-Schwarz inequality we can show that

$$\|\rho_0 - \rho_1\|_1^2 = [\mathrm{Tr}(|\rho_0 - \rho_1|\mathbb{I})]^2 \leq \mathrm{Tr}[(\rho_0 - \rho_1)^2]\mathrm{Tr}[\mathbb{I}^2] = 2^n \mathrm{Tr}[(\rho_0 - \rho_1)^2], \quad (3.1)$$

or

$$\|\rho_0 - \rho_1\|_1 \leq \sqrt{2^n \mathrm{Tr}[(\rho_0 - \rho_1)^2]}.$$

A simple calculation shows that

$$\mathrm{Tr}[(\rho_0 - \rho_1)^2] = \frac{4}{2^n |\mathcal{B}|}.$$

The theorem follows from the previous equation, together with Theorem 2.2.2 and Eq. (3.1). □

3.3.3 AND function

One of the simplest functions to consider is the AND function. Recall, that we always assume that Bob has no a priori knowledge about the outcome of the function. In the case of the AND, this means that we are considering a very specific prior: with probability 1/2 Alice will choose the only string x for which AND(x) = 1. Without any post-measurement information, Bob can already compute the AND quite well.

3.3.4. THEOREM. *Let* $P_X(x) = 1/(2(2^n-1))$ *for all* $x \in \{0,1\}^n \setminus \{1\ldots1\}$ *and* $P_X(1\ldots1) = \frac{1}{2}$. *Let* $\mathcal{B} = \{+,\times\}$ *with* $U_+ = \mathbb{I}^{\otimes n}$, $U_\times = H^{\otimes n}$ *and* $P_B(+) = P_B(\times) = 1/2$. *Then Bob succeeds at* STAR(AND) *with probability at most*

$$p = \begin{cases} \frac{1}{2} + \frac{1}{2\sqrt{2}} & \text{if } n = 1, \\ 1 - \frac{1}{2(2^n-1)} & \text{if } n \geq 2. \end{cases} \tag{3.2}$$

There exists a strategy for Bob that achieves p.

Proof. Let $|c_1\rangle = |1\rangle^{\otimes n}$ and $|h_1\rangle = [H|1\rangle]^{\otimes n}$. Eq. (3.2) is obtained by substituting

$$\begin{aligned} \rho_0 &= \frac{1}{2}\left[\frac{\mathbb{I} - |c_1\rangle\langle c_1|}{2^n-1} + \frac{\mathbb{I} - |h_1\rangle\langle h_1|}{2^n-1}\right], \\ \rho_1 &= \frac{|c_1\rangle\langle c_1| + |h_1\rangle\langle h_1|}{2}, \end{aligned}$$

and $q = 1/2$ in Theorem 2.2.2. □

In Theorem 3.4.3, we show an optimal bound for the case that Bob does indeed receive the extra information. By comparing the previous equation with Eq. (3.4) later on, we can see that for $n = 1$ announcing the basis does not help. However, for $n > 1$ we will observe an improvement of $[2(2^n + 2^{n/2} - 2)]^{-1}$.

3.3.4 XOR function

The XOR function provides an example of a Boolean function where we observe both the largest advantage as well as the smallest advantage in receiving post-measurement information: For strings of even length we show that without the extra information Bob can never do better than guessing the basis. For strings of odd length, however, he can do quite a bit better. Interestingly, it turns out that in this case the post-measurement information is completely useless to him. We first investigate how well Bob does at STAR(XOR) for two bases:

3.3.5. THEOREM. *Let* $P_X(x) = \frac{1}{2^n}$ *for all* $x \in \{0,1\}^n$. *Let* $\mathcal{B} = \{+,\times\}$ *with* $U_+ = \mathbb{I}^{\otimes n}$, $U_\times = H^{\otimes n}$ *and* $P_B(+) = P_B(\times) = 1/2$. *Then Bob succeeds at* STAR(XOR) *with probability at most*

$$p = \begin{cases} \frac{3}{4} & \text{if } n \text{ is even}, \\ \frac{1}{2}\left(1 + \frac{1}{\sqrt{2}}\right) & \text{if } n \text{ is odd}. \end{cases}$$

There exists a strategy for Bob that achieves p.

Proof. Our proof works by induction on n. The case of $n = 1$ was addressed in Lemma 3.3.1. Now, consider $n = 2$: Let $\sigma_0^{(2)} = \frac{1}{2}(\rho_{0+}^{(2)} + \rho_{0\times}^{(2)})$ and $\sigma_1^{(2)} = \frac{1}{2}(\rho_{1+}^{(2)} + \rho_{1\times}^{(2)})$, where $\rho_{0+}^{(2)}$ and $\rho_{1+}^{(2)}$ are defined as $\rho_{yb}^{(n)} = \frac{1}{2^{n-1}} \sum_{x\in\{0,1\}^n, x\in \mathrm{XOR}^{-1}(y)} U_b|x\rangle\langle x|U_b^\dagger$ with $y \in \{0,1\}$ and $b \in \mathcal{B} = \{+,\times\}$. A straightforward calculation shows that $\|\sigma_0^{(2)} - \sigma_1^{(2)}\|_1 = 1$.

We now show that the trace distance does not change when we go from strings of length n to strings of length $n+2$: Note that we can write

$$\begin{aligned}
\rho_{0+}^{(n+2)} &= \frac{1}{2}(\rho_{0+}^{(n)} \otimes \rho_{0+}^{(2)} + \rho_{1+}^{(n)} \otimes \rho_{1+}^{(2)}) \\
\rho_{0\times}^{(n+2)} &= \frac{1}{2}(\rho_{0\times}^{(n)} \otimes \rho_{0\times}^{(2)} + \rho_{1\times}^{(n)} \otimes \rho_{1\times}^{(2)}) \\
\rho_{1+}^{(n+2)} &= \frac{1}{2}(\rho_{0+}^{(n)} \otimes \rho_{1+}^{(2)} + \rho_{1+}^{(n)} \otimes \rho_{0+}^{(2)}) \\
\rho_{1\times}^{(n+2)} &= \frac{1}{2}(\rho_{0\times}^{(n)} \otimes \rho_{1\times}^{(1)} + \rho_{1\times}^{(n)} \otimes \rho_{0\times}^{(2)}).
\end{aligned} \tag{3.3}$$

Let $\sigma_0^{(n)} = \frac{1}{2}(\rho_{0+}^{(n)} + \rho_{0\times}^{(n)})$ and $\sigma_1^{(n)} = \frac{1}{2}(\rho_{1+}^{(n)} + \rho_{1\times}^{(n)})$. A small calculation shows that

$$\begin{aligned}
\sigma_0^{(n+2)} - \sigma_1^{(n+2)} = \frac{1}{8}\big[&(\rho_{0+}^{(n)} + \rho_{0\times}^{(n)} - \rho_{1+}^{(n)} - \rho_{1\times}^{(n)}) \otimes |\Phi^+\rangle\langle\Phi^+| \\
&- (\rho_{0+}^{(n)} + \rho_{0\times}^{(n)} - \rho_{1+}^{(n)} - \rho_{1\times}^{(n)}) \otimes |\Psi^-\rangle\langle\Psi^-| \\
&+ (\rho_{0+}^{(n)} + \rho_{1\times}^{(n)} - \rho_{1+}^{(n)} - \rho_{0\times}^{(n)}) \otimes |\Phi^-\rangle\langle\Phi^-| \\
&- (\rho_{0+}^{(n)} + \rho_{1\times}^{(n)} - \rho_{1+}^{(n)} - \rho_{0\times}^{(n)}) \otimes |\Psi^+\rangle\langle\Psi^+|\big]
\end{aligned}$$

We then get that

$$\|\sigma_0^{(n+2)} - \sigma_1^{(n+2)}\|_1 = \frac{1}{2}\left(\|\sigma_0^{(n)} - \sigma_1^{(n)}\|_1 + \|\tilde{\sigma}_0^{(n)} - \tilde{\sigma}_1^{(n)}\|_1\right),$$

where $\tilde{\sigma}_0^{(n)} = \frac{1}{2}(\rho_{0+}^{(n)} + \rho_{1\times}^{(n)})$ and $\tilde{\sigma}_1^{(n)} = \frac{1}{2}(\rho_{1+}^{(n)} + \rho_{0\times}^{(n)})$. Consider the unitary $U = \sigma_x^{\otimes n}$ if n is odd, and $U = \sigma_x^{\otimes n-1} \otimes \mathbb{I}$ if n is even. It is easy to verify that $\sigma_0^{(n)} = U\tilde{\sigma}_0^{(n)}U^\dagger$ and $\sigma_1^{(n)} = U\tilde{\sigma}_1^{(n)}U^\dagger$. We thus have that $\|\sigma_0^{(n)} - \sigma_1^{(n)}\|_1 = \|\tilde{\sigma}_0^{(n)} - \tilde{\sigma}_1^{(n)}\|_1$ and therefore

$$\|\sigma_0^{(n+2)} - \sigma_1^{(n+2)}\|_1 = \|\sigma_0^{(n)} - \sigma_1^{(n)}\|_1.$$

It then follows from Helstrom's Theorem 2.2.2 that the maximum probability to distinguish $\sigma_0^{(n+2)}$ from $\sigma_1^{(n+2)}$ and thus compute the XOR of the $n+2$ bits is given by

$$\frac{1}{2} + \frac{\|\sigma_0^{(n)} - \sigma_1^{(n)}\|_1}{4},$$

which gives the claimed result. □

A similar argument is possible, if we use three mutually unbiased bases. Intuitively, one might expect Bob's chance of success to drop as we had more bases. Interestingly, however, we obtain the same bound of 3/4 if n is even.

3.3.6. THEOREM. *Let $P_X(x) = \frac{1}{2^n}$ for all $x \in \{0,1\}^n$. Let $\mathcal{B} = \{+, \times, \odot\}$ with $U_+ = \mathbb{I}^{\otimes n}$, $U_\times = H^{\otimes n}$, and $U_\odot = K^{\otimes n}$ with $P_B(+) = P_B(\times) = P_B(\odot) = 1/3$. Then Bob succeeds at* STAR(XOR) *with probability at most*

$$p = \begin{cases} \frac{3}{4} & \text{if } n \text{ is even,} \\ \frac{1}{2}\left(1 + \frac{1}{\sqrt{3}}\right) & \text{if } n \text{ is odd.} \end{cases}$$

There exists a strategy for Bob that achieves p.

Proof. Our proof is very similar to the case of only 2 mutually unbiased bases. The case of $n = 1$ follows from Lemma 3.3.2. This time, we have for $n = 2$: $\sigma_0^{(2)} = \frac{1}{3}(\rho_{0+}^{(2)} + \rho_{0\times}^{(2)} + \rho_{0\odot}^{(2)})$ and $\sigma_1^{(2)} = \frac{1}{3}(\rho_{1+}^{(2)} + \rho_{1\times}^{(2)} + \rho_{1\odot}^{(2)})$. We have $\|\sigma_0^{(2)} - \sigma_1^{(2)}\|_1 = 1$.

We again show that the trace distance does not change when we go from strings of length n to strings of length $n+2$. We use the definitions from Eq. (3.3) and let

$$\begin{aligned} \rho_{0\odot}^{(n+2)} &= \frac{1}{2}(\rho_{0\odot}^{(n)} \otimes \rho_{0\odot}^{(2)} + \rho_{1\odot}^{(n)} \otimes \rho_{1\odot}^{(2)}), \\ \rho_{1\odot}^{(n+2)} &= \frac{1}{2}(\rho_{0\odot}^{(n)} \otimes \rho_{1\odot}^{(2)} + \rho_{1\odot}^{(n)} \otimes \rho_{0\odot}^{(2)}). \end{aligned}$$

We can compute

$$\begin{aligned} \sigma_0^{(n+2)} - \sigma_1^{(n+2)} = \frac{1}{4}\big[&(\bar{\sigma}_1^{(n)} - \bar{\sigma}_0^{(n)}) \otimes |\Phi^+\rangle\langle\Phi^+| \\ &- (\hat{\sigma}_1^{(n)} - \hat{\sigma}_0^{(n)}) \otimes |\Psi^-\rangle\langle\Psi^-| \\ &+ (\tilde{\sigma}_1^{(n)} - \tilde{\sigma}_0^{(n)}) \otimes |\Phi^-\rangle\langle\Phi^-| \\ &- (\sigma_1^{(n)} - \sigma_0^{(n)}) \otimes |\Psi^+\rangle\langle\Psi^+|\big], \end{aligned}$$

where $\bar{\sigma}_1^{(n)} = (\rho_{0+}^{(n)} + \rho_{0\times}^{(n)} + \rho_{1\odot}^{(n)})/3$, $\bar{\sigma}_0^{(n)} = (\rho_{1+}^{(n)} + \rho_{1\times}^{(n)} + \rho_{0\odot}^{(n)})/3$, $\hat{\sigma}_1^{(n)} = (\rho_{0+}^{(n)} + \rho_{1\times}^{(n)} + \rho_{0\odot}^{(n)})/3$, $\hat{\sigma}_0^{(n)} = (\rho_{1+}^{(n)} + \rho_{0\times}^{(n)} + \rho_{1\odot}^{(n)})/3$, $\tilde{\sigma}_0^{(n)} = (\rho_{1+}^{(n)} + \rho_{0\times}^{(n)} + \rho_{0\odot}^{(n)})/3$, and $\tilde{\sigma}_0^{(n)} = (\rho_{0+}^{(n)} + \rho_{1\times}^{(n)} + \rho_{1\odot}^{(n)})/3$. Consider the unitaries $\bar{U} = \sigma_y^{\otimes n}$, $\hat{U} = \sigma_x^{\otimes n}$, and $\tilde{U} = \sigma_z^{\otimes n}$ if n is odd, and $\bar{U} = \sigma_y^{\otimes n-1} \otimes \mathbb{I}$, $\hat{U} = \sigma_x^{\otimes n-1} \otimes \mathbb{I}$, and $\tilde{U} = \sigma_z^{\otimes n-1} \otimes \mathbb{I}$ if n is even. It is easily verified that $\sigma_0^{(n)} = \bar{U}\bar{\sigma}_0^{(n)}\bar{U}^\dagger$, $\sigma_1^{(n)} = \bar{U}\bar{\sigma}_1^{(n)}\bar{U}^\dagger$, $\sigma_0^{(n)} = \hat{U}\hat{\sigma}_0^{(n)}\hat{U}^\dagger$, $\sigma_1^{(n)} = \hat{U}\hat{\sigma}_1^{(n)}\hat{U}^\dagger$, $\sigma_0^{(n)} = \tilde{U}\tilde{\sigma}_0^{(n)}\tilde{U}^\dagger$, and $\sigma_1^{(n)} = \tilde{U}\tilde{\sigma}_1^{(n)}\tilde{U}^\dagger$. We then get that

$$\|\sigma_0^{(n+2)} - \sigma_1^{(n+2)}\|_1 = \|\sigma_0^{(n)} - \sigma_1^{(n)}\|_1,$$

from which the claim follows. □

Surprisingly, if Bob does have some a priori knowledge about the outcome of the XOR the problem becomes much harder for Bob. By expressing the states in the Bell basis and using Helstrom's result, it is easy to see that if Alice chooses $x \in \{0,1\}^2$ such that with probability q, $\mathrm{XOR}(x) = 0$, and with probability $(1-q)$, $\mathrm{XOR}(x) = 1$, Bob's probability of learning $\mathrm{XOR}(x)$ correctly is minimized for $q = 1/3$. In that case, Bob succeeds with probability at most $2/3$, which can be achieved by the trivial strategy of ignoring the state he received and always outputting 1. This is an explicit example where making a measurement does not help in state discrimination. It has previously been noted by Hunter [Hun03] that such cases can exist in mixed-state discrimination.

3.4 Using post-measurement information

We are now ready to advance to the core of our problem. We first consider the case where Bob does receive post-measurement information, but still has no quantum memory at his disposal. Consider an instance of PI_0-STAR with a function $f : \mathcal{X} \to \mathcal{Y}$ and $m = |\mathcal{B}|$ bases, and some priors P_X and P_B on the sets $\mathcal{X}$ and $\mathcal{B}$. If Bob cannot store any quantum information, all his nontrivial actions are contained in the first measurement, which must equip him with possible outputs $o_i \in \mathcal{Y}$ for each basis $i = 1, \ldots, m$. In other words, his most general strategy is a POVM with $|\mathcal{Y}|^m$ outcomes, each labeled by the strings $o_1, \ldots, o_m$ for $o_i \in \mathcal{Y}$ and $m = |\mathcal{B}|$. Once Alice has announced b, Bob outputs $\hat{Y} = o_b$. Here we first prove a general lower bound on the usefulness of post-measurement information that beats the guessing bound. Then, we analyze in detail the AND and the XOR function on n bits.

3.4.1 A lower bound for balanced functions

We first give a lower bound on Bob's success probability for any balanced function and any number of mutually unbiased bases, by constructing an explicit measurement that achieves it. Without loss of generality, we assume in this section that $\mathcal{B} = \{0, \ldots, m-1\}$, as otherwise we could consider a lexicographic ordering of $\mathcal{B}$.

3.4.1. Theorem. *Let $f : \mathcal{X} \to \mathcal{Y}$ be a balanced function, and let P_X and P_B be the uniform distributions over $\mathcal{X}$ and $\mathcal{B}$ respectively. Let the set of unitaries $\{U_b | b \in \mathcal{B}\}$ give rise to $|\mathcal{B}|$ mutually unbiased bases, and choose an encoding such that $\forall x, x' \in \mathcal{X} : \langle x|x'\rangle = \delta_{xx'}$. Then Bob succeeds at* PI_0-$\mathrm{STAR}(f)$ *with*

probability at least

$$p = p_{guess} + \begin{cases} \frac{|\mathcal{Y}|-1}{|\mathcal{Y}|(|\mathcal{Y}|+3)} & \text{if } m = 2, \\ \frac{4(|\mathcal{Y}|^2-1)}{3|\mathcal{Y}|(2+|\mathcal{Y}|(|\mathcal{Y}|+6))} & \text{if } m = 3, \\ -\frac{2}{2|\mathcal{Y}|} + \frac{2(|\mathcal{Y}|+m-1)}{|\mathcal{Y}|^2+3|\mathcal{Y}|(m-1)+m^2-3m+2} & \text{if } m \geq 4. \end{cases}$$

where p_{guess} is the probability that Bob can achieve by guessing the basis as given in Lemma 3.2.3. In particular, we always have $p > p_{guess}$.

Proof. Our proof works by constructing a square-root type measurement that achieves the lower bound. As explained above, Bob's strategy for learning $f(x)$ is to perform a measurement with $|\mathcal{Y}|^m$ possible outcomes, labeled by the strings $o_1, \ldots, o_m$ for $o_i \in \mathcal{Y}$ and $m = |\mathcal{B}|$. Once Alice has announced b, Bob outputs $f(x) = o_b$.

Take the projector $P_{yb} = \sum_{x \in f^{-1}(y)} |\Phi_b^x\rangle\langle\Phi_b^x|$ and $\rho_{yb} = \frac{1}{k} P_{yb}$, where $k = |f^{-1}(y)| = |\mathcal{X}|/|\mathcal{Y}|$. Let $M_{o_1,\ldots,o_m}$ denote the measurement operator corresponding to outcome $o_1, \ldots, o_m$. Note that outcome $o_1, \ldots, o_m$ is the correct outcome for input state ρ_{yb} if and only if $o_b = y$. We can then write Bob's probability of success as

$$\frac{1}{m|\mathcal{Y}|} \sum_{o_1,\ldots,o_m \in \mathcal{Y}} \mathrm{Tr}\left(M_{o_1,\ldots,o_m} \left(\sum_{b \in \mathcal{B}} \rho_{o_b b} \right) \right).$$

We make use of the following measurement:

$$M_{o_1,\ldots,o_m} = S^{-\frac{1}{2}} \left(\sum_{b \in \mathcal{B}} P_{o_b b} \right)^3 S^{-\frac{1}{2}}, \text{ with } S = \sum_{o_1,\ldots,o_m \in \mathcal{Y}} \left(\sum_{b \in \mathcal{B}} P_{o_b b} \right)^3.$$

Clearly, we have $\sum_{o_1,\ldots,o_m \in \mathcal{Y}} M_{o_1,\ldots,o_m} = \mathbb{I}$ and $\forall o_1, \ldots, o_m \in \mathcal{Y} : M_{o_1,\ldots,o_m} \geq 0$ by construction and thus we indeed have a valid measurement. We first show that $S = c_m \mathbb{I}$:

$$\begin{aligned} S &= \sum_{o_1,\ldots,o_m \in \mathcal{Y}} \left(\sum_{b \in \mathcal{B}} P_{o_b b} \right)^3 \\ &= \sum_{o_1,\ldots,o_m \in \mathcal{Y}} \sum_{b,b',b'' \in \mathcal{B}} P_{o_b b} P_{o_{b'} b'} P_{o_{b''} b''} \\ &= \sum_{o_1,\ldots,o_m \in \mathcal{Y}} \Bigg(\sum_{b} P_{o_b b} + 2 \sum_{bb', b \neq b'} P_{o_b b} P_{o_{b'} b'} \\ &\qquad + \sum_{bb', b \neq b'} P_{o_b b} P_{o_{b'} b'} P_{o_b b} + \sum_{bb'b'', b \neq b', b \neq b'', b' \neq b''} P_{o_b b} P_{o_{b'} b'} P_{o_{b''} b''} \Bigg) \\ &= \left[m|\mathcal{Y}|^{m-1} + 2m(m-1)|\mathcal{Y}|^{m-2} + m(m-1)|\mathcal{Y}|^{m-2} + m(m-1)(m-2)|\mathcal{Y}|^{m-3} \bar{\delta}_{2m} \right] \mathbb{I}, \end{aligned}$$

where $\bar{\delta}_{2m} = 1-\delta_{2m}$ and we have used the definition that for any b, $P_{o_b b}$ is a projector and $\sum_{x\in\mathcal{X}} |\Phi_b^x\rangle\langle\Phi_b^x| = \mathbb{I}$ which gives $\sum_{o_i\in\mathcal{Y}} P_{o_i b_i} = \sum_{o_i\in\mathcal{Y}} \sum_{x\in f^{-1}(y)} |\Phi_b^x\rangle\langle\Phi_b^x| = \mathbb{I}$. We can then write Bob's probability of success using this particular measurement as

$$\frac{1}{c_m k m |\mathcal{Y}|} \sum_{o_1,\ldots,o_m\in\mathcal{Y}} \mathrm{Tr}\left(\left(\sum_{b\in\mathcal{B}} P_{o_b b}\right)^4\right).$$

It remains to evaluate this expression. Using the circularity of the trace, we obtain

$$\begin{aligned}
&\sum_{o_1,\ldots,o_m\in\mathcal{Y}} \mathrm{Tr}\left(\left(\sum_{b\in\mathcal{B}} P_{o_b b}\right)^4\right) \\
&= \sum_{o_1,\ldots,o_m\in\mathcal{Y}} \mathrm{Tr}\Bigg(\sum_b P_{o_b b} + 6\sum_{bb',b\neq b'} P_{o_b b}P_{o_{b'}b'} \\
&\quad + 4\sum_{bb'b'',b\neq b',b\neq b'',b'\neq b''} P_{o_b b}P_{o_{b'}b'}P_{o_{b''}b''} + 2\sum_{bb'b'',b\neq b',b\neq b'',b'\neq b''} P_{o_b b}P_{o_{b'}b'}P_{o_b b}P_{o_{b''}b''} \\
&\quad + \sum_{bb'b''\tilde{b},b\neq b',b\neq b'',b\neq\tilde{b},b'\neq b'',b'\neq\tilde{b},b''\neq\tilde{b}} P_{o_b b}P_{o_{b'}b'}P_{o_{b''}b''}P_{o_{\tilde{b}}\tilde{b}} + \sum_{bb',b\neq b'} P_{o_b b}P_{o_{b'}b'}P_{o_b b}P_{o_{b'}b'}\Bigg) \\
&\geq \big[m|\mathcal{Y}|^{m-1} + 6m(m-1)|\mathcal{Y}|^{m-2} + 6m(m-1)(m-2)|\mathcal{Y}|^{m-3}\bar{\delta}_{2m} \\
&\quad + m(m-1)(m-2)(m-3)|\mathcal{Y}|^{t(m-4)}\bar{\delta}_{2m}\bar{\delta}_{3m}\big]\mathrm{Tr}(\mathbb{I}) + m(m-1)|\mathcal{Y}|^{m-2}k,
\end{aligned}$$

where we have again used the assumption that for any b, $P_{o_b b}$ is a projector and $\sum_{x\in\mathcal{X}} |\Phi_b^x\rangle\langle\Phi_b^x| = \mathbb{I}$ with $\mathrm{Tr}(\mathbb{I}) = |\mathcal{X}|$. For the last term we have used the following: Note that $\mathrm{Tr}(P_{o_b b}P_{o_{b'}b'}) = k^2/|\mathcal{X}|$, because we assumed mutually unbiased bases. Let $r = \mathrm{rank}(P_{o_b b}P_{o_{b'}b'})$. Using Cauchy-Schwarz, we can then bound $\mathrm{Tr}((P_{o_b b}P_{o_{b'}b'})^2) = \sum_i^r \lambda_i(P_{o_b b}P_{o_{b'}b'})^2 \geq k^4/(|\mathcal{X}|^2 r) \geq k^3/|\mathcal{X}|^2 = k/|\mathcal{Y}|^2$, where $\lambda_i(A)$ is the i-th eigenvalue of a matrix A, by noting that $r \leq k$ since $\mathrm{rank}(P_{o_b b}) = \mathrm{rank}(P_{o'_b b'}) = k$. Putting things together we obtain

$$p \geq \frac{1}{c_m m}\left[G_m(1) + \left(6 + \frac{1}{|\mathcal{Y}|}\right)G_m(2) + 6G_m(3) + G_m(4)\right],$$

where $m = |\mathcal{B}|$, $c_m = G_m(1) + 3G_m(2) + G_m(3)$ and function $G_m : \mathbb{N} \to \mathbb{N}$ defined as $G_m(i) = \frac{m!}{(m-i)!}|\mathcal{Y}|^{m-i}\prod_{j=2}^{i-1}\bar{\delta}_{mj}$. This expression can be simplified to obtain the claimed result. □

Note that we have only used the assumption that Alice uses mutually unbiased bases in the very last step to say that $\mathrm{Tr}(P_{o_b b}P_{o_{b'}b'}) = k^2/|\mathcal{X}|$. One could generalize our argument to other cases by evaluating $\mathrm{Tr}(P_{o_b b}P_{o_{b'}b'})$ approximately.

In the special case $m = |\mathcal{Y}| = 2$ (i.e. binary function, with two bases) we obtain:

3.4.2. COROLLARY. *Let $f : \{0,1\}^n \to \{0,1\}$ be a balanced function and let $P_X(x) = 2^{-n}$ for all $x \in \{0,1\}^n$. Let $\mathcal{B} = \{0,1\}$ with $U_0 = \mathbb{I}^{\otimes n}$, $U_1 = H^{\otimes n}$ and $P_B(0) = P_B(1) = 1/2$. Then Bob succeeds at* PI_0-STAR(f) *with probability $p \geq 0.85$.*

Observe that this almost attains the upper bound of $\approx .853$ of Lemma 3.3.1 in the case of no post-measurement information. In Section 3.5.2 we show that indeed this bound can always be achieved when post-measurement information is available.

It is perhaps interesting to note that our general bound depends only on the number of function values $|\mathcal{Y}|$ and the number of bases m. The number of function inputs $|\mathcal{X}|$ itself does not play a direct role.

3.4.2 Optimal bounds for the AND and XOR function

We now show that for some specific functions, the probability of success can even be much larger. We hereby concentrate on the case where Alice uses two or three mutually unbiased bases to encode her input. Our proofs thereby lead to explicit measurements. In the following, we again assume that Bob has no a priori knowledge of the function value. It turns out that the optimal measurement directly lead us to the essential idea underlying our algebraic framework of Section 3.5.1.

AND function

3.4.3. THEOREM. *Let $P_X(x) = 1/(2(2^n-1))$ for all $x \in \{0,1\}^n \setminus \{1\ldots1\}$ and $P_X(1\ldots1) = \frac{1}{2}$. Let $\mathcal{B} = \{+,\times\}$ with $U_+ = \mathbb{I}^{\otimes n}$, $U_\times = H^{\otimes n}$ and $P_B(+) = P_B(\times) = 1/2$. Then Bob succeeds at* PI_0-STAR(AND) *with probability at most*

$$p = \frac{1}{2}\left[2 + \frac{1}{2^n + 2^{n/2} - 2} - \frac{1}{2^n - 1}\right]. \tag{3.4}$$

There exists a strategy for Bob that achieves p.

Proof. To learn the value of AND(x), Bob uses the same strategy as in Section 3.4.1: he performs a measurement with 4 possible outcomes, labeled by the strings $o_+, o_\times$ with $o_+, o_\times \in \{0,1\}$. Once Alice has announced her basis choice $b \in \{+,\times\}$, Bob outputs $\mathrm{AND}(x) = o_b$. Note that without loss of generality we can assume that Bob's measurement has only 4 outcomes, i.e. Bob only stores 2 bits of classical information because he will only condition his answer on the value of b later on.

Following the approach in the last section, we can write Bob's optimal probability of success as a semidefinite program:

$$\begin{array}{ll} \text{maximize} & \frac{1}{4}\sum_{o_+,o_\times\in\{0,1\}} \mathrm{Tr}[b_{o_+o_\times} M_{o_+o_\times}] \\ \text{subject to} & \forall o_+, o_\times \in \{0,1\} : M_{o_+o_\times} \geq 0, \\ & \sum_{o_+,o_\times\in\{0,1\}} M_{o_+o_\times} = \mathbb{I}, \end{array}$$

where

$$b_{00} = \rho_{0+} + \rho_{0\times}, \quad b_{01} = \rho_{0+} + \rho_{1\times},$$
$$b_{10} = \rho_{1+} + \rho_{0\times}, \quad b_{11} = \rho_{1+} + \rho_{1\times},$$

with $\forall y \in \{0,1\}, b \in \{+,\times\} : \rho_{yb} = \frac{1}{|AND^{-1}(y)|} \sum_{x \in \mathrm{AND}^{-1}(y)} U_b|x\rangle\langle x|U_B^\dagger$. Consider $\mathcal{H}_2$, the 2-dimensional Hilbert space spanned by $|c_1\rangle \stackrel{\text{def}}{=} |1\rangle^{\otimes n}$ and $|h_1\rangle \stackrel{\text{def}}{=} |1_\times\rangle^{\otimes n}$. Let $|c_0\rangle \in \mathcal{H}_2$ and $|h_0\rangle \in \mathcal{H}_2$ be the state vectors orthogonal to $|c_1\rangle$ and $|h_1\rangle$ respectively. They can be expressed as:

$$\begin{aligned} |c_o\rangle &= \frac{(-1)^{n+1}|c_1\rangle + 2^{n/2}|h_1\rangle}{\sqrt{2^n - 1}}, \\ |h_o\rangle &= \frac{2^{n/2}|c_1\rangle + (-1)^{n+1}|h_1\rangle}{\sqrt{2^n - 1}}. \end{aligned}$$

Then $\Pi_\parallel = |c_0\rangle\langle c_0| + |c_1\rangle\langle c_1| = |h_0\rangle\langle h_0| + |h_1\rangle\langle h_1|$ is a projector onto $\mathcal{H}_2$. Let $\Pi_\perp$ be a projector onto the orthogonal complement of $\mathcal{H}_2$. Note that the $b_{o_+o_\times}$ are all composed of two blocks, one supported on $\mathcal{H}_2$ and the other on its orthogonal complement. We can thus write

$$\begin{aligned} b_{00} &= \frac{2\Pi_\perp}{2^n - 1} + \frac{|c_0\rangle\langle c_0| + |h_0\rangle\langle h_0|}{2^n - 1}, \\ b_{01} &= \frac{\Pi_\perp}{2^n - 1} + \left[\frac{|c_0\rangle\langle c_0|}{2^n - 1} + |h_1\rangle\langle h_1|\right], \\ b_{10} &= \frac{\Pi_\perp}{2^n - 1} + \left[\frac{|h_0\rangle\langle h_0|}{2^n - 1} + |c_1\rangle\langle c_1|\right], \\ b_{11} &= 0 + |c_1\rangle\langle c_1| + |h_1\rangle\langle h_1|. \end{aligned} \tag{3.5}$$

We give an explicit measurement that achieves p and then show that it is optimal. Take

$$\begin{aligned} M_{00} &= \Pi_\perp \\ M_{o_+o_\times} &= \lambda_{o_+o_\times}|\psi_{o_+o_\times}\rangle\langle\psi_{o_+o_\times}|, \end{aligned}$$

with $\lambda_{01} = \lambda_{10} = (1+\eta)^{-1}$ where

$$\begin{aligned} \eta &= \left|\frac{1 - 2\beta^2 + (-1)^{n+1} 2\beta\sqrt{1-\beta^2}\sqrt{2^n - 1}}{2^{n/2}}\right|, \\ |\psi_{01}\rangle &= \alpha|c_0\rangle + \beta|c_1\rangle, \\ |\psi_{10}\rangle &= \alpha|h_0\rangle + \beta|h_1\rangle, \end{aligned}$$

with α and β real and satisfying $\alpha^2 + \beta^2 = 1$. We also set $M_{11} = \mathbb{I} - M_{00} - M_{01} - M_{10}$. We take

$$\beta = (-1)^n \frac{1}{\sqrt{2^{2n} + 2^{\frac{3}{2}n+1} - 2^{\frac{n}{2}+1}}}.$$

Putting it all together, we thus calculate Bob's probability of success:

$$p = \frac{1}{2}\left[2 + \frac{1}{2^n + 2^{n/2} - 2} - \frac{1}{2^n - 1}\right].$$

We now show that this is in fact the optimal measurement for Bob. For this we consider the dual of our semidefinite program above:

$$\begin{array}{ll} \text{minimize} & \mathrm{Tr}(Q) \\ \text{subject to} & \forall o_+, o_\times \in \{0,1\} : Q \geq \frac{b_{o_+ o_\times}}{4}. \end{array}$$

Our goal is now to find a Q such that $p = \mathrm{Tr}(Q)$ and Q is dual feasible. We can then conclude from the duality of SDP that p is optimal. Consider

$$\begin{aligned} Q = \; & \tfrac{\Pi_\perp}{2(2^n-1)} + \tfrac{1}{4}\left(\tfrac{2-2^{1+n/2}+2^{3n/2}}{2-3\cdot 2^{n/2}+2^{3n/2}}\right)(|c_1\rangle\langle c_1| + |h_1\rangle\langle h_1|) \\ & -(-1)^n \tfrac{1}{4(2^{1-\frac{n}{2}}+2^n-3)}(|c_1\rangle\langle h_1| + |c_1\rangle\langle h_1|). \end{aligned}$$

Now we only need to show that the Q above satisfies the constraints, i.e. $\forall o_+, o_\times \in \{0,1\} : Q \geq b_{o_+ o_\times}/4$. Let $Q_\perp = \Pi_\perp Q \Pi_\perp$ and $Q_\parallel = \Pi_\parallel Q \Pi_\parallel$. By taking a look at Eq. (3.5) one can easily see that $Q_\perp \geq \frac{\Pi_\perp b_{o_+ o_\times} \Pi_\perp}{4}$, so that it is only left to show that

$$Q_\parallel \geq \frac{\Pi_\parallel b_{o_+ o_\times} \Pi_\parallel}{4}, \text{ for } o_+ o_\times \in \{0,1\}, o_+ o_\times \neq 00.$$

These are 2×2 matrices and this can be done straightforwardly. We thus have $\mathrm{Tr}(Q) = p$ and the result follows from the duality of semidefinite programming. □

It also follows that if Bob just wants to learn the value of a single bit, he can do no better than what he could achieve without waiting for Alice's announcement of the basis b:

3.4.4. Corollary. *Let $x \in \{0,1\}$, $P_X(x) = \frac{1}{2}$ and $f(x) = x$. Let $\mathcal{B} = \{+, \times\}$ with $U_+ = \mathbb{I}$ and $U_\times = H$. Then Bob succeeds at* PI_0-STAR(f) *with probability at most*

$$p = \frac{1}{2} + \frac{1}{2\sqrt{2}}.$$

There exists a strategy for Bob that achieves p.

The AND function provides an intuitive example of how Bob can compute the value of a function perfectly by storing just a single qubit. Consider the measurement with elements $\{\Pi_\parallel, \Pi_\perp\}$ from the previous section. It is easy to see that the outcome $\perp$ has zero probability if $\mathrm{AND}(x) = 1$. Thus, if Bob obtains that outcome he can immediately conclude that $\mathrm{AND}(x) = 0$. If Bob obtains outcome

$\|$ then the post-measurement states live in a 2-dimensional Hilbert space ($\mathcal{H}_2$), and can therefore be stored in a single qubit. Thus, by keeping the remaining state we can calculate the AND perfectly once the basis is announced. Our proof in Section 3.5.2, which shows that in fact *all* Boolean functions can be computed perfectly if Bob can store only a single qubit, makes use of a very similar effect to the one we observed here explicitly.

XOR function

We now examine the XOR function. This will be useful in order to gain some insight into the usefulness of post-measurement information later. For strings of even length, there exists a simple strategy for Bob even when three mutually unbiased bases are used.

3.4.5. Theorem. *Let $n \in \mathbb{N}$ be even, and let $P_X(x) = \frac{1}{2^n}$ for all $x \in \{0,1\}^n$. Let $\mathcal{B} = \{+, \times, \odot\}$ with $U_+ = \mathbb{I}^{\otimes n}$, $U_\times = H^{\otimes n}$ and $U_\odot = K^{\otimes n}$, where $K = (\mathbb{I} + i\sigma_x)/\sqrt{2}$. Then there is a strategy where Bob succeeds at* PI_0-STAR(XOR) *with probability $p = 1$.*

Proof. We first construct Bob's measurement for the first 2 qubits, which allows him to learn $x_1 \oplus x_2$ with probability 1. Note that the 12 possible states that Alice sends can be expressed in the Bell basis as follows:

$$
\begin{aligned}
|00\rangle &= \frac{1}{\sqrt{2}}(|\Phi^+\rangle + |\Phi^-\rangle) \quad H^{\otimes 2}|00\rangle = \tfrac{1}{\sqrt{2}}(|\Phi^+\rangle + |\Psi^+\rangle)\\
|01\rangle &= \frac{1}{\sqrt{2}}(|\Psi^+\rangle + |\Psi^-\rangle) \quad H^{\otimes 2}|01\rangle = \tfrac{1}{\sqrt{2}}(|\Phi^-\rangle + |\Psi^-\rangle)\\
|10\rangle &= \frac{1}{\sqrt{2}}(|\Psi^+\rangle - |\Psi^-\rangle) \quad H^{\otimes 2}|10\rangle = \tfrac{1}{\sqrt{2}}(|\Phi^-\rangle - |\Psi^-\rangle)\\
|11\rangle &= \frac{1}{\sqrt{2}}(|\Phi^+\rangle - |\Phi^-\rangle) \quad H^{\otimes 2}|11\rangle = \tfrac{1}{\sqrt{2}}(|\Phi^+\rangle - |\Psi^+\rangle)
\end{aligned}
$$

$$
\begin{aligned}
K^{\otimes 2}|00\rangle &= \frac{1}{\sqrt{2}}(|\Phi^-\rangle + i|\Psi^+\rangle)\\
K^{\otimes 2}|01\rangle &= \frac{1}{\sqrt{2}}(i|\Phi^+\rangle + |\Psi^-\rangle)\\
K^{\otimes 2}|10\rangle &= \frac{1}{\sqrt{2}}(i|\Phi^+\rangle - |\Psi^-\rangle)\\
K^{\otimes 2}|11\rangle &= -\frac{1}{\sqrt{2}}(|\Phi^-\rangle - i|\Psi^+\rangle).
\end{aligned}
$$

Bob now simply measures in the Bell basis and records his outcome. If Alice now announces that she used the computational basis, Bob concludes that $x_1 \oplus x_2 = 0$

if the outcome is one of $|\Phi^{\pm}\rangle$ and $x_1 \oplus x_2 = 1$ otherwise. If Alice announces she used the Hadamard basis, Bob concludes that $x_1 \oplus x_2 = 0$ if the outcome was one of $\{|\Phi^+\rangle, |\Psi^+\rangle\}$ and $x_1 \oplus x_2 = 1$ otherwise. Finally, if Alice announces that she used the $\odot$ basis, Bob concludes that $x_1 \oplus x_2 = 0$ if the outcome was one of $\{|\Phi^-\rangle, |\Psi^+\rangle\}$ and $x_1 \oplus x_2 = 1$ otherwise. Bob can thus learn the XOR of two bits with probability 1. To learn the XOR of the entire string, Bob applies this strategy to each two bits individually and then computes the XOR of all answers. □

Analogously to the proof of Theorem 3.4.5, we obtain:

3.4.6. Corollary. *Let $n \in \mathbb{N}$ be even, and let $P_X(x) = \frac{1}{2^n}$ for all $x \in \{0,1\}^n$. Let $\mathcal{B} = \{+, \times\}$ with $U_+ = \mathbb{I}^{\otimes n}$ and $U_\times = H^{\otimes n}$. Then there is a strategy where Bob succeeds at* PI_0-STAR(XOR) *with probability $p = 1$.*

Interestingly, there is no equivalent strategy for Bob if n is odd. In fact, as we show in the next section, in this case the post-measurement information gives no advantage to Bob at all.

3.4.7. Theorem. *Let $n \in \mathbb{N}$ be odd, and let $P_X(x) = \frac{1}{2^n}$ for all $x \in \{0,1\}^n$. Let $\mathcal{B} = \{+, \times\}$ with $U_+ = \mathbb{I}^{\otimes n}$, $U_\times = H^{\otimes n}$ and $P_B(+) = P_B(\times) = 1/2$. Then Bob succeeds at* PI_0-STAR(XOR) *with probability at most*

$$p = \frac{1}{2}\left(1 + \frac{1}{\sqrt{2}}\right).$$

There exists a strategy for Bob that achieves p.

Proof. Similar to the proof of the AND function, we can write Bob's optimal probability of success as the following semidefinite program in terms of the length of the input string, n:

$$\begin{array}{ll} \text{maximize} & \frac{1}{4}\sum_{o_+,o_\times\in\{0,1\}} \mathrm{Tr}[b^{(n)}_{o_+o_\times} M_{o_+o_\times}] \\ \text{subject to} & \forall o_+, o_\times \in \{0,1\} : M_{o_+o_\times} \geq 0, \\ & \sum_{o_+,o_\times\in\{0,1\}} M_{o_+o_\times} = \mathbb{I}, \end{array}$$

where

$$b^{(n)}_{o_+o_\times} = \rho^{(n)}_{o_++} + \rho^{(n)}_{o_\times\times},$$

and $\rho^{(n)}_{o_b b} = \frac{1}{2^{n-1}} \sum_{x\in\{0,1\}^n, x\in \mathrm{XOR}^{-1}(o_b)} U_b|x\rangle\langle x|U_b^\dagger$. The dual can be written as

$$\begin{array}{ll} \text{minimize} & \frac{1}{4}\mathrm{Tr}(Q^{(n)}) \\ \text{subject to} & \forall o_+, o_\times \in \{0,1\} : Q^{(n)} \geq b^{(n)}_{o_+o_\times}. \end{array}$$

Our proof is now by induction on n. For $n = 1$, let $Q^{(1)} = 2p\mathbb{I}$. It is easy to verify that $\forall o_+, o_\times \in \{0,1\} : Q^{(1)} \geq b^{(1)}_{o_+ o_\times}$ and thus $Q^{(1)}$ is a feasible solution of the dual program.

We now show that for $n+2$, $Q^{(n+2)} = Q^{(n)} \otimes \frac{1}{4}\mathbb{I}$ is a feasible solution to the dual for $n+2$, where $Q^{(n)}$ is a solution for the dual for n. Note that the XOR of all bits in the string can be expressed as the XOR of the first $n-2$ bits XORed with the XOR of the last two. Recall Eq. (3.3) and note that we can write

$$\rho^{(2)}_{0+} = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|) = \frac{1}{2}(|\Phi^+\rangle\langle\Phi^+| + |\Phi^-\rangle\langle\Phi^-|)$$
$$\rho^{(2)}_{1+} = \frac{1}{2}(|01\rangle\langle 01| + |10\rangle\langle 10|) = \frac{1}{2}(|\Psi^+\rangle\langle\Psi^+| + |\Psi^-\rangle\langle\Psi^-|).$$

It is easy to see that $\rho^{(2)}_{0\times} = H\rho^{(2)}_{0+}H = \frac{1}{2}(|\Phi^+\rangle\langle\Phi^+| + |\Psi^+\rangle\langle\Psi^+|)$ and $\rho^{(2)}_{1\times} = H\rho^{(2)}_{1+}H = \frac{1}{2}(|\Phi^-\rangle\langle\Phi^-| + |\Psi^-\rangle\langle\Psi^-|)$. By substituting from the above equation we then obtain

$$\begin{aligned} b^{(n+2)}_{00} = \rho^{(n+2)}_{0+} + \rho^{(n+2)}_{0\times} = \frac{1}{4}\big((\rho^{(n)}_{0+} + \rho^{(n)}_{0\times}) \otimes |\Phi^+\rangle\langle\Phi^+| + (\rho^{(n)}_{0+} + \rho^{(n)}_{1\times}) \otimes |\Phi^-\rangle\langle\Phi^-| \\ (\rho^{(n)}_{1+} + \rho^{(n)}_{0\times}) \otimes |\Psi^+\rangle\langle\Psi^+| + (\rho^{(n)}_{1+} + \rho^{(n)}_{1\times}) \otimes |\Psi^-\rangle\langle\Psi^-|\big) \\ \leq \frac{1}{4}Q^{(n)} \otimes \mathbb{I}, \end{aligned}$$

where we have used the fact that $Q^{(n)}$ is a feasible solution for the dual for n and that $|\Phi^+\rangle\langle\Phi^+| + |\Phi^-\rangle\langle\Phi^-| + |\Psi^+\rangle\langle\Psi^+| + |\Psi^-\rangle\langle\Psi^-| = \mathbb{I}$. The argument for $b^{(n+2)}_{01}$, $b^{(n+2)}_{10}$ and $b^{(n+2)}_{11}$ is analogous. Thus $Q^{(n+2)}$ satisfies all constraints.

Putting things together, we have for odd n that $\mathrm{Tr}(Q^{(n+2)}) = \mathrm{Tr}(Q^{(n)}) = \mathrm{Tr}(Q^{(1)})$ and since the dual is a minimization problem we know that

$$p \leq \frac{1}{4}\mathrm{Tr}(Q^{(1)}) = c$$

as claimed. Clearly, there exists a strategy for Bob that achieves $p = c$. He can compute the XOR of the first $n-1$ bits perfectly, as shown in Theorem 3.4.6. By Corollary 3.4.4 he can learn the value of the remaining n-th bit with probability $p = c$. □

We obtain a similar bound for three bases:

3.4.8. Theorem. *Let $n \in \mathbb{N}$ be odd, and let $P_X(x) = \frac{1}{2^n}$ for all $x \in \{0,1\}^n$. Let $\mathcal{B} = \{+, \times, \odot\}$ with $U_+ = \mathbb{I}^{\otimes n}$, $U_\times = H^{\otimes n}$ and $U_\odot = K^{\otimes n}$, where $K = (\mathbb{I} + i\sigma_x)/\sqrt{2}$, with $P_B(+) = P_B(\times) = P_B(\odot) = 1/3$. Then Bob succeeds at* PI_0-STAR(XOR) *with probability at most*

$$p = \frac{1}{2}\left(1 + \frac{1}{\sqrt{3}}\right).$$

There exists a strategy for Bob that achieves p.

Proof. The proof follows the same lines as Theorem 3.4.7. Bob's optimal probability of success is:

$$\begin{array}{ll} \text{maximize} & \frac{1}{6}\displaystyle\sum_{o_+,o_\times,o_\odot\in\{0,1\}} \mathrm{Tr}[b^{(n)}_{o_+o_\times o_\odot} M_{o_+o_\times o_\odot}] \\ \text{subject to} & \forall o_+,o_\times,o_\odot \in \{0,1\}\in\{0,1\} : M_{o_+o_\times o_\odot} \geq 0, \\ & \displaystyle\sum_{o_+,o_\times,o_\odot\in\{0,1\}} M_{o_+o_\times o_\odot} = \mathbb{I}, \end{array}$$

where

$$b^{(n)}_{o_+o_\times o_\odot} = \sum_{b\in\mathcal{B}} \rho_{o_b b},$$

and

$$\rho_{o_b b} = \frac{1}{2^{n-1}} \sum_{x\in XOR(o_b)} U_b|x\rangle\langle x|U_b^\dagger.$$

The dual can be written as

$$\begin{array}{ll} \text{minimize} & \frac{1}{6}\mathrm{Tr}(Q^{(n)}) \\ \text{subject to} & \forall o_+,o_\times,o_\odot \in \{0,1\} : Q^{(n)} \geq b^{(n)}_{o_+o_\times o_\odot}. \end{array}$$

Again, the proof continues by induction on n. For $n = 1$, let $Q^{(1)} = 3p\mathbb{I}$. It is easy to verify that $\forall o_+,o_\times,o_\odot \in \{0,1\} : Q^{(1)} \geq b^{(1)}_{o_+o_\times o_\odot}$ and thus $Q^{(1)}$ is a feasible solution of the dual program. The rest of the proof is done exactly in the same way as in Theorem 3.4.7 using that

$$\rho^{(2)}_{0\odot} = \frac{1}{2}(|\Phi^-\rangle\langle\Phi^-| + |\Psi^+\rangle\langle\Psi^+|)$$
$$\rho^{(2)}_{1\odot} = \frac{1}{2}(|\Psi^-\rangle\langle\Psi^-| + |\Phi^+\rangle\langle\Phi^+|).$$

□

3.5 Using post-measurement information and quantum memory

3.5.1 An algebraic framework for perfect prediction

So far, we had assumed that Bob is not allowed to store any qubits and can only use the additional post-measurement information to improve his guess. Now, we investigate the case where he has a certain amount of quantum memory at his disposal. In particular, we present a general algebraic approach to determine the minimum dimension 2^q of quantum memory needed to succeed with probability

1 at an instance of PI_q-STAR($\mathcal{E}$), for any ensemble $\mathcal{E} = \{p_{yb}, \rho_{yb}\}$ as long as the individual states for different values of y are mutually orthogonal for a fixed b, i.e., $\forall y \neq z \in \mathcal{Y}\ \mathrm{Tr}(\rho_{yb}, \rho_{zb}) = 0$. In particular, we are looking for an instrument consisting of a family of completely positive maps $\rho \mapsto A\rho A^\dagger$, adding up to a trace preserving map, such that $\mathrm{rank}(A) \leq 2^q$. This ensures that the post-measurement state "fits" into q qubits, and thus takes care of the memory bound. The fact that after the announcement of b the remaining state $A\rho_{yb}A^\dagger$ gives full information about y is expressed by demanding orthogonality of the different post-measurement states:

$$\forall b \in \mathcal{B}, \forall y \neq z \in \mathcal{Y} \quad A\rho_{yb}A^\dagger A\rho_{zb}A^\dagger = 0. \tag{3.6}$$

Note that here we explicitly allow the possibility that, say, $A\rho_{zb}A^\dagger = 0$: this means that if Bob obtains outcome A and later learns b, he can exclude the output value z. What Eq. (3.6) also implies is that for all states $|\psi\rangle$ and $|\varphi\rangle$ in the support of ρ_{yb} and ρ_{zb}, respectively, one has $A|\psi\rangle\langle\psi|A^\dagger A|\varphi\rangle\langle\varphi|A^\dagger = 0$. Hence, introducing the support projectors P_{yb} of the ρ_{yb}, we can reformulate Eq. (3.6) as

$$\forall b \in \mathcal{B}, \forall y \neq z \in \mathcal{Y} \quad AP_{yb}A^\dagger AP_{zb}A^\dagger = 0,$$

which can equivalently be expressed as

$$\forall b \in \mathcal{B}, \forall y \neq z \in \mathcal{Y} \quad \mathrm{Tr}\big(A^\dagger AP_{yb}A^\dagger AP_{zb}\big) = 0, \tag{3.7}$$

by noting that $A^\dagger A$ as well as the projectors are positive-semidefinite operators. As expected, we see that only the POVM operators $M = A^\dagger A$ of the instrument play a role in this condition. Our conditions can therefore also be written as $MP_{yb}MP_{zb} = 0$. From this condition, we now derive the following lemma.

3.5.1. Lemma. *Bob, using a POVM with operators* $\{M_i\}$*, succeeds at* PI_q-STAR *with probability* 1*, if and only if*

1. *for all* i*,* $\mathrm{rank}(M_i) \leq 2^q$*,*
2. *for all* $y \in \mathcal{Y}$ *and* $b \in \mathcal{B}$*,* $[M, P_{yb}] = 0$*, where* P_{yb} *is the projection on the support of* ρ_{yb}*.*

Proof. We first show that these two conditions are necessary. Note that only the commutation condition has to be proved. Let M be a measurement operator from a POVM succeeding with probability 1. Then, for any y, b, we have by Eq. (3.7) that

$$\mathrm{Tr}\big(MP_{yb}M(\mathbb{I} - P_{yb})\big) = 0, \text{ hence } \mathrm{Tr}\big(MP_{yb}MP_{yb}\big) = \mathrm{Tr}\big(MP_{yb}M\big).$$

Thus, by the positivity of the trace on positive operators, the cyclicity of the trace, and $P_{yb}^2 = P_{yb}$ we have that

$$\begin{aligned} 0 &\leq \mathrm{Tr}\big([M, P_{yb}]^\dagger [M, P_{yb}]\big) \\ &= \mathrm{Tr}\big(-(MP_{yb} - P_{yb}M)^2\big) \\ &= \mathrm{Tr}\big(-MP_{yb}MP_{yb} - P_{yb}MP_{yb}M + P_{yb}M^2P_{yb} + MP_{yb}^2M\big) = 0. \end{aligned}$$

But that means that the commutator $[M, P_{yb}]$ has to be 0.

Sufficiency is easy: since the measurement operators commute with the states' support projectors P_{yb}, and these are orthogonal to each other for fixed b, the post-measurement states of these projectors, $\propto \sqrt{M}P_{yb}\sqrt{M}$ are also mutually orthogonal for fixed b. Thus, if Bob learns b, he can perform a measurement to distinguish the different values of y perfectly. The post-measurement states are clearly supported on the support of M, which can be stored in q qubits. Since Bob's strategy succeeds with probability 1, it succeeds with probability 1 for any states supported in the range of the P_{yb}. □

Note that the operators M of the instrument need not commute with the originally given states ρ_{yb}. Nevertheless, the measurement preserves the orthogonality of ρ_{yb} and ρ_{zb} with $y \neq z$ for fixed b, i.e., $\mathrm{Tr}(\rho_{yb}\rho_{zb}) = 0$. Now that we know that the POVM operators of the instrument have to commute with all the states' support projectors P_{yb}, we can invoke some well-developed algebraic machinery to find the optimal such instrument.

Looking at Appendix B, we see that M has to come from the commutant of the operators P_{yb}. These themselves generate a $*$-subalgebra $\mathcal{A}$ of the full operator algebra $\mathbb{B}(\mathcal{H})$ of the underlying Hilbert space $\mathcal{H}$, and the structure of such algebras and their commutants in finite dimension is well understood. We know from Theorem B.4.7 that the Hilbert space $\mathcal{H}$ has a decomposition (i.e., there is an isomorphism which we write as an equality)

$$\mathcal{H} = \bigoplus_j \mathcal{J}_j \otimes \mathcal{K}_j \tag{3.8}$$

into a direct sum of tensor products such that the $*$-algebra $\mathcal{A}$ and its commutant algebra $\mathrm{Comm}(\mathcal{A}) = \big\{M : \forall P \in \mathbb{B}(\mathcal{H})\ [P, M] = 0\big\}$ can be written

$$\mathcal{A} \cong \bigoplus_j \mathcal{B}(\mathcal{J}_j) \otimes \mathbb{I}_{\mathcal{K}_j}, \tag{3.9}$$

$$\mathrm{Comm}(\mathcal{A}) \cong \bigoplus_j \mathbb{I}_{\mathcal{J}_j} \otimes \mathcal{B}(\mathcal{K}_j). \tag{3.10}$$

Koashi and Imoto [KI02], in the context of finding the quantum operations which leave a set of states invariant, have described an algorithm to find the

commutant $\mathrm{Comm}(\mathcal{A})$, and more precisely the Hilbert space decomposition of Eq. (3.8), of the states $P_{yb}/\mathrm{Tr}P_{yb}$. They show that for this decomposition, there exist states $\sigma_{j|i}$ on $\mathcal{J}_j$, a conditional probability distribution $\{q_{j|i}\}$, and states ω_j on $\mathcal{K}_j$ which are independent of i, such that we can write them as

$$\forall i \quad \sigma_i = \bigoplus_j q_{j|i}\sigma_{j|i} \otimes \omega_j,$$

Looking at Eq. (3.10), we see that the smallest rank operators $M \in \mathrm{Comm}(\mathscr{A})$ are of the form $\mathbb{I}_{\mathcal{J}_j} \otimes |\psi\rangle\langle\psi|$ for some j and $|\psi\rangle \in \mathcal{K}_j$, and that they are all admissible. Since we need a family of operators M that are closed to a POVM (i.e., their sum is equal to the identity), we know that all j have to occur. Hence, the minimal quantum memory requirement is

$$\min 2^q = \max_j \dim \mathcal{J}_j. \tag{3.11}$$

The strategy Bob has to follow is this: For each j, pick a basis $\{|e_{k|j}\rangle\}$ for $\mathcal{K}_j$ and measure the POVM $\{\mathbb{I}_{\mathcal{J}_j} \otimes |e_{k|j}\rangle\langle e_{k|j}|\}$, corresponding to the decomposition

$$\mathcal{H} = \bigoplus_{jk} \mathcal{J}_j \otimes |e_{k|j}\rangle\langle e_{k|j}|,$$

which commutes with the P_{yb}. For each outcome, he can store the post-measurement state in q qubits [as in Eq. (3.11)], preserving the orthogonality of the states for different y but fixed b. Once he learns b he can thus obtain y with certainty.

Of course, carrying out the Koashi-Imoto algorithm may not be a straightforward task in a given situation. We now consider two explicit examples that one can understand as two special cases of this general method: First, we show that in fact *all* Boolean functions with two bases (mutually unbiased or not) can be computed perfectly when Bob is allowed to store just a single qubit. Second, however, we show that there exist three bases such that for *any balanced* function, Bob must store *all* qubits to compute the function perfectly. We also give a recipe how to construct such bases.

3.5.2 Using two bases

For two bases, Bob needs to store only a single qubit to compute any Boolean function perfectly. As outlined in Section 3.5.1, we need to show that there exists a measurement with the following properties: First, the post-measurement states of states corresponding to strings x such that $f(x) = 0$ are orthogonal to the post-measurement states of states corresponding to strings y such that $f(y) = 1$. Indeed, if this is true and we keep the post-measurement state, then after the basis is announced, we can distinguish perfectly between both types of states. Second, of course, we need that the post-measurement states are supported in subspaces of

dimension at most 2. The following little lemma shows that this is the case for any Boolean function. The same statement has been shown independently many times before in a variety of different contexts. For example, Masanes and also Toner and Verstraete have shown the same in the context of non-local games [Mas06, TV06]. The key ingredient is also present in Bathia's textbook [Bha97]. Indeed, there is a close connection between the amount of post-measurement information we require, and the amount of entanglement we need to implement measurements in the setting of non-local games. We return to this question in Chapter 6.

3.5.2. LEMMA. *Let* $f : \{0,1\}^n \to \{0,1\}$ *and* $P_{0b} = \sum_{x \in f^{-1}(0)} U_b|x\rangle\langle x|U_b^\dagger$ *where* $U_0 = \mathbb{I}$ *and* $U_1 = U$*, then there exists a direct sum decomposition of the Hilbert space*

$$\mathcal{H} = \bigoplus_{i=1}^{m} \mathcal{H}_i, \text{ with } \dim \mathcal{H}_i \leq 2,$$

such that P_{00} *and* P_{01} *can be expressed as*

$$P_{00} = \sum_{i=1}^{m} \Pi_i P_{00} \Pi_i,$$
$$P_{01} = \sum_{i=1}^{m} \Pi_i P_{01} \Pi_i,$$

where Π_i *is the orthogonal projector onto* $\mathcal{H}_i$*.*

Proof. There exists a basis so that P_{00} and P_{01} can be written as

$$P_{00} = \begin{bmatrix} \mathbb{I}_{n_0} & 0_{n_0 \times n_1} \\ 0_{n_1 \times n_0} & 0_{n_1 \times n_1} \end{bmatrix}, P_{01} = \begin{bmatrix} A^{00}_{n_0 \times n_0} & A^{01}_{n_0 \times n_1} \\ (A^{01})^\dagger_{n_1 \times n_0} & A^{11}_{n_1 \times n_1} \end{bmatrix},$$

where $n_y = |f^{-1}(y)|$ is the number of strings x such that $f(x) = y$, and we have specified the dimensions of the matrix blocks for clarity. In what follows these dimensions will be omitted. We assume without loss of generality that $n_0 \leq n_1$. It is easy to check that, since P_{01} is a projector, it must satisfy

$$\begin{aligned} A^{00}(\mathbb{I}_{n_0} - A^{00}) &= A^{01} A^{01\dagger}, \\ A^{11}(\mathbb{I}_{n_1} - A^{11}) &= A^{01\dagger} A^{01}. \end{aligned} \tag{3.12}$$

Consider a unitary of the following form

$$V = \begin{bmatrix} V_0 & 0 \\ 0 & V_1 \end{bmatrix},$$

where V_0 and V_1 are $n_0 \times n_0$ and $n_1 \times n_1$ unitaries respectively. Under such a unitary, P_{00} and P_{01} are transformed to:

$$\begin{aligned} V P_{00} V^\dagger &= P_{00}, \\ V P_{01} V^\dagger &= \begin{bmatrix} V_0 A^{00} V_0^\dagger & V_0 A^{01} V_1^\dagger \\ (V_0 A^{01} V_1^\dagger)^\dagger & V_1 A^{11} V_1^\dagger \end{bmatrix}. \end{aligned} \tag{3.13}$$

We now choose V_0 and V_1 from the singular value decomposition (SVD, [HJ85, Theorem 7.3.5]) of $A^{01} = V_0^\dagger D V_1$ which gives

$$D = V_0 A^{01} V_1^\dagger = \sum_{k=1}^{n_0} d_k |u_k\rangle\langle v_k|,$$

where $d_k \geq 0$, $\langle u_k|u_l\rangle = \langle v_k|v_l\rangle = \delta_{kl}$. Since $(A^{01})^\dagger A^{01}$ and $A^{01}(A^{01})^\dagger$ are supported in orthogonal subspaces, it also holds that $\forall k,l : \langle u_k|v_l\rangle = 0$. Eqs. (3.12) and (3.13) now give us

$$\begin{aligned} V_0 A^{00} V_0^\dagger (\mathbb{I}_{n_0} - V_0 A^{00} V_0^\dagger) &= \textstyle\sum_{k=1}^{n_0} d_k^2 |u_k\rangle\langle u_k|, \\ V_1 A^{11} V_1^\dagger (\mathbb{I}_{n_1} - V_1 A^{11} V_1^\dagger) &= \textstyle\sum_{k=1}^{n_0} d_k^2 |v_k\rangle\langle v_k|. \end{aligned}$$

Suppose for the time being that all the d_k are different. Since they are all non-negative, all the d_k^2 will also be different and it must hold that

$$\begin{aligned} V_0 A^{00} V_0^\dagger &= \sum_{k=1}^{n_0} a_k^0 |u_k\rangle\langle u_k|, \\ V_1 A^{11} V_1^\dagger &= \sum_{k=1}^{n_0} a_k^1 |v_k\rangle\langle v_k| + \sum_{k=n_0+1}^{n_1} a_k^1 |\tilde{v}_k\rangle\langle \tilde{v}_k| \end{aligned}$$

for some a_k^0, a_k^1 and $|\tilde{v}_k\rangle$ with $1 \leq k \leq n_1$.. Note that we can choose $|\tilde{v}_k\rangle$ such that $\forall k,k', k \neq k' : \langle \tilde{v}_k|\tilde{v}_{k'}\rangle = 0$ and $\forall k,l : \langle u_k|\tilde{v}_l\rangle = 0$. We can now express $V P_{01} V^\dagger$ as

$$\begin{aligned} &V P_{01} V^\dagger = \\ &= \sum_{k=1}^{n_0} \left[a_k^0 |u_k\rangle\langle u_k| + d_k(|u_k\rangle\langle v_k| + |v_k\rangle\langle u_k|) + a_k^1 |v_k\rangle\langle v_k| \right] + \sum_{k=n_0+1}^{n_1} a_k^1 |\tilde{v}_k\rangle\langle \tilde{v}_k|. \end{aligned}$$

It is now clear that we can choose all $\mathcal{H}_k = \mathrm{span}\{|u_k\rangle, |v_k\rangle\}$, and $\mathcal{H}_{k'} = \mathrm{span}\{|\tilde{v}_{k'}\rangle\}$ which are orthogonal and together add up to $\mathcal{H}$.

In the case that all the d_k are not different, there is some freedom left in choosing $|u_k\rangle$ and $|v_k\rangle$ that still allows us to make $V_0 A^{00} V_0^\dagger$ and $V_1 A^{11} V_1^\dagger$ diagonal so that the rest of the proof follows in the same way. □

In particular, the previous lemma implies that the post-measurement states corresponding to strings x for which $f(x) = 0$ are orthogonal to those corresponding to strings x for which $f(x) = 1$, which is expressed in the following lemma.

3.5.3. Lemma. *Suppose one performs the measurement given by $\{\Pi_i : i \in [m]\}$. If the outcome of the measurement is i and the state was $U_b|x\rangle$, then the post-measurement state is*

$$|x,i,b\rangle := \frac{\Pi_i U_b |x\rangle}{\sqrt{\langle x| U_b^\dagger \Pi_i U_b |x\rangle}}.$$

The post-measurement states satisfy

$$\forall x \in f^{-1}(0),\ x' \in f^{-1}(1),\ i \in [m] : \langle x, i, b | x', i, b \rangle = 0.$$

Proof. The proof follows straightforwardly from that fact that the Π_i commute with both P_{00} and P_{01} (which follows from Lemma 3.5.2). □

Now we are ready to prove the main theorem of this section.

3.5.4. Theorem. *Let $|\mathcal{Y}| = |\mathcal{B}| = 2$, then there exists a strategy for Bob such that he succeeds at* PI_1-STAR($\mathcal{E}$) *with probability $p = 1$, for any function f and prior P_X on $\mathcal{X}$.*

Proof. The strategy that Bob uses is the following:

- Bob performs the measurement given by $\{\Pi_i : i \in [m]\}$.
- He obtains an outcome $i \in [m]$ and stores the post-measurement state which is supported in the at most two-dimensional subspace $\mathcal{H}_i$.
- After the basis $b \in \{0, 1\}$ is announced, he measures $\{P_{0b}, P_{1b}\}$ and reports the outcome of this measurement.

By Lemma 3.5.3 this leads to success probability 1. □

Our result also gives us a better lower bound for all Boolean functions than what we had previously obtained in Section 3.4.1. Instead of storing the qubit, Bob now measures it immediately along the lines of Lemma 3.3.1. It is not too difficult to convince yourself that for one qubit the worst-case post-measurement states to distinguish are in fact those in Lemma 3.3.1.

3.5.5. Corollary. *Let $|\mathcal{Y}| = |\mathcal{B}| = 2$, then Bob succeeds at* PI_0-STAR($\mathcal{E}$) *with probability at least $p \geq (1 + 1/\sqrt{2})/2$.*

In particular, our result implies that for the task of constructing Rabin-OT in [DFSS05] it is essential for Alice to choose a random function f from a larger set, which is initially unknown to Bob.

As a final remark, note that the prior distributions do not play any role. Likewise, it is not actually important that the states ρ_{yb} are proportional to projectors: we only require that for all $b \in \{0, 1\}$, the states ρ_{0b} and ρ_{1b} are orthogonal.

3.5.3 Using three bases

We have just shown that Bob can compute any Boolean function perfectly when two bases are used. However, we now show that for any balanced Boolean function there exist three bases, such that Bob needs to store *all* qubits in order to compute the function perfectly. The idea behind our proof is that for a particular choice of three bases, any measurement operator that satisfies the conditions set out in Lemma 3.5.1 must be proportional to the identity. This means that we cannot reduce the number of qubits to be stored by a measurement and must keep everything. First, we prove the following lemma which we need in our main proof.

3.5.6. Lemma. *Let M be a self-adjoint matrix which is diagonal in two mutually unbiased bases, then M must be proportional to the identity.*

Proof. Let $|x\rangle$ $|u_x\rangle$ $x \in \{1, \ldots, d\}$ be the two MUBs and let m_x and m'_x be the eigenvalues corresponding to $|x\rangle$ and $|u_x\rangle$ respectively, then we can write

$$M = \sum_{x=1}^{d} m_x |x\rangle\langle x| = \sum_{x'=1}^{d} m'_{x'} |u_{x'}\rangle\langle u_{x'}|.$$

From the previous equation, it follows that

$$\langle x|M|x\rangle = m_x = \sum_{x'=1}^{d} m'_{x'} |\langle u_{x'}|x\rangle|^2 = \frac{1}{d}\mathrm{Tr}M,$$

which implies the desired result. □

We are now ready to prove the main result of this section.

3.5.7. Theorem. *Let $|\mathcal{Y}| = 2$ and $|\mathcal{B}| = 3$, then for any balanced function f and prior P_X on $\mathcal{X}$ which is uniform on the pre-images $f^{-1}(y)$, there exist three bases such that Bob succeeds at* PI_q-STAR($\mathcal{E}$) *with probability $p = 1$ if and only if $q = \log d$.*

Proof. Let $P_{00} = \sum_{x \in f^{-1}(0)} |x\rangle\langle x|$, $P_{01} = U_1 P_{00} U_1^\dagger$ and $P_{02} = U_2 P_{00} U_2^\dagger$. Also, let $s : f^{-1}(0) \to f^{-1}(1)$ be a bijective map, and let $s_x = s(x)$. By a reordering of the basis, P_{00}, U_1 and U_2 can be written as

$$P_{00} = \begin{bmatrix} \mathbb{I} & 0 \\ 0 & 0 \end{bmatrix}, U_1 = \begin{bmatrix} U_1^{00} & U_1^{01} \\ U_1^{10} & U_1^{11} \end{bmatrix}, U_2 = \begin{bmatrix} U_2^{00} & U_2^{01} \\ U_2^{10} & U_2^{11} \end{bmatrix},$$

where all the blocks are of size $(d/2) \times (d/2)$. P_{01} and P_{02} then take the following form:

$$P_{01} = \begin{bmatrix} U_1^{00} U_1^{00\dagger} & U_1^{00} U_1^{10\dagger} \\ (U_1^{00} U_1^{10\dagger})^\dagger & U_1^{10} U_1^{10\dagger} \end{bmatrix}, P_{02} = \begin{bmatrix} U_2^{00} U_2^{00\dagger} & U_2^{00} U_2^{10\dagger} \\ (U_2^{00} U_2^{10\dagger})^\dagger & U_2^{10} U_2^{10\dagger} \end{bmatrix}.$$

It follows from Lemma 3.5.1, that we only have to prove that $[M, P_{00}] = [M, P_{01}] = [M, P_{02}] = 0$ implies that M must be proportional to the identity. Write

$$M = \begin{bmatrix} M^{00} & M^{01} \\ (M^{01})^\dagger & M^{11} \end{bmatrix}.$$

Commutation with P_{00} implies $M^{01} = 0$. Commutation with P_{01} and P_{02} implies

$$\begin{aligned} [M^{00}, U_1^{00} U_1^{00\dagger}] &= [M^{00}, U_2^{00} U_2^{00\dagger}] = 0, && (3.14) \\ [M^{11}, U_1^{10} U_1^{10\dagger}] &= [M^{11}, U_2^{10} U_2^{10\dagger}] = 0, && (3.15) \\ M^{00}(U_1^{00} U_1^{10\dagger}) &= (U_1^{00} U_1^{10\dagger}) M^{11}, && (3.16) \\ M^{00}(U_2^{00} U_2^{10\dagger}) &= (U_2^{00} U_2^{10\dagger}) M^{11}. && (3.17) \end{aligned}$$

We choose U_1 and U_2 in the following way:

$$\begin{aligned} U_1 &= \sum_{x \in f^{-1}(0)} \left[a_x(|x\rangle\langle x| + |s_x\rangle\langle s_x|) + \sqrt{1 - a_x^2}(|x\rangle\langle s_x| - |s_x\rangle\langle x|) \right], \\ U_2 &= \sum_{x \in f^{-1}(0)} \left[a_x(|u_x\rangle\langle u_x| + |v_x\rangle\langle v_x|) + \sqrt{1 - a_x^2}(|u_x\rangle\langle v_x| - |v_x\rangle\langle u_x|) \right], \end{aligned}$$

with $a_x \in [0, 1]$, satisfying $a_x = a_{x'}$ if and only if $x = x'$. Furthermore, choose $|u_x\rangle$ and $|v_x\rangle$ such that

$$\forall x, x' \in f^{-1}(0),\ \langle x|v_{x'}\rangle = \langle s_x|u_{x'}\rangle = 0,\ |\langle x|u_{x'}\rangle|^2 = |\langle s_x|v_{x'}\rangle|^2 = 2/d.$$

With this choice for U_1 and U_2 we have that

$$\begin{aligned} U_1^{00} U_1^{00\dagger} &= \sum_{x \in f^{-1}(0)} a_x^2 |x\rangle\langle x|, \\ U_2^{00} U_2^{00\dagger} &= \sum_{x \in f^{-1}(0)} a_x^2 |u_x\rangle\langle u_x|, \end{aligned}$$

i.e., $\{|x\rangle\}$ and $\{|u_x\rangle\}$ form an eigenbasis for $U_1^{00} U_1^{00\dagger}$ and $U_2^{00} U_2^{00\dagger}$ respectively. Furthermore, since all the a_x^2 are different, the eigenbases are unique. Now, using Eq. (3.14), we see that M^{00} must commute with both $U_1^{00} U_1^{00\dagger}$ and $U_2^{00} U_2^{00\dagger}$, and since their eigenbases are unique, it must be true that M^{00} is diagonal in both $\{|x\rangle\}$ and $\{|u_x\rangle\}$. Using the result of Lemma 3.5.6 it follows that $M^{00} = m_0 \mathbb{I}_{d/2}$. In exactly the same way we can prove that $M^{11} = m_1 \mathbb{I}_{d/2}$ using Eq. (3.15). It remains to prove that $m_0 = m_1$, which follows directly from either Eq. (3.16) or Eq. (3.17). □

From our proof it is clear how to construct U_1 and U_2. For P_{00} as defined above, we could choose vectors of the form $|x\rangle = |0\rangle|\hat{x}\rangle$ and $|s_x\rangle = |1\rangle|\hat{x}\rangle$ where

$\hat{x} \in \{0,1\}^{n-1}$ to construct U_1. For U_2 we could then pick $|u_x\rangle = |0\rangle H^{\otimes n-1}|\hat{x}\rangle$ and analogously $|v_x\rangle = |1\rangle H^{\otimes n-1}|\hat{x}\rangle$. As we will see in Chapter 6, our example shows that for non-local games we cannot hope to prove a statement analogous to [Mas06, TV06] for three measurement settings where each measurement has two outcomes.

Note, however, that whereas we know that for such unitaries Bob must store all qubits in order to compute the value of the function perfectly, it remains unclear how close he can get to computing the function perfectly when storing fewer qubits. In particular, he can always choose two of the three bases, and employ the strategy outlined in the previous section: he stores the one qubit that allows him to succeed with probability 1 for two of the bases. If he gets the third basis then he just flips a coin. In this case, he is correct with probability $2/3 + 1/(3 \cdot 2) = 5/6$ for a balanced function and a uniform prior. It remains an important open question to address the approximate case.

3.6 Conclusion

We have introduced a new state discrimination problem, motivated by cryptography: discrimination with extra information about the state after the measurement, or, more generally, after a quantum memory bound applies. We have left most general questions open, but we found fairly complete results in the case of guessing $y = f(x)$ with mutually unbiased encodings.

We have shown that storing just a single qubit allows Bob to succeed at PI-STAR perfectly for *any* Boolean function and any two bases. In contrast, we showed how to construct *three* bases such that Bob needs to store *all* qubits in order to compute the function perfectly. We have also given an explicit strategy for two functions, namely the AND and the XOR. More generally, it would be interesting to determine, how many qubits Bob needs to store to compute $f(x)$ perfectly for any function $f : \mathcal{X} \to \mathcal{Y}$ in terms of the number of outputs $|\mathcal{Y}|$ and the number of bases $|\mathcal{B}|$. It should be clear that the algebraic techniques of Section 3.5.1 allow us to answer these questions for any given function in principle. However, so far, we have not been able to obtain explicit structures for wider classes of functions. Our results imply that in existing protocols in the bounded quantum storage model [DFSS05] we cannot restrict ourselves to a single fixed function f to perform privacy amplification. Note that our algebraic framework can also address the question of using more than one function, where f is also announced after the memory bound applies [DFSS05]: we merely obtain a larger problem. Yet, it is again difficult to determine a general bound.

In the important case of two mutually unbiased bases and balanced functions, we have shown (Theorem 3.3.3 and Corollary 3.5.5) that there exists a clear separation between the case where Bob gets the post-measurement information (PI-STAR) and when he does not (STAR). Namely, for any such function, Bob's

optimal success probability is never larger than $(1 + 1/\sqrt{2})/2 \approx 0.853$ for STAR and always at least as large as the same number for PI-STAR.

In some cases the gap between STAR and PI-STAR can be more dramatic. The XOR function on strings of even length with two mutually unbiased bases is one of these cases. We have shown that in this case the advantage can be maximal. Namely, *without* the extra information Bob can never do better than guessing the basis, *with* it however, he can compute the value of the function perfectly. This contrasts with the XOR function on strings of odd length, where the optimal success probabilities of STAR and PI-STAR are both $(1 + 1/\sqrt{2})/2$ and the post-measurement information is completely useless for Bob. It would be interesting to see, how large the gap between STAR and PI-STAR can be for any function $f : \{0,1\}^n \to \{0,1\}^k$ where $k > 2$. We return to this question in Chapter 6.4.

It would also be nice to show a general lower bound for non-balanced functions or a non-uniform prior. As the example for 3 bases showed, the uniform prior is not necessarily the one that leads to the largest gap, and thus the prior can play an important role. Another generalization would be to consider functions of the form $f : [d]^n \to [d]^k$.

We now turn our attention to uncertainty relations. These will play an important role in locking in Chapter 5. In the problem of locking, we also distinguish measurement *with* basis information, analogous to our PI_q-STAR with $q = n$, and *without* corresponding to PI_0-STAR. So far, our objective has been to obtain an accurate guess of a value, e.g. $y = f(x)$. In Chapter 5, we are interested in a slightly different problem: How can we maximize the classical mutual information? In particular, can we use mutually unbiased bases to obtain locking effects?

Chapter 4

Uncertainty relations

Uncertainty relations lie at the very core of quantum mechanics. Intuitively, they quantify how much we can learn about different properties of a quantum system simultaneously. Some properties lead to very strong uncertainty relations: if we decide to learn one, we remain entirely ignorant about the others. But what characterizes such properties? In this chapter, we first investigate whether choosing our measurements to be mutually unbiased bases allows us to obtain strong uncertainty relations. Sadly, it turns out that mutual unbiasedness is not sufficient. Instead, we need to consider anti-commuting measurements.

4.1 Introduction

Heisenberg first realized that quantum mechanics leads to uncertainty relations for conjugate observables such as position and momentum [Hei27]. Uncertainty relations are probably best known in the form given by Robertson [Rob29], who extended Heisenberg's result to any two observables A and B. Robertson's relation states that if we prepare many copies of the state $|\psi\rangle$, and measure each copy individually using either A or B, we have

$$\Delta A \Delta B \geq \frac{1}{2}|\langle\psi|[A,B]|\psi\rangle|$$

where $\Delta X = \sqrt{\langle\psi|X^2|\psi\rangle - \langle\psi|X|\psi\rangle^2}$ for $X \in \{A, B\}$ is the standard deviation resulting from measuring $|\psi\rangle$ with observable X. Recall from Chapter 2, that classically we always have $[A, B] = 0$, and there is no such limiting lower bound. Hence, uncertainty relations are another characteristic that sets apart quantum theory. The consequences are rather striking: even if we had a perfect measurement apparatus, we are nevertheless limited!

Entropic uncertainty relations are an alternative way to state Heisenberg's uncertainty principle. They are frequently a more useful characterization, because the "uncertainty" is lower bounded by a quantity that does not depend on the

state to be measured [Deu83, Kra87]. Recently, entropic uncertainty relations have gained importance in the context of quantum cryptography in the bounded storage model, where proving the security of such protocols ultimately reduces to establishing such relations [DFR+07]. Proving new entropic uncertainty relations could thus give rise to new protocols. Intuitively, it is clear that uncertainty relations have a significant impact on what kind of protocols we can obtain in the quantum settings. Recall the cryptographic task of oblivious transfer from Chapter 1: the receiver should be able to extract information about one particular property of a system, but should learn as little as possible about all other properties. It is clear that, without placing any additional restrictions on the receiver, uncertainty relations intuitively quantify how well we are able to implement such a primitive.

Entropic uncertainty relations were first introduced by Bialynicki-Birula and Mycielski [BBM75]. For our purposes, we will be interested in uncertainty relations in the form put forward by Deutsch [Deu83]. Following a conjecture by Kraus [Kra87], Maassen and Uffink [MU88] have shown that if we measure the state $|\psi\rangle$ with observables A and B determined by the bases $\mathcal{A} = \{|a_1\rangle, \ldots, |a_d\rangle\}$ and $\mathcal{B} = \{|b_1\rangle, \ldots, |b_d\rangle\}$ respectively, we have

$$\frac{1}{2}\left(H(\mathcal{A}||\psi\rangle) + H(\mathcal{B}||\psi\rangle)\right) \geq -\log c(\mathcal{A}, \mathcal{B}),$$

where $c(\mathcal{A}, \mathcal{B}) = \max\{|\langle a|b\rangle| \mid |a\rangle \in \mathcal{A}, |b\rangle \in \mathcal{B}\}$, and

$$H(\mathcal{X}||\psi\rangle) = -\sum_{i=1}^{d} |\langle\psi|x_i\rangle|^2 \log |\langle\psi|x_i\rangle|^2$$

is the Shannon entropy [Sha48] arising from measuring the state $|\psi\rangle$ in the basis $\mathcal{X} = \{|x_1\rangle, \ldots, |x_d\rangle\}$. In fact, Maassen and Uffink provide a more general statement which also leads to uncertainty relations for higher order Rényi entropies. Such relations have also been shown by Bialynicki-Birula [BB06] for special sets of observables. Note that the above relation achieves our initial goal: the lower bound no longer depends on the state $|\psi\rangle$, but only on A and B itself. What is the strongest possible relation we could obtain? That is, which choices of $\mathcal{A}$ and $\mathcal{B}$ maximize $-\log c(\mathcal{A}, \mathcal{B})$? It is not hard to see that choosing $\mathcal{A}$ and $\mathcal{B}$ to be mutually unbiased (see Section 2.4) provides us with a lower bound of $(\log d)/2$ which is the strongest possible uncertainty relation: If we have no entropy for one of the bases, then the entropy for the other bases must be maximal. For example, in case of a one qubit system of $d = 2$ choosing $\mathcal{A} = \{|0\rangle, |1\rangle\}$ and $\mathcal{B} = \{|+\rangle, |-\rangle\}$ to be the computational and the Hadamard basis respectively, we obtain a lower bound of $1/2$.

Can we derive a similar relation for measurements using three or more observables? Surprisingly, very little is known for a larger number of measurement settings [Aza04]. Sanchez-Ruiz [San93, SR95] (using results of Larsen [Lar90])

has shown that for measurements using all $d+1$ mutually unbiased bases, we can obtain strong uncertainty relations. Here, we provide an elementary proof of his result in dimension $d = 2^n$. Given the fact that mutually unbiased bases seem to be a good choice if we use only two or $d+1$ measurement settings, it may be tempting to conclude that choosing our measurements to be mutually unbiased always gives us good uncertainty relations for which the lower bound is as large as possible. Numerical results for MUBs in prime dimensions up to 29 indicate that MUBs may indeed be a good choice [DHL$^+$04]. However, we show that merely being mutually unbiased is not sufficient to obtain strong uncertainty relations. To this end, we prove tight entropic uncertainty relations for measurements in a large number of mutually unbiased bases (MUBs) in square dimensions. In particular, we consider any MUBs derived from mutually orthogonal Latin squares [WB05], and *any* set of MUBs obtained from the set of unitaries of the form $\{U \otimes U^*\}$, where $\{U\}$ gives rise to a set of MUBs in dimension s when applied to the basis elements of the computational basis. For any s, there are at most $s+1$ such MUBs in a Hilbert space of dimension $d = s^2$: recall from Section 2.4 that we can have at most $s+1$ MUBs in a space of dimension s. Let $\mathbb{B}$ be the set of MUBs coming from one of these two constructions. We prove that for any subset $\mathbb{T} \subseteq \mathbb{B}$ of these bases we have

$$\min_{|\psi\rangle} \sum_{\mathcal{B}\in\mathbb{T}} H(\mathcal{B}||\psi\rangle) = \frac{|\mathbb{T}|}{2} \log d.$$

Our result shows that one needs to be careful to think of "maximally incompatible" measurements as being necessarily mutually unbiased. When we take entropic uncertainty relations as our measure of "incompatibility", mutually unbiased measurements are not always the most incompatible when considering more than two observables. In particular, it has been shown [HLSW04] that if we choose approximately $(\log d)^4$ bases uniformly at random, then with high probability $\min_{|\psi\rangle}(1/|\mathbb{T}|)\sum_{\mathcal{B}\in\mathbb{T}} H(\mathcal{B}||\psi\rangle) \geq \log d - 3$. This means that there exist $(\log d)^4$ bases for which this sum of entropies is very large, i.e., measurements in such bases are very incompatible. However, we show that when d is large, there exist $\sqrt{d}$ mutually unbiased bases that are much less incompatible according to this measure. When considering entropic uncertainty relations as a measure of "incompatibility", we must therefore look for different properties for the bases to define incompatible measurements.

Luckily, we are able to obtain maximally strong uncertainty relations for two-outcome measurements for *anti-commuting* observables. In particular, we obtain for $\Gamma_1, \ldots, \Gamma_K$ with $\{\Gamma_i, \Gamma_j\} = 0$ that

$$\min_{\rho} \frac{1}{K} \sum_{j=1}^{K} H(\Gamma_j|\rho) = 1 - \frac{1}{K},$$

where $H(\Gamma_j|\rho) = -\sum_{b\in\{0,1\}} \mathrm{Tr}(\Gamma_j^b \rho) \log \mathrm{Tr}(\Gamma_j^b \rho)$ and Γ_j^0, Γ_j^1 are projectors onto

the positive and negative eigenspace of Γ_j respectively. Thus, if we have zero entropy for one of the terms, we must have maximal entropy for all others. For the collision entropy we obtain something slightly suboptimal

$$\min_{\rho} \frac{1}{K} \sum_{j=1}^{K} H_2(\Gamma_j, \rho) \approx 1 - \frac{\log e}{K}$$

for large K, where $H_2(\Gamma_j|\rho) = -\log \sum_{b\in\{0,1\}} \mathrm{Tr}(\Gamma_j^b \rho)^2$. Especially our second uncertainty relation is of interest for cryptographic applications.

4.2 Limitations of mutually unbiased bases

We first prove tight entropic uncertainty for measurements in MUBs in square dimensions. We need the result of Maassen and Uffink [MU88] mentioned above:

4.2.1. Theorem (Maassen and Uffink). *Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be two orthonormal basis in a Hilbert space of dimension d. Then for all pure states $|\psi\rangle$*

$$\frac{1}{2}\left(H(\mathcal{B}_1||\psi\rangle) + H(\mathcal{B}_2||\psi\rangle)\right) \geq -\log c(\mathcal{B}_1, \mathcal{B}_2),$$

where $c(\mathcal{B}_1, \mathcal{B}_2) = \max\{|\langle b_1|b_2\rangle| \mid |b_1\rangle \in \mathcal{B}_1, |b_2\rangle \in \mathcal{B}_2\}$.

The case when $\mathcal{B}_1$ and $\mathcal{B}_2$ are MUBs is of special interest for us. More generally, when one has a set of MUBs a trivial application of Theorem 4.2.1 leads to the following corollary also noted in [Aza04].

4.2.2. Corollary. *Let $\mathbb{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_m\}$ be a set of MUBs in a Hilbert space of dimension d. Then*

$$\frac{1}{m} \sum_{t=1}^{m} H(\mathcal{B}_t||\psi\rangle) \geq \frac{\log d}{2}.$$

Proof. Using Theorem 4.2.1, one gets that for any pair of MUBs $\mathcal{B}_t$ and $\mathcal{B}_{t'}$ with $t \neq t'$

$$\frac{1}{2}\left[H(\mathcal{B}_t|\psi) + H(\mathcal{B}_{t'}|\psi)\right] \geq \frac{\log d}{2}.$$

Adding up the resulting equation for all pairs $t \neq t'$ we get the desired result. □

We now show that this bound can in fact be tight for a large set of MUBs.

4.2.1 MUBs in square dimensions

Corollary 4.2.2 gives a lower bound on the average of the entropies of a set of MUBs. But how good is this bound? We show that the bound is indeed tight when we consider product MUBs in a Hilbert space of square dimension.

4.2.3. THEOREM. *Let $\mathbb{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_m\}$ with $m \geq 2$ be a set of MUBs in a Hilbert space $\mathcal{H}$ of dimension s. Let U_t be the unitary operator that transforms the computational basis to $\mathcal{B}_t$. Then $\mathbb{V} = \{\mathcal{V}_1, \ldots, \mathcal{V}_m\}$, where*

$$\mathcal{V}_t = \{U_t|k\rangle \otimes U_t^*|l\rangle \mid k, l \in [s]\},$$

is a set of MUBs in $\mathcal{H} \otimes \mathcal{H}$, and it holds that

$$\min_{|\psi\rangle} \frac{1}{m} \sum_{t=1}^{m} H(\mathcal{V}_t||\psi\rangle) = \frac{\log d}{2},$$

where $d = \dim(\mathcal{H} \otimes \mathcal{H}) = s^2$.

Proof. It is easy to check that $\mathbb{V}$ is indeed a set of MUBs. Our proof works by constructing a state $|\psi\rangle$ that achieves the bound in Corollary 4.2.2. It is easy to see that the maximally entangled state

$$|\psi\rangle = \frac{1}{\sqrt{s}} \sum_{k=1}^{s} |kk\rangle,$$

satisfies $U \otimes U^*|\psi\rangle = |\psi\rangle$ for any $U \in \mathrm{U}(d)$. Indeed,

$$\begin{aligned}
\langle\psi|U \otimes U^*|\psi\rangle &= \frac{1}{s} \sum_{k,l=1}^{s} \langle k|U|l\rangle\langle k|U^*|l\rangle \\
&= \frac{1}{s} \sum_{k,l=1}^{s} \langle k|U|l\rangle\langle l|U^\dagger|k\rangle \\
&= \frac{1}{s}\mathrm{Tr}UU^\dagger = 1.
\end{aligned}$$

Therefore, for any $t \in [m]$ we have that

$$\begin{aligned}
H(\mathcal{V}_t||\psi\rangle) &= -\sum_{kl} |\langle kl|U_t \otimes U_t^*|\psi\rangle|^2 \log |\langle kl|U_t \otimes U_t^*|\psi\rangle|^2 \\
&= -\sum_{kl} |\langle kl|\psi\rangle|^2 \log |\langle kl|\psi\rangle|^2 \\
&= \log s = \frac{\log d}{2}.
\end{aligned}$$

Taking the average of the previous equation over all t we obtain the result. □

4.2.2 MUBs based on Latin squares

We now consider mutually unbiased bases based on Latin squares [WB05] as described in Section 2.4.1. Our proof again follows by providing a state that achieves the bound in Corollary 4.2.2, which turns out to have a very simple form.

4.2.4. Lemma. *Let $\mathbb{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_m\}$ with $m \geq 2$ be any set of MUBs in a Hilbert space of dimension $d = s^2$ constructed on the basis of Latin squares. Then*

$$\min_{|\psi\rangle} \frac{1}{m} \sum_{\mathcal{B}\in\mathbb{B}} H(\mathcal{B}||\psi\rangle) = \frac{\log d}{2}.$$

Proof. Consider the state $|\psi\rangle = |1,1\rangle$ and fix a basis $\mathcal{B}_t = \{|v^t_{i,j}\rangle | i,j \in [s]\} \in \mathbb{B}$ coming from a Latin square. It is easy to see that there exists exactly one $j \in [s]$ such that $\langle v^t_{1,j}|1,1\rangle = 1/\sqrt{s}$. Namely this will be the $j \in [s]$ at position $(1,1)$ in the Latin square. Fix this j. For any other $\ell \in [s], \ell \neq j$, we have $\langle v^t_{1,\ell}|1,1\rangle = 0$. But this means that there exist exactly s vectors in $\mathcal{B}$ such that $|\langle v^t_{i,j}|1,1\rangle|^2 = 1/s$, namely exactly the s vectors derived from $|v^t_{1,j}\rangle$ via the Hadamard matrix. The same argument holds for any such basis $\mathcal{B} \in \mathbb{T}$. We get

$$\begin{aligned}
\sum_{\mathcal{B}\in\mathbb{T}} H(\mathcal{B}||1,1\rangle) &= \sum_{\mathcal{B}\in\mathbb{T}} \sum_{i,j\in[s]} |\langle v^t_{i,j}|1,1\rangle|^2 \log |\langle v^t_{i,j}|1,1\rangle|^2 \\
&= |\mathbb{T}| s \frac{1}{s} \log \frac{1}{s} \\
&= |\mathbb{T}| \frac{\log d}{2}.
\end{aligned}$$

The result then follows directly from Corollary 4.2.2. □

4.2.3 Using a full set of MUBs

We now provide an alternative proof of an entropic uncertainty relation for a full set of mutually unbiased bases. This has previously been proved in [San93, SR95]. We already provided an alternative proof using the fact that the set of all mutually unbiased bases forms a 2-design [BW07]. Here, we provide a new alternative proof for dimension $d = 2^n$ which has the advantage that it neither requires the introduction of 2-designs, nor the results of [Lar90] that were used in the previous proof by Sanchez-Ruiz [San93, SR95]. Instead, our proof is extremely simple: After choosing a convenient parametrization of quantum states, the statement follows immediately using only elementary Fourier analysis.

For the parametrization, we first introduce a basis for the space of $2^n \times 2^n$ matrices with the help of mutually unbiased bases. Recall from Section 2.4 that in dimension 2^n, we can find exactly $2^n + 1$ MUBs.

4.2.5. Lemma. *Consider the Hermitian matrices*

$$S_b^j = \sum_{x\in\{0,1\}^n} (-1)^{j\cdot x}|x_b\rangle\langle x_b|,$$

for $b \in [d+1]$, $j \in \{0,\ldots,d-1\}$ and for all $x, x' \in \{0,1\}^n$ and $b \neq b' \in [d+1]$ we have $|\langle x_b|x'_{b'}\rangle|^2 = 1/d$. Then the set $\{\mathbb{I}\} \cup \{S_b^j \mid b \in [d+1], j \in \{0,\ldots,d-1\}\}$ forms a basis for the space of $d \times d$ matrices, where for all j and b, S_b^j is traceless and $(S_b^j)^2 = \mathbb{I}$.

Proof. First, note that we have $(d+1)(d-1)+1 = d^2$ matrices. We now show that they are all orthogonal. Note that

$$\mathrm{Tr}(S_b^j) = \sum_{x\in\{0,1\}^n} (-1)^{j\cdot x} = 0,$$

since $j \neq 0$, and hence S_b^j is traceless. Hence $\mathrm{Tr}(\mathbb{I}S_b^j) = 0$. Furthermore,

$$\mathrm{Tr}(S_b^j S_{b'}^{j'}) = \sum_{x,x'\in\{0,1\}^n} (-1)^{j\cdot x}(-1)^{j'\cdot x'}|\langle x_b|x'_{b'}\rangle|^2. \tag{4.1}$$

For $b \neq b'$, Eq. (4.1) gives us $\mathrm{Tr}(S_b^j S_{b'}^{j'}) = (1/d)\left(\sum_x (-1)^{j\cdot x}\right)\left(\sum_{x'}(-1)^{j'\cdot x'}\right) = 0$, since $j, j' \neq 0$. For $b = b'$, but $j \neq j'$, we get $\mathrm{Tr}(S_b^j S_{b'}^{j'}) = \sum_x (-1)^{(j\oplus j')\cdot x} = 0$ since $j \oplus j' \neq 0$.

Finally, $\left(S_b^j\right)^2 = \sum_{xx'}(-1)^{j\cdot x}(-1)^{j\cdot x'}|x_b\rangle\langle x_b||x'_b\rangle\langle x'_b| = \mathbb{I}$. □

Since $\{\mathbb{I}, S_b^j\}$ form a basis for the $d \times d$ matrices, we can thus express the state ρ of a d-dimensional system as

$$\rho = \frac{1}{d}\left(\mathbb{I} + \sum_{b\in[d+1]}\sum_{j\in\{0,\ldots,d-1\}} s_b^j S_b^j\right),$$

for some coefficients $s_b^j \in \mathbb{R}$. It is now easy to see that

4.2.6. Lemma. *Let ρ be a pure state parametrized as above. Then*

$$\sum_{b\in[d+1]}\sum_{j\in\{0,\ldots,d-1\}} (s_b^j)^2 = d-1.$$

Proof. If ρ is a pure state, we have $\mathrm{Tr}(\rho^2) = 1$. Hence

$$\begin{aligned}\mathrm{Tr}(\rho^2) &= \frac{1}{d^2}\left(\mathrm{Tr}(\mathbb{I}) + \sum_{b\in[d+1]}\sum_{j\in\{0,\ldots,d-1\}} (s_b^j)^2\mathrm{Tr}(\mathbb{I})\right)\\ &= \frac{1}{d}\left(1 + \sum_b\sum_j (s_b^j)^2\right) = 1,\end{aligned}$$

from which the claim follows. □

Suppose now that we are given a set of $d+1$ MUBs $\mathcal{B}_1, \ldots, \mathcal{B}_{d+1}$ with $\mathcal{B}_b = \{|x_b\rangle \mid x \in \{0,1\}^n\}$. Then the following simple observation lies at the core of our proof:

4.2.7. Lemma. *Let $|x_b\rangle$ be the x-th basis vector of the b-th MUB. Then for any state ρ*

$$\mathrm{Tr}(|x_b\rangle\langle x_b|\rho) = \frac{1}{d}\left(1 + \sum_{j\in\{0,\ldots,d-1\}} (-1)^{j\cdot x} s_b^j\right).$$

Proof. We have

$$\mathrm{Tr}(|x_b\rangle\langle x_b|\rho) = \frac{1}{d}\left(\mathrm{Tr}(|x_b\rangle\langle x_b|) + \sum_{b',j} s_{b'}^j \mathrm{Tr}(S_{b'}^j |x_b\rangle\langle x_b|)\right)$$

Suppose $b \neq b'$. Then $\mathrm{Tr}(S_{b'}^j |x_b\rangle\langle x_b|) = (1/d)\sum_{x'}(-1)^{j\cdot x'} = 0$, since $j \neq 0$. Suppose $b = b'$. Then $\mathrm{Tr}(S_{b'}^j |x_b\rangle\langle x_b|) = \sum_{x'}(-1)^{j\cdot x'}|\langle x_b|x'_b\rangle|^2 = (-1)^{j\cdot x}$, from which the claim follows. □

We are now ready to prove an entropic uncertainty relation for N mutually unbiased bases.

4.2.8. Theorem. *Let $\mathcal{S} = \{\mathcal{B}_1, \ldots, \mathcal{B}_N\}$ be a set of mutually unbiased bases. Then*

$$\frac{1}{N}\sum_{b\in[N]} H_2(\mathcal{B}_b, |\Psi\rangle) \geq -\log\frac{N+d-1}{dN}.$$

Proof. First, note that we can define functions $f_b(j) = s_b^j$ for $j \in \{0,\ldots,d-1\}$ and $f_b(0) = 1$. Then $\hat{f}_b(x) = (1/\sqrt{d})(\sum_{j\in\{0,\ldots,d-1\}}(-1)^{j\cdot x}s_b^j)$ is the Fourier transform of f_b and $(1/\sqrt{d})\hat{f}_b(x) = \mathrm{Tr}(|x_b\rangle\langle x_b|)$ by Lemma 4.2.7. Thus

$$\begin{aligned}
\frac{1}{N}\sum_{b\in[N]} H_2(\mathcal{B}_b, |\Psi\rangle) &= -\frac{1}{N}\sum_{b\in[N]} \log \sum_{x\in\{0,1\}^n} |\langle x_b|\Psi\rangle|^4 \\
&\geq -\log\frac{1}{dN}\sum_b\sum_x \hat{f}_b(x)^2 \\
&= -\log\frac{1}{dN}\sum_b (1 + \sum_j (s_b^j)^2) \\
&= -\log\frac{1}{dN}(N+d-1),
\end{aligned}$$

where the first inequality follows from Jensen's inequality and the concavity of log. The next equality follows from Parseval's equality, and the last follows from the fact that $|\Psi\rangle$ is a pure state and Lemma 4.2.6. □

4.2.9. COROLLARY. *Let $\mathcal{S} = \{\mathcal{B}_1, \ldots, \mathcal{B}_N\}$ be a set of mutually unbiased bases. Then*

$$\frac{1}{N} \sum_{b \in [N]} H(\mathcal{B}_b || \Psi\rangle) \geq -\log \frac{N + d - 1}{dN}.$$

In particular, for a full set of $N = d + 1$ MUBs we have $(1/N) \sum_b H(\mathcal{B}_b||\Psi\rangle) \geq \log((d+1)/2)$.

Proof. This follows immediately from Theorem 4.2.8 and the fact that $H(\cdot) \geq H_2(\cdot)$. □

It is interesting to note that this bound is the same that arises from interpolating between the results of Sanchez-Ruiz [San93, SR95] and Maassen and Uffink [MU88] as was done by Azarchs [Aza04].

4.3 Good uncertainty relations

As we saw, merely choosing our measurements to be mutually unbiased is not sufficient to obtain good uncertainty relations. However, we now investigate measurements using *anti-commuting* observables for which we do obtain maximally strong uncertainty relations! In particular, we consider the matrices $\Gamma_1, \ldots, \Gamma_{2n}$, satisfying the anti-commutation relations

$$\Gamma_i \Gamma_j = -\Gamma_j \Gamma_i, \ \Gamma_i^2 = \mathbb{I} \tag{4.2}$$

for all $i, j \in [2n]$. Such operators $\Gamma_1, \ldots, \Gamma_{2n}$ form generators for the Clifford algebra, which we explain in more detail in Appendix C.

Intuitively, these operators have a property that is very similar to being mutually unbiased: Recall from Appendix C that we can write for all $j \in [2n]$

$$\Gamma_j = \Gamma_j^0 - \Gamma_j^1,$$

where Γ_j^0 and Γ_j^1 are projectors onto the positive and negative eigenspace of Γ_j respectively. We also have that for all $i, j \in [2n]$ with $i \neq j$

$$\mathrm{Tr}(\Gamma_i \Gamma_j) = \frac{1}{2} \mathrm{Tr}(\Gamma_i \Gamma_j + \Gamma_j \Gamma_i) = 0.$$

Hence the positive and negative eigenspaces of such operators are similarly mutually unbiased as bases can be: from

$$\mathrm{Tr}(\Gamma_i \Gamma_j^0) = \mathrm{Tr}(\Gamma_i \Gamma_j^1),$$

we immediately see that if we would pick a vector lying in the positive or negative eigenspace of Γ_j and perform a measurement with Γ_i, the probability to obtain outcome Γ_i^0 or outcome Γ_i^1 must be the same. Thus, one might intuitively hope to obtain good uncertainty relations for measurements using such operators. We now show that this is indeed the case.

4.3.1 Preliminaries

Before we can turn to proving our uncertainty relations, we recall a few simple observations from Appendix C. The operators $\Gamma_1, \ldots, \Gamma_{2n}$ have a unique (up to unitary) representation in terms of the matrices

$$\begin{aligned} \Gamma_{2j-1} &= \sigma_y^{\otimes(j-1)} \otimes \sigma_x \otimes \mathbb{I}^{\otimes(n-j)}, \\ \Gamma_{2j} &= \sigma_y^{\otimes(j-1)} \otimes \sigma_z \otimes \mathbb{I}^{\otimes(n-j)}, \end{aligned}$$

for $j = 1, \ldots, n$. We now fix this representation. The product $\Gamma_0 := i\Gamma_1\Gamma_2 \cdots \Gamma_{2n}$ is also called the pseudo-scalar. A particularly useful fact is that the collection of operators

$$\begin{aligned} \mathbb{I} & \\ \Gamma_j & \qquad (1 \leq j \leq 2n) \\ \Gamma_{jk} &= i\Gamma_j\Gamma_k \ (1 \leq j < k \leq 2n) \\ \Gamma_{jk\ell} &= \Gamma_j\Gamma_k\Gamma_\ell \ (1 \leq j < k < \ell \leq 2n) \\ &\vdots \\ \Gamma_{12\ldots(2n)} &= \Gamma_0 \end{aligned}$$

forms an orthogonal basis for the $d \times d$ complex matrices for $d = 2^n$, where in the definition of the above operators we introduce a factor of i to all with an even number of indices to make the whole set a basis for the Hermitian operators with real valued coefficients. Hence we can write every state $\rho \in \mathcal{H}$ as

$$\rho = \frac{1}{d}\left(\mathbb{I} + \sum_j g_j\Gamma_j + \sum_{j<k} g_{jk}\Gamma_{jk} + \ldots + g_0\Gamma_0\right). \tag{4.3}$$

The real valued coefficients $(g_1, \ldots, g_{2n})$ in this expansion are called "vector" components, the ones belonging to higher degree products of Γ's are "tensor" or "k-vector" components.

Recall that we may think of the operators $\Gamma_1, \ldots, \Gamma_{2n}$ as the basis vectors of a $2n$-dimensional real vector space. Essentially, we can then think of the positive and negative eigenspace of such operators as the positive and negative direction of the basis vectors. We can visualize the $2n$ basis vectors with the help of a $2n$-dimensional hypercube. Each basis vector determines two opposing faces of the hypercube[1], where we can think of the two faces as corresponding to the positive and negative eigenspace of each operator as illustrated in Figures 4.1 and 4.2.

Finally, recall that within the Clifford algebra two vectors are orthogonal if and only if they anti-commute. Hence, if we transform the generating set of Γ_j linearly,

$$\Gamma'_k = \sum_j T_{jk}\Gamma_j,$$

[1]Note that the face of an $2n$-dimensional hypercube is a $2n-1$ dimensional hypercube itself.

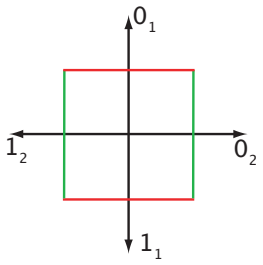


Figure 4.1: 2n = 2-cube

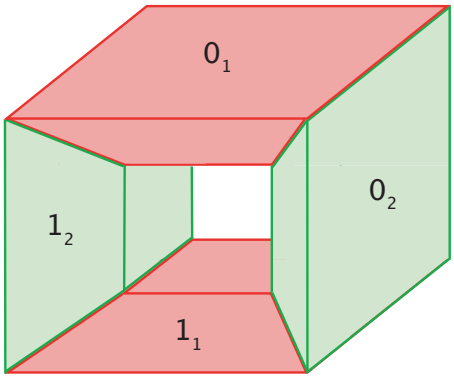


Figure 4.2: 2n = 4-cube

the set $\{\Gamma'_1, \ldots, \Gamma'_{2n}\}$ satisfies the anti-commutation relations if and only if $(T_{jk})_{jk}$ is an orthogonal matrix. In that case there exists a matching unitary $U(T)$ of $\mathcal{H}$ which transforms the operator basis as

$$\Gamma'_j = U(T)\Gamma_j U(T)^\dagger.$$

We thus have an $\mathrm{O}(2n)$ symmetry of the generating set $\Gamma_1, \ldots, \Gamma_{2n}$. Indeed, this can be extended to a $\mathrm{SO}(2n+1)$ symmetry by viewing Γ_0 as an additional "vector": It is not difficult to see that Γ_0 anti-commutes with $\Gamma_1, \ldots, \Gamma_{2n}$. We are thus free to remove one of these operators from the generating set and replace it with Γ_0 to obtain a new set of generators. Evidently, we may also view these as basis vectors. This observation forms the basis of the following little lemma, which allows us to prove our uncertainty relations:

4.3.1. Lemma. *The linear map $\mathbb{P}$ taking ρ as in Eq. (4.3) to*

$$\mathbb{P}(\rho) := \frac{1}{d}\left(\mathbb{I} + \sum_{j=0}^{2n} g_j \Gamma_j\right) \tag{4.4}$$

is positive. I.e., if ρ is a state, then so is $\mathbb{P}(\rho)$, and in this case $\sum_{j=0}^{2n} g_j^2 \leq 1$. Conversely, if $\sum_{j=0}^{2n} g_j^2 \leq 1$, then

$$\sigma = \frac{1}{d}\left(\mathbb{I} + \sum_{j=0}^{2n} g_j \Gamma_j\right)$$

is positive semidefinite, hence a state.

Proof. First, we show that there exists a unitary U such that $\rho' = U\rho U^\dagger$ has no pseudo-scalar Γ_0, and only one nonzero vector component, say at Γ_1. Hence, our goal is to find the transformation U that rotates $g = \sum_{j=0}^{2n} g_j \Gamma_j$ to the vector $b = \sqrt{\ell}\Gamma_1$, where we let $\ell := \sum_{j=0}^{2n} g_j^2 = (g_1')^2$. Finding such a transformation for only the first $2n$ generators can easily be achieved, as we saw in Appendix C. The challenge is thus to include Γ_0. To this end we perform three individual operations: First, we rotate $g' = \sum_{j=1}^{2n} g_j \Gamma_j$ onto the vector $b' = \sqrt{\ell'}\Gamma_1$ with $\ell' := \sum_{j=1}^{2n} g_j^2$. Second, we exchange Γ_2 and Γ_0. And finally we rotate the vector $g'' = \sqrt{\ell'}\Gamma_1 + g_0\Gamma_2$ onto the vector $b = \sqrt{\ell}\Gamma_1$.

First, we rotate $g' = \sum_{j=1}^{2n} g_j \Gamma_j$ onto the vector $b' = \sqrt{\ell'}\Gamma_1$: This is exactly analogous to the transformation constructed in Appendix C. Consider the vector $\hat{g} = \frac{1}{\sqrt{\ell'}} g'$. We have $\hat{g}^2 = |\hat{g}|^2 \mathbb{I} = \mathbb{I}$ and thus the vector is of length 1. Let $m = \hat{g} + \Gamma_1$ denote the vector lying in the plane spanned by Γ_1 and $\hat{g}$ located exactly halfway between Γ_1 and $\hat{g}$. Let $\hat{m} = c(\hat{g} + \Gamma_1)$ with $c = 1/\sqrt{2(1 + g_1/\sqrt{\ell'})}$. It is easy to verify that $\hat{m}^2 = \mathbb{I}$ and hence the vector $\hat{m}$ has length 1. To rotate the vector g' onto the vector b', we now need to first reflect g' around the plane perpendicular to $\hat{m}$, and then around the plane perpendicular to Γ_1. Hence, we now define $R = \Gamma_1 \hat{m}$. Evidently, R is unitary since $RR^\dagger = R^\dagger R = \mathbb{I}$. First of all, note that

$$\begin{aligned}
Rg' &= \Gamma_1 \hat{m} g' \\
&= c\Gamma_1 \left(\frac{1}{\sqrt{\ell'}} g' + \Gamma_1\right) g' \\
&= c\left(\Gamma_1 \frac{g'^2}{\sqrt{\ell'}} + \Gamma_1^2 g'\right) \\
&= c\sqrt{\ell'}\left(\Gamma_1 + \frac{1}{\sqrt{\ell'}} g'\right) \\
&= \sqrt{\ell'}\hat{m}.
\end{aligned}$$

Hence,

$$Rg'R^\dagger = \sqrt{\ell'}\hat{m}\hat{m}\Gamma_1 = \sqrt{\ell'}\Gamma_1 = b',$$

as desired. Using the geometry of the Clifford algebra, one can see that k-vectors remain k-vectors when transformed with the rotation R (see Appendix C). Similarly, it is easy to see that Γ_0 is untouched by the operation R

$$R\Gamma_0 R^\dagger = \Gamma_0 R R^\dagger = \Gamma_0,$$

since $\{\Gamma_0, \Gamma_j\} = 0$ for all $j \in \{1, \ldots, 2n\}$. We can thus conclude that

$$R\rho R^\dagger = \frac{1}{d}\left(\mathbb{I} + \sqrt{\ell'}\Gamma_1 + g_0\Gamma_0 + \sum_{j<k} g'_{jk}\Gamma_{jk} + \ldots\right),$$

for some coefficients g'_{jk} and similar for the terms involving higher products.

Second, we exchange Γ_2 and Γ_0: To this end, recall that $\Gamma_2, \ldots, \Gamma_{2n}, \Gamma_0$ is also a generating set for the Clifford algebra. Hence, we can now view Γ_0 itself as a vector with respect to the new generators. To exchange Γ_0 and Γ_2, we now simply rotate Γ_0 onto Γ_2. Essentially, this corresponds to a rotation about 90 degrees in the plane spanned by vectors Γ_0 and Γ_2. Consider the vector $n = \Gamma_0 + \Gamma_2$ located exactly halfway between both vectors. Let $\hat{n} = n/\sqrt{2}$ be the normalized vector. Let $R' = \Gamma_2\hat{n}$. A small calculation analogous to the above shows that

$$R'\Gamma_0 R'^\dagger = \Gamma_2 \text{ and } R'\Gamma_2 R'^\dagger = -\Gamma_0.$$

We also have that $\Gamma_1, \Gamma_3, \ldots, \Gamma_{2n}$ are untouched by the operation: for $j \neq 0$ and $j \neq 2$, we have that

$$R'\Gamma_j R'^\dagger = \Gamma_j,$$

since $\{\Gamma_0, \Gamma_j\} = \{\Gamma_2, \Gamma_j\} = 0$. How does R' affect the k-vectors in terms of the original generators $\Gamma_1, \ldots, \Gamma_{2n}$? Using the anti-commutation relations and the definition of Γ_0 it is easy to convince yourself that all k-vectors are mapped to k'-vectors with $k' \geq 2$ (except for Γ_0 itself). Hence, the coefficient of Γ_1 remains untouched. We can thus conclude that

$$R'R\rho R^\dagger R'^\dagger = \frac{1}{d}\left(\mathbb{I} + \sqrt{\ell'}\Gamma_1 + g_0\Gamma_2 + \sum_{j<k} g''_{jk}\Gamma_{jk} + \ldots\right),$$

for some coefficients g''_{jk} and so on.

Finally, we now rotate the vector $g'' = \sqrt{\ell'}\Gamma_1 + g_0\Gamma_2$ onto the vector b. Note that $(g'')^2 = (\ell + g_0^2)\mathbb{I} = \ell\mathbb{I}$. Let $\hat{g}'' = g''/\sqrt{\ell}$ be the normalized vector. Our rotation is derived exactly analogous to the first step: Let $k = \hat{g}'' + \Gamma_1$, and let $\hat{k} = k/\sqrt{2(1 + \sqrt{\ell'}/\sqrt{\ell})}$. Let $R'' = \Gamma_1\hat{k}$. A simple calculation analogous to the above shows that

$$R''g''R''^\dagger = \sqrt{\ell}\Gamma_1,$$

as desired. Again, we have $R''\Gamma_k R''^\dagger = \Gamma_k$ for $k \neq 1$ and $k \neq 2$. Furthermore, k-vectors remain k-vectors under the actions of R'' [DL03]. Summarizing, we obtain

$$R''R'R\rho R^\dagger R'^\dagger R''^\dagger = \frac{1}{d}\left(\mathbb{I} + \sqrt{\ell}\Gamma_1 + \sum_{j<k} g'''_{jk}\Gamma_{jk} + \dots\right),$$

for some coefficients g'''_{jk} and so on. Thus, we can take $U = R''R'R$ to arrive at a new, simpler looking, state

$$\begin{aligned}\rho' &= U\rho U^\dagger \\ &= \frac{1}{d}\left(\mathbb{I} + g'_1\Gamma_1 + \sum_{j<k} g'''_{jk}\Gamma_{jk} + \dots + 0\,\Gamma_0\right),\end{aligned}$$

for some g'''_{jk}, etc.

Similarly, there exist of course orthogonal transformations F_j that take Γ_k to $(-1)^{\delta_{jk}}\Gamma_k$. Such transformations flip the sign of a chosen Clifford generator. In a similar way to the above, it is easy to see that $F_j = \Gamma_0\Gamma_j$ fulfills this task: we rotate Γ_j by 90 degrees in the plane given by Γ_0 and Γ_j as in the example we examined in Appendix C. Now, consider

$$\rho'' = \frac{1}{2}\left(\rho' + F_j\rho' F_j^\dagger\right),$$

for $j > 1$. Clearly, if ρ' was a state, ρ'' is a state as well. Note that we no longer have terms involving Γ_j in the basis expansion: Note that if we flip the sign of precisely those terms that have an index j (i.e., they have a factor Γ_j in the definition of the operator basis), and then the coefficients cancel with those of ρ'.

We now iterate this map through $j = 2, 3, \dots, 2n$, and we are left with a final state $\hat{\rho}$ of the form

$$\hat{\rho} = \frac{1}{d}\left(\mathbb{I} + g'_1\Gamma_1\right).$$

By applying $U^\dagger = (R''R'R)^\dagger$ from above, we now transform $\hat{\rho}$ to $U^\dagger\hat{\rho}U = \mathbb{P}(\rho)$, which is the first part of the lemma.

Looking at $\hat{\rho}$ once more, we see that it can be positive semidefinite only if $g'_1 \leq 1$, i.e., $\sum_{j=0}^{2n} g_j^2 \leq 1$. Evidently, $\mathrm{Tr}(\hat{\rho}) = 1$ and hence $\hat{\rho}$ is a state.

Conversely, if $\sum_{j=0}^{2n} g_j^2 \leq 1$, then the (Hermitian) operator $A = \sum_j g_j\Gamma_j$ has the property

$$A^2 = \sum_{jk} g_j g_k \Gamma_j\Gamma_k = \sum_j g_j^2\mathbb{I} \leq \mathbb{I},$$

i.e. $-\mathbb{I} \leq A \leq \mathbb{I}$, so $\sigma = \frac{1}{d}(\mathbb{I} + A) \geq 0$. □

4.3.2 A meta-uncertainty relation

We now first use the above tools to prove a "meta"-uncertainty relation, from which we will then derive two new entropic uncertainty relations. Evidently, we have immediately from the above that

4.3.2. LEMMA. *Let $\rho \in \mathcal{H}$ with $\dim \mathcal{H} = 2^n$ be a quantum state, and consider $K \leq 2n+1$ anti-commuting observables Γ_j. Then,*

$$\sum_{j=0}^{K-1} \big(\mathrm{Tr}(\rho\Gamma_j)\big)^2 \leq \sum_{j=0}^{2n} \big(\mathrm{Tr}(\rho\Gamma_j)\big)^2 = \sum_{j=0}^{2n} g_j^2 \leq 1.$$

Our result is essentially a generalization of the Bloch sphere picture to higher dimensions: For $n = 1$ ($d = 2$) the state is parametrized by $\rho = \frac{1}{2}(\mathbb{I} + g_1\Gamma_1 + g_2\Gamma_2 + g_0\Gamma_0)$ where $\Gamma_1 = X$, $\Gamma_2 = Z$ and $\Gamma_0 = Y$ are the familiar Pauli matrices. Lemma 4.3.2 tells us that $g_0^2 + g_1^2 + g_2^2 \leq 1$, i.e., the state must lie inside the Bloch sphere (see Figure 2.1). Our result may be of independent interest, since it is often hard to find conditions on the coefficients $g_1, g_2, \ldots$ such that ρ is a state.

Notice that the $g_j = \mathrm{Tr}(\rho\Gamma_j)$ are directly interpreted as the expectations of the observables Γ_j. Indeed, g_j is precisely the bias of the ± 1-variable Γ_j:

$$\Pr[\Gamma_j = 1|\rho] = \frac{1+g_j}{2}.$$

Hence, we can interpret Lemma 4.3.2 as a form of uncertainty relation between the observables Γ_j: if one or more of the observables have a large bias (i.e., they are more precisely defined), this limits the bias of the other observables (i.e., they are closer to uniformly distributed).

4.3.3 Entropic uncertainty relations

It turns out that Lemma 4.3.2 has strong consequences for the Rényi and von Neumann entropic averages

$$\frac{1}{K}\sum_{j=0}^{K-1} H_\alpha\left(\Gamma_j|\rho\right),$$

where $H_\alpha(\Gamma_j|\rho)$ is the Rényi entropy at α of the probability distribution arising from measuring the state ρ with observable Γ_j. The minima over all states ρ of such expressions can be interpreted as giving entropic uncertainty relations, as we shall now do for $\alpha = 2$ (the collision entropy) and $\alpha = 1$ (the Shannon entropy).

4.3.3. THEOREM. *Let $\dim \mathcal{H} = 2^n$, and consider $K \leq 2n+1$ anti-commuting observables as defined in Eq. (4.2). Then,*

$$\min_\rho \frac{1}{K}\sum_{j=0}^{K-1} H_2\left(\Gamma_j|\rho\right) = 1 - \log\left(1 + \frac{1}{K}\right) \sim 1 - \frac{\log e}{K},$$

where $H_2(\Gamma_j|\rho) = -\log \sum_{b\in\{0,1\}} \mathrm{Tr}(\Gamma_j^b \rho)^2$*, and the minimization is taken over all states* ρ*. The latter holds asymptotically for large* K*.*

Proof. Using the fact that $\Gamma_j^b = (\mathbb{I} + (-1)^b \Gamma_j)/2$ we can first rewrite

$$\begin{aligned}\frac{1}{K}\sum_{j=0}^{K-1} H_2\left(\Gamma_j|\rho\right) &= -\frac{1}{K}\sum_{j=0}^{K-1} \log\left[\frac{1}{2}\left(1+\mathrm{Tr}(\rho\Gamma_j)^2\right)\right] \\ &\geq -\log\left(\frac{1}{2K}\sum_{j=0}^{K-1}\left(1+g_j^2\right)\right) \\ &\geq 1-\log\left(1+\frac{1}{K}\right),\end{aligned}$$

where the first inequality follows from Jensen's inequality and the concavity of the log, and the second from Lemma 4.3.2. Clearly, the minimum is attained if all $g_j = \mathrm{Tr}(\rho\Gamma_j) = \sqrt{\frac{1}{K}}$. It follows from Lemma 4.3.1 that our inequality is tight. Via the Taylor expansion of $\log\left(1+\frac{1}{K}\right)$ we obtain the asymptotic result for large K. □

For the Shannon entropy ($\alpha = 1$) we obtain something even nicer:

4.3.4. Theorem. *Let* $\dim \mathcal{H} = 2^n$*, and consider* $K \leq 2n+1$ *anti-commuting observables as defined in Eq. (4.2). Then,*

$$\min_\rho \frac{1}{K}\sum_{j=0}^{K-1} H(\Gamma_j|\rho) = 1 - \frac{1}{K},$$

where $H(\Gamma_j|\rho) = -\sum_{b\in\{0,1\}} \mathrm{Tr}(\Gamma_j^b\rho)\log \mathrm{Tr}(\Gamma_j^b\rho)$*, and the minimization is taken over all states* ρ*.*

Proof. To see this, note that by rewriting our objective as above, we observe that we need to minimize the expression

$$\frac{1}{K}\sum_{j=0}^{K-1} H\left(\frac{1\pm\sqrt{t_j}}{2}\right),$$

subject to $\sum_j t_j \leq 1$ and $t_j \geq 0$, via the identification $t_j = (\mathrm{Tr}(\rho\Gamma_j))^2$. An elementary calculation shows that the function $f(t) = H\left(\frac{1\pm\sqrt{t}}{2}\right)$ is concave in $t \in [0,1]$:

$$f'(t) = \frac{1}{4\ln 2}\frac{1}{\sqrt{t}}\big(\ln(1-\sqrt{t}) - \ln(1+\sqrt{t})\big),$$

and so

$$f''(t) = \frac{1}{8\ln 2}\frac{1}{t^{3/2}}\left(\ln\frac{1+\sqrt{t}}{1-\sqrt{t}} - \frac{2\sqrt{t}}{1-t}\right).$$

Since we are only interested in the sign of the second derivative, we ignore the (positive) factors in front of the bracket, and are done if we can show that

$$\begin{aligned} g(t) &:= \ln\frac{1+\sqrt{t}}{1-\sqrt{t}} - \frac{2\sqrt{t}}{1-t} \\ &= \ln(1+\sqrt{t}) + \frac{1}{1+\sqrt{t}} - \ln(1-\sqrt{t}) - \frac{1}{1-\sqrt{t}} \end{aligned}$$

is non-positive for $0 \leq t \leq 1$. Substituting $s = 1 - \sqrt{t}$, which is also between 0 and 1, we rewrite this as

$$h(s) = -\ln s - \frac{1}{s} + \ln(2-s) + \frac{1}{2-s},$$

which has derivative

$$h'(s) = (1-s)\left(\frac{1}{s^2} - \frac{1}{(2-s)^2}\right),$$

and this is clearly positive for $0 < s < 1$. In other words, h increases from its value at $s = 0$ (where it is $h(0) = -\infty$) to its value at $s = 1$ (where it is $h(1) = 0$), so indeed $h(s) \leq 0$ for all $0 \leq s \leq 1$. Consequently, also $f''(t) \leq 0$ for $0 \leq t \leq 1$.

Hence, by Jensen's inequality, the minimum is attained with one of the t_j being 1 and the others 0, giving just the lower bound of $1 - \frac{1}{K}$. □

We have shown that anti-commuting Clifford observables obey the strongest possible uncertainty relation for the von Neumann entropy. It is interesting that in the process of the proof, however, we have found three uncertainty type inequalities (the sum of squares bound, the bound on H_2, and finally the bound on H_1), and all three have a different structure of attaining the limit. The sum of squares bound can be achieved in every direction (meaning for every tuple satisfying the bound we get one attaining it by multiplying all components by some appropriate factor), the H_2 expression requires all components to be equal, while the H_1 expression demands exactly the opposite.

4.4 Conclusion

We showed that merely choosing our measurements to be mutually unbiased does not lead to strong uncertainty relations. However, we were able to identify another property which does lead to optimal entropic uncertainty relations for two outcome measurements! *Anti-commuting* Clifford observables obey the strongest

possible uncertainty relation for the von Neumann entropy: if we have no uncertainty for one of the measurements, we have maximum uncertainty for all others. We also obtain a slightly suboptimal uncertainty relation for the collision entropy which is strong enough for all cryptographic purposes. Indeed, one could use our entropic uncertainty relation in the bounded quantum storage setting to construct, for example, 1-K oblivious transfer protocols analogous to [DFR$^+$07]. Here, instead of encoding a single bit into either the computational or Hadamard basis, which gives us a 1-2 OT, we now encode a single bit into the positive or negative eigenspace of each of these K operators. It is clear from the representation of such operators discussed earlier, that such an encoding can be done experimentally as easily as encoding a single bit into three mutually unbiased basis given by σ_x, σ_y, σ_z. Indeed, our construction can be seen as a direct extension of such an encoding: we obtain the uncertainty relation for the three MUBs previously proved by Sanchez [San93, SR95] as a special case of our analysis for $K = 3$. It is perhaps interesting to note that the same operators also play a prominent role in the setting of non-local games as discussed in Chapter 6.3.2.

Sadly, strong uncertainty relations for measurements with more than two outcomes remain inaccessible to us. It has been shown [Feh07] that uncertainty relations for more outcomes can be obtained via a coding argument from uncertainty relations as we construct them here. Yet, these are far from optimal. A natural choice would be to consider the generators of a generalized Clifford algebra, yet such an algebra does not have such nice symmetry properties which enabled us to implement operations on the vector components above. It remains an exciting open question whether such operators form a good generalization, or whether we must continue our search for new properties.

Chapter 5

Locking classical information

Locking classical correlations in quantum states [DHL+04] is an exciting feature of quantum information, intricately related to entropic uncertainty relations. In this chapter, we will investigate whether good locking effects can be obtained using mutually unbiased bases.

5.1 Introduction

Consider a two-party protocol with one or more rounds of communication. Intuitively, one would expect that in each round the amount of correlation between the two parties cannot increase by much more than the amount of data transmitted. For example, transmitting 2ℓ classical bits or ℓ qubits (and using superdense coding) should not increase the amount of correlation by more than 2ℓ bits, no matter what the initial state of the two-party system was. This intuition is accurate when we take the classical mutual information $\mathcal{I}_c$ as our correlation measure, and require all communication to be classical. However, when quantum communication was possible at some point during the protocol, everything changes: there exist two-party mixed quantum states, such that transmitting just a single extra bit of classical communication can result in an arbitrarily large increase in $\mathcal{I}_c$ [DHL+04]. The magnitude of this increase thereby only depends on the dimension of the initial mixed state. Since then similar locking effects have been observed, also for other correlation measures [CW05b, HHHO05]. Such effects play a role in very different scenarios: they have been used to explain physical phenomena related to black holes [SO06], but they are also important in cryptographic applications such as quantum key distribution [KRBM07] and quantum bit string commitment that we will encounter in Chapter 10. We are thus interested in determining how exactly we can obtain locking effects, and how dramatic they can be.

5.1.1 A locking protocol

The correlation measure considered here, is the classical mutual information of a bipartite quantum state ρ_{AB}, which is the maximum classical mutual information that can be obtained by local measurements $M_A \otimes M_B$ on the state ρ_{AB} (see Chapter 2):

$$\mathcal{I}_c(\rho_{AB}) = \max_{M_A \otimes M_B} \mathcal{I}(A, B). \tag{5.1}$$

Recall from Chapter 2 that the mutual information is defined as $\mathcal{I}(A, B) = H(P_A) + H(P_B) - H(P_{AB})$ where H is the Shannon entropy. P_A, P_B, and P_{AB} are the probability distributions corresponding to the individual and joint outcomes of measuring the state ρ_{AB} with $M_A \otimes M_B$. The mutual information between A and B is a measure of the information that B contains about A. This measure of correlation is of particular relevance for quantum bit string commitments in Chapter 10. Furthermore, the first locking effect was observed for this quantity in the following protocol between two parties: Alice (A) and Bob (B). Let $\mathbb{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_m\}$ with $\mathcal{B}_t = \{|b_1^t\rangle, \ldots, |b_d^t\rangle\}$ be a set of m MUBs in $\mathbb{C}^d$. Alice picks an element $k \in \{1, \ldots, d\}$ and a basis $\mathcal{B}_t \in \mathbb{B}$ uniformly at random. She then sends $|b_k^t\rangle$ to Bob, while keeping t secret. Such a protocol gives rise to the joint state

$$\rho_{AB} = \frac{1}{md} \sum_{k=1}^{d} \sum_{t=1}^{m} (|k\rangle\langle k| \otimes |t\rangle\langle t|)_A \otimes (|b_k^t\rangle\langle b_k^t|)_B.$$

Clearly, if Alice told her basis choice t to Bob, he could measure in the right basis and obtain the correct k. Alice and Bob would then share $\log d + \log m$ bits of correlation, which is also their mutual information $\mathcal{I}_c(\sigma_{AB})$, where σ_{AB} is the state obtained from ρ_{AB} after the announcement of t. But, how large is $\mathcal{I}_c(\rho_{AB})$, when Alice does *not* announce t to Bob? It was shown [DHL$^+$04] that in dimension $d = 2^n$, using the two MUBs given by the unitaries $U_+ = \mathbb{I}^{\otimes n}$ and $U_\times = H^{\otimes n}$ applied to the computational basis we have $\mathcal{I}_c(\rho_{AB}) = (1/2) \log d$ (see Figure 5.1, where $|x_b\rangle = U_b|x\rangle$). This means that the single bit of basis information Alice transmits to Bob "unlocks" $(1/2) \log d$ bits: *without* this bit, the mutual information is $(1/2) \log d$, but *with* this bit it is $\log d + 1$. To get a good locking protocol, we want to use only a small number of bases, i.e., m should be as small as possible, while at the same time forcing $\mathcal{I}_c(\rho_{AB})$ to be as low as possible. That is, we want $\log m / (\log d - \mathcal{I}_c(\rho_{AB}))$ to be small.

It is also known that if Alice and Bob randomly choose a large set of unitaries from the Haar measure to construct $\mathbb{B}$, then $\mathcal{I}_c(\rho_{AB})$ can be brought down to a small constant [HLSW04]. However, no explicit constructions with more than two bases are known that give good locking effects. Based on numerical studies for spaces of prime dimension $3 \leq d \leq 30$, one might hope that adding a third MUB would strengthen the locking effect and give $\mathcal{I}_c(\rho_{AB}) \approx (1/3) \log d$ [DHL$^+$04].

Here, however, we show that this intuition fails us. We prove that for three MUBs given by $\mathbb{I}^{\otimes n}$, $H^{\otimes n}$, and $K^{\otimes n}$ where $K = (\mathbb{I} + i\sigma_x)/\sqrt{2}$ and dimension

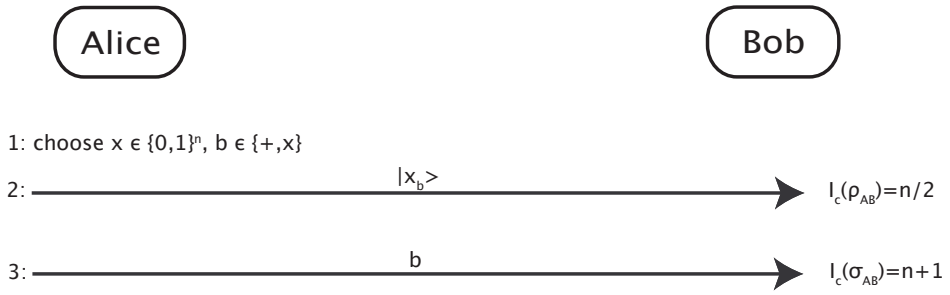


Figure 5.1: A locking protocol for 2 bases.

$d = 2^n$ for some even integer n, we have

$$\mathcal{I}_c(\rho_{AB}) = \frac{1}{2}\log d, \tag{5.2}$$

the same locking effect as with two MUBs. We also show that for any subset of the MUBs based on Latin squares and the MUBs in square dimensions based on generalized Pauli matrices [BBRV02], we again obtain Eq. (5.2), i.e., using two or all $\sqrt{d}$ of them makes no difference at all! Finally, we show that for any set of MUBs $\mathbb{B}$ based on generalized Pauli matrices in *any* dimension, $\mathcal{I}_c(\rho_{AB}) = \log d - \min_{|\phi\rangle}(1/|\mathbb{B}|)\sum_{\mathcal{B}\in\mathbb{B}} H(\mathcal{B}||\phi\rangle)$, i.e., it is enough to determine a bound on the entropic uncertainty relation to determine the strength of the locking effect. Although bounds for general MUBs still elude us, our results show that merely choosing the bases to be mutually unbiased is not sufficient and we must look elsewhere to find bases which provide good locking.

5.1.2 Locking and uncertainty relations

We first explain the connection between locking and entropic uncertainty relations. In particular, we will see that for MUBs based on generalized Pauli matrices, we only need to look at such uncertainty relations to determine the exact strength of the locking effect.

In order to determine how large the locking effect is for some set of mutually unbiased bases $\mathbb{B}$, and the shared state

$$\rho_{AB} = \sum_{t=1}^{|\mathbb{B}|}\sum_{k=1}^{d} p_{t,k}(|k\rangle\langle k| \otimes |t\rangle\langle t|)_A \otimes (|b_k^t\rangle\langle b_k^t|)_B, \tag{5.3}$$

we must find the value of $\mathcal{I}_c(\rho_{AB})$ or at least a good upper bound. That is, we must find a POVM $M_A \otimes M_B$ that maximizes Eq. (5.1). Here, $\{p_{t,k}\}$ is a probability distribution over $\mathbb{B} \times [d]$. It has been shown in [DHL$^+$04] that we can

restrict ourselves to taking M_A to be the local measurement determined by the projectors $\{|k\rangle\langle k| \otimes |t\rangle\langle t|\}$. It is also known that we can limit ourselves to take the measurement M_B consisting of rank one elements $\{\alpha_i|\Phi_i\rangle\langle\Phi_i|\}$ only [Dav78], where $\alpha_i \geq 0$ and $|\Phi_i\rangle$ is normalized. Maximizing over M_B then corresponds to maximizing Bob's accessible information as defined in Chapter 2 for the ensemble $\mathcal{E} = \{p_{k,t}, |b_k^t\rangle\langle b_k^t|\}$

$$\mathcal{I}_{acc}(\mathcal{E}) = \max_{M_B} \left(-\sum_{k,t} p_{k,t} \log p_{k,t} + \sum_i \sum_{k,t} p_{k,t} \alpha_i \langle\Phi_i|\rho_{k,t}|\Phi_i\rangle \log \frac{p_{k,t}\langle\Phi_i|\rho_{k,t}|\Phi_i\rangle}{\langle\Phi_i|\mu|\Phi_i\rangle} \right), \tag{5.4}$$

where $\mu = \sum_{k,t} p_{k,t}\rho_{k,t}$ and $\rho_{k,t} = |b_k^t\rangle\langle b_k^t|$. Therefore, we have $\mathcal{I}_c(\rho_{AB}) = \mathcal{I}_{acc}(\mathcal{E})$. As we saw in Chapter 2, maximizing the accessible information is often a very hard task. Nevertheless, for our choice of MUBs, the problem will turn out to be quite easy in the end.

5.2 Locking using mutually unbiased bases

5.2.1 An example

We now determine how well we can lock information using specific sets of mutually unbiased bases. We first consider a very simple example with only three MUBs that provides the intuition behind the remainder of our proof. The three MUBs we consider now are generated by the unitaries $\mathbb{I}$, H and $K = (\mathbb{I} + i\sigma_x)/\sqrt{2}$ when applied to the computational basis. For this small example, we also investigate the role of the prior over the bases and the encoded basis elements. It turns out that this does not affect the strength of the locking effect positively, i.e., we do not obtain a stronger locking affect using a non-uniform prior. Actually, it is possible to show the same for encodings in many other bases. However, we do not consider this case in full generality as to not obscure our main line of argument.

5.2.1. Lemma. *Let $U_1 = \mathbb{I}^{\otimes n}$,$U_2 = H^{\otimes n}$, and $U_3 = K^{\otimes n}$, and take $k \in \{0,1\}^n$ where n is an even integer. Let $\{p_t\}$ with $t \in [3]$ be a probability distribution over the set $\mathcal{S} = \{U_1, U_2, U_3\}$. Suppose that $p_1, p_2, p_3 \leq 1/2$ and let $\{p_{t,k}\}$ with $p_{t,k} = p_t/d$ be the joint distribution over $\mathcal{S} \times \{0,1\}^n$. Consider the ensemble $\mathcal{E} = \{p_t \frac{1}{d}, U_t|k\rangle\langle k|U_t^\dagger\}$, then*

$$\mathcal{I}_{acc}(\mathcal{E}) = \frac{n}{2}.$$

If, on the other hand, there exists a $t \in [3]$ such that $p_t > 1/2$, then $\mathcal{I}_{acc}(\mathcal{E}) > n/2$.

Proof. We first give an explicit measurement strategy and then prove a matching upper bound on $\mathcal{I}_{acc}$. Consider the Bell basis vectors $|\Gamma_{00}\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$,

$|\Gamma_{01}\rangle = (|00\rangle - |11\rangle)/\sqrt{2}$, $|\Gamma_{10}\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$, and $|\Gamma_{11}\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$. Note that we can write for the computational basis

$$\begin{aligned} |00\rangle &= \frac{1}{\sqrt{2}}(|\Gamma_{00}\rangle + |\Gamma_{01}\rangle), \\ |01\rangle &= \frac{1}{\sqrt{2}}(|\Gamma_{10}\rangle + |\Gamma_{11}\rangle), \\ |10\rangle &= \frac{1}{\sqrt{2}}(|\Gamma_{10}\rangle - |\Gamma_{11}\rangle), \\ |11\rangle &= \frac{1}{\sqrt{2}}(|\Gamma_{00}\rangle - |\Gamma_{01}\rangle). \end{aligned}$$

The crucial fact to note is that if we fix some k_1, k_2, then there exist exactly two Bell basis vectors $|\Gamma_{i_1 i_2}\rangle$ such that $|\langle\Gamma_{i_1 i_2}|k_1, k_2\rangle|^2 = 1/2$. For the remaining two basis vectors the inner product with $|k_1, k_2\rangle$ will be zero. A simple calculation shows that we can express the two-qubit basis states of the other two mutually unbiased bases analogously: for each two qubit basis state there are exactly two Bell basis vectors such that the inner product is zero and for the other two the inner product squared is $1/2$.

We now take the measurement given by $\{|\Gamma_i\rangle\langle\Gamma_i|\}$ with $|\Gamma_i\rangle = |\Gamma_{i_1 i_2}\rangle \otimes \ldots \otimes |\Gamma_{i_{n-1} i_n}\rangle$ for the binary expansion of $i = i_1 i_2 \ldots i_n$. Fix a $k = k_1 k_2 \ldots k_n$. By the above argument, there exist exactly $2^{n/2}$ strings $i \in \{0,1\}^n$ such that $|\langle\Gamma_i|k\rangle|^2 = 1/2^{n/2}$. Putting everything together, Eq. (5.4) now gives us for any prior distribution $\{p_{t,k}\}$ that

$$-\sum_i \langle\Gamma_i|\mu|\Gamma_i\rangle \log\langle\Gamma_i|\mu|\Gamma_i\rangle - \frac{n}{2} \leq \mathcal{I}_{acc}(\mathcal{E}). \tag{5.5}$$

For our particular distribution we have $\mu = \mathbb{I}/d$ and thus

$$\frac{n}{2} \leq \mathcal{I}_{acc}(\mathcal{E}).$$

We now prove a matching upper bound that shows that our measurement is optimal. For our distribution, we can rewrite Eq. (5.4) for the POVM given by $\{\alpha_i|\Phi_i\rangle\langle\Phi_i|\}$ to

$$\begin{aligned} \mathcal{I}_{acc}(\mathcal{E}) &= \max_M \left(\log d + \sum_i \frac{\alpha_i}{d} \sum_{k,t} p_t |\langle\Phi_i|U_t|k\rangle|^2 \log |\langle\Phi_i|U_t|k\rangle|^2 \right) \\ &= \max_M \left(\log d - \sum_i \frac{\alpha_i}{d} \sum_t p_t H(\mathcal{B}_t||\Phi_i\rangle) \right), \end{aligned}$$

for the bases $\mathcal{B}_t = \{U_t|k\rangle \mid k \in \{0,1\}^n\}$.

It follows from Corollary 4.2.2 that $\forall i \in \{0,1\}^n$ and $p_1, p_2, p_3 \leq 1/2$

$$\begin{aligned}(1/2 - p_1)[H(\mathcal{B}_2||\Phi_i\rangle) + H(\mathcal{B}_3||\Phi_i\rangle)] \quad &+ \\ (1/2 - p_2)[H(\mathcal{B}_1||\Phi_i\rangle) + H(\mathcal{B}_3||\Phi_i\rangle)] \quad &+ \\ (1/2 - p_3)[H(\mathcal{B}_1||\Phi_i\rangle) + H(\mathcal{B}_2||\Phi_i\rangle)] \quad &\geq n/2,\end{aligned}$$

where we used the fact that $p_1 + p_2 + p_3 = 1$. Reordering the terms we now get $\sum_{t=1}^{3} p_t H(\mathcal{B}_t||\Phi_i\rangle) \geq n/2$. Putting things together and using the fact that $\sum_i \alpha_i = d$, we obtain

$$\mathcal{I}_{acc}(\mathcal{E}) \leq \frac{n}{2},$$

from which the result follows.

If, on the other hand, there exists a $t \in [3]$ such that $p_t > 1/2$, then by measuring in the basis $\mathcal{B}_t$ we obtain $\mathcal{I}_{acc}(\mathcal{E}) \geq p_t n > n/2$, since the entropy will be 0 for basis $\mathcal{B}_t$ and we have $\sum_t p_t = 1$. □

Above, we have only considered a non-uniform prior over the set of *bases*. In Chapter 3, we observed that when we want to guess the XOR of a string of length 2 encoded in one (unknown to us) of these three bases, the uniform prior on the strings is not the one that gives the smallest probability of success. This might lead one to think that a similar phenomenon could be observed in the present setting, i.e., that one might obtain better locking with three basis for a non-uniform prior on the strings. In what follows, however, we show that this is not the case.

Let $p_t = \sum_k p_{k,t}$ be the marginal distribution on the basis, then the difference in Bob's knowledge between receiving only the quantum state and receiving the quantum state *and* the basis information, where we will ignore the basis information itself, is given by

$$\Delta(p_{k,t}) = H(p_{k,t}) - \mathcal{I}_{acc}(\mathcal{E}) - H(p_t),$$

Consider the post-measurement state $\nu = \sum_i \langle\Gamma_i|\mu|\Gamma_i\rangle|\Gamma_i\rangle\langle\Gamma_i|$. Using Eq. (5.5) we obtain

$$\Delta(p_{k,t}) \leq H(p_{k,t}) - S(\nu) + n/2 - H(p_t), \tag{5.6}$$

where S is the von Neumann entropy. Consider the state

$$\rho_{12} = \sum_{k=1}^{d} \sum_{t=1}^{3} p_{k,t} (|t\rangle\langle t|)_1 \otimes (U_t|k\rangle\langle k|U_t^\dagger)_2,$$

for which we have that

$$\begin{aligned}S(\rho_{12}) = H(p_{k,t}) \quad &\leq S(\rho_1) + S(\rho_2) \\ &= H(p_t) + S(\mu) \\ &\leq H(p_t) + S(\nu).\end{aligned}$$

Using Eq. (5.6) and the previous equation we get

$$\Delta(p_{k,t}) \leq n/2,$$

for any prior distribution. This bound is saturated by the uniform prior and therefore we conclude that the uniform prior results in the largest gap possible.

5.2.2 MUBs from generalized Pauli matrices

We now consider MUBs based on the generalized Pauli matrices X_d and Z_d as described in Chapter 2.4.2. We consider a uniform prior over the elements of each basis and the set of bases. Choosing a non-uniform prior does not lead to a better locking effect.

5.2.2. Lemma. *Let $\mathbb{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_m\}$ be any set of MUBs constructed on the basis of generalized Pauli matrices in a Hilbert space of prime power dimension $d = p^N$. Consider the ensemble $\mathcal{E} = \{\frac{1}{dm}, |b_k^t\rangle\langle b_k^t|\}$. Then*

$$\mathcal{I}_{acc}(\mathcal{E}) = \log d - \frac{1}{m} \min_{|\psi\rangle} \sum_{\mathcal{B}_t \in \mathbb{B}} H(\mathcal{B}_t||\psi\rangle).$$

Proof. We can rewrite Eq. (5.4) for a POVM M_B of the form $\{\alpha_i|\Phi_i\rangle\langle\Phi_i|\}$ as

$$\begin{aligned} \mathcal{I}_{acc}(\mathcal{E}) &= \max_{M_B} \left(\log d + \sum_i \frac{\alpha_i}{dm} \sum_{k,t} |\langle\Phi_i|b_k^t\rangle|^2 \log |\langle\Phi_i|b_k^t\rangle|^2 \right) \\ &= \max_{M_B} \left(\log d - \sum_i \frac{\alpha_i}{d} \sum_t p_t H(\mathcal{B}_t||\Phi_i\rangle) \right). \end{aligned}$$

For convenience, we split up the index i into $i = a, b$ with $a = a_1, \ldots, a_N$ and $b = b_1, \ldots, b_N$, where $a_\ell, b_\ell \in \{0, \ldots, p-1\}$ in the following.

We first show that applying generalized Pauli matrices to the basis vectors of a MUB merely permutes those vectors.

1. Claim. *Let $\mathcal{B}_t = \{|b_1^t\rangle, \ldots, |b_d^t\rangle\}$ be a basis based on generalized Pauli matrices (Chapter 2.4.2) with $d = p^N$. Then $\forall a, b \in \{0, \ldots, p-1\}^N, \forall k \in [d]$ we have that $\exists k' \in [d]$, such that $|b_{k'}^t\rangle = X_d^{a_1} Z_d^{b_1} \otimes \ldots \otimes X_d^{a_N} Z_d^{b_N} |b_k^t\rangle$.*

Proof. Let $\mathcal{T}_p^i$ for $i \in \{0, 1, 2, 3\}$ denote the generalized Pauli's $\mathcal{T}_p^0 = \mathbb{I}_p$, $\mathcal{T}_p^1 = X_p$, $\mathcal{T}_p^3 = Z_p$, and $\mathcal{T}_p^2 = X_p Z_p$. Note that $X_p^u Z_p^v = \omega^{uv} Z_p^v X_p^u$, where $\omega = e^{2\pi i/p}$. Furthermore, define $\mathcal{T}_p^{i,(x)} = \mathbb{I}^{\otimes(x-1)} \otimes \mathcal{T}_p^i \otimes \mathbb{I}^{N-x}$ to be the Pauli operator $\mathcal{T}_p^i$ applied to the x-th qupit. Recall from Section 2.4.2 that there exist sets of Pauli operators C_t such that the basis $\mathcal{B}_t$ is the unique simultaneous eigenbasis of the set of operators in C_t, i.e., for all $k \in [d]$ and $f, g \in [N]$,

$|b_k^t\rangle \in \mathcal{B}_t$ and $c_{f,g}^t \in C_t$, we have $c_{f,g}^t|b_k^t\rangle = \lambda_{k,f,g}^t|b_k^t\rangle$ for some value $\lambda_{k,f,g}^t$. Note that any vector $|v\rangle$ that satisfies this equation is proportional to a vector in $\mathcal{B}_t$. To prove that any application of one of the generalized Paulis merely permutes the vectors in $\mathcal{B}_t$ is therefore equivalent to proving that $\mathcal{T}_p^{i,(x)}|b_k^t\rangle$ are eigenvectors of $c_{f,g}^t$ for any $f, g \in [k]$ and $i \in \{1, 3\}$. This can be seen as follows: Note that $c_{f,g}^t = \bigotimes_{n=1}^N \left(\mathcal{T}_p^{1,(n)}\right)^{f_N} \left(\mathcal{T}_p^{3,(n)}\right)^{g_N}$ for $f = (f_1, \ldots, f_N)$ and $g = (g_1, \ldots, g_N)$ with $f_N, g_N \in \{0, \ldots, p-1\}$ [BBRV02]. A calculation then shows that

$$c_{f,g}^t \mathcal{T}_p^{i,(x)}|b_k^t\rangle = \tau_{f_x,g_x,i}\lambda_{k,f,g}^t \mathcal{T}_p^{i,(x)}|b_k^t\rangle,$$

where $\tau_{f_x,g_x,i} = \omega^{g_x}$ for $i = 1$ and $\tau_{f_x,g_x,i} = \omega^{-f_x}$ for $i = 3$. Thus $\mathcal{T}_p^{i,(x)}|b_k^t\rangle$ is an eigenvector of $c_{f,g}^t$ for all t, f, g and i, which proves our claim. □

Suppose we are given $|\psi\rangle$ that minimizes $\sum_{\mathcal{B}_t \in \mathbb{T}} H(\mathcal{B}_t||\psi\rangle)$. We can then construct a full POVM with d^2 elements by taking $\{\frac{1}{d}|\Phi_{ab}\rangle\langle\Phi_{ab}|\}$ with $|\Phi_{ab}\rangle = (X_d^{a_1} Z_d^{b_1} \otimes \ldots \otimes X_d^{a_N} Z_d^{b_N})^\dagger|\psi\rangle$. However, it follows from our claim above that $\forall a, b, k, \exists k'$ such that $|\langle\Phi_{ab}|b_k^t\rangle|^2 = |\langle\psi|b_{k'}^t\rangle|^2$, and thus $H(\mathcal{B}_t||\psi\rangle) = H(\mathcal{B}_t||\Phi_{ab}\rangle)$ from which the result follows. □

Determining the strength of the locking effects for such MUBs is thus equivalent to proving bounds on entropic uncertainty relations. We thus obtain as a corollary of Theorem 4.2.3 and Lemma 5.2.2, that, for dimensions which are the square of a prime power (i.e. $d = p^{2N}$), using any product MUBs based on generalized Paulis does not give us any better locking than just using 2 MUBs.

5.2.3. Corollary. *Let $\mathbb{S} = \{\mathcal{S}_1, \ldots, \mathcal{S}_m\}$ with $m \geq 2$ be any set of MUBs constructed on the basis of generalized Pauli matrices in a Hilbert space of prime (power) dimension $s = p^N$. Define U_t as the unitary that transforms the computational basis into the t-th MUB, i.e., $\mathcal{S}_t = \{U_t|1\rangle, \ldots, U_t|s\rangle\}$. Let $\mathbb{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_m\}$ be the set of product MUBs with $\mathcal{B}_t = \{U_t \otimes U_t^*|1\rangle, \ldots, U_t \otimes U_t^*|d\rangle\}$ in dimension $d = s^2$. Consider the ensemble $\mathcal{E} = \{\frac{1}{dm}, |b_k^t\rangle\langle b_k^t|\}$. Then*

$$\mathcal{I}_{acc}(\mathcal{E}) = \frac{\log d}{2}.$$

Proof. The claim follows from Theorem 4.2.3 and the proof of Lemma 5.2.2, by constructing a similar measurement formed from vectors $|\hat{\Phi}_{\hat{a}\hat{b}}\rangle = K_{a^1b^1} \otimes K_{a^2b^2}^*|\psi\rangle$ with $\hat{a} = a^1a^2$ and $\hat{b} = b^1b^2$, where a^1, a^2 and b^1, b^2 are defined like a and b in the proof of Lemma 5.2.2, and $K_{ab} = (X_d^{a_1} Z_d^{b_1} \otimes \ldots \otimes X_d^{a_N} Z_d^{b_N})^\dagger$ from above. □

The simple example we considered above is in fact a special case of Corollary 5.2.3. It shows that if the vector that minimizes the sum of entropies has certain symmetries, the resulting POVM can even be much simpler. For example, the Bell states are vectors which such symmetries.

5.2.3 MUBs from Latin squares

At first glance, one might think that maybe the product MUBs based on generalized Paulis are not well suited for locking just because of their product form. Perhaps MUBs with entangled basis vectors do not exhibit this problem? Let's examine how well MUBs based on Latin squares can lock classical information in a quantum state. All such MUBs are highly entangled, with the exception of the two extra MUBs based on non-Latin squares. Surprisingly, it turns out, however, that *any* set of at least two MUBs based on Latin squares, does equally well at locking as using just 2 such MUBs. Thus such MUBs perform equally "badly", i.e., we cannot improve the strength of the locking effect by using more MUBs of this type.

5.2.4. Lemma. *Let* $\mathbb{B} = \{\mathcal{B}_1, \ldots, \mathcal{B}_m\}$ *with* $m \geq 2$ *be any set of MUBs in a Hilbert space of dimension* $d = s^2$ *constructed on the basis of Latin squares. Consider the ensemble* $\mathcal{E} = \{\frac{1}{dm}, |b_k^t\rangle\langle b_k^t|\}$. *Then*

$$\mathcal{I}_{acc}(\mathcal{E}) = \frac{\log d}{2}.$$

Proof. Note that we can again rewrite $\mathcal{I}_{acc}(\mathcal{E})$ as in the proof of Lemma 5.2.2. Consider the simple measurement in the computational basis $\{|i,j\rangle\langle i,j| \mid i,j \in [s]\}$. The result then follows by the same argument as in Lemma 4.2.4. □

Intuitively, our measurement outputs one sub-square of the Latin square used to construct the MUBs as depicted in Figure 5.2.3. As we saw in the construction of MUBs based on Latin squares in Chapter 2.4.1, each entry "occurs" in exactly $\sqrt{d} = s$ MUBs.

1	2	3
2	3	1
3	1	2

Figure 5.2: Measurement for $|1,1\rangle$.

5.3 Conclusion

We have shown tight bounds on locking for specific sets of mutually unbiased bases. Surprisingly, it turns out that using more mutually unbiased basis does not

always lead to a better locking effect. It is interesting to consider what may make these bases so special. The example of three MUBs considered in Lemma 5.2.1 may provide a clue. These three bases are given by the common eigenbases of $\{\sigma_x \otimes \sigma_x, \sigma_x \otimes \mathbb{I}, \mathbb{I} \otimes \sigma_x\}$, $\{\sigma_z \otimes \sigma_z, \sigma_z \otimes \mathbb{I}, \mathbb{I} \otimes \sigma_z\}$ and $\{\sigma_y \otimes \sigma_y, \sigma_y \otimes \mathbb{I}, \mathbb{I} \otimes \sigma_y\}$ respectively [BBRV02]. However, $\sigma_x \otimes \sigma_x$, $\sigma_z \otimes \sigma_z$ and $\sigma_y \otimes \sigma_y$ commute and thus also share a common eigenbasis, namely the Bell basis. This is exactly the basis we will use as our measurement. For all MUBs based on generalized Pauli matrices, the MUBs in prime power dimensions are given as the common eigenbasis of similar sets consisting of strings of Paulis. It would be interesting to determine the strength of the locking effect on the basis of the commutation relations of elements of *different* sets. Furthermore, perhaps it is possible to obtain good locking from a subset of such MUBs where none of the elements from different sets commute.

It is also worth noting that the numerical results of [DHL$^+$04] indicate that at least in dimension p using more than three bases does indeed lead to a stronger locking effect. It would be interesting to know, whether the strength of the locking effect depends not only on the number of bases, but also on the dimension of the system in question.

Whereas general bounds still elude us, we have shown that merely choosing mutually unbiased bases is not sufficient to obtain good locking effects. We thus have to look for different properties. Sadly, whereas we were able to obtain good uncertainty relations in Chapter 4.3, the same approach does not work here: To obtain good locking we must not only find good uncertainty relations, but also find a way to encode many bits using only a small number of encodings.

Part III

Entanglement

Chapter 6

Introduction

Entanglement is possibly the most intriguing element of quantum theory. It plays a crucial role in quantum algorithms, quantum cryptography and the understanding of quantum mechanics itself. It enables us to perform quantum teleportation, as well as superdense coding [NC00]. In this part, we investigate one particular aspect of quantum entanglement: the violation of Bell-inequalities, and their implications for classical protocols. But first, let's take a brief look at the history of entanglement, and introduce the essential ingredients we need later.

6.1 Introduction

In 1935, Einstein, Podolsky and Rosen (EPR) identified one of the striking consequences of what latter became known as entanglement. In their seminal article [EPR35] "Can Quantum Mechanical Description of Physical Reality Be Considered Complete?" the authors define "elements of reality" as follows:

> *If, without in any way disturbing a system, we can predict with certainty (i.e. with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity.*

EPR call a theory that satisfies this condition *complete*. They put forward the now famous EPR-Paradox, here stated informally using discrete variables as put forward by Bohm [Per93]. EPR assume that if we have a state shared between two spatially separated systems, Alice and Bob, that do not interact at the time of a measurement,

> *no real change can take place in the second system as a consequence of anything that may be done to the first system.*

That means that Alice and Bob cannot use the shared state itself to transmit information. We will also refer to this as the *no-signaling* condition. Now consider

the shared state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(\underbrace{|0\rangle}_{Alice}\underbrace{|0\rangle}_{Bob} + \underbrace{|1\rangle}_{Alice}\underbrace{|1\rangle}_{Bob}) = \frac{1}{\sqrt{2}}(\underbrace{|+\rangle}_{Alice}\underbrace{|+\rangle}_{Bob} + \underbrace{|-\rangle}_{Alice}\underbrace{|-\rangle}_{Bob}). \tag{6.1}$$

Suppose that we measure Alice's system in the computational basis to obtain outcome c_A. Note that we can now predict the outcome of a measurement of Bob's system in the computational basis with certainty: $c_B = c_A$, without having disturbed Bob's system in any way. Thus c_B is an "element of physical reality". However, we might as well have measured Alice's system in the Hadamard basis to obtain outcome h_A. Likewise, we can now predict with certainty the outcome of measuring Bob's system in the Hadamard basis, $h_B = h_A$, again without causing any disturbance to the second system. Thus h_B should also be an "element of physical reality". But as we saw in Chapter 4, quantum mechanics forbids us to assign exact values to both c_B and h_B simultaneously, as measurements in the computational and Hadamard basis are non-commutative. Indeed, in Chapter 4.2, we saw that these two measurements give the strongest entropic uncertainty relation for two measurements. EPR thus conclude

> *that the quantum mechanical description of reality given by the wave function is not complete.*

EPR's article spurred a flurry of discussion that continues up to the present day. Shortly after the publication of their article, Schrödinger published two papers in which he coined the term entanglement (German: Verschränkung) [Sch35b, Sch35a] and investigated this phenomenon which he described as "not one, but rather *the* characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought" [Sch35b]. One point of discussion in the ensuing years was whether the fact that quantum mechanics is not complete, means that there might exist a more detailed description of nature which *is* complete. Even though, these more detailed descriptions also called "hidden variables" had remained inaccessible to us so far: a better theory and better technology might enable us to learn them. Thus quantum mechanical observations would merely appear to be probabilistic in the absence of our knowledge of such hidden variables.

6.1.1 Bell's inequality

This idea was put to rest by Bell [Bel65] in 1964, when he proposed conditions that *any* classical theory, i.e. any theory based on local hidden variables, has to satisfy, and which can be verified experimentally. These conditions are known as *Bell inequalities*. Intuitively, Bell inequalities measure the strength of non-local correlations attainable in any classical theory. Non-local correlations arise as the result of measurements performed on a quantum system shared between

two spatially separated parties. Imagine two parties, Alice and Bob, who are given access to a shared quantum state $|\Psi\rangle$, but cannot communicate. In the simplest case, each of them is able to perform one of two possible measurements. Every measurement has two possible outcomes labeled ± 1. Alice and Bob now measure $|\Psi\rangle$ using an independently chosen measurement setting and record their outcomes. In order to obtain an accurate estimate for the correlation between their measurement settings and the measurement outcomes, they perform this experiment independently many times using an identically prepared state $|\Psi\rangle$ in each round.

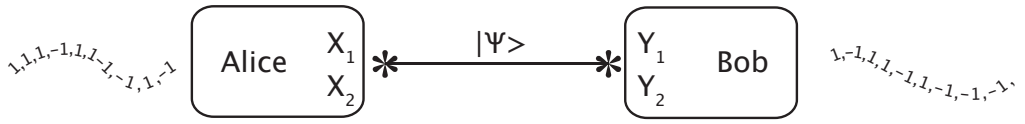


Figure 6.1: Alice and Bob measure many copies of $|\Psi\rangle$

Both classical and quantum theories impose limits on the strength of non-local correlations. In particular, both should not violate the non-signaling condition of special relativity as put forward by EPR above. That is, the local choice of the measurement setting does not allow Alice and Bob to transmit information. Limits on the strength of correlations which are possible in the framework of any *classical* theory are the Bell inequalities. The best known Bell inequality is the Clauser, Horne, Shimony and Holt (CHSH) inequality [CHSH69]

$$\langle CHSH\rangle_c = |\langle X_1Y_1\rangle + \langle X_1Y_2\rangle + \langle X_2Y_1\rangle - \langle X_2Y_2\rangle| \leq 2, \tag{6.2}$$

where X_1, X_2 and Y_1, Y_2 are the observables representing the measurement settings of Alice and Bob respectively and we use $\langle X_iY_j\rangle = \langle\Psi|X_i \otimes Y_j|\Psi\rangle$ to denote the mean value of X_i and Y_j. Quantum mechanics allows for a violation of the CHSH inequality, and is thus indeed non-classical: If we take the shared state $|\Psi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ and let $X_1 = \sigma_x$, $X_2 = \sigma_z$, $Y_1 = (\sigma_x + \sigma_z)/2$, and $Y_2 = (\sigma_x - \sigma_z)/2$ we obtain

$$\langle CHSH\rangle_q = |\langle X_1Y_1\rangle + \langle X_1Y_2\rangle + \langle X_2Y_1\rangle - \langle X_2Y_2\rangle| = 2\sqrt{2}.$$

Most importantly, this violation can be experimentally verified allowing us to test the validity of the theory. The first such tests were performed by Clauser [Cla76] and Aspect, Dalibard, Grangier, and Roger [AGR82, ADR82]. Over the years these tests have been refined considerably, ruling out many loopholes present in the initial experiments such as for example detector inefficiency [RKM$^+$01]. Yet,

no conclusive test has been achieved so far. Unfortunately, such experimental concerns are outside the scope of this thesis, and we merely point to an overview of such issues [Asp99].

6.1.2 Tsirelson's bound

Curiously, even quantum mechanics itself still limits the strength of non-local correlations. Tsirelson's bound [Tsi80] says that for quantum mechanics

$$|\langle X_1 Y_1\rangle + \langle X_1 Y_2\rangle + \langle X_2 Y_1\rangle - \langle X_2 Y_2\rangle| \leq 2\sqrt{2},$$

and thus the above measurements are optimal. We provide a simple proof of this fact in Chapter 7. It is interesting to consider what would happen if quantum mechanics allowed for more powerful non-local correlations. To this end, it is convenient to rewrite the CHSH inequality from Eq. (6.2) in the form

$$\sum_{x,y\in\{0,1\}} \Pr[a_x \oplus b_y = x \cdot y] \leq 3.$$

Here, $x \in \{0,1\}$ and $y \in \{0,1\}$ denote the choice of Alice's and Bob's measurement, $a_x \in \{0,1\}$ and $b_y \in \{0,1\}$ the respective binary outcomes, and $\oplus$ addition modulo 2 (see Section 6.2.3 for details). In this form, quantum mechanics allows a violation up to the maximal value of $2+\sqrt{2}$. Since special relativity would even allow a violation of Tsirelson's bound, Popescu and Rohrlich [PR94, PR96, PR97] raised the question why nature is not more 'non-local'? That is, why does quantum mechanics not allow for a stronger violation of the CHSH inequality up to the maximal value of 4? To gain more insight into this question, they constructed a toy-theory based on non-local boxes. Each such box takes inputs $x, y \in \{0,1\}$ from Alice and Bob respectively and always outputs measurement outcomes a_x,b_y such that $x \cdot y = a_x \oplus b_y$. Alice and Bob still cannot use this box to transmit any information. However, since for all x and y, $\Pr[a_x \oplus b_y = x \cdot y] = 1$, the above sum equals 4 and thus non-local boxes lead to a maximum violation of the CHSH inequality.

Van Dam [vD05, vD00] has shown that having access to such non-local boxes allows Alice and Bob to perform any kind of distributed computation by transmitting only a *single* bit of information. This is even true for slightly less perfect boxes achieving weaker correlations [BBL$^+$06]. In [BCU$^+$06], we showed that given any non-local boxes, Alice and Bob could perform bit commitment and oblivious transfer, which is otherwise known to be impossible. Thus, such cryptographic principles are in principle compatible with the theory of non-signaling: non-signaling itself does not prevent us from implementing them.

Looking back to the uncertainty relations in Chapter 4, which rest at the heart of the EPR paradox, we might suspect that the violation of the CHSH inequality likewise depends on the commutation relations between the local measurements

of Alice and Bob. Indeed, it has been shown by Landau [Lan87], and Khalfin and Tsirelson [KT87], there exists a state $|\Psi\rangle$ such that

$$|\langle X_1 Y_1\rangle + \langle X_1 Y_2\rangle + \langle X_2 Y_1\rangle - \langle X_2 Y_2\rangle| = 2\sqrt{1 + 4\| \, [X_1^0, X_2^0][Y_1^0, Y_2^0] \, \|},$$

for any $X_1 = X_1^0 - X_1^1$, $X_2 = X_2^0 - X_2^1$ and $Y_1 = Y_1^0 - Y_1^1$, $Y_2 = Y_2^0 - Y_2^1$, where we use the superscripts '0' and '1' to denote the projectors onto the positive and negative eigenspace respectively. Thus, given any observables X_1, X_2 and Y_1, Y_2, the CHSH inequality is violated if and only if $[X_1^0, X_2^0][Y_1^0, Y_2^0] \neq 0$.

6.2 Setting the stage

6.2.1 Entangled states

The state given in Eq. (6.1) is just one possible example of an entangled state. Recall from Chapter 2 that if $|\Psi\rangle \in \mathcal{H}^A \otimes \mathcal{H}^B$ is a pure state, we say that $|\Psi\rangle$ is *separable* if and only if there exist states $|\Psi^A\rangle \in \mathcal{H}^A$ and $|\Psi^B\rangle \in \mathcal{H}^B$ such that $|\Psi\rangle = |\Psi^A\rangle \otimes |\Psi^B\rangle$. A separable pure state is also called a *product state*. A state that is not separable is called *entangled*. For mixed states the definition is slightly more subtle. Let $\rho \in \mathcal{S}(\mathcal{H}^A \otimes \mathcal{H}^B)$ be a mixed state. Then ρ is called a *product state* if there exist $\rho^A \in \mathcal{S}(\mathcal{H}^A)$ and $\rho^B \in \mathcal{S}(\mathcal{H}^B)$ such that $\rho = \rho^A \otimes \rho^B$. The state ρ is called *separable*, if there exists an ensemble $\mathcal{E} = \{p_j, |\Psi_j\rangle\langle\Psi_j|\}$ such that $|\Psi_j\rangle = |\Psi_j^A\rangle \otimes |\Psi_j^B\rangle$ with $|\Psi_j^A\rangle \in \mathcal{H}^A$ and $|\Psi_j^B\rangle \in \mathcal{H}^B$ for all j, such that

$$\rho = \sum_j p_j |\Psi_j\rangle\langle\Psi_j| = \sum_j p_j |\Psi_j^A\rangle\langle\Psi_j^A| \otimes |\Psi_j^B\rangle\langle\Psi_j^B|.$$

Intuitively, if ρ is separable then ρ corresponds to a mixture of separable pure states according to a joint probability distribution $\{p_j\}$, a purely classical form of correlation. Given a description of a mixed state ρ it is an NP-hard problem to decide whether ρ is separable [Gur03]. However, many criteria and approximation algorithms have been proposed [DPS02, DPS04, DPS05, IT06, ITCE04]. It is an interesting question to determine the maximal violation of a given Bell-inequality for a fixed state ρ [LD07]. Here, we only concern ourselves with *maximal* violations of Bell inequalities, and refer to [Ioa07] for an overview of the separability problem. Generally, the maximal violation is obtained by using the maximally entangled state. However, there are cases for which the maximal violation is achieved by a non maximally entangled state [CGL$^+$02]. Note that we can never observe a Bell inequality violation for a separable state: it is no more than a classical mixture of separable pure states. On the other hand, any two-qubit *pure* state that is entangled violates the CHSH inequality [Gis91]. However, not all entangled *mixed* states violate the CHSH inequality! A counterexample was given

by Werner [Wer81] with the so-called Werner-state

$$\rho_W = p\frac{2}{d^2+d}P_{sym} + (1-p)\frac{2}{d^2-d}P_{asym},$$

where P_{sym} and P_{asym} are projectors onto the symmetric and the anti-symmetric subspace respectively. For $p \geq 1/2$ this state is separable, but it is entangled for $p < 1/2$. Yet, the CHSH inequality is not violated. A lot of work has been done to quantify the amount of entanglement in quantum states, and we refer to [Ter99, Eis01, Chr05] for an overview.

6.2.2 Other Bell inequalities

The CHSH inequality we encountered above is by no means the only Bell inequality. Recall that non-local correlations arise as the result of measurements performed on a quantum system shared between two spatially separated parties. Let x and y be the variables corresponding to Alice and Bob's choice of measurement. Let a and b denote the corresponding outcomes[1]. Let $\Pr[a,b|x,y]$ be the probability of obtaining outcomes a,b given settings x,y. What values are allowed for $\Pr[a,b|x,y]$? Clearly, we want that for all x,y,a,b we have that $\Pr[a,b|x,y] \geq 0$ and $\sum_{a,b}\Pr[a,b|x,y] = 1$. From the no-signaling condition we furthermore obtain that the marginals obey $\Pr[a|x] = \Pr[a|x,y] = \sum_b \Pr[a,b|x,y]$ and likewise for $\Pr[b|y]$, i.e. the probability of Alice's measurement outcome is independent of Bob's choice of measurement setting, and vice versa. For n players, who each perform one of N measurements with k possible outcomes, we have $(Nk)^n$ such probabilities to assign, giving us a $(Nk)^n$ dimensional vector. To find all Bell inequalities, we now look for inequalities that bound the classically accessible region (a convex polytope) for such assignments. It is clear that we can find a huge number of such inequalities. Of course, often the most interesting inequalities are the ones that are satisfied only classically, but where we can find a better quantum strategy. Much work has been done to identify such inequalities, and we refer to [WW01b] for an excellent overview. In the following chapters, we are interested in the following related question: Given an inequality, what is the optimal quantum measurement strategy that maximizes the inequality?

6.2.3 Non-local games

It is often convenient to view Bell experiments as a game between two, or more, distant players, who cooperate against a special party. We call this special party the *verifier*. In a two player game with players Alice and Bob, the verifier picks two questions, say s_1 and s_2, and hands them to Alice and Bob respectively, who now need to decide answers a_1 and a_2. To this end, they may agree on any

[1]For simplicity, we assume that the set of possible outcomes is the same for each setting.

strategy beforehand, but can no longer communicate once the game starts. The verifier then decides according to a fixed set of public rules, whether Alice and Bob win by giving answers a_1, a_2 to questions s_1, s_2. In a quantum game, Alice and Bob may perform measurements on a shared entangled state to determine their answers. We can thus think of the questions as measurement settings and the answers as measurement outcomes.

More formally, we consider games among N players $P_1, \ldots, P_N$. Let $S_1, \ldots, S_N$ and $A_1, \ldots, A_N$ be finite sets corresponding to the possible questions and answers respectively. Let π be a probability distribution on $S_1 \times \ldots \times S_N$, and let V be a predicate on $A_1 \times \ldots \times A_N \times S_1 \times \ldots \times S_N$. Then $G = G(V, \pi)$ is the following N-player cooperative game: A set of questions $(s_1, \ldots, s_N) \in S_1 \times \ldots \times S_N$ is chosen at random according to the probability distribution π. Player P_j receives question s_j, and then responds with answer $a_j \in A_j$. The players win if and only if $V(a_1, \ldots, a_N, s_1, \ldots, s_N) = 1$. We write $V(a_1, \ldots, a_N | s_1, \ldots, s_N) = V(a_1, \ldots, a_N, s_1, \ldots, s_N)$ to emphasize the fact that $a_1, \ldots, a_N$ are the answers given questions $s_1, \ldots, s_N$.

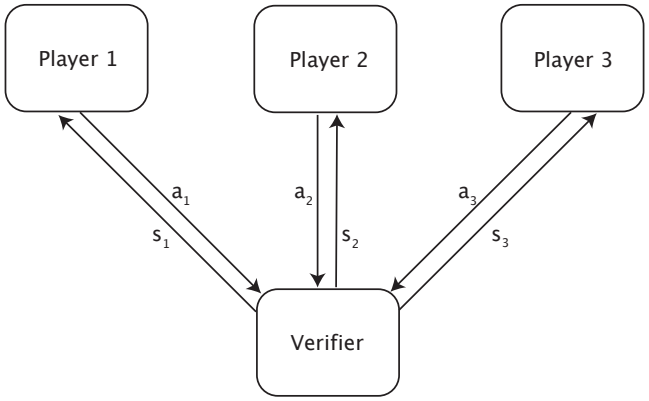


Figure 6.2: Multiplayer non-local games.

The *value* of the game $\omega(G)$ is the probability that the players win the game, maximized over all possible strategies. We use $\omega_c(G)$ and $\omega_q(G)$ to differentiate between the value of the game in the classical and quantum case respectively. Classically, $\omega_c(G)$ can always be attained by a deterministic strategy [CHTW04a]. We can thus write

$$\omega_c(G) = \max_{f_1,\ldots,f_N} \sum_{s_1,\ldots,s_N} \pi(s_1, \ldots, s_N) V(f_1(s_1), \ldots, f_N(s_N) | s_1, \ldots, s_N), \qquad (6.3)$$

where the maximization is taken over all functions $f_j : S_j \to A_j$ that determine

answers $a_j = f_j(s_j)$.

Quantumly, the strategy of the players consists of their choice of measurements and shared entangled state. Let $|\Psi\rangle$ denote the players' choice of state, and let $X_{s_j}^{[j]} = \{X_{s_j}^{a_j,[j]} \mid a_j \in A_j\}$ denote the POVM of player P_j for question $s_j \in S_j$. Here, we always assume that the underlying Hilbert space is finite-dimensional. The value of the quantum game is then

$$\omega_q(G) = \max_{X^{[1]},\ldots,X^{[N]}} \sum_{s_1,\ldots,s_N} \pi(s_1,\ldots,s_N) \sum_{a_1,\ldots,a_N} \langle\Psi|X_{s_1}^{a_1,[1]} \otimes \ldots \otimes X_{s_N}^{a_N,[N]}|\Psi\rangle, \quad (6.4)$$

where the maximization is taken over all POVMS $X_{s_j}^{[j]}$ for all $j \in [N]$ and $s_j \in S_j$. In the following, we say that a set of measurement operators *achieves* p, if

$$p = \sum_{s_1,\ldots,s_N} \pi(s_1,\ldots,s_N) \sum_{a_1,\ldots,a_N} \langle\Psi|X_{s_1}^{a_1,[1]} \otimes \ldots \otimes X_{s_N}^{a_N,[N]}|\Psi\rangle.$$

Of particular relevance in the next chapters is a special class of two-player games known as XOR-games [CHTW04a]: Here, $N = 2$ and we assume that $A_1 = A_2 = \{0, 1\}$. The two players P_1 (Alice) and P_2 (Bob) each have only two possible measurement outcomes. Furthermore, the winning condition only depends on the XOR of answers a_1 and a_2 and thus we write $V(c|s_1, s_2)$ with $c = a_1 \oplus a_2$. It can be shown [CHTW04a] that the optimal POVM in this case consists only of projectors. We can thus write $X_{s_1}^{[1]}$ and $X_{s_2}^{[2]}$ as observables with two eigenvalues: $X_{s_1}^{[1]} = X_{s_1}^{0,[1]} - X_{s_1}^{1,[1]}$ and $X_{s_2}^{[2]} = X_{s_2}^{0,[2]} - X_{s_2}^{1,[2]}$ where $s_1 \in S_1$ and $s_2 \in S_2$. A small calculation using the fact that $X_{s_1}^{0,[1]} + X_{s_1}^{1,[1]} = \mathbb{I}$ and $X_{s_2}^{0,[2]} + X_{s_2}^{1,[2]} = \mathbb{I}$ shows that we can rewrite the optimal value of a quantum XOR-game as

$$\omega_q(G) = \quad (6.5)$$

$$\max_{X^{[1]},X^{[2]}} \frac{1}{2} \sum_{s_1,s_2} \pi(s_1, s_2) \sum_{c\in\{0,1\}} V(c|s_1, s_2)(1 + (-1)^c \langle\Psi|X_{s_1}^{[1]} \otimes X_{s_2}^{[2]}|\Psi\rangle). \quad (6.6)$$

From the above, we can see that XOR-games correspond to correlation inequalities with two-outcome measurements. We will see in Chapter 7 that this reformulation enables us to determine the optimal measurements for such XOR-games in a very simple manner. Indeed, the CHSH inequality can be rephrased as a simple quantum XOR-game. Here, Alice and Bob win if and only if given questions s_1, s_2 they return answers a_1, a_2 such that $s_1 \cdot s_2 = a_1 \oplus a_2$, i.e. we have $V(c|s_1, s_2) = 1$ if and only if $s_1 \cdot s_2 = c$. Recalling Eq. (6.2) we can write

$$\omega(CHSH) = \frac{1}{2}\left(1 + \frac{\langle CHSH\rangle}{4}\right),$$

from which we obtain $\omega_q(CHSH) = 1/2 + 1/(2\sqrt{2})$ vs. $\omega_c(CHSH) = 3/4$.

6.3 Observations

In the following chapters, we are concerned with finding the optimal quantum measurement strategies for Bell inequalities. To this end, we first make a few simple observations that help us understand the structural properties of our problem. In particular, this also enables us to understand the relation between Bell inequalities and the problem of post-measurement information in Chapter 6.4. We then present a theorem by Tsirelson [Tsi80, Tsi87] that plays a crucial role in the subsequent chapters.

6.3.1 Simple structural observations

Suppose we are given a set of measurements for Alice and Bob and a shared state ρ. Can we reduce the dimension of Alice and Bob's measurements operators and the thereby amount of entanglement they need? As we saw in Chapter 2, we can often simplify our problem by identifying its classical and quantum part. Indeed, this is also the case here.

6.3.1. Lemma. *Let $\mathcal{H} = \mathcal{H}^A \otimes \mathcal{H}^B$ and let $\mathcal{A} = \{X_s^a \in \mathbb{B}(\mathcal{H}^A)\}$ and $\mathcal{B} = \{Y_t^b \in \mathbb{B}(\mathcal{H}^B)\}$ be the set of Alice and Bob's measurement operators respectively. Let $\rho \in \mathcal{S}(\mathcal{H})$ be the state shared by Alice and Bob. Suppose that for such operators we have*

$$q = \sum_{s\in S,t\in T} \pi(s,t) \sum_{a\in A,b\in B} V(a,b|s,t)\mathrm{Tr}(X_s^a \otimes Y_t^b \rho).$$

Then there exist measurement operators $\tilde{\mathcal{A}} = \{\tilde{X}_s^a\}$ and $\tilde{\mathcal{B}} = \{\tilde{Y}_t^b\}$ and a state $\tilde{\rho}$ such

$$q \leq \sum_{s\in S,t\in T} \pi(s,t) \sum_{a\in A,b\in B} V(a,b|s,t)\mathrm{Tr}(\tilde{X}_s^a \otimes \tilde{Y}_t^b \tilde{\rho}).$$

and the C^*-algebra generated by $\tilde{\mathcal{A}}$ and $\tilde{\mathcal{B}}$ is simple.

Proof. Let $\mathscr{A} = \langle \mathcal{A} \rangle$ and $\mathscr{B} = \langle \mathcal{B} \rangle$. If $\mathscr{A}$ and $\mathscr{B}$ are simple, we are done. If not, we know from Lemma B.4.1 and Lemma B.4.4 that there exists a decomposition $\mathcal{H}^A \otimes \mathcal{H}^B = \bigoplus_{jk} \mathcal{H}_j^A \otimes \mathcal{H}_k^B$. Consider $\mathrm{Tr}((M^A \otimes M^B)\rho)$, where $M^A \otimes M^B \in \mathscr{A} \otimes \mathscr{B}$. It follows from the above that $M^A \otimes M^B = \bigoplus_{jk} (\Pi_j^A \otimes \Pi_k^B) M^A \otimes M^B (\Pi_j^A \otimes \Pi_k^B)$, where Π_j^A and Π_k^B are projectors onto $\mathcal{H}_j^A$ and $\mathcal{H}_k^B$ respectively. Let $\hat{\rho} = \bigoplus_{jk} (\Pi_j^A \otimes \Pi_k^B)\rho(\Pi_j^A \otimes \Pi_k^B)$. Clearly,

$$\begin{aligned} \mathrm{Tr}((M^A \otimes M^B)\hat{\rho}) &= \sum_{jk} \mathrm{Tr}\left((\Pi_j^A \otimes \Pi_k^B) M^A \otimes M^B (\Pi_j^A \otimes \Pi_k^B)\rho\right) \\ &= \mathrm{Tr}((M^A \otimes M^B)\rho). \end{aligned}$$

The statement now follows immediately by convexity: Alice and Bob can now measure ρ using $\{\Pi_j^A \otimes \Pi_k^B\}$ and record the classical outcomes j, k. The new

measurements are then $\tilde{A}^a_{s,j} = \Pi^A_j X^a_s \Pi^A_j$ and $\tilde{B}^b_{t,k} = \Pi^B_k Y^b_t \Pi^B_k$ on state $\tilde{\rho}_{jk} = (\Pi^A_j \otimes \Pi^B_k)\rho(\Pi^A_j \otimes \Pi^B_k)/\mathrm{Tr}((\Pi^A_j \otimes \Pi^B_k)\rho)$. By construction, $\tilde{\mathscr{A}}_j = \{\tilde{A}^a_{s,j}\}$ and $\tilde{\mathscr{B}}_k = \{\tilde{B}^b_{t,k}\}$ are simple.

Let q_{jk} denote the probability that we obtain outcomes j, k, and let

$$r_{jk} = \sum_{s\in S, t\in T} \pi(s,t) \sum_{a\in A, b\in B} V(a,b|s,t)\mathrm{Tr}(\tilde{A}^a_{s,j} \otimes \tilde{B}^b_{t,j}\tilde{\rho}_{jk}).$$

Then $q = \sum_{jk} q_{jk} r_{jk} \leq \max_{jk} r_{jk}$. Let u, v be such that $r_{u,v} = \max_{jk} r_{jk}$. Hence, we can skip the initial measurement and instead use measurements $\tilde{X}^a_s = \tilde{A}^a_{s,u}$, $\tilde{Y}^b_t = \tilde{B}^b_{t,v}$ and state $\tilde{\rho} = \tilde{\rho}_{u,v}$. □

It also follows immediately from the above proof that

6.3.2. Corollary. $\dim(\tilde{\rho}) \leq \dim(\mathcal{H}^A_u)\dim(\mathcal{H}^B_v)$

We can thus assume without loss of generality, that the algebra generated by Alice and Bob's optimal measurements is always simple. We also immediately see why we can simulate the quantum measurement classically if Alice or Bob's measurements commute locally. Indeed, the above proof tells us how to construct the appropriate classical strategy:

6.3.3. Corollary. *Let $\mathcal{H} = \mathcal{H}^A \otimes \mathcal{H}^B$ and let $\mathcal{A} = \{X^a_s \in \mathbb{B}(\mathcal{H}^A)\}$ and $\mathcal{B} = \{Y^b_t \in \mathbb{B}(\mathcal{H}^B)\}$ be the set of Alice and Bob's measurement operators respectively. Let $\rho \in \mathcal{S}(\mathcal{H})$ be the state shared by Alice and Bob. Let p be the value of the non-local game achieved using these measurements. Suppose that for all s, s',and a, a' we have that $[X^a_s, X^{a'}_{s'}] = 0$ (or for all t, t', b, b' $[Y^b_t, Y^{b'}_{t'}] = 0$). Then there exists a classical strategy for Alice and Bob that achieves p.*

Proof. Our conditions imply that either $\mathscr{A}$ or $\mathscr{B}$ is abelian. Suppose wlog that $\mathscr{A}$ is abelian. Hence, by the above proof we have $\max_j \dim(\mathcal{H}^A_j) = 1$. Again, Alice and Bob perform the measurements determined by Π^A_j and Π^B_k and record their outcomes j, k. Since $\dim(\mathcal{H}^A_j) = 1$, Alice's post-measurement state is in fact classical, and we have no further entanglement between Alice and Bob. □

To violate a Bell inequality, Alice and Bob must thus use measurements which do not commute locally. However, since Alice and Bob are spatially separated, we can write Alice and Bob's measurement operators as $X = \hat{X} \otimes \mathbb{I}$ and $Y = \mathbb{I} \otimes \hat{Y}$ respectively as for any ρ we can write $\mathrm{Tr}(\rho(X \otimes Y)) = \mathrm{Tr}(\rho(\hat{X} \otimes \mathbb{I})(\mathbb{I} \otimes \hat{Y}))$. Thus $[X, Y] = 0$. Thus from a bipartite structure we obtain certain commutation relations. How about the converse? As it turns out, in any finite-dimensional C^*-algebra[2], these two notions are equivalent: From commutation we immediately obtain a bipartite structure! We encounter this well-known, rather beautiful observation in Appendix B.

[2]or indeed any Type-I von Neumann algebra

6.3.2 Vectorizing measurements

In Chapter 7, we show how to obtain the optimal measurements for any bipartite correlation inequality. At first sight, this may appear to be a daunting problem: We must simultaneously maximize Eq. (6.5) over the state ρ as well as measurement operators of the form $X \otimes Y$, a problem which is clearly not convex. Yet, the following brilliant observation by Tsirelson [Tsi80, Tsi87] greatly simplifies our problem.

6.3.4. Theorem (Tsirelson). *Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ be observables with eigenvalues in the interval $[-1, 1]$. Then for any state $|\Psi\rangle \in \mathcal{H}^A \otimes \mathcal{H}^B$ and for all $s \in [n]$, $t \in [m]$, there exist real unit vectors $x_1, \ldots, x_n$,$y_1, \ldots, y_m \in \mathbb{R}^{n+m}$ such that*

$$\langle \Psi | X_s \otimes Y_t | \Psi \rangle = x_s \cdot y_t,$$

where $x_s \cdot y_t$ is the standard inner product. Conversely, let $x_s, y_t \in \mathbb{R}^N$ be real unit vectors. Let $|\Psi\rangle \in \mathcal{H}^A \otimes \mathcal{H}^B$ be any maximally entangled state where $\dim(\mathcal{H}^A) = \dim(\mathcal{H}^B) = 2^{\lfloor N/2 \rfloor}$. Then for all $s \in [n]$, $t \in [m]$ there exist observables X_s on $\mathcal{H}^A$ and Y_t on $\mathcal{H}^B$ with eigenvalues in $\{-1, 1\}$ such that

$$x_s \cdot y_t = \langle \Psi | X_s \otimes Y_t | \Psi \rangle.$$

In fact, by limiting ourselves onto the space spanned by the vectors $x_1, \ldots, x_n$ or $y_1, \ldots, y_m$, we could further decrease the dimension of the vectors to $N = \min\{n, m\}$ [Tsi87]. The result was proven by Tsirelson in a more general form for any finite-dimensional C^*-algebra. Here, we do not consider this more abstract argument, but instead simply sketch how to obtain the vectors and state how to find the corresponding measurement operators in turn [Tsi93]. To find vectors x_s and y_t, we merely need to consider the vectors

$$x_s = X_s \otimes \mathbb{I} |\Psi\rangle \text{ and } y_t = \mathbb{I} \otimes Y_t |\Psi\rangle,$$

where may take the vectors to be real [Tsi80]. Recall that we are only interested in the inner products. But clearly we can then bound the dimension of our vectors as the number of our vectors is strictly limited and thus cannot span a space of dimension larger than N.

To construct observables corresponding to a given set of vectors, consider the generators of a Clifford algebra $\Gamma_1, \ldots, \Gamma_N$ with N even[3] that we already encountered in Section 4.3, i.e., we have that for all $j \neq k \in [N]$, $\{\Gamma_j, \Gamma_k\} = 0$ and $\Gamma_j^2 = \mathbb{I}$. Note that we also have $\mathrm{Tr}(\Gamma_j \Gamma_k) = \delta_{jk}$ as the two matrices anti-commute. Consider two vectors $x_s, y_t \in \mathbb{R}^N$ with $x_s = (x_s^1, \ldots, x_s^N)$ and $y_t = (y_t^1, \ldots, y_t^N)$. Define $X_s = \sum_{j \in [N]} x_s^j \Gamma_j^T$ and $Y_t = \sum_{j \in [N]} y_t^j \Gamma_j$ and let $|\Psi\rangle = (1/\sqrt{d}) \sum_k |k\rangle |k\rangle$ with $d = 2^{\lfloor N/2 \rfloor}$ be the maximally entangled state. We then have

$$\langle \Psi | X_s \otimes Y_t | \Psi \rangle = \frac{1}{d} \sum_{jk} x_s^j y_t^k \mathrm{Tr}(\Gamma_j \Gamma_k) = \frac{1}{d} \sum_j x_s^j y_t^k \mathrm{Tr}(\mathbb{I}) = x_s \cdot y_t.$$

[3] If N is odd, we obtain one additional element from Γ_0.

Note that in principle we could have chosen any set of orthogonal operators $\Gamma_1, \ldots, \Gamma_N$ to obtain the stated equality. However, we obtain from their anti-commutation that

$$X_s^2 = \sum_{jk} x_s^j x_s^k \Gamma_j \Gamma_k = \frac{1}{2} \sum_{jk} x_s^j x_s^k \{\Gamma_j, \Gamma_k\} = \sum_j (x_s^j)^2 \mathbb{I} = \mathbb{I},$$

since $||x_s|| = 1$. Hence, X_s has eigenvalues in $\{-1, 1\}$ as desired. Curiously, $\Gamma_1, \ldots, \Gamma_N$ were also the right choice of operators to obtain good uncertainty relations in Chapter 4.3.

6.4 The use of post-measurement information

Looking back to Chapter 3, we see that we have already encountered the same structure in the context of post-measurement information. Recall that there our goal was to determine y given some $\rho_{yb} \in \{\rho_{yb} \mid y \in \mathcal{Y} \text{ and } b \in \mathcal{B}\}$ after receiving additional post-measurement information b. In particular, as we explain in more detail in Chapter 8 we see that the question of how much post-measurement information is required is the same as the following: given a set of observables, how large does our quantum state have to be in order to implement the resulting non-local game? However, we can further exploit the relationship between these two problems to prove a gap between the optimal success probability in the setting of state discrimination (STAR) and the setting of state discrimination *with* post-measurement information (PI-STAR). In particular, we show that for some problems, if we can succeed perfectly in the setting of PI-STAR without keeping any qubits at all, our success at STAR can in fact be bounded by a Bell-type inequality! Of course, PI-STAR itself is not a non-local problem. However, as we saw in Appendix B, the commutation relations which are necessary for Bob to succeed at PI-STAR perfectly in Lemma 3.5.1, do induce a bipartite structure. We now exploit the structural similarity of the two problems.

We first consider the very simple case of two bases and a Boolean function. Here, it turns out that we can bound the value of the STAR problems using the CHSH inequality. We do this by showing a bound on the average of two equivalent STAR problems, illustrated in Figures 6.3 and 6.4. The XOR function considered in Chapter 3 is an example of such a problem. Below we construct a generalization of the CHSH inequality which allows us to make more general statements. We state our result in the notation introduced in Chapter 3. For simplicity, we use indices $+$ and $\times$ to denote two *arbitrary* bases and use the notation $\text{STAR}(\rho_0, \ldots, \rho_{n-1})$ to refer to a state discrimination problem between n different states.

6.4.1. Lemma. *Let $P_X(x) = 1/2^n$ for all $x \in \{0,1\}^n$ and let $f : \{0,1\}^n \to \{0,1\}$ be any Boolean function. Let $\mathcal{B} = \{+, \times\}$ denote a set of two bases, and suppose*

there exists a unitary U such that $\rho_{0+} = U\rho_{0+}U^\dagger$, $\rho_{1+} = U\rho_{1+}U^\dagger$, $\rho_{1\times} = U\rho_{0\times}U^\dagger$ and $\rho_{0\times} = U\rho_{1\times}U^\dagger$. Suppose Bob succeeds at PI-STAR$_0(f)$ with probability $p = 1$. Then he succeeds at STAR(ρ_0, ρ_1) with probability at most 3/4, where $\rho_0 = (\rho_{0+} + \rho_{0\times})/2$, $\rho_1 = (\rho_{1+} + \rho_{1\times})/2$.

Proof. Let P_{0+}, P_{1+}, $P_{0\times}$ and $P_{1\times}$ be projectors onto the support of ρ_{0+}, ρ_{1+}, $\rho_{0\times}$ and $\rho_{1\times}$ respectively. Suppose that Bob succeeds with probability p at STAR(ρ_0, ρ_1). Then there exists a strategy for Alice and Bob to succeed at the CHSH game with probability p, where Alice's measurements are given by $\{P_{0+}, P_{1+}\}$ and $\{P_{0\times}, P_{1\times}\}$:

Let $\hat{\rho}_0 = (\rho_{0+} + \rho_{1\times})/2$ and $\hat{\rho}_1 = (\rho_{1+} + \rho_{0\times})/2$. Note that since there exists such a U, we have that Bob succeeds at STAR$(\hat{\rho}_0, \hat{\rho}_1)$ with probability p as well. Suppose that Alice and Bob share the maximally entangled state $|\Psi_{AB}\rangle^{\otimes n}$ with $|\Psi_{AB}\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$. With probability 1/2 Alice chooses measurement setting $x = 0$ and then her measurement is given by $\{P_{0+}, P_{1+}\}$. Let a denote her measurement outcome. Bob's system is now in the state ρ_{a+}. Similarly, with probability 1/2 Alice sets $x = 1$ and measures $\{P_{0\times}, P_{1\times}\}$, which leaves Bob's system in the state $\rho_{a\times}$. Let y denote Bob's measurement setting. The CHSH game now requires Bob to obtain a measurement outcome b such that $x \cdot y = a \oplus b$. Thus, for $y = 0$, Bob always tries to obtain $b = a$ which means he wants to solve STAR(ρ_0, ρ_1). For $y = 1$, Bob tries to obtain $b = a$ for $x = 0$ but $b = 1 - a$ for $x = 1$, i.e., he wants to solve STAR$(\hat{\rho}_0, \hat{\rho}_1)$. Since Bob chooses $y \in \{0, 1\}$ uniformly at random, we obtain the stated result.

Now suppose that Bob succeeds at PI-STAR$_0(f)$ with probability $p = 1$. We know from Lemma 3.5.1 that for all $y, y' \in \mathcal{Y}$ and $b, b' \in \mathcal{B}$ we have $[P_{yb}, P_{y'b'}] = 0$ where P_{yb} is a projector onto the support of ρ_{yb}. Now, suppose that on the contrary he succeeds at STAR(ρ_0, ρ_1) with probability greater than 3/4. Then we know from the above argument that there exists a strategy for Alice and Bob to succeed at the CHSH game with probability greater than 3/4 where Alice measures two *commuting* observables, which contradicts Corollary 6.3.3. □

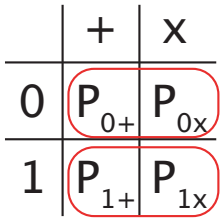


Figure 6.3: Original problem

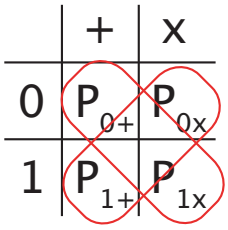


Figure 6.4: Derived problem

It may appear unrealistic to assume that the two STAR problems are essentially equal. Note however, that this is indeed the case in the example of

the XOR function and two mutually unbiased bases (e.g. computational and Hadamard). The unitary here is just $U = \sigma_z \otimes \mathbb{I}^{\otimes n-1}$, as σ_z acts as a bit-flip in the Hadamard basis, but leaves the computational basis invariant. We saw in the proof of Theorem 3.3.5 that such unitaries exist for any choice of two bases from the computational, Hadamard and K-basis. Indeed, for the XOR function on a string of length n with n even we saw that for STAR the optimal probability is $p = 3/4$, whereas for PI-STAR we obtained $p = 1$, as expected.

To generalize this approach, we need to consider more complicated inequalities. In general, there are many possibilities for such inequalities, and one should choose an inequality that reflects the equivalences of possible STAR problems: For example, for the XOR function the CHSH inequality is a good choice as we could identify $U = \sigma_z \otimes \mathbb{I}^{\otimes n-1}$ to give us an equivalence between the two problems. Of course, we would like to ensure that for one of Bob's measurement settings he needs to solve the original STAR problem where Bob's goal is to determine Alice's measurement outcome. At the same time, we would like to minimize the number of possibly inequivalent additional STAR problems created in a similar proof, i.e. we would like to find an inequality where Bob has only a small number of measurement settings. As an example, we consider the following easy way to extend the CHSH inequality. Here, we assume that Alice has equally many measurement outcomes as she has measurement settings. In the language of PI-STAR that means we have wlog $A = \mathcal{Y} = S = \mathcal{B}$. We fix the number of Bob's measurement settings to 2, but allow an arbitrarily large number of settings $|S| = |\mathcal{B}|$ for Alice. Wlog we use $T = \{0, 1\}$ and $S = \{0, \ldots, |\mathcal{B}| - 1\}$. We now define the predicate V with the help of the function τ_t^s for $s \in S$ and $t \in T$. Let $\tau_0^s(y) = y$ for all $s \in S$ and let $\tau_1^0(y) = y$ and $\tau_1^s(y) = \sigma^s(y)$ for all $s \in \{1, \ldots, |\mathcal{B}| - 1\}$, where $\sigma = (1, \ldots, |\mathcal{B}| - 1)$ is the cyclic permutation. We now define the inequality as a non-local game with predicate $V(a, b|s, t) = 1$ if and only if $b = \tau_t^s(a)$. Intuitively, this means that if Bob chooses setting $t = 0$ he is required to solve the original STAR problem, where he tries to guess Alice's measurement outcome. For the setting $t = 1$ he has to solve the problem where the values of y are shifted depending on the basis. Note that the CHSH inequality is a special case of this inequality. Recall from Section 6.2.3 that the optimal value of a classical game can always be attained by a deterministic strategy. Let $f_A : S \to \mathcal{Y}$ and $f_B : T \to \mathcal{Y}$ denote the functions implementing this strategy for Alice and Bob respectively. Looking at Eq. (6.3) we see that we can write

$$\omega_c(G) = \max \frac{1}{2|\mathcal{B}|} \sum_{t,s} [\tau_t^s(f_A(s)) = f_B(t)],$$

where $[x = y] = 1$ if and only if $x = y$. It is easy to see that for a uniform choice of Alice and Bob's measurements, the best thing Bob can do is answer $f_B(t) = x$ for all t where we choose any fixed $x \in \mathcal{Y}$ and let $x = f_A(s)$ for all $s \in S$, i.e. Alice and Bob agree on a particular outcome which will always be their answer regardless

of their setting. For $t = 0$ this means that Bob is always correct, and for $t = 1$ he will be correct if Alice obtained $a = 0$. We then have $\omega_c(G) = (|\mathcal{B}| + 1)/(2|\mathcal{B}|)$. For the CHSH case, this gives us $\omega_c(G) = 3/4$ as expected. It is now possible to make a similar statement then in Lemma 6.4.1 for a bigger PI-STAR problem.

The connection to Bell inequalities helped us understand the case where there exists a clear gap between the two problems. Here, post-measurement information was extremely helpful to us. However, as we saw in Chapter 3 there do exist cases where post-measurement information is entirely useless: we can do equally well without it, if we cannot store any quantum information. Interestingly, in the example of the XOR function on an odd number of input bits this happens exactly when the corresponding states correspond to a measurement that maximally violates CHSH. We have thus reached an extremal point of our problem. Is it possible to find conditions on a set of states which determine when post-measurement information is indeed useful?

6.5 Conclusion

As we saw, entanglement is an inherent aspect of quantum theory. We can experimentally violate Bell's inequality, because we can indeed measure non-commuting observables. The existence of such violations is, next to uncertainty relations and locking, another consequence of the existence of non-commuting measurements within quantum theory. This illustrates their close link to uncertainty relations, locking and even post-measurement information we encountered in the preceding chapters. In essence, in all these tasks we are faced with exactly the same problem: what are the consequences of non-commuting measurements? And how can we find maximally "incompatible" measurements?

In the following chapters, we examine entanglement from a variety of viewpoints. In Chapter 7, we first consider Bell inequalities, and show how to find upper bounds on their violation in a quantum setting. Our approach allows us to find the optimal measurements for any bipartite correlation inequality with two-outcome measurements in a very easy manner. We then consider more general multipartite inequalities. Sadly, our method does not easily apply for more general inequalities. In fact, it is not even clear how large our optimization problem would have to be. We therefore consider a related problem in Chapter 8: Given a probability distribution over measurement outcomes, how large a state do we need to implement such a strategy? We prove a very weak lower bound on the dimension on the resulting state for a very restricted class of games. Finally, we consider the effects that entanglement has on classical protocols in Chapter 9. To this end we examine interactive proof systems where the two provers are allowed to share entanglement. Surprisingly, it turns out that two such provers can be simulated by just a single quantum prover.

Chapter 7

Finding optimal quantum strategies

In the previous chapter, we encountered the CHSH inequality and its generalizations in the guise of quantum games. Tsirelson has proven an upper bound on the CHSH inequality that can be achieved using a quantum strategy. But how can we prove upper bounds for more general inequalities? Or actually, how can we find the optimal measurement strategy? In this chapter, we answer these questions for a restricted class of inequalities by presenting a method that yields the optimal strategy for any two-player correlation inequality with n measurement settings and two measurement outcomes, i.e. an XOR-game.

7.1 Introduction

Optimal strategies for generalized inequalities not only have applications in computer science with regard to interactive proof systems, but may also be important to ensure security in cryptographic protocols. From a physical perspective finding such bounds may also be helpful. As Braunstein and Caves [BC90b] have shown, it is interesting to consider inequalities based on many measurement settings, in particular, the chained CHSH inequality in Eq. 7.1 below: Here, the gap between the classical and the quantum bound is larger than for the original CHSH inequality with only two measurement settings. This can be helpful in real experiments that inevitably include noise, as this inequality leads to a larger gap achieved by the optimal classical and the quantum strategy, and may thus lead to a better test. However, determining bounds on the correlations that *quantum* theory allows remains a difficult problem [BM05]. All Tsirelson-type bounds are known for correlation inequalities with two measurement settings and two outcomes for both Alice and Bob [Tsi93]. Landau [Lan88] has taken a step towards finding Tsirelson-type bounds by considering when two-party correlations of two measurement settings for both Alice and Bob can be realized using quantum measurements. Filipp and Svozil [FS04] have considered the case of three measurement settings analytically and conducted numerical studies for a

larger number of settings. Werner and Wolf [WW01a] also considered obtaining Tsirelson-type bounds for two-outcome measurements for multiple parties and studied the case of three and four settings explicitly. However, their method is hard to apply to general inequalities. Finally, Buhrman and Massar have shown a bound for a generalized CHSH inequality using three measurement settings with three outcomes each [BM05]. It is not known whether this bound can be attained.

Our approach is based on semidefinite programming in combination with Tsirelson's seminal results [Tsi80, Tsi87, Tsi93] as outlined in Section 6.3.2. See Appendix A for a brief introduction to semidefinite programming. It is very easy to apply and gives tight bounds as we can find the optimal measurements explicitly. Let X and Y be Alice's and Bob's observables, and let $|\Psi\rangle$ be a state shared by Alice and Bob. The key benefit we derive from Tsirelson's construction is that it saves us from the need to maximize over all states $|\Psi\rangle$ and observables. Instead, we can replace any terms of the form $\langle\Psi|X\otimes Y|\Psi\rangle$ with the inner product of two real unit vectors $x\cdot y$, and then maximize over all such vectors instead. Our method is thereby similar to methods used in computer science for the two-way partitioning problem [BV04] and the approximation algorithm for MAXCUT by Goemans and Williamson [GW95]. Semidefinite programming allows for an efficient way to approximate Tsirelson's bounds for any CHSH-type inequalities numerically. However, it can also be used to prove Tsirelson type bounds analytically. As an illustration, we first give an alternative proof of Tsirelson's original bound using semidefinite programming. We then prove a new Tsirelson-type bound for the following generalized CHSH inequality [Per93, BC90b]. Classically, it can be shown that

$$|\sum_{i=1}^{n}\langle X_iY_i\rangle+\sum_{i=1}^{n-1}\langle X_{i+1}Y_i\rangle-\langle X_1Y_n\rangle|\leq 2n-2. \tag{7.1}$$

Here, we show that for quantum mechanics

$$|\sum_{i=1}^{n}\langle X_iY_i\rangle+\sum_{i=1}^{n-1}\langle X_{i+1}Y_i\rangle-\langle X_1Y_n\rangle|\leq 2n\cos\left(\frac{\pi}{2n}\right),$$

where $\{X_1,\ldots,X_n\}$ and $\{Y_1,\ldots,Y_n\}$ are observables with eigenvalues ±1 employed by Alice and Bob respectively, corresponding to their n possible measurement settings. It is well known that this bound can be achieved [Per93, BC90b] for a specific set of measurement settings if Alice and Bob share a singlet state. Here, we show that this bound is indeed optimal for *any* state $|\Psi\rangle$ and choice of measurement settings. This method generalizes to other CHSH inequalities, for example, the inequality considered by Gisin [Gis99].

7.2 A simple example: Tsirelson's bound

To illustrate our approach we first give a detailed proof of Tsirelson's bound using semidefinite programming. This proof is more complicated than Tsirelson's original proof. However, it serves as a good introduction to the following section. Let X_1, X_2 and Y_1, Y_2 denote the observables with eigenvalues ± 1 used by Alice and Bob respectively. Our goal is now to show an upper bound for

$$|\langle X_1 Y_1 \rangle + \langle X_1 Y_2 \rangle + \langle X_2 Y_1 \rangle - \langle X_2 Y_2 \rangle|.$$

From Theorem 6.3.4 we know that there exist real unit vectors $x_s, y_t \in \mathbb{R}^4$ such that for all $s, t \in \{0, 1\}$ $\langle X_s Y_t \rangle = x_s \cdot y_t$. In order to find Tsirelson's bound, we thus want to solve the following problem: maximize $x_1 \cdot y_1 + x_1 \cdot y_2 + x_2 \cdot y_1 - x_2 \cdot y_2$, subject to $\| x_1 \| = \| x_2 \| = \| y_1 \| = \| y_2 \| = 1$. Note that we can drop the absolute value since any set of vectors maximizing the above equation, simultaneously leads to a set of vectors minimizing it by taking $-y_1, -y_2$ instead. We now phrase this as a semidefinite program. Let $G = [g_{ij}]$ be the Gram matrix of the vectors $\{x_1, x_2, y_1, y_2\} \subseteq \mathbb{R}^4$ with respect to the inner product:

$$G = \begin{pmatrix} x_1 \cdot x_1 & x_1 \cdot x_2 & x_1 \cdot y_1 & x_1 \cdot y_2 \\ x_2 \cdot x_1 & x_2 \cdot x_2 & x_2 \cdot y_1 & x_2 \cdot y_2 \\ y_1 \cdot x_1 & y_1 \cdot x_2 & y_1 \cdot y_2 & y_1 \cdot y_2 \\ y_2 \cdot x_1 & y_2 \cdot x_2 & y_2 \cdot y_1 & y_2 \cdot y_2 \end{pmatrix}.$$

G can thus be written as $G = B^T B$ where the columns of B are the vectors $\{x_1, x_2, y_1, y_2\}$. By [HJ85, Theorem 7.2.10] we can write $G = B^T B$ if and only if G is positive semidefinite. We thus impose the constraint that $G \geq 0$. To make sure that we obtain unit vectors, we add the constraint that all diagonal entries of G must be equal to 1. Define

$$W = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \\ 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \end{pmatrix}.$$

Note that the choice of order of the vectors in B is not unique, however, a different order only leads to a different W and does not change our argument. We can now rephrase our optimization problem as the following SDP:

$$\begin{array}{ll} \text{maximize} & \frac{1}{2}\text{Tr}(GW) \\ \text{subject to} & G \geq 0 \text{ and } \forall i, g_{ii} = 1 \end{array}$$

Analogous to Appendix A, we can then write for the Lagrangian

$$L(G, \lambda) = \frac{1}{2}\text{Tr}(GW) - \text{Tr}(\text{diag}(\lambda)(G - I)),$$

where $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$. The dual function is then

$$\begin{aligned} g(\lambda) &= \sup_G \mathrm{Tr}\left(G\left(\frac{1}{2}W - \mathrm{diag}(\lambda)\right)\right) + \mathrm{Tr}(\mathrm{diag}(\lambda)) \\ &= \begin{cases} \mathrm{Tr}(\mathrm{diag}(\lambda)) & \text{if } \frac{1}{2}W - \mathrm{diag}(\lambda) \leq 0 \\ \infty & \text{otherwise} \end{cases} \end{aligned}$$

We then obtain the following dual formulation of the SDP

$$\begin{array}{ll} \text{minimize} & \mathrm{Tr}(\mathrm{diag}(\lambda)) \\ \text{subject to} & -\frac{1}{2}W + \mathrm{diag}(\lambda) \geq 0 \end{array}$$

Let p' and d' denote optimal values for the primal and Lagrange dual problem respectively. From weak duality it follows that $d' \geq p'$. For our example, it is not difficult to see that this is indeed true as we show in Appendix A.

In order to prove Tsirelson's bound, we now exhibit an optimal solution for both the primal and dual problem and then show that the value of the primal problem equals the value of the dual problem. The optimal solution is well known [Tsi80, Tsi87, Per93]. Alternatively, we could easily guess the optimal solution based on numerical optimization by a small program for Matlab[1] and the package SeDuMi [SA] for semidefinite programming. Consider the following solution for the primal problem

$$G' = \begin{pmatrix} 1 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 1 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 1 & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 1 \end{pmatrix},$$

which gives rise to the primal value $p' = \frac{1}{2}\mathrm{Tr}(G'W) = 2\sqrt{2}$. Note that $G' \geq 0$ since all its eigenvalues are non-negative [HJ85, Theorem 7.2.1], and all its diagonal entries are 1. Thus all constraints are satisfied. The lower left quadrant of G' is in fact the same as the well known correlation matrix for 2 observables [Tsi93, Equation 3.16]. Next, consider the following solution for the dual problem

$$\lambda' = \frac{1}{\sqrt{2}}(1, 1, 1, 1).$$

The dual value is then $d' = \mathrm{Tr}(\mathrm{diag}(\lambda')) = 2\sqrt{2}$. Because $-W + \mathrm{diag}(\lambda') \geq 0$, λ' satisfies the constraint. Since $p' = d'$, G' and λ' are in fact optimal solutions for the primal and dual respectively. We can thus conclude that

$$|\langle X_1 Y_1\rangle + \langle X_1 Y_2\rangle + \langle X_2 Y_1\rangle - \langle X_2 Y_2\rangle| \leq 2\sqrt{2},$$

which is Tsirelson's bound [Tsi80]. By Theorem 6.3.4, this bound is achievable.

[1] See http://www.cwi.nl/~wehner/tsirel/ for the Matlab example code.

7.3 The generalized CHSH inequality

We now show how to obtain bounds for inequalities based on more than 2 observables for both Alice and Bob. In particular, we prove a bound for the chained CHSH inequality for the quantum case. It is well known [Per93] that it is possible to choose observables $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$, and the maximally entangled state, such that

$$|\sum_{i=1}^{n}\langle X_i Y_i\rangle + \sum_{i=1}^{n-1}\langle X_{i+1}Y_i\rangle - \langle X_1 Y_n\rangle| = 2n\cos\left(\frac{\pi}{2n}\right).$$

We now show that this is optimal. Our proof is similar to the last section. However, it is more difficult to show feasibility for all n.

7.3.1. THEOREM. *Let $\rho \in \mathcal{A}\otimes\mathcal{B}$ be an arbitrary state, where $\mathcal{A}$ and $\mathcal{B}$ denote the Hilbert spaces of Alice and Bob. Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ be observables with eigenvalues ± 1 on $\mathcal{A}$ and $\mathcal{B}$ respectively. Then*

$$|\sum_{i=1}^{n}\langle X_i Y_i\rangle + \sum_{i=1}^{n-1}\langle X_{i+1}Y_i\rangle - \langle X_1 Y_n\rangle| \leq 2n\cos\left(\frac{\pi}{2n}\right),$$

Proof. By Theorem 6.3.4, our goal is to find the maximum value for $x_1\cdot y_1 + x_2\cdot y_1 + x_2\cdot y_2 + x_3\cdot y_2 + \ldots + x_n\cdot y_n - x_1\cdot y_n$, for real unit vectors $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}^{2n}$. As above we can drop the absolute value. Let $G = [g_{ij}]$ be the Gram matrix of the vectors $\{x_1, \ldots, x_n, y_1, \ldots, y_n\} \subseteq \mathbb{R}^{2n}$. As before, we can thus write $G = B^T B$, where the columns of B are the vectors $\{x_1, \ldots, x_n, y_1, \ldots, y_n\}$, if and only if $G \geq 0$. To ensure we obtain unit vectors, we again demand that all diagonal entries of G equal 1. Define $n \times n$ matrix A and $2n \times 2n$ matrix W by

$$A = \begin{pmatrix} 1 & 1 & 0 & \ldots & 0 \\ 0 & 1 & 1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & 1 & 1 \\ -1 & 0 & \ldots & 0 & 1 \end{pmatrix}, \; W = \begin{pmatrix} 0 & A^\dagger \\ A & 0 \end{pmatrix}.$$

We can now phrase our maximization problem as the following SDP:

$$\begin{array}{ll} \text{maximize} & \frac{1}{2}\mathrm{Tr}(GW) \\ \text{subject to} & G \geq 0 \text{ and } \forall i, g_{ii} = 1 \end{array}$$

Analogous to the previous section, the dual SDP is then:

$$\begin{array}{ll} \text{minimize} & \mathrm{Tr}(\mathrm{diag}(\lambda)) \\ \text{subject to} & -\frac{1}{2}W + \mathrm{diag}(\lambda) \geq 0 \end{array}$$

Let p' and d' denote optimal values for the primal and dual problem respectively. As before, $d' \geq p'$.

Primal We now show that the vectors suggested in [Per93] are optimal. For $k \in [n]$, choose unit vectors $x_k, y_k \in \mathbb{R}^{2n}$ to be of the form

$$\begin{aligned} x_k &= (\cos(\phi_k), \sin(\phi_k), 0, \ldots, 0), \\ y_k &= (\cos(\psi_k), \sin(\psi_k), 0, \ldots, 0), \end{aligned}$$

where $\phi_k = \frac{\pi}{2n}(2k-2)$ and $\psi_k = \frac{\pi}{2n}(2k-1)$. The angle between x_k and y_k is given by $\psi_k - \phi_k = \frac{\pi}{2n}$ and thus $x_k \cdot y_k = \cos\left(\frac{\pi}{2n}\right)$. The angle between x_{k+1} and y_k is $\phi_{k+1} - \psi_k = \frac{\pi}{2n}$ and thus $x_{k+1} \cdot y_k = \cos\left(\frac{\pi}{2n}\right)$. Finally, the angle between $-x_1$ and y_n is $\pi - \psi_n = \frac{\pi}{2n}$ and so $-x_1 \cdot y_n = \cos\left(\frac{\pi}{2n}\right)$. The value of our primal problem is thus given by

$$p' = \sum_{k=1}^{n} x_k \cdot y_k + \sum_{k=1}^{n-1} x_{k+1} \cdot y_k - x_1 \cdot y_n = 2n \cos\left(\frac{\pi}{2n}\right).$$

Let G' be the Gram matrix constructed from all vectors x_k, y_k as described earlier. Note that our constraints are satisfied: $\forall i : g_{ii} = 1$ and $G' \geq 0$, because G' is symmetric and of the form $G' = B^T B$.

Dual Now consider the $2n$-dimensional vector

$$\lambda' = \cos\left(\frac{\pi}{2n}\right)(1, \ldots, 1).$$

In order to show that this is a feasible solution to the dual problem, we have to prove that $-\frac{1}{2}W + \mathrm{diag}(\lambda') \geq 0$ and thus the constraint is satisfied. To this end, we first show that

2. Claim. *The eigenvalues of A are given by $\gamma_s = 1 + e^{i\pi(2s+1)/n}$ with $s = 0, \ldots, n-1$.*

Proof. Note that if the lower left corner of A were 1, A would be a circulant matrix [Gra71], i.e. each row of A is constructed by taking the previous row and shifting it one place to the right. We can use ideas from circulant matrices to guess eigenvalues γ_s with eigenvectors

$$u_s = (\rho_s^{n-1}, \rho_s^{n-2}, \rho_s^{n-3}, \ldots, \rho_s^1, \rho_s^0),$$

where $\rho_s = e^{-i\pi(2s+1)/n}$ and $s = 0, \ldots, n-1$. By definition, $u = (u_1, u_2, \ldots, u_n)$ is an eigenvector of A with eigenvalue γ if and only if $Au = \gamma u$. Here, $Au = \gamma u$ if and only if

$$\begin{aligned} &(i) \quad \forall j \in \{1, \ldots, n-1\} : u_j + u_{j+1} = \gamma u_j, \\ &(ii) \quad -u_1 + u_n = \gamma u_n. \end{aligned}$$

Since for any $j \in \{1, \ldots, n-1\}$

$$\begin{aligned} u_j + u_{j+1} &= \rho_s^{n-j} + \rho_s^{n-j-1} = \\ &= e^{-i(n-j)\pi(2s+1)/n}(1 + e^{i\pi(2s+1)/n}) = \\ &= \rho_s^{n-j}\gamma_s = \gamma_s u_j, \end{aligned}$$

(i) is satisfied. Furthermore (ii) is satisfied, since

$$\begin{aligned} -u_1 + u_n &= -\rho_s^{n-1} + \rho_s^0 = \\ &= -e^{-i\pi(2s+1)}e^{i\pi(2s+1)/n} + 1 = \\ &= 1 + e^{i\pi(2s+1)/n} = \\ &= \gamma_s \rho_s^0 = \gamma_s u_n. \end{aligned}$$

□

3. Claim. *The largest eigenvalue of W is given by $\gamma = 2\cos\left(\frac{\pi}{2n}\right)$.*

Proof. By [HJ85, Theorem 7.3.7], the eigenvalues of W are given by the singular values of A and their negatives. It follows from Claim 2 that the singular values of A are

$$\sigma_s = \sqrt{\gamma_s \gamma_s^*} = \sqrt{2 + 2\cos\left(\frac{\pi(2s+1)}{n}\right)}.$$

Considering the shape of the cosine function, it is easy to see that the largest singular value of A is given by $\sqrt{2 + 2\cos(\pi/n)} = \sqrt{4\cos^2(\pi/(2n))}$, the largest eigenvalue of W is $\sqrt{2 + 2\cos(\pi/n)} = 2\cos(\pi/(2n))$. □

Since $-\frac{1}{2}W$ and $\mathrm{diag}(\lambda')$ are both Hermitian, Weyl's theorem [HJ85, Theorem 4.3.1] implies that

$$\gamma_{min}\left(-\frac{1}{2}W + \mathrm{diag}(\lambda')\right) \geq \gamma_{min}\left(-\frac{1}{2}W\right) + \gamma_{min}\left(\mathrm{diag}(\lambda')\right),$$

where $\gamma_{min}(M)$ is the smallest eigenvalue of a matrix M. It then follows from the fact that $\mathrm{diag}(\lambda')$ is diagonal and Claim 3 that

$$\gamma_{min}\left(-\frac{1}{2}W + \mathrm{diag}(\lambda')\right) \geq -\frac{1}{2}\left(2\cos\left(\frac{\pi}{2n}\right)\right) + \cos\left(\frac{\pi}{2n}\right) = 0.$$

Thus $-\frac{1}{2}W + \mathrm{diag}(\lambda') \geq 0$ and λ' is a feasible solution to the dual problem. The value of the dual problem is then

$$d' = \mathrm{Tr}(\mathrm{diag}(\lambda')) = 2n\cos\left(\frac{\pi}{2n}\right).$$

Because $p' = d'$, G' and λ' are optimal solutions for the primal and dual respectively, which completes our proof. □

Note that for the primal problem we are effectively dealing with 2-dimensional vectors, x_k, y_k. As we saw in Section 6.3.2, it follows from Tsirelson's construction [Tsi93] that in this case we just need a single EPR pair such that we can find observables that achieve this bound. In fact, these vectors just determine the measurement directions as given in [Per93].

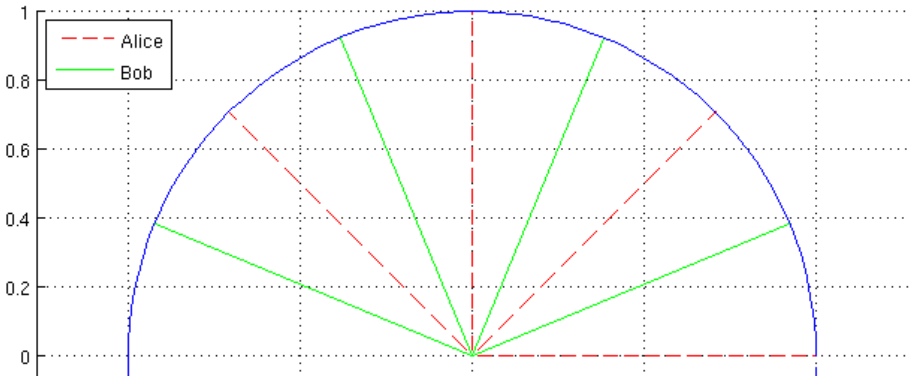


Figure 7.1: Optimal vectors for $n = 4$ obtained numerically using Matlab.

7.4 General approach and its applications

7.4.1 General approach

Our approach can easily be generalized to other correlation inequalities. For another inequality, we merely use a different matrix A in W. For example, for Gisin's CHSH inequality [Gis99], A is the matrix with 1's in the upper left half and on the diagonal, and -1's in the lower right part. Otherwise our approach stays exactly the same, and thus we do not consider this case here. Numerical results provided by our Matlab example code suggest that Gisin's observables are optimal. Given the framework of semidefinite programming, the only difficulty in proving bounds for other inequalities is to determine the eigenvalues of the corresponding A, a simple matrix. All bounds found this way are tight, as we can always implement the resulting strategy using a maximally entangled state as shown in Section 6.3.2.

With respect to finding numerical bounds, we see that the optimal strategy can be found in time exponential in the number of measurement settings: The size of the vectors scales exponentially with the number of settings, however, we

can fortunately find the optimal vectors in time polynomial in the length of the vectors using well-known algorithms for semidefinite programming [BV04].

7.4.2 Applications

In Chapter 9, we will see that the mere existence of such a semidefinite program has implications for the computational complexity of interactive proof systems with entanglement. Cleve, Høyer, Toner and Watrous [CHTW04a] have also remarked during their presentation at CCC'04 that Tsirelson's constructions leads to an approach by semidefinite programming in the context of multiple interactive proof systems with entanglement, but never gave an explicit argument.

The above semidefinite program has also been used to prove results about compositions of quantum games, in particular, parallel repetitions of quantum XOR-games [CSUU07]. One particular type of composition studied by Cleve, Slofstra, Unger and Upadhyay [CSUU07] is the XOR-composition of non-local games. For example, an XOR-composition of a CHSH game is a new game where Alice and Bob each have n inputs $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ with $x_j, y_j \in \{0, 1\}$ and must give answers a and b such that $a \oplus b = \bigoplus x_j \cdot y_j$. In terms of our semidefinite program, this is indeed easy to analyze. The matrix defining the game is now given by

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad W = \begin{pmatrix} 0 & A^{\otimes n} \\ A^{\otimes n} & 0 \end{pmatrix}.$$

Note that the eigenvalues of W are given by $\pm\sqrt{\gamma(A)\gamma(A)^*}$ where $\gamma(A) = \pm(\sqrt{2})^n$ is an eigenvalue of $A^{\otimes n}$. Consider the matrix $G = \mathbb{I} + W/(\sqrt{2})^n$. Clearly, $W/(\sqrt{2})^n$ has eigenvalues ± 1 so we have $G \geq 0$. Thus G is a valid solution to our primal problem, for which we obtain $p = \mathrm{Tr}(GW)/2 = (2\sqrt{2})^k$. Consider $\lambda = (1, \ldots, 1)((\sqrt{2})^n/2)$. Clearly, it is a valid solution to our dual problem as $-W/2 + \mathrm{diag}(\lambda) \geq 0$, again using Weyl's theorem. This gives for our dual problem $d = \mathrm{Tr}(\mathrm{diag}(\lambda)) = (2\sqrt{2})^n = p$ and thus our primal solution is optimal. For more general problems, such a composition may be more complicated as the dual solution is not immediately related to the eigenvalues of W. Nevertheless, it can be readily evaluated using Schur's complement trick [CSUU07]. By rewriting, one can then relate such compositions to the questions of parallel repetition: Given multiple runs of the game, does there exist a better quantum measurement than executing the optimal strategy of each round many times? It is very interesting that this is in fact not true for XOR-games [CSUU07]. However, there exist inequalities and specific quantum states for which collective measurements are better. Such examples can be found in the works of Peres [Per96] and Liang and Doherty [LD06]. Sadly, our approach fails here.

7.5 Conclusion

We have provided a simple method to obtain the optimal measurements for any bipartite correlation inequality, i.e. any two-player XOR game. Our method easily allows us to obtain bounds using numerical analysis, but also suits itself to construct analytical proofs as demonstrated by our examples. However, the above discussion immediately highlights the shortcomings of our approach. How can we find the optimal strategies for more generalized inequalities, where we have more than two players or a non-correlation inequality? Or more than two measurement outcomes? How can we find the optimal strategy for a fixed quantum state that is given to Alice and Bob? To address more than two measurement outcomes, we can rescale the observables such that they have eigenvalues in the interval $[-1, 1]$. Indeed, examining Tsirelson's proof, it is easy to see that we could achieve the same by demanding that the vectors have a length proportional to the number of settings. However, it is clear that the converse of Tsirelson's theorem that allows us to construct measurement from the vectors can no longer hold. Indeed, any matrix $M = \sum_j m_j X_j$ that can be written as a sum of anti-commuting matrices $X_1, \ldots, X_n$ with $X_j^2 = \mathbb{I}$ and $\sum_j m_j^2 = 1$ must have eigenvalues ± 1 itself since $M^2 = \sum_{jk} m_j m_k X_j X_k = \left(\sum_j m_j^2\right) \mathbb{I} = \mathbb{I}$.

Since the completion of this work, exciting progress has been made to answer the above questions. Liang and Doherty [LD07] have shown how to obtain lower and upper bounds on the optimal strategy achievable using a fixed quantum state using semidefinite programming relaxations. Kempe, Kobayashi, Matsumoto, Toner and Vidick [KKM+07] have since shown that there exist three-player games for which the optimal quantum strategy cannot be computed using a semidefinite program that is exponential in the number of measurement settings unless P=NP. Finally, Navascués, Pironio and Acín [NPA07] have shown how to obtain bounds for general two-party inequalities with more measurement outcomes using semidefinite programming, inspired by Landau [Lan88]. Their beautiful approach used successive hierarchies of semidefinite programs to obtain better and better bounds. In their approach, they consider whether a given distribution over outcomes can be obtained using a quantum strategy. Sadly, it does not give a general method to construct actual measurements and thus show that an obtained bound is tight. A similar result obtained using an approach that is essentially dual to [NPA07] has been obtained in [DLTW08], which also proves a convergence result for such a hierarchy.

One of the difficulties we face when trying to find tight bounds for more general inequalities is to determine how large our optimization problem has to be. But even if we are given some distribution over possible outcomes, how can we decide how large our system has to be in order to implement a quantum strategy? In general, this is a tricky problem which we will consider in the next chapter.

Chapter 8

Bounding entanglement in NL-games

In the previous chapter, we provided a simple method to determine the optimal quantum strategy for two-outcome XOR games. However, when trying to find the optimal strategies for more general games, we are faced with a fundamental issue: How large do we have to choose our state and measurements such that we can achieve the optimal quantum value?

8.1 Introduction

Determining an upper bound and the amount of entanglement we need, given the description of the game alone, turns out to be a tricky problem in the general case. Hence, we address an intermediate problem: Given the description of a non-local game and associated probabilities, how large a state do Alice and Bob need to implement such a strategy? Navascués, Pironio and Acín [NPA07] and also [DLTW08] have shown how to obtain upper bounds for the violation of more general quantum games using multiple hierarchies of semidefinite programs. However, their method does not provide us with an explicit strategy, and it remains unclear how many levels of the hierarchy we need to consider in order to obtain a tight bound. Yet, from their method we can obtain a probability distribution over measurement outcomes. Using our approach, we can then determine an extremely weak lower bound on the dimension of the quantum state we would need in order to implement a corresponding quantum strategy.

The idea behind our approach is to transform a non-local game into a random access code. A random access code is an encoding of a string into a quantum state such that we can retrieve at least one entry of our choice from this string with some probability. Intuitively, Alice's measurements will create an encoding. Bob's choice of measurement then determines which bit of this "encoding" he wants to retrieve. We prove a general lower bound for any independent one-to-one non-local game among n players, where a one-to-one non-local game is a game where for each possible measurement setting there exists exactly one correct

measurement outcome. In particular, we show that in any one-to-one non-local game where player P_j obtains the correct outcome $a \in A_j$ for any measurement setting $s \in S_j$ with probability p the dimension d of player P_j's state obeys

$$d \geq 2^{(\log |A_j| - H(p) - (1-p)\log(|A_j|-1))|S_j|}.$$

Even though our bound is very weak, and the class of games very restricted, we are hopeful that our approach may lead to stronger results in the future. Finally, we discuss how we could obtain upper bounds from the description of the non-local game alone without resorting to probability distributions.

8.2 Preliminaries

Before we can prove our lower bound, we first introduce the notion of a random access code. For our purposes, we need to generalize the existing results on random access codes. We use $M(\rho)$ to denote the random variable corresponding to the outcome of a measurement M on a state ρ. We also use A^n to denote an n-element string where each element is chosen from an alphabet A. We will also use the notation $\vec{s}_{-j}$ to denote the string $\vec{s} = (s_1, \ldots, s_n)$ without the element s_j.

8.2.1 Random access codes

A quantum (n, m, p)-random access code (RAC) [ANTV99, Nay99] over a binary alphabet is an encoding of an n-bit string x into an m-qubit state ρ_x such that for any $i \in [n]$ we can retrieve x_i from ρ_x with probability p. Note that we are only interested in retrieving a single bit of the original string x from ρ_x. In general, it is unlikely we will be able to retrieve more than a single bit. For such codes the following lower bound has been shown [Nay99, Theorem 2.3], where it is assumed that the original strings x are chosen uniformly at random:

8.2.1. Theorem (Nayak). *Any (n, m, p)-random access code has $m \geq (1 - H(p))n$.*

In the following, we make use of a generalization of random access codes to larger alphabets. We also need two additional generalizations: First, we also want to obtain a bound on such a RAC encoding if the string x is chosen from a slightly more general, possibly non-uniform, distribution. Let P_{X_t} be a probability distribution over Σ and let $P_X = P_{X_1} \times \ldots \times P_{X_n}$ be a probability distribution over Σ^n. That is, a particular string x is chosen with probability $P_X(x) = \Pi_{t=1}^n P_{X_t}(x_t)$. Note that we assume that the individual entries of x are chosen independently.

Second, we allow for unbalanced random access codes, where each entry of the string x may have a different probability of being decoded correctly. We define

8.2.2. Definition. An $(n, m, (p_1, \ldots, p_n))_{|\Sigma|}$-unbalanced random access code (URAC) over a finite alphabet Σ is an encoding of an n-element string $x \in \Sigma^n$ into an m-qubit state ρ_x such that for any $t \in [n]$, there exists a measurement M_t with outcomes Σ such that for all $x \in \Sigma^n$ we have $\Pr[M_t(\rho_x) = x_t] \geq p_t$.

Fortunately, it is straightforward to extend the analysis of Nayak [Nay99] to this setting. We extend the proof by Nayak as opposed to other known proofs of this lower bound in order to deal with unbalanced random access codes more easily.

8.2.3. Lemma. *Let $P_X = P_{X_1} \times \ldots \times P_{X_n}$ be a probability distribution over Σ^n. Then any $(n, m, (p_1, \ldots, p_n))_{|\Sigma|}$-unbalanced random access code has*

$$m \geq \sum_{t=1}^{n} H(X_t) - H(p_t) - (1 - p_t) \log(|\Sigma| - 1)$$

where X_t is a random variable chosen from Σ according to the probability distribution P_{X_t}.

Proof. The proof follows along the same lines as Lemma 4.1 and Claim 4.6 of [Nay99]. We state the adaption for clarity:

We first consider decoding a single element. Let σ_a with $a \in \Sigma$ be density matrices, and let P be a probability distribution over Σ. Define $\sigma = \sum_{a \in \Sigma} P(a)\sigma_a$. Let M be a measurement with outcomes Σ that given any state σ_a gives the correct outcome a with average probability p. Let X be a random variable over Σ chosen according to probability distribution P, and let Z be a random variable over Σ corresponding to the outcome of the measurement. It now follows from Fano's inequality (see for example [Hay06, Theorem 2.2]) that $\mathcal{I}(X, Z) = H(X) - H(X|Z) \geq H(X) - H(p) - (1 - p)\log(|\Sigma| - 1)$. Using Holevo's bound, we then have $S(\sigma) \geq \sum_{a \in \Sigma} P(a)S(\sigma_a) + H(X) - H(p) - (1 - p)\log(|\Sigma| - 1)$.

We now consider an entire string x encoded as a state ρ_x. Consider k with $n \geq k \geq 0$ and define $\rho_y = \sum_{z \in \Sigma^{n-k}} q_z \rho_{zy}$ with $q_z = \Pi_{j=n-k}^{n} P_{X_j}(z_j)$ where we used indices $z = z_n, \ldots, z_{n-k}$ and P_{X_j} to denote the probability distribution over Σ according to which the j-th entry was encoded. We now claim that $S(\rho_y) \geq \sum_{a \in \Sigma} P_{X_{n-k}}(a)S(\rho_{ay}) + H(X_{n-k}) - H(p_{n-k}) - (1 - p_{n-k})\log(|\Sigma| - 1)$. The proof follows by downward induction over k: Consider $n = k$, clearly $S(\rho_y) \geq 0$ and the claim is valid. Now suppose our claim holds for $k + 1$. Note that we have $\rho_y = \sum_{a \in \Sigma} P_{X_{n-k}}(a)\rho_{ay}$. Note that strings encoded by the density matrices ρ_{ay} only differ by one element $a \in \Sigma$. We can therefore distinguish them with probability p_{n-k}. From the above discussion we have that $S(\rho_y) \geq \sum_{a \in \Sigma} P_{X_{n-k}}(a)S(\rho_{ay}) + H(X_{n-k}) - H(p_{n-k}) - (1 - p_{n-k})\log(|\Sigma| - 1)$.

Using the inductive hypothesis, letting y be the empty string and using the fact that $S(\rho) \leq \log d = m$ then completes the proof. □

8.2.2 Non-local games and state discrimination

For our purpose, we need to think of non-local games as a special form of state discrimination. When each subset of players performs a measurement on their part of the state, they effectively prepare a certain state on the system of the remaining players. Let $\chi_{\vec{a}_{-j}}^{\vec{s}_{-j}}$ denote the state of Player P_j if the remaining players chose measurement settings $\vec{s}_{-j}$ and obtained outcomes $\vec{a}_{-j}$. Note that the probability that player P_j holds $\chi_{\vec{a}_{-j}}^{\vec{s}_{-j}}$ is $\Pr[\vec{a}_{-j}, \vec{s}_{-j}] = \Pr[\vec{a}_{-j}|\vec{s}_{-j}]\Pi_{\ell=1,\ell\neq j}^{N}\pi_\ell(s_\ell)$. Define the state

$$\zeta_{a_j}^{s_j} = \frac{1}{q_{a_j}^{s_j}} \left(\sum_{\vec{s}_{-j}} \sum_{\vec{a}_{-j}} V(\vec{a}|\vec{s}) \Pr[\vec{a}_{-j}|\vec{s}_{-j}] \chi_{\vec{a}_{-j}}^{\vec{s}_{-j}} \right)$$

where $q_{a_j}^{s_j} = \sum_{\vec{s}_{-j}} \sum_{\vec{a}_{-j}} V(\vec{a}|\vec{s}) \Pr[\vec{a}_{-j}|\vec{s}_{-j}]$ to ensure normalization. We call a game *independent*, if the sets of probabilities $\{q_{a_j}^u \mid a_j \in A_j\}$ and $\{q_{a_k}^v \mid a_k \in A_j\}$ are uncorrelated for all measurement settings $u, v \in S_j$ with $u \neq v$. Note that $q_{a_j}^{s_j}$ is the probability that player P_j holds state $\zeta_{a_j}^{s_j}$, and that $\sum_{a_j \in A_j} q_{a_j}^{s_j} = 1$ since the game is one-to-one. If player P_j now chooses measurement setting s_j he is effectively trying to solve a state discrimination problem, given the ensemble $\{q_{a_j}^{s_j}, \zeta_{a_j}^{s_j} | a_j \in A_j\}$.

Note that we already encountered this viewpoint in Chapter 6.4. Consider the simple case of the CHSH game. Here, Alice (Player 1) and Bob (Player 2) had to give answers a_1 and a_2 for settings s_1 and s_2 such that $s_1 \cdot s_2 = a_1 \oplus a_2$. Let $\zeta_{a_1}^{s_1}$ denote Bob's state if Alice chose measurement setting s_1 and obtained outcome a_1. If Bob chooses setting $s_2 = 0$, he has to solve the state discrimination problem described by Figure 6.3: he must answer $a_2 = a_1$, and hence his goal is to learn a_1. That is, he must solve the state discrimination problem given by $\rho_0 = (\zeta_0^0 + \zeta_0^1)/2$ and $\rho_1 = (\zeta_1^0 + \zeta_1^1)/2$. For $s_2 = 1$, he has to solve the problem given by Figure 6.4: For $s_1 = 0$, he must answer $a_2 = a_1$, but for $s_1 = 1$ he must answer $a_2 \neq a_1$. Hence, he must solve the state discrimination problem given by $\tilde{\rho}_0 = (\zeta_0^0 + \zeta_1^1)/2$ and $\tilde{\rho}_1 = (\zeta_1^0 + \zeta_0^1)/2$.

8.3 A lower bound

We now show how to obtain a random access encoding from a one-to-one non-local game. This enables us to find a lower bound on the dimension of the quantum state necessary for any player P_j to implement particular non-local strategies. Recall that we are trying to give a bound given all parameters of the game. In particular, we are given the probabilities $\Pr[\vec{a}_{-j}|\vec{s}_{-j}]$ that the remaining players obtain outcomes $\vec{a}_{-j}$ for their measurement settings $\vec{s}_{-j}$, as well as the value of the game. Note that we do not need to know an actual state and measurement strategy for the players. We just want to give a lower bound for a chosen set of parameters, whether these can be obtained or not.

8.3.1. THEOREM. *Any one-to-one independent non-local game where player P_j obtains the correct outcome $a_j \in A_j$ for measurement setting $s_j \in S_j$ with probability p_{s_j} for all $\vec{s}_{-j} \in S_1 \times \ldots \times S_{j-1} \times S_{j+1} \times \ldots \times S_N$ and $\vec{a}_{-j} \in A_1 \times \ldots \times A_{j-1} \times A_{j+1} \times \ldots \times A_N$ is a $(|S_j|, m, (p_1, \ldots, p_{|S_j|}))_{|A_j|}$-unbalanced random access code.*

Proof. To encode a string, the other players choose measurement settings $\vec{s}_{-j}$ and measure their part of the state as in the non-local game to obtain outcomes $\vec{a}_{-j}$. Note that the string is chosen randomly by the measurement. Since our game was one-to-one we can define a function

$$g(\vec{s}_{-j}, \vec{a}_{-j}) = f_1(\vec{s}_{-j}, \vec{a}_{-j}), \ldots, f_{|S_j|}(\vec{s}_{-j}, \vec{a}_{-j}).$$

Let $x = g(\vec{s}_{-j}, \vec{a}_{-j})$ be the encoded string and note that $\rho_x = \chi^{\vec{s}_{-j}}_{\vec{a}_{-j}}$. We have $P_{X_t}(c) = q_c^{s_j}$, since our game is one-to-one. Since our game is independent, we have that P_X is a product distribution. To retrieve the t-th entry of x, player P_j then has to distinguish $\zeta_{a_j}^{s_j}$ as in the non-local game which he can do with probability p_{s_j} by assumption. □

Now that we can obtain a random access code from a non-local game, we can easily give a lower bound on the dimension of the state from a lower bound of the size of the random access code. It follows immediately from Theorem 8.3.1 and Lemma 8.2.3 that

8.3.2. COROLLARY. *In any one-to-one independent non-local game where player P_j obtains the correct outcome $a \in A_j$ for measurement setting $s \in S_j$ with probability p_s for all measurement settings $\vec{s}_{-j} \in S_1 \times \ldots \times S_{j-1} \times S_{j+1} \times \ldots \times S_N$ and outcomes $\vec{a}_{-j} \in A_1 \times \ldots \times A_{j-1} \times A_{j+1} \times \ldots \times A_N$ of the other players, the dimension d of player P_j's state obeys*

$$d \geq 2^{\sum_{t=1}^{|S_j|} H(X_t) - H(p_t) - (1-p_t)\log(|A_j|-1)},$$

where X_t is a random variable chosen from A_j where $\Pr[X_t = a] = q_a^t$.

For almost all known games, we can obtain a simplified bound as each player will choose a measurement setting uniformly at random. Likewise, in most cases we can assume that the probability that the players obtain certain outcomes is also uniform. Indeed, if we do not know a particular measurement strategy for a given game, we can find a bound if we assume that the distribution over the outcomes given the choice of measurement settings is uniform. In this case, we also assume that the probability of giving the correct answer is the same for each possible choice of measurement settings and is equal to the value of the game. We then obtain

8.3.3. Corollary. *In any one-to-one independent non-local game where player P_j obtains the correct outcome $a \in A_j$ for any measurement setting $s \in S_j$ with probability p where $q_a^t = 1/|A_j|$ for all $t \in S_j$ and measurement settings $\vec{s}_{-j} \in S_1 \times \ldots \times S_{j-1} \times S_{j+1} \times \ldots \times S_N$ and outcomes $\vec{a}_{-j} \in A_1 \times \ldots \times A_{j-1} \times A_{j+1} \times \ldots \times A_N$ of the other players, the dimension d of player P_j's state obeys*

$$d \geq 2^{(\log(|A_j|) - H(p) - (1-p)\log(|A_j|-1))|S_j|}.$$

Note that if we are willing to assume that the optimal value of the game is achieved when the players share a maximally entangled state, we can improve this bound to $d \geq \max_j 2^{(\log(|A_j|) - H(p) - (1-p)\log(|A_j|-1))|S_j|}$.

Let's look at a small example which illustrates the proof. Consider the CHSH inequality. Here, we have only two players, Alice (Player 1) and Bob (Player 2). Bob's goal is to obtain an outcome a_2 such that $s_1 \cdot s_1 = a_1 + a_2 \mod 2$. This means we define the function $g(s_1, a_1) = x$ as $g(0,0) = 0,0$, $g(1,0) = 1,1$, $g(0,1) = 1,0$ and $g(1,1) = 0,1$. For the lower bound we do not need to consider a specific encoding, however, for the well-known CHSH state and measurements we would have an encoding of $\rho_{00} = |0\rangle\langle 0|$, $\rho_{01} = |-\rangle\langle -|$, $\rho_{10} = |+\rangle\langle +|$, and $\rho_{11} = |1\rangle\langle 1|$ and $q^1_{x_1} = q^2_{x_2} = 1/2$ for all $x_1, x_2 \in \{0,1\}$. How many qubits does Bob need to use if he wants to give the correct answers with probability $p = 1/2 + 1/(2\sqrt{2})$? Since everything is uniform we obtain $\log d \geq (1 - H(p))2 \approx 0.8$, i.e., Bob needs to keep at least one qubit.

Our bound contains a tradeoff between the probability p of giving the correct answer, the number of measurement settings, and the number of possible outcomes. Clearly, our bound will only be good, if the number of measurement settings is large. It is also clear that it performs badly as p approaches $1/2$ and $|A_j|$ is large, and thus for most cases our bound will be very unsatisfactory. The following figures illustrate the tradeoff between the different parameters of Corollary 8.3.3.

8.4 Upper bounds

Ideally, we would find an upper bound on the amount of entanglement we need purely from the description of the game alone. Clearly, Tsirelson's construction from Chapter 6.3.2 tells us that for any XOR game the local dimension of Alice's and Bob's system is $d \leq 2^{N/2}$, where N is the number of measurement settings. Similarly to XOR games, we can consider mod k-games. Here, Alice and Bob have to give answers a_1, a_2 given questions s_1, s_1 such that $f(s_1, s_2) = a_1 + a_2 \mod k$ for some function $f : S_1 \times S_2 \to \{0, \ldots, k-1\}$. One may hope that for mod k-games, similarly than for XOR-games, the following holds:

8.4.1. Conjecture. *For any mod k-game, the dimension of Alice's and Bob's systems obeys $d \leq k^{N/2}$, where N is the number of measurement settings for Alice and Bob.*

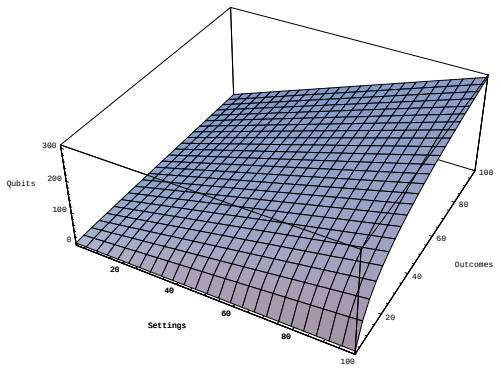


Figure 8.1: Tradeoff for $p = 0.6$.

An alternative approach to bounding the dimension would be to consider how far we can reduce the size of an existing state and observables using Lemma 6.3.1. Suppose that Alice has only two measurement settings $X_0 = X_0^0 - X_0^1$ and $X_1 = X_1^0 - X_1^1$ with $X_0^0 + X_0^1 = \mathbb{I}$ and $X_1^0 + X_1^1 = \mathbb{I}$. We know from Lemma 3.5.2 that there exist projectors Π_j such that we can decompose X_s as $X_s = \sum_j \Pi_j X_s^a \Pi_j$ for $s, b \in 0, 1$, where $\text{rank}(\Pi_j) \leq 2$. Hence, we can immediately conclude from Lemma 6.3.1 that if Alice only measures two possible observables with two outcomes each, the dimension of her state does not need to exceed $d = 2$. This has previously been proved by Masanes [Mas06]. Could we prove something similar for three measurement settings? Sadly, Theorem 3.5.7 tells us that this is not possible! There do exist three measurements for which no such decomposition exists. It is not hard to see that the question of how large Alice's entangled state has to be given a specific set of measurement operators is essentially equivalent to the question of how many qubits we need to store in the problem of post-measurement information to achieve perfect success. In both settings we are interested in reducing the dimension by finding a way to block-diagonalize the matrices.

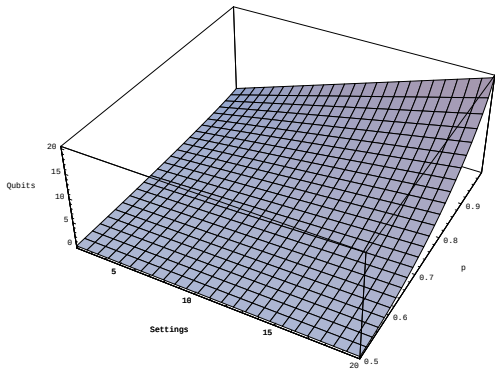


Figure 8.2: Tradeoff for 2 outcomes.

8.5 Conclusion

Bounding the amount of entanglement that we need to implement the optimal strategy in non-local games remains a tricky problem. We have given a simple lower bound on the amount of entanglement necessary for an extremely restricted class of games. The CHSH game forms an instance of such a game. Even though our bound is very weak, and the class of games quite restricted, we are hopeful that our approach may lead to stronger statements in the future. We also showed how our earlier considerations and Tsirelson's construction led to an upper bound for the specific case of XOR-games. Sadly, better bounds still elude us so far.

Chapter 9

Interactive Proof Systems

As we saw in the past chapters, two spatially separated parties, Alice and Bob, can use entanglement to obtain correlations that are impossible to achieve classically, without any additional communication. However, there do exist classical systems whose strength, or security, indeed depends crucially on the fact that specific parties cannot communicate during the course of the protocol. How are such systems affected by the presence of entanglement? Can Alice and Bob use their shared entanglement to gain a significant advantage? Here, we study interactive proof systems which are a specific case of such a classical system. Surprisingly, it turns out that the space-like separation is lost alltogether and we can simulate two classical parties with just a single quantum one.

9.1 Introduction

9.1.1 Classical interactive proof systems

Before getting to the heart of the matter, we first need to take a closer look at interactive proof systems. Classical interactive proof systems have received considerable attention [BFL91, BOGKW88, CCL90, Fei91, LS91, FL92] since their introduction by Babai [Bab85] and Goldwasser, Micali and Rackoff [GMR89] in 1985. An interactive proof system takes the form of a protocol of one or more rounds between two parties, a verifier and a prover. Whereas the prover is computationally unbounded, the verifier is limited to probabilistic polynomial time. Both the prover and the verifier have access to a common input string x. The goal of the prover is to convince the verifier that x belongs to a pre-specified language L. The verifier's aim, on the other hand, is to determine whether the prover's claim is indeed valid. In each round, the verifier sends a polynomial (in x) size query to the prover, who returns a polynomial size answer. At the end of the protocol, the verifier decides to accept, and conclude $x \in L$, or reject based on the messages exchanged and his own private randomness. A language

has an interactive proof if there exists a verifier V and a prover P such that: If $x \in L$, the prover can always convince V to accept. If $x \notin L$, no strategy of the prover can convince V to accept with non-negligible probability. IP denotes the class of languages having an interactive proof system. Watrous [Wat99] first considered the notion of *quantum* interactive proof systems. Here, the prover has unbounded quantum computational power whereas the verifier is restricted to quantum polynomial time. In addition, the two parties can now exchange quantum messages. QIP is the class of languages having a quantum interactive proof system. Classically, it is known that IP = PSPACE [Sha92, She92], where PSPACE is the class of languages decidable using only polynomial space. For the quantum case, it has been shown that PSPACE $\subseteq$ QIP $\subseteq$ EXP [Wat99, KW00]. If, in addition, the verifier is given polynomial size quantum advice, the resulting class QIP/qpoly contains all languages [Raz05]. Let QIP(k) denote the class where the prover and verifier are restricted to exchanging k messages. It is known that QIP = QIP(3) [KW00] and QIP(1) $\subseteq$ PP [Vya03, MW05], where PP is the class of all problems solvable by a probabilistic machine in polynomial time. We refer to [MW05] for an overview of the extensive work done on QIP(1), also known as QMA. Very little is known about QIP(2) and its relation to either PP or PSPACE.

In multiple-prover interactive proof systems the verifier can interact with multiple, computationally unbounded provers. Before the protocol starts, the provers are allowed to agree on a joint strategy, however they can no longer communicate during the execution of the protocol. Let MIP denote the class of languages having a *multiple*-prover interactive proof system. Here, we are especially interested in two-prover interactive proof systems as introduced by Ben-Or, Goldwasser, Kilian and Widgerson [BOGKW88]. Babai, Fortnow and Lund [BFL91], and Feige and Lovász [FL92] have shown that a language is in NEXP if and only if it has a two-prover one-round proof system, i.e., MIP[2] = NEXP. Feige and Lovász have also shown that a system using more than two-provers is thus no more powerful than a system with only two provers, i.e., MIP[2] = MIP. Let $\oplus$MIP[2] denote the restricted class where the verifier's output is a function of the XOR of two binary answers. Even for such a system $\oplus$MIP[2] = NEXP, for certain soundness and completeness parameters [CHTW04a]. Classical multiple-prover interactive proof systems are thus more powerful than classical proof systems based on a single prover, assuming PSPACE $\neq$ NEXP.

9.1.2 Quantum multi-prover interactive proof systems

Given the advent of quantum computing, one can also consider quantum interactive proof systems with *multiple* provers. These can be grouped into two categories: First, one can consider provers and a verifier that are quantum themselves and can exchange quantum messages. Kobayashi and Matsumoto have considered such *quantum* multiple-prover interactive proof systems which form

an extension of quantum single prover interactive proof systems as described above. Let QMIP denote the resulting class. In particular, they showed that QMIP = NEXP if the provers do *not* share quantum entanglement [KM03]. If the provers share at most polynomially many entangled qubits the resulting class is contained in NEXP [KM03].

Secondly, one can consider proof systems where all communication remains classical, but the provers can share any entangled state as part of their strategy on which they are allowed to perform arbitrary measurements. Cleve, Høyer, Toner and Watrous [CHTW04a] have raised the question whether a *classical* two-prover system is weakened in such a setting. We write MIP* if the provers share entanglement. The authors provide a number of examples which demonstrate that the soundness condition of a classical proof system can be compromised, i.e. the interactive proof system is weakened, when entanglement is used. In their paper, it is proved that $\oplus$MIP$^*[2] \subseteq$ NEXP. Later, the same authors also showed that $\oplus$MIP$^*[2] \subseteq$ EXP using semidefinite programming [CHTW04b]. Entanglement thus clearly weakens an interactive proof system, assuming EXP $\neq$ NEXP.

Intuitively, entanglement allows the provers to coordinate their answers, even though they cannot use it to communicate. By measuring the shared entangled state the provers can generate correlations which they can use to deceive the verifier. Tsirelson [Tsi80, Tsi87] has shown that even quantum mechanics limits the strength of such correlations, as we saw in Chapter 6. Recall that Popescu and Roehrlich [PR94, PR96, PR97] have raised the question why nature imposes such limits. To this end, they constructed a toy-theory based on non-local boxes [PR94, vD00], which are hypothetical "machines" generating correlations stronger than possible in nature. In their full generalization, non-local boxes can give rise to any type of correlation as long as they cannot be used to signal. Preda [Pre05] showed that sharing non-local boxes allows two provers to coordinate their answers perfectly and obtained $\oplus\text{MIP}_{\text{NL}} = \text{PSPACE}$, where we write $\oplus\text{MIP}_{\text{NL}}$ to indicate that the two provers share non-local boxes.

Kitaev and Watrous [KW00] mention that it is unlikely that a single-prover *quantum* interactive proof system can simulate multiple classical provers, because then from QIP $\subseteq$ EXP and MIP = NEXP it follows that EXP = NEXP.

Surprisingly, it turns out that when the provers are allowed to share entanglement it can be possible to simulate two such classical provers by one quantum prover. This indicates that entanglement among provers truly leads to a weaker proof system. In particular, we show that a two-prover one-round interactive proof system where the verifier computes the XOR of two binary answers and the provers are allowed to share an arbitrary entangled state, can be simulated by a single quantum interactive proof system with two messages: $\oplus$MIP$^*[2] \subseteq$ QIP(2). Since very little is known about QIP(2) so far [KW00], we hope that our result may help shed some light on its relation to PP or PSPACE. Our result also leads to a proof that $\oplus$MIP$^*[2] \subseteq$ EXP.

9.2 Proof systems and non-local games

9.2.1 Non-local games

For our proof, it is necessary to link interactive proof systems to non-local games, as we described in Chapter 6.2.3. Since we consider only two parties, we omit unnecessary indices and use separate letters to refer to the sets of possible questions and answers. We briefly recap our setup, summarized in Figure 9.1: Let S, T, A and B be finite sets, and π a probability distribution on $S \times T$. Let V be a predicate on $S \times T \times A \times B$. Then $G = G(V, \pi)$ is the following two-person cooperative game[1]: A pair of questions $(s, t) \in S \times T$ is chosen at random according to the probability distribution π. Then s is sent to player 1, henceforth called Alice, and t to player 2, which we call Bob. Upon receiving s, Alice has to reply with an answer $a \in A$. Likewise, Bob has to reply to question t with an answer $b \in B$. They win if $V(s, t, a, b) = 1$ and lose otherwise. Alice and Bob may agree on any kind of strategy beforehand, but they are no longer allowed to communicate once they have received questions s and t. The value $\omega(G)$ of a game G is the maximum probability that Alice and Bob win the game. We write $V(a, b|s, t)$ instead of $V(s, t, a, b)$ to emphasize the fact that a and b are answers given questions s and t.

Here, we are particularly interested in non-local games. Alice and Bob are allowed to share an arbitrary entangled state $|\Psi\rangle$ to help them win the game. Let $\mathcal{H}^A$ and $\mathcal{H}^B$ denote the Hilbert spaces of Alice and Bob respectively. The state $|\Psi\rangle \in \mathcal{H}^A \otimes \mathcal{H}^B$ is part of the quantum strategy that Alice and Bob can agree on beforehand. This means that for each game, Alice and Bob can choose a specific $|\Psi\rangle$ to maximize their chance of success. In addition, Alice and Bob can agree on quantum measurements. For each $s \in S$, Alice has a projective measurement described by $\{X_s^a \mid a \in A\}$ on $\mathcal{H}^A$. For each $t \in T$, Bob has a projective measurement described by $\{Y_t^b \mid b \in B\}$ on $\mathcal{H}^B$. For questions $(s, t) \in S \times T$, Alice performs the measurement corresponding to s on her part of $|\Psi\rangle$ which gives her outcome a. Likewise, Bob performs the measurement corresponding to t on his part of $|\Psi\rangle$ with outcome b. Both send their outcome, a and b, back to the verifier. The probability that Alice and Bob answer $(a, b) \in A \times B$ is then given by

$$\langle \Psi | X_s^a \otimes Y_t^b | \Psi \rangle.$$

The probability that Alice and Bob win the game is now given by

$$\Pr[\text{Alice and Bob win}] = \sum_{s,t} \pi(s, t) \sum_{a,b} V(a, b|s, t) \langle \Psi | X_s^a \otimes Y_t^b | \Psi \rangle. \tag{9.1}$$

The *quantum value* $\omega_q(G)$ of a game G is the maximum probability over all possible quantum strategies that Alice and Bob win. Recall that *XOR game* is

[1] Players 1 and 2 collaborate against the verifier

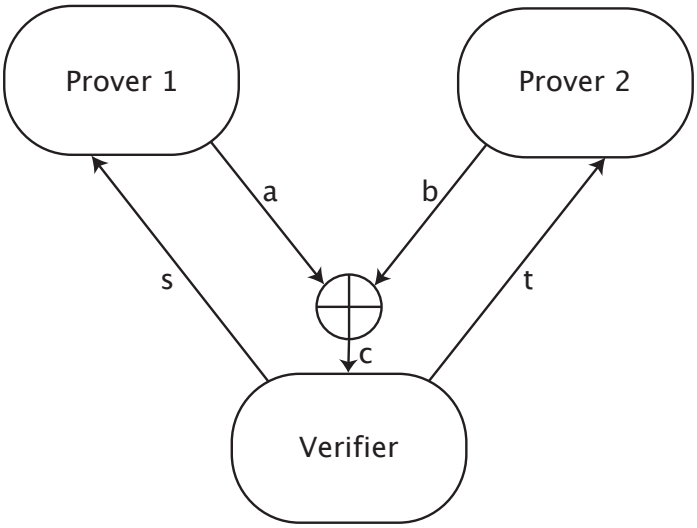


Figure 9.1: A one-round XOR proof system.

a game where the value of V only depends on $c = a \oplus b$ and not on a and b independently. For XOR games we write $V(c|s,t)$ instead of $V(a,b|s,t)$. Here, we are only interested in the case that $a \in \{0,1\}$ and $b \in \{0,1\}$ and XOR games. Alice and Bob's measurements are then described by $\{X_s^0, X_s^1\}$ for $s \in S$ and $\{Y_t^0, Y_t^1\}$ for $t \in T$ respectively. Note that $X_s^0 + X_s^1 = \mathbb{I}$ and $Y_t^0 + Y_t^0 = \mathbb{I}$ and thus these measurements can be expressed in the form of observables X_s and Y_t with eigenvalues ± 1: $X_s = X_s^0 - X_s^1$ and $Y_t = Y_t^0 - Y_t^1$. Recall from Chapter 6.3.2 that Tsirelson [Tsi80, Tsi87] has shown that for any $|\Psi\rangle \in \mathcal{H}^A \otimes \mathcal{H}^B$ there exists real unit vectors $x_s, y_t \in \mathbb{R}^N$ with $N = |S| + |T|$ such that $\langle\Psi|X_s \otimes Y_t|\Psi\rangle = \langle x_s|y_t\rangle$. It is then easy to see from Eq. (9.1) that for XOR games we can express the maximum winning probability as

$$\omega_q(G) = \max_{x_s, y_t} \frac{1}{2} \sum_{s,t} \pi(s,t) \sum_c V(c|s,t) \left(1 + (-1)^c \langle x_s|y_t\rangle\right), \tag{9.2}$$

where the maximization is taken over all unit vectors $x_s, y_t \in \mathbb{R}^N$.

9.2.2 Multiple classical provers

It is well known [CHTW04a, FL92], that two-prover one-round interactive proof systems with classical communication can be modeled as (non-local) games. Here,

Alice and Bob take the role of the two provers. The verifier now poses questions s and t, and evaluates the resulting answers. A proof system associates with each string x a game G_x, where $\omega(G_x)$ determines the probability that the verifier accepts (and thus concludes $x \in L$). The string x, and thus the nature of the game G_x is known to both the verifier and the provers. Ideally, for all $x \in L$ the value of $\omega(G_x)$ is close to one, and for $x \notin L$ the value of $\omega(G_x)$ is close to zero. It is possible to extend the game model for MIP[2] to use a randomized predicate for the acceptance predicate V. This corresponds to V taking an extra input string chosen at random by the verifier. However, known applications of MIP[2] proof systems do not require this extension [Fei95]. Our argument in Section 9.3 can easily be extended to deal with randomized predicates. Since V is not a randomized predicate in [CHTW04a], we follow this approach.

We concentrate on proof systems involving two provers, one round of communication, and single-bit answers. The provers are computationally unbounded, but limited by the laws of quantum physics. However, the verifier is probabilistic polynomial time bounded. As defined by Cleve, Høyer, Toner and Watrous [CHTW04a],

9.2.1. Definition. For $0 \leq s < c \leq 1$, let $\oplus\mathrm{MIP}_{c,s}[2]$ denote the class of all languages L recognized by a classical two-prover interactive proof system of the following form:

- They operate in one round, each prover sends a single bit in response to the verifier's question, and the verifier's decision is a function of the parity of those two bits.
- If $x \in L$ then there exists a strategy for the provers for which the probability that the verifier accepts is at least c (the *completeness* probability).
- If $x \notin L$ then, whatever strategy the two provers follow, the probability that the verifier accepts is at most s (the *soundness* probability).

9.2.2. Definition. For $0 \leq s < c \leq 1$, let $\oplus\mathrm{MIP}^*_{c,s}[2]$ denote the class corresponding to a modified version of the previous definition: all communication remains classical, but the provers may share prior quantum entanglement, which may depend on x, and perform quantum measurements.

We generally omit indices c, s, unless they are explicitly relevant.

In Chapter 7, we discussed how to find the optimal strategies for XOR-games. In particular, we saw that we can determine the optimal value of $\omega_q(G_x)$ in time exponential in $\min(|S|, |T|)$ using semidefinite programming. This implies immediately that $\oplus\mathrm{MIP}^* \subseteq \mathrm{EXP}$, as was shown by Cleve, Høyer, Toner and Watrous [CHTW04a] during their presentation at CCC'04. Here, we show something stronger, namely that $\oplus\mathrm{MIP}^* \subseteq \mathrm{QIP}(2)$.

9.2.3 A single quantum prover

Instead of two classical provers, we can also consider a system consisting of a single quantum prover P_q and a quantum polynomial time verifier V_q as defined by Watrous [Wat99]. Again, the quantum prover P_q is computationally unbounded, however, he is limited by the laws of quantum physics. The verifier and the prover can communicate over a quantum channel. In this thesis, we are only interested in one round quantum interactive proof systems: the verifier sends a single quantum message to the prover, who responds with a quantum answer. We here express the definition of QIP(2) [Wat99] in a form similar to the definition of $\oplus$MIP*:

9.2.3. Definition. Let QIP$(2, c, s)$ denote the class of all languages L recognized by a quantum one-prover one-round interactive proof system of the following form:

- If $x \in L$ then there exists a strategy for the quantum prover for which the probability that the verifier accepts is at least c.
- If $x \notin L$ then, whatever strategy the quantum prover follows, the probability that the quantum verifier accepts is at most s.

9.3 Simulating two classical provers with one quantum prover

We now show that an interactive proof system where the verifier bases his decision only on the XOR of two binary answers is in fact no more powerful than a system based on a single quantum prover. The main idea behind our proof is to combine two classical queries into one quantum query, and thereby simulate the classical proof system with a single quantum prover. Similar techniques have been used to prove results about classical locally decodable codes [KW03, WdW05]. Recall that the two provers can use an arbitrary entangled state as part of their strategy.

For our proof we make use of the fact that we can write the optimal value of the game as in Eq. (9.2).

9.3.1. Theorem. *For all s and c such that $0 \leq s < c \leq 1$, $\oplus\text{MIP}^*_{c,s}[2] \subseteq \text{QIP}(2, c, s)$.*

Proof. Let $L \in \oplus\text{MIP}^*_{c,s}[2]$ and let V_e be a verifier witnessing this fact. Let P_e^1 (Alice) and P_e^2 (Bob) denote the two provers sharing entanglement. Fix an input string x. As mentioned above, interactive proof systems can be modeled as games indexed by the string x. It is therefore sufficient to show that there exists a verifier V_q and a quantum prover P_q, such that $\omega_{sim}(G_x) = \omega_q(G_x)$, where $\omega_{sim}(G_x)$ is the value of the simulated game.

Let s,t be the questions that V_e sends to the two provers P_e^1 and P_e^2 in the original game. The new verifier V_q now constructs the following state in $\mathcal{V} \otimes \mathcal{M}$

$$|\Phi_{init}\rangle = \frac{1}{\sqrt{2}}(\underbrace{|0\rangle}_{\mathcal{V}}\underbrace{|0\rangle|s\rangle}_{\mathcal{M}} + \underbrace{|1\rangle}_{\mathcal{V}}\underbrace{|1\rangle|t\rangle}_{\mathcal{M}}),$$

and sends register $\mathcal{M}$ to the single quantum prover P_q.

We first consider the honest strategy of the prover. Let a and b denote the answers of the two classical provers to questions s and t respectively. The quantum prover now transforms the state to

$$|\Phi_{honest}\rangle = \frac{1}{\sqrt{2}}((-1)^a \underbrace{|0\rangle}_{\mathcal{V}}\underbrace{|0\rangle|s\rangle}_{\mathcal{M}} + (-1)^b \underbrace{|1\rangle}_{\mathcal{V}}\underbrace{|1\rangle|t\rangle}_{\mathcal{M}}),$$

and returns register $\mathcal{M}$ back to the verifier. The verifier V_q now performs a measurement on $\mathcal{V} \otimes \mathcal{M}$ described by the following projectors

$$\begin{aligned} P_0 &= |\Psi_{st}^+\rangle\langle\Psi_{st}^+| \otimes I \\ P_1 &= |\Psi_{st}^-\rangle\langle\Psi_{st}^-| \otimes I \\ P_{reject} &= I - P_0 - P_1, \end{aligned}$$

where $|\Psi_{st}^{\pm}\rangle = (|0\rangle|0\rangle|s\rangle \pm |1\rangle|1\rangle|t\rangle)/\sqrt{2}$. If he obtains outcome "reject", he immediately aborts and concludes that the quantum prover is cheating. If he obtains outcome $m \in \{0,1\}$, the verifier concludes that $c = a \oplus b = m$. Note that $\Pr[m = a \oplus b|s,t] = \langle\Phi_{honest}|P_{a\oplus b}|\Phi_{honest}\rangle = 1$, so the verifier can reconstruct the answer perfectly.

We now consider the case of a dishonest prover. In order to convince the verifier, the prover applies a transformation on $\mathcal{M} \otimes \mathcal{P}$ and send register $\mathcal{M}$ back to the verifier. We show that for any such transformation the value of the resulting game is at most $\omega_q(G_x)$: Note that the state of the total system in $\mathcal{V} \otimes \mathcal{M} \otimes \mathcal{P}$ can now be described as

$$|\Phi_{dish}\rangle = \frac{1}{\sqrt{2}}(|0\rangle|\phi_s\rangle + |1\rangle|\phi_t\rangle)$$

where $|\phi_s\rangle = \sum_{u \in S' \cup T'} |u\rangle|\alpha_u^s\rangle$ and $|\phi_t\rangle = \sum_{v \in S' \cup T'} |v\rangle|\beta_v^t\rangle$ with $S' = \{0s|s \in S\}$ and $T' = \{1t|t \in T\}$. Any transformation employed by the prover can be described this way. We now have that

$$\Pr[m = 0|s,t] = \langle\Phi_{dish}|P_0|\Phi_{dish}\rangle = \frac{1}{4}(\langle\alpha_s^s|\alpha_s^s\rangle + \langle\beta_t^t|\beta_t^t\rangle) + \frac{1}{2}\Re(\langle\alpha_s^s|\beta_t^t\rangle) \quad (9.3)$$

$$\Pr[m = 1|s,t] = \langle\Phi_{dish}|P_1|\Phi_{dish}\rangle = \frac{1}{4}(\langle\alpha_s^s|\alpha_s^s\rangle + \langle\beta_t^t|\beta_t^t\rangle) - \frac{1}{2}\Re(\langle\alpha_s^s|\beta_t^t\rangle) \quad (9.4)$$

The probability that the prover wins is given by

$$\Pr[\text{Prover wins}] = \sum_{s,t} \pi(s,t) \sum_{c \in \{0,1\}} V(c|s,t) \Pr[m = c|s,t].$$

The prover will try to maximize his chance of success by maximizing $\Pr[m = 0|s,t]$ or $\Pr[m = 1|s,t]$. We can therefore restrict ourselves to considering real unit vectors for which $\langle \alpha_s^s | \alpha_s^s \rangle = 1$ and $\langle \beta_t^t | \beta_t^t \rangle = 1$, as the dimension of our vectors is directly determined by their number. Hence, we may also assume that $|\alpha_s^{s'}\rangle = 0$ iff $s \neq s'$ and $|\beta_t^{t'}\rangle = 0$ iff $t \neq t'$: any other strategy can lead to rejection and thus to a lower probability of success. By substituting into Eqs. (9.3) and (9.4), it follows that the probability that the quantum prover wins the game (when he avoids rejection) is

$$\frac{1}{2} \sum_{s,t,c} \pi(s,t) V(c|s,t)(1 + (-1)^c \langle \alpha_s^s | \beta_t^t \rangle). \tag{9.5}$$

In order to convince the verifier, the prover's goal is to choose real vectors $|\alpha_s^s\rangle$ and $|\beta_t^t\rangle$ which maximize Eq. (9.5). Since in $|\phi_s\rangle$ and $|\phi_t\rangle$ we sum over $|S'|+|T'| = |S|+|T|$ elements respectively, the dimension of $\mathcal{P}$ need not exceed $N = |S|+|T|$. Thus, it is sufficient to restrict the maximization to vectors in $\mathbb{R}^{|S|+|T|}$. Given Eq. (9.5), we thus have

$$\omega_{sim}(G_x) = \max_{\alpha_s^s, \beta_t^t} \frac{1}{2} \sum_{s,t,c} \pi(s,t) V(c|s,t)(1 + (-1)^c \langle \alpha_s^s | \beta_t^t \rangle),$$

where the maximization is taken over vectors $\{\alpha_s^s \in \mathbb{R}^N : s \in S\}$, and $\{\beta_t^t \in \mathbb{R}^N : t \in T\}$. However, we know from Eq. (9.2) that

$$\omega_q(G_x) = \max_{x_s, y_t} \frac{1}{2} \sum_{s,t,c} \pi(s,t) V(c|s,t)(1 + (-1)^c \langle x_s | y_t \rangle)$$

where the maximization is taken over unit vectors $\{x_s \in \mathbb{R}^N : s \in S\}$ and $\{y_t \in \mathbb{R}^N : t \in T\}$. We thus have

$$\omega_{sim}(G_x) = \omega_q(G_x)$$

which completes our proof. □

9.3.2. Corollary. *For all s and c such that $0 \leq s < c \leq 1$, $\oplus\mathrm{MIP}^*_{c,s}[2] \subseteq \mathrm{EXP}$.*

Proof. This follows directly from Theorem 9.3.1 and the result that $\mathrm{QIP}(2) \subseteq \mathrm{EXP}$ [KW00]. □

9.4 Conclusion

As we have shown, the strength of classical systems can be weakened considerably in the presence of entanglement. In our example above, we showed that the systems can be weakened so much that all space-like separation is lost: we saw that two classical parties with entanglement are as powerful as a single quantum party.

It would be interesting to show that this result also holds for a proof system where the verifier is not restricted to computing the XOR of both answers, but some other Boolean function. However, the approach based on vectors from Tsirelson's results does not work for binary games. Whereas it is easy to construct a single quantum query which allows the verifier to compute an arbitrary function of the two binary answers with some advantage, it thus remains unclear how the value of the resulting game is related to the value of a binary game. Furthermore, mere classical tricks trying to obtain the value of a binary function from XOR itself seem to confer extra cheating power to the provers.

Examples of non-local games with longer answers [CHTW04a], such as the Kochen-Specker or the Magic Square game, seem to make it even easier for the provers to cheat by taking advantage of their entangled state. Furthermore, existing proofs that MIP = NEXP break down if the provers share entanglement. It is therefore an open question whether $\mathrm{MIP}^* = \mathrm{NEXP}$ or, $\mathrm{MIP}^* \subseteq \mathrm{EXP}$.

As described, non-locality experiments between two space-like separated observers, Alice and Bob, can be cast in the form of non-local games. For example, the experiment based on the well known CHSH inequality [CHSH69], is a non-local game with binary answers of which the verifier computes the XOR [CHTW04a]. Our result implies that this non-local game can be simulated in superposition by a single prover/observer: Any strategy that Alice and Bob might employ in the non-local game can be mirrored by the single prover in the constructed "superposition game", and also vice versa, due to Tsirelson's constructions [Tsi80, Tsi87] mentioned earlier. This means that the "superposition game" corresponding to the non-local CHSH game is in fact limited by Tsirelson's inequality [Tsi80], even though it itself has no non-local character. Whereas this may be purely coincidental, it would be interesting to know its physical interpretation, if any. Perhaps it may be interesting to ask whether Tsirelson-type inequalities have any consequences on local computations in general, beyond the scope of these very limited games.

Part IV

Consequences for Crytography

Chapter 10

Limitations

Finally, we turn our attention to cryptographic protocols directly. As we saw in Chapter 1, it is impossible to implement bit commitment even in the quantum setting! In the face of the negative results, what can we still hope to achieve?

10.1 Introduction

Here, we consider the task of committing to an entire string of n bits at once when both the honest player and the adversary have unbounded resources. Since perfect bit commitment is impossible, perfect bit string commitment is clearly impossible as well. Curiously, however, we can still make interesting statements in the quantum setting, if we give both Alice and Bob a limited ability to cheat. That is, we allow Alice to change her mind about the committed string within certain limited parameters, and allow Bob to gain some information about the committed string. It turns out that it matters crucially how we measure Bob's information gain.

First, we introduce a framework for the classification of bit string commitments in terms of the length n of the string, Alice's ability to cheat on at most a bits and Bob's ability to acquire at most b bits of information before the reveal phase. We say that Alice can cheat on a bits if she can reveal up to 2^a strings successfully. Bob's security definition is crucial to our investigation: If b determines a bound on his probability to guess Alice's string, then we prove that $a + b$ is at least n. This implies that the trivial protocol, where Alice's commitment consists of sending b bits of her string to Bob, is optimal. If, however, b is a bound on the accessible information that the quantum states contain about Alice's string, then we show that non-trivial schemes exist. More precisely, we construct schemes with $a = 4 \log n + O(1)$ and $b = 4$. This is impossible classically. We also present a simple, implementable, protocol, that achieves $a = 1$ and $b = n/2$. This protocol can furthermore be made cheat-sensitive. Quantum commitments of strings have previously been considered by Kent [Ken03], who pointed out that in the

quantum world useful bit string commitments could be possible despite the no-go theorem for bit commitment. His scenario differs significantly from ours and imposes an additional constraint, which is not present in our work: Alice does not commit to a superposition of strings.

10.2 Preliminaries

10.2.1 Definitions

We first formalize the notion of quantum string commitments in a quantum setting.

10.2.1. Definition. *An* (n, a, b)-Quantum Bit String Commitment (QBSC) *is a quantum communication protocol between two parties, Alice (the committer) and Bob (the receiver), which consists of two phases:*

- *(Commit Phase) Assume that both parties are honest. Alice chooses a string* $x \in \{0,1\}^n$ *with probability* p_x*. Alice and Bob communicate and at the end Bob holds state* ρ_x*.*
- *(Reveal Phase) If both parties are honest, Alice and Bob communicate and at the end Bob outputs* x*. Bob accepts.*

We have the following two security requirements:

- *(Concealing) If Alice is honest, then for any strategy of Bob*

$$\sum_{x\in\{0,1\}^n} p^B_{x|x} \leq 2^b,$$

  where $p^B_{x|x}$ *is the probability that Bob correctly guesses* x *before the reveal phase.*

- *(Binding) If Bob is honest, then for any strategy of Alice*

$$\sum_{x\in\{0,1\}^n} p^A_x \leq 2^a,$$

  where p^A_x *is the probability that Alice successfully reveals* x *(Bob accepts the opening of* x*).*

Bob thereby accepts the opening of a string x, if he performs a test depending on the individual protocol to check Alice's honesty and concludes that she was indeed honest. Note that quantumly, Alice can always commit to a superposition of different strings without being detected. Thus even for a perfectly binding bit

string commitment (i.e. $a = 0$) we only demand that $\sum_{x\in\{0,1\}^n} p_x^A \leq 1$, whereas classically one wants that $p_{x'}^A = \delta_{x,x'}$. Note that our concealing definition reflects Bob's a priori knowledge about x. We choose an a priori uniform distribution (i.e. $p_x = 2^{-n}$) for (n, a, b)-QBSCs, which naturally comes from the fact that we consider n-bit strings. A generalization to any (P_X, a, b)-QBSC where P_X is an arbitrary distribution is possible but omitted in order not to obscure our main line of argument.

Instead of Bob's guessing probability, one can take any information measure B to express the security against Bob. In general, we consider an (n, a, b)-QBSC$_B$ where the new Concealing-condition reads

- (General Concealing) If Alice is honest, then for any ensemble $\mathcal{E} = \{p_x, \rho_x\}$ that Bob can obtain by a cheating strategy $B(\mathcal{E}) \leq b$.

Later, we will show that for B being the *accessible information*, non-trivial protocols, i.e. protocols with $a+b \ll n$, do exist. Recall that the accessible information was defined in Section 2.3.2 as $\mathcal{I}_{acc}(\mathcal{E}) = \max_M I(X, Y)$, where P_X is the prior distribution of the random variable X, Y is the random variable of the outcome of Bob's measurement on $\mathcal{E}$, and the maximization is taken over all measurements M.

10.2.2 Model

We work in the model of two-party non-relativistic quantum protocols of Yao [Yao95], simplified by Lo and Chau [LC97] which is usually adopted in this context. Here, any two-party quantum protocol can be regarded as a pair of quantum machines (Alice and Bob), interacting through a quantum channel. Consider the product of three Hilbert spaces $\mathcal{H}_A$, $\mathcal{H}_B$ and $\mathcal{H}_C$ of bounded dimensions, representing the Hilbert spaces of Alice's and Bob's machines and the channel, respectively. Without loss of generality, we assume that each machine is initially in a specified pure state. Alice and Bob perform a number of rounds of communication over the channel. Each such round can be modeled as a unitary transformation on $\mathcal{H}_A \otimes \mathcal{H}_C$ and $\mathcal{H}_B \otimes \mathcal{H}_C$ respectively. Since the protocol is known to both Alice and Bob, they know the set of possible unitary transformations used in the protocol. We assume that Alice and Bob are in possession of both a quantum computer and a quantum storage device. This enables them to add ancillae to the quantum machine and use reversible unitary operations to replace measurements. The state of this ancilla can then be read off only at the end of the protocol, and by doing so, Alice and Bob can effectively delay any measurements until the very end. The resulting protocol will be equivalent to the original and thus we can limit ourselves to protocols where both parties only measure at the very end. Moreover, any classical computation or communication that may occur can be simulated by a quantum computer.

10.2.3 Tools

We now gather the essential ingredients for our proof. First, we now show that every (n, a, b)-QBSC is an (n, a, b)-QBSC$_\xi$. The security measure $\xi(\mathcal{E})$ is defined by

$$\xi(\mathcal{E}) := n - H_2(\rho_{AB}|\rho), \tag{10.1}$$

where $\rho_{AB} = \sum_x p_x |x\rangle\langle x| \otimes \rho_x$ and $\rho = \sum_x p_x \rho_x$ are only dependent on the ensemble $\mathcal{E} = \{p_x, \rho_x\}$. $H_2(\cdot|\cdot)$ is an entropic quantity defined in [Ren05]

$$H_2(\rho_{AB}|\rho) := -\log \mathrm{Tr}\left(\left[\mathbb{I} \otimes \rho^{-\frac{1}{2}})\rho_{AB}\right]^2\right).$$

Interestingly, this quantity is directly connected to Bob's maximal average probability of successfully guessing the string:

10.2.2. Lemma. *Bob's maximal average probability of successfully guessing the committed string, i.e.* $\sup_M \sum_x p_x p_{x|x}^{B,M}$ *where* $M = \{M_x\}$ *ranges over all measurements and* $p_{y|x}^{B,M} = \mathrm{Tr}(M_y \rho_x)$ *is the conditional probability of outputting* y *given* ρ_x*, obeys*

$$\sup_M \sum_x p_x p_{x|x}^{B,M} \geq 2^{-H_2(\rho_{AB}|\rho)}.$$

Proof. By definition, the maximum average guessing probability is lower bounded by the average guessing probability for a particular measurement strategy. We choose the *square-root measurement* which has operators

$$M_x = p_x \rho^{-\frac{1}{2}} \rho_x \rho^{-\frac{1}{2}}.$$

We use $p_{x|x}^B = \mathrm{Tr}(M_x \rho_x)$ to denote the probability that Bob guesses x given ρ_x, hence

$$\begin{aligned}
\log \sum_x p_x p_{x|x}^{B,\max} &\geq \log \sum_x p_x^2 \mathrm{Tr}(\rho^{-\frac{1}{2}} \rho_x \rho^{-\frac{1}{2}} \rho_x) \\
&= \log \mathrm{Tr}\left(\sum_x p_x^2 |x\rangle\langle x| \otimes \rho^{-\frac{1}{2}} \rho_x \rho^{-\frac{1}{2}} \rho_x\right) \\
&= \log \mathrm{Tr}\left(\left[(\mathbb{I} \otimes \rho^{-\frac{1}{2}})\rho_{AB}\right]^2\right) \\
&= -H_2(\rho_{AB}|\rho)
\end{aligned}$$

□

Related estimates were derived in [BK02].

Furthermore, we make use of the following theorem, known as *privacy amplification against a quantum adversary* with two-universal hash functions, which

we state in a form that is most convenient for our purposes in this chapter. A class $\mathcal{F}$ of functions $f : \{0,1\}^n \to \{0,1\}^\ell$ is thereby called *two-universal* if for all $x \neq y \in \{0,1\}^n$ and for uniformly at random chosen $f \in \mathcal{F}$ we have $\Pr[f(x) = f(y)] \leq 2^{-\ell}$. For example, the set of all affine functions[1]. from $\{0,1\}^n$ to $\{0,1\}^\ell$ is two-universal [CW79]. The following theorem expresses how hash functions can decrease Bob's knowledge about a random variable when he holds some quantum information. In our case, Bob will hold some quantum memory and privacy amplification is used to find Alice's attack.

10.2.3. Theorem (Th. 5.5.1 in [Ren05] (see also [KMR05])). *Let $\mathcal{G}$ be a class of two-universal hash functions from $\{0,1\}^n$ to $\{0,1\}^s$. Application of $g \in \mathcal{G}$ to the random variable X maps the ensemble $\mathcal{E} = \{p_x, \rho_x\}$ to $\mathcal{E}_g = \{q_y^g, \sigma_y^g\}$ with probabilities $q_y^g = \sum_{x \in g^{-1}(y)} p_x$ and quantum states $\sigma_y^g = \sum_{x \in g^{-1}(y)} p_x \rho_x$. Then*

$$\frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} d(\mathcal{E}_g) \leq \frac{1}{2} 2^{-\frac{1}{2}[H_2(\rho_{AB}|\rho) - s]}, \tag{10.2}$$

where $d(\mathcal{E}) := D\big(\sum_x p_x |x\rangle\langle x| \otimes \rho_x, \mathbb{I}/2^n \otimes \rho\big)$ (and similarly for $d(\mathcal{E}_g)$).

Finally, the following reasoning that is used to prove the impossibility of quantum bit commitment [LC97, May96b] will be essential: Suppose ρ_0 and ρ_1 are density operators that correspond to the state of Bob's system if Alice committed a "0" or a "1" respectively. Let $|\phi_0\rangle$ and $|\phi_1\rangle$ be the corresponding purifications on the joint system of Alice and Bob: Alice holds the purification of ρ_0 and ρ_1. If ρ_0 equals ρ_1 then Alice can find a local unitary transformation U that Alice can apply to her part of the system such that $|\phi_1\rangle = U \otimes \mathbb{I}|\phi_0\rangle$. This enables Alice to change the total state from $|\phi_0\rangle$ to $|\phi_1\rangle$ and thus cheat using entanglement! This reasoning also holds in an approximate sense [May96b], here used in the following form:

10.2.4. Lemma. *Let $D(\rho_0, \rho_1) \leq \epsilon$ and assume that the bit-commitment protocol is error-free if both parties are honest. Then there exists a method for Alice to cheat such that the probability of successfully revealing a 0 during the reveal phase, given that she honestly committed herself to a 1 during the commit phase, is at least $1 - \sqrt{2\epsilon}$.*

Proof. $D(\rho_0, \rho_1) \leq \epsilon$ implies $\max_U |\langle\phi_0|U \otimes \mathbb{I}|\phi_1\rangle| \geq 1 - \varepsilon$ by Uhlmann's theorem [Uhl76]. Here, $|\phi_0\rangle$ and $|\phi_1\rangle$ correspond to the joint states after the commit phase if Alice committed to a '0' or '1' respectively where the maximization ranges over all unitaries U on Alice's (i.e. the purification) side. Let $|\psi_0\rangle = U \otimes \mathbb{I}|\phi_1\rangle$ for a U achieving the maximization, be the state that Alice prepares by applying

[1] Geometrically, an affine function is a linear function plus a translation

U to the state on her side when she wants to reveal a '1', given a prior honest commitment to '0'. We then have

$$\begin{aligned} D(|\phi_0\rangle\langle\phi_0|, |\psi_0\rangle\langle\psi_0|) &= \sqrt{1-|\langle\phi_0|\psi_0\rangle|^2} \\ &\leq \sqrt{1-(1-\epsilon)^2} \\ &\leq \sqrt{2\epsilon}. \end{aligned}$$

If Bob is honest, the reveal phase can be regarded as a measurement resulting in a distribution P_Y (or P_Z) if $|\phi_0\rangle$ (or $|\psi_0\rangle$) was the state before the reveal phase. The random variables Y and Z can take values $\{0,1\}$ (corresponding to the opened bit) or the value 'reject (r)'. Since the trace distance does not increase under measurements, $D(P_Y, P_Z) \leq D(|\phi_0\rangle\langle\phi_0|, |\psi_0\rangle\langle\psi_0|) \leq \sqrt{2\epsilon}$. Hence $\frac{1}{2}(|P_Y(0) - P_Z(0)| + |P_Y(1) - P_Z(1)| + |P_Y(r) - P_Z(r)|) \leq \sqrt{2\epsilon}$. Since $|\phi_0\rangle$ corresponds to Alice's honest commitment to 0 we have $P_Y(0) = 1$, $P_Y(1) = P_Y(r) = 0$ and hence $P_Z(0) \geq 1 - \sqrt{2\epsilon}$. □

10.3 Impossibility of quantum string commitments

As we saw above, any (n, a, b)-QBSC is also an (n, a, b)-QBSC$_\xi$ with the security measure $\xi(\mathcal{E})$ defined in Eq. (10.1). To prove our impossibility result we now prove that an (n, a, b)-QBSC$_\xi$ can only exist for values a, b and n obeying $a + b + c \geq n$, where c is a small constant independent of a, b and n. This in turn implies the impossibility of an (n, a, b)-QBSC for such parameters. Finally, we show that if we execute the protocol many times in parallel, the protocol can only be secure if $a + b \geq n$.

The intuition behind our proof is simple: To cheat, Alice first chooses a two-universal hash function g. She then commits to a superposition of all strings for which $g(x) = y$ for a specific y. We now know from the privacy amplification theorem above, that even though Bob may gain some knowledge about x, he is entirely ignorant about y. But then Alice can change her mind and reveal a string from a different set of strings for which $g(x) = y'$ with $y \neq y'$ as we saw above! The following figure illustrates this idea.

10.3.1. THEOREM. *(n, a, b)-QBSC$_\xi$ schemes with $a + b + c < n$ do not exist, where $c = 5 \log 5 - 4 \approx 7.61$ is a constant.*

Proof. Consider an (n, a, b)-QBSC$_\xi$ and the case where both Alice and Bob are honest. Alice committed to x. We denote the joint state of the system Alice-Bob-Channel $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ after the commit phase by $|\phi_x\rangle$ for input state $|x\rangle$. Let ρ_x be Bob's reduced density matrix, and let $\mathcal{E} = \{p_x, \rho_x\}$ where $p_x = 2^{-n}$.

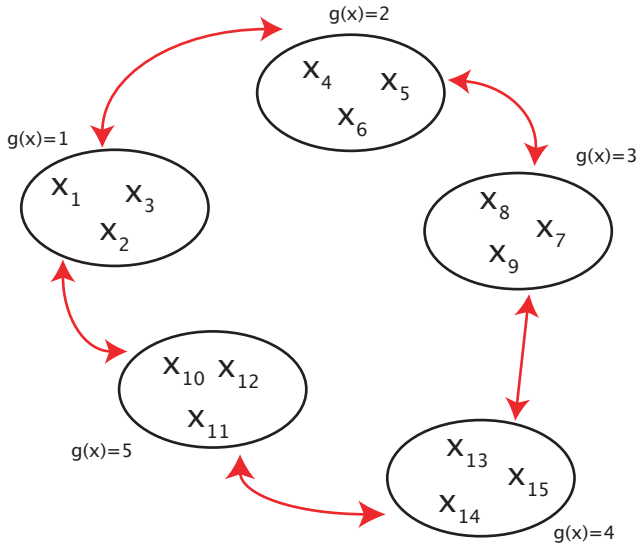


Figure 10.1: Moving from a set of string with $g(x) = y$ to a set of strings with $g(x) = (y \mod 5) + 1$.

Assuming that Bob is honest, we will give a cheating strategy for Alice in the case where $a + b + 5\log 5 - 4 < n$. The strategy will depend on the two-universal hash function $g : \mathcal{X} = \{0,1\}^n \to \mathcal{Y} = \{0,1\}^{n-m}$, for appropriately chosen m. Alice picks a $y \in \mathcal{Y}$ and constructs the state $(\sum_{x \in g^{-1}(y)} |x\rangle|x\rangle)/\sqrt{|g^{-1}(y)|}$. She then gives the second half of this state as input to the protocol and stays honest for the rest of the commit phase. The joint state of Alice and Bob at the end of the commit phase is thus $|\psi_y^g\rangle = (\sum_{x \in g^{-1}(y)} |x\rangle|\phi_x\rangle)/\sqrt{|g^{-1}(y)|}$. The reduced states on Bob's side are $\sigma_y^g = \frac{1}{q_y^g} \sum_{x \in g^{-1}(y)} p_x \rho_x$ with probability $q_y^g = \sum_{x \in g^{-1}(y)} p_x$. We denote this ensemble by $\mathcal{E}_g$. Let $\sigma^g = \sum_y q_y^g \sigma_y^g$.

We now apply Theorem 10.2.3 with $s = n - m$ and $\xi(\mathcal{E}) \leq b$ and obtain $\frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} d(\mathcal{E}_g) \leq \varepsilon$ where $\varepsilon = \frac{1}{2} 2^{-\frac{1}{2}(m-b)}$. Hence, there is at least one g such that $d(\mathcal{E}_g) \leq \varepsilon$. Intuitively, this means that Bob knows only very little about the value of $g(x)$. This g defines Alice's cheating strategy. It is straightforward to verify that $d(\mathcal{E}_g) \leq \varepsilon$ implies

$$2^{-(n-m)} \sum_{y \in \mathcal{Y}} D(\sigma^g, \sigma_y^g) \leq 2\varepsilon. \tag{10.3}$$

We therefore assume without loss of generality that Alice chooses $y_0 \in \mathcal{Y}$ with $D(\sigma_{y_0}^g, \sigma^g) \leq 2\varepsilon$.

We first observe that the probability to successfully reveal some x in $g^{-1}(y)$

given $|\psi_y^g\rangle$ is one[2]. We say that Alice reveals y if she reveals an x such that $y = g(x)$. We then also have that the probability for Alice to reveal y given $|\psi_y^g\rangle$ successfully is one. Let $\tilde{p}_x$ and $\tilde{q}_y^g$ denote the probabilities to successfully reveal x and y respectively and $\tilde{p}_{x|y}^g$ be the conditional probability to successfully reveal x, given y. We have

$$\sum_x \tilde{p}_x = \sum_y \tilde{q}_y^g \sum_{x \in g^{-1}(y)} \tilde{p}_{x|y}^g \geq \sum_y \tilde{q}_y^g,$$

where the inequality follows from our observation above.

As in the impossibility proof of bit commitment, Alice can now transform $|\psi_{y_0}^g\rangle$ approximately into $|\psi_y^g\rangle$ if $\sigma_{y_0}^g$ is sufficiently close to σ_y^g by using only local transformations on her part. Indeed, Lemma 10.2.4 tells us how to bound the probability of revealing y, given that the state was really $|\psi_{y_0}\rangle$. Since this reasoning applies to all y, on average, we have

$$\begin{aligned}
\sum_y \tilde{q}_y^g &\geq \sum_y \left(1 - \sqrt{2}\sqrt{D(\sigma_{y_0}^g, \sigma_y^g)}\right) \\
&\geq 2^{n-m} - 2^{n-m}\sqrt{2}\sqrt{2^{m-n}\sum_y D(\sigma_{y_0}^g, \sigma_y^g)} \\
&\geq 2^{n-m}\left(1 - \sqrt{2}\sqrt{2^{m-n}\left(\sum_y D(\sigma_{y_0}^g, \sigma^g) + D(\sigma^g, \sigma_y^g)\right)}\right) \\
&\geq 2^{n-m}(1 - 2\sqrt{2\varepsilon}),
\end{aligned}$$

where the first inequality follows from Lemma 10.2.4, the second from Jensen's inequality and the concavity of the square root function, the third from the triangle inequality and the fourth from Eq. (10.3) and $D(\sigma_{y_0}^g, \sigma^g) \leq 2\varepsilon$. Recall that to be secure against Alice, we require $2^a \geq 2^{n-m}(1 - 2\sqrt{2\varepsilon})$. We insert $\varepsilon = \frac{1}{2}2^{-\frac{1}{2}(m-b)}$, define $m = b + \gamma$ and take the logarithm on both sides to get

$$a + b + \delta \geq n, \tag{10.4}$$

where $\delta = \gamma - \log(1 - 2^{-\gamma/4+1})$. Keeping in mind that $1 - 2^{-\gamma/4+1} > 0$ (or equivalently $\gamma > 4$), we find that the minimum value of δ for which Eq. (10.4) is satisfied is $\delta = 5\log 5 - 4$ and arises from $\gamma = 4(\log 5 - 1)$. Thus, no (n, a, b)-QBSC$_\xi$ with $a + b + 5\log 5 - 4 < n$ exists. □

It follows immediately that the same restriction holds for an (n, a, b)-QBSC:

[2]Alice learns x, but can't pick it: she committed to a superposition and x is chosen randomly by measurement.

10.3.2. COROLLARY. *(n, a, b)-QBSC schemes, with $a + b + c < n$ do not exist, where $c = 5 \log 5 - 4 \approx 7.61$ is a constant.*

Proof. For the uniform distribution $p_x = 2^{-n}$, we have from the concealing condition that $\sum_x p^B_{x|x} \leq 2^b$, which by Lemma 10.2.2 implies $\xi(\mathcal{E}) \leq b$. Thus, a (n, a, b)-QBSC is an (n, a, b)-QBSC$_\xi$ from which the result follows. □

Since the constant c does not depend on a, b and n, multiple parallel executions of the protocol can only be secure if $a + b \geq n$. This follows by considering m parallel executions of the protocol as a single execution with a string of length mn.

10.3.3. COROLLARY. *Let P be an (n, a, b)-QBSC with P^m an (mn, ma, mb)-QBSC. Then $n < a + b + c/m$. In particular, no (n, a, b)-QBSC with $a + b < n$ can be executed securely an arbitrary number of times in parallel.*

Thus, we can indeed hope to do no better than the trivial protocol. It follows directly from [KMP04] that the results in this section also hold in the presence of superselection rules, where, very informally, quantum actions are restricted to act only on certain subspaces of a larger Hilbert space.

10.4 Possibility

Surprisingly, if one is willing to measure Bob's ability to learn x using a weaker measure of information, the accessible information, non-trivial protocols become possible. These protocols are based on a discovery known as "locking of classical information in quantum states" which we already encountered in Chapter 5.

The protocol, which we call LOCKCOM$(n, \mathcal{U})$, uses this effect and is specified by a set $\mathcal{U} = \{U_1, \ldots, U_{|\mathcal{U}|}\}$ of unitaries. We have

Protocol 1: LOCKCOM$(n, \mathcal{U})$

1: Commit phase: Alice has the string $x \in \{0,1\}^n$ and randomly chooses $r \in \{1, \ldots, |\mathcal{U}|\}$. She sends the state $U_r|x\rangle$ to Bob, where $U_r \in \mathcal{U}$.

2: Reveal phase: Alice announces r and x. Bob applies $U_r^\dagger$ and measures in the computational basis to obtain x'. He accepts if and only if $x' = x$.

We now first show that our protocol is secure with respect to Definition 10.2.1 if Alice is dishonest. Note that our proof only depends on the number of unitaries used, and is independent of a concrete instantiation of the protocol.

10.4.1. LEMMA. *For any LOCKCOM$(n, \mathcal{U})$ protocol the security against a dishonest Alice is bounded by $2^a \leq |\mathcal{U}|$,*

Proof. Let $\tilde{p}_x$ denote the probability that Alice reveals x successfully. Then, $\tilde{p}_x \leq \sum_r \tilde{p}_{x,r}$, where $\tilde{p}_{x,r}$ is the probability that x is accepted by Bob when the reveal information was r. Let ρ denote the state of Bob's system. Summation over x now yields

$$\begin{aligned}\sum_x \tilde{p}_x &\leq \sum_{x,r} \tilde{p}_{x,r} \\ &= \sum_{x,r} \mathrm{Tr}|x\rangle\langle x|U_r^\dagger \rho U_r \\ &= \sum_r \mathrm{Tr}\rho = 2^a.\end{aligned}$$

□

In order to examine security against a dishonest Bob, we have to consider the actual form of the unitaries. We first show that there do indeed exist interesting protocols. Secondly, we present a simple, implementable, protocol. To see that interesting protocols can exist, let Alice choose a set of $O(n^4)$ unitaries independently according to the Haar measure (approximately discretized) and announce the resulting set $\mathcal{U}$ to Bob. They then perform LOCKCOM(n,$\mathcal{U}$). Following the work of [HLSW04], we now show that this variant is secure against Bob with high probability. That is, there exist $O(n^4)$ unitaries that bring Bob's accessible information down to a constant: $\mathcal{I}_{acc}(\mathcal{E}) \leq 4$:

10.4.2. Theorem. *For $n \geq 3$, there exist $(n, 4\log n + O(1), 4)$-$QBSC_{\mathcal{I}_{acc}}$ protocols.*

Proof. Let $\mathcal{U}_{ran}$ denote the set of m randomly chosen bases and consider the LOCKCOM(n, a, b) scheme using unitaries $\mathcal{U} = \mathcal{U}_{ran}$. Security against Alice is again given by Lemma 10.4.1. We now need to show that this choice of unitaries achieves the desired locking effect and thus security against Bob. Again, let $d = 2^n$ denote the dimension. As we saw in Section 5.2.1 we have that

$$\mathcal{I}_{acc}(\mathcal{E}) \leq \log d + \max_{|\phi\rangle} \sum_j \frac{1}{m} H(X_j),$$

where X_j denotes the outcome of the measurement of $|\phi\rangle$ in basis j and the maximum is taken over all pure states $|\phi\rangle$. According to [HLSW04, Appendix B] there is a constant $C' > 0$ such that

$$\begin{aligned}\Pr[\inf_\phi \frac{1}{m}\sum_{j=1}^m H(X_j) &\leq (1-\varepsilon)\log d - 3] \\ &\leq \left(\frac{10}{\varepsilon}\right)^{2d} 2^{-m\left(\frac{\varepsilon C' d}{2(\log d)^2} - 1\right)},\end{aligned}$$

for $d \geq 7$ and $\varepsilon \leq 2/5$. Set $\varepsilon = \frac{1}{\log d}$. The RHS of the above equation then decreases provided that $m > \frac{8}{C'}(\log d)^4$. Thus with $d = 2^n$ and $\log m = 4\log n + O(1)$, the accessible information of the ensemble corresponding to the commitment is then $\mathcal{I}_{acc}(\mathcal{E}) \leq \log d - (1-\varepsilon)\log d + 3 = \varepsilon \log d + 3 = 4$ for our choice of ε. □

Unfortunately, the protocol is inefficient both in terms of computation and communication. It remains open to find an efficient constructive scheme with those parameters.

In contrast, for only two bases, an efficient construction exists and uses the identity and the Hadamard transform as unitaries. For this case, the security of the standard LOCKCOM protocol follows immediately from the locking arguments of Chapter 5. It has been shown that this protocol can be made cheat-sensitive [Chr05].

10.4.3. Theorem. *LOCKCOM$(n, 1, n/2)$ using $\mathcal{U} = \{\mathbb{I}^{\otimes n}, H^{\otimes n}\}$ is a $(n, 1, n/2)$-QBSC$_{\mathcal{I}_{acc}}$ protocol.*

Proof. The result follows immediately from Lemma 10.4.1 and the fact that by Corollary 5.2.3 $\mathcal{I}_{acc}(\mathcal{E}) \leq n/2$ for Bob. □

We can thus obtain non-trivial protocols by exploiting the locking effects discussed in Chapter 5. Note, however, that the security parameters are very weak. Indeed, if Alice uses only two possible bases chosen with equal probability then Bob is always able to obtain the encoded string with probability at least 1/2: he simply guesses the basis and performs the corresponding measurement.

10.5 Conclusion

We have introduced a framework for quantum commitments to a string of bits. Even if we consider string commitments that are weaker than bit commitments, no non-trivial protocols can exist if we choose a very strong measure of security. A property of quantum states known as *locking*, however, allowed us to propose meaningful protocols for a much weaker security demand. One could extend our method to the case of weak secure function evaluation as was done for the original bit commitment protocol in [Lo97]. After completion of our work, Jain [Jai05] has also shown using a different method that QBSC$_\chi$ protocols with $a + 16b + 31 < n$ cannot exist.

A drawback of weakening the security requirement is that LOCKCOM protocols are not necessarily composable. Thus, if LOCKCOM is used as a sub-protocol in a larger protocol, the security of the resulting scheme has to be evaluated on a case by case basis. However, LOCKCOM protocols are secure when executed

in parallel. This is a consequence of the definition of Alice's security parameter and the additivity of the accessible information (see Chapter 2), and sufficient for many cryptographic purposes.

However, two important open questions remain: First, how can we construct efficient protocols using more than two bases? It may be tempting to conclude that we could simply use a larger number of mutually unbiased bases, such as given by the identity and Hadamard transform. Yet, as we saw in Chapter 4 using more mutually unbiased bases does not necessarily lead to a better locking effect and thus better string commitment protocols. Finally, are there any novel applications for this weak quantum string commitment?

Fortunately, it turns out that we can implement protocol with very strong security parameters if we are willing to introduce additional assumptions. We now show how to obtain oblivious transfer from the assumption that qubits are affected by noise during storage.

Chapter 11

Possibilities: Exploiting storage errors

Given the negative results from the last chapter, what can we still hope to achieve? Fortunately, the situation is not quite as bleak if we are taking advantage of the technical limitation that quantum storage is necessarily noisy. Here, the very problem that still prevents us from implementing a quantum computer can actually be turned to our advantage! As we saw in Chapter 1 the primitive of oblivious transfer allows us to implement essentially all cryptographic protocols among two mutually distrustful players, and hence we focus on this primitive.

11.1 Introduction

As outlined in Chapter 1, it was recently shown that secure OT is possible when the receiver Bob has a limited amount of quantum memory [DFSS05, DFR$^+$07] at his disposal. Within this 'bounded-quantum-storage model' OT can be implemented securely as long as a dishonest receiver Bob can store at most $n/4 - O(1)$ qubits coherently, where n is the number of qubits transmitted from Alice to Bob. The problem with this approach is that it assumes an explicit limit on the physical number of qubits (or more precisely, the rank of the adversary's quantum state). However, at present we do not know of any practical physical situation which enforces such a limit for quantum information. On the other hand it is a fact that currently and in the near-future storing photonic qubits is noisy. We therefore propose an alternative model of *noisy-quantum storage* inspired by present-day physical implementations: We require no explicit memory bound, but we assume that any qubit that is placed into quantum storage undergoes a certain amount of noise. Here, we take the 1-2 OT protocol from [DFR$^+$07] as our starting point, and analyze it in this model. This simple 1-2 OT protocol can be implemented using photonic qubits (using polarization or phase-encoding) with standard BB84 quantum key distribution [BB84, GRTZ02] hardware, only with different classical post-processing.

Our adversary model is that of collective attacks (in analogy with collective eavesdropping attacks in the quantum key distribution setting). More precisely:

- Bob may choose to (partially) measure (a subset of) his qubits immediately upon reception using an error-free *product* measurement.
- Bob may store each incoming qubit, or post-measurement state from a prior partial measurement, separately and wait until he gets additional information from Alice (at Step 3 in Protocol 1).
- Once he obtained the additional information he may perform an arbitrary coherent measurement on his stored qubits and stored classical data.

We assume that a qubit q_i undergoes some noise while in storage, where we denote the combined channel given by Bob's initial (partial) measurement, followed by the noise by super-operator $\mathcal{S}_i$. The source of noise can be due to the transfer of qubit onto a different physical carrier, such as an atomic ensemble or atomic state for example, or into an error-correcting code with fidelity less than 1. In addition, the (encoded) qubit will undergo noise once it has been transferred into 'storage'. Hence, the quantum operation $\mathcal{S}_i$ in any real world setting will necessarily include some form of noise. Note that such noise is typically much larger than the noise experienced by honest players who only need to make immediate complete measurements in the BB84 basis.

First of all, we show that for any initial measurement by Bob, and any noisy superoperator $\mathcal{S}_i$ the 1-2 OT protocol is secure if the honest players can perform *perfect* noise-free quantum operations. As an explicit example, we consider depolarizing noise for which reduce the set of optimal attacks to two simple ones: measure in the so-called Breidbart basis or let the qubits undergo depolarizing noise. This allows us to obtain an explicit tradeoff between the amount of noise in storage and the security of the protocol.

In a real implementation using photonic qubits the execution of the protocol by the honest players is imperfect: their quantum operations can be inaccurate or noisy, weak laser pulses instead of single photon sources are used and qubits undergo decoherence in transmission. Note, however, that unlike in QKD, we also want to execute such protocols over very short distances (for example in banking applications) such that the depolarization rate during transmission in free-space is very low. Our practical 1-2 OT-protocol is a small modification of the perfect protocol, so that we can separately deal with erasure errors (i.e. photon loss) and the rate of these errors does not affect the security of the protocol. We then show for this practical protocol how one can derive trade-offs between the amount of storage noise, the amount of noise for the operations performed by the honest players, and the security of the protocol. At the end, we discuss the issue of analyzing fully coherent attacks for our protocol. Indeed, there is a close relation between the 1-2 OT protocol and BB84 quantum key distribution.

Our security analysis can in principle be carried over to obtain a secure identification scheme in the noisy-quantum-storage model analogous to [DFSS07]. This scheme achieves password-based identification and is of particular practical relevance as it can be used for banking applications.

11.1.1 Related work

Precursors of the idea of basing the security of 1-2 OT on storage-noise are already present in [BBCS92b] which laid the foundations for the protocol in [DFR$^+$07], but no rigorous analysis was carried through in that paper. Furthermore, it was pointed out in [Sch07, DFSS08] how the original bounded-quantum-storage analysis applies in the case of noise levels which are so large that the rank of a dishonest player's quantum storage is reduced to $n/4$. In contrast, we are able to give an explicit security tradeoff even for small amounts of noise. We furthermore note that our security proof is not exploiting the noise in the communication channel (which has been done in the classical setting to achieve cryptographic tasks, see e.g. [CK88, Cré97, CMW04]), but is solely based on the fact that the dishonest receiver's quantum storage is noisy. Another technical limitation has been considered in [Sal98] where a bit-commitment scheme was shown secure under the assumption that the dishonest committer can only measure a limited number of qubits coherently. Our analysis differs in that we allow any coherent measurement at the very end. Furthermore, the security analysis of our protocol is considerably simpler and more promising to be extended to cover more general cases.

11.2 Preliminaries

11.2.1 Definitions

We start by introducing some tools, definitions and technical lemmas. To define the security of 1-2 OT, we need to express what it means for a dishonest quantum player not to gain any information. Let ρ_{XE} be a state that is part classical, part quantum, i.e. a cq-state $\rho_{XE} = \sum_{x\in\mathcal{X}} P_X(x)|x\rangle\langle x| \otimes \rho_E^x$. Here, X is a classical random variable distributed over the finite set $\mathcal{X}$ according to distribution P_X. In this Chapter, we will write the *non-uniformity* of X given $\rho_E = \sum_x P_X(x)\rho_E^x$ as

$$d(X|\rho_E) := \frac{1}{2}\left|\left| \mathbb{I}/|\mathcal{X}| \otimes \rho_E - \sum_x P_X(x)|x\rangle\langle x| \otimes \rho_E^x \right|\right|_1 .$$

Intuitively, if $d(X|\rho_E) \leq \varepsilon$ the distribution of X is ε-close to uniform even given ρ_E, i.e., ρ_E gives hardly any information about X. A simple property of the non-uniformity which follows from its definition is that for any cq-state of the form

$\rho_{XED} = \rho_{XE} \otimes \rho_D$, we have

$$d(X|\rho_{ED}) = d(X|\rho_E)\,. \tag{11.1}$$

We prove the security of a randomized version of OT. In such a protocol, Alice does not choose her input strings herself, but instead receives two strings S_0, $S_1 \in \{0,1\}^\ell$ chosen uniformly at random by the protocol. Randomized OT (ROT) can easily be converted into OT: after the ROT protocol is completed, Alice uses her strings S_0, S_1 obtained from ROT as one-time pads to encrypt her original inputs $\hat{S}_0$ and $\hat{S}_1$, i.e. she sends an additional classical message consisting of $\hat{S}_0 \oplus S_0$ and $\hat{S}_1 \oplus S_1$ to Bob. Bob can retrieve the message of his choice by computing $S_C \oplus (\hat{S}_C \oplus S_C) = \hat{S}_C$. He stays completely ignorant about the other message $\hat{S}_{1-C}$ since he is ignorant about S_{1-C}. The security of a quantum protocol implementing ROT is defined in [DFSS05, DFR$^+$07] for a standalone setting. A more involved definition allowing for composability can be found in [WW07]. In the following, we use ρ_B to denote the complete quantum state of Bob's lab at the end of the protocol including any additional classical information he may have received directly from Alice. Similarly, we use $\rho_{CS_0'S_1'A}$ and $\rho_{S_0'S_1'A}$ to denote the c-q states corresponding to the state of Alice's lab at the end of the protocol including her classical information about Bob's choice bit C and outputs S_0' and S_1' as defined below.

11.2.1. Definition. An ε-secure 1-2 ROT$^\ell$ is a protocol between Alice and Bob, where Bob has input $C \in \{0,1\}$, and Alice has no input. For any distribution of C:

- (Correctness) If both parties are honest, Alice gets output $S_0, S_1 \in \{0,1\}^\ell$ and Bob learns $Y = S_C$ except with probability ε.
- (Receiver-security) If Bob is honest and obtains output Y, then for any cheating strategy of Alice resulting in her state ρ_A, there exist random variables S_0' and S_1' such that $\Pr[Y = S_C'] \geq 1 - \varepsilon$ and C is ε-independent of S_0',S_1' and ρ_A, i.e., $D(\rho_{CS_0',S_1'A}, \rho_C \otimes \rho_{S_0',S_1'A}) \leq \varepsilon$.
- (Sender-security) If Alice is honest, then for any cheating strategy of Bob resulting in his state ρ_B, there exists a random variable $C' \in \{0,1\}$ such that $d(S_{1-C'}|S_{C'}C'\rho_B) \leq \varepsilon$.

Note that cheating Bob may of course not choose a C beforehand. Intuitively, our requirement for security states that whatever Bob does, he will be ignorant about at least one of Alice's inputs. This input is determined by his cheating strategy. Our requirement for receiver security states that C is independent of Alice's output, and hence Alice learns nothing about C.

The protocol makes use of two-universal hash functions that are used for privacy amplification similar as in QKD, which we already encountered in Section 10.2.3. For the remainder of this Chapter, we first define

11.2.2. Definition. For a measurement M with POVM elements $\{M_x\}_{x\in\mathcal{X}}$ let $p^M_{y|x} = \mathrm{Tr}(M_y\rho^x_E)$ the probability of outputting guess y given ρ^x_E. Then

$$P_g(X|\rho_E) := \sup_M \sum_x P_X(x) p^M_{x|x}$$

is the maximal average success probability of guessing $x \in \mathcal{X}$ given the reduced state ρ_E of the cq-state ρ_{XE}.

We will employ privacy amplification in the form of the following Lemma, which is an immediate consequence of Lemma 10.2.2 and Theorem 10.2.3 (Theorem 5.5.1 in [Ren05]):

11.2.3. Lemma. *Let $\mathcal{F}$ be a class of two-universal hash functions from $\{0,1\}^n$ to $\{0,1\}^\ell$. Let F be a random variable that is uniformly and independently distributed over $\mathcal{F}$, and let ρ_{XE} be a cq-state. Then,*

$$d(F(X)|F,\rho_E) \leq 2^{\frac{\ell}{2}-1}\sqrt{P_g(X|\rho_E)}\,.$$

If we have an additional k bits of classical information D about X,

$$d(F(X)|F,D,\rho_E) \leq 2^{\frac{\ell+k}{2}-1}\sqrt{P_g(X|\rho_E)}.$$

Furthermore, we will need the following lemma which states that the optimal strategy to guess $X = x \in \{0,1\}^n$ given individual quantum information about the bits of X is to measure each register individually.

11.2.4. Lemma. *Let ρ_{XE} be a cq-state with uniformly distributed $X = x \in \{0,1\}^n$ and $\rho^x_E = \rho^{x_1}_{E_1} \otimes \ldots \otimes \rho^{x_n}_{E_n}$. Then the maximum probability of guessing x given state ρ_E is $P_g(X|\rho_E) = \Pi_{i=1}^n P_g(X_i|\rho_{E_i})$, which can be achieved by measuring each register separately.*

Proof. For simplicity, we will assume that each bit is encoded using the same states $\rho_0 = \rho^0_{E_i}$ and $\rho_1 = \rho^1_{E_i}$. The argument for different encodings is analogous, but harder to read. First of all, note that we can phrase the problem of finding the optimal probability of distinguishing two states as a semi-definite program (SDP)

$$\begin{array}{ll} \text{maximize} & \frac{1}{2}\left(\mathrm{Tr}(M_0\rho_0) + \mathrm{Tr}(M_1\rho_1)\right) \\ \text{subject to} & M_0, M_1 \geq 0 \\ & M_0 + M_1 = \mathbb{I} \end{array}$$

with the dual program

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\mathrm{Tr}(Q) \\ \text{subject to} & Q \geq \rho_0 \\ & Q \geq \rho_1. \end{array}$$

Let p_* and d_* denote the optimal values of the primal and dual respectively. From the weak duality of SDPs, we have $p_* \leq d_*$. Indeed, since $M_0, M_1 = \mathbb{I}/2$ are feasible solutions, we even have strong duality: $p_* = d_*$ [VB96].

Of course, the problem of determining the entire string x from $\hat{\rho}_x := \rho_E^x$ can also be phrased as a SDP:

$$\begin{array}{ll} \text{maximize} & \frac{1}{2^n}\sum_{x\in\{0,1\}^n} \mathrm{Tr}(M_x\hat{\rho}_x) \\ \text{subject to} & \forall x, M_x \geq 0 \\ & \sum_{x\in\{0,1\}^n} M_x = \mathbb{I} \end{array}$$

with the corresponding dual

$$\begin{array}{ll} \text{minimize} & \frac{1}{2^n}\mathrm{Tr}(\hat{Q}) \\ \text{subject to} & \forall x, \hat{Q} \geq \hat{\rho}_x. \end{array}$$

Let $\hat{p}_*$ and $\hat{d}_*$ denote the optimal values of this new primal and dual respectively. Again, $\hat{p}_* = \hat{d}_*$.

Note that when trying to learn the entire string x, we are of course free to measure each register individually and thus $(p_*)^n \leq \hat{p}_*$. We now show that $\hat{d}_* \leq (d_*)^n$ by constructing a dual solution $\hat{Q}$ from the optimal solution to the dual of the single-register case, Q_*: Take $\hat{Q} = Q_*^{\otimes n}$. Since $Q_* \geq \rho_0$ and $Q_* \geq \rho_1$ it follows that $\forall x, Q_*^{\otimes n} \geq \hat{\rho}_x$. Thus $\hat{Q}$ is satisfies the dual constraints. Clearly, $2^{-n}\mathrm{Tr}(\hat{Q}) = (2^{-1}\mathrm{Tr}(Q_*))^n$ and thus we have $\hat{d}_* \leq (d_*)^n$ as promised. But from $(p_*)^n \leq \hat{p}_*$, $\hat{p}_* = \hat{d}_*$, and $p_* = d_*$ we immediately have $\hat{p}_* = (p_*)^n$. □

The next tool we need is an uncertainty relation for noisy channels and measurements. Let $\sigma_{0,+} = |0\rangle\langle 0|$, $\sigma_{1,+} = |1\rangle\langle 1|$, $\sigma_{0,\times} = |+\rangle\langle +|$ and $\sigma_{1,\times} = |-\rangle\langle -|$ denote the BB84-states corresponding to the encoding of a bit $z \in \{0,1\}$ into basis $b \in \{+,\times\}$ (computational resp. Hadamard basis). Let $\sigma_+ = (\sigma_{0,+} + \sigma_{1,+})/2$ and $\sigma_\times = (\sigma_{0,\times} + \sigma_{1,\times})/2$. Consider the state $\mathcal{S}(\sigma_{z,b})$ for some super-operator $\mathcal{S}$. Note that $P_g(X|\mathcal{S}(\sigma_b))$ (see Lemma 11.2.4) denotes the maximal average success probability for guessing a uniformly distributed X when $b = +$ or $b = \times$. An uncertainty relation for such success probabilities can be stated as

$$P_g(X|\mathcal{S}(\sigma_+)) \cdot P_g(X|\mathcal{S}(\sigma_\times)) \leq \Delta(\mathcal{S})^2, \tag{11.2}$$

where Δ is a function from the set of superoperators to the real numbers. For example, when $\mathcal{S}$ is a quantum measurement $\mathcal{M}$ mapping the state $\sigma_{z,b}$ onto purely classical information it can be argued (e.g. by using a purification argument and Corollary 4.15 in [Sch07]) that $\Delta(\mathcal{M}) \equiv \frac{1}{2}(1 + 2^{-1/2})$ which can be achieved

by a measurement in the Breidbart basis, where the Breidbart basis is given by $\{|0\rangle_B, |1\rangle_B\}$ with

$$\begin{aligned} |0\rangle_B &= \cos(\pi/8)|0\rangle + \sin(\pi/8)|1\rangle \\ |1\rangle_B &= \sin(\pi/8)|0\rangle - \cos(\pi/8)|1\rangle \end{aligned}$$

It is clear that for a unitary superoperator U we have $\Delta(U)^2 = 1$ which can be achieved. It is not hard to show that

11.2.5. LEMMA. *The only superoperators $\mathcal{S}\colon \mathcal{H}_{in} \to \mathcal{H}_{out}$ with $\dim(\mathcal{H}_{in}) = 2$ for which $P_g(X|\mathcal{S}(\sigma_+)) \cdot P_g(X|\mathcal{S}(\sigma_\times)) = 1$ are reversible operations.*

Proof. Using Helstrom's formula [Hel67] we have that $P_g(Z|\mathcal{S}(\sigma_b)) = \frac{1}{2}[1 + ||\mathcal{S}(\sigma_{0,b}) - \mathcal{S}(\sigma_{1,b})||_1/2]$ and thus for $\Delta(\mathcal{S}) = 1$ we need that for both $b \in \{\times, +\}$, $||\mathcal{S}(\sigma_{0,b}) - \mathcal{S}(\sigma_{1,b})||_1/2 = 1$. This implies that $\mathcal{S}(\sigma_{0,b})$ and $\mathcal{S}(\sigma_{1,b})$ are states which have support on orthogonal sub-spaces for *both* b. Let $\mathcal{S}(\sigma_{0,+}) = \sum_k p_k |\psi_k\rangle\langle\psi_k|$ and $\mathcal{S}(\sigma_{1,+}) = \sum_k q_k |\psi_k^\perp\rangle\langle\psi_k^\perp|$ where for all k, l $\langle\psi_k^\perp|\psi_l\rangle = 0$. Consider the purification of $\mathcal{S}(\sigma_{i,b})$ using an ancillary system i.e. $|\phi_{i,b}\rangle = U_{\mathcal{S}}|i\rangle_b|0\rangle$. We can write $|\phi_{0,+}\rangle = \sum_k \sqrt{p_k}|\psi_k, k\rangle$ and $|\phi_{1,+}\rangle = \sum_k \sqrt{q_k}|\psi_k^\perp, k\rangle$. Hence $U_{\mathcal{S}}|0\rangle_\times|0\rangle = \frac{1}{\sqrt{2}}(|\phi_{0,+}\rangle + |\phi_{1,+}\rangle)$ and similar for $U_{\mathcal{S}}|1\rangle_\times|0\rangle$. So we can write

$$||\mathcal{S}(\sigma_{0,\times}) - \mathcal{S}(\sigma_{1,\times})||_1 = \left|\left|\sum_k \sqrt{p_k q_k}(|\psi_k\rangle\langle\psi_k^\perp| + |\psi_k^\perp\rangle\langle\psi_k|)\right|\right|_1 \leq 2\sum_k \sqrt{p_k q_k}.$$

For this quantity to be equal to 2 we observe that it is necessary that $p_k = q_k$. Thus we set $p_k = q_k$. We observe that if any of the states $|\psi_k\rangle$ (or $\psi_k^\perp$) are non-orthogonal, i.e. $|\langle\psi_k|\psi_l\rangle| > 0$, then we have $||\sum_k p_k(|\psi_k\rangle\langle\psi_k^\perp| + |\psi_k^\perp\rangle\langle\psi_k|)||_1 < 2$.

Let S_k be the two-dimensional subspace spanned by the orthogonal vectors $|\psi_k\rangle$ and $|\psi_k^\perp\rangle$. By the arguments above, the spaces S_k are mutually orthogonal. We can reverse the super-operator $\mathcal{S}$ by first projecting the output into one of the orthogonal subspaces S_k and then applying a unitary operator U_k that maps $|\psi_k\rangle$ and $|\psi_k^\perp\rangle$ onto the states $|0\rangle$ and $|1\rangle$. □

Finally, we need the following little technical lemma:

11.2.6. LEMMA. *For any $\frac{1}{2} \leq p_i \leq 1$ with $\prod_{i=1}^n p_i \leq p^n$, we have*

$$\frac{1}{2^n}\prod_{i=1}^n (1 + p_i) \leq p^{\log(4/3)n}. \tag{11.3}$$

Proof. With $\lambda := \log(4/3)$, it is easy to verify that $p_i^{-\lambda} + p_i^{1-\lambda} \leq 2$ for $1/2 \leq p_i \leq 1$ and therefore,

$$\frac{1}{2^n}\prod_{i=1}^n (1 + p_i) = \frac{1}{2^n}\prod_{i=1}^n p_i^\lambda \left(p_i^{-\lambda} + p_i^{1-\lambda}\right) \leq \frac{1}{2^n} \cdot p^{\lambda n} \cdot 2^n.$$

□

11.3 Protocol and analysis

11.3.1 Protocol

We use $\in_R$ to denote the uniform choice of an element from a set. We further use $x_{|\mathcal{T}}$ to denote the string $x = x_1, \ldots, x_n$ restricted to the bits indexed by the set $\mathcal{T} \subseteq \{1, \ldots, n\}$. For convenience, we take $\{+, \times\}$ instead of $\{0, 1\}$ as domain of Bob's choice bit C and denote by $\overline{C}$ the bit different from C.

Protocol 2: 1-2 ROT$^{\ell}(C, T)$[DFR$^+$07]

1: Alice picks $X \in_R \{0,1\}^n$ and $\Theta \in_R \{+, \times\}^n$. Let $\mathcal{I}_b = \{i \mid \Theta_i = b\}$ for $b \in \{+, \times\}$. At time $t = 0$, she sends $\sigma_{X_1,\Theta_1} \otimes \ldots \otimes \sigma_{X_n,\Theta_n}$ to Bob.

2: Bob measures all qubits in the basis corresponding to his choice bit $C \in \{+, \times\}$. This yields outcome $X' \in \{0,1\}^n$.

3: Alice picks two hash functions $F_+, F_\times \in_R \mathcal{F}$, where $\mathcal{F}$ is a class of two-universal hash functions. At time $t = T$, she sends $\mathcal{I}_+$,$\mathcal{I}_\times$, F_+,$F_\times$ to Bob. Alice outputs $S_+ = F_+(X_{|\mathcal{I}_+})$ and $S_\times = F_\times(X_{|\mathcal{I}_\times})$ [a].

4: Bob outputs $S_C = F_C(X'_{|\mathcal{I}_C})$.

[a] If $X_{|\mathcal{I}_b}$ is less than n bits long Alice pads the string $X_{|\mathcal{I}_b}$ with 0's to get an n bit-string in order to apply the hash function to n bits.

11.3.2 Analysis

We now show that this protocol is secure according to Definition 11.2.1.

(i) Correctness: It is clear that the protocol is correct. Bob can determine the string $X_{|\mathcal{I}_C}$ (except with negligible probability 2^{-n} the set $\mathcal{I}_C$ is non-empty) and hence obtains S_C.

(ii) Security against dishonest Alice: this holds in the same way as shown in [DFR$^+$07]. As the protocol is non-interactive, Alice never receives any information from Bob at all, and Alice's input strings can be extracted by letting her interact with an unbounded receiver.

(iii) Security against dishonest Bob: Our goal is to show that there exists a $C' \in \{+, \times\}$ such that Bob is completely ignorant about $S_{\overline{C'}}$. In our adversary model, Bob's collective storage cheating strategy can be described by some superoperator

$$\mathcal{S} = \bigotimes_{i=1}^{n} \mathcal{S}_i$$

that is applied on the qubits between the time they arrive at Bob's and the time T that Alice sends the classical information. We define the choice bit C'

as a fixed function of Bob's cheating strategy $\mathcal{S}$. Formally, we set $C' \equiv +$ if $\prod_{i=1}^n P_g(X_i|\mathcal{S}_i(\sigma_+)) \geq \prod_{i=1}^n P_g(X_i|\mathcal{S}_i(\sigma_\times))$ and $C' \equiv \times$ otherwise.

Due to the uncertainty relation for each $\mathcal{S}_i$ (from Eq. (11.2)) it then holds that

$$\prod_i P_g(X_i|\mathcal{S}_i(\sigma_{\overline{C'}})) \leq \prod_i \Delta(\mathcal{S}_i) \leq (\Delta_{\max})^n$$

where $\Delta_{\max} := \max_i \Delta(\mathcal{S}_i)$. This will be used in the proof below.

In the remainder of this section, we show that the non-uniformity

$$\delta_{\text{sec}} := d(S_{\overline{C'}}|S_{C'}C'\rho_B)$$

is negligible in n for a collective attack. Here ρ_B is the complete quantum state of Bob's lab at the end of the protocol including the classical information $\mathcal{I}_+, \mathcal{I}_\times, F_+, F_\times$ he got from Alice and his quantum information $\bigotimes_{i=1}^n \mathcal{S}_i(\sigma_{X_i,\Theta_i})$. Expressing the non-uniformity in terms of the trace-distance allows us to observe that $\delta_{\text{sec}} = 2^{-n} \sum_{\theta\in\{+,\times\}^n} d(S_{\overline{C'}}|\Theta = \theta, S_{C'}C'\rho_B)$. Now, for fixed $\Theta = \theta$, it is clear from the construction that $S_{C'}, C', F_{C'}$ and $\bigotimes_{i\in\mathcal{I}_{C'}} \mathcal{S}_i(\sigma_{X_i,C'})$ are independent of $S_{\overline{C'}} = F_{\overline{C'}}(X_{|\mathcal{I}_{\overline{C'}}})$ and we can use Eq. (11.1). Hence, one can bound the non-uniformity as in Lemma 11.2.3, i.e. by the square-root of the probability of correctly guessing $X_{|\mathcal{I}_{\overline{C'}}}$ given the state $\bigotimes_{i\in\mathcal{I}_{\overline{C'}}} \mathcal{S}_i(\sigma_{X_i,\overline{C'}})$. Lemma 11.2.4 tells us that to guess X, Bob can measure each remaining qubit individually and hence we obtain

$$\begin{aligned}
\delta_{\text{sec}} &\leq 2^{\frac{\ell}{2}-1} \cdot 2^{-n} \sum_{\theta\in\{+,\times\}^n} \sqrt{\prod_{i\in\mathcal{I}_{\overline{C'}}} P_g(X_i|\mathcal{S}_i(\sigma_{\overline{C'}}))} \\
&\leq 2^{\frac{\ell}{2}-1} \sqrt{2^{-n} \sum_{\theta\in\{+,\times\}^n} \prod_{i\in\mathcal{I}_{\overline{C'}}} P_g(X_i|\mathcal{S}_i(\sigma_{\overline{C'}}))} \\
&\leq 2^{\frac{\ell}{2}-1} \sqrt{2^{-n} \prod_{i=1}^n \left(1 + P_g(X_i|\mathcal{S}_i(\sigma_{\overline{C'}}))\right)},
\end{aligned}$$

where we used the concavity of the square-root function in the last inequality. Lemma 11.2.6 together with the bound $\prod_i P_g(X_i|\mathcal{S}_i(\sigma_{\overline{C'}})) \leq (\Delta_{\max})^n$ lets us conclude that

$$\delta_{\text{sec}} \leq 2^{\frac{\ell}{2}-1} \cdot (\Delta_{\max})^{\frac{\log(4/3)}{2} n}.$$

Lemma 11.2.5 shows that for essentially any noisy superoperator $\Delta(\mathcal{S}) < 1$. This shows that for any collective attacks there exists an n which yields arbitrarily high security.

11.4 Practical oblivious transfer

In this section, we prove the security of a ROT protocol that is robust against noise for the honest parties. Our protocol is thereby a small modification of the

protocol considered in [Sch07]. Note that for our analysis, we have to assume a worst-case scenario where a dishonest receiver Bob has access to a perfect noise-free quantum channel and only experiences noise during storage.

First, we consider erasure noise (in practice corresponding to photon loss) during preparation, transmission and measurement of the qubits by the honest parties. Let $1 - p_{\text{erase}}$ be the total constant probability for an honest Bob to measure and detect a photon in the $\{+, \times\}$ basis given that an honest Alice prepares a qubit (or weak laser pulse) in her lab and sends it to him. The probability p_{erase} is determined among others by the mean photon number in the pulse, the loss on the channel and the quantum efficiency of the detector. In our protocol we assume that the (honest) erasure rate p_{erase} is *independent* of whether qubits were encoded or measured in the $+$- or $\times$-basis. This assumption is necessary to guarantee the correctness and the security against a cheating *Alice* only. Fortunately, this assumption is well matched with physical capabilities.

Any other noise source during preparation, transmission and measurement can be characterized as an effective classical noisy channel resulting in the output bits X' that Bob obtains at Step 3 of Protocol 11.4. For simplicity, we model this compound noise source as a classical binary symmetric channel acting independently on each bit of X. Typical noise sources for polarization-encoded qubits are depolarization during transmission, dark counts in Bob's detector and misaligned polarizing beam-splitters. Let the effective bit-error probability of this binary symmetric channel be $p_{\text{error}} < 1/2$.

Before engaging in the actual protocol, Alice and Bob agree on the system parameters p_{erase} and p_{error} similarly to Step 1 of the protocol in [BBCS92b]. Furthermore, they agree on a family $\{C_n\}$ of linear error correcting codes of length n capable of efficiently correcting $n \cdot p_{\text{error}}$ errors. For any string $x \in \{0,1\}^n$, error correction is done by sending the syndrome information $syn(x)$ to Bob from which he can correctly recover x if he holds an output $x' \in \{0,1\}^n$ obtained by flipping each bit of x independently with probability p_{error}. It is known that for large enough n, the code C_n can be chosen such that its rate is arbitrarily close to $1 - h(p_{\text{error}})$ and the syndrome length (the number of parity check bits) are asymptotically bounded by $|syn(x)| < h(p_{\text{error}})n$ [Cré97], where $h(p_{\text{error}})$ is the binary Shannon entropy. We assume the players have synchronized clocks. In each time slot, Alice sends one qubit (laser pulse) to Bob.

Protocol 3: Noise-Protected Photonic 1-2 ROT$^{\ell}(C,T)$

1: Alice picks $X \in_R \{0,1\}^n$ and $\Theta \in_R \{+,\times\}^n$.

2: For $i = 1,\ldots,n$: In time slot $t = i$, Alice sends σ_{X_i,Θ_i} as a phase- or polarization-encoded weak pulse of light to Bob.

3: In each time slot, Bob measures the incoming qubit in the basis corresponding to his choice bit $C \in \{+,\times\}$ and records whether he detects a photon or not. He obtains some bit-string $X' \in \{0,1\}^m$ with $m \leq n$.

4: Bob reports back to Alice in which time slots he received a qubit. Alice restricts herself to the set of $m \leq n$ bits that Bob did not report as missing. Let this set of qubits be S_{remain} with $|S_{\text{remain}}| = m$.

5: Let $\mathcal{I}_b = \{i \in S_{\text{remain}} \mid \Theta_i = b\}$ for $b \in \{+,\times\}$ and let $m_b = |\mathcal{I}_b|$. Alice aborts the protocol if either m_+ or $m_\times \leq (1-p_{\text{erase}})n/2 - O(\sqrt{n})$. If this is not the case, Alice picks two hash functions $F_+, F_\times \in_R \mathcal{F}$, where $\mathcal{F}$ is a set of two-universal hash functions. At time $t = n+T$, Alice sends $\mathcal{I}_+$,$\mathcal{I}_\times$, F_+,$F_\times$, and the syndromes $syn(X_{|\mathcal{I}_+})$ and $syn(X_{|\mathcal{I}_\times})$ according to codes of appropriate length m_b to Bob. Alice outputs $S_+ = F_+(X_{|\mathcal{I}_+})$ and $S_\times = F_\times(X_{|\mathcal{I}_\times})$.

6: Bob uses $syn(X_{|\mathcal{I}_C})$ to correct the errors on his output $X'_{|\mathcal{I}_C}$. He obtains the corrected bit-string X_{cor} and outputs $S'_C = F_C(X_{\text{cor}})$.

Let us consider the security and correctness of this modified protocol.
(i) Correctness: By assumption, p_{erase} is independent of the basis in which Alice sent the qubits. Thus, S_{remain} is with high probability a random subset of the transmitted qubits of of size $m \approx (1-p_{\text{erase}})n \pm O(\sqrt{n})$ qubits independent of the value of bases Θ. This implies that in Step 5 the protocol is aborted with a probability exponentially small in m, and hence in n. The codes are chosen such that Bob can decode except with negligible probability. These facts imply that if both parties are honest the protocol is correct (i.e. $S_C = S'_C$) with exponentially small probability of error.
(ii) Security against dishonest Alice: Even though in this scenario Bob *does* communicate to Alice, the information stating which qubits were erased is by assumption independent of the basis in which he measured and thus of his choice bit C. Hence Alice does not learn anything about his choice bit C. Her input strings can be extracted as in Protocol 1.
(iii) Security against dishonest Bob: Our analysis is essentially identical to our analysis for Protocol 1 where we address the error-correcting properties as in [Sch07]. First of all, we note that Bob can always make Alice abort the protocol by reporting back an insufficient number of received qubits. If this is not the case, then we define C' as in the analysis of Protocol 1 and we need to bound the non-uniformity

δ_{sec} as before. Let us for simplicity assume that $m_b = m/2$ (this is true with high probability, up to a factor of $O(\sqrt{n})$ which becomes negligible for large n) with $m \approx (1 - p_{\text{erase}})n$. We perform the same analysis, where we restrict ourselves to the set of remaining qubits. We first follow through the same steps simplifying the non-uniformity using that the total attack superoperator $\mathcal{S}$ is a product of superoperators. Then we use the bound in Lemma 11.2.3 for each $\theta \in \{+, \times\}^n$ where we now have to condition on the additional information $syn(X_{|\mathcal{I}_{\overline{C'}}})$ which is $mh(p_{\text{error}})/2$ bits long. Note that Bob does not gain any information when Alice aborts the protocol, since her decision to abort is a function of the bits Bob reported as being erased and he can thus compute Alice's decision himself. Using the second part of Lemma 11.2.3 and following identical steps in the remainder of the proof implies

$$\delta_{\text{sec}} \leq 2^{\frac{\ell}{2} - 1 + h(p_{\text{error}})\frac{m}{4}} (\Delta_{\max})^{\frac{\log(4/3)}{2} m} . \tag{11.4}$$

From this expression it is clear that the security depends crucially on the value of $\Delta_{\max}$ versus the binary entropy $h(p_{\text{error}})$. The trade-off in our bound is not extremely favorable for security as we will see.

11.5 Example: depolarizing noise

Let us now consider the security in an explicit example, where Bob's storage is affected by depolarizing noise, and he is not able to encode the incoming qubits into a higher-dimensional system such as an error correcting code.

Again, we first address the simpler setting where the honest players experience no noise themselves. In order to explicitly bound $\Delta(\mathcal{S}_i)$ we should allow for intermediate strategies of Bob in which he partially measures the incoming qubits leaving some quantum information undergoing depolarizing noise. To model this noise we let $\mathcal{S}_i = \mathcal{N} \circ \mathcal{P}_i$, where $\mathcal{P}_i$ is any noiseless quantum operation of Bob's choosing from one qubit to one qubit that generates some classical output. For example, $\mathcal{P}_i$ could be a partial measurement providing Bob with some classical information and a slightly disturbed quantum state, or just a unitary operation. Let

$$\mathcal{N}(\rho) := r\rho + (1 - r)\frac{\mathbb{I}}{2}$$

be the fixed depolarizing 'quantum storage' channel that Bob cannot influence (see Figure 11.1).

To determine δ_{sec}, we have to find an uncertainty relation similar to Eq. (11.2) by optimizing over all possible partial measurements $\mathcal{P}_i$,

$$\Delta_{\max}^2 = \max_{\mathcal{S}_i} \Delta(\mathcal{S}_i)^2 = \max_{\mathcal{P}_i} P_g(X|\mathcal{S}_i(\sigma_+)) \cdot P_g(X|\mathcal{S}_i(\sigma_\times)). \tag{11.5}$$

We solve this problem for depolarizing noise using the symmetries inherent in our problem. In Section 11.5.1 we prove the following.

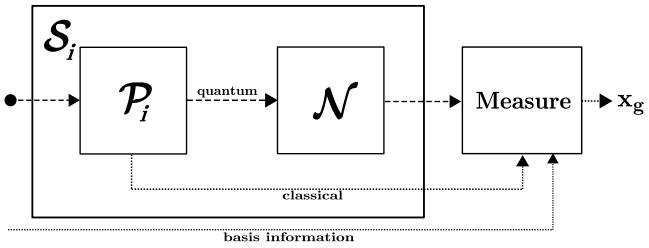


Figure 11.1: Bob performs a partial measurement $\mathcal{P}_i$, followed by noise $\mathcal{N}$, and outputs a guess bit x_g depending on his classical measurement outcome, the remaining quantum state, and the additional basis information.

11.5.1. THEOREM. *Let $\mathcal{N}$ be the depolarizing channel and let $\max_{\mathcal{S}_i} \Delta(\mathcal{S}_i)$ be defined as above. Then*

$$\max_{\mathcal{S}_i} \Delta(\mathcal{S}_i) = \begin{cases} \frac{1+r}{2} & \text{for } r \geq \frac{1}{\sqrt{2}} \\ \frac{1}{2} + \frac{1}{2\sqrt{2}} & \text{for } r < \frac{1}{\sqrt{2}} \end{cases}$$

Our result shows that for $r < 1/\sqrt{2}$ a direct measurement $\mathcal{M}$ in the Breidbart basis is the best attack Bob can perform. For this measurement, we have $\Delta(\mathcal{M}) = 1/2 + 1/(2\sqrt{2})$. If the depolarizing noise is low ($r \geq 1/\sqrt{2}$), then our result states that the best strategy for Bob is to simply store the qubit as is.

11.5.1 Optimal cheating strategy

We now prove Theorem 11.5.1 in a series of steps. Recall, that to determine the security bound, we have to find an uncertainty relation similar to Eq. (11.2) by optimizing over all possible partial measurements $\mathcal{P}$ and final measurements $\mathcal{M}$ as in Eq. 11.5. To improve readability, we will drop the index i and use $\mathcal{S}$ in place of $\mathcal{S}_i$ to denote the cheating operation acting on a single qubit. For our analysis, it will be convenient to think of $\mathcal{P}$ as a partial measurement of the incoming qubit. Note that this corresponds to letting Bob perform an arbitrary CPTP map from the space of the incoming qubit to the space carrying the stored qubit. It will furthermore be convenient to consider the maximizing the sum instead:

$$\Gamma(\mathcal{S}) = \max_{\mathcal{M},\mathcal{P}} P_g(X|\mathcal{S}(\sigma_+)) + P_g(X|\mathcal{S}(\sigma_\times)).$$

This immediately gives us the bound $\Delta(\mathcal{S}) \leq \Gamma(\mathcal{S})/2$. In the following, we will use the shorthand

$$\begin{aligned} p_+ &:= P_g(X|\mathcal{S}(\sigma_+)) \\ p_\times &:= P_g(X|\mathcal{S}(\sigma_\times)) \end{aligned}$$

for the probabilities that Bob correctly decodes the bit after Alice has announced the basis information.

Any measurement Bob may perform can be characterized by a set of measurement operators $\{F_k\}$ such that $\sum_k F_k^\dagger F_k = \mathbb{I}$. The probability that Bob succeeds in decoding the bit after the announcement of the basis is simply the average over the probability that he correctly decodes the bit, conditioned on the fact that he obtained outcome k. I.e., for $b \in \{+, \times\}$

$$\begin{aligned} p_b &= \sum_k p_{k|b} \left(\frac{1}{2} + \frac{1}{4} || p_{0|kb} N(\tilde{\sigma}^k_{0,b}) - p_{1|kb} N(\tilde{\sigma}^k_{1,b}) ||_1 \right) \\ &= \frac{1}{2} + \frac{1}{4} \sum_k p_{k|b} || r(p_{0|kb} \tilde{\sigma}^k_{0,b} - p_{1|kb} \tilde{\sigma}^k_{1,b}) + (1-r)(p_{0|kb} - p_{1|kb}) \mathbb{I}/2 ||_1, \end{aligned}$$

where

$$p_{k|b} = \mathrm{Tr}\left(F_k \frac{\sigma_{0,b} + \sigma_{1,b}}{2} F_k^\dagger \right) = \frac{1}{2} \mathrm{Tr}(F_k F_k^\dagger)$$

is the probability of obtaining measurement outcome k conditioned on the fact that the basis was b (and we even see from the above that it is actually independent of b), $\tilde{\sigma}^k_{0,b} = F_k \sigma_{0,b} F_k^\dagger / p_{k|0b}$ is the post-measurement state for outcome k, and $p_{0|kb}$ is the probability that we are given this state. Definitions are analogous for the bit 1.

We now show that Bob's optimal strategy is to measure in the Breidbart basis for $r < 1/\sqrt{2}$, and to simply store the qubit for $r \geq 1/\sqrt{2}$. This then immediately allows us to evaluate $\Delta_{\max}$. To prove our result, we proceed in three steps: First, we will simplify our problem considerably until we are left with a single Hermitian measurement operator over which we need to maximize. Second, we show that the optimal measurement operator is diagonal in the Breidbart basis. And finally, we show that depending on the amount of noise, this measurement operator is either proportional to the identity, or proportional to a rank one projector. Our individual claims are indeed very intuitive.

For any measurement $M = \{F_k\}$, let $B(M) = p_+^M + p_\times^M$ for the measurement M, where p_+^M and $p_\times^M$ are the success probabilities similar to Eq. (11.6), but restricted to using the measurement M. First of all, note that we can easily combine two measurements. Intuitively, the following statement says that if we choose one measurement with probability α, and the other with probability β our average success probability will be the average of the success probabilities obtained via the individual measurements:

4. Claim. *Let $M_1 = \{F_k^1\}$ and $M_2 = \{F_k^2\}$ be two measurements. Then $B(\alpha M_1 + \beta M_2) = \alpha B(M_1) + \beta B(M_2)$, where where $\alpha M_1 + \beta M_2 = \{\sqrt{\alpha} F_k^1\} \cup \{\sqrt{\beta} F_k^2\}$ for $\alpha, \beta \geq 0$ and $\alpha + \beta = 1$.*

Proof. Let $F = \{F_k\}_{k=1}^f$ and $G = \{G_k\}_{k=1}^g$ be measurements, $0 \leq \alpha \leq 1$ and $M := \{\sqrt{\alpha} F_k\}_{k=1}^f \cup \{\sqrt{1-\alpha} G_k\}_{k=f+1}^{f+g}$ be the measurement F with probability

α and measurement G with probability $1-\alpha$. We denote by $p^F_\cdot, p^G_\cdot, p^M_\cdot$ the probabilities corresponding to measurements F, G, M respectively. Observe that for $1 \leq k \leq f$, $p^M_{k|b} = \frac{1}{2}\mathrm{Tr}(\alpha F_k F_k^\dagger) = \alpha p^F_{k|b}$ and analogously for $f+1 \leq k \leq f+g$, we have $p^M_{k|b} = (1-\alpha)p^G_{k|b}$. We observe furthermore that for $1 \leq k \leq f$ and $x \in \{0,1\}$, α cancels out by the normalization, $\tilde{\sigma}^{k,M}_{x,b} = \frac{\alpha F_k \sigma_{x,b} F_k^\dagger}{p^M_{k|xb}} = \frac{F_k \sigma_{x,b} F_k^\dagger}{p^F_{k|xb}} = \tilde{\sigma}^{k,F}_{x,b}$ and similarly for $f+1 \leq k \leq f+g$. Finally, we can convince ourselves that $p^M_{x|kb} = p^F_{x|kb} = p^G_{x|(k-f)b}$, as the probability to be given state $\tilde{\sigma}^k_{0,b}$ is the same when the measurement outcome and the basis is fixed. Putting everything together, we obtain

$$\begin{aligned}
p^M_b &= \sum_{k=1}^{f+g} p^M_{k|b}\left(\frac{1}{2} + \frac{1}{4}||p^M_{0|kb}N(\tilde{\sigma}^{k,M}_{0,b}) - p^M_{1|kb}N(\tilde{\sigma}^{k,M}_{1,b})||_1\right) \\
&= \sum_{k=1}^{f} \alpha p^F_{k|b}\left(\frac{1}{2} + \frac{1}{4}||p^F_{0|kb}N(\tilde{\sigma}^{k,F}_{0,b}) - p^F_{1|kb}N(\tilde{\sigma}^{k,F}_{1,b})||_1\right) \\
&\quad + \sum_{k=f+1}^{g} (1-\alpha)p^G_{k|b}\left(\frac{1}{2} + \frac{1}{4}||p^G_{0|kb}N(\tilde{\sigma}^{k,G}_{0,b}) - p^G_{1|kb}N(\tilde{\sigma}^{k,G}_{1,b})||_1\right) \\
&= \alpha p^F_b + (1-\alpha)p^G_b\,.
\end{aligned}$$

□

We can now make a series of observations.

5. Claim. *Let $M = \{F_k\}$ and $G = \{\mathbb{I}, X, Z, XZ\}$. Then for all $g \in G$ we have $B(M) = B(gMg^\dagger)$.*

Proof. This claim follows immediately from that fact that for the trace norm we have $||UAU^\dagger||_1 = ||A||_1$ for all unitaries U, and by noting that for all $g \in G$, g can at most exchange the roles of 0 and 1. I.e., we perform a bit flip before the measurement which we can correct for afterwards by applying classical post-processing: we have for all $g \in G$ that

$$\begin{aligned}
&p_{k|b}\left|\left|p_{0|kb}N\left(\frac{F_k g\sigma_{0,b}g^\dagger F_k^\dagger}{p_{k|0b}}\right) - p_{1|kb}N\left(\frac{F_k g\sigma_{1,b}g^\dagger F_k^\dagger}{p_{k|1b}}\right)\right|\right|_1 \\
&= p_{k'|b}\left|\left|p_{0|kb}N\left(\frac{F_k \sigma_{0,b} F_k^\dagger}{p_{k|0b}}\right) - p_{1|kb}N\left(\frac{F_k \sigma_{1,b} F_k^\dagger}{p_{k|1b}}\right)\right|\right|_1 .
\end{aligned}$$

□

It also follows that

11.5.2. Corollary. *For all k we have for all $b \in \{+, \times\}$ and $g \in G$ that*

$$\left\| p_{0|kb} N\left(\frac{F_k \sigma_{0,b} F_k^\dagger}{p_{k|0b}}\right) - p_{1|kb} N\left(\frac{F_k \sigma_{1,b} F_k^\dagger}{p_{k|1b}}\right) \right\|_1$$
$$= \left\| p_{0|kb} N\left(\frac{F_k g \sigma_{0,b} g^\dagger F_k^\dagger}{p_{k|0b}}\right) - p_{1|kb} N\left(\frac{F_k g \sigma_{1,b} g^\dagger F_k^\dagger}{p_{k|1b}}\right) \right\|_1 .$$

Proof. This follows from the proof of Claim 5. □

6. Claim. *Let $G = \{\mathbb{I}, X, Z, XZ\}$. There exists a measurement operator F such that the maximum of $B(M)$ over all measurements M is achieved by a measurement proportional to $\{gFg^\dagger \mid g \in G\}$.*

Proof. Let $M = \{F_k\}$ be a measurement. Let $K = |M|$ be the number of measurement operators. Clearly, $\hat{M} = \{\hat{F}_{g,k}\}$ with

$$\hat{F}_{g,k} = \frac{1}{4} g F_k g^\dagger,$$

is also a quantum measurement since $\sum_{g,k} \hat{F}_{g,k}^\dagger \hat{F}_{g,k} = \mathbb{I}$. It follows from Claims 4 and 5 that $B(M) = B(\hat{M})$. Define operators

$$N_{g,k} = \frac{1}{\sqrt{2\mathrm{Tr}(F_k^\dagger F_k)}} g F_k g^\dagger.$$

Note that

$$\sum_{g \in G} N_{g,k} = \frac{1}{\sqrt{2\mathrm{Tr}(F_k^\dagger F_k)}} \sum_{u,v \in \{0,1\}} X^u Z^v F_k^\dagger F_k Z^v X^u = \mathbb{I}.$$

(see for example Hayashi [Hay06]). Hence $M_k = \{N_{g,k}\}$ is a valid quantum measurement. Now, note that $\hat{M}$ can be obtained from $M_1, \ldots, M_K$ by averaging. Hence, by Claim 4 we have

$$B(M) = B(\hat{M}) \leq \max_k B(M_k).$$

Let M^* be the optimal measurement. Clearly, $m = B(M^*) \leq \max_k B(M_k^*) \leq m$ by the above and Corollary 11.5.2 from which our claim follows. □

Note that Claim 6 also gives us that we have at most 4 measurement operators. Wlog, we will take the measurement outcomes to be labeled $1, 2, 3, 4$.

Finally, we note that we can restrict ourselves to optimizing over positive-semidefinite (and hence Hermitian) matrices only.

7. Claim. *Let F be a measurement operator, and let*

$$g(F) := 1 + \sum_{b,k} p_{k|b} \left|\left|p_{0|b} N(\tilde{\sigma_{0,b}}) - p_{1|b} N(\tilde{\sigma_{1,b}})\right|\right|_1$$

with $\tilde{\sigma_{0,b}} = F\sigma_{0,b}F^\dagger/\mathrm{Tr}(F\sigma_{0,b}F^\dagger)$ and $\tilde{\sigma_{1,b}} = F\sigma_{1,b}F^\dagger/\mathrm{Tr}(F\sigma_{1,b}F^\dagger)$. Then there exists a Hermitian operator $\hat{F}$, such that $g(F) = g(\hat{F})$.

Proof. Let $F^\dagger = \hat{F}U$ be the polar decomposition of $F^\dagger$, where $\hat{F}$ is positive semidefinite and U is unitary [HJ85, Corollary 7.3.3]. Evidently, since the trace is cyclic, all probabilities remain the same. It follows immediately from the definition of the trace norm that $||UAU^\dagger||_1 = ||A||_1$ for all unitaries U, which completes our proof. □

To summarize, our optimization problem can now be simplified to

$$\begin{aligned}
&\max_M B(M) = \max_M p_+^M + p_\times^M \leq \\
&\max_F 1 + \sum_{b,k} p_{k|b} \left|\left|p_{0|b} N(\tilde{\sigma_{0,b}}) - p_{1|b} N(\tilde{\sigma_{1,b}})\right|\right|_1 \\
&= 1 + 2\sum_b \left|\left| r(F(\sigma_{0,b} - \sigma_{1,b})F) \qquad + (1-r)\mathrm{Tr}(F(\sigma_{0,b} - \sigma_{1,b})F)\frac{\mathbb{I}}{2} \right|\right|_1
\end{aligned}$$

where the maximization is now taken over a single operator F, and we have used the fact that we can write $p_{0|kb} = p_{k|0b}/(2p_{k|b})$ and we have 4 measurement operators.

F is diagonal in the Breidbart basis

Now that we have simplified our problem already considerably, we are ready to perform the actual optimization. Since we are in dimension $d = 2$ and F is Hermitian, we may express F as

$$F = \alpha|\phi\rangle\langle\phi| + \beta|\phi^\perp\rangle\langle\phi^\perp|,$$

for some state $|\phi\rangle$ and real numbers α, β. We first of all note that from $\sum_k F_k F_k^\dagger = \mathbb{I}$, we obtain that

$$\begin{aligned}
\mathrm{Tr}\left(\sum_k F_k F_k^\dagger\right) = \sum_k \mathrm{Tr}(F_k F_k) = \\
\sum_{g\in\{\mathbb{I},X,Z,XZ\}} \mathrm{Tr}(gFgg^\dagger Fg^\dagger) = 4\mathrm{Tr}(FF) = \mathrm{Tr}(\mathbb{I}) = 2,
\end{aligned}$$

and hence $\mathrm{Tr}(FF) = \alpha^2 + \beta^2 = 1/2$. Furthermore using that $|\phi\rangle\langle\phi| + |\phi^\perp\rangle\langle\phi^\perp| = \mathbb{I}$ we then have

$$F = \beta\mathbb{I} + (\alpha - \beta)|\phi\rangle\langle\phi|, \tag{11.6}$$

with $\beta = \sqrt{1-\alpha^2}$. Our first goal is now to show that $|\phi\rangle$ is a Breidbart vector (or the bit-flipped version thereof). To this end, we first formalize our intuition that we may take $|\phi\rangle$ to lie in the XZ plane of the Bloch sphere only. Since we are only interested in the trace-distance term of $B(M)$, we restrict ourselves to considering

$$C(F) := \sum_b \quad \left|\left| r(F(\sigma_{0,b} - \sigma_{1,b})F) + (1-r)\mathrm{Tr}(F(\sigma_{0,b} - \sigma_{1,b})F)\frac{\mathbb{I}}{2} \right|\right|_1 .$$

8. Claim. *Let F be the operator that maximizes $C(F)$, and write F as in Eq.(11.6). Then $|\phi\rangle$ lies in the XZ plane of the Bloch sphere. (i.e. $\mathrm{Tr}(FY) = 0$).*

Proof. We first parametrize the state in terms of its Bloch vector:

$$|\phi\rangle\langle\phi| = \frac{\mathbb{I} + xX + yY + zZ}{2}.$$

Since $|\phi\rangle$ is pure we can write $y = \sqrt{1-x^2-z^2}$. Hence, we can express F as

$$F = \frac{1}{2}\left((\alpha+\beta)\mathbb{I} + (\alpha-\beta)(xX + yY + zZ)\right).$$

Noting that $\sigma_{0,+} - \sigma_{1,+} = Z$ and $\sigma_{0,\times} - \sigma_{1,\times} = X$ we can compute for the computational basis

$$\begin{aligned} P &:= r(FZF) + (1-r)\mathrm{Tr}(FZF)\frac{\mathbb{I}}{2} \\ &= \frac{1}{2}\left(\left(2\alpha^2 - \frac{1}{2}\right) z\mathbb{I} + r\left((\alpha-\beta)^2 xzX \;+\; (\alpha-\beta)^2 yzY + \left((\alpha-\beta)^2 z^2 + 2\alpha\beta\right) Z\right)\right), \end{aligned}$$

and for the Hadamard basis:

$$\begin{aligned} T &:= r(FXF) + (1-r)\mathrm{Tr}(FXF)\frac{\mathbb{I}}{2} \\ &= \frac{1}{2}\left(\left(2\alpha^2 - \frac{1}{2}\right) x\mathbb{I} + r\left(\left((\alpha-\beta)^2 x^2 + 2\alpha\beta\right) X\right)\right. \\ &+ \left.(\alpha-\beta)^2 xyY + (\alpha-\beta)^2 xzZ\right) \end{aligned}$$

Note that $||P||_1 = \sum_j |\lambda_j(P)|$, where λ_j is the j-th eigenvalue of P. A lengthy computation (using Mathematica), and plugging in $\beta = \sqrt{1/2 - \alpha^2}$ and $y = \sqrt{1-x^2-z^2}$ shows that we have

$$\begin{aligned} \lambda_1(P) &= \frac{1}{4}\left(\left(4\alpha^2 - 1\right) z - r\sqrt{z^2 + 8\alpha^2(2\alpha^2-1)(z^2-1)}\right) \\ \lambda_2(P) &= \frac{1}{4}\left(\left(4\alpha^2 - 1\right) z + r\sqrt{z^2 + 8\alpha^2(2\alpha^2-1)(z^2-1)}\right) \end{aligned}$$

Similarly, we obtain for the Hadamard basis that

$$\begin{aligned}\lambda_1(T) &= \frac{1}{4}\left(\left(4\alpha^2-1\right)x - r\sqrt{x^2+8\alpha^2(2\alpha^2-1)(x^2-1)}\right)\\ \lambda_2(T) &= \frac{1}{4}\left(\left(4\alpha^2-1\right)x + r\sqrt{x^2+8\alpha^2(2\alpha^2-1)(x^2-1)}\right)\end{aligned}$$

We define

$$\begin{aligned}f(\alpha,x) &:= \left(\alpha^2-\frac{1}{4}\right)x\\ g(\alpha,x) &:= \frac{1}{4}\sqrt{x^2+8\alpha^2(2\alpha^2-1)(x^2-1)}.\\ h(\alpha,x,r) &:= |f(\alpha,x)+rg(\alpha,x)|+|f(\alpha,x)-rg(\alpha,x)|\end{aligned}$$

Note that our optimization problem now takes the form

$$\begin{array}{ll}\text{maximize} & h(\alpha,x,r)+h(\alpha,z,r)\\ \text{subject to} & x^2+z^2\leq 1\\ & 0\leq x\leq 1\\ & 0\leq z\leq 1,\end{array}$$

where we can introduce the last two inequality constraints without loss of generality, since the remaining three measurement operators will be given by XFX, ZFZ, and $XZFZX$.

To show that we can let $y=0$ for the optimal solution, we have to show that for all α and all r, the function $h(\alpha,x,r)$ is increasing on the interval $0\leq x\leq 1$ (and indeed Mathematica will convince you in an instant that this is the case). Our analysis is is further complicated by the absolute values. We therefore first consider

$$h(\alpha,x,r)^2 = 2(f(\alpha,x)^2+r^2g(\alpha,x)^2+|f(\alpha,x)^2-r^2g(\alpha,x)^2|,$$

where we have used the fact that f and g are real valued functions. In principle, we can now analyze $h_+(\alpha,x,r)^2 = 2(f(\alpha,x)^2+r^2g(\alpha,x)^2+f(\alpha,x)^2-r^2g(\alpha,x)^2$ and $h_-(\alpha,x,r)^2 = 2(f(\alpha,x)^2+r^2g(\alpha,x)^2-f(\alpha,x)^2+r^2g(\alpha,x)^2$ separately on their respective domains. By rewriting, we obtain

$$h_+(\alpha,x,r)^2 = \frac{1}{4}r^2(x^2+8\alpha^2(2\alpha^2-1)(x^2-1)),$$

and

$$h_-(\alpha,x,r)^2 = 4\left(\alpha^2-\frac{1}{4}\right)^2x^2.$$

Luckily, the first derivatives of h_+ and h_- turns out to be positive everywhere for our choice of parameters $0\leq\alpha\leq 1/\sqrt{2}$, and $0\leq r,z\leq 1$. Hence, by further

inspection at the transitional points we can conclude that h is an increasing function of x. But this means that to maximize our target expression, we must choose x and z as large as possible. Hence, choosing $y = 0$ is the best choice and our claim follows. □

We can now immediately extend this analysis to find

9. Claim. *Let F be the operator that maximizes $C(F)$, and write F as in Eq.(11.6). Then*

$$|\phi\rangle = g(\cos(\pi/8)|0\rangle + \sin(\pi/8)|1\rangle),$$

for some $g \in \{\mathbb{I}, X, Z, XZ\}$.

Proof. Extending our analysis from the previous proof, we can compute the second derivative of both functions. It turns out that also the second derivatives are positive, and hence h is convex in x. By Claim 8, we can rewrite our optimization problem as

$$\begin{array}{ll} \text{maximize} & h(\alpha, x, r) + h(\alpha, z, r) \\ \text{subject to} & x^2 + z^2 = 1 \\ & 0 \leq x \leq 1 \\ & 0 \leq z \leq 1 \end{array}$$

It now follows from the fact that h is convex in x and the constraint $x^2 + z^2 = 1$ (by computing the Lagrangian of the above optimization problem), that for the optimal solution we must have $x = z$, and our claim follows. □

Optimality of the trivial strategies

Now that we have shown that F is in fact diagonal in the Breidbart basis (or the bit flipped version thereof) we have only a single parameter left in our optimization problem. We must now optimize over all operators F of the form

$$F = \alpha|\phi\rangle\langle\phi| + \sqrt{1/2 - \alpha^2}|\phi^\perp\rangle\langle\phi^\perp|,$$

where we may take $|\phi\rangle$ to be $|0\rangle_B$ or $|1\rangle_B$. Our aim is now to show that either F is the identity, or $F = |\phi\rangle\langle\phi|$ depending on the value of r.

10. Claim. *Let F be the operator that maximizes $C(F)$. Then $F = c\mathbb{I}$ (for some $c \in \mathbb{R}$) for $r \geq 1/\sqrt{2}$, and $F = |\phi\rangle\langle\phi|$ for $r < 1/\sqrt{2}$, where*

$$|\phi\rangle = g(\cos(\pi/8)|0\rangle + \sin(\pi/8)|1\rangle),$$

for some $g \in \{\mathbb{I}, X, Z, XZ\}$.

Proof. We can now plug in $x = z = 1/\sqrt{2}$ in the expressions for the eigenvalues in our previous proof. Ignoring the constant positive factors which do not contribute to our argument, we can then write

$$\begin{aligned} \lambda_1(P) &= \left(4\alpha^2 - 1\right) - r\sqrt{1 - 16\alpha^4 + 8\alpha^2}, \\ \lambda_2(P) &= \left(4\alpha^2 - 1\right) + r\sqrt{1 - 16\alpha^4 + 8\alpha^2}. \end{aligned}$$

And similarly for the Hadamard basis. We again define functions

$$\begin{aligned} f(\alpha) &:= \left(4\alpha^2 - 1\right) \\ g(\alpha) &:= \sqrt{1 - 16\alpha^4 + 8\alpha^2} \\ h(\alpha, r) &:= |f(\alpha, x) + rg(\alpha, x)| + |f(\alpha, x) - rg(\alpha, x)| \end{aligned}$$

Note that our optimization problem now takes the form

$$\begin{array}{ll} \text{maximize} & 2h(\alpha, r) \\ \text{subject to} & 0 \leq \alpha \leq \frac{1}{\sqrt{2}} \end{array}$$

Since we are maximizing, we might as well consider the square of our target function and ignore the leading constant as it is irrelevant for our argument.

$$h(\alpha, r)^2 = 2(f(\alpha)^2 + r^2 g(\alpha)^2 + |f(\alpha)^2 - r^2 g(\alpha)^2|,$$

To deal with the absolute value, we now perform a case analysis similar to the one above. Computing the zeros crossings of the function $f(\alpha)^2 - r^2 g(\alpha)^2$, we analyze each interval separately. Computing the first and second derivatives on the intervals we find that $h(\alpha, r)^2$ has exactly two peaks: The first at $\alpha = 0$, and the second at $\alpha = 1/2$. We have that $h(0, r)^2 = 2$ for all r, and $h(1/2, r)^2 = 4r^2$. Hence, we immediately see that the maximum is located at $\alpha = 0$ for $r \leq 1/\sqrt{2}$, and at $\alpha = 1/2$ for $r \geq 1/\sqrt{2}$. □

Theorem 11.5.1 now follows directly from Claim 10: Bob either measures in the Breidbart basis, or stores the qubit as is. We believe that a similar analysis can be done for the dephasing channel, by first symmetrizing the noise by applying a rotation over $\pi/4$ to our input states.

11.5.2 Noise tradeoff

We now consider the more practical setting, where the honest parties also experience noise. Clearly, there is a strict tradeoff between the noise p_{error} on the channel experienced by the honest parties, and the noise experienced by dishonest Bob. Our practical security bound is fairly weak. In the near-future we may anticipate

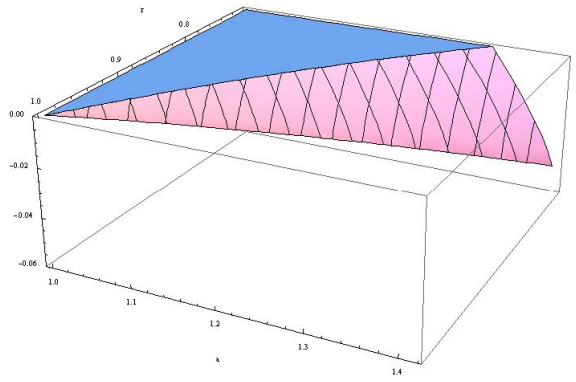


Figure 11.2: $h((1-ar)/2)/4+\log(\frac{1+r}{2})\log(4/3)/2$, where we only show the region below 0, i.e., where security can be attained.

that storage is better than direct measurement if good photonic memories become available. However, we are free in our protocol to stretch the waiting time T between Bob's reception of the qubits and his reception of the classical basis information, say, to seconds, which means that one has to consider the overall noise rate on a qubit that is stored for seconds.

We again consider the case of depolarizing noise during storage. For $r < 1/\sqrt{2}$ (when it is better for Bob to measure in the Breidbart basis), we obtain that our protocol is secure as long as

$$h(p_{\text{error}}) < 2\log\left(\frac{1}{2}+\frac{1}{2\sqrt{2}}\right)\log(3/4).$$

Hence, we require that $p_{\text{error}} \lessapprox 0.029$. This puts a strong restriction on the noise rate of the honest protocol. Yet, since our protocols are particularly interesting at short distances (e.g. in the case of secure identification), we can imagine very short free-space implementations such that depolarization noise during transmission is negligible and the main depolarization noise source is due to Bob's honest measurements.

For $r \geq 1/\sqrt{2}$ (when it's better for Bob to store the qubit as is) we also obtain a tradeoff involving r. As an example, suppose that the qubits in the honest protocol are also subjected to depolarizing noise at rate $1-r_{d,\text{honest}}$. The effective classical error rate for a depolarizing channel is then simply $p_{\text{error}} = (1-r_{d,\text{honest}})/2$. Thus we can consider when the function $h(p_{\text{error}})/4 + \log(\frac{1+r}{2})\log(4/3)/2$ goes below 0. If we assume that $r_{d,\text{honest}} = ar$, for some scaling factor $1 \leq a \leq 1/r$ (i.e., the honest party never has more noise than the dishonest party), we obtain a clear tradeoff between a and r depicted in Figure 11.2.

11.6 Conclusion

We have introduced the model of noisy-quantum storage. In this model, we have determined security bounds for a perfect ROT protocol given collective storage attacks by Bob. Furthermore, we showed how to construct a practical ROT where we do allow the honest parties to experience noise during transmissions and their operations as well. We provided an explicit security tradeoff between the noise affecting the honest parties, and the noise during storage for a dishonest Bob.

Ideally, we would like to show security against general coherent noisy attacks. The problem with analyzing a coherent attack of Bob described by some super-operator $\mathcal{S}$ affecting all his incoming qubits is not merely a technical one: one first needs to determine a realistic noise model in this setting. It may be possible using variations of de Finetti theorems as in the proof of QKD [Ren05] to prove for a symmetrized version of our protocol that any coherent attack by Bob is equivalent to a collective attack. Yet, the present scenario differs in that it is not as straightforward to achieve a symmetrization of the protocol. However, one can in fact analyze a specific type of coherent noise, one that essentially corresponds to an eavesdropping attack in QKD. Note that the 1-2 OT protocol can be seen as two runs of QKD interleaved with each other. The strings $f(x_{|\mathcal{I}_+})$ and $f(x_{|\mathcal{I}_\times})$ are then the two keys generated. The noise must be such that it leaves Bob with exactly the same information as the eavesdropper Eve in QKD. In this case, it follows from the security of QKD that the dishonest Bob (learning exactly the same information as the eavesdropper Eve) does not learn anything about the two keys.

In terms of long-term security, *fault-tolerant* photonic computation (e.g., with the KLM scheme [KLM01]) might allow a dishonest Bob to encode the incoming quantum information into a fault-tolerant quantum memory. This implies that in storage, the effective noise rate can be made arbitrarily small. However, the encoding of a single unknown state is *not* a fault-tolerant quantum operation: already the encoding process introduces errors whose rates cannot be made arbitrarily small with increasing effort. Hence, even in the presence of a quantum computer, there is a residual storage noise rate due to the unprotected encoding operations. The question of security then becomes a question of a trade-off between this residual noise rate versus the intrinsic noise rate. Finally, it remains to address composability of the protocol within our model, which has already been considered for the bounded-quantum-storage model [WW07].

Appendix A

Linear algebra and semidefinite programming

Semidefinite programming is a useful tool to solve optimization problems. Since we employed semidefinite programming in Chapters 3, 7, and 11, we briefly state the most important notions. We refer to [BV04] for an in-depth introduction.

A.1 Linear algebra prerequisites

Before turning to semidefinite programming in the next section, we first briefly recall some elementary definitions from linear algebra. We thereby assume the reader is familiar with basic concepts, such as matrix multiplication and addition. Unless explicitly indicated, all vector spaces V considered here are over the field of complex numbers. We use $V = \mathbb{C}^d$ to denote a d-dimensional complex vector space, and $\mathbb{C}^{d\times d}$ to denote the space of complex $d \times d$ matrices. A set of vectors $|v_1\rangle, \ldots, |v_d\rangle \in V$ is *linearly independent* if $\sum_{i=1}^{d} a_i|v_i\rangle = 0$ implies that $a_1 = \ldots = a_d = 0$. A *basis* of a d-dimensional vector space V is a set of linearly independent vectors $|v_1\rangle, \ldots, |v_d\rangle \in V$, the *basis vectors*, such that any vector $|u\rangle \in V$ can be written as a linear combination of basis vectors. We use $\langle v|$ to denote the conjugate transpose of a vector $|v\rangle$. If there exists a vector $|v\rangle \in V$ with $|v\rangle \neq 0$ such that $A|v\rangle = \lambda|v\rangle$, we say that $|v\rangle$ is an *eigenvector* of A and the scalar λ the corresponding *eigenvalue*.

The *inner product* of two vectors $|u\rangle, |v\rangle \in V$ with $|u\rangle = (u_1, \ldots, u_d)$ and $|v\rangle = (v_1, \ldots, v_d)$ is given by $\langle u|v\rangle = \sum_i u_i^* v_i$. The 2-*norm* of a vector is given by $|||v\rangle|| = \sqrt{\langle v|v\rangle}$. Unless otherwise indicated, all norms of a vector are 2-norms in this text. We also use $|||v\rangle||_V$ to denote emphasize that the norm is defined on a vector space V. Two vectors $|u\rangle, |v\rangle \in V$ such that $\langle u|v\rangle = 0$ are *orthogonal*. If, in addition, $|||u\rangle|| = |||v\rangle|| = 1$ then they are also called *orthonormal*.

A *Hilbert space* is defined as a vector space V with an inner product, where the vector space is complete. We refer to [Con90] for a formal definition of the

notion of completeness and merely note that informally a vector space is complete if for any sequence of vectors in said space approaching a limit, the limit is also an element of the vector space. A *bounded operator* is an operator $A : V \to V'$ such that there exists a $c \in \mathbb{R}$ satisfying $||A|v\rangle||_{V'} \leq c|||v\rangle||_V$ for all $v \in V$. The smallest such c is also called the *operator norm* of A.

The *transpose* of a matrix A is written as A^T and given by $A^T_{ij} = A_{ji}$, where A_{ij} denotes the entry of the matrix A at column i and row j. Similarly, the conjugate transpose $A^\dagger$ of A is of the form $A^\dagger_{ij} = A^*_{ji}$. We use $\mathbb{I}$ to denote the identity matrix defined as $\mathbb{I} = [\mathbb{I}_{ij}]$ with $\mathbb{I}_{ij} = \delta_{ij}$. A matrix U is called *unitary* if $UU^\dagger = U^\dagger U = \mathbb{I}$. Furthermore, M is called *Hermitian* if and only if $M = M^\dagger$. Any Hermitian matrix can be decomposed in terms of its eigenvalues λ_j and eigenvectors $|u_j\rangle$ as $M = \sum_j \lambda_j |u_j\rangle\langle u_j|$, where $|u_j\rangle\langle u_j|$ is a projector onto the vector $|u_j\rangle$. We also call this the *eigendecomposition* of M. The *support* of M is the space spanned by all its eigenvectors with non-zero eigenvalue.

The *tensor product* of an $m \times n$-matrix A and an $m' \times n'$ matrix B is given by the $mm' \times nn'$-matrix

$$A \otimes B = \begin{pmatrix} A_{11}B & \dots & A_{1n}B \\ A_{21}B & \dots & A_{2n}B \\ & \ddots & \\ A_{n1}B & \dots & A_{nn}B \end{pmatrix}.$$

The tensor product is also defined for two vector spaces V and V'. In particular, if the basis of the d-dimensional vector space V is given by $\{|v_1\rangle, \dots, |v_d\rangle\}$ and the basis of the d'-dimensional vector space V' is given by $\{|v'_1\rangle, \dots, |v'_{d'}\rangle\}$, then $W = V \otimes V'$ denotes the $d \cdot d'$-dimensional vector space W with basis $\{|v_i\rangle \otimes |v'_j\rangle \mid i \in [d], j \in [d']\}$.

The *direct sum* of an $m \times n$-matrix A and an $m' \times n'$ matrix B is given by the $m + m' \times n + n'$ matrix

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}.$$

Two vector spaces V and V' defined as above can also be composed in an analogous fashion yielding a $d + d'$ dimensional vector space $W = V \oplus V'$, where any $|w\rangle \in W$ can be written as $|w\rangle = |v\rangle \oplus |v'\rangle$ for some $|v\rangle \in V$ and $|v'\rangle \in V'$ with $|v\rangle \oplus |v'\rangle = (v_1, \dots, v_d, v'_1, \dots, v'_{d'})$ for $|v\rangle = (v_1, \dots, v_d)$ and $|v'\rangle = (v'_1, \dots, v'_{d'})$.

The *trace* of a matrix A is given by the sum of its diagonal entries $\mathrm{Tr}(A) = \sum_i A_{ii}$. Note that $\mathrm{Tr}(A + B) = \mathrm{Tr}(A) + \mathrm{Tr}(B)$, and $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$. If A is an Hermitian matrix, then $\mathrm{Tr}(A)$ is the sum of its eigenvalues.

Finally, the *rank* of a matrix A is denoted as $\mathrm{rank}(A)$ and given by the maximal number of linearly independent columns (or rows) of A.

A.2 Definitions

We now turn to the definitions relevant for our discussion of semidefinite programming. A Hermitian matrix M is *positive semidefinite* if and only if all of its eigenvalues are non-negative [HJ85, Theorem 7.2.1]. Throughout this text, we use $M \geq 0$ to indicate that M is positive semidefinite. We know from [HJ85, Theorem 7.2.11]:

A.2.1. Proposition. *For a Hermitian matrix $M \in \mathbb{C}^{d\times d}$ the following three statements are equivalent:*

1. $M \geq 0$,
2. $x^\dagger M x \geq 0$ *for all vectors* $x \in \mathbb{C}^d$,
3. $M = G^\dagger G$ *for some matrix* $G \in \mathbb{C}^{d\times d}$.

M is called *positive definite* if and only if all of its eigenvalues are positive: we have $x^\dagger M x > 0$ for all vectors $x \in \mathbb{C}^d$. We use $M > 0$ to indicate that M is positive definite. We also encounter projectors, where a Hermitian matrix M is a *projector* if and only if $M^2 = M$. Note that this implies that $M \geq 0$. We say that two projectors M_1 and M_2 are *orthogonal projectors* if and only if $M_1 M_2 = 0$.

Furthermore, we use $\mathcal{S}^d$ to denote the set of all Hermitian matrices, $\mathcal{S}^d = \{X \in \mathbb{C}^{d\times d} \mid X = X^\dagger\}$, and $\mathcal{S}^d_+ = \{X \in \mathcal{S}^d \mid X \geq 0\}$ for the set of all positive semidefinite matrices. A set $\mathcal{T}$ is a *cone*, if for any $\alpha \geq 0$ and $T \in \mathcal{T}$ we have $\alpha T \in \mathcal{T}$. A set $\mathcal{T}$ is *convex*, if for any $\alpha \in [0,1]$ and $T_1, T_2 \in \mathcal{T}$ we have $\alpha T_1 + (1-\alpha)T_2 \in \mathcal{T}$. A set $\mathcal{T}$ is called a *convex cone*, if $\mathcal{T}$ is convex and a cone: for any $\alpha_1, \alpha_2 \geq 0$ and $T_1, T_2 \in \mathcal{T}$ we must have that $\alpha_1 T_1 + \alpha_2 T_2 \in \mathcal{T}$. Note that $\mathcal{S}^d_+$ is a convex cone: Let $\alpha_1, \alpha_2 \geq 0$, and $A, B \in \mathcal{S}^d_+$. Then for any $x \in \mathbb{C}^d$ we have

$$x^\dagger(\alpha_1 A + \alpha_2 B)x = \alpha_1 x^\dagger A x + \alpha_2 x^\dagger B x \geq 0.$$

Hence, $\alpha_1 A + \alpha_2 B \in \mathcal{S}^d_+$. The following will be of use in Chapter 3.

A.2.2. Proposition. *Let $A, B \in S^d$. Then $A \geq 0$ if and only if for all $B \geq 0$* $\mathrm{Tr}(AB) \geq 0$.

Proof. Suppose that $A \geq 0$. Note that we can decompose $B = \sum_j \lambda_j |u_j\rangle\langle u_j|$ where for all j $\lambda_j \geq 0$ since $B \geq 0$. Hence, $\mathrm{Tr}(AB) = \sum_j \lambda_j \mathrm{Tr}(A|u_j\rangle\langle u_j|) = \sum_j \lambda_j \langle u_j|A|u_j\rangle \geq 0$, since $A \geq 0$.

To prove the converse, suppose on the contrary that for all $B \geq 0$ we have $\mathrm{Tr}(AB) \geq 0$, but $A \not\geq 0$. If $A \not\geq 0$, then there exists some vector $|v\rangle$ such that $\langle v|A|v\rangle < 0$. Let $B = |v\rangle\langle v|$. Clearly, $B \geq 0$ and $\mathrm{Tr}(AB) = \langle v|A|v\rangle < 0$ which is a contradiction. □

A.3 Semidefinite programming

Semidefinite programming is a special case of convex optimization. Its goal is to solve the following semidefinite program (SDP) in terms of the variable $M \in \mathcal{S}^d$

$$\begin{array}{ll} \text{maximize} & \mathrm{Tr}(CM) \\ \text{subject to} & \mathrm{Tr}(A_i M) = b_i, i = 1, \ldots, p, \text{ and } M \geq 0 \end{array}$$

for given matrices $C, A_1, \ldots, A_p \in \mathcal{S}^d$. The above form is called the *standard form* of an SDP, and any SDP can be cast in this fashion (possibly at the expense of additional variables) [BV04]. To gain some geometric intuition about this task, note that $M \geq 0$ means that M must lie in the cone $\mathcal{S}^d_+$. The constraints $Tr(A_i M) = b_i$ determine a set of hyperplanes which further limit our possible solutions. A matrix M is called *feasible*, if it satisfies all constraints.

An important aspect of semidefinite programming is duality. Intuitively, the idea behind Lagrangian duality is to extend the objective function (here $\mathrm{Tr}(CM)$) with a weighted sum of the constraints in such a way, that we will be penalized if the constraints are not fulfilled. The weights then correspond to the dual variables. Optimizing over these weights then gives rise to the *dual problem*. The original problem is called the *primal problem*. For the above SDP in standard form, we can write down the Lagrangian as

$$\begin{aligned} L(M, \lambda_1, \ldots, \lambda_p, K) &= \mathrm{Tr}(CM) + \sum_{i=1}^{p} \lambda_i (b_i - \mathrm{Tr}(A_i M)) + \mathrm{Tr}(KM) \\ &= \mathrm{Tr}((C - \sum_i \lambda_i A_i + K)M) + \sum_i \lambda_i b_i, \end{aligned}$$

where $K \geq 0$. The dual function is then

$$\begin{aligned} g(\lambda_1, \ldots, \lambda_p, K) &= \sup_M \left(\mathrm{Tr}((C - \sum_i \lambda_i A_i + K)M) + \sum_i \lambda_i b_i \right) \\ &= \begin{cases} \sum_i \lambda_i b_i & \text{if } C - \sum_i \lambda_i A_i + K = 0 \\ \infty & \text{otherwise} \end{cases} \end{aligned}$$

From $C - \sum_i \lambda_i A_i + K = 0$ and $K \geq 0$, we obtain that $K = -C + \sum_i \lambda_i A_i \geq 0$. This gives us the dual problem as

$$\begin{array}{ll} \text{minimize} & \sum_i \lambda_i b_i \\ \text{subject to} & \sum_i \lambda_i A_i \geq C, \end{array}$$

where the optimization is now over the dual variables λ_i.

We generally use d^* to denote the optimal value of the dual problem, and p^* for the optimal value of the primal problem. Weak duality says that $d^* \geq p^*$.

Let's see why this is true in the above construction of the dual problem. Let $M^{(*)}$ and $\{\lambda_i^{(*)}\}$ be the optimal solutions to the primal and dual problem respectively. In particular, this means that $M^{(*)}$ and $\{\lambda_i^{(*)}\}$ must satisfy the constraints. Then

$$\begin{aligned} d^* - p^* &= \sum_i \lambda_i^{(*)} b_i - \mathrm{Tr}(CM^{(*)}) \\ &= \sum_i \lambda_i^{(*)} \mathrm{Tr}(A_i M^{(*)}) - \mathrm{Tr}(CM^{(*)}) \\ &= \mathrm{Tr}\left(\left(-C + \sum_i \lambda_i^{(*)} A_i\right) M^{(*)}\right) \geq 0, \end{aligned}$$

by Proposition A.2.2 since $M^{(*)} \geq 0$ and $\sum_i \lambda_i^{(*)} A_i \geq C$. An important consequence of weak duality, is that if we have $d^* = p^*$ for a feasible dual and primal solution respectively, we can conclude that both solutions are optimal. If solutions exist such that $d^* = p^*$, we also speak of strong duality. We know from Slater's conditions [BV04], that strong duality holds if there exists a feasible solution to the primal problem which also satisfies $M > 0$.

A.4 Applications

In many quantum problems, we want to optimize over states, or measurement operators. Evidently, semidefinite programming is very well suited to this case: When optimizing over a state ρ, we ask that $\rho \geq 0$ and $\mathrm{Tr}(\rho) = 1$. When optimizing over measurement operators $M_1, \ldots, M_k$ belonging to one POVM, we ask that $M_j \geq 0$ for all $j \in [k]$ and $\sum_j M_j = \mathbb{I}$. Concrete examples can be found in Chapters 3, 7, and 11.

Appendix B

C^*-Algebra

As C^*-algebras are not usually encountered in computer science, we briefly state the most important results we will refer to for convenience. In particular, they help us understand the framework of post-measurement information we encountered in Chapter 3 as well as the structure of bipartite non-local games in Chapter 6.

B.1 Introduction

Instead of starting out with the usual axioms of quantum states and their evolutions, any physical system can be characterized by a C^*-algebra $\mathscr{A}$ of observables. States of this system are now identified purely by means of measurements of these observables. This starting point is rather beautiful in its abstraction: So far, nothing has been said how we can represent elements of this algebra. Yet, it turns out that all the usual axioms can be derived from this abstract structure: we can represent observables as operators and states as vectors in a Hilbert space. In fact, any such algebra $\mathscr{A}$ is isomorphic to an algebra of bounded operators on a Hilbert space. So why should we bother adopting this abstract viewpoint? It turns out that C^*-algebras often make it easier to understand the fundamental differences between the classical and the quantum setting. If the algebra $\mathscr{A}$ is abelian, we have a classical system. Otherwise, our system is inherently quantum. Commutativity leads to several nice structural properties of an algebra which have been exploited to answer many central questions in quantum information: When can we clone physical states? What information can be extracted without disturbing the system? That is, what part of a system is in fact classical and what is truly quantum?

Here, we will mere scratch the surface of this formalism. In particular, we will focus on finite-dimensional C^*-algebras only, which is all we will need in Chapters 3 and 6. For more information, consult any textbook on the topic [Tak79, BR02, Arv76]. We assume that the reader is familiar with the basic concepts such

as a Hilbert space and refer to [Con90] for an introduction. First, we need to introduce some essential definitions in Section B.2. We then examine states and observables, and their familiar representation in a Hilbert space in Section B.3.2. In Section B.4, we then concentrate on commutation: We will sketch how from commutation relations we in fact obtain a bipartite structure. It turns out that commutation relations also play an important role in determining which operations leave states invariant. Looking at the structure of the problem, it turns out that in fact many problems ranging from cloning to post-measurement information and bipartite non-local games are quite closely related.

B.2 Some terminology

A Banach algebra $\mathscr{A}$ is a linear associative algebra[1] which is also a Banach space, with the property that for all A and $B \in \mathscr{A}$ we have

$$\| AB \| \leq \| A \| \| B \|.$$

The norm $\| A \|$ of A is thereby a real number satisfying the usual requirements that for all $A \in \mathscr{A}$ we have $\| A \| \geq 0$ where $\| A \| = 0$ if and only if $A = 0$, $\| \alpha A \| = \alpha \| A \|$, $\| A + B \| \leq \| A \| + \| B \|$, and $\| AB \| \leq \| A \| \| B \|$. $\mathscr{A}$ is called a $*$-algebra if it has the additional property that it admits an involution $A \to A^\dagger \in \mathscr{A}$ such that for all A and $B \in \mathscr{A}$ the following holds: $(A^\dagger)^\dagger = A$, $(A + B)^\dagger = A^\dagger + B^\dagger$, $(\alpha A)^\dagger = \bar{\alpha} A^\dagger$, and $(AB)^\dagger = B^\dagger A^\dagger$. A C^*-algebra is now an even more special case: in addition we also have that $\| A^\dagger A \| = \| A \|^2$ for all $A \in \mathscr{A}$. This also gives us $\| A^\dagger \| = \| A \|$. In the following we will simply use the term "algebra" to refer to a C^*-algebra. The trick is not to be intimidated. It is easier to have a more concrete picture in mind: For example, the algebra $\mathbb{B}(\mathcal{H})$ of all bounded operators on a Hilbert space $\mathcal{H}$ is a C^*-algebra, when we take sums and products of operators in the usual way and take our norm to be the operator norm $\| A \| = \sup(\| Av \| \mid v \in \mathcal{H}, \| v \| = 1)$, where $\| v \|^2 = \langle v | v \rangle$ for the inner product $\langle \cdot | \cdot \rangle$ of the Hilbert space. This algebra is closed under all the usual operations such as addition, multiplication, and multiplication by scalars[2] and the involution operation. This involution is now the adjoint operation $A \to A^\dagger$, which in physics is usually denoted by $\dagger$ instead of $*$. In some physics papers, you will therefore also find the name $\dagger$-algebra instead. As in the example of post-measurement information, we are also often interested in the $*$-algebra generated by a given set of operators. Any operator X in a Hilbert space $\mathcal{H}$ determines a C^*-algebra $\mathscr{A}$ which we will denote by $\mathscr{A} = \langle X \rangle$. This is the smallest C^*-algebra which contains both X and the identity, i.e. $\langle X \rangle = \bigcap_{X, \mathbb{I} \in \mathscr{B}} \mathscr{B}$. What's included in $\langle X \rangle$? Recall that $\mathscr{A}$ is closed under the adjoint operation so we definitely

[1] An associative algebra over the complex numbers is a vector space over the complex numbers with a multiplication that is associative.

[2] We will take the underlying field to be $\mathbb{C}$.

have $X^\dagger$. In addition, our conditions above imply that we will see all possible polynomials in X and $X^\dagger$. For example, $X+X^\dagger$ and $XX^\dagger$ are also elements of the algebra. We use $\langle X_1, \ldots, X_k \rangle$ to denote the C^*-algebra generated by operators $X_1, \ldots, X_k$, and $\langle \mathcal{S} \rangle$ to denote the algebra generated by operators from the set $\mathcal{S}$.

If an algebra $\mathscr{B}$ satisfies $\mathscr{B} \subseteq \mathscr{A}$, we call $\mathscr{B}$ a *subalgebra* of $\mathscr{A}$. An algebra $\mathscr{A}$ is *unital* if it contains the identity. We will always use $\mathbb{I}$ to denote the identity element. Since we restrict ourselves to the finite-dimensional case, we can assume that any C^*-algebra is in fact unital [Tak79]. We will always take $\mathscr{A}$ to be unital here. An element $A \in \mathscr{A}$ of a Banach algebra $\mathscr{A}$ is called *invertible* if there exists some $A' \in \mathscr{A}$ such that $AA' = A'A = \mathbb{I}$. Furthermore, for a C^*-algebra $\mathscr{A}$, the *spectrum* of $A \in \mathscr{A}$ is given by $\mathrm{Sp}_{\mathscr{A}}(A) = \{\lambda \in \mathbb{C} \mid A - \lambda\mathbb{I} \text{ is not invertible}\}$. Note that for any $A \in \mathbb{B}(\mathcal{H})$, this is just the spectrum of the operator relative to $\mathbb{B}(\mathcal{H})$ in the usual sense.

A *left ideal* in some algebra $\mathscr{A}$ is a subalgebra $\mathscr{B} \subseteq \mathscr{A}$ such that for any elements $B \in \mathscr{B}$ and $A \in \mathscr{A}$ we have that $AB \in \mathscr{B}$. Similarly, $\mathscr{B}$ is called a *right ideal* if $BA \in \mathscr{B}$. A *two-sided ideal* or simply *ideal* has both properties: $\mathscr{B}$ is both a left and right ideal of $\mathscr{A}$. An algebra $\mathscr{A}$ is called *simple* if its only ideals are $\{0\}$ and $\mathscr{A}$ itself. An algebra $\mathscr{A}$ is called *semisimple*, if it can be written as the direct sum of simple algebras. To get a better feeling for what this actually means, it is perhaps again helpful to think of a particular representation of the algebra in terms of bounded operators on a Hilbert space. In terms of representations, being simple means that the representation is irreducible. Being semisimple then means that the representation is completely reducible: i.e. for the representation π of A we can express $\pi(A)$ as a sum of irreducible representations. We will examine this decomposition in more detail in Section B.4.1.

B.3 Observables, states and representations

B.3.1 Observables and states

A physical system is characterized by a set of measurable quantities, i.e. observables. As mentioned above, we will assume that a physical system is in fact described by a C^*-algebra $\mathscr{A}$ of observables. As we will see below, we can take the observables to live in a Hilbert space $\mathcal{H}$, and $\mathscr{A} \subseteq \mathbb{B}(\mathcal{H})$. Where do the states come in? In the language of C^*-algebras, states are *positive linear functionals* on $\mathscr{A}$: A linear functional on an algebra is a function $f : \mathscr{A} \to \mathbb{C}$ such that for all $A, B \in \mathscr{A}$ we have $f(A+B) = f(A) + f(B)$ and $f(\alpha A) = \alpha f(A)$ where $\alpha \in \mathbb{C}$ is a scalar. A linear functional is called *positive* if $f(A) \geq 0$ for any $A \in \mathscr{A}$ whenever $A \geq 0$. A *state* on $\mathscr{A}$ is a positive linear functional f on $\mathscr{A}$ with the additional property that it has norm 1, i.e., $f(\mathbb{I}) = 1$. The set of states is a convex set of linear functionals and its extreme elements are called *pure* states.

The set of all states on an algebra $\mathscr{A}$ is also called the *state space*, often denoted by $\mathcal{E}(\mathscr{A})$. Any observable $A \in \mathscr{A}$ in our algebra is uniquely characterized by the expectation of all states when we measure A: So the value of $f(A)$ for all states $f \in \mathcal{E}(\mathscr{A})$ in our state space uniquely characterizes any element A of our algebra. The converse is also true: the value of $f(A)$ for all $A \in \mathscr{A}$ completely characterizes the state f. To get a better feeling for this, it is again helpful to think of an algebra $\mathscr{A} \subseteq \mathbb{B}(\mathcal{H})$. Given a vector v living in the Hilbert space $\mathcal{H}$, we can construct a linear functional on $\mathscr{A}$ by letting $f(A) = \langle v|Av\rangle$. The same is true if we consider any abstract $\mathscr{A}$ and its representation π on a Hilbert space, by letting $f(A) = \langle v|\pi(A)v\rangle$ given $v \in \mathcal{H}$.

B.3.2 Representations

We now examine how an abstract C^*-algebra can be represented by a set of operators on a Hilbert space, via the famous construction by Gelfand, Naimark and Segal. An account of this construction can be found in any standard textbook on C^*-algebra [Tak79, BR02, Arv76]. For completeness, we here give a heavily annotated, largely self-contained, explanation of the GNS construction. As it turns out, by the GNS construction, any C^*-algebra is isomorphic to an algebra of bounded operators, a result which we will merely state here. When trying to find a representation of a C^*-algebra $\mathscr{A}$, our goal is to find a pair $(\pi, \mathcal{H})$ where $\mathcal{H}$ is a Hilbert space and $\pi : \mathscr{A} \to \mathscr{B}(\mathcal{H})$ is a $*$-homomorphism which maps any element of our algebra to a bounded operator in the chosen Hilbert space.

B.3.1. Theorem (GNS). *Let $\mathscr{A}$ be a unital C^*-algebra, and let f be a positive linear functional on $\mathscr{A}$. Then there exists a representation $(\mathcal{H}_f, \pi_f)$ of $\mathscr{A}$ with a Hilbert space $\mathcal{H}_f$, a $*$-homomorphism[3] $\pi_f : \mathscr{A} \to \mathbb{B}(\mathcal{H}_f)$ and a vector $\Phi_f \in \mathcal{H}_f$ such that for all $A \in \mathscr{A}$*

$$f(A) = \langle \Phi_f | \pi_f(A) \Phi_f \rangle.$$

Proof. First, we construct the Hilbert space $\mathcal{H}_f$. Since $\mathscr{A}$ is a Banach space, we can turn it into a pre-Hilbert space[4] by defining the positive semidefinite sesquilinear form

$$\langle A|B\rangle_f = f(A^\dagger B),$$

for all $A, B \in \mathscr{A}$. Note that this form may be degenerate[5]. In order to eliminate this degeneracy, consider

$$\mathcal{I}_f = \{A \mid A \in \mathscr{A} \text{ and } f(A^\dagger A) = 0\}.$$

[3] A homomorphism that preserves the $*$.

[4] We take a *pre-Hilbert space* to be a vector space with a positive semidefinite sesquilinear form, and a *strict pre-Hilbert space* to be a vector space with an inner product.

[5] Such a form is nondegenerate if and only if: $\langle A|B\rangle_f = 0$ for all $B \in \mathscr{A}$ implies that $A = 0$.

Note that $\mathcal{I}_f$ is a linear subspace of $\mathscr{A}$ since for all $I, J \in \mathcal{I}_f$ we have $f((I + J)^\dagger(I + J)) = f(I^\dagger I) + f(J^\dagger I) + f(I^\dagger J) + f(J^\dagger J) \leq 2\sqrt{f(J^\dagger J)f(I^\dagger I)} = 0$, where we used the Cauchy-Schwarz inequality[6].

We now show that $\mathcal{I}_f$ is a left ideal of $\mathscr{A}$: Let $I \in \mathcal{I}_f$ and $A, B \in \mathscr{A}$. We then need to show that $AI \in \mathcal{I}_f$. Indeed, from $(AI)^\dagger(AI) \geq 0$ we have

$$0 \leq f((AI)^\dagger(AI)) = f(I^\dagger A^\dagger AI) \leq \sqrt{f(I^\dagger I)f((A^\dagger AI)^\dagger(A^\dagger AI))} = 0,$$

where the inequality follows from the Cauchy-Schwarz inequality.

The Hilbert space $\mathcal{H}_f$ is then constructed by completing the quotient space $\mathscr{A}/\mathcal{I}_f$. This works as follows: Define the equivalence classes

$$\Psi_A = \{A + I \mid I \in \mathcal{I}_f\}.$$

Note that these equivalence classes constitute a complex vector space on their own, where addition and scalar multiplication are defined via the following operations inherited from $\mathscr{A}$. We have $\Psi_{A+B} = \Psi_A + \Psi_B$ and $\Psi_{\alpha A} = \alpha\Psi_A$. We can then define the inner product

$$\langle\Psi_A|\Psi_B\rangle = \langle A|B\rangle_f = f(A^\dagger B).$$

Note that Ψ_A and Ψ_B of course depend on f. One can verify that this a correct definition. Indeed, the inner product does not depend on our choice of representative from each equivalence class: Let $I_1, I_2 \in \mathcal{I}_f$, and let $A, B \in \mathscr{A}$. Then

$$f((A + I_1)^\dagger(B + I_2)) = f(A^\dagger B) + f(A^\dagger I_2) + f(I_1^\dagger B) + f(I_1^\dagger I_2) = f(A^\dagger B),$$

where the last equality follows again from the Cauchy-Schwarz inequality. We can now obtain $\mathcal{H}_f$ by forming the completion of this space. It is well-known in functional analysis that any strict pre-Hilbert space can be embedded as a dense subspace of a Hilbert space in such a way that the inner product is preserved.

Second, we must construct π_f. We first define the action of $\pi_f(A)$ on the vectors constructed above as

$$\pi_f(A)\Psi_B = \Psi_{AB}.$$

Note that this definition is again independent of our choice of representative from each equivalence class since for all $A, B \in \mathscr{A}$ we have

$$\pi_f(A)\Psi_{B+I} = \Psi_{A(B+I)} = \Psi_{AB+AI} = \Psi_{AB} = \pi_f(A)\Psi_B,$$

since $\mathcal{I}_f$ is a left ideal of $\mathscr{A}$ and we already saw that $AI \in \mathcal{I}_f$. It remains to show that π_f is a homomorphism and that $\pi_f(A)$ is indeed bounded. To see that π_f is a homomorphism, note that

$$\pi_f(AB)\Psi_C = \Psi_{ABC} = \pi_f(A)\pi_f(B)\Psi_C$$

[6]In this context the CS-inequality gives us that for all $A, B \in \mathscr{A}$ we have $|f(A^\dagger B)|^2 \leq f(A^\dagger A)f(B^\dagger B)$

and

$$\pi_f(\lambda A + \gamma B)\Psi_C = \Psi_{\lambda A+\gamma B} = \lambda\Psi_A + \gamma\Psi_B = (\lambda\pi_f(A) + \gamma\pi_f(B))\Psi_C,$$

as desired. To see that $\pi_f(A)$ is bounded, consider

$$\begin{aligned} \| \pi_f(A)\Psi_B \|^2 &= \langle \Psi_{AB} | \Psi_{AB} \rangle = f((AB)^\dagger(AB)) \\ &= f(B^\dagger A^\dagger AB) \leq \| A \|^2 f(B^\dagger B) \leq \| A \|^2 \| \Psi_B \|^2, \end{aligned}$$

where we used the fact that from $B^\dagger A^\dagger AB \leq \| A \|^2 B^\dagger B$ we have $f(B^\dagger A^\dagger AB) \leq \| A \|^2 f(B^\dagger B)$ (see for example [Tak79]).

Finally, we need to construct the vector Φ_f. Since $\mathscr{A}$ is unital we can take $\Phi_f = \Psi_{\mathbb{I}}$. This gives us $\langle \Phi_f | \pi_f(A)\Phi_f \rangle = \langle \Psi_{\mathbb{I}} | \pi_f(A)\Psi_{\mathbb{I}} \rangle = \langle \Psi_{\mathbb{I}} | \Psi_A \rangle = f(\mathbb{I}^\dagger A) = f(A)$. Note that $\pi_f(A)\Psi_{\mathbb{I}} = \Psi_A$, i.e., $\Phi_f = \Psi_{\mathbb{I}}$ is cyclic for $(\mathcal{H}_f, \pi_f)$. □

The resulting representation is irreducible if and only if f is pure [BR02, Theorem 2.3.19]. By considering a family of states F, and applying the GNS construction to all $f \in F$ and taking the direct sum of representations it is then possible to show that:

B.3.2. THEOREM. *(GN) Let $\mathscr{A}$ be a unital C^*-algebra. Then $\mathscr{A}$ is isomorphic to an algebra of bounded operators on a Hilbert space $\mathcal{H}$.*

B.4 Commuting operators

$\mathscr{A}$ is abelian if and only if the physical system corresponding to this algebra is classical. Thus to distinguish the quantum from the classical problems, commutation will be central to our discussion. In fact, it leads to very nice structural properties which we already exploited in Chapter 3. First, however, we will need a bit more terminology. The *commutator* of two operators A and B is given by $[A, B] = AB - BA$. For quantum applications, two observables A and B are called *compatible* if they commute, i.e., $[A, B] = 0$. Conversely, A and B are called *complementary* if $[A, B] \neq 0$. The *center* $\mathscr{Z}_{\mathscr{A}}$ of an algebra $\mathscr{A}$ is the set of all elements in $\mathscr{A}$ that commute with all elements of $\mathscr{A}$, i.e.

$$\mathscr{Z}_{\mathscr{A}} = \{Z \mid Z \in \mathscr{A}, \forall A \in \mathscr{A} : [Z, A] = 0\}.$$

It is easy to see that if $\mathscr{A}$ only has a trivial center, i.e. $\mathscr{Z}_{\mathscr{A}} = \{c\mathbb{I} \mid c \in \mathbb{C}\}$, $\mathscr{A}$ is simple [Tak79]. If $\mathscr{A} \subseteq \mathbb{B}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$, then the *commutant* of $\mathscr{A}$ in $\mathbb{B}(\mathcal{H})$ is

$$\mathrm{Comm}(\mathscr{A}) = \{X \mid X \in \mathbb{B}(\mathcal{H}), \forall A \in \mathscr{A} : [X, A] = 0\}.$$

We have $\mathscr{Z}_{\mathscr{A}} = \mathscr{A} \cap \mathrm{Comm}(\mathscr{A})$.

B.4.1 Decompositions

In any of our problems, the interesting case is when the algebra $\mathscr{A}$ under consideration is in fact simple: that is, "fully quantum". In all problems we will consider, it will turn out that we can always break down the problem into smaller components by decomposing any $\mathscr{A}$ into a sum of simple algebras.[7] Luckily, such a decomposition always exists in the finite-dimensional case:

B.4.1. Lemma. *Let $\mathscr{A}$ be a finite-dimensional C^*-algebra. Then there exists a decomposition*

$$\mathscr{A} = \bigoplus_j \mathscr{A}_j,$$

such that $\mathscr{A}_j$ is simple.

Proof. Let $\mathscr{Z}_A$ be the center of $\mathscr{A}$. Clearly, since $\mathscr{A}$ is finite-dimensional, $\mathscr{Z}_A$ is a finite-dimensional abelian C^*-algebra. Since $\mathscr{Z}_A$ is finite, there exist a finite set of positive linear functionals $\{f_1, \ldots, f_m\}$, such that $f_j(AB) = f_j(A)f_j(B)$ and $f_j(A) \in \mathrm{Sp}_{\mathscr{Z}_A}(A)$ for all $A, B \in \mathscr{Z}_A$.[8] For all $1 \leq k \leq m$, choose $\Pi_k \in \mathscr{Z}_A$ such that $f_j(\Pi_k) = \delta_{jk}$ for all j. Note that $\Pi_1, \ldots, \Pi_m$ are projectors and $\sum_j \Pi_j = \mathbb{I}$ since for all j we have $f_j(\Pi_k\Pi_\ell) = f_j(\Pi_k)f_j(\Pi_\ell)$ since $\mathscr{Z}_A$ is abelian. Now we have

$$\mathscr{A} = \mathbb{I}\mathscr{A}\mathbb{I} = \sum_{jk=1}^{m} \Pi_j \mathscr{A} \Pi_k = \sum_{j=1}^{m} \Pi_j \mathscr{A} \Pi_j,$$

since for all $A \in \mathscr{A}$ we have $\Pi_j A \Pi_k = \Pi_j \Pi_k A = 0$ since $\Pi_j, \Pi_j \in \mathscr{Z}_A$. Note that $\mathscr{A}_j = \Pi_j \mathscr{A} \Pi_j$ only has a trivial center: its only elements that commute with any element of $\mathscr{A}_j$ are scalar multiples of Π_j. Hence, $\mathscr{A}_j$ is simple. □

In fact, it is possible to show that [Tak79]:

B.4.2. Corollary. *Let $\mathscr{A}$ be a finite-dimensional C^*-algebra Then there exists $\mathcal{H}$ and a decomposition*

$$\mathcal{H} = \bigoplus_j \mathcal{H}_j,$$

such that

$$\mathscr{A} \cong \bigoplus_j \mathbb{B}(\mathcal{H}_j),$$

Note that this means that any element $A \in \mathscr{A}$ can be written as $A = \sum_j \Pi_j A \Pi_j$ where Π_j is a projection onto $\mathcal{H}_j$.

[7] Recall that we only consider the finite-dimensional case.

[8] For a matrix algebra these are just the eigenvectors with equal eigenvalue

B.4.2 Bipartite structure

As we saw in Chapter 3, commutation relations induce a beautiful structure captured by the Double Commutant theorem. We here sketch a proof of the parts of this theorem which is interesting for understanding non-local games: Consider a bipartite system $\mathcal{H}^1 \otimes \mathcal{H}^2$, and operators $A = \hat{A} \otimes \mathbb{I}^{[2]}$ and $B = \mathbb{I}^{[1]} \otimes \hat{B}$ with $\hat{A} \in \mathbb{B}(\mathcal{H}^1)$ and $\hat{B} \in \mathbb{B}(\mathcal{H}^2)$. Clearly, $[A, B] = 0$ since A and B act on two different subsystems. Curiously, however, we can essentially reverse the argument: A set of commutation relations gives rise to a bipartite structure itself!

B.4.3. Lemma. *Let $\mathcal{H}$ be a finite-dimensional Hilbert space, and let $\{X_s^a \in \mathbb{B}(\mathcal{H}) \mid s \in S\}$ and $\{Y_t^b \in \mathbb{B}(\mathcal{H}) \mid s \in T\}$. Then the following two statements are equivalent:*

1. *For all $s \in S$, $t \in T$, $a \in A$ and $b \in B$ it holds that $[X_s^a, Y_t^b] = 0$.*
2. *There exist Hilbert spaces $\mathcal{H}^A, \mathcal{H}^B$ such that $\mathcal{H} = \mathcal{H}^A \otimes \mathcal{H}^B$ and for all $s \in S$, $a \in A$ we have $X_s^a \in \mathbb{B}(\mathcal{H}^A)$ and for all $t \in T$, $b \in B$ we have $Y_t^b \in \mathbb{B}(\mathcal{H}^B)$.*

This statement can easily be extended to more than two players. Here, we will only address the finite-dimensional case.

First of all, recall that by Lemma 6.3.1, we can greatly simplify our problem for non-local games and restrict ourselves to C^*-algebras that are simple. As we saw earlier in Lemma B.4.1, it is well known that we can decompose any finite dimensional algebra into the sum of simple algebras. We furthermore need that for any simple algebra, the following holds:

B.4.4. Lemma. *[Tak79] Let $\mathcal{H}$ be a Hilbert space, and let $\mathscr{A} \subseteq \mathbb{B}(\mathcal{H})$ be simple. Then $\mathcal{H} = \mathcal{H}^A \otimes \mathcal{H}^B$ and $\mathscr{A} \cong \mathbb{B}(\mathcal{H}^A) \otimes \mathbb{I}^B$.*

We are now ready to prove Lemma B.4.3. First, we examine the case where we are given a simple algebra $\mathscr{A} \in \mathbb{B}(\mathcal{H})$, for some Hilbert space $\mathcal{H}$. We will need the following version of Schur's lemma.

B.4.5. Lemma. *Let $\mathcal{Z}$ be the center of $\mathbb{B}(\mathcal{H})$. Then $\mathcal{Z} = \{c\mathbb{I} | c \in \mathbb{C}\}$.*

Proof. Let $C \in Z$ and let $d = \dim(\mathcal{H})$. Let $\mathcal{B} = \{E_{ij} | i, j \in [d]\}$ be a basis for $\mathbb{B}(\mathcal{H})$, where $E_{ij} = |i\rangle\langle j|$ is the matrix of all 0's and a 1 at position (i, j). Since $C \in \mathcal{Z}$ and $E_{ij} \in \mathbb{B}(\mathcal{H})$ we have for all $i \in [d]$

$$CE_{ii} = E_{ii}C.$$

Note that CE_{ii} (or $E_{ii}C$) is the matrix of all 0's but the ith column (or row) is determined by the elements of C. Hence all off diagonal elements of C must be 0. Now consider

$$C(E_{ij} + E_{ji}) = (E_{ij} + E_{ji})C.$$

Note that $C(E_{ij} + E_{ji})$ (or $(E_{ij} + E_{ji})C$) is the matrix in which the ith and jth columns (rows) of C have been swapped and the remaining elements are 0. Hence all diagonal elements of C must be equal. Thus there exists some $c \in \mathbb{C}$ such that $C = c\mathbb{I}$. □

Using this Lemma, we can now show that

B.4.6. Lemma. *Let $C \in \mathbb{B}(\mathcal{H}^A \otimes \mathcal{H}^B)$ such that for all $B \in \mathbb{B}(\mathcal{H}^B)$ we have*

$$[C, (\mathbb{I}^A \otimes B)] = 0$$

Then there exists an $A \in \mathbb{B}(\mathcal{H}^A)$ such that $C = A \otimes \mathbb{I}^B$.

Proof. Let $d_A = \dim(\mathcal{H}^A)$ and $d_B = \dim(\mathcal{H}^B)$. Note that we can write any C as

$$C = \begin{pmatrix} C_{11} & \dots & C_{1d_A} \\ \vdots & & \vdots \\ C_{d_A 1} & \dots & C_{d_A d_A} \end{pmatrix},$$

for $d_A \times d_A$ matrices A_{ij}. We have $C(\mathbb{I}^A \otimes B) = (\mathbb{I}^A \otimes B)C$ if and only if for all $i, j \in [d_A]$ $C_{ij}B = BC_{ij}$, i.e. $[C_{ij}, B] = 0$. Since this must hold for all $B \in \mathbb{B}(\mathcal{H}^B)$, we have by Lemma B.4.5 that there exists some $a_{ij} \in \mathbb{C}$ such that $C_{ij} = a_{ij}\mathbb{I}^B$. Hence $C = A \otimes \mathbb{I}^B$ with $A = [a_{ij}]$. □

For the case that the algebra generated by Alice and Bob's measurement operators is simple, Lemma B.4.3 now follows immediately:

Proof. [Proof of Lemma B.4.3 if $\mathscr{A}$ is simple] Let $\mathscr{A} = \langle \{X_s^a\} \rangle \subseteq \mathbb{B}(\mathcal{H})$ be the algebra generated by Alice's measurement operators. If $\mathscr{A}$ is simple, it follows from Lemma B.4.4 that $\mathscr{A} \cong \mathbb{B}(\mathcal{H}^A) \otimes \mathbb{I}^B$ for $\mathcal{H} = \mathcal{H}^A \otimes \mathcal{H}^B$. It then follows from Lemma B.4.6 that for all $t \in T$ and $b \in B$ we must have $Y_t^b \in \mathbb{B}(\mathcal{H}^B)$. □

Thus, we obtain a tensor product structure! Recall that Lemma 6.3.1 states that for non-local games this is all we need.

In general, what happens if $\mathscr{A}$ is not simple? We now sketch the argument in the case the $\mathscr{A}$ is semisimple, which by Lemma B.4.1 we may always assume in the finite-dimensional case. Fortunately, we can still assume that our commutation relations leave us with a bipartite structure. We can essentially infer this from van Neumann's famous Double Commutant Theorem [Tak79, BR02], partially stated here.

B.4.7. Theorem. *Let $\mathscr{A}$ be a finite-dimensional C^*-algebra. Then there exists $\mathcal{H} = \mathcal{H}^A \otimes \mathcal{H}^B$ and a decomposition*

$$\mathcal{H} = \bigoplus_j \mathcal{H}_j^A \otimes \mathcal{H}_j^B$$

such that

$$\mathscr{A} \cong \bigoplus_j \mathbb{B}(\mathcal{H}_j^A) \otimes \mathbb{I}_j^B$$

and

$$Comm(\mathscr{A}) \cong \bigoplus_j \mathbb{I}_j^A \otimes \mathbb{B}(\mathcal{H}_j^B). \tag{B.1}$$

Proof. (Sketch) We already now from Lemma B.4.1 that $\mathscr{A}$ can be decomposed into a sum of simple algebras. Clearly, the RHS of Eq. B.1 is an element of $\mathrm{Comm}(\mathscr{A})$. To see that the LHS is contained in the RHS, consider the projection Π_j^A onto $\mathcal{H}_j^A$. Note that $\Pi_j^A \in \mathscr{A}$, and thus for any $X \in \mathrm{Comm}(\mathscr{A})$ we have $[X, \Pi_j^A] = 0$. Hence, we can write $X = \sum_j (\Pi_j^A \otimes \mathbb{I}^B) X (\Pi_j^A \otimes \mathbb{I}^B)$, and thus we can restrict ourselves to considering each factor individually. The result then follows immediately from Lemma B.4.6. □

If we have more than two players, the argument is essentially analogous, and we merely sketch it in the relevant case when the algebra generated by the players's measurements is simple, since Lemma 6.3.1 directly extends to more than two players as well. Suppose we have N players $\mathcal{P}_1, \ldots, \mathcal{P}_N$ and let $\mathcal{H}$ denote their joint Hilbert space. Let $\mathscr{A}$ be the algebra generated by all measurement operators of players $\mathcal{P}_1, \ldots . \mathcal{P}_{N-1}$ respectively. Then it follows from Lemma B.4.6 and Lemma B.4.4 that $\mathcal{H} = \mathcal{H}^{1,\ldots,N-1} \otimes \mathcal{H}^N$ where $\mathscr{A} \cong \mathbb{B}(\mathcal{H}^{1,\ldots,N-1})$ and for all measurement operators M of player $\mathcal{P}_N$ we have that $M \in \mathbb{B}(\mathcal{H}^N)$. By applying Lemma B.4.6 recursively we obtain that there exists a way to partition the Hilbert space into subsystems $\mathcal{H} = \mathcal{H}^1 \otimes \ldots \otimes \mathcal{H}^N$ such that the measurement operators of player $\mathcal{P}_j$ act on $\mathcal{H}^j$ alone.

In quantum mechanics, we will always obtain such a tensor product structure from commutation relations, even if the Hilbert space is infinite dimensional [Sum90]. Here, we start out with a type-I algebra, the corresponding Hilbert space and operators can then be obtained by the famous GNS construction [Tak79], an approach which is rather beautiful in its abstraction. In quantum statistical mechanics and quantum field theory, we will also encounter factors of type-II and type-III. As it turns out, the above argument does not generally hold in this case, however, there are a number of conditions that can lead to a similar structure. Unfortunately, we cannot consider this case here and merely refer to the survey article by Summers [Sum90].

B.4.3 Invariant observables and states

As we saw in Chapter 3, expressing our problem in terms of commutation relations enables us to exploit their structural consequences. Particularly interesting is also the fact that we can characterize the set of states which are invariant by a quantum channel by means of such relations, repeated here for convenience sake:

B.4.8. Lemma. *(HKL) [HKL03] Let $\Lambda : \mathcal{H} \to \mathcal{H}$ be a unital quantum channel with $\Lambda(\rho) = \sum_m V_m \rho V_m^\dagger$, and let $\mathcal{S}$ be a set of quantum states. Then*

$$\forall \rho \in \mathcal{S}, \Lambda(\rho) = \rho \text{ if and only if } \forall m \forall \rho \in \mathcal{S}, [V_m, \rho] = 0.$$

Let's see what this means for a specific unital channel $\Lambda(\rho) = \sum_m V_m \rho V_m^\dagger$ and a particular ensemble given by states $\rho_1, \ldots, \rho_n \in \mathcal{H}$. As in Chapter 3, we now consider the $*$-algebra generated by $\rho_1, \ldots, \rho_n$. Let $\mathscr{A}$ denote the resulting algebra. By Theorem B.4.7, we know that we can write

$$\mathscr{A} \cong \bigoplus_j \mathbb{B}(\mathcal{H}_j^1) \otimes \mathbb{I}_j^{[2]}$$

and

$$\mathrm{Comm}(\mathscr{A}) \cong \bigoplus_j \mathbb{I}_j^{[1]} \otimes \mathbb{B}(\mathcal{H}_j^2).$$

Clearly, we have from the above that if Λ leaves our ensemble of states untouched, we must have $V_m \in \mathrm{Comm}(\mathscr{A})$ for all m. Thus we know that V_m must be of the form $\mathbb{I}_j^{[1]} \otimes V_j^{[2]}$ on each factor. What does this mean operationally? Suppose we can write $\mathcal{H} = \bigoplus_j \mathcal{H}_j$ such that $\rho_k = \sum_j \Pi_j \rho_k \Pi_j$ for each ρ_k, where Π_j is a projector onto $\mathcal{H}_j$. That is, we can simultaneously block-diagonalize all ρ_k. Then we know that V_m must be equal to the identity on each factor $\mathcal{H}_j$, i.e. V_m must be of the form $\bigoplus_j c_j \Pi_j$ for some c_j with $|c_j| = 1$. Another nice application of this viewpoint is an algebraic no-cloning theorem, as put forward by Lindblad [Lin99].

B.5 Conclusion

Even though the sheer number of new definitions may appear daunting, we saw that the language of C^*-algebras can help us get a grip on some of the fundamental properties of quantum states quite easily. Of course, the language of C^*-algebras is not the most convenient one for all problems. Yet, there are many cases for which the language of C^*-algebras is especially useful. As we saw earlier, one of these cases is when we consider measurements performed by two parties on a bipartite system. Another class of problems deals with questions of the following forms: Which operations leave a given set of states invariant? How much can we learn from a given state without disturbing it? What part of a state is "truly" quantum and which parts can we consider to be classical? How can we encode our states such that they are left untouched by a set of operations?

For example, another application is the compression of quantum states. Koashi and Imoto consider how a quantum state can be decomposed into a quantum, a classical and a redundant part to aid compression. In their paper, they provide an algorithm which in fact allows us to compute (with a lot of pain) the decomposition of an algebra and its commutant algebra [KI02]. It is probably

not so surprising by now that other tasks involving invariance under operations are also closely related: Choi and Kribs [CK06] have phrased the principle of decoherence-free subspaces in terms of what they call algebraic noise commutant formalism. In this text, we have exploited C^*-algebras to investigate the use of post-measurement information in Chapter 3. As we saw in Chapter 8, the question of how much post-measurement information is needed is in fact closely related to how much entanglement we need to succeed in non-local games. Whereas these two problem may appear unrelated at first sight, their structural similarities show their close connection. Likewise, these similarities also enabled us in Chapter 6 to investigate how much we can really gain by receiving additional post-measurement information. Finally, the close connection of C^*-algebras and Clifford algebras discussed in Appendix C was one of the factors that led us to discover the uncertainty relations of Chapter 4. Hence, C^*-algebras sometimes help us to understand the similarities between problems, and aid our intuition.

Appendix C

Clifford Algebra

Similar to C^*-algebra, Clifford algebra plays little role in computer science even though it has recently found numerous applications in the area of computer graphics. Here, we informally summarize the most important facts we need in this text. Our aim is merely to provide the reader with some intuition underlying our uncertainty relation in Chapter 4, and refer to [Lou01] for an in-depth introduction.

C.1 Introduction

Clifford algebra is closely related to C^*-algebra. Yet, it exhibits many beautiful geometrical aspects which remain inaccessible to us otherwise. In particular, we will see that commutation and anti-commutation carries a geometric meaning within this algebra.

For any integer n, the unital associative algebra generated by $\Gamma_1, \ldots, \Gamma_{2n}$, subject to the anti-commutation relations

$$\Gamma_i \Gamma_j = -\Gamma_j \Gamma_i, \ \Gamma_i^2 = \mathbb{I}$$

is called *Clifford algebra*. It has a unique representation by Hermitian matrices on n qubits (up to unitary equivalence) which we fix henceforth. This representation can be obtained via the famous Jordan-Wigner transformation [JW28]:

$$\begin{aligned} \Gamma_{2j-1} &= \sigma_y^{\otimes(j-1)} \otimes \sigma_x \otimes \mathbb{I}^{\otimes(n-j)}, \\ \Gamma_{2j} &= \sigma_y^{\otimes(j-1)} \otimes \sigma_z \otimes \mathbb{I}^{\otimes(n-j)}, \end{aligned}$$

for $j = 1, \ldots, n$. A Clifford algebra of n generators is isomorphic to a C^*-algebra of matrices of size $2^{n/2} \times 2^{n/2}$ for n even and to the direct sum of two C^*-algebras of matrices of size $2^{(n-1)/2} \times 2^{(n-1)/2}$ for n odd [Tsi87].

C.2 Geometrical interpretation

The crucial advantage of the Clifford algebra is that we can view the operators $\Gamma_1, \ldots, \Gamma_{2n}$ as $2n$ orthogonal vectors forming a basis for a $2n$-dimensional real vector space $\mathbb{R}^{2n}$. Each vector $a = (a_1, \ldots, a_{2n}) \in \mathbb{R}^{2n}$ can then be written as linear combination of basis elements as $a = \sum_j a_j \Gamma_j$. The *Clifford product* of two vectors a and b is given by

$$ab = a \cdot b + a \wedge b,$$

where $a \cdot b = \sum_j a_j b_j \mathbb{I}$ is the inner product of two vectors and $a \wedge b$ is the outer product, as given below. We will write scalars as scalar multiples of the identity element whose matrix representation is simply the identity matrix. If we represent $\Gamma_1, \ldots, \Gamma_{2n}$ using the matrices from above, then the Clifford product is simply the matrix product of the resulting matrices. Hence, we will now adopt this viewpoint with the representation in mind. Note that the Clifford product satisfies $a^2 = |a|^2 \mathbb{I} = \sum_j a_j^2 \mathbb{I}$, where $|a| = ||a||_2 = \sqrt{\sum_j a_j^2}$ is the 2-norm of the vector a which we refer to as the *length* of a vector.

C.2.1 Inner and outer product

We can see immediately from the definition of the Clifford product that the inner product of two vectors $a, b \in \mathbb{R}^{2n}$ as depicted in Figure C.1 is given by $a \cdot b = |a||b| \cos\psi \mathbb{I}$, and can be expressed as:

$$a \cdot b = \frac{1}{2} \{a, b\} = \frac{1}{2}(ab + ba).$$

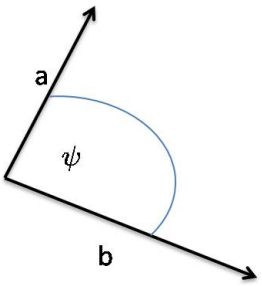


Figure C.1: Two vectors

Hence, anti-commutation takes a geometric meaning within the algebra: two vectors anti-commute if and only if they are orthogonal!

Similarly, we can write

$$a \wedge b = \frac{1}{2} [a, b] = \frac{1}{2}(ab - ba).$$

Geometrically, this means that two vectors are parallel if and only if they commute.

To gain some intuition, let's look at the simple example of $\mathbb{R}^2$: Here, we have $a = a_1\Gamma_1 + a_2\Gamma_2$ and $b = b_1\Gamma_1 + b_2\Gamma_2$. The Clifford product of a and b is now given as

$$ab = \sum_{jk} a_j b_k \Gamma_j \Gamma_k = (a_1 b_1 + a_2 b_2)\mathbb{I} + (a_1 b_2 - b_1 a_2)\Gamma_1\Gamma_2.$$

The element $a \wedge b = (a_1 b_2 - b_1 a_2)\Gamma_1\Gamma_2$ represents the oriented plane segment of the parallelogram determined by a and b in Figure C.2 below. The area of this parallelogram is exactly $|a \wedge b| = |a_1 b_2 - b_1 a_2|$. Note that we have $a \wedge b = -b \wedge a$, as shown in Figure C.3. Thus $a \wedge b$ not only gives us the area but also encodes a direction.

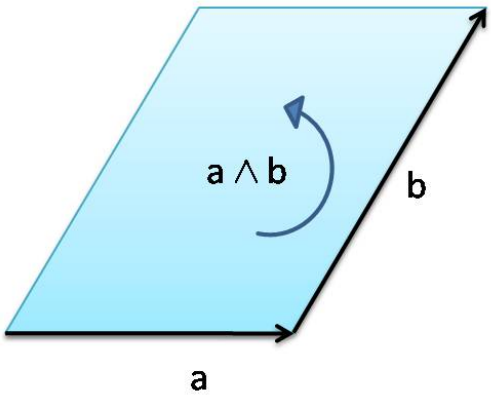


Figure C.2: $a \wedge b$

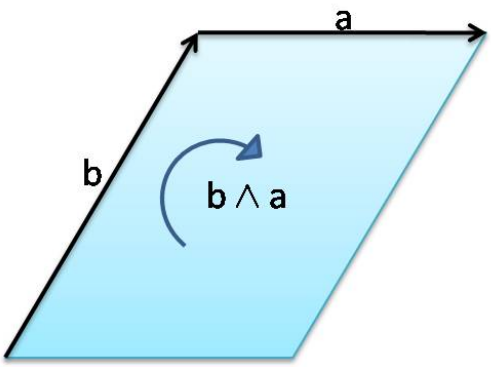


Figure C.3: $b \wedge a$

In higher dimensions, the elements generated by $a \wedge b \wedge c$ etc similarly correspond to oriented plane or volume segments. Note that we have $\Gamma_i \wedge \Gamma_j = \Gamma_i\Gamma_j$ for all basis vectors Γ_i and Γ_j. We will refer to products of k elements of the form $\Gamma_{i_1} \dots \Gamma_{i_k}$ as *k-vectors*.

C.2.2 Reflections

The power of the Clifford algebra mainly lies in the fact that we can express geometrical operations involving any k-vector in an extremely easy fashion using the Clifford product. Here, we will only be concerned with performing operations on 1-vectors.

Consider the projection of a vector a onto a vector m as depicted in Figure C.4. Let $a_\parallel$ be the part of a that is parallel to m, and $a_\perp$ the part of a that lies perpendicular to m. Clearly, we may write $a = a_\parallel + a_\perp$. Using the definition of

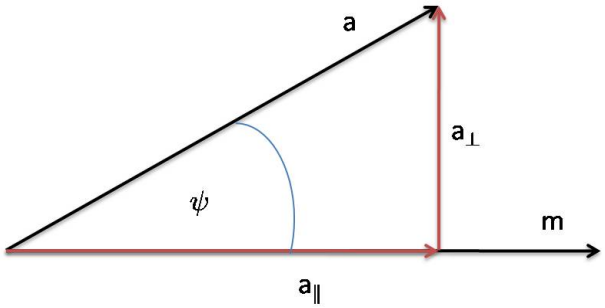


Figure C.4: Projections onto a vector

the Clifford product, we may write

$$a_{\parallel} = |a| \cos\psi \frac{m}{|m|} = (a \cdot m)m^{\dagger},$$

where we define $m^{\dagger} = m/|m|^2$ to be the inverse of m. Indeed, we have $mm^{\dagger} = \mathbb{I}$. If m is a unit vector, then in terms of the matrix representation given above $m^{\dagger}$ is the adjoint of the matrix m. For the product of two vectors we define $(nm)^{\dagger} = m^{\dagger}n^{\dagger}$. We can also write

$$a_{\perp} = a - a_{\parallel} = a - (a \cdot m)m^{\dagger} = (am - (a \cdot m))m^{\dagger} = (a \wedge m)m^{\dagger}.$$

We can now easily determine the reflection of a around the vector m, as depicted in Figure C.5:

$$t = a_{\parallel} - a_{\perp} = (a \cdot m - a \wedge m)m^{\dagger} = (m \cdot a + m \wedge a)m^{\dagger} = mam^{\dagger}.$$

Consider $n = 1$. Then the 2-dimensional real vector space is given by basis vectors $\Gamma_1 = X$ and $\Gamma_2 = Z$. Indeed, this is the familiar XZ-plane of the Bloch sphere depicted in Figure 2.1. Consider the Hadamard transform $H = (X + Z)/\sqrt{2}$. Figure C.7 demonstrates that H plays exactly this role: it reflects X around the vector H to obtain $HXH = Z$. Given t, we can also easily derive the vector obtained by reflecting a around the plane perpendicular to m (in 0), as shown in Figure C.6.

$$-t = -mam^{\dagger}.$$

C.2.3 Rotations

From reflections we may now obtain rotations as successive reflections. Suppose we are given vectors m and n as shown in Figure C.8. To rotate the vector a by

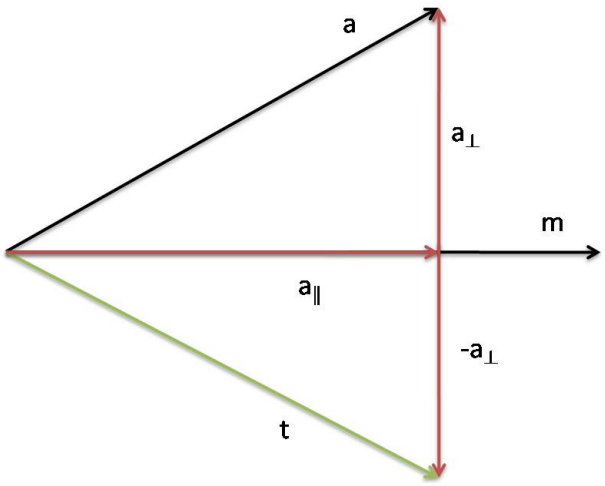


Figure C.5: Reflection of a around m

an angle that is twice the angle between m and n, we now first reflect a around b to obtain $b = mam^\dagger$. We then reflect b around n to obtain

$$c = nbn^\dagger = nmam^\dagger n^\dagger = RaR^\dagger,$$

where we let $R = nm$. As desired, R rotates a by an angle of $2(\psi + \phi)$.

We can easily convince ourselves that R does not affect any vector d that is orthogonal to both n and m.

$$RdR^\dagger = nmdm^\dagger n^\dagger = dnmm^\dagger n^\dagger = d,$$

where we have used the fact that two vectors anti-commute if and only if they are orthogonal. Note also that $RR^\dagger = \mathbb{I}$. It can be shown that if V is a k-vector, then $RVR^\dagger$ is also a k-vector for any rotation R [DL03]. Indeed, this is easy to see, for the k-vector formed by orthogonal basis vectors:

$$\begin{aligned} R(\Gamma_{i_1} \wedge \ldots \wedge \Gamma_{i_k})R^\dagger &= R(\Gamma_{i_1} \ldots \Gamma_{i_k})R^\dagger = R\Gamma_{i_1}R^\dagger \ldots R\Gamma_{i_k}R^\dagger \\ &= R\Gamma_{i_1}R^\dagger \wedge \ldots \wedge R\Gamma_{i_k}R^\dagger, \end{aligned}$$

where we have used the fact that rotations preserve the angles between vectors. We will need this fact in our proof in Chapter 4.

Clifford algebra offers a very convenient way to express rotations around arbitrary angles in the plane $m \wedge n$ [Lou01]. In Chapter 4, however, we will only need to understand how we can find the rotation R that takes us from a given vector $g = \sum_j g_j \Gamma_j$ with length $|g|$ to the vector $|g|\Gamma_1$. Indeed, our strategy works for finding the rotation of any vector g to a target vector t of the same length. Consider Figure C.9.

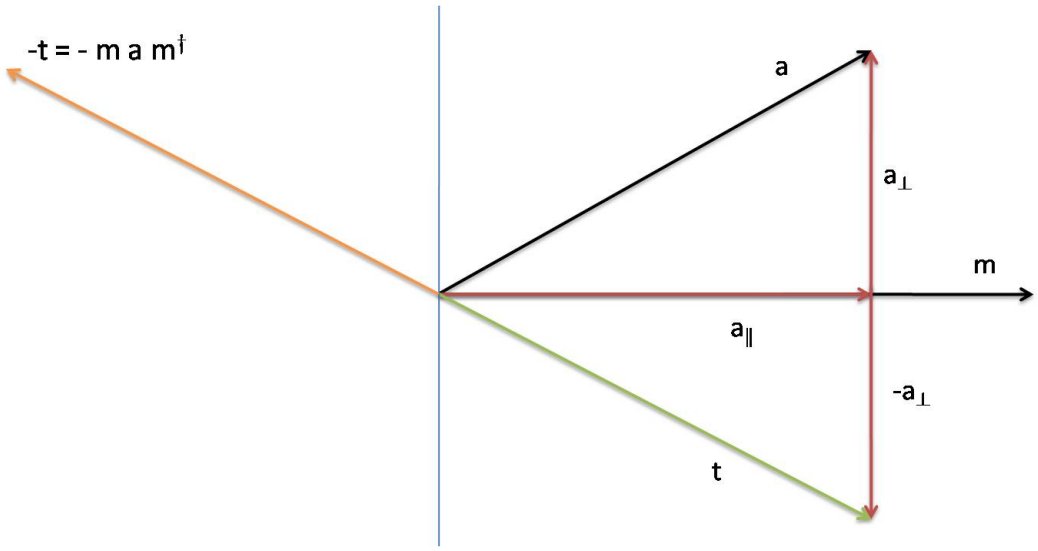


Figure C.6: Reflection of a plane perpendicular to m

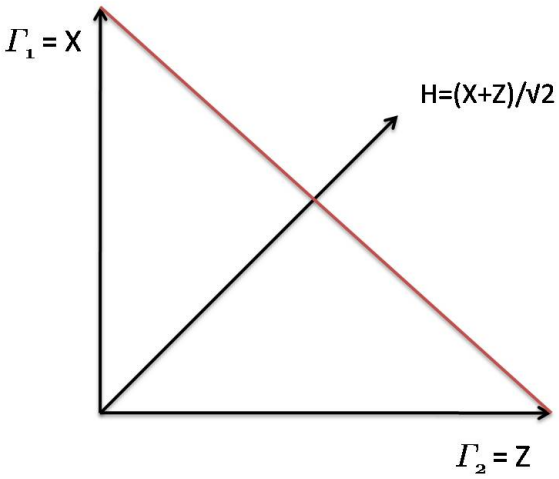


Figure C.7: Hadamard transform as reflection

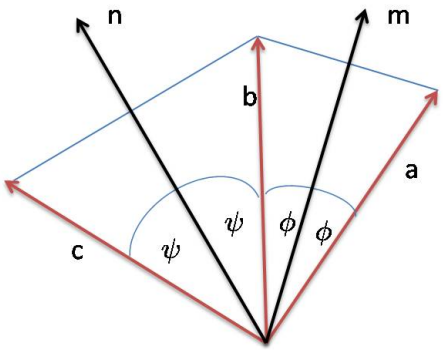


Figure C.8: Rotating in the plane $m \wedge n$.

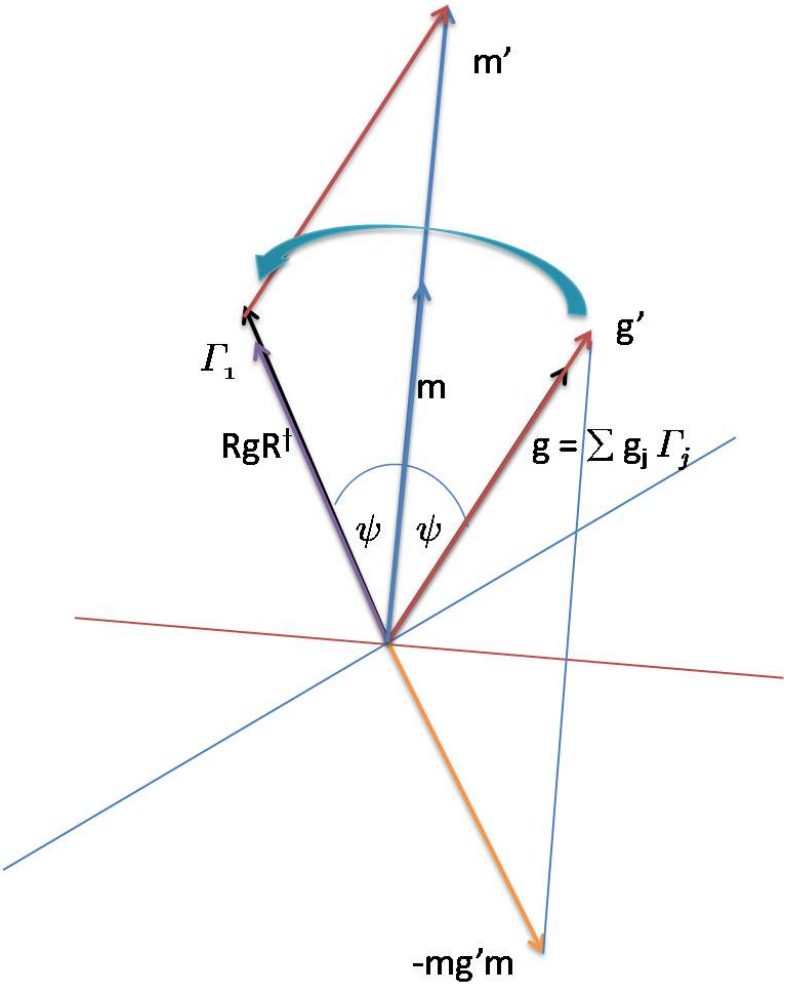


Figure C.9: Rotating g to $|g|\Gamma_1$.

For convenience, we first normalize g to obtain the vector

$$g' = \frac{g}{|g|}.$$

We then compute the vector m' lying exactly half-way between g' and our target vector Γ_1 and normalize it to obtain

$$m = \frac{g' + \Gamma_1}{|g' + \Gamma_1|} = \frac{g' + \Gamma_1}{\sqrt{2(1 + g_1/|g|)}}.$$

We now first reflect g' around the plane perpendicular to the vector m to obtain $-mg'm$, followed by a reflection around the plane perpendicular to the target vector Γ_1:

$$-\Gamma_1(-mg'm)\Gamma_1 = \Gamma_1 mg'm\Gamma_1 = Rg'R^\dagger,$$

with $R = \Gamma_1 m$, where we have used the fact that both Γ_1 and m have unit length and hence $m = m^\dagger$ and $\Gamma_1 = \Gamma_1^\dagger$. Evidently,

$$\begin{aligned} Rg' = \Gamma_1 mg' &= \frac{1}{|g' + \Gamma_1|}\Gamma_1(g' + \Gamma_1)g' \\ &= \frac{1}{|g' + \Gamma_1|}\Gamma_1(g'^2 + \Gamma_1 g') = \frac{1}{|g' + \Gamma_1|}(\Gamma_1 + g') = m, \end{aligned}$$

and hence

$$Rg'R^\dagger = mm\Gamma_1 = \Gamma_1.$$

We then also have that

$$RgR^\dagger = |g|Rg'R^\dagger = |g|\Gamma_1$$

as desired. We will employ a similar rotation in Chapter 4.

C.3 Application

Here, the primary benefit which we gain by considering a Clifford algebra, is that we can parametrize matrices in terms of its generators, and products thereof. Suppose we are given some matrix

$$\rho = \frac{1}{d}\left(\mathbb{I} + \sum_j b_j B_j\right),$$

where $\mathbb{I} \cup \{B_j\}$ form a basis for the $d \times d$ complex matrices, such that for all $j \neq j'$ we have $\mathrm{Tr}(B_j B_{j'}) = 0$, $\mathrm{Tr}(B_j) = 0$, $B_j^2 = \mathbb{I}$, and $b_j \in \mathbb{R}$. We saw in Chapter 4 how to construct such a basis for $d = 2^n$ based on mutually unbiased bases. In fact, this gives us the well-known Pauli basis, given by the 2^{2n} elements of the

form $B_j = B_j^1 \otimes \ldots \otimes B_j^n$ with $B_j^i \in \{\mathbb{I}, \sigma_x, \sigma_y, \sigma_z\}$. When solving optimization problems within quantum information, we are often faced with the following problem: When is ρ a quantum state? That is, what are the necessary and sufficient conditions for the coefficients b_j such that $\rho \geq 0$?

For $d = 2$, this is an easy problem: We can write $\rho = (\mathbb{I} + \sum_{j\in\{x,y,z\}} r_j\sigma_j)/2$ where $\vec{r} = (r_x, r_y, r_z)$ is the Bloch vector we encountered in Chapter 2. We have that $\rho \geq 0$ if and only if $-\mathbb{I} \leq \sum_j r_j\sigma_j \leq \mathbb{I}$, i.e.

$$\left(\sum_j r_j\sigma_j\right)^2 = \frac{1}{2}\sum_{j,j'} r_j r_{j'}\{\sigma_j, \sigma_{j'}\} = \left(\sum_j r_j^2\right)\mathbb{I} \leq \mathbb{I}.$$

Thus, we have $\rho \geq 0$ if and only if $\sum_j r_j^2 \leq 1$. Geometrically, this means that any point on or inside the Bloch sphere corresponds to a valid quantum state as illustrated in Figure 2.1. Sadly, when we consider $d > 2$, our task becomes considerably more difficult. Clearly, since $\mathrm{Tr}(\rho^2) \leq 1$ for any quantum state, we can always say that

$$\begin{aligned}
\mathrm{Tr}(\rho^2) &= \frac{1}{d^2}\left(\mathrm{Tr}(\mathbb{I}) + 2\sum_j b_j\mathrm{Tr}(B_j) + \sum_{jj'} b_j b_{j'}\mathrm{Tr}(B_j B_{j'})\right) \\
&= \frac{1}{d^2}\left(d + \sum_j b_j^2\mathrm{Tr}(\mathbb{I})\right) \\
&= \frac{1}{d}\left(1 + \sum_j b_j^2\right) \leq 1,
\end{aligned}$$

giving us $\sum_j b_j^2 \leq d - 1$. Unfortunately, this condition is too weak for almost all practical applications. There exist many matrices which obey this condition, but nevertheless do not correspond to valid quantum states. Luckily, we can say something much stronger using the Clifford algebra.

Let's consider the operators $\Gamma_1, \ldots, \Gamma_{2n}$ themselves. Evidently, each operator Γ_i has exactly two eigenvalues ± 1: Let $|\eta\rangle$ be an eigenvector of Γ_i with eigenvalue λ. From $\Gamma_i^2 = \mathbb{I}$ we have that $\lambda^2 = 1$. Furthermore, we have $\Gamma_i(\Gamma_j|\eta\rangle) = -\lambda\Gamma_j|\eta\rangle$. Thus, if λ is an eigenvalue of Γ_i then so is $-\lambda$. We can therefore express each Γ_i as

$$\Gamma_i = \Gamma_i^0 - \Gamma_i^1,$$

where Γ_i^0 and Γ_i^1 are projectors onto the positive and negative eigenspace of Γ_i respectively. Furthermore, note that we have for all i, j with $i \neq j$

$$\mathrm{Tr}(\Gamma_i\Gamma_j) = \frac{1}{2}\mathrm{Tr}(\Gamma_i\Gamma_j + \Gamma_j\Gamma_i) = 0,$$

that is all such operators are orthogonal with respect to the Hilbert-Schmidt inner product. We now use the fact that the collection of operators

$$\begin{aligned}
&\mathbb{I} \\
\Gamma_j &\qquad (1 \le j \le 2n) \\
\Gamma_{jk} &:= i\Gamma_j\Gamma_k \ (1 \le j < k \le 2n) \\
\Gamma_{jk\ell} &:= \Gamma_j\Gamma_k\Gamma_\ell \ (1 \le j < k < \ell \le 2n) \\
&\vdots \\
\Gamma_{12\ldots(2n)} &:= i\Gamma_1\Gamma_2\cdots\Gamma_{2n} =: \Gamma_0
\end{aligned}$$

forms an orthogonal basis for the $d \times d$ matrices with $d = 2^n$ [Die06]. By counting, the above operators form a complete operator basis with respect to the Hilbert-Schmidt inner product. In fact, by working out the individual basis elements with respect to the representation above, we see that this basis is in fact equal to the Pauli basis. Notice that the products with an odd number of factors are Hermitian, while the ones with an even number of factors are skew-Hermitian, so in the definition of the above operators we introduce a factor of i to all with an even number of indices to make the whole set a basis for the Hermitian operators. Hence we can write every state $\rho \in \mathcal{H}$ as

$$\rho = \frac{1}{d}\left(\mathbb{I} + \sum_j g_j\Gamma_j + \sum_{j<k} g_{jk}\Gamma_{jk} + \ldots + g_0\Gamma_0\right),$$

with real coefficients $g_j, g_{jk}, \ldots$.

It is clear from the above that if we transform the generating set of Γ_j linearly,

$$\Gamma'_k = \sum_j T_{jk}\Gamma_j,$$

then the set $\{\Gamma'_1, \ldots, \Gamma'_{2n}\}$ satisfies the anti-commutation relations if and only if $(T_{jk})_{jk}$ is an orthogonal matrix: these are exactly the operations which preserve the inner product. In that case there exists a matching unitary $U(T) \in \mathbb{B}(\mathcal{H})$ which transforms the operator basis as

$$\Gamma'_j = U(T)\Gamma_j U(T)^\dagger.$$

As an importance consequence, it can be shown [Die06] that any operation $U(T)$ transforms the state ρ as

$$U(T)\rho U(T)^\dagger = \frac{1}{d}\left(\mathbb{I} + T(g) + \sum_{j<k} g'_{jk}\Gamma_{jk} + \ldots + g'_0\Gamma_0\right),$$

where we write $T(g)$ to indicate the transformation of the vector $g = \sum_j g_j \Gamma_j$ by T. For example, for the rotation R constructed earlier, we may immediately write

$$\begin{aligned} R\rho R^\dagger &= \frac{1}{d}\left(\mathbb{I} + RgR^\dagger + \sum_{j<k} g_{jk} R\Gamma_{jk}R^\dagger + \ldots + g_0 R\Gamma_0 R^\dagger\right), \\ &= \frac{1}{d}\left(\mathbb{I} + |g|\Gamma_1 + \sum_{j<k} g'_{jk}\Gamma_{jk} + \ldots + g'_0\Gamma_0\right), \end{aligned}$$

Thus, we can think of the 1-vector components of ρ as vectors in a generalized Bloch sphere. In Chapter 4, we will extend this approach to include the Γ_0 as an additional "vector". There, we use these facts to prove a useful statement which leads to our uncertainty relations:

C.3.1. Lemma (Lemma 4.3.2). *For any state ρ, we have $\sum_j g_j^2 \leq 1$.*

With respect to our discussion above, this is indeed a generalization of what we observed for the Bloch sphere in $d = 2$. Note that we obtain a whole range of such statements as we can find different sets of $2n$ anti-commuting matrices within the entire set of 2^{2n} basis elements above.

C.4 Conclusion

Luckily, we made some progress to give a characterization of quantum states in terms of their basis coefficients that was sufficient to prove our uncertainty relation from Chapter 4. Parametrizing states using Clifford algebra elements provides us with additional structure to characterize quantum states that is not at all obvious when looking at them from a linear algebra point of view alone. We hope that parametrizing states in this fashion could enable us to make even stronger statements in the future. It is also interesting to think about standard quantum gates as geometrical operations within the Clifford algebra. Indeed, this is possible to a large extent, but lies outside the scope of this text.

Clearly, the subspace spanned by the elements $\Gamma_1, \ldots, \Gamma_{2n}$ plays a special role. Note that when considering the state minimizing our uncertainty relation, only its 1-vector coefficients played any role. The other coefficients do not contribute at all to the minimization problem. It is interesting to observe that we have in fact already seen a similar effect in Chapter 6. Recall that we used Tsirelson's construction to turn vectors $a, b \in \mathbb{R}^{2n}$ back into observables by letting $A = \sum_j a_j \Gamma_j$ and $B = \sum_j b_j \Gamma_j$. The optimal strategy of Alice and Bob could then be implemented using the maximally entangled state of local dimension $d = 2^n$

$$|\Psi\rangle\langle\Psi| = \frac{1}{d}\left(\mathbb{I} + \sum_j g_j \Gamma_j \otimes \Gamma_j + \sum_j r_j R_j \otimes R_j\right),$$

where $g_j = \pm 1$ and we used the R_j simply as a remainder term. Clearly, the coefficients r_j do not contribute to the term $\langle\Psi|A \otimes B|\Psi\rangle$ at all, and only the coefficients g_j matter. However, in dimension $d = 2^n$ we have only $2n$ such terms. Curiously, the remaining terms are only needed to ensure that $\rho \geq 0$. Numerical feasibility analysis using semidefinite programming for $d = 4$ and $d = 8$ reveals that we do indeed need to take the maximally entangled state, and cannot omit any of the remaining terms.

Bibliography

[ABRD04] A. Ambainis, H. Buhrman, H. Roehrig, and Y. Dodis. Multiparty quantum coin flipping. In *Proceedings of 19th IEEE Conference on Computational Complexity*, pages 250–259, 2004. pages 17

[ADR82] A. Aspect, J. Dalibard, and G. Roger. Experimental test of Bell's inequality using time-varying analyzers. *Physical Review Letters*, 49(25):1804–1807, 1982. pages 107

[AE01] Y. Aharonov and B. G. Englert. The mean king's problem: Prime degrees of freedom. *Physics Letters A*, 284:1–5, 2001. pages 47

[AGR82] A. Aspect, P. Grangier, and G. Roger. Experimental realization of Einstein-Podolsky-Rosen Gedankenexperiment - a new violation of Bell inequalities. *Physical Review Letters*, 49(2):91–94, 1982. pages 107

[Amb01] A. Ambainis. A new protocol and lower bounds for quantum coin flipping. In *Proceedings of 33rd ACM STOC*, pages 134–142, 2001. pages 11, 16

[AMTdW00] A. Ambainis, M. Mosca, A. Tapp, and R. de Wolf. Private quantum channels. In *Proceedings of 41st IEEE FOCS*, pages 547–553, 2000. pages 19

[ANTV99] A. Ambainis, A. Nayak, A. Ta-Shma, and U. Vazirani. Quantum dense coding and a lower bound for 1-way quantum finite automata. In *Proceedings of 31st ACM STOC*, pages 376–383, 1999. quant-ph/9804043. pages 132

[Arv76] W. Arveson. *An invitation to C^*-algebra*. Springer, 1976. pages 193, 196

[AS83] B. Alpern and F. B. Schneider. Key exchange using 'keyless cryptography'. *Information Processing Letters*, 16:79–1, 1983. pages 18

[AS04] A. Ambainis and A. Smith. Small pseudo-random families of matrices: Derandomizing approximate quantum encryption. In *Proceedings of RANDOM 2004*, volume 3122 of *Lecture Notes in Computer Science*, pages 249–260. Springer, 2004. pages 19

[Asp99] A. Aspect. Bell's inequality test: more ideal than ever. *Nature*, 398:189–190, 1999. pages 108

[ATSVY00] D. Aharonov, A. Ta-Shma, U. Vazirani, and A. Yao. Quantum bit escrow. In *Proceedings of the 32th ACM STOC*, pages 705–714, 2000. pages 11, 16

[Aza04] A. Azarchs. Entropic uncertainty relations for incomplete sets of mutually unbiased observables. quant-ph/0412083, 2004. pages 76, 78, 83

[Bab85] L. Babai. Trading group theory for randomness. In *Proceedings of 17th ACM STOC*, pages 421–429, 1985. pages 139

[BB83] C. H. Bennett and G. Brassard. Quantum cryptography and its application to provably secure key expansion, public-key distribution and coin tossing. In *Proceedings of IEEE ISIT 83*, page 91, 1983. pages 3

[BB84] C. H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing*, pages 175–179, 1984. pages 3, 10, 163

[BB89] C. H. Bennett and G. Brassard. The dawn of a new era for quantum cryptography: The experimental prototype is working. *Sigact News*, 20(4):78–82, 1989. pages 4

[BB06] I. Bialynicki-Birula. Formulation of the uncertainty relations in terms of the rényi entropies. *Physical Review A*, 74:052101, 2006. pages 76

[BBB+92] C.H. Bennett, F. Bessette, G. Brassard, L. Salvail, and J. Smolin. Experimental quantum cryptography. *Journal of Cryptology*, 5(1):3–28, 1992. pages 4

[BBBW82] C.H Bennett, G. Brassard, S. Breidbart, and S. Wiesner. Quantum cryptography, or unforgeable subway tokens. In *Advances in Cryptology CRYPTO '82*, pages 267–275, 1982. pages 3, 4, 8, 19

[BBCS92a] C. H. Bennett, G. Brassard, C. Crépeau, and H. Skubiszewska. Practical quantum oblivious transfer. In *Advances in Cryptology — CRYPTO '91*, volume 576 of *Lecture Notes in Computer Science*, pages 351–366. Springer, 1992. pages 11

[BBCS92b] C. H. Bennett, G. Brassard, C. Crépeau, and M.-H. Skubiszewska. Practical quantum oblivious transfer. In *CRYPTO '91: Proceedings of the 11th Annual International Cryptology Conference on Advances in Cryptology*, pages 351–366. Springer-Verlag, 1992. pages 9, 165, 172

[BBF+] G. Brassard, A. Broadbent, J. Fitzsimons, S. Gambs, and A. Tapp. Anonymous quantum communication. arXiv:0706.2356. pages 19

[BBL+06] G. Brassard, H. Buhrman, N. Linden, A. Methot, A. Tapp, and F. Unger. A limit on nonlocality in any world in which communication complexity is not trivial. *Physical Review Letters*, 96:250401, 2006. pages 108

[BBM75] I. Bialynicki-Birula and J. Mycielski. Uncertainty relations for information entropy in wave mechanics. *Communications in Mathematical Physics*, 44:129–132, 1975. pages 76

[BBRV02] S. Bandyopadhyay, P.O. Boykin, V.P. Roychowdhury, and F. Vatan. A new proof for the existence of mutually unbiased bases. *Algorithmica*, 34(4):512–528, 2002. pages 39, 41, 42, 95, 100, 102

[BC90a] G. Brassard and C. Crépeau. Quantum bit commitment and coin tossing protocols. In *Advances in Cryptology – Proceedings of Crypto '90*, pages 49–61, 1990. pages 10

[BC90b] S.L. Braunstein and C.M. Caves. Wringing out better Bell inequalities. *Annals of Physics*, 202:22–56, 1990. pages 121, 122

[BCJL93] G. Brassard, C. Crépeau, R. Jozsa, and D. Langlois. A quantum bit commitment protocol provavly unbreakable by both parties. In *Proceedings of 34th IEEE FOCS*, pages 362–371, 1993. pages 10

[BCMS97] G. Brassard, C. Crépeau, D. Mayers, and L. Salvail. A brief review on the impossibility of quantum bit commitment. quant-ph/9712023, 1997. pages 10

[BCU+06] H. Buhrman, M. Christandl, F. Unger, S. Wehner, and A. Winter. Implications of superstrong nonlocality for cryptography. *Proceedings of the Royal Society A*, 462(2071):1919–1932, 2006. pages 11, 108

[Bel65] J. S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1:195–200, 1965. pages 106

[Ben92] C. Bennett. Quantum cryptography using any two nonorthogonal states. *Physical Review Letters*, 68:3121–3124, 1992. pages 4

[BFL91] L. Babai, L. Fortnow, and C. Lund. Non-deterministic exponential time has two-prover interactive protocols. *Computational Complexity*, 1(1):3–40, 1991. pages 139, 140

[BH05] J. Bergou and M. Hillery. Quantum-state filtering applied to the discrimination of boolean functions. *Physical Review A*, 72:012302, 2005. pages 46

[Bha97] R. Bhatia. *Matrix Analysis*. Springer, 1997. pages 67

[BHH03] J. Bergou, U. Herzog, and M. Hillery. Quantum state filtering and discrimination between sets of boolean functions. *Physical Review Letters*, 90:257901, 2003. pages 46

[BHH04] J. Bergou, U. Herzog, and M. Hillery. Discrimination of quantum states. In M. Paris and J. Rehacek, editors, *Quantum State Estimation*, volume 3, pages 417–465. Springer, Berlin, 2004. pages 47

[BHH05] J. Bergou, U. Herzog, and M. Hillery. Optimal unambiguous filtering of a quantum state: An instance in mixed state discrimination. *Physical Review A*, 71:042314, 2005. pages 46

[BK02] H. Barnum and E. Knill. Reversing quantum dynamics with near-optimal quantum and classical fidelity. *Journal of Mathematical Physics*, 43:2097, 2002. pages 154

[Bla79] G. R. Blakley. Safeguarding cryptography keys. In *Proceedings of the National Computer Conference 48*, pages 313–317, 1979. pages 17

[Blu83] M. Blum. Coin flipping by telephone a protocol for solving impossible problems. *SIGACT News*, 15(1):23–27, 1983. pages 10, 15

[BM05] H. Buhrman and S. Massar. Causality and Cirel'son bounds. *Physical Review A*, 72:052103, 2005. pages 121, 122

[BOGKW88] M. Ben-Or, S. Goldwasser, J. Kilian, and A. Wigderson. Multi prover interactive proofs: How to remove intractability. In *Proceedings of 20th ACM STOC*, pages 113–131, 1988. pages 139, 140

[BOHL+05] M. Ben-Or, M. Horodecki, D.W. Leung, D. Mayers, and J. Oppenheim. The universal composable security of quantum key distribution. In *Proceedings of the 2nd Theory of Cryptography Conference*, volume 3378 of *Lecture Notes in Computer Science*, pages 386–406. Springer, 2005. pages 19, 20

[BOM04] M. Ben-Or and D. Mayers. General security definition and composability for quantum and classical protocols. quant-ph/0409062, 2004. pages 20

[Boy02] P. Boykin. *Information Security and Quantum Mechanics: Security of Quantum Protocols*. PhD thesis, University of California, Los Angeles, 2002. pages 18

[BP03] H. Bechmann-Pasquinucci. Quantum seals. *International Journal of Quantum Information*, 1(2):217–224, 2003. pages 19

[BPDM05] H. Bechmann-Pasquinucci, G.M. D'Ariano, and C. Macchiavello. Impossibility of perfect quantum sealing of classical information. *International Journal of Quantum Information*, 3:435–440, 2005. pages 19

[BR02] O. Bratteli and D. Robinson. *Operator Algebras and Quantum Statistical Mechanics I.* Springer, 2002. pages 193, 196, 198, 201

[BR03] P. O. Boykin and V. Roychowdhury. Optimal encryption of quantum bits. *Physical Review A*, page 042317, 2003. quant-ph/0003059. pages 19

[Bra05] G. Brassard. Brief history of quantum cryptography: A personal perspective. In *Proceedings of IEEE Informtion Theory Workshop on Theory and Practise in Information Theoretic Security*, pages 19–23, 2005. pages 3

[BS05] J. Bouda and J. Sprojcar. Anonymous transmission of quantum information. quant-ph/0512122, 2005. pages 19

[BV04] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. pages 122, 129, 187, 190, 191

[BW07] M. Ballester and S. Wehner. Entropic uncertainty relations and locking: tight bounds for mutually unbiased bases. *Physical Review A*, 75:022319, 2007. pages 80

[Cac97] C. Cachin. *Entropy Measures and Unconditional Security in Cryptography*. PhD thesis, ETH Zurich, Switzerland, 1997. pages 37

[CBH03] R. Clifton, J. Bub, and H. Halvorson. Characterizing quantum theory in terms of information-theoretic constraints. *Foundations of Physics*, 33:1561–1591, 2003. pages 20

[CCL90] J. Cai, A. Condon, and R. Lipton. On bounded round multiprover interactive proof systems. In *Proceedings of the Fifth Annual Conference on Structure in Complexity Theory*, pages 45–54, 1990. pages 139

[CCM98] C. Cachin, C. Crépeau, and J. Marcil. Oblivious transfer with a memory-bounded receiver. In *Proceedings of 39th IEEE FOCS*, pages 493–502, 1998. pages 10

[CG96] R. Canetti and R. Gennaro. Incoercible multiparty computation (extended abstract). In *Proceedings of 37th IEEE FOCS*, pages 504–513, 1996. pages 18

[CGL99] R. Cleve, D. Gottesman, and H-K. Lo. How to share a quantum secret. *Physical Review Letters*, 83(3):648–651, 1999. pages 17

[CGL+02] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu. Bell inequalities for arbitrarily high dimensional systems. *Physical Review Letters*, 88:040404, 2002. pages 109

[CGS02] C. Crépeau, D. Gottesman, and A. Smith. Secure multiparty quantum computation. In *Proceedings of 34th ACM STOC*, 2002. pages 17, 19, 20

[CGS05] C. Crépeau, D. Gottesman, and A. Smith. Approximate quantum error-correcting codes and secret sharing schemes. In *Proceedings of Advances in Cryptology - EUROCRYPT '05*, volume 3494 of *Lecture Notes in Computer Science*, pages 285–301. Springer, 2005. pages 17

[Cha81] D. Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM*, 24(2):84–88, 1981. pages 18

[Cha88] D. Chaum. The dining cryptographers problem: Unconditional sender and recipient untraceability. *Journal of Cryptology*, 1:65–75, 1988. pages 18

[Cha03] H.F. Chau. Sealing quantum meassage by quantum code. quant-ph/0308146, 2003. pages 19

[Chi05] A. Childs. Secure assisted quantum computation. *Quantum Information and Computation*, 5:456, 2005. pages 19

[Chr05] M. Christandl. *The structure of bipartite quantum states - Insights from group theory and cryptography.* PhD thesis, University of Cambridge, 2005. quant-ph/0604183. pages 110, 161

[CHSH69] J. Clauser, M. Horne, A. Shimony, and R. Holt. Proposed experiment to test local hidden-variable theories. *Physical Review Letters*, 23:880–884, 1969. pages 107, 148

[CHTW04a] R. Cleve, P. Høyer, B. Toner, and J. Watrous. Consequences and limits of nonlocal strategies. In *Proceedings of 19th IEEE Conference on Computational Complexity*, pages 236–249, 2004. pages 111, 112, 129, 140, 141, 143, 144, 148

[CHTW04b] R. Cleve, P. Høyer, B. Toner, and J. Watrous. Consequences and limits of nonlocal strategies. Presentation at 19th IEEE Conference on Computational Complexity, 2004. pages 141

[CK88] C. Crépeau and J. Kilian. Achieving oblivious transfer using weakened security assumptions. In *Proceedings of 29th IEEE FOCS*, pages 42–52, 1988. pages 6, 10, 165

[CK06] M. Choi and D. Kribs. A method to find quantum noiseless subsystems. *Physical Review Letters*, 96:050501, 2006. pages 204

[CL98] H.F. Chau and H-K. Lo. Making an empty promise with a quantum computer. *Fortsch. Phys.*, 46:507–520, 1998. Republished in 'Quantum Computing, where do we want to go tomorrow?' edited by S. Braunstein,Wiley-VCH, Berlin, 1999. pages 10

[Cla76] J.F. Clauser. Experimental investigation of a polarization correlation anomaly. *Physical Review Letters*, 36(21):1223–1226, 1976. pages 107

[CMW04] Claude Crépeau, Kirill Morozov, and Stefan Wolf. Efficient unconditional oblivious transfer from almost any noisy channel. In *International Conference on Security in Communication Networks*

(SCN), volume 4 of *Lecture Notes in Computer Science*, 2004. pages 165

[Col07] R. Colbeck. An entanglement-based protocol for strong coin tossing with bias 1/4. *Physics Letters A*, 362(5):309–392, 2007. pages 16

[Con90] J. B. Conway. *A course in functional anlysis*. Springer, 1990. pages 187, 194

[Cra99] R. Cramer. Introduction to secure computation. In *Lectures on Data Security - Modern Cryptography in Theory and Practise*, volume 1561 of *Lecture Notes in Computer Science*, pages 16–62, 1999. pages 17

[Cré94] C. Crépeau. Quantum oblivious transfer. *Journal of Modern Optics*, 41(12):2455–2466, 1994. pages 4, 9, 11

[Cré97] C. Crépeau. Efficient cryptographic protocols based on noisy channels. In *Advances in Cryptology – Proceedings of EUROCRYPT '97*, 1997. pages 165, 172

[CSUU07] R. Cleve, W. Slofstra, F. Unger, and S. Upadhyay. Strong paralell repetition theorem for quantum xor proof systems. In *Proceedings of the 22nd IEEE Conference on Computational Complexity*, pages 109–114, 2007. pages 129

[CvdGT95] C. Crépeau, J. van de Graaf, and A. Tapp. Committed oblivious transfer and private multi-party computation. In *CRYPTO '95: Proceedings of the 15th Annual International Cryptology Conference on Advances in Cryptology*, pages 110–123. Springer-Verlag, 1995. pages 10, 13

[CW79] J. L. Carter and M. N. Wegman. Universal classes of hash functions. *Journal of Computer and System Sciences*, 18:143–154, 1979. pages 155

[CW05a] M. Christandl and S. Wehner. Quantum anonymous transmissions. In *Proceedings of 11th ASIACRYPT*, volume 3788 of *LNCS*, pages 217–235, 2005. pages 18, 20

[CW05b] M. Christandl and A. Winter. Uncertainy, monogamy and locking of quantum correlations. *IEEE Transactions on Information Theory*, 51(9):3159–3165, 2005. pages 93

[Dav78] E. Davies. Information and quantum measurement. *IEEE Transactions on Information Theory*, 24(5):596–599, 1978. pages 38, 96

[Deu83] D. Deutsch. Uncertainty in quantum measurements. *Phys. Rev. Lett.*, 50:631–633, 1983. pages 76

[DFMS04] I. Damgård, S. Fehr, K. Morozov, and L. Salvail. Unfair noisy channels and oblivious transfer. In *Proceedings of TCC 2004*, volume 2951 of *Lecture Notes in Computer Science*, pages 355–373. Springer, 2004. pages 10

[DFR+07] I. Damgård, S. Fehr, R. Renner, L. Salvail, and C. Schaffner. A tight high-order entropic uncertainty relation with applications in the bounded quantum-storage model. Proceedings of CRYPTO 2007, 2007. pages 11, 76, 92, 163, 165, 166, 170

[DFSS05] I. Damgård, S. Fehr, L. Salvail, and C. Schaffner. Cryptography in the Bounded Quantum-Storage Model. In *Proceedings of 46th IEEE FOCS*, pages 449–458, 2005. pages 6, 11, 43, 45, 46, 48, 69, 72, 163, 166

[DFSS07] I. Damgård, S. Fehr, L. Salvail, and C. Schaffner. Secure identification and QKD in the bounded-quantum-storage model. In *Proceedings of CRYPTO 2007*, pages 342–359, 2007. pages 165

[DFSS08] I. B. Damgård, S. Fehr, L. Salvail, and C. Schaffner. Cryptography in the bounded-quantum-storage model. *special issue of SIAM Journal of Computing*, 2008. To appear. pages 165

[DHL+04] D. DiVincenzo, M. Horodecki, D. Leung, J. Smolin, and B. Terhal. Locking classical correlation in quantum states. *Physical Review Letters*, 92(067902), 2004. pages 77, 93, 94, 95, 102

[DHT03] D. P. DiVincenzo, P. Hayden, and B. Terhal. Hiding quantum data. *Foundations of Physics*, 33(11):1629–1647, 2003. pages 17

[Die06] K. Dietz. Generalized bloch spheres for m-qubit states. *Journal of Physics A: Math. Gen.*, 36(6):1433–1447, 2006. pages 214

[DKS99] I. Damgård, J. Kilian, and L. Salvail. On the (im)possibility of basing oblivious transfer and bit committment on weakened security assumptions. In *Advances in Cryptology - EUROCRYPT '99*, volume 1592 of *Lecture Notes in Computer Science*, pages 56–73. Springer, 1999. pages 10

[DKSW06] G. D'Ariano, D. Kretschmann, D. Schlingemann, and R.F. Werner. Quantum bit commitment revisited: the possible and the impossible. quant-ph/0605224, 2006. pages 10

[DL03] C. Doran and A. Lasenby. *Geometric algebra for physicists*. Cambridge University Press, 2003. pages 88, 209

[DLT02] D. P. DiVincenzo, D. W. Leung, and B. M. Terhal. Quantum data hiding. *IEEE Trans. Inf Theory*, 48(3):580–599, 2002. arXiv e-print quant-ph/0103098. pages 17, 38

[DLTW08] A. C. Doherty, Y-C. Liang, B. Toner, and S. Wehner. The quantum moment problem. Submitted., 2008. pages 130, 131

[DN06] P. A. Dickinson and A. Nayak. Approximate randomization of quantum states with fewer bits of key. In *Quantum Computing Back Action, IIT Kanpur*, volume 864 of *AIP Conference Proceedings*, pages 18–36. Springer, 2006. pages 19

[DPS02] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri. Distinguishing separable and entangled states. *Physical Review Letters*, 88(18):187904, 2002. pages 109

[DPS04] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri. A complete family of separability criteria. *Physical Review A*, 69:022308, 2004. pages 109

[DPS05] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri. Detecting multipartite entanglement. *Physical Review A*, 71:032333, 2005. pages 109

[EF01] Y. Eldar and G. Forney. On quantum detection and the square-root measurement. *IEEE Transactions on Information Theory*, 47:858–872, 2001. pages 46

[EGL85] S. Even, O. Goldreich, and A. Lempel. A randomized protocol for signing contracts. *Communications of the ACM*, 28(6):637–647, 1985. pages 12

[Eis01] J. Eisert. *Entanglement in quantum information theory*. PhD thesis, University of Potsdam, 2001. quant-ph/0610253. pages 110

[Eke91] A. Ekert. Quantum cryptography based on Bell's theorem. *Physical Review Letters*, 67:661–663, 1991. pages 4

[Eld03] Y. Eldar. A semidefinite programming approach to optimal unambiguous discrimination of quantum states. *IEEE Transactions on Information Theory*, 49:446–456, 2003. pages 35, 46

[EMV03] Y. Eldar, A. Megretski, and G. Verghese. Designing optimal quantum detectors via semidefinite programming. *IEEE Transactions on Information Theory*, 49:1017–1012, 2003. pages 46

[EMV04] Y. Eldar, A. Megretski, and G. Verghese. Optimal detection of symmetric mixed quantum states. *IEEE Transactions on Information Theory*, 50:1198–1207, 2004. pages 46

[EPR35] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47:777–780, 1935. pages 105

[EW02] T. Eggeling and R.F. Werner. Hiding classical data in multi-partite quantum states. *Physical Review Letters*, 89(9):097905, 2002. pages 17

[Feh07] S. Fehr. Personal communication, 2007. pages 92

[Fei91] U. Feige. On the success probability of two provers in one-round proof systems. In *Proceedings of the Sixth Annual Conference on Structure in Complexity Theory*, pages 116–123, 1991. pages 139

[Fei95] U. Feige. Error reduction by parallel repetition - the state of the art. Technical Report CS95-32, Weizmann Institute, 1, 1995. pages 144

[FL92] U. Feige and L. Lovász. Two-prover one-round proof systems: their power and their problems. In *Proceedings of 24th ACM STOC*, pages 733–744, 1992. pages 139, 140, 143

[FS04] S. Filipp and K. Svozil. Tracing the bounds on Bell-type inequalities. In *Proceedings of Foundations of Probability and Physics-3*, pages 87–94, 2004. pages 121

[Fuc95] C. A. Fuchs. *Distinguishability and Accessible Information in Quantum Theory*. PhD thesis, University of New Mexico, Albuquerque, 1995. quant-ph/9601020. pages 35, 38

[GC01] D. Gottesman and I. Chuang. Quantum signatures. quant-ph/0105032, 2001. pages 19

[Gis91] N. Gisin. Bell's inequality holds for all non-product states. *Physics Letters A*, 154:201–202, 1991. pages 109

[Gis99] N. Gisin. Bell inequality for arbitrary many settings of the analyzers. *Physics Letters A*, 260:1–3, 1999. pages 122, 128

[GKK+06] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf. Exponential separations for one-way quantum communication complexity. In *Proceedings of 39th ACM STOC*, pages 516–525, 2006. pages 20, 47

[GMR89] S. Goldwasser, S. Micali, and C. Rackoff. The knowledge complexity of interactive proof systems. *SIAM Journal on Computing*, 1(18):186–208, 1989. pages 139

[Gol01] O. Goldreich. *Foundations of Cryptography*, volume Basic Tools. Cambridge University Press, 2001. pages 10, 15, 16, 17

[Got00] D. Gottesman. On the theory of quantum secret sharing. *Physical Review A*, 61:042311, 2000. pages 17

[Gra71] R. M. Gray. *Toeplitz and Circulant Matrices: A review*. 1971. pages 126

[Gra04] M. Grassl. On SIC-POVMs and MUBs in dimension 6. In *Proceedings ERATO Conference on Quantum Information Science*, pages 60–61, 2004. pages 39

[GRTZ02] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden. Quantum cryptography. *Reviews of Modern Physics*, 74:pp. 145–195, 2002. pages 163

[Gur03] L. Gurvits. Classical deterministic complexity of Edmund's problem and quantum entanglement. In *Proceedings of 35th ACM STOC*, pages 10–19, 2003. pages 109

[GW95] M.X. Goemans and D.P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. Assoc. Comput. Mach.*, 42:1115–1145, 1995. pages 122

[Hay06] M. Hayashi. *Quantum Information - An introduction*. Springer, 2006. pages 31, 33, 35, 42, 133, 178

[HBB99] M. Hillery, V. Buzek, and A. Bethiaume. Quantum secret sharing. *Physical Review A*, 3:1829–1834, 1999. pages 17

[Hei27] W. Heisenberg. Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik. *Zeitschrift für Physik*, 43:172–198, 1927. pages 75

[Hel67] C. W. Helstrom. Detection theory and quantum mechanics. *Information and Control*, 10(1):254–291, 1967. pages 33, 169

[HHHO05] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim. Locking entanglement measures with a single qubit. *Physical Review Letters*, 94:200501, 2005. pages 93

[HJ85] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985. pages 68, 123, 124, 127, 179, 189

[HK04] L. Hardy and A. Kent. Cheat sensitive quantum bit commitment. *Physical Review Letters*, 92(157901), 2004. pages 11

[HKL03] J.A. Holbrook, D. W. Kribs, and R. Laflamme. Noiseless subsystems and the structure of the commutant in quantum error correction. *Quantum Information and Computation*, 2(5):381–419, 2003. pages 31, 203

[HLS05] P. Hayden, D. Leung, and G. Smith. Multiparty data hiding of quantum information. *Physical Review A*, 71:062339, 2005. pages 17

[HLSW04] P. Hayden, D. Leung, P. Shor, and A. Winter. Randomizing quantum states: Constructions and applications. *Communications in Mathematical Physics*, 250(2):371–391, 2004. pages 77, 94, 160

[Hol73] A. S. Holevo. Information theoretical aspects of quantum measurements. *Probl. Inf. Transm.*, 9:110–118, 1973. pages 38

[HR07] I. Haitner and O. Reingold. Statistically-hiding commitment from any one-way function. In *Proceedings of 39th ACM STOC*, pages 1–10, 2007. pages 10

[HRP+06] P.A. Hiskett, D. Rosenberg, C.G. Peterson, R.J. Hughes, S. Nam, A.E. Lita, A.J. Miller, and J.E. Nordholt. Long-distance quantum key distribution in optical fibre. *New Journal of Physics*, 8:193, 2006. pages 4

[Hun03] K. Hunter. Measurement does not always aid state discrimination. *Physical Review A*, 6:012306, 2003. pages 54

[HW94] P. Hausladen and W. Wootters. A pretty good measurement for distinguishing quantum states. *Journal of Modern Optics*, 41:2385–2390, 1994. pages 46

[Ioa07] L. Ioannou. *Computing finite-dimensional bipartite quantum separability*. PhD thesis, University of Cambridge, 2007. cs/0504110. pages 109

[IT06] L. M. Ioannou and B. C. Travaglione. Quantum separability and entanglement detection via entanglement-witness search and global optimization. *Physical Review A*, 73:052314, 2006. pages 109

[ITCE04] L. M. Ioannou, B. C. Travaglione, D. Cheung, and A. Ekert. Improved algorithm for quantum separability and entanglement detection. *Physical Review A*, 70:060303, 2004. pages 109

[Jai05] R. Jain. Stronger impossibility results for quantum string commitment. quant-ph/0506001, 2005. pages 161

[JW28] P. Jordan and E. Wigner. Über das paulische äquivalenzverbot. *Zeitschrift für Physik*, 47:631, 1928. pages 205

[Kah96] D. Kahn. *The Codebreakers: The comprehensive history of secret communication from ancient times to the internet*. Simon and Schuster, 1996. pages 3

[Ken99] A. Kent. Secure classical bit commitment using fixed capacity communication channels. *Journal of Cryptology*, 18(4):313–335, 1999. pages 11

[Ken03] A. Kent. Quantum bit string commitment. *Physical Review Letters*, 90(237901), 2003. pages 151

[KI02] M. Koashi and N. Imoto. Operations that do not disturb partially known quantum states. *Physical Review A*, 66:022318, 2002. pages 46, 65, 203

[Kil88] J. Kilian. Founding cryptography on oblivious transfer. In *Proceedings of 20th ACM STOC*, pages 20–31, 1988. pages 13

[Kit02] A. Kitaev. Quantum coin flipping. Talk at QIP 2002, 2002. pages 15

[KKM+07] J. Kempe, H. Kobayashi, K. Matsumoto, B. Toner, and T. Vidick. On the power of entangled provers: immunizing games against entanglement. arXiv:0704.2903, 2007. pages 130

[KLM01] E. Knill, R. Laflamme, and G. Milburn. A scheme for efficient quantum computation with linear optics. *Nature*, 409:46–52, 2001. pages 185

[KM03] H. Kobayashi and K. Matsumoto. Quantum multi-prover interactive proof systems with limited prior entanglement. *Journal of Computer and Systems Sciences*, 66(3):429–450, 2003. pages 141

[KMP04] A. Kitaev, D. Mayers, and J. Preskill. Superselection rules and quantum protocols. *Physical Review A*, 69:052326, 2004. pages 7, 10, 159

[KMR05] Robert Koenig, Ueli Maurer, and Renato Renner. On the power of quantum memory. *IEEE Transactions on Information Theory*, 51(7):2391–2401, 2005. pages 155

[KN04] I. Kerenidis and A. Nayak. Weak coin flipping with small bias. *Information Processing Letters*, 89(3):131–135, 2004. pages 16

[KR03] A. Klappenecker and M. Rötteler. Constructions of mutually unbiased bases. In *Finite Fields and Applications: 7th international conference Fq7*, pages 137–144. Lecture Notes in Computer Science, 2003. pages 39

[KR04] A. Klappenecker and M. Rötteler. Constructions of mutually unbiased bases. In *International Conference on Finite Fields and Applications (Fq7)*, volume 2948 of *Lecture Notes in Computer Science*, pages 137–144. Springer, 2004. pages 39

[KR05] A. Klappenecker and M. Rötteler. New Tales of the Mean King. quant-ph/0502138, 2005. pages 47

[Kra87] K. Kraus. Complementary observables and uncertainty relations. *Physical Review D*, 35(10):3070–3075, 1987. pages 76

[KRBM07] R. Koenig, R. Renner, A. Bariska, and U. Maurer. Small accessible quantum information does not imply security. *Physical Review Letters*, 98:140502, 2007. pages 93

[KT87] L.A. Khalfin and B.S. Tsirelson. A quantitative criterion of the applicability of the classical description within the quantum theory. In *Symposium on the Foundations of Modern Physics*, pages 369–401, 1987. pages 109

[KW00] A. Kitaev and J. Watrous. Parallelization, amplification, and exponential time simulation of quantum interactive proof systems. In *Proceedings of 32nd ACM STOC*, pages 608–617, 2000. pages 140, 141, 147

[KW03] I. Kerenidis and R. de Wolf. Exponential lower bound for 2-query locally decodable codes via a quantum argument. In *Proceedings of 35th ACM STOC*, pages 106–115, 2003. pages 145

[Lan87] L.J. Landau. On the violation of bell's inequality in quantum theory. *Physics Letters A*, 123(3):115–118, 1987. pages 109

[Lan88] L.J. Landau. Empirical two-point correlation functions. *Foundations of Physics*, 18(4):449–460, 1988. pages 121, 130

[Lar90] U. Larsen. Superspace geometry: the exact uncertainty relationship between complementary aspects. *J. Phys. A: Math. Gen.*, 23:1041–1061, 1990. pages 76, 80

[LBZ02] J. Lawrence, C. Brukner, and A. Zeilinger. Mutually unbiased binary observable sets on n qubits. *Physical Review A*, 65:032320, 2002. pages 42

[LC96] H-K. Lo and H.F. Chau. Why quantum bit commitment and ideal quantum coin tossing are impossible. In *Proceedings of PhysComp96*, 1996. quant-ph/9605026. pages 10

[LC97] H-K. Lo and H. F. Chau. Is quantum bit commitment really possible? *Physical Review Letters*, 78:3410, 1997. pages 10, 153, 155

[LC98] H-K. Lo and H.F. Chau. Why quantum bit commitment and ideal quantum coin tossing are impossible. *Physica D*, 120:177–187, 1998. pages 15

[LC99] H-K. Lo and H. F. Chau. Unconditional security of quantum key distribution over arbitrarily long distances. *Science*, 283:2050–2056, 1999. pages 4

[LD06] Y-C. Liang and A.C. Doherty. Better bell inequality violation by collective measurements. *Physical Review A*, 73:052116, 2006. pages 129

[LD07] Y-C. Liang and A.C. Doherty. Bounds on quantum correlations in bell inequality experiments. *Physical Review A*, 75:042103, 2007. pages 109, 130

[Lin99] G. Lindblad. A general no-cloning theorem. *Letters in Mathematical Physics*, 47:189–196, 1999. pages 203

[Lo97] H-K. Lo. Insecurity of quantum secure computations. *Physical Review A*, 56:1154, 1997. pages 12, 14, 161

[Lou01] P. Lounesto. *Clifford Algebras and Spinors*. Cambridge University Press, 2001. pages 205, 209

[LS91] D. Lapidot and A. Shamir. Fully parallelized multi prover protocols for NEXP-time. In *Proceedings of 32nd FOCS*, pages 13–18, 1991. pages 139

[Mas06] L. Masanes. Asymptotic violation of bell inequalities and distillability. *Physical Review Letters*, 97:050503, 2006. pages 67, 72, 137

[Mau92] U. Maurer. Conditionally-perfect secrecy and a provably-secure randomized cipher. *Journal of Cryptology*, 5(1):53–66, 1992. pages 6, 10

[May96a] D. Mayers. Quantum key distribution and string oblivious transfer in noisy channels. In *Proceedings of Advances in Cryptology - CRYPTO '96*, pages 343–357, 1996. pages 4

[May96b] D. Mayers. The trouble with quantum bit commitment. quant-ph/9603015, 1996. pages 10, 155

[May97] D. Mayers. Unconditionally secure quantum bit commitment is impossible. *Physical Review Letters*, 78:3414–3417, 1997. quant-ph/9605044. pages 10

[Moc04] C. Mochon. Quantum weak coin-flipping with bias of 0.192. In *Proceedings of 45th IEEE FOCS*, pages 2–11, 2004. pages 16

[Moc05] C. Mochon. A large family of quantum weak coin-flipping protocols. *Physical Review A*, 72:022341, 2005. pages 16

[Moc07a] C. Mochon. A family of generalized 'pretty good' measurements and the minimal-error pure-state discrimination problems for which they are optimal. *Physical Review A*, 75:042313, 2007. pages 46

[Moc07b] C. Mochon. Quantum weak coin flipping with arbitrarily small bias. 2007. arXiv:0711.4114. pages 16

[MSC99] D. Mayers, L. Salvail, and Y. Chiba-Kohno. Unconditionally secure quantum coin tossing. quant-ph/9904078, 22 Apr 1999. pages 15

[MU88] H. Maassen and J. Uffink. Generalized entropic uncertainty relations. *Physical Review Letters*, 60(1103), 1988. pages 76, 78, 83

[MvOV97] A. Menezes, P. van Oorschot, and S. Vanstone. *Handbook of Applied Cryptography*. CRC Press, 1997. pages 17

[MW05] C. Marriott and J. Watrous. Quantum Arthur-Merlin games. cs.CC/0506068, 2005. pages 140

[Nao91] M. Naor. Bit commitment using pseudorandomness. *Journal of Cryptology*, 4(2):151–158, 1991. pages 10

[Nay99] A. Nayak. Optimal lower bounds for quantum automata and random access codes. In *Proceedings of 40th IEEE FOCS*, pages 369–376, 1999. quant-ph/9904093. pages 132, 133

[NC00] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000. pages 35, 36, 37, 38, 42, 105

[NPA07] M. Navascues, S. Pironio, and A. Acin. Bounding the set of quantum correlations. *Physical Review Letters*, 98:010401, 2007. pages 130, 131

[Per93] A. Peres. *Quantum Theory: Concepts and Methods*. Kluwer Academic Publishers, 1993. pages 25, 38, 42, 105, 122, 124, 125, 126, 128

[Per96] A. Peres. Collective tests for quantum nonlocality. *Physical Review A*, 54:2685, 1996. pages 129

[PR94] S. Popescu and D. Rohrlich. Quantum nonlocality as an axiom. *Foundations of Physics*, 24(3):379–385, 1994. pages 108, 141

[PR96] S. Popescu and D. Rohrlich. Nonlocality as an axiom for quantum theory. In *The dilemma of Einstein, Podolsky and Rosen, 60 years later: International symposium in honour of Nathan Rosen*, 1996. pages 108, 141

[PR97] S. Popescu and D. Rohrlich. Causality and nonlocality as axioms for quantum mechanics. In *Proceedings of the Symposium of Causality and Locality in Modern Physics and Astronomy: Open Questions and Possible Solutions*, 1997. pages 108, 141

[Pre05] D. Preda. Non-local multi-prover interactive proofs. CWI Seminar, 21 June, 2005. pages 141

[Qua] ID Quantique. http://www.idquantique.com. pages 4

[Rab81] M. Rabin. How to exchange secrets by oblivious transfer. Technical report, Aiken Computer Laboratory, Harvard University, 1981. Technical Report TR-81. pages 12, 13

[Raz05] R. Raz. Quantum information and the PCP theorem. In *Proceedings of 46th FOCS*, pages 459–468, 2005. pages 140

[Reg03] O. Regev. New lattice-based cryptographic constructions. *Journal of the ACM*, 51(6):899–942, 2003. pages 7

[Rén60] A. Rényi. On measures of information and entropy. In *Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability*, pages 547–561, 1960. pages 37

[Ren05] R. Renner. *Security of Quantum Key Distribution*. PhD thesis, ETH Zurich, 2005. quant-ph/0512258. pages 4, 19, 154, 155, 167, 185

[Riv99] R. L. Rivest. Unconditionally secure commitment and oblivious transfer schemes using private channels and a trusted initializer. http://people.csail.mit.edu/rivest/Rivest-commitment.pdf, 1999. pages 6

[RK05] R. Renner and R. Koenig. Universally composable privacy amplification against quantum adversaries. In *Proceedings of TCC 2005*, volume 3378 of *Lecture Notes in Computer Science*, pages 407–425. Springer, 2005. pages 19

[RKM+01] M. A. Rowe, D. Kielpinski, V. Meyer, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Wineland. Experimental violation of a bell's inequality with efficient detection. *Nature*, 409:791–794, 2001. pages 107

[Rob29] H.P. Robertson. The uncertainty principle. *Physical Review*, 34:163–164, 1929. pages 75

[SA] J. Sturm and AdvOL. SeDuMi. http://sedumi.mcmaster.ca/. pages 124

[SA99] F. Stajano and R. J. Anderson. The cocaine auction protocol: On the power of anonymous broadcast. In *Information Hiding*, pages 434–447, 1999. pages 18

[Sal98] L. Salvail. Quantum bit commitment from a physical assumption. In *Proceedings of CRYPTO'98*, volume 1462 of *Lecture Notes in Computer Science*, pages 338–353, 1998. pages 7, 11, 165

[San93] J. Sanchez. Entropic uncertainty and certainty relations for complementary observables. *Physics Letters A*, 173:233–239, 1993. pages 76, 80, 83, 92

[Sch35a] E. Schrödinger. Die gegenwärtige Situation der Quantenmechanik. *Naturwissenschaften*, 23:807,823,840, 1935. pages 106

[Sch35b] E. Schrödinger. Discussion of probability relations between separated systems. *Proceedings of the Cambridge Philosophical Society*, 31:555–563, 1935. pages 106

[Sch07] C. Schaffner. *Cryptography in the Bounded-Quantum Storage Model*. PhD thesis, University of Aarhus, 2007. pages 165, 168, 172, 173

[Sha48] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 623–656, 1948. pages 76

[Sha79] A. Shamir. How to share a secret. *Communications of the ACM*, 22(2):612–613, 1979. pages 17

[Sha92] A. Shamir. IP = PSPACE. *Journal of the ACM*, 39(4):869–877, 1992. pages 140

[She92] A. Shen. IP = PSPACE: simplified proof. *Journal of the ACM*, 39(4):878–880, 1992. pages 140

[Sho97] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997. Earlier version in FOCS'94. pages 4, 7

[SIGA05] V. Scarani, S. Iblisdir, N. Gisin, and A. Acín. Quantum cloning. *Reviews in Modern Physics*, 77:1225, 2005. pages 27

[SO06] J. Smolin and J. Oppenheim. Information locking in black holes. *Physical Review Letters*, 96:091302, 2006. pages 93

[SP00] P.W. Shor and J. Preskill. Simple proof of security of the bb84 quantum key distribution protocol. *Physical Review Letters*, 85(2):441–444, 2000. pages 4

[SR95] J. Sanchez-Ruiz. Improved bounds in the entropic uncertainty and certainty relations for complementary observables. *Physics Letters A*, 201:125–131, 1995. pages 76, 80, 83, 92

[SR02a] R. Spekkens and T. Rudolph. Degrees of concealment and bindingness in quantum bit commitment protocols. *Physical Review A*, 65(012310), 2002. pages 11, 16

[SR02b] R. Spekkens and T. Rudolph. A protocol for cheat-sensitive weak coin flipping. *Physical Review Letters*, 89:227901, 2002. pages 16

[SS05] S. Singh and R. Srikanth. Quantum seals. *Physica Scripta*, 71:433, 2005. pages 19

[Sum90] S.J. Summers. On the independence of local algebras in quantum field theory. *Reviews in Mathematical Physics*, 2(2):201–247, 1990. pages 202

[Tak79] M. Takesaki. *Theory of Operator Algebras I.* Springer, 1979. pages 193, 195, 196, 198, 199, 200, 201, 202

[Tec] MagicQ Technologies. http://www.magicqtech.com. pages 4

[Ter99] B. M. Terhal. *Quantum Algorithms and Quantum Entanglement.* PhD thesis, CWI and University of Amsterdam, 1999. pages 110

[THLD02] B. Terhal, M. Horodecki, D.W. Leung, and D.P.DiVincenzo. The entanglement of purification. *J. Math. Phys.*, 43:4286–4298, 2002. pages 36

[Tsi80] B. Tsirelson. Quantum generalizations of Bell's inequality. *Letters in Mathematical Physics*, 4:93–100, 1980. pages xvi, 108, 113, 115, 122, 124, 141, 143, 148

[Tsi87] B. Tsirelson. Quantum analogues of Bell inequalities: The case of two spatially separated domains. *Journal of Soviet Mathematics*, 36:557–570, 1987. pages xvi, 113, 115, 122, 124, 141, 143, 148, 205

[Tsi93] B. Tsirelson. Some results and problems on quantum Bell-type inequalities. *Hadronic Journal Supplement*, 8(4):329–345, 1993. pages 115, 121, 122, 124, 128

[TV06] B. Toner and F. Verstraete. Monogamy of bell correlations and tsirelson's bound. quant-ph/0611001, 2006. pages 67, 72

[Uhl76] A. Uhlmann. The "transition probability" in the state space of a $*$-algebra. *Rep. Math. Phys.*, 9(2):273–279, 1976. pages 155

[Unr04] D. Unruh. Simulatable security for quantum protocols. quant-ph/0409125, 2004. pages 20

[UTSM$^+$] R. Ursin, F. Tiefenbacher, T. Schmitt-Manderbach, H. Weier, T. Scheidl, M. Lindenthal, B. Blauensteiner, T. Jennewein, J. Perdigues, P. Trojek, B. Oemer, M. Fuerst, M. Meyenburg, J. Rarity, Z. Sodnik, C. Barbieri, H. Weinfurter, and A. Zeilinger. Free-space distribution of entanglement and single photons over 144 km. quant-ph/0607182. pages 4

[VB96] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM review*, 38:49, 1996. pages 168

[vD00] W. van Dam. *Nonlocality & Communication Complexity.* PhD thesis, University of Oxford, Department of Physics, 2000. pages 108, 141

[vD05] W. van Dam. Impossible consequences of superstrong nonlocality. quant-ph/0501159, 2005. pages 108

[vdG98] J. van de Graaf. *Towards a formal definition of security for quantum protocols*. PhD thesis, partment d'informatique et de r.o., Universit de Montral, 1998. http://www.cs.mcgill.ca/ crepeau/PS/thesejeroen.ps. pages 19

[Vya03] M. Vyalyi. QMA=PP implies that PP contains PH. *Electronic Colloquium on Computational Complexity*, TR03-021, 2003. pages 140

[Wat99] J. Watrous. PSPACE has constant-round quantum interactive proof systems. In *Proceedings of 40th IEEE FOCS*, pages 112–119, 1999. cs.CC/9901015. pages 140, 145

[WB05] P. Wocjan and T. Beth. New construction of mutually unbiased bases in square dimensions. *Quantum Information and Computation*, 5(2):93–101, 2005. pages 39, 40, 77, 80

[WdW05] S. Wehner and R. de Wolf. Improved lower bounds for locally decodable codes and private information retrieval. In *Proceedings of the 32nd ICALP*, volume 3580 of *LNCS*, pages 1424–1436, 2005. pages 145

[Wer81] R.F. Werner. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Physical Review A*, 40:4277–4281, 1981. pages 110

[WF89] W.K. Wootters and B. Fields. Optimal state-determination by mutually unbiased measurements. *Ann. Phys.*, 191(368), 1989. pages 39, 41

[Wie83] S. Wiesner. Conjugate coding. *Sigact News*, 15(1), 1983. pages 3, 12, 19

[Wik] Wikibooks. History of cryptography. http://wikibooks.org/wiki/Cryptography:History_of_Cryptography. pages 3

[Wul07] J. Wullschleger. Oblivious-transfer amplification. In *Advances in Cryptology — EUROCRYPT '07*, Lecture Notes in Computer Science, pages 555–572. Springer, 2007. pages 14

[WW01a] R.F. Werner and M.M. Wolf. All-multipartite bell-correlation inequalities for two dichotomic observables per site. *Physical Review A*, 64:032112, 2001. pages 122

[WW01b] R.F. Werner and M.M. Wolf. Bell inequalities and entanglement. *Quantum Information and Computation*, 1(3), 2001. pages 110

[WW07] S. Wehner and J. Wullschleger. Security in the bounded quantum storage model. arxiv:0710.1185, 2007. pages 11, 20, 166, 185

[WY06] M. Wang and F. Yan. Conclusive quantum state classification. quant-ph/0605127, 2006. pages 46

[Yao82] A. C. Yao. Protocols for secure computations. In *Proceedings of the 23rd Annual IEEE FOCS*, pages 160–164, 1982. pages 11

[Yao95] A. C.-C. Yao. Security of quantum protocols against coherent measurements. In *Proceedings of 20th ACM STOC*, pages 67–75, 1995. pages 9, 153

[YKL75] H. P. Yuen, R. S. Kennedy, and M. Lax. Optimum testing of multiple hypotheses in quantum detection theory. *IEEE Transactions on Information Theory*, 21, 1975. pages 34

[Zau99] G. Zauner. *Quantendesigns - Grundzüge einer nichtkommutativen Designtheorie*. PhD thesis, Universität Wien, 1999. pages 39

[ZLG00] Y-S. Zhang, C-F. Li, and G-C. Guo. Unconditionally secure quantum coin tossing via entanglement swapping. quant-ph/0012139, 2000. pages 15

Index

$*$-algebra, 194
C^*-Algebra, 193, 194
$H_2(\cdot \mid \cdot)$, 154
$\oplus$MIP[2], 144
$\oplus$MIP*, 141
$\oplus$MIP$_{\text{NL}}$, 141
$\oplus$MIP*[2], 144, 145
k-vector, 207
1-2 OT, 13
    practical, 171
1-2 oblivious transfer, 13

accessible information, 38
adversary, 6
adversary model, 164
algebra
    $*$-algebra, 194
    C^*-algebra, 194
    Banach, 194
    center, 198
    commutant, 198
    GNS construction, 196
    ideal, 195
    invertible, 195
    observable, 195
    positive linear functional, 195
    pure state, 195
    semisimple, 195
    simple, 195
    spectrum, 195
    state, 195
    state space, 196
    unital, 195
anonymous transmissions, 18
anti-commuting observables, 83, 205

Banach algebra, 194
basis, 187
    Breidbart, 169
    computational, 26
    Hadamard, 26
    measurement in, 30
    MUB, 39, 81
    mutually unbiased, 39
basis vectors, 187
BB84, 3, 163
BC, 10
Bell inequalities, 110
Bell inequality, 106
    CHSH, 107
binary entropy, 36
bipartite state, 27
    mutual information, 36
bipartite structure, 200
bit commitment, 9, 10
    impossibility, 155
    limitations, 151
bit string commitment, 11
    impossibility, 156
    possibility, 159

Bloch sphere, 33
Bloch vector, 33
bounded operator, 188
bounded storage
  classical, 6
  composability, 20
  due to noise, 163
  quantum, 11, 43, 163
Boykin, P. O., 18
Breidbart basis, 169

center, 198
cheat-sensitive, 8, 11, 31
CHSH inequality, 107, 123
  generalized, 125
classical ensemble, 27
classical information, 31
classical mutual information, 36
Clifford algebra, 83, 205
  generators, 205
  geometrical properties, 206
  product, 206
  uncertainty relation, 83
Clifford product, 206
coin tossing
  strong, 15
  weak, 16
collision entropy, 37
commitment
  bit, 10
  quantum bit string, 152
commutant, 198
compatible, 198
complementary, 198
complete, 105
complex Hadamard matrix, 40
composability, 20
  bounded storage, 20
computational basis, 26
computational security, 7
conditional entropy, 36
cone, 189
conjugate transpose, 188
convex cone, 189
CP map, 30
CPTP map, 31
cq-state, 166
cryptography, 3

data hiding, 17
density matrix, 25
density operator, 25
depolarizing channel, 174
depolarizing noise, 174
dishonest, 6
distance
  from uniform, 165
  trace, 32
  variational, 35
distinguishability, 32
disturb, 31
dual, 190
dual function, 190
duality, 190
  strong, 191
  weak, 190

E91, 4
eigendecomposition, 188
eigenvalue, 187
eigenvector, 187
encoding, 26
ensemble, 26
  classical, 27
entanglement, 8, 105
entropy
  binary, 36
  collision, 37
  conditional, 36
  joint, 36
  min-entropy, 37
  Rényi, 37
  Shannon, 36
  von Neumann, 37
EPR, 105
  pair, 106

Paradox, 105
state, 106

feasible, 190
fidelity, 35

games
non-local, 110
optimal strategies, 121
superposition, 148
XOR, 112
XOR, winning probability, 143
generalized Pauli matrices, 41
GNS construction, 196

Hadamard basis, 26
Hadamard transform, 29
Hermitian, 188
hidden variables, 106
Hilbert space, 187
bounded operators, 194
Holevo quantity, 38
honest, 5
honest-but-curious, 5, 32

ideal, 195
impossibility
bit commitment, 155
bit string commitment, 156
information
accessible, 38
mutual, 36
information theoretic security, 6
inner product, 187
instrument, 31
interactive proof system, 139
invertible, 195
IP, 140

joint entropy, 36
Jordan Wigner transformation, 205

K-transform, 29
Kraus operator, 30

Lagrangian, 190
Latin square, 39
mutually orthogonal, 39
left ideal, 195
LOCKCOM, 159
locking, 8, 93
application, 159
mutually unbiased bases, 96
protocol, 94
uncertainty relations, 95

measurement, 29
delay, 8
in a basis, 30
observable, 29
operators, 29
post-measurement state, 29
POVM, 30
projective, 29
min-entropy, 37
MIP, 140
MIP[2], 140
mixed state, 26
mixture, 26
MOLS, 39
MUB, 39
from Latin squares, 39
from Pauli matrices, 41
mutual information, 36
classical, 36
mutually unbiased basis, 39
from Latin squares, 39
from Pauli matrices, 41

no-cloning, 7, 27
noise
depolarizing, 174
noisy channel, 6
noisy quantum storage model, 163
noisy storage, 163
non-local games, 110
non-uniformity, 165

oblivious transfer, 163

1-2, 13
practical, 171
Rabin, 13
observable, 29, 195
algebra, 195
anti-commuting, 83
compatible, 198
complementary, 198
operator norm, 188
orthogonal projector, 189
OT, 163
1-2, 13
Rabin, 13
outer product, 207

Pauli matrices, 29
generalized, 41
strings of, 41
perfect secrecy, 6
physical reality, 105
PI-STAR, 47, 116
algebraic framework, 63
AND, 57
gap, 116
lower bound, 54
no quantum memory, 54
perfect prediction, 63
quantum memory, 63
three bases, 70
to succeed at, 48
two bases, 66
XOR, 60
player, 5
positive definite, 189
positive linear functional, 195
positive semidefinite, 189
post-measurement information, 43, 47
algebraic framework, 63
AND, 57
gap, 116
lower bound, 54
no quantum memory, 54
perfect prediction, 63
quantum memory, 63
three bases, 70
two bases, 66
XOR, 60
post-measurement state, 29
POVM, 30
primal, 190
privacy amplification, 154, 155
private input, 5
private shared input, 5
private shared randomness, 5
projective measurement, 29
projector, 189
orthogonal, 189
public information, 5
pure state, 25
purification, 28

QBSC, 152
impossibility, 156
possibility, 159
QIP, 140
QIP(2), 141, 145
QKD, 3, 9
quantum bit string commitment, 152
quantum channel, 31
trace preserving, 30
unital, 31
quantum key distribution, 3, 9
quantum seals, 19
quantum signatures, 19
quantum strategy, 121
qubit, 25

RAC, 132
random access code, 132
unbalanced, 133
randomized oblivious transfer, 166
rank, 188
reduced state, 28
right ideal, 195
ROT, 166
Rényi entropy, 37

SDP, 190
  dual, 190
  dual function, 190
  duality, 190
  example, 35, 57, 59, 61
  feasible, 190
  Lagrangian, 190
  primal, 190
  Slater's conditions, 191
  standard form, 190
  strong duality, 191
  weak duality, 190
secret
  classical, 17
  quantum, 17
secret sharing, 17
secure function evaluation, 12
secure multi-party computation, 16
semidefinite program (SDP), 190
semidefinite programming, 187
semisimple, 195
  decomposition, 199
SFE, 12
Shannon entropy, 36
simple, 195
SMP, 16
spectrum, 195
STAR, 47, 116
  AND, 50
  Boolean functions, 50
  to succeed at, 48
  XOR, 51
state
  algebra, 195
  bipartite, 27
  discrimination, 32
  discrimination with post-measurement information, 43
  EPR pair, 106
  invariant, 202
  mixed, 26
  post-measurement, 29
  pure, 25, 195
  reduced, 28
  von Neumann entropy, 37
  Werner, 110
state discrimination, 47
  AND, 50
  Boolean functions, 50
  with post-measurement information, 47
  XOR, 51
state space, 196
string commitment, 151
string of Pauli matrices, 41
strong duality, 191
superoperator, 31
superposition game, 148
superselection rules, 7, 10

tensor product, 188
trace, 188
trace distance, 32
trace norm, 33
trace out, 28
traceless, 18
transpose, 188
trusted initializer, 6
Tsirelson
  bound, 108, 123
  construction, 115, 122
  vectors, 115
two-party protocols, 153
two-sided ideal, 195
two-universal hash function, 154

uncertainty, 75
uncertainty relation, 8, 75
  Clifford algebra, 83
  for all MUBs, 82
  for anti-commuting observables, 83
  for Latin square MUBs, 80
  for MUBs in square dimensions, 79
  for two bases, 78
  Heisenberg, 75

Maassen and Uffink, 78
meta, 89
unconditional security, 6
unforgeable subway tokens, 3, 19
unital
algebra, 195
quantum channel, 31
unitary, 188
unitary evolution, 29
URAC, 133

variational distance, 35
verifier, 110
von Neumann entropy, 37

weak duality, 190
Werner state, 110

List of Figures

1.1 Encrypted pottery glaze formula, Mesopotamia 1500 BC 4
1.2 QKD today 4
1.3 Schematic run of a BC protocol when Alice and Bob are honest. . 9
1.4 Schematic run of a 1-2 OT protocol. 13

2.1 Bloch vector $(r_x, r_y, r_z) = (\cos\psi \sin\theta, \sin\psi \sin\theta, \cos\theta)$ 34
2.2 Latin Square (LS) 40
2.3 Mutually Orthogonal LS 40

3.1 Using post-measurement information. 44

4.1 2n = 2-cube 85
4.2 2n = 4-cube 85

5.1 A locking protocol for 2 bases. 95
5.2 Measurement for $|1, 1\rangle$. 101

6.1 Alice and Bob measure many copies of $|\Psi\rangle$ 107
6.2 Multiplayer non-local games. 111
6.3 Original problem 117
6.4 Derived problem 117

7.1 Optimal vectors for $n = 4$ obtained numerically using Matlab. . . 128

8.1 Tradeoff for $p = 0.6$. 137
8.2 Tradeoff for 2 outcomes. 138

9.1 A one-round XOR proof system. 143

10.1 Moving from a set of string with $g(x) = y$ to a set of strings with $g(x) = (y \mod 5) + 1$. 157

11.1 Bob performs a partial measurement $\mathcal{P}_i$, followed by noise $\mathcal{N}$, and outputs a guess bit x_g depending on his classical measurement outcome, the remaining quantum state, and the additional basis information. . . 175
11.2 $h((1-ar)/2)/4 + \log(\frac{1+r}{2})\log(4/3)/2$, where we only show the region below 0, i.e., where security can be attained. 184

C.1 Two vectors 206
C.2 $a \wedge b$ 207
C.3 $b \wedge a$ 207
C.4 Projections onto a vector 208
C.5 Reflection of a around m 209
C.6 Reflection of a plane perpendicular to m 210
C.7 Hadamard transform as reflection 210
C.8 Rotating in the plane $m \wedge n$. 211
C.9 Rotating g to $|g|\Gamma_1$. 211

Symbols

Symbol		Page
log	binary logarithm	
ln	natural logarithm	
a^*	complex conjugate of a	
$\|a\|$	absolute value of a	
$\|\mathcal{S}\|$	number of elements of the set $\mathcal{S}$	
$\mathbb{N}$	set of natural numbers $1, 2, 3, \ldots$	
$\mathbb{R}$	set of real numbers	
$\mathbb{C}$	set of complex numbers	
$[n]$	set of numbers $\{1, \ldots, n\}$	
$x_{\|\mathcal{S}}$	string x restricted to the indices in $\mathcal{S}$	
δ_{ij}	$\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise	
$\mathbb{I}$	identity matrix	
$\mathbb{I}_d$	$d \times d$ identity matrix	
$A^{[1]}$	matrix A acting on subsystem 1	
A^{-1}	inverse of the matrix A	
A^T	transpose of the matrix A	
A^*	conjugate of the matrix A	
$A^\dagger$	conjugate transpose of the matrix A	
$[a_{ij}]$	matrix whose entry in the i-th row and j-th column is a_{ij}	
A_{ij}	the entry in the i-th row and j-th column of the matrix A	
$\mathrm{Tr}(A)$	trace of $A = [a_{ij}]$ given by $\sum_j a_{jj}$	
$\mathrm{rank}(A)$	rank of the matrix A	
$A > 0$	A is positive definite	189
$A \geq 0$	A is positive semidefinite	189
$\|\|A\|\|_1$	trace norm of A, given by $\mathrm{Tr}\sqrt{A^\dagger A}$	
$\vec{a}$	real vector $\vec{a} = (a_1, \ldots, a_d)$	
$\|\Psi\rangle$	complex vector $\|\Psi\rangle = (\alpha_1, \ldots, \alpha_d)$	
$\langle\Psi\|$	conjugate transpose of the vector $\|\Psi\rangle$	

$\|x_b\rangle$	string x encoded in basis b	
$\langle\Psi\|\Phi\rangle$	inner product of $\|\Psi\rangle$ and $\|\Phi\rangle$	
$x \cdot y$	standard inner product of real vectors x and y	
$\|\Psi\rangle\langle\Phi\|$	outer product of $\|\Psi\rangle$ and $\|\Phi\rangle$	
$\|\Psi\rangle\langle\Psi\|$	projector onto the vector $\|\Psi\rangle$	
$\|\|\|\Psi\rangle\|\|$	2-norm given by $\sqrt{\langle\Psi\|\Psi\rangle}$	
$\mathcal{H}$	a Hilbert space	
$\mathbb{B}(\mathcal{H})$	set of all bounded operators on $\mathcal{H}$	
$\mathcal{S}(\mathcal{H})$	set of states on $\mathcal{H}$	
$[A, B]$	commutator $AB - BA$	
$\{A, B\}$	anti-commutator $AB + BA$	
$\mathrm{Comm}(\mathscr{A})$	commutant of the algebra $\mathscr{A}$	198
$\mathcal{Z}_{\mathscr{A}}$	center of the algebra $\mathscr{A}$	198
$\langle A_1, \ldots, A_n\rangle$	algebra generated by $A_1, \ldots, A_n$	194
$\langle\mathcal{S}\rangle$	algebra generated by operators from the set $\mathcal{S}$	194
$D(\rho, \sigma)$	trace distance of ρ and σ	32
$F(\rho, \sigma)$	fidelity of ρ and σ	35
$d(X\|\rho)$	distance from uniform of r.v. X given state ρ	165
$h(p)$	binary entropy	36
$H(X, Y)$	joint entropy of X and Y	36
$H(X\|Y)$	conditional entropy of X given Y	36
$\mathcal{I}(X, Y)$	mutual information of X and Y	36
$\mathcal{I}_c(\rho_{AB})$	classical mutual information of ρ_{AB}	36
$\mathcal{I}_{acc}(\mathcal{E})$	accessible information of an ensemble $\mathcal{E}$	38
$S(\rho)$	von Neumann entropy of the state ρ	37
$\chi(\rho)$	Holevo quantity	38
$H_\infty(X)$	min-entropy	37
$H_2(X)$	collision entropy	37
$H_2(\rho_{AB}\|\rho_B)$	collision entropy of ρ_{AB} given ρ_B	154

Samenvatting

Quantum computing heeft een grote invloed op cryptografie gehad. Met de ontdekking van Shors quantum algorithme voor het factoriseren van grote getallen kunnen opeens bijna alle klassieke systemen gebroken worden zodra een quantum computer is gebouwd. Het is daarom belangrijk om andere manieren te verzinnen om veilige cryptografische protocollen te kunnen implementeren. Dit proefschrift draagt ertoe bij om zowel de fysieke beperkingen, als ook de mogelijkheden van cryptographie in een quantum omgeving beter te begrijpen. Wij bekijken eerst twee aspecten die een cruciale rol spelen voor de veiligheid van quantum protocollen: onzekerheidsrelaties en quantum entanglement. Hoe kunnen wij goede onzekerheidsrelaties voor een groot aantal meetinstellingen vinden? Wat is het effect van entanglement op klassieke protocollen? En, welke beperkingen legt entanglement quantum protocollen op? Ten slotte, kunnen wij deze beperkingen omzeilen onder realitische aanames?

Informatie in quantum toestanden

In dit deel houden wij ons bezig met het extraheren van informatie uit quantum toestanden. Een van de meest fundamentele doelen is het onderscheiden van quantum toestanden. Gegeven een set van mogelijke toestanden, wat is de toestand die wij op dit moment voor handen hebben? Wij bestuderen een variant van dit probleem dat van belang is voor de veiligheid van protocollen in het bounded quantum storage model. We ontvangen na de meting, of meer algemeen nadat een quantum memory bound toegepast wordt, nog extra informatie. Wij introduceren een algemeen algebraisch raamwerk, dat het mogelijk maakt om dit probleem voor elke set van toestanden op te lossen en geven twee voorbeelden.

Verder onderzoeken wij entropische onzekerheidsrelaties, die een andere manier vormen om Heisenberg's onzekerheids principe te beschrijven. Dit is meestal een beter manier om "onzekerheid" te beschrijven aangezien de ondergrens niet afhangt van een bepaald toestand maar alleen van de metingen zelf. Entropische

onzekerheidsrelaties hebben recentelijk meer invloed gekregen binnen het veld van quantum cryptografie in het bounded storage model, waar de veiligheid van protocollen uiteindelijk afhangt van zulke onzekerheidsrelaties. Dus nieuwe onzekerheidsrelaties kunnen tot nieuwe protocollen leiden.

Onzekerheidrelaties zijn bekend voor twee of $d+1$ wederzijds "unbiased measurements". Wij bewijzgn eerst nauwe entropische onzekerheidsrelaties voor metingen met een groot aantal "mutually unbiased bases" (MUBs) in dimensionen $d = s^2$. Wij laten ook zien dat MUBs geen goede keuze zijn voor "locking" van klassieke informatie in quantum toestanden; ook als wij meer dan twee van zulke MUBs gebruiken neemt het locking effect niet toe.

Onze resultaten laten zien dat men heel voorzichtig dient te zijn om "maximaal incompatibele" metingen als wederzijds "unbiased" te veronderstellen. Maar welke eigenschappen moeten een meting hebben om heel 'incompatibel' te zijn? Gelukkig kunnen wij zulke eigenschappen vinden voor metingen met twee uitkomsten. Voor anti-commuterende metingen die generatoren van een Clifford algebra vormen, bewijzen wij optimale onzekerheidsrelaties voor de Shannon entropie, en bijna optimale relaties voor de collision entropie. Onze resultaten kunnen worden toegepast op quantum cryptographie.

Entanglement

In dit deel onderzoeken wij quantum entanglement. Allereerst, kijken wij naar Tsirelson inequalities. Wij laten zien hoe wij de optimale strategie voor spelletjes met twee uitkomsten met behulp van semidefinite programming kunnen bepalen. Als voorbeeld laten wij een upper bound voor de gegeneraliseerde CHSH ongelijkheid zien.

Verder laten wij zien hoe klassieke interactieve bewijssystemen met twee spelers (provers) kunnen veranderen als de spelers entanglement kunnen delen. Dit is een voorbeeld van hoe de veiligheid van klassieke systemen kan veranderen, ook al is het alleen mogelijk een beperkt soort quantum operaties uit te voeren: Het bewijssysteem wordt significant verzwakt ook al hebben de spelers geen toegang tot een quantum computer.

Applicaties voor de cryptografie

In deel IV onderzoeken wij de consequenties van onzekerheidsrelaties en entanglement in quantum systemen voor de cryptografie. Traditioneel houdt de cryptografie zich vooral bezig met het veilig versturen van berichten. Maar met de opkomst van het internet zijn nieuwe taken van belang geworden. Wij willen protocollen creeren voor het elektronisch stemmen, online veilingen, ondertekenen van contracten en vele andere applicaties, waarbij de deelnemers elkaar niet

vertrouwen. De focus ligt daarbij op twee primitieven, met behulp waarvan wij al deze problemen kunnen oplossen: bit commitment en oblivious transfer. Klassieke protocollen voor deze primitieven zijn gebaseerd op computationele aanames die met behulp van een quantum computer gebroken kunnen worden. Helaas is het bekend dat zelfs in de quantum wereld deze primitieven niet helemaal zonder aannames geimplementeerd kunnen worden. Wat hopen wij dan wel te kunnen bereiken?

Als bit commitment onmogelijk is, kunnen wij misschien de taak een klein beetje aanpassen en dan nuttige protocollen vinden? Hier bekijken wij commitments van een hele string van bits tegelijk, waar de tegenstander niet is beperkt. Als bit commitment onmogelijk is, is perfecte string commitment ook niet mogelijk. Maar wij geven elke tegenstander de mogelijkheid om een beetje vals te spelen. Wij geven een raamwerk voor een familie van string commitment protocollen. Hoe wij informatie meten blijkt een cruciale rol te spelen; voor een heel sterke maat van informatie laten wij zien dat zelfs deze imperfecte string commitments niet mogelijk zijn. Maar voor een zwakkere manier om informatie te meten construeren wij toch niet-triviale protocollen die klassiek niet mogelijk zijn.

Ten slotte laten wij zien dat bit commitment en oblivious transfer wel mogelijk worden, indien wij de tegenstander realistische beperkingen opleggen. Wij introduceeren het noisy-storage model, dat nauw is gerelateerd aan het bounded-storage model. Wij laten zien dat het mogelijk is om oblivious transfer te implementeren, zolang de tegenstander qubits niet zonder fouten kan opslaan. Gegeven de status van de experimentele mogelijkheden vandaag de dag, lijkt dit een realistische aanname, maar is afhankelijk van de implementatie moeilijk te bepalen. Dezelfde problemen die het ook zo moeilijk maken om een quantum computer te bouwen komen ons hier ten goede!

Summary

Quantum computing had a profound impact on cryptography. Shor's discovery of an efficient quantum algorithm for factoring large integers implies that nearly all existing classical systems based on computational assumptions can be broken, once a quantum computer is built. It is therefore imperative to find other means of implementing secure protocols. This thesis aims to contribute to the understanding of both the physical limitations, as well as the possibilities of cryptography in the quantum setting. To this end, we first investigate two notions that are crucial to the security of quantum protocols: uncertainty relations and entanglement. How can we find good uncertainty relations for a large number of measurement settings? How does the presence of entanglement affect classical protocols? And, what limitations does it impose on implementing quantum protocols? Finally, can we circumvent some of those limitations using realistic assumptions?

Information in Quantum States

In this part, we start by investigating how to extract information from quantum states. One of the most basic tasks is the discrimination of quantum states. Given an ensemble of known quantum states, which one do we hold in our hands? We study a variant of this problem which is of central importance for the security of protocols in the bounded-quantum-storage model. Here, we are given additional information about the state after the measurement or, more generally, after a quantum memory bound is applied. We prove general bounds on the success probability which answer in the negative the question whether deterministic privacy amplification is possible in all known protocols in the bounded-quantum-storage model. To this end, we introduce a general algebraic framework which allows us to solve this problem for any set of states and provide two explicit examples.

We then turn to examine entropic uncertainty relations, which are an alternative way to state Heisenberg's uncertainty principle. They are frequently a more useful characterization, because the "uncertainty" is lower bounded by a quantity that does not depend on the state to be measured. Recently, entropic uncertainty relations have gained importance in the context of quantum cryptography in the bounded-storage model, where proving the security of protocols ultimately reduces to bounding such relations. Proving new entropic uncertainty relations could thus give rise to new protocols. Such relations are known for two or $d + 1$ mutually unbiased measurements. We prove tight entropic uncertainty relations for measurements in a large number of specific mutually unbiased bases (MUBs) in square dimensions. In a similar way, we show that such MUBs are unsuitable for locking classical correlations in quantum states: Using 2 or all of them does not increase the locking effect.

Our result shows that one needs to be careful about thinking of "maximally incompatible" measurements as being necessarily mutually unbiased. But what properties do measurements need to have in order to give strong uncertainty relations? We find very strong uncertainty relations from the generators of a Clifford algebra. In particular, we prove that for k such anti-commuting observables $X_1, \ldots, X_k$ we obtain optimal uncertainty relations for the Shannon entropy and nearly optimal relations for the collision entropy. Our results have immediate applications to quantum cryptography in the bounded-storage model.

Entanglement

In this part, we investigate the intriguing notion of quantum entanglement. We demonstrate how to find the optimal quantum strategies for correlation inequalities where each measurement has exactly two outcomes using semidefinite programming. As an example, we prove a tight upper bound for a well-known generalized CHSH inequality.

Furthermore, we consider how a classical two-prover interactive proof system changes if the provers are allowed to share entanglement. In this setting, a polynomial time bounded verifier is allowed to ask questions to two unbounded provers, who are trying to convince the verifier of the validity of a specific statement, even if the statement is false. The provers may thereby agree on any strategy ahead of time, but can no longer communicate once the protocol starts. Surprisingly, it turns out that, when the provers are allowed to share entanglement, it is possible to simulate two such classical provers using a single quantum prover. This indicates that entanglement among provers truly weakens the proof system, and provides an example of how classical systems can be affected, even if we only allow a very limited set of quantum operations.

Applications to Cryptography

In this part, we consider the consequences of uncertainty relations and entanglement in quantum systems to cryptography. Traditional cryptography is concerned with the secure and reliable transmission of messages. With the advent of widespread electronic communication and the internet, however, new cryptographic tasks have become increasingly important. We would like to construct secure protocols for electronic voting, online auctions, contract signing and many other applications where the protocol participants themselves do not trust each other. main focus is on two primitives, which form an important building block for constructing multi-party protocols: bit commitment and oblivious transfer. Classical protocols for such problems are usually based on computational assumptions which do not stand up to a quantum computer. Unfortunately, it has been shown that even quantum computers do not help in this case and that perfect quantum bit commitment and oblivious transfer are impossible. In the face of such negative statements, what can we still hope to achieve?

Given that perfect bit commitment is impossible, perhaps we can alter the task slightly and obtain useful protocols? Here, we considered commitments to an entire string of bits at once, when the attacker has unbounded resources at his disposal. Evidently, if perfect bit commitment is impossible, perfect string commitment is also impossible as well. However, we showed that we can obtain non-trivial quantum protocols, where the participants have a small ability to cheat. To this end, we introduced a framework for the classification of string commitment protocols. In particular, we proved that the measure of information is crucial to the security: For a very strong notion of security, we showed that even slightly imperfect quantum string commitments are also impossible. Nevertheless, we showed that for a weaker measure of information we can indeed obtain nontrivial protocols, which are impossible in a classical world.

Luckily, it turns out that we can implement oblivious transfer if we are willing to assume that storing qubits is noisy. We introduce the model of noisy quantum storage, which is similar to the model of bounded quantum storage. Here, however, we consider an explicit noise model inspired by present day technology. If the honest parties can perform perfect quantum operations, then we show that the protocol is secure for any amount of noise. In case the honest participants are only able to perform noisy operations themselves, we analyze a practical protocol that can be implemented using present-day hardware. We show how to derive explicit tradeoffs between the amount of storage noise, the amount of noise in the operations performed by the honest participants and the security of the protocol. Here, the very problem that makes it so hard to implement a quantum computer can actually be turned to our advantage.

Titles in the ILLC Dissertation Series:

ILLC DS-2001-01: **Maria Aloni**
Quantification under Conceptual Covers

ILLC DS-2001-02: **Alexander van den Bosch**
Rationality in Discovery - a study of Logic, Cognition, Computation and Neuropharmacology

ILLC DS-2001-03: **Erik de Haas**
Logics For OO Information Systems: a Semantic Study of Object Orientation from a Categorial Substructural Perspective

ILLC DS-2001-04: **Rosalie Iemhoff**
Provability Logic and Admissible Rules

ILLC DS-2001-05: **Eva Hoogland**
Definability and Interpolation: Model-theoretic investigations

ILLC DS-2001-06: **Ronald de Wolf**
Quantum Computing and Communication Complexity

ILLC DS-2001-07: **Katsumi Sasaki**
Logics and Provability

ILLC DS-2001-08: **Allard Tamminga**
Belief Dynamics. (Epistemo)logical Investigations

ILLC DS-2001-09: **Gwen Kerdiles**
Saying It with Pictures: a Logical Landscape of Conceptual Graphs

ILLC DS-2001-10: **Marc Pauly**
Logic for Social Software

ILLC DS-2002-01: **Nikos Massios**
Decision-Theoretic Robotic Surveillance

ILLC DS-2002-02: **Marco Aiello**
Spatial Reasoning: Theory and Practice

ILLC DS-2002-03: **Yuri Engelhardt**
The Language of Graphics

ILLC DS-2002-04: **Willem Klaas van Dam**
On Quantum Computation Theory

ILLC DS-2002-05: **Rosella Gennari**
Mapping Inferences: Constraint Propagation and Diamond Satisfaction

ILLC DS-2002-06: **Ivar Vermeulen**
A Logical Approach to Competition in Industries

ILLC DS-2003-01: **Barteld Kooi**
Knowledge, chance, and change

ILLC DS-2003-02: **Elisabeth Catherine Brouwer**
Imagining Metaphors: Cognitive Representation in Interpretation and Understanding

ILLC DS-2003-03: **Juan Heguiabehere**
Building Logic Toolboxes

ILLC DS-2003-04: **Christof Monz**
From Document Retrieval to Question Answering

ILLC DS-2004-01: **Hein Philipp Röhrig**
Quantum Query Complexity and Distributed Computing

ILLC DS-2004-02: **Sebastian Brand**
Rule-based Constraint Propagation: Theory and Applications

ILLC DS-2004-03: **Boudewijn de Bruin**
Explaining Games. On the Logic of Game Theoretic Explanations

ILLC DS-2005-01: **Balder David ten Cate**
Model theory for extended modal languages

ILLC DS-2005-02: **Willem-Jan van Hoeve**
Operations Research Techniques in Constraint Programming

ILLC DS-2005-03: **Rosja Mastop**
What can you do? Imperative mood in Semantic Theory

ILLC DS-2005-04: **Anna Pilatova**
A User's Guide to Proper names: Their Pragmatics and Semanics

ILLC DS-2005-05: **Sieuwert van Otterloo**
A Strategic Analysis of Multi-agent Protocols

ILLC DS-2006-01: **Troy Lee**
Kolmogorov complexity and formula size lower bounds

ILLC DS-2006-02: **Nick Bezhanishvili**
Lattices of intermediate and cylindric modal logics

ILLC DS-2006-03: **Clemens Kupke**
Finitary coalgebraic logics

ILLC DS-2006-04: **Robert Špalek**
Quantum Algorithms, Lower Bounds, and Time-Space Tradeoffs

ILLC DS-2006-05: **Aline Honingh**
The Origin and Well-Formedness of Tonal Pitch Structures

ILLC DS-2006-06: **Merlijn Sevenster**
Branches of imperfect information: logic, games, and computation

ILLC DS-2006-07: **Marie Nilsenova**
Rises and Falls. Studies in the Semantics and Pragmatics of Intonation

ILLC DS-2006-08: **Darko Sarenac**
Products of Topological Modal Logics

ILLC DS-2007-01: **Rudi Cilibrasi**
Statistical Inference Through Data Compression

ILLC DS-2007-02: **Neta Spiro**
What contributes to the perception of musical phrases in western classical music?

ILLC DS-2007-03: **Darrin Hindsill**
It's a Process and an Event: Perspectives in Event Semantics

ILLC DS-2007-04: **Katrin Schulz**
Minimal Models in Semantics and Pragmatics: Free Choice, Exhaustivity, and Conditionals

ILLC DS-2007-05: **Yoav Seginer**
Learning Syntactic Structure

ILLC DS-2008-01: **Stephanie Wehner**
Cryptography in a Quantum World