Cryptography from Quantum Uncertainty
in the Presence of Quantum Side Information

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. P.F. van der Heijden,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 18 december 2012
klokke 10.00 uur

door

Niek Johannes Bouman

geboren te Son en Breugel
in 1983

Samenstelling van de promotiecommissie

Promotor: Prof. dr. Ronald Cramer (CWI & Universiteit Leiden)

Co-Promotor: Dr. Serge Fehr (CWI)

Overige leden:

Prof. dr. Richard Gill (Universiteit Leiden)
Prof. dr. Renato Renner (ETH Zürich)
Prof. dr. Louis Salvail (Université de Montréal)
Dr. ir. Berry Schoenmakers (Technische Universiteit Eindhoven)
Prof. dr. Peter Stevenhagen (Universiteit Leiden)

Het onderzoek is uitgevoerd bij Centrum Wiskunde & Informatica (CWI) te Amsterdam. CWI is het nationale onderzoekscentrum voor wiskunde en informatica.

Dit werk maakt deel uit van het onderzoeksprogramma Vrije competitie Exacte Wetenschappen (EW) dat gefinancierd is door de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO).

Cryptography from Quantum Uncertainty

in the Presence of Quantum Side Information

Bouman, Niek Johannes

Cryptography from Quantum Uncertainty,
in the Presence of Quantum Side Information

Printed by Ipskamp Drukkers
AMS MSC2010 classification: 81P94 (Quantum Cryptography)
NUR: 919
ISBN: 978-94-6191-536-8

Typeset with $\LaTeX\ 2_\varepsilon$.
Cover Photo and Design: Niek J. Bouman

to Ægle Hoekstra, my inspirational math teacher from high school, who died of non-Hodgkins lymphoma in 1999, at the age of 46

Acknowledgements

Firstly, I would like to thank my thesis supervisor Serge Fehr for his guidance. Serge was always available for my questions and always willing to proofread my write-ups. He returned those with valuable feedback, often within one or two days. Due to his detailed feedback the quality of my writing has significantly improved over the past years, and under his guidance my "engineer's understanding" of mathematics has transformed into understanding and appreciating the rigorous approach to mathematics.

I would like to thank Ronald Cramer for proofreading my thesis near the end of my PhD, and for giving not only technical remarks about my thesis but also more philosophical remarks about research and academic writing in general. Thanks also go to Christian Schaffner, who was, beyond one of my co-authors, a coach for me. He also proofread some of my work, for which I am grateful. I would also like to thank Carlos González-Guillén for co-authorship.

I want to thank the members of the thesis reading committee for their involvement. Special thanks go to Renato Renner, who invited me to visit his group at ETH Zürich, to Berry Schoenmakers and Louis Salvail for their extensive feedback on the thesis, and to Peter Stevenhagen, for his help to accelerate the process of planning a date for the defense. Ronald de Wolf, thank you for proofreading the *Nederlandse samenvatting*. I want to thank the following academic colleagues for interesting discussions, answering questions and/or providing feedback on our work: Dejan Dukaric, Anthony Leverrier, Xin Li, Krzysztof Pietrzak, Marco Tomamichel, Dominique Unruh, Salil Vadhan, Daniel Wichs, Severin Winkler, Jürg Wullschleger, David Zuckerman.

Groeten aan huidige en oud-groepsleden, en aan mijn parttime-kamergenoot Wieb Bosma. Susanne van Dam, dank voor al jouw hulp, variërend van een pleister voor een bloedende vinger tot naam-badges voor een event, en nog meer dank voor jouw vriendelijkheid en vrolijkheid. Bikkie Aldeias, dankzij jou is de hoofdingang van CWI een gezellige stek. En jouw warme ontvangst op de dag dat ik, met voedselvergiftiging, toch maar ging solliciteren is onvergetelijk. Ook wil ik graag de ondersteunende diensten van CWI bedanken, in het bijzonder de volgende personen die mij met allerlei zaken hebben geholpen: Huib van den Berg, Edwin de Boer, Margriet Brouwer, Maarten Dijkema, Peter Dijkhuis, Karin van Gemert, Minnie Middelberg, Rick Ooteman, Michaël Smeding, Mike Zonsveld.

Vrienden, ik tref 't maar met jullie! Een groet aan jullie allemaal, en dank voor jullie steun. Tot gauw!

Henk en Mieke, zonder jullie zou ik — en dus ook dit boekje — er nooit zijn geweest. Dit dankwoord vind ik een mooie plek om stil te staan bij jullie onvoorwaardelijke steun. Het geeft een fijn gevoel om te weten dat ik altijd op jullie kan terugvallen. Van de onbezorgde jeugd die jullie mij hebben gegeven, en van de relaxte studententijd die jullie mogelijk hebben gemaakt, heb ik nog iedere dag veel profijt.

Niek Bouman,
Amsterdam, november 2012

Contents

Acknowledgements **7**

1 Introduction **13**

1.1 What is Cryptography? **14**

1.1.1 Public-Key Encryption 14

1.1.2 Secure Two-Party Computation 15

1.1.3 Cryptography as a Science 18

1.2 What is Quantum Mechanics? **19**

1.2.1 Superposition 19

1.2.2 Entanglement 21

1.2.3 The Uncertainty Principle 23

1.2.4 Implications to Information Theory and Computer Science 24

1.3 The Research Field of Quantum Cryptography **25**

1.3.1 Concrete Example: Quantum Key Distribution 26

1.3.2 Beyond QKD: Secure Two-Party Computation 31

1.3.3 Beyond QKD: Position-Based Quantum Cryptography . . 33

1.3.4 A Brief History 34

1.4 Thesis Outline, Contributions and Open Problems **36**

1.4.1 Chapter 2: Preliminaries 36

1.4.2 Chapter 3: Random Sampling from a Quantum Population 37

1.4.3 Chapter 4: Authentication from a Weak Key with a Privacy Requirement 38

1.4.4 Chapter 5: Hybrid Security of Password-Based Identification 39

2 Preliminaries **41**

2.1 Basic Notation **42**

2.2 Probability Theory **43**

2.2.1 Random Variables 44

2.2.2 Some Important Distributions 45

2.2.3 The Bias of a Binary Random Variable 46

2.2.4 Distance between Distributions 47

2.2.5 Jensen's Inequality 49

2.2.6 Hoeffding's Inequality 50

2.3 Classical Measures of Uncertainty . . . 51
2.3.1 Rényi Entropies . . . 52
2.3.2 Conditional Entropy . . . 53
2.4 Privacy Amplification . . . 55
2.4.1 Universal Hashing . . . 55
2.4.2 Extractors . . . 58
2.5 Statistically Secure Encryption and Authentication . . . 61
2.5.1 The One-Time Pad . . . 61
2.5.2 Message Authentication . . . 61
2.6 Hilbert Spaces . . . 62
2.6.1 Dirac's Braket Notation . . . 63
2.6.2 Operators . . . 64
2.6.3 Tensor Products . . . 65
2.6.4 Vector and Matrix Representations . . . 68
2.7 Quantum Systems and Operations . . . 70
2.7.1 Postulates of Quantum Mechanics . . . 70
2.7.2 States . . . 72
2.7.3 Computational and Hadamard Basis . . . 73
2.7.4 Partial Trace and Purification . . . 73
2.7.5 The Schmidt Decomposition . . . 74
2.7.6 Quantum Operations . . . 75
2.7.7 Alternative Descriptions of Measurements . . . 77
2.7.8 Entanglement . . . 78
2.7.9 Hybrid Systems . . . 79
2.7.10 Distance between States . . . 81
2.8 The One-Time Pad in a Quantum Setting . . . 83
2.9 Measures of Uncertainty for Density Matrices . . . 83
2.10 Extractors against Quantum Side Information . . . 88
2.11 Quantum Identification . . . 89
2.11.1 The Basic QID Protocol . . . 91
2.11.2 Security against Man-in-the-Middle Attacks . . . 93

3 Random Sampling from a Quantum Population 95
3.1 Introduction . . . 96
3.1.1 Proposed Framework . . . 97
3.1.2 Applications . . . 97
3.1.3 Notation . . . 99
3.2 Sampling from a Classical Population . . . 99
3.2.1 Sampling Strategies . . . 99

3.2.2 The Error Probability 102
3.2.3 Error Probabilities of the Example Sampling Strategies . . 103
3.3 Sampling from a Quantum Population 108
3.3.1 Analyzing Sampling Strategies in the Quantum Setting . . 109
3.3.2 The Quantum vs. the Classical Error Probability 111
3.3.3 Superpositions with a Small Number of Terms 115
3.4 A Security Proof for Quantum Oblivious Transfer 118
3.4.1 Security against Dishonest Alice 120
3.4.2 The Harder Case: Security against Dishonest Bob 122
3.5 An Accessible Proof for Quantum Key Distribution 125
3.5.1 Survey of Existing QKD Proofs 126
3.5.2 Our Proof 127
3.6 Conclusion 132

4 Authentication from a Weak Key with a Privacy Requirement 133
4.1 Introduction 134
4.1.1 Related Work 135
4.1.2 Motivation 135
4.1.3 Contributions 136
4.1.4 The Fuzzy Case 136
4.2 Security Definition 137
4.3 Dodis and Wichs' Authentication Protocol 137
4.3.1 Towards Achieving Key-Privacy 139
4.4 Our Construction 141
4.5 Proofs of Security and Privacy 143
4.6 Instantiating the Building Blocks 148
4.6.1 Look-Ahead Extractors against Classical Side Information 148
4.6.2 Look-Ahead Extractors and Quantum Side Information? . 155
4.6.3 Security and Instantiation of the MAC 156
4.6.4 Instantiating Protocol AUTH 156
4.7 The Fuzzy Case 159
4.8 Application: Password-Based Identification 160
4.9 Open Problem 161

5 Hybrid Security of Password-Based Identification 163
5.1 Introduction 164
5.1.1 A New Uncertainty Relation 164
5.1.2 Related Work 168
5.2 An All-But-One Entropic Uncertainty Relation 169

5.2.1 Constructing Good Families of Bases 175
5.3 A New Quantum Identification Protocol 175
5.3.1 Description of Our New Protocol 177
5.4 (Unconditional) Server Security 178
5.5 User Security in the BQSM 179
5.6 User Security in the Single-Qubit-Operations Model 182
5.6.1 The Model 182
5.6.2 No Privacy Amplification 183
5.6.3 Single-Qubit Measurements 183
5.6.4 User Security of `NEWQID` 187
5.6.5 Attack against `NEWQID` using Operations on Pairs of Qubits 192
5.7 Conclusion 194

A Notation 195

Bibliography 197

Index 209

Nederlandse samenvatting 213

Curriculum Vitae 215

1

Introduction

Chapter Contents

1.1 What is Cryptography? 14

1.1.1 Public-Key Encryption 14

1.1.2 Secure Two-Party Computation 15

1.1.3 Cryptography as a Science 18

1.2 What is Quantum Mechanics? 19

1.2.1 Superposition 19

1.2.2 Entanglement 21

1.2.3 The Uncertainty Principle 23

1.2.4 Implications to Information Theory and Computer Science 24

1.3 The Research Field of Quantum Cryptography 25

1.3.1 Concrete Example: Quantum Key Distribution . . . 26

1.3.2 Beyond QKD: Secure Two-Party Computation . . . 31

1.3.3 Beyond QKD: Position-Based Quantum Cryptography 33

1.3.4 A Brief History 34

1.4 Thesis Outline, Contributions and Open Problems 36

1.4.1 Chapter 2: Preliminaries 36

1.4.2 Chapter 3: Random Sampling from a Quantum Population 37

1.4.3 Chapter 4: Authentication from a Weak Key with a Privacy Requirement 38

1.4.4 Chapter 5: Hybrid Security of Password-Based Identification 39

1.1 What is Cryptography?

The word *cryptography* stems from the Greek words *kryptos* ("hidden") and *graphein* ("writing") and used to be a synonym for *encryption* (private communication).

In the past, encryption was mainly used for military-strategic and diplomatic communication. Julius Caesar already used encryption to protect messages of military significance. And in World War II, each of the opposing forces used encryption.[1] Today, encryption is mainstream technology, and this is probably mainly due to the advent of the Internet era. Modern Internet browsers, for example, support multiple encryption standards. Examples include the *Advanced Encryption Standard* (AES) [DR00], which is a method for *symmetric-key* encryption (meaning that encryption and decryption are performed using the same key) and *RSA* [RSA78], which provides *public-key* encryption (explained below). RSA is widely used to, for example, secure internet connections and to securely perform online credit-card payments.

1.1.1 Public-Key Encryption

A milestone in the post-war history of encryption is *public-key cryptography*, in which each of the parties has its own *pair* of keys: a *private key* and an accompanying *public key*. As its name implies, the public key is not secret, but is publicly announced. This public key can be used by anyone to encrypt a message that is intended for the holder of the accompanying private key; only the latter person can decrypt this *ciphertext* (i.e., the encrypted message). The idea of public-key cryptography is commonly attributed to Diffie and Hellman [DH76], and to Merkle [Mer78]. Diffie and Hellman's contribution was to devise a key-agreement system based on the presumed difficulty of computing the logarithm in a large finite field (the "*discrete-log problem*"). This key-agreement system produces a shared *symmetric* key, which can then be used to encrypt data using a symmetric encryption scheme. In 1977, Rivest, Adleman and Shamir invented a public-key encryption scheme that became known as *RSA* [RSA78]. RSA is related to the presumed computational difficulty of performing integer factorization.[2] *ElGamal* [ElG85] is another widely known public-key encryption method, which was invented in 1984 and is named after its inventor.

[1] The Nazis used the *Enigma* machine to encrypt their strategic communications. Fortunately, in 1932 Polish codebreakers had already found structural weaknesses in an early version of Enigma, which gave the Allies an important advantage during the war. See also [Sin99].

[2] It turns out that RSA was invented already in 1973 by researchers of the Government Communication Headquarters (GCHQ) in the UK, but this remained classified until 1997.

The idea of public-key cryptography also led to the invention of *digital signatures* [DH76]. A digital signature is a publicly verifiable proof of authenticity of a message. When a person, say Alice, wants to *sign* a message m, she uses her private key to compute the signature from the message. No one else can *sign* a message "under Alice's name," since this requires Alice's private key, but everyone else can *verify* this signature on a message m' using Alice's public key. A successful verification proves that m' equals the message that Alice has signed.

Another way of obtaining public-key cryptographic schemes is using elliptic curves over large finite fields; this is referred to as *elliptic-curve cryptography* (ECC). ECC was first described in 1985, independently by Koblitz [Kob87] and Miller [Mil86]. ECC requires shorter key lengths and produces shorter signatures than for instance RSA to achieve a comparable security level.

Lattice-based encryption [AD96, MR09] uses yet another way to obtain public-key cryptographic schemes, and is currently quite popular in the academic cryptography community. In lattice-based encryption, security is based on the presumed computational intractability of certain problems on high-dimensional integer lattices, such as the shortest-vector problem. Although typical key sizes are quite large, the arithmetic operations required to perform encryption and decryption are very lightweight in comparison to, for example, ECC. Unlike RSA or ECC, lattice-based encryption is not known to be vulnerable to quantum-computer attacks (to which we will come back later).

1.1.2 Secure Two-Party Computation

Although the design and study of encryption schemes is still an important part of modern cryptography, a great amount of research is dedicated to problems that are fundamentally different from encryption. An important example that we want to discuss here is *secure two-party computation* (2PC), which itself is a particular case of *secure multi-party computation*, which is also called *secure computation* or *secure function evaluation*.

The central problem in secure 2PC, as first described by Yao in 1982 [Yao82], is the following. Consider two parties, Alice and Bob, and suppose that each of them holds a piece of private information, respectively X_A and X_B. Alice and Bob have agreed on two (possibly randomized[3]) functions, one for Alice (f_A) and one for Bob (f_B), and both functions take as argument (X_A, X_B). The goal for Alice and Bob is

[3] A *randomized function* is a function that takes an additional (but usually implicit) argument containing independent randomness (independent from all other possible arguments to the function).

to *jointly* compute their functions by means of executing an interactive *protocol* (i.e., performing a prescribed sequence of local computations and exchanging messages), such that Alice learns $f_A(X_A, X_B)$ and Bob learns $f_B(X_A, X_B)$, and such that they learn *nothing* beyond this. That is, both functions should be computed correctly, and we require that X_A should remain secret to Bob and X_B should remain secret to Alice, up to the information that is revealed by $f_B(X_A, X_B)$ and $f_A(X_A, X_B)$ respectively. Moreover, this should be achieved *even* if one of the players is actively trying to a) manipulate the outcome(s) of the computation and/or b) learn more information than he or she is supposed to. Hence, unlike encryption, where one aims to protect against a malicious *outsider* (the *adversary*), secure 2PC deals with scenarios where one of the *insiders* (Alice or Bob) could be dishonest. (If both parties are dishonest at the same time, then nothing can be achieved.)

An insightful instance of the secure-2PC problem is Yao's *Millionaires' problem*: two millionaires want to find out who is richer, without revealing how rich they are (except when they turn out to be equally rich). To be precise, the millionaires both want to learn the function

$$f(X_1, X_2) = \begin{cases} -1 & \text{if} \quad X_1 < X_2 \\ 0 & \text{if} \quad X_1 = X_2 \\ 1 & \text{if} \quad X_1 > X_2 \end{cases}$$

where X_1 and X_2 represent their respective fortunes.

To formally define *security* in secure 2PC, one considers an "ideal solution" where the parties send their inputs to an imaginary trusted party, called the *ideal functionality*, which then computes both functions and returns each output to the appropriate party. For a protocol to be *secure*, we require that at the end of its execution, the views of the players can be accurately simulated[4] using the views of the players in the ideal solution (i.e., when the users interact with the ideal functionality). Note that there are multiple ways to measure this accuracy, but this goes beyond the scope of this introduction.

More Instances of the Secure-2PC Problem

Suppose that Alice and Bob want to securely compute the *equality function*. This particular instance of the secure-2PC problem has a direct application to *password-based identification*. For example, when someone (the *user*) wants to withdraw cash from an ATM (the *server*), the user first has to announce his identity to the

[4] Depending on the *model* (see Section 1.1.3), additional complexity-theoretic requirements on the simulation might be necessary, for example that the simulator runs in polynomial time.

server (by means of inserting his debit card). In order to verify the user's identity, the server then prompts the user for a password and checks whether the password that the user entered matches the password that the server has stored in its record assigned to that user's identity.

The problem of this approach to password-based identification is that in case the user is interacting with a *fake* server (which does *not* know the user's password in advance), the user reveals his password to this fake server.

With the help of secure 2PC, the user and the server can compare the two passwords without revealing them by securely computing the equality function, i.e.

$$f(X_1, X_2) = \begin{cases} 0 & \text{if} \quad X_1 \neq X_2 \\ 1 & \text{if} \quad X_1 = X_2, \end{cases}$$

with the respective passwords as input. As a result, not only the server is protected against fake users (who do not know the password), but the (honest) user is now also protected against fake servers, which attempt to learn information about the password by interacting with the user.

An instance of the secure-2PC problem that is of special importance is *oblivious transfer* (OT) [Rab81, EGL85]. Oblivious transfer is known to be *complete* [Kil88] for secure 2PC, which means that any secure-2PC functionality can be built from sufficiently many OTs. In the most common form of OT, called "one-out-of-two"-OT (1-2 OT, for short), Alice holds two messages, m_1 and m_2. Bob may choose and view only one of the two messages; he learns nothing about the other message. Alice, in turn, remains ignorant about which message Bob chose to view.

Coin flipping [Blu81] is an instance of the secure-2PC problem where privacy of the inputs is not a concern. Instead, the emphasis here is on computing the function correctly. Alice and Bob, living far apart, are talking on the phone about their late grandmother's inheritance. They decide to flip a coin about who gets grandmother's golden watch. However, since they are far apart, how can they flip a coin such that both of them are convinced that the coin was tossed fairly? Hence, in this instance of the secure-2PC problem, Alice and Bob want to securely compute the randomized function that returns a random bit. This function does not take arguments from the players (only the implicit argument that provides the independent randomness; see footnote 3). We will come back to coin flipping in the context of quantum cryptography in Section 1.3.2.

1.1.3 Cryptography as a Science

Claude Shannon was the first to approach cryptography from a mathematical perspective [Sha49], thereby initiating the study of cryptography as a scientific discipline. His work can be appreciated as an early example of what we now call *provable security*. By provable security, we mean cryptographic research following the scientific methodology of rigorous analysis [Gol06]. The security analysis of a cryptographic scheme[5] consists of a formal *definition* that captures what it should mean for a scheme to be secure, a *model* that specifies what the adversary is capable of, and a *security proof*, which proves that in the specified model, the scheme satisfies the security definition.

Finding appropriate security definitions is often non-trivial. On the one hand, they should be as strong as possible. On the other hand, it should still be possible to *achieve* the definition by means of some protocol. By now, we have several established security definitions for a wide range of cryptographic problems. (By "established," we mean generally accepted by the cryptographic community as being "appropriate.") However, even those established definitions should not be viewed as set in stone; e.g., it may happen that a definition that is reasonable given the *anticipated* use of the scheme, turns out to be unsuitable for the *actual* way in which the scheme is used.

The *model* specifies the restrictions that we put on the adversary. Needless to say, those restrictions should be reasonable, otherwise it would not make sense to assume that a real adversary operates within those restrictions. Nevertheless, the meaning of "reasonable" may change over time. Furthermore, we prefer the restrictions to be as mild as possible. A common restriction is that the adversary has bounded computing power. (Note that this can be formalized using complexity theory.) When the latter restriction is the *only* restriction that comprises the model, then we speak of the *standard model*. In Maurer's *bounded-storage model* [Mau90], we restrict the amount of data that the adversary can store. Finding alternative restrictions is an ongoing challenge.

The proof technique employed in the *security proof* often strongly depends on the model. For example, in the standard model, security proofs are typically *conditional* on some unproven intractability assumption (like the presumed intractability of integer factorization), in the form of a *reduction*. In a reduction, one proves that the ability to efficiently break the cryptographic scheme can be used to efficiently solve the underlying mathematical problem. The contrapositive statement then

[5] A cryptographic scheme is a suite of related (cryptographic) protocols.

expresses that: "because solving the underlying mathematical problem is assumed to be intractable, we can assume that breaking the cryptographic scheme is intractable as well." In contrast, in the bounded-storage model one uses information-theoretic techniques to prove security. Moreover, those proofs (and also the proofs presented in this thesis) are *unconditional*; they do not rely on some unproven intractability assumption.

When a cryptographic scheme has been "proven secure," it does not mean that the scheme is unbreakable in practice. The main cause of this paradox is that often several details of a physical implementation of the scheme are missing in the scheme's theoretical description for which the proof holds. These gaps between the theoretical description and physical implementation of the scheme can thus be exploited for attacks. *Leakage-resilient cryptography* [DP08] is a line of research within cryptography to address a specific class of such attacks, called *side-channel attacks*, which try to exploit leakage of private information through radiation, power consumption, sound, time delays, etcetera.

To arrive at a scheme that is accompanied by a security proof, it is usually *not* a successful approach to first design a "secure-looking" scheme and then attempt to prove its security. Instead, one typically designs a cryptographic scheme "along with its proof," i.e., one has some proof strategy in mind, and then builds the scheme such that this proof strategy applies.

1.2 What is Quantum Mechanics?

Quantum mechanics[6] describes the behavior of energy and matter on a small (atomic and sub-atomic) scale. To illustrate why quantum mechanics can be useful in cryptography, we will discuss some examples of typical quantum behavior.

1.2.1 Superposition

A main concept in quantum mechanics is *superposition*. In this introduction, we want to give an example of an optical experiment where the presence of superposition can be observed [SS98]. For our experiment, we use a source that is capable of producing a single photon,[7] beam splitters, mirrors and single-photon detectors.

[6] Quantum mechanics is sometimes called *quantum dynamics* or simply *quantum theory*.

[7] A *photon* is the elementary particle for electromagnetic radiation. Whenever we speak of a photon in this thesis, we mean a photon with a frequency that lies in (or near) the visible spectrum.

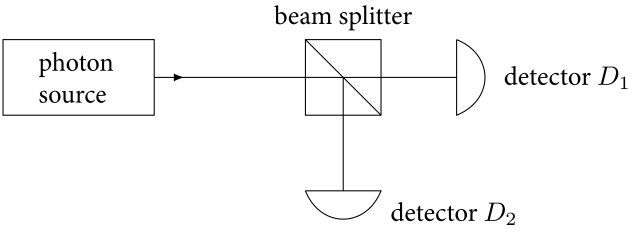

Figure 1.1: Setup showing the particle-like behavior of photons. In this setup, each incoming photon will be detected *either* at D_1 or D_2, both with probability one-half; there will never be a single photon that is detected at both detectors simultaneously. Hence, the photon behaves like a particle that randomly chooses one of the two paths emanating from the beam splitter, and this is due to the nature of the setup: the photon is immediately observed after passing though the beam splitter.

We will actually start by analyzing a "reference" experiment where superposition does *not* occur (or, is immediately destroyed before it can be observed, if you want). Consider the setup illustrated in Figure 1.1, which will demonstrate the *particle-like behavior* of light. Suppose that the source emits single photons at a constant rate (think of one photon per second, but this is actually irrelevant for our discussion). These photons pass through a beam splitter, and at each of the two outputs of the beam splitter we have placed a detector. When performing this experiment, we will observe that for each photon (i.e., every second) *only one* of the detectors "clicks" (i.e., reports a detection of a single photon); we will never observe an event where both detectors click simultaneously (of course, under the assumption of an idealized noise-free setting). This leads us to conclude that the photon behaves like a particle that randomly takes one of the two outputs of the beam splitter.

Let us now analyze Figure 1.2, which shows a *Mach–Zehnder interferometer*. Here, each photon encounters two beam splitters along its path. The two paths emanating from the outputs of the left beam splitter merge again at the right beam splitter. From what we have seen in the reference experiment (see Figure 1.1), it is tempting to conclude that the photons will behave like particles also in this setup, such that the photon appears either at D_1 or D_2, both with probability $1/2$. This is however not the case: *every photon will be detected at* D_1*!* In this setup, it turns out that the photon behaves *wave-like*, i.e., it takes *both paths simultaneously* and interferes at the right beam splitter constructively towards D_1 (and destructively towards D_2). The paths are said to be in *superposition*. According to quantum mechanics, the superposition exists (or, is preserved) in this experiment because the setup does

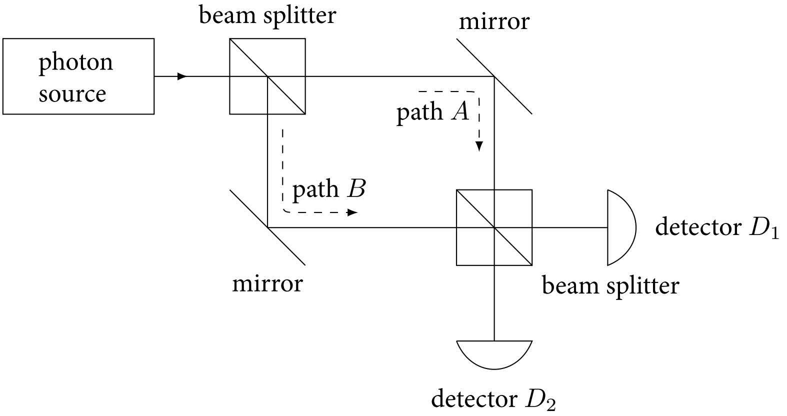

Figure 1.2: Mach–Zehnder interferometer showing quantum interference: a single photon from the source will always be detected at D_1. Note that this apparent asymmetry is due to the location of the photon source; if the photon source is placed above the first beam splitter, then every photon will be detected at D_2.

not reveal whether the photon traveled via path A or path B.

If, on the other hand, we regard the photon as a particle and measure by some means which path it takes, then the superposition *collapses* (gets destroyed) and the photon will start behaving particle-like again. In this case, the probabilities of detecting the photon at D_1 and D_2 will coincide with those of the reference experiment, i.e., both equal to one-half. In short, the photon's behavior *depends on whether it is observed or not.* Note that one possible way to determine which path the photon takes is to make the length of one of the paths longer (by an amount much larger than the wavelength of the photon) and analyze the photon's traveling time.

1.2.2 Entanglement

Entanglement is a special kind of superposition of states of a joint quantum system, where the latter is a physical system that consists of multiple subsystems (e.g., particles). When performing measurements on *entangled* particles, where these measurements are separated in time by a space-like interval,[8] the outcomes can

[8] By saying that measurements are separated in time by a space-like interval, we mean that for a given distance d between the particles on which those measurements are performed, the time interval between the two measurements is small enough to rule out a causal relationship between the

be correlated in a stronger sense than classical (i.e., non-quantum) physics would allow.

Before we try to explain how entanglement can exhibit itself, we discuss the following reference experiment. Suppose that Charlie prepares two boxes in the following manner. First, Charlie flips a coin. If heads comes up, he puts a ball in each box; otherwise he leaves both boxes empty. Then, he closes the boxes and sends one of them to Alice and the other to Bob, who are living far apart. Now, if Alice and Bob open their boxes, it will be no surprise that either they both see a ball (we will call this event $\mathcal{B}$) or they both see just an empty box (event $\mathcal{E}$). Furthermore, if we assume that Charlie does not tell the outcome of the coinflip to Alice and/or Bob, the events $\mathcal{B}$ and $\mathcal{E}$ occur randomly in their view. Nevertheless, it is important to note that someone (in this example: Charlie) can already tell what Alice and Bob will observe before they open the boxes themselves.

Let us now give an example of entanglement in terms of balls and boxes by means of the following thought experiment. (It is a thought experiment because macroscopic objects such as balls and boxes cannot be entangled in the sense described here.) Suppose that Charlie prepares the boxes in a different way: instead of flipping a coin to decide on whether to put a ball in each box or to leave both boxes empty, he creates a *superposition* over these two options. In some sense, the (closed) boxes then simultaneously both contain a ball and are both empty. We will now call the boxes *entangled*.[9] As a consequence of preparing the boxes in this peculiar way, Charlie cannot tell anymore what Alice and Bob will observe when they open the boxes. Let us assume that Alice first opens her box. Upon opening, the superposition collapses, fully randomly, to one of the two possibilities (a ball in each box, or both boxes empty). I.e., it is as if an imaginary coin is tossed at the moment of opening the box, and furthermore Alice's *local* operation (opening her box) has a *global* consequence: Bob's box will also change from a box that is "containing a ball and simultaneously empty" into an ordinary box that is *either* containing a ball *or* empty, depending on Alice's observation. Then, even when Bob opens his box within a space-like time interval after Alice, he will observe the same outcome as Alice. Since the (binary) outcome is only determined upon the first measurement, it is *only* known to Alice and Bob who observe their individual measurement outcome (when assuming that the boxes were indeed entangled in the sense described above).

measurement outcomes (where "small enough" of course depends on d.)

[9] To be precise, not every superposition over states of a joint quantum system would be called an entangled state. In the example, we speak of entanglement because the set of *global* options $\{bb, ee\}$ (where b and e represent a box with a ball and an empty box, respectively) cannot be written as a Cartesian product of *local* options, i.e., $\{b, e\} \times \{b, e\} = \{bb, be, eb, ee\} \neq \{bb, ee\}$.

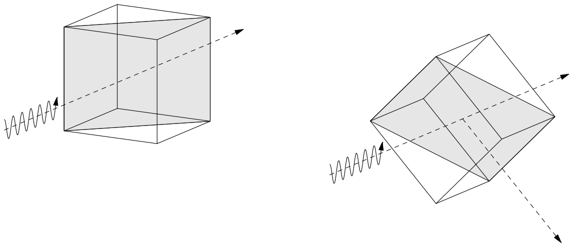

Figure 1.3: A demonstration of Heisenberg's uncertainty principle by means of incompatible photon-polarization measurements. We see two instances of a situation where a vertically polarized photon encounters a polarizing beam splitter. In both instances, we assume that single-photon detectors are placed at both outputs of the PBS, to enforce particle-like behavior. Left: the PBS is aligned to the photon's polarization and hence the photon is transmitted with certainty. Right: the PBS is oriented diagonally (rotated over 45 degrees) with respect to the photon's polarization, hence the photon is randomly transmitted or reflected, which means that the measurement outcome is maximally uncertain.

In terms of entangled particles (e.g., two photons that are entangled with respect to their polarizations), we can say that a measurement on one particle affects the outcome of a future measurement on the other (remote) particle as well. This remote "influence" is *instantaneous*; it is not limited by the speed of light. *Information*, on the other hand, cannot travel faster than light. Nonetheless, there is no contradiction here because entanglement alone cannot be used to transmit information.

1.2.3 The Uncertainty Principle

The *uncertainty principle*, as first described by Heisenberg, states that for any quantum system there exist pairs of so-called incompatible measurements, meaning that at least one of those measurements will produce an outcome that is somewhat uncertain. For example, it is well known that we cannot accurately measure both the position and the momentum of a moving particle (such as an electron).

In order to give an example of a pair of incompatible optical measurements, we first need to explain one of the building blocks. A *polarizing beam splitter* (PBS) turns an incoming polarized photon into a superposition of photons that are polarized along the axes of the PBS' own coordinate system. When a PBS is placed in a setup like Figure 1.1, the detectors will induce an immediate collapse of the superposition. In this case, we speak of a measurement. The photon is then either transmitted or reflected, depending on the polarization of the photon relative to the alignment of the PBS: if the photon's polarization is exactly parallel to the beam splitter's alignment, the photon is transmitted; when it is exactly perpendicular, the photon is reflected. For intermediary angle differences, the choice between transmission and reflection becomes random, where the probability of transmission is given by $\cos^2 \alpha$ for angle difference α. The polarization of the transmitted (reflected) photon will thus always be parallel (perpendicular) to the beam splitter's alignment.

Now, let us turn to the example. Suppose that a photon with a vertical polarization is entering a vertically-oriented PBS, see Figure 1.3 (left illustration). We assume that single-photon detectors are placed at both outputs of the PBS, such that we can speak of a measurement. Because the beam splitter is perfectly aligned to the photon's polarization, the photon will always be transmitted though the PBS. In the right illustration of Figure 1.3 we see the same measurement setup, but now rotated over 45 degrees (the photon still has a vertical polarization, though). Hence, the photon will be randomly (with equal probabilities) transmitted or reflected. If you are willing to accept that the optical measurements shown in Figure 1.3 indeed form an incompatible pair (showing this is beyond the scope of this example), then we see that the uncertainty principle supports our analysis: the measurement outcome in the left setup contains no uncertainty, whereas the outcome in the right setup is maximally uncertain.

Uncertainty relations are quantitative expressions of the uncertainty principle. As we will see later, some of these uncertainty relations are valuable tools for proving security in (quantum) cryptography. (In Chapter 5 we will come back to these uncertainty relations.) Although the uncertainty principle is stated for *pairs* of measurements, by now multiple uncertainty relations are known which hold for more than two measurements.

1.2.4 Implications to Information Theory and Computer Science

The mathematical theory of information, *information theory* [Sha48, CT06], is built on probability theory. Quantum mechanics gives rise to a generalization of information theory, called *quantum information theory* [NC00]. In particular, quantum mechanics generalizes the notion of *information*. *Quantum information* has some peculiar properties by which it clearly distinguishes itself from *classical* information. The foremost example is the *non-cloneability* of quantum information: while we can always copy a classical bit, we cannot in general copy an unknown quantum state.

Richard Feynman realized in 1982 [Fey82] that quantum mechanics can also be exploited for *computation*. Research indicates that computing with the help of quantum-mechanical effects, called *quantum computing*, solves certain problems more efficiently than classical computing (defined by the *Turing-machine* model of computation).

As a result of theoretical research into quantum computing, we know several *quantum algorithms* (algorithms designed for quantum computers). For example, Shor [Sho97] has invented an efficient quantum algorithm for integer factorization. On the practical side, however, the technology of quantum computers is still in its infancy. Although researchers have succeeded in building quantum computers with just a few qubits—they even managed to run Shor's algorithm on it, to factor the number 15—their designs do not yet scale well to a lot of qubits. Hence, up to now no one has succeeded in building a large enough quantum computer (with respect to the number of qubits) to solve instances of the factoring problem that cannot yet be solved in reasonable time using classical computers.

1.3 The Research Field of Quantum Cryptography

After having briefly introduced cryptography as well as quantum mechanics, we can now characterize *quantum cryptography* as a generalization of cryptography, in which a) the classical notion of information is replaced by quantum information, and b) the Turing-machine model of computation is replaced by quantum computation.

It will be helpful to subdivide quantum cryptography into multiple classes, based on whether the cryptographic task, as well as the protocol that realizes this task, is "classical" or "quantum." An example of a cryptographic task that is "quantum" is encrypting a quantum message; a "classical task" is for example establishing a common classical key. The latter is the goal of *quantum key distribution* (QKD) [BB84]. The word "quantum" in QKD refers to the *protocol*, which uses quantum communication to achieve the classical cryptographic task. A *quantum protocol* is a protocol that uses quantum communication and typically performs one or more measurements on these quantum states. A quantum protocol may also perform quantum computations.

To the best of our knowledge, a quantum task necessarily requires a quantum protocol for its realization, hence we will distinguish three classes (instead of all four combinations). Table 1.1 shows these three classes: (1) a quantum task realized by a quantum protocol, (2) a classical task realized by a quantum protocol, and (3) a classical task realized by a classical protocol. In all three classes, we assume the adversary (as well as malicious parties) to be "quantum," i.e., capable of storing and processing quantum information. In both the first and second class, the security

	"Fully Quantum"	Quantum Protocols for a Classical Task	"Post-Quantum Cryptography"
Cryptographic Task	*Quantum*	*Classical*	*Classical*
Protocol	*Quantum*	*Quantum*	*Classical*
Adversary	*Quantum*	*Quantum*	*Quantum*

Table 1.1: A helpful way to classify works in quantum cryptography.

is typically based on quantum-mechanical effects, which make it (under certain circumstances) possible to achieve very strong security guarantees. In this thesis, we will mainly focus on problems that belong the second class. The third class, in which merely the adversary is assumed to be quantum, is known as *post-quantum cryptography*.

For some tasks, like key distribution, quantum cryptography can provide security solely based on the laws of quantum mechanics, thus without further restrictions on the adversary. For other tasks, like secure 2PC [Yao82] or position-based cryptography [KMS11], this is not the case and some restrictions have to be included in the model, like restricting the adversary's quantum storage capabilities. In both of the above cases, quantum cryptography allows for *unconditional* security proofs, where one does not have to rely on unproven assumptions from complexity theory. Moreover, quantum cryptography typically provides *everlasting security*, which means that the restriction need only hold *during* the execution of the scheme, but not anymore after its execution; this is in contrast to classical cryptography based on a computational assumption, where the security of a scheme usually breaks down even if the attacker gains sufficient computing power only *after* the execution of the scheme.

1.3.1 Concrete Example: Quantum Key Distribution

Let us discuss QKD on a high level to give a concrete example of quantum cryptography.

We consider two parties at different locations, Alice and Bob, as well as a potential attacker, Eve. We require that Alice and Bob can communicate over an insecure quantum channel, as well as over a public authentic classical channel. By saying that the (quantum) channel is *insecure* we mean that Eve has complete control over it: she can capture, block, insert, modify or delay messages. The *public* property of the classical channel means that Eve can read every transmitted message sent over this channel, nonetheless, because it is *authentic*, she cannot inject or modify messages.

The goal of QKD is to establish a common classical secure key between Alice and Bob, where *secure* means that it is (essentially) uniformly distributed on its range and independent from Eve's (quantum) information. A QKD protocol is called *secure* if, except with negligible probability, it either aborts or establishes identical secure keys for Alice and Bob. Note that this is the best we can hope for, since Eve can always enforce the protocol to abort (e.g., by blocking all quantum communication).

The security of QKD is based on a consequence of the uncertainty principle: Eve cannot eavesdrop on the quantum channel without disturbing some of the quantum states that are transmitted over this channel. Alice and Bob can detect these disturbances, and Eve will get caught.[10]

Note that in case Alice and Bob merely have access to a classical channel that is not authentic, they can turn this channel into an authentic one by, e.g., using a message authentication code, for which they need an initial short shared key. Hence, a more appropriate name for QKD would be *quantum key expansion*, because in a typical practical setting Alice and Bob will not have an authentic classical channel and thus need a short authentication key to start with.

BB84 Quantum Key Distribution

We proceed by describing the well-known *BB84 protocol* due to Bennett and Brassard [BB84]. BB84 requires an optical quantum channel between Alice and Bob and is a so-called "prepare-and-measure" protocol. Alice is capable of preparing and sending polarized photons. For example, she sends these photons through a fiber or through free space. Bob has a measurement device that allows him to measure the polarization of incoming photos in two different polarization bases (characterized below). The BB84 protocol consists of four phases.

1. *Quantum Communication:* Alice sends polarized photons: for each photon she uses two random bits to choose the polarization. The first bit selects the basis, i.e., either the rectilinear basis (consisting of polarization angles $\{0^\circ, 90^\circ\}$) or the diagonal basis (angles $\{45^\circ, 135^\circ\}$); the second bit selects between the two angles within the selected basis. The latter bit will be called the *information bit*, which is said to be *encoded* in either the rectilinear or diagonal basis.

[10] Alternatively, we can say that by the *no-cloning* property of unknown quantum states (see Section 2.7.6), Eve cannot make a perfect copy of the (non-orthogonal) states that are sent over the quantum channel.

Bob chooses a random measurement basis for each incoming photon, records the classical binary measurement outcomes, and confirms receipt.

2. *Sifting and Error Estimation:* Alice and Bob announce over the public classical channel the bases they used, so that both parties can determine for which indices their bases coincide. For each such index (up to some errors), Bob's measurement outcome is supposed to coincide with the information bit; the bits belonging to other indices (where the bases do not coincide) are discarded. Then, Alice and Bob determine the error rate between their remaining bit strings by publicly announcing (and thus sacrificing) a random subset of the positions. A high error rate is an indication of the presence of Eve: when the error rate is above a certain threshold, the protocol aborts.

3. *Information Reconciliation:* Alice sends error-correction information to Bob, which allows him to correct the errors with respect to the remaining information bits, at the expense of leaking some information to Eve.[11] The result is called the *raw key*. Note that Eve has some (quantum) side information about this raw key: beyond the classical information leaked during information reconciliation, Eve may have quantum information obtained from tampering with the quantum channel.

   The following method for information reconciliation is based on the use of a binary linear code $\mathcal{C}$. Let us assume for simplicity that the codeword length of the code equals the raw-key length. Alice computes s_A, which is the syndrome of the raw key k_{raw} with respect to $\mathcal{C}$, and sends s_A to Bob. Bob computes s_B, the syndrome of his noisy version of the raw key, k_{noise}, with respect to $\mathcal{C}$, and then computes the sum (vector addition over $\mathbb{F}_2$) of the syndromes $s := s_A \oplus s_B$. Let v be the error vector of lowest Hamming weight corresponding to s (with respect to $\mathcal{C}$). Bob computes the reconciled raw key as $k_{\text{noise}} \oplus v$ (where addition is over $\mathbb{F}_2$).

4. *Privacy Amplification:* Alice and Bob have a common raw key about which Eve has some (quantum) side information. Alice and Bob do not know what Eve's side information looks like: Eve may know some values of individual bits, or certain parities of bits, her own quantum system can be *entangled* to certain bits, etc. Nevertheless, Alice and Bob can compute an upper bound on the *amount* of information that Eve can possibly have about the raw key,

[11] There also exist variants in which Bob sends the error-correction information to Alice (known as *reverse reconciliation*), or in which Alice and Bob interactively perform information reconciliation (e.g., the Cascade protocol [BS93]).

as a function of: a) the amount of information leakage caused by information reconciliation, and b) the error rate that they have determined in step 2. This then enables them to perform *privacy amplification*, which converts the raw key into a shorter key that is essentially secret to Eve. Given the amount of randomness in the raw key conditional on Eve's side information, the *privacy amplification theorem* (Theorem 2.65) relates the length of the produced secret key to the achieved level of security.

Key Rate versus Distance

With respect to QKD over optical fiber, several experimental and field setups have been demonstrated throughout the world [ECP$^+$05, PPM08, CHZ$^+$09, SFI$^+$11]. Recall that the *error rate* is an important parameter in QKD: for a given level of security it essentially determines the ratio between the raw-key length and the secret-key length. The noise inherent to an optical fiber contributes to this error rate, and the amount of optical noise is mainly determined by the length of the fiber. Hence, the overall performance of a QKD system is usually characterized by a key-rate versus distance curve, including the level of security.

It is infeasible to properly compare the performance of different existing QKD experiments (like the ones cited above) for several reasons. Firstly, to obtain a relation between the key-rate and the level of security, one needs a QKD proof that gives a *finite-key bound*,[12] which is not yet available for every QKD protocol used in those experiments. Secondly, not every experimenter uses the same security definition; for example, sometimes security is merely claimed against individual attacks,[13] and often the inferior accessible-information-based security notion is used, instead of the trace-distance-based notion. Thirdly, most publications about experimental QKD setups lack a thorough theoretical analysis of the achieved level of security.

Still, to give a rough indication of the state of the art as of 2012 (without focusing on the actual security level), key rates up to 1 Mbit/s over a distance of 50 km have been demonstrated in the lab, and roughly 300 kbit/s over an actual 45-km link in Japan. At larger distances of around 100 km, the key rate typically drops to the order of 1 kbit/s. Note that these figures are extremely low compared to typical bitrates of data networks (currently, in the order of gigabits per second). Furthermore, the

[12] A bound on the security of a key given in terms of parameters of the protocol, such as the key's length, instead of merely an asymptotic security claim.

[13] For a formal definition of *individual attacks*, *collective attacks* and *coherent attacks*, we refer to [BM10].

maximum allowable distance (for which one still gets a reasonable key rate) is in the order of 100 km and is not expected to increase significantly, because it is strongly related to intrinsic parameters such as the attenuation of the fiber. With respect to *free-space QKD*, Hughes *et al.* [HNDP02] have demonstrated a key-rate of 200 bits/s over 10 km.

Continuous-Variable QKD

Instead of using qubits[14] as carriers of quantum information, one could make use of higher-dimensional or infinite-dimensional systems to build a QKD protocol. This is the basic idea behind *continuous-variable QKD* (CV-QKD). For a detailed review of CV-QKD, we refer the reader to [WPGP$^+$12].

An advantage of CV-QKD is that standard telecom components can be used as detectors. More precisely, CV-QKD uses homodyne or heterodyne detection instead of photon counting, and this enables the use of PIN photodiodes, which are both cheaper and faster than the avalanche photodiodes that are used for single-photon detection [Hug04, Lev09].

Existing proofs for qubit-based protocols like BB84 do not automatically apply to the continuous-variable case. In 2009, Renner and Cirac [RC09] proved security of CV-QKD against the most general attacks, via an extension of a de Finetti theorem to infinite-dimensional systems. A more recent proof by Furrer *et al.* [FFB$^+$11] gives a tighter bound on the key rate, and is based on an extension of a recently discovered entropic uncertainty relation to infinite-dimensional systems.

Attacks on Implementations and Device-Independent QKD

Although QKD is often presented as being "unconditionally secure," in fact several conditions need to be met to guarantee security. The main conditions are that Alice and Bob need to have secure laboratories, they should have complete control over their devices, they need a trusted source of local randomness, and they can trust their local computers [Hän10].

It is often implicitly assumed that those conditions are met, while this is not always the case: successful attacks on (commercial) *implementations* of quantum-cryptographic protocols have already been reported. Those attacks typically violate one (ore more) of the conditions above, or they exploit gaps between an actual imple-

[14] A *qubit* is an elementary (two-dimensional) quantum system. An example of a qubit is the polarization of a single photon.

mentation of a protocol and its theoretical description, for which the mathematical proofs hold.

An example of such a gap is the difference between preparing states in the BB84 protocol and in practice; in the former, the states are prepared using a single-qubit source. At the time of this writing, the technology of single-photon sources—for example, using spontaneous parametric down-conversion—is not very mature yet. In many QKD implementations, the photon source is a strongly attenuated laser pulse, which has a Poisson-distributed number of output photons. These photons all have the same state, and hence a *photon-number splitting attack* [BLMS00] can be performed, by which the security is compromised. Hwang [Hwa03] proposed a method to specifically counter these attacks, using *decoy states*.

Another recent class of attacks takes full control over Bob's detector by "blinding" it with a laser pulse that can be injected *remotely* (outside Alice's and Bob's lab) into the fiber [LWW$^+$10]. The affected detector type is the *passively-quenched avalanche photodiode*, which is often used in practice.[15] To combat these *blinding attacks*, several countermeasures have been proposed.

Of course, adding ad-hoc countermeasures each time a new attack has been found is not a desirable approach. This motivates research towards *device-independent QKD* (DI-QKD), which aims at reducing the number of above-mentioned conditions (the conditions mentioned at the beginning of this section) to a minimum. In particular, the goal of DI-QKD is to remove most of the conditions related to Alice's and Bob's devices, and to design QKD schemes whose security relies on experimentally verifiable properties of the devices [Hän10].

1.3.2 Beyond QKD: Secure Two-Party Computation

In quantum key distribution, we consider two main parties that cooperate to establish a key that is secure against external parties. In this section we focus on a situation involving two mutually-distrustful parties. As discussed before in Section 1.1.2, the problem in secure two-party computation is to devise some protocol that enables Alice and Bob, having inputs X_1 and X_2 respectively, to compute their outputs, $f_1(X_1, X_2)$ and $f_2(X_1, X_2)$ respectively, for known and possibly randomized functions f_1 and f_2, but such that neither party learns anything beyond this.

[15] "Passively quenched" means that the photodiode is brought and kept in its electrical working region by means of a series resistor. By injecting a bright light pulse into the fiber, the behavior of the photodiode can be fully controlled, because the resulting photocurrent changes the biasing voltage over the photodiode through the presence of the resistor.

Impossibility of General Secure Two-Party Computation

An important fact that has been formalized and proven in several ways, is that without restrictions on the adversary (beyond the limitations imposed by the laws of quantum mechanics) secure quantum protocols for general secure two-party computation cannot exist [May97, Lo97, LC97, Col07, SSS09, BCS12]. From a theoretical point of view it is interesting to note that *without any restrictions on the adversary,* a quantum protocol can achieve stronger security properties than any classical protocol for certain functionalities (e.g., coin flipping and oblivious transfer). Nevertheless, those "stronger security properties" are often not strong enough to be of practical use.

Circumventing the Impossibility Results

The impossibility results mentioned above can be circumvented by adding restrictions to the model. The *bounded-quantum-storage model* (BQSM) restricts the adversary in the number of qubits that he can store. Given the current state of technology, it is reasonable to assume that a real adversary stays within the limits of this model (for an appropriately chosen upper bound). In the BQSM, it is possible to devise protocols for several cryptographic functionalities, and with strong security guarantees (strong enough to be of practical use). An important example is the protocol in the BQSM for 1-2 oblivious transfer. Other examples of protocols that have been proven secure in the BQSM include *Rabin oblivious transfer*,[16] identification, bit commitment and quantum key distribution [Sch07, DFSS07]. Although quantum key distribution is not a protocol involving mutually distrustful parties and can even be shown to be secure without any restrictions on the adversary, the BQSM allows for a QKD protocol with built-in authentication, where the (separate) authentication key can be re-used (this rules out a particular authentication-key exhaustion attack, more about this can be found in Section 2.11.2).

Coin Flipping

Since Blum [Blu81] formulated the problem of secure coin flipping in 1981, it has been an active area of research in (quantum) cryptography. We distinguish two types of coin tossing, *weak* and *strong*.

[16] In *Rabin oblivious transfer* Alice inputs a message and with probability one-half Bob receives this message perfectly; otherwise he remains completely ignorant about it. Alice remains ignorant about whether Bob actually received the message or not. A Rabin oblivious transfer is also called a secure erasure channel with erasure probability one half.

Weak coin tossing is sufficient for a setting where the two players are aware of each other's preferred outcome. E.g., in the example with grandmother's golden watch (page 17), it is known to both Alice and Bob that each of them wants to win the watch. And moreover, they will of course not cheat in a way such that the *other* party's chance of winning increases. Strong coin tossing tries to prevent cheaters from biasing the coin in *both* directions. Hence, strong coin tossing is applicable to a scenario where the particular outcome that a player prefers is *unknown* to the other player.

It is known that there cannot exist a *classical* protocol for weak (and hence no strong) coin tossing that provides any kind of protection against cheating in an *unconditional* setting (i.e., without assumptions) [DK02]. When relying on a computational intractability assumption one can achieve arbitrary small bias for weak as well as strong coin tossing.

In 2007, Mochon [Moc07] came up with a *quantum* protocol for weak coin flipping that achieves arbitrary small bias without any restrictions on the adversary. In 2009, Chailloux and Kerenidis [CK09] gave an optimal quantum protocol for strong coin flipping that achieves constant bias without any restrictions on the adversary, thereby matching an existing lower bound for strong coin flipping by Kitaev.

1.3.3 Beyond QKD: Position-Based Quantum Cryptography

Position-based cryptography uses the *geographical location* of a party (Alice) as its sole credential. A typical task in this context is position verification, in which Alice has to prove to a set of verifiers that she is indeed present at the position where she claims to be.

A first attempt to achieve position verification is a protocol where the verifiers each send a message to the prover, the prover computes a function of those messages and sends the result back to each verifier. Finally, the verifiers jointly estimate the position of the prover from the arrival times of the messages. Note that by special relativity, the messages sent in the protocol cannot travel faster than with the speed of light in vacuum. For simplicity, it is usually assumed that computation is instantaneous.

Chandran *et al.* [CGMO09] show that (classical) relativistic protocols for position verification cannot be *unconditionally* secure: coalitions of fake provers can always break such protocols. They also propose a position-verification protocol that is secure under an assumption that is similar to the bounded-storage and bounded-retrieval models, nevertheless, this assumption is not very practical.

In the search for different assumptions, researchers tried to exploit the no-cloning property of quantum states for position verification, by sending quantum messages between the prover and the verifier, where the prover is supposed to perform some quantum computation on these messages. Though, a recent impossibility result by Buhrman *et al.* [BCF^{+}11] shows that a very general class of protocols for *position-based quantum cryptography* can be broken if the coalition of fake provers pre-share a lot of entanglement: a double exponential (in the qubit size of the joint state shared by a pair of fake provers) number of EPR pairs. Note that the amount of entanglement required for this attack was recently reduced to single exponential [BK11]. The main challenge is to show that this exponential amount is really necessary to break these kind of protocols. In other words, can there exist protocols which can be easily executed by honest parties, while dishonest provers need *at least* an exponential amount of shared entanglement to break it? For initial work into this direction, see [BFSS11].

1.3.4 A Brief History

Quantum cryptography was invented in the late sixties by Stephen Wiesner. In his paper titled *Conjugate Coding*, Wiesner had come up with an idea for quantum money, which had the benefit of being impossible to counterfeit. Furthermore, he described the notion of a *multiplexing channel*—which would be re-invented about a decade later by Rabin under the name of *oblivious transfer*—as well as an implementation for it. Wiesner tried to publish his ideas but unfortunately his manuscript got rejected. He also told Charles Bennett about his idea, but neither Bennett nor himself succeeded in raising other people's interest. This changed when Gilles Brassard and Seth Breidbart joined Wiesner and Bennett; together they submitted a manuscript to the *CRYPTO '82* conference [BBBW82], in which they exposed Wiesner's idea applied to subway tokens instead of money, and introduced the term *quantum cryptography*. This time, the paper got accepted and this led to a renewed interest in the topic. This in turn enabled Wiesner to also publish his original manuscript [Wie83].

At the time, quantum cryptography was not considered to be practical: the quantum money (or subway tokens) required tiny and robust single-photon storage registers, which were—and still are—not available. This changed when Bennett, together with Gilles Brassard, realized that it would be more practical to transmit photons between remote locations than to store them in some register. After this realization, it still took them a while before they had found the "killer application," which turned out to be key agreement. This lead to the famous BB84 quantum key distribution protocol [BB84].[17]

From today's perspective, the initial description of BB84 was rather incomplete; it did not yet include information reconciliation (error correction) and privacy amplification. The latter was first described in [BBR88]. A more complete description (from today's perspective) of BB84 appeared in [BBB$^+$92]. Nonetheless, a proper security proof against the most general attacks (coherent attacks) was lacking; the authors merely claimed security against individual attacks (where the attacker operates on transmitted photons separately).

Meanwhile, different protocols for QKD had been proposed, for example Ekert's entanglement-based QKD protocol [Eke91], of which a simpler variant was shown to be equivalent to BB84 with respect to its security [BBM92]. Note that another well-known protocol for QKD is the *six-state QKD protocol* due to Bruss [Bru98].

Also, quantum protocols had appeared for *secure two-party computation* (2PC): coin tossing, bit commitment [BB84, BC90] and oblivious transfer [BBCS91]. Security of some of these protocols against attacks more general than the individual attacks was merely conjectured.

In 1996, Mayers [May97] and, independently, Lo and Chau [LC97] proved that any quantum bit commitment protocol, or, in fact, all quantum protocols for secure 2PC can be broken by a party with unlimited quantum storage and processing capabilities.

On the positive side, Mayers showed around the same year on the first security proof for QKD against the most general attacks [May95, May96] (see also [May01]). Some years later, Biham *et al.* gave a proof for QKD as well [BBB$^+$00]. Furthermore, Lo and Chau gave a security proof for a more complex QKD scheme (which required the honest parties to perform quantum computations). However, none of the proofs were fully satisfactory; the proofs of Mayers and Biham *et al.* were quite complicated, whereas Lo and Chau's proof was easier to understand but merely covered an impractical variant of QKD. In 2000, Shor and Preskill [SP00] gave a simple and intuitive proof for the BB84 protocol, which many regard as the first satisfactory QKD proof. It is based on the proof by Lo and Chau, but adapted such that Alice and Bob do not need quantum computers to execute the protocol, thereby making it compatible with BB84.

Until 2005, the standard way to formally define security in quantum cryptography was in terms of the *accessible information*. In 2005, König and Renner [RK05] (see

[17] BB84 should actually have been called BB83, since the protocol was first mentioned on a single-page *ISIT '83* abstract.

also [KRBM07]) and, independently, Ben-Or *et al.* [BOHL$^+$05] discovered that security definitions based on the accessible information do not achieve universal composability, due to a *locking* [DHL$^+$04] property of the accessible information notion. Both groups of authors then proposed a trace-distance based security definition, which provably does guarantee universal composability. In his PhD thesis [Ren05], Renner gave a QKD proof using the "right" trace-distance-based security definition.

Also around 2005, Damgård *et al.*, inspired by Maurer's (classical) bounded-storage model [Mau90], discovered that by restricting the adversary's quantum storage capabilities, Mayers' and Lo and Chau's impossibility results could be circumvented [DFSS05]. In the so-called *bounded-quantum-storage model* (BQSM), various 2PC functionalities can be proven secure, for example oblivious transfer, bit commitment and password-based identification. More recently, the BQSM was generalized to the *noisy-quantum-storage model* [WST08].

1.4 Thesis Outline, Contributions and Open Problems

Below, we will introduce the remaining chapters of this thesis. Chapter 2 covers basic results, Chapter 3, 4 and 5 are editions of work published earlier, respectively [BF10, BF11, BFGS12]. Chapter 3 and Chapter 5 are weakly related in that they both present a new quantum-information-theoretic tool, but apart from this the latter three chapters are rather "orthogonal" in the sense that they cover distinct topics.

1.4.1 Chapter 2: Preliminaries

Chapter 2 provides a theoretical foundation for the remainder of the thesis: we give introductions of probability theory, information theory, functional analysis, quantum mechanics and quantum information theory. Furthermore, we discuss several important concepts from cryptography like privacy amplification, authentication, extractors and identification.

Chapter 2 mainly consists of existing results, which are (up to some exceptions) stated without proof. Furthermore, the following results are well-known, but are hard to find in the literature, hence their proof is explicitly included: Proposition 2.53, Proposition 2.56, Proposition 2.58 and Proposition 2.62.

1.4.2 Chapter 3: Random Sampling from a Quantum Population

Chapter 3 belongs to a line of research aimed at finding new quantum-information-theoretic tools. In the context of quantum cryptography, such tools give rise to new quantum-cryptographic protocols, and are useful to *rigorously*[18] analyze the security of those and other quantum-cryptographic protocols. Recent examples of quantum-information-theoretic tools include a quantum-generalization of conditional min-entropy, a privacy amplification theorem that is based on this particular generalized min-entropy notion [Ren05, RK05] and various entropic uncertainty relations (see also Chapter 5). In Chapter 3, we present a tool to analyze random sampling in a quantum setting.

Random sampling allows us to learn information about a population (a collection of objects) by inspecting only a relatively small number of objects from that population. For example, an exit poll usually gives a very good prediction for the outcome of an election. In fact, classical sampling theory *guarantees* that a sample gives an accurate prediction for an entire population, except with a small error probability that can often be shown to decrease *exponentially* in the sample size (provided that the sample is selected appropriately, e.g., uniformly random over the collection).

In Chapter 3, we study the above problem in a quantum setting: when sampling some parts from a large quantum state (where sampling here means performing a *quantum measurement* on those parts), what can we conclude about the entire state?

We present a formal analysis of this problem, from which it becomes clear what exactly can be deduced from the measurement outcomes about the entire state. Additionally, we show a simple relation between the "error probability" in the quantum setting and the error probability from classical sampling theory. In particular, this relation implies that to find the error probability of any quantum sampling problem, it suffices to find the error probability of corresponding classical sampling problem, for which several good and useful bounds are known.

Many quantum-cryptographic protocols make use of random sampling to verify that a quantum state has some desired property. Hence, our results can be regarded as a useful quantum-information-theoretic tool to analyze such protocols. In particular, we present two new rigorous security proofs that make use of our new sampling tool: one for BB84 quantum key distribution, and one for a *quantum reduction* from

[18] With respect to QKD as well as other quantum protocols for basic cryptographic primitives (like OT), many of the early security proofs were not rigorous, but merely consisted of handwaving arguments.

oblivious transfer to bit commitment, by which we mean a quantum protocol for OT that requires black-box access to a bit-commitment primitive. Note that by the various impossibility results for secure 2PC, it is not possible to come up with a "stand-alone" (i.e., without relying on another primitive, such as bit commitment) quantum OT protocol that is secure against unrestricted quantum adversaries. Both proofs show security against coherent attacks (see footnote 13 on page 29), and they are relatively short, easy to understand, and non-asymptotic (i.e., they provide explicit security bounds).

1.4.3 Chapter 4: Authentication from a Weak Key with a Privacy Requirement

Chapter 4 studies a problem related to message authentication. The well-known method for non-interactive statistically-secure message authentication as described by Wegman and Carter requires a *uniformly random* authentication key. However, such a key may not always be available; e.g., the source of randomness used to produce the key might be imperfect, or an adversary may have gained partial knowledge about the key. Recently, the problem of authentication from a key with small min-entropy, also called a *weak key*, has received ample attention.[19] Maurer and Wolf [MW97] showed that when using a particular strongly universal family of functions, non-interactive message authentication is secure whenever the min-entropy rate[20] of the key is larger than $1/2$. On the other hand, for min-entropy rates below $1/2$, non-interactive secure message authentication is impossible [DS02, DW09]. Renner and Wolf [RW03], however, show by construction that *interactive* secure message authentication is possible for arbitrary min-entropy rates below $1/2$. Their construction requires a linear number of rounds of interaction (linear in the bit-length of the message). Subsequent work focused on reducing the number of rounds, the communication complexity and the entropy loss. Recently, this line of work has been closed by Xin Li [Li12], who presents a two-round message-authentication protocol with asymptotically optimal entropy loss and communication complexity.[21]

[19] A typical definition of a *weak key* is a random variable over bit strings of some finite length, where this length is denoted as n, whose min-entropy (conditional on the adversary's information) is an arbitrarily small fraction of n.

[20] The *min-entropy rate* of a random variable X whose range is the set of bit strings of length n, is defined as $H_{\min}(X)/n$.

[21] Actually, [Li12] presents an optimal privacy amplification protocol (optimal with respect to number of rounds, entropy loss and communication complexity). However, the protocol is actually an interactive message-authentication protocol, which is used to authenticate the seed for a randomness extractor. It is this extractor that turns the message-authentication protocol into a privacy-amplification protocol.

In Chapter 4, we study the problem of authentication from a weak key in a different but related scenario, in which the weak key is a one-time *session key* that is derived from a public source of randomness with the help of a *long-term* key (e.g., a password). This scenario occurs naturally in, e.g., Maurer's *bounded-storage model* [Mau90], where the long-term key is used to select a small number of bits from the huge bit string, which then together form the session key, as well as in the quantum setting, where the long-term key determines the measurement basis for a quantum measurement, whose classical outcome is used as session key. Our goal now is to authenticate a message using the weak session key, in such a way that nearly no information about the long-term key is leaked to the adversary. Ensuring privacy of the long-term key is vital for the long-term key to be re-usable. Previous work has not considered such a privacy issue, and previous solutions do not seem to satisfy this requirement.

We propose a new four-round protocol for message authentication from a weak (session) key. Given a secure look-ahead extractor, we prove that our protocol satisfies *security* against an active adversary and *long-term-key privacy*, which means that the protocol leaks essentially no information about the long-term key. For the setting where the adversary's side information about the session key is classical, we can use an existing construction for a secure look-ahead extractor. For the general case, in which this side information is a quantum state, we were not able to show the existence of a secure look-ahead extractor, and leave this as an open problem. The existence of the latter object has a direct implication to a problem related to identification in the bounded-quantum-storage model (this is discussed in more detail in Section 4.8).

1.4.4 Chapter 5: Hybrid Security of Password-Based Identification

In Chapter 5 we consider the task of password-based identification. Since password-based identification is an instance of secure 2PC, it is well-known that it is impossible to obtain a secure quantum protocol for identification without further restrictions on the adversary. In [DFSS07], Damgård *et al.* propose an identification scheme in the bounded-quantum-storage model, and show its security (see also Section 2.11). When using a security model like the BQSM to prove a scheme's security, a practical question arises, namely how to set the parameter(s) of the model (in case of the BQSM, this would be the number of qubits that the adversary can store) such that the behavior of a real adversary is likely to stay within the model. In particular, a wrong choice of such a model parameter might make a scheme lose all of its security guarantees. One particular approach to partly circumvent this problem is to devise schemes that are secure in two or more security models *simultaneously*. An example of this approach from the literature is [DFL$^+$09], which proposes a "compiler" which can add a layer of computational security (security based on the presumed computational difficulty of certain problems) to any quantum protocol for secure 2PC that consists of a BB84-quantum-communication phase followed by classical communication.

Chapter 5 also follows this approach and proposes a new quantum identification protocol that is secure in two security models simultaneously, i.e., in the BQSM and in a new model, called the *single-qubit-operations model* (SQOM).

In the BQSM, the security proof is based on a new *entropic uncertainty relation*, which is another main contribution of Chapter 5. Entropic uncertainty relations are quantitative characterizations of Heisenberg's uncertainty principle, which make use of an entropy measure to quantify uncertainty. In quantum cryptography, they are often used as convenient tools in security proofs. The new entropic uncertainty relation presented in Chapter 5 is the first uncertainty relation that lower bounds the uncertainty in the measurement outcome for *all but one* measurements, chosen from an arbitrary (and in particular an arbitrarily *large*) set of possible measurements. Besides this, it uses the *min-entropy* as entropy measure, rather than the Shannon entropy. The uncertainty relation might very well be useful in other quantum-cryptographic applications as well.

The SQOM models an adversary that has unbounded storage capabilities but is restricted to non-adaptive single-qubit operations. Hence, our new identification scheme also offers security in case the bounded-quantum-storage assumption fails to hold. The scheme by Damgård *et al.*, on the other hand, is completely insecure against an adversary in the SQOM. The security proof in the SQOM relies on a minimum-distance property of a random binary matrix and a XOR inequality by Diaconis and Shahshahani (Theorem 2.8). Due to the restriction to non-adaptive operations, the SQOM is not general enough to be practically relevant, hence our result of achieving security in the SQOM should be regarded as a stepping stone towards the open problem of achieving security in a more general restricted-operations model (which does, in particular, not restrict to non-adaptive operations).

2

Preliminaries

Chapter Contents

2.1 **Basic Notation** **42**

2.2 **Probability Theory** **43**

2.2.1 Random Variables 44

2.2.2 Some Important Distributions 45

2.2.3 The Bias of a Binary Random Variable 46

2.2.4 Distance between Distributions 47

2.2.5 Jensen's Inequality 49

2.2.6 Hoeffding's Inequality 50

2.3 **Classical Measures of Uncertainty** **51**

2.3.1 Rényi Entropies 52

2.3.2 Conditional Entropy 53

2.4 **Privacy Amplification** **55**

2.4.1 Universal Hashing 55

2.4.2 Extractors 58

2.5 **Statistically Secure Encryption and Authentication** **61**

2.5.1 The One-Time Pad 61

2.5.2 Message Authentication 61

2.6 **Hilbert Spaces** **62**

2.6.1 Dirac's Braket Notation 63

2.6.2 Operators 64

2.6.3 Tensor Products 65

2.6.4 Vector and Matrix Representations 68

2.7 **Quantum Systems and Operations** **70**

2.7.1 Postulates of Quantum Mechanics 70

2.7.2 States 72
2.7.3 Computational and Hadamard Basis 73
2.7.4 Partial Trace and Purification 73
2.7.5 The Schmidt Decomposition 74
2.7.6 Quantum Operations 75
2.7.7 Alternative Descriptions of Measurements 77
2.7.8 Entanglement 78
2.7.9 Hybrid Systems 79
2.7.10 Distance between States 81
2.8 The One-Time Pad in a Quantum Setting 83
2.9 Measures of Uncertainty for Density Matrices 83
2.10 Extractors against Quantum Side Information 88
2.11 Quantum Identification 89
2.11.1 The Basic QID Protocol 91
2.11.2 Security against Man-in-the-Middle Attacks 93

2.1 Basic Notation

We use $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Z}$, $\mathbb{N}$, $\mathbb{F}_q$, $\mathbb{F}_q^*$ for respectively the field of real numbers, the field of complex numbers, the ring of rational integers, the set of strictly positive integers, the finite field of order $q \in \mathbb{N}$, where $q = p^n$ for $p, n \in \mathbb{N}$ with p prime, and the multiplicative group of $\mathbb{F}_q$. If $q = p$, then we let $\mathbb{F}_q$ be the field $\mathbb{Z}/p\mathbb{Z}$. We adopt the following notational convention from computer science: in the particular case of $\mathbb{F}_{2^n}$ for arbitrary $n \in \mathbb{N}$, we use the $\oplus$ symbol to denote addition, and we also use the $\oplus$ symbol for vector addition in the vector space $\mathbb{F}_2^n$ for arbitrary $n \in \mathbb{N}$. For $x \in \mathbb{R}$ such that $x > 0$, $\log x$ denotes the *binary* logarithm of x, unless stated otherwise. We use e to denote the base of the natural logarithm (sometimes also called Euler's number). For any $\alpha \in \mathbb{C}$, $\bar{\alpha}$ denotes the complex conjugate of α. Let $[a, b]$ for any $a, b \in \mathbb{R}$ such that $b \geq a$ denote the closed real interval $\{x \in \mathbb{R} : a \leq x \leq b\}$. For $n \in \mathbb{N}$, we write $[n]$ for the set of integers from 1 to n, i.e., $[n] := \{1, \ldots, n\}$. For a matrix A with entries in $\mathbb{C}$, A^{T} and $A^{\dagger}$ respectively denote the transpose and conjugate transpose of A. Let K be a field, let V be a K-vector space and let $\mathcal{S}$ be a non-empty set of vectors from V. The K-linear span of $\mathcal{S}$ is denoted as $\mathrm{span}(\mathcal{S})$. (Usually, the field K will be clear from context.) For $i, j \in \mathbb{Z}$, δ_{ij} denotes the *Kronecker delta symbol*: $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise. Let $f : \mathbb{N} \to \mathbb{R}$ and $g : \mathbb{N} \to \mathbb{R}$ be arbitrary functions. We say that $f = O(g)$ if there exists a real constant $c > 0$ and a number $n_0 \in \mathbb{N}$ such that $|f(n)| \leq c \cdot g(n)$ for all $n \in \mathbb{N}$ such that $n \geq n_0$. We say that $f = \Theta(g)$ if there exist real constants $c > 0$ and $d > 0$ and a number $n_0 \in \mathbb{N}$ such that $c \cdot g(n) \leq f(n) \leq d \cdot g(n)$ for all $n \in \mathbb{N}$ such that $n \geq n_0$.

2.2 Probability Theory

A *finite probability space* is a pair (Ω, Pr), where Ω is a non-empty finite set called *sample space* and Pr is a probability function

$$\mathrm{Pr} : \Omega \to [0, 1],$$

with $\sum_{\omega \in \Omega} \mathrm{Pr}(\omega) = 1$. Note that we only consider finite probability spaces in this thesis. Subsets in Ω are called *events*. For any event $\mathcal{A} \subseteq \Omega$, the probability $\mathrm{Pr}[\mathcal{A}]$ of the event is given by $\mathrm{Pr}[\mathcal{A}] := \sum_{\omega \in \mathcal{A}} \mathrm{Pr}(\omega)$, where by convention $\mathrm{Pr}[\emptyset] = 0$.

The probability of an event $\mathcal{E} \subseteq \Omega$ *conditioned on an event* $\mathcal{A}$ with $\mathrm{Pr}[\mathcal{A}] > 0$ is given by

$$\mathrm{Pr}[\mathcal{E}|\mathcal{A}] := \frac{\mathrm{Pr}[\mathcal{E} \cap \mathcal{A}]}{\mathrm{Pr}[\mathcal{A}]}.$$

Sometimes, we write a comma instead of the set-intersection symbol ("$\cap$") to denote a joint event, i.e., $\mathrm{Pr}[\mathcal{E}, \mathcal{A}] = \mathrm{Pr}[\mathcal{E} \cap \mathcal{A}]$.

A consequence of the definition above is the *product rule* for events:

$$\mathrm{Pr}[\mathcal{E} \cap \mathcal{A}] = \mathrm{Pr}[\mathcal{E}|\mathcal{A}]\,\mathrm{Pr}[\mathcal{A}].$$

The following upper bound will be useful:

$$\mathrm{Pr}[\mathcal{E}|\mathcal{A}] = \frac{\mathrm{Pr}[\mathcal{A} \cap \mathcal{E}]}{\mathrm{Pr}[\mathcal{A}]} \leq \frac{\mathrm{Pr}[\mathcal{E}]}{\mathrm{Pr}[\mathcal{A}]}.$$

Another simple yet powerful inequality that we will use frequently is the *union bound*. For a finite set of events $\{\mathcal{A}_1, \ldots, \mathcal{A}_n\}$ where $\mathcal{A}_i \subseteq \Omega$ for any $i \in [n]$ it holds that

$$\mathrm{Pr}\Big[\bigcup_{i \in [n]} \mathcal{A}_i\Big] \leq \sum_{i \in [n]} \mathrm{Pr}[\mathcal{A}_i].$$

This means that the probability that at least one of the events $\mathcal{A}_i$ happens is no greater than the sum of the probabilities of all individual events.

2.2.1 Random Variables

Let $(\Omega, \Pr)$ be a finite probability space and let $\mathcal{X}$ be a non-empty finite set. A *random variable* X is a function

$$X : \Omega \to \mathcal{X}.$$

The set $\mathcal{X}$ is called the *range* of the random variable, and we say that X is a random variable *over* $\mathcal{X}$. An element $x \in \mathcal{X}$ with non-zero probability is also called an *outcome* or *realization* of the random variable X.

The *distribution* of X is the function $P_X : \mathcal{X} \to [0,1]$ defined as

$$P_X(x) := \Pr[X = x] \quad \text{for all } x \in \mathcal{X},$$

where $X = x$ is a shorthand for the event $\{\omega \in \Omega : X(\omega) = x\}$ and should be read as "the event that the random variable X takes on the value x." Note that it holds that $\sum_{x \in \mathcal{X}} P_X(x) = 1$; by definition, P_X inherits this property from $\Pr$. The *support* of a distribution is the set of elements from the range which have non-zero probability: $\mathrm{supp}(P_X) := \{x \in \mathcal{X} : P_X(x) > 0\}$.

We will often define events in terms of random variables. The shorthand notation $X = x$ used above is not limited to the equality function but extends also to other operations defined on $\mathcal{X}$, e.g., events like $X \neq x$ are defined similarly.

Convention 2.1 *In this thesis, we will often make a statement in which several random variables occur. Unless stated otherwise, these random variables are defined in the same probability space.*

The *joint distribution* of two (or more) random variables X and Y is denoted by P_{XY}, i.e., $P_{XY}(x,y) = \Pr[X{=}x, Y{=}y]$. The pair XY is the random variable

$$\begin{aligned} XY : \Omega &\to \mathcal{X} \times \mathcal{Y} \\ \omega &\mapsto (X(\omega), Y(\omega)) \end{aligned}$$

Given a joint distribution P_{XY}, the distribution for X (resp. Y) alone is obtained by *marginalizing* over Y (X),

$$P_X(x) = \sum_{y \in \mathcal{Y}} P_{XY}(x,y) \quad \text{for all } x \in \mathcal{X},$$

and P_X is then called a *marginal distribution*, and similarly for P_Y.

The *conditional distribution* of X conditioned on an event $\mathcal{A}$ with $\Pr[\mathcal{A}] > 0$ is given by

$$P_{X|\mathcal{A}}(x) := \frac{\Pr[X = x, \mathcal{A}]}{\Pr[\mathcal{A}]} \quad \text{for all } x \in \mathcal{X}.$$

If $\mathcal{A}$ is the event $Y = y$, then we denote the conditional distribution as $P_{X|Y=y}(x)$ or as $P_{X|Y}(x|y)$; these two forms of notation are used interchangeably.

Convention 2.2 *In writing equations with distributions, we often shorten the notation by omitting the parentheses containing the function arguments. Note that we will only use this shorthand if the omitted arguments can be reconstructed again without ambiguity from the subscripts of the distributions. For example, the product rule for distributions is written compactly as* $P_{XY} = P_{X|Y}P_Y$*, and this should be understood as* $P_{XY}(x, y) = P_{X|Y}(x|y)P_Y(y)$ *for any* $(x, y) \in \mathcal{X} \times \mathcal{Y}$ *for which* $P_Y(y) > 0$. *In case the quantification is not mentioned explicitly, it is assumed to be over all inputs for which all involved conditional probabilities are well-defined.*

Random variables X and Y are *independent* if $P_{XY} = P_X P_Y$. For n random variables $X_1, X_2, \ldots, X_n$ for n integer and $n > 2$, we say that they are independent (or: *mutually independent*) if $P_{X_1 X_2 \cdots X_n} = P_{X_1} P_{X_2} \cdots P_{X_n}$. An ordered sequence of random variables (X, Y, Z) forms a *Markov chain*, denoted as $X \leftrightarrow Y \leftrightarrow Z$ if $P_{XZ|Y} = P_{X|Y}P_{Z|Y}$. It is easy to verify that the latter expression is equivalent to $P_{Z|XY} = P_{Z|Y}$ as well as to $P_{X|YZ} = P_{X|Y}$.

We will also use the notion *distribution* without associating it to a particular random variable. In this case, we mean any non-negative real function $p : \mathcal{X} \to [0, 1]$, where $\mathcal{X}$ is a non-empty finite set, such that $p(x) \geq 0$ for all $x \in \mathcal{X}$ and $\sum_{x \in \mathcal{X}} p(x) = 1$. Any distribution can be understood as a *distribution of* some random variable X over $\mathcal{X}$ in some probability space $(\Omega, \Pr)$: a trivial construction for this probability space and random variable X is given by $\Omega := \mathcal{X}$, with $\Pr(\omega) := p(\omega)$ and $X(\omega) = \omega$ for any $\omega \in \Omega$. Hence, from here we will always define a random variable (or several random variables in the same probability space) by specifying the (joint) probability distribution, and leave the probability space $(\Omega, \Pr)$ implicit.

2.2.2 Some Important Distributions

Let X be a random variable over $\mathcal{X}$. We say that X *is uniformly distributed over* $\mathcal{X}$ if its distribution P_X equals

$$P_X(x) := |\mathcal{X}|^{-1} \quad \text{for all } x \in \mathcal{X}.$$

The distribution P_X is called the *uniform distribution* over $\mathcal{X}$. When we say that X is *random*, we actually mean it to be uniformly distributed over $\mathcal{X}$. We write $x \overset{\mathtt{r}}{\leftarrow} \mathcal{X}$

to denote that the element x is picked independently and uniformly at random from the set $\mathcal{X}$.

More generally, for a non-empty subset $\mathcal{I} \subseteq \mathcal{X}$, we say that X has a *flat distribution* on $\mathcal{I}$ if for all $x \in \mathcal{X}$

$$P_X(x) := \begin{cases} |\mathcal{I}|^{-1} & \text{if } x \in \mathcal{I}, \\ 0 & \text{otherwise.} \end{cases}$$

When we say that X "has a flat distribution" (thus without mentioning the subset $\mathcal{I}$), we mean that there *exists* a set $\mathcal{I}$ such that X has a flat distribution on $\mathcal{I}$.

We say that X is a *binary random variable* if its range $\mathcal{X} = \mathbb{F}_2$ and we say that it has a *Bernoulli distribution* with parameter p if $P_X(1) = p$.

The *binomial distribution* gives the probability that k out of n independent and identically distributed binary random variables X_i, $\forall i \in [n]$ (all having the same parameter p) take the value one. Let $S := \sum_{i \in [n]} X_i$ (where the sum is over the integers). Then, the binomial distribution is given by

$$\Pr[S = k] = \binom{n}{k} p^k (1-p)^{n-k}.$$

2.2.3 The Bias of a Binary Random Variable

The *bias* of a binary random variable X is defined as

$$\text{bias}(X) := \big|P_X(0) - P_X(1)\big|.$$

This also naturally defines the bias of X conditioned on an event $\mathcal{E}$ as

$$\text{bias}(X|\mathcal{E}) := \big|P_{X|\mathcal{E}}(0) - P_{X|\mathcal{E}}(1)\big|.$$

The bias thus ranges between 0 and 1 and can be understood as a degree of predictability of a bit: if the bias is small then the bit is close to random, and if the bias is large (i.e., approaches 1) then the bit has essentially no uncertainty.

Lemma 2.3 *For a sum of two independent binary random variables X_1 and X_2, the bias of the sum is the product of the individual biases:*

$$\text{bias}(X_1 \oplus X_2) = \text{bias}(X_1)\text{bias}(X_2).$$

Proof. Let us prove the case where $P_{X_1}(0) \geq P_{X_2}(1)$ and $P_{X_2}(0) \geq P_{X_2}(1)$. (The other cases follow similarly.)

$$\begin{aligned}
&\mathrm{bias}(X_1)\mathrm{bias}(X_2)\\
&\quad = |P_{X_1}(0) - P_{X_1}(1)| \cdot |P_{X_2}(0) - P_{X_2}(1)|\\
&\quad = (P_{X_1}(0) - P_{X_1}(1)) \cdot (P_{X_2}(0) - P_{X_2}(1))\\
&\quad = P_{X_1}(0)P_{X_2}(0) + P_{X_1}(1)P_{X_2}(1) - P_{X_1}(0)P_{X_2}(1) - P_{X_1}(1)P_{X_2}(0)\\
&\quad = |P_{X_1}(0)P_{X_2}(0) + P_{X_1}(1)P_{X_2}(1) - P_{X_1}(0)P_{X_2}(1) - P_{X_1}(1)P_{X_2}(0)|\\
&\quad = \mathrm{bias}(X_1 \oplus X_2).
\end{aligned}$$

□

2.2.4 Distance between Distributions

Let $\mathcal{P}_{\mathcal{X}}$ be the set of non-negative real-valued functions on $\mathcal{X}$. *Statistical distance* is the function defined as

$$\begin{aligned}
\mathrm{SD} : \mathcal{P}_{\mathcal{X}} \times \mathcal{P}_{\mathcal{X}} &\to \mathbb{R}\\
(p, q) &\mapsto \tfrac{1}{2} \sum_{x \in \mathcal{X}} |p(x) - q(x)|.
\end{aligned}$$

The name *statistical distance* stems from its typical use as a distance measure for probability distributions.

For random variables X and Y that have the same range, we may also write $\mathrm{SD}(X, Y)$, where the latter should be understood as the statistical distance between their distributions P_X and P_Y.

The statistical distance is a metric, i.e., it has the following properties for all $p, q, r \in \mathcal{P}_{\mathcal{X}}$,

1. *Non-negativity:* $\mathrm{SD}(p, q) \geq 0$;
2. *Identity of indiscernibles:* $\mathrm{SD}(p, q) = 0 \iff p = q$;
3. *Symmetry:* $\mathrm{SD}(p, q) = \mathrm{SD}(q, p)$;
4. *Triangle inequality:* $\mathrm{SD}(p, r) \leq \mathrm{SD}(p, q) + \mathrm{SD}(q, r)$.

In the literature, the statistical distance is sometimes defined without the factor $\frac{1}{2}$; we include it deliberately because then the statistical distance equals the *distinguishing advantage*, i.e., the maximum difference in probability that p and q assign to the same event (where the maximum is taken over all events).

Theorem 2.4 *For all distributions p and q on $\mathcal{X}$, it holds that*

$$\mathrm{SD}(p,q) = \max_{\mathcal{A}\subseteq\mathcal{X}} \Big(p[\mathcal{A}] - q[\mathcal{A}]\Big)$$

This theorem is well known, for a proof see for example Theorem 6.15 in [Sho05].

Definition 2.5 Let P_X and P_Y be distributions on $\mathcal{X}$. A (joint) distribution p on $\mathcal{X} \times \mathcal{X}$ is a *coupling* of (P_X, P_Y) if its two marginal distributions are P_X and P_Y respectively, that is,

$$\sum_{y\in\mathcal{X}} p(x,y) = P_X(x) \quad \forall x \in \mathcal{X}, \quad \text{and} \quad \sum_{x\in\mathcal{X}} p(x,y) = P_Y(y) \quad \forall y \in \mathcal{X}.$$

The following two theorems comprise the "coupling interpretation" of statistical distance.

Theorem 2.6 (Coupling Inequality, see, e.g., [Lin92, Ch. 1, (2.6)]) *Let P_X and P_Y be distributions on $\mathcal{X}$. For any coupling P_{VW} of (P_X, P_Y) it holds that*

$$\mathrm{SD}(P_X, P_Y) \leq \Pr[V \neq W].$$

Theorem 2.7 (Maximal Coupling, see, e.g., [Lin92, Ch. 1, Thm. 5.2]) *Let P_X and P_Y be distributions on $\mathcal{X}$. Then there exists a unique coupling $P_{\widetilde{V}\widetilde{W}}$ of (P_X, P_Y) such that*

$$\mathrm{SD}(P_X, P_Y) = \Pr[\widetilde{V} \neq \widetilde{W}],$$

and

$$P_{\widetilde{V}\widetilde{W}|\widetilde{V}\neq\widetilde{W}} = P_{\widetilde{V}|\widetilde{V}\neq\widetilde{W}} P_{\widetilde{W}|\widetilde{V}\neq\widetilde{W}}.$$

The following theorem gives a useful upper bound on the statistical distance between a random variable X and a uniform random variable U (having the same range), in terms of the biases of all $\mathbb{F}_2$ linear functions with binary outputs applied to X.

Theorem 2.8 (Diaconis and Shahshahani) *Let X be a random variable over $\mathcal{X}$ with distribution P_X, where $\mathcal{X} := \mathbb{F}_2^n$, and let $U_{\mathcal{X}}$ be an independent random variable that is uniformly distributed over $\mathcal{X}$. Then, the following holds,*

$$\mathrm{SD}(P_X, U_{\mathcal{X}}) \leq \frac{1}{2}\Big[\sum_{f\in\mathbb{F}_2^n\setminus\{\mathbf{0}\}} \mathrm{bias}(f\bullet X)^2\Big]^{\frac{1}{2}}.$$

where $\mathbf{0}$ denotes the zero vector in $\mathbb{F}_2^n$ and $f \bullet X$ means the standard inner product on $\mathbb{F}_2^n$ between f and X.

The original version of Theorem 2.8 appeared in [Dia88], where it is expressed in the language of representation theory. The version above is due to [NN93].

2.2.5 Jensen's Inequality

Definition 2.9 A real-valued function $f : \mathcal{I} \to \mathbb{R}$ on an arbitrary interval[1] $\mathcal{I}$ in $\mathbb{R}$ is said to be *convex* if for all $x_1, x_2 \in \mathcal{I}$ and for all $\lambda \in [0,1] \subset \mathbb{R}$ it holds that

$$\lambda f(x_1) + (1-\lambda) f(x_2) \geq f(\lambda x_1 + (1-\lambda) x_2).$$

The function f is called *strictly convex* if equality holds only at the endpoints (i.e., when $\lambda \in \{0,1\}$) or when $x_1 = x_2$. The function f is (strictly) *concave* if $-f$ is (strictly) convex.

In other words, a function f is convex if and only if all chords lie above or on the graph of f.

Examples of strictly *concave* functions include the square-root function: $\{x \in \mathbb{R} : x \geq 0\} \to \{x \in \mathbb{R} : x \geq 0\}$, $x \mapsto \sqrt{x}$ and the logarithm function: $\{x \in \mathbb{R} : x > 0\} \to \mathbb{R}$, $x \mapsto \log x$ for arbitrary base strictly larger than 1. Straight lines in the plane $\mathbb{R}^2$ are both convex and concave.

Theorem 2.10 (Jensen's Inequality, see, e.g., [CT06][2]) *Let $f : \mathcal{I} \to \mathbb{R}$ be a convex function on an arbitrary interval $\mathcal{I} \subset \mathbb{R}$. Then, for any $x_1, \ldots, x_n \in \mathcal{I}$ and any $p_1, \ldots, p_n \in \mathbb{R}$ such that $p_i \geq 0$ for all $i \in [n]$ and $\sum_i p_i = 1$, it holds that*

$$\sum_i p_i f(x_i) \geq f\Big(\sum_i p_i x_i\Big).$$

If f is strictly convex and $p_i > 0$ for all $i \in [n]$, then equality holds if and only if $x_1 = \ldots = x_n$. If f is concave then the inequality should be reversed.

A *real* random variable is a random variable whose range is contained in $\mathbb{R}$. The *expectation* (or *expected value*) of a real random variable X is defined as

$$\mathbb{E}[X] := \sum_{x \in \mathcal{X}} P_X(x) \cdot x.$$

In this case, we can rewrite Jensen's inequality using the expectation as follows,

$$\mathbb{E}[f(X)] \geq f(\mathbb{E}[X]).$$

[1] open, closed, or neither.

[2] Although [CT06] states Jensen's inequality for a convex function on an *open* interval, their proof does not make use of this.

2.2.6 Hoeffding's Inequality

When one repeatedly flips a fair coin, it is well known that the empirical frequencies of obtaining heads and tails, divided by the total number of coin flips, will ultimately both approach one half. In general, this phenomenon is known as the *law of large numbers*, and is also called *concentration of measure*. The phenomenon can be formalized in many ways, in weaker and stronger forms. A particularly convenient version, which is strong enough for us, is Hoeffding's inequality [Hoe63, DP09]. We will make use of it in several proofs. We state it here for mutually independent binary random variables.

Theorem 2.11 (Hoeffding's Inequality, Mutually Independent Random Variables) *Let $X_1, X_2, \ldots, X_n$ be mutually independent binary random variables, each distributed according to the Bernoulli distribution with the same parameter $\mu \in [0,1]$, and let $\bar{X} := \frac{1}{n}\big|\{i \in [n] : X_i = 1\}\big|$. Then for all $t \in \mathbb{R}$ such that $0 \leq t \leq 1 - \mu$ it holds that*

$$\Pr[\bar{X} - \mu \geq t] \leq \exp(-2nt^2).$$

For a proof, the reader is referred to the original paper by Hoeffding [Hoe63], or to a recent book on concentration of measure by Dubhashi and Panconesi [DP09]. An easy observation is that by applying Theorem 2.11 to the random variable $1 - X_i$, we obtain the same upper bound as above for $\Pr[-\bar{X} + \mu \geq t]$. Hence we get the following corollary.

Corollary 2.12 (Two-Sided Version of Theorem 2.11) *Let $\{X_i\}_{i \in [n]}$, $\bar{X}$, μ and t be as in Theorem 2.11. Then, the following holds*

$$\Pr[|\bar{X} - \mu| \geq t] \leq 2\exp(-2nt^2).$$

Random Variables with a Particular Type of Dependence

Above we have stated Hoeffding's inequality for mutually independent random variables. We will also make use of a version that applies to random variables that are dependent in the following sense.

Theorem 2.13 (Hoeffding's Inequality, RVs with a Particular Type of Dependence) *Let $b = (b_1, \ldots, b_\ell) \in \mathbb{F}_2^\ell$ be a bit string of length $\ell > 0$. Let $n \in \mathbb{N}$ such that $n \leq \ell$. Let $I \subset [\ell]$ be a uniformly distributed random variable over all size-n subsets of $[\ell]$. Let $I_1 < I_2 < \ldots < I_n$ denote the elements of I. Let $Y_i := b_{I_i}$ for all $i \in [n]$.*

Furthermore, let $\bar{Y} := \frac{1}{n}\left|\{i \in [n] : Y_i = 1\}\right|$ *and* $\mu := \frac{1}{\ell}\left|\{j \in [\ell] : b_j = 1\}\right|$. *Then, for any* $t \in \mathbb{R}$ *such that* $0 \leq t \leq 1 - \mu$*, it holds that*

$$\Pr[\bar{Y} - \mu \geq t] \leq \exp(-2nt^2).$$

The proof of Theorem 2.13 can be found in Section 6 of [Hoe63], and in [DP09].

Similarly to Corollary 2.12, we get the following corollary.

Corollary 2.14 (Two-Sided Version of Theorem 2.13) *Let* $\{Y_i\}_{i\in[n]}$, $\bar{Y}$, μ *and* t *be as in Theorem 2.13. Then, the following holds*

$$\Pr[|\bar{Y} - \mu| \geq t] \leq 2\exp(-2nt^2).$$

Serfling [Ser74] proves that the bound from Theorem 2.13 can be strengthened as follows.

Theorem 2.15 (Serfling's Inequality) *Let* $b = (b_1, \ldots, b_\ell) \in \mathbb{F}_2^\ell$ *be a bit string of length* $\ell > 0$. *Let* $n \in \mathbb{N}$ *such that* $n \leq \ell$. *Let* $I \subset [\ell]$ *be a uniformly distributed random variable over all size-*n *subsets of* $[\ell]$. *Let* $I_1 < I_2 < \ldots < I_n$ *denote the elements of* I. *Let* $Y_i := b_{I_i}$ *for all* $i \in [n]$. *Furthermore, let* $\bar{Y} := \frac{1}{n}\left|\{i \in [n] : Y_i = 1\}\right|$ *and* $\mu := \frac{1}{\ell}\left|\{j \in [\ell] : b_j = 1\}\right|$. *Then for all* $t \in \mathbb{R}$ *such that* $t \geq 0$,

$$\Pr\big[\bar{Y} - \mu \geq t\big] \leq \exp\big(-\tfrac{2t^2 n\ell}{\ell - n + 1}\big).$$

This one-sided bound implies the following two-sided bound:

$$\Pr\big[|\bar{Y} - \mu| \geq t\big] \leq 2\exp\big(-\tfrac{2t^2 n\ell}{\ell - n + 1}\big),$$

which follows in the same way as Corollary 2.12.

2.3 Classical Measures of Uncertainty

In this section we define some notions from classical information theory, where "classical" means "non-quantum." Many of these notions will be generalized to the quantum case later, where we will also give more properties. Nonetheless, having an understanding of the notions in the classical case helps to understand them in the quantum case.

Definition 2.16 The *Shannon entropy* [Sha48] of a distribution $p : \mathcal{X} \to [0,1]$ is defined as

$$H(p) := -\sum_{x\in\mathrm{supp}(p)} p(x)\log p(x),$$

Since $\log$ expresses the binary logarithm, the Shannon entropy is expressed in bits. By the convention $0 \log 0 = 0$, which is justified by the fact that $\lim_{y \to 0} y \log y = 0$, we can slightly simplify the definition of Shannon entropy and write $H(p) := -\sum_{x \in \mathcal{X}} p(x) \log p(x)$, where $\mathcal{X}$ is the domain of p.

We write $H(X)$ for the Shannon entropy of (the distribution of) a random variable X, i.e., $H(X)$ should be understood as $H(P_X)$.

One interpretation of the Shannon entropy of a random variable X is the number of bits that are needed *on average* to encode an outcome $x \in \mathcal{X}$. Note that there are also many other important interpretations of Shannon entropy.

The *binary entropy function* $h : [0,1] \to [0,1]$ is defined as

$$h(p) := -(p \log(p) + (1-p) \log(1-p))$$

for $0 \leq p \leq 1$. (Recall the convention $0 \log 0 = 0$.)

2.3.1 Rényi Entropies

The Shannon entropy is a special case of a more general class of entropy measures, called the *entropies of order* α [Rén61].

Definition 2.17 For any $\alpha \in \mathbb{R}$ such that $\alpha > 0$ and $\alpha \neq 1$, the *Rényi entropy of order* α of a distribution $p : \mathcal{X} \to [0,1]$ is defined as

$$H_\alpha(p) := \frac{1}{1-\alpha} \log \sum_{x \in \operatorname{supp}(p)} p(x)^\alpha.$$

Similar to the Shannon entropy we write $H_\alpha(X)$ for the Rényi entropy of a random variable X, but we stress that strictly speaking the entropy is a function of the distribution P_X. Below, X denotes an arbitrary random variable over arbitrary $\mathcal{X}$ with arbitrary distribution P_X.

When taking the limit $\alpha \to 1$, we obtain the Shannon entropy:

$$H_1(X) := \lim_{\alpha \to 1} H_\alpha(X) = H(X).$$

The case $\alpha = 2$ is called the *collision entropy*

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

For $\alpha \to \infty$ we obtain the *min-entropy*:

$$H_{\min}(X) = H_\infty(X) := \lim_{\alpha \to \infty} H_\alpha(X) = -\log \max_{x \in \mathcal{X}} P_X(x).$$

It will be convenient to define the *collision probability* and the *guessing probability* of a random variable X as respectively

$$p_{\mathsf{col}} := \sum_{x\in\mathcal{X}} P_X(x)^2 \quad \text{and} \quad p_{\mathsf{guess}}(X) := \max_{x\in\mathcal{X}} P_X(x),$$

such that we can alternatively define the collision entropy and min-entropy of a random variable X as

$$H_2(X) := -\log p_{\mathsf{col}}(X) \quad \text{and} \quad H_{\min}(X) := -\log p_{\mathsf{guess}}(X).$$

For $\alpha \to 0$ we obtain the *max-entropy*,

$$H_{\max}(X) = H_0(X) := \lim_{\alpha\to 0} H_\alpha(X) = \log|\mathrm{supp}(P_X)|.$$

A useful operational meaning of the max-entropy of a random variable X is the number of bits needed to store a single realization of X.

The following proposition states an important property of Rényi entropy.

Proposition 2.18 *For all $\alpha, \beta \in \mathbb{R}$ such that $0 \leq \alpha < \beta$ and for all random variables X, it holds that*

$$H_\alpha(X) \geq H_\beta(X),$$

with equality if and only if X has a flat distribution.

A proof of this statement can be found in [Cac97].[3] Applied to the notions that we have just defined, we get

$$H_{\min}(X) \leq H_2(X) \leq H(X) \leq H_{\max}(X).$$

2.3.2 Conditional Entropy

In the following, let X and Y be random variables over $\mathcal{X}$ and $\mathcal{Y}$ respectively with joint probability distribution P_{XY}, and let $\mathcal{A}$ be an event such that $\Pr[\mathcal{A}] > 0$.

The Shannon entropy naturally extends to the Shannon entropy *conditional on an event $\mathcal{A}$*, i.e.,

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

[3] In [Cac97], Proposition 2.18 is stated slightly differently due to a different definition of $H_0(X)$. Nonetheless, the proof applies.

Recall that $Y = y$ for $y \in \mathcal{Y}$ represents an event, hence writing $H(X|Y = y)$ is now well-defined. Sometimes, we want to express the uncertainty of a random variable X *when given a random variable* Y, without fixing that random variable to a particular outcome y. For this purpose, the *conditional entropy* is defined as

$$H(X|Y) := \sum_{y \in \mathcal{Y}} P_Y(y) H(X|Y = y).$$

Note that we will also use the term *conditional entropy* for the entropy conditional on an event. By the above, the following is also naturally defined,

$$H(X|Y, \mathcal{A}) = \sum_{y \in \mathcal{Y}} P_{Y|\mathcal{A}}(y) H(X|Y = y, \mathcal{A}).$$

The *chain rule* for Shannon entropy states that

$$H(XY) = H(X) + H(Y|X),$$

for arbitrary random variables X and Y. Using the chain rule, the conditional Shannon entropy can alternatively be defined as $H(X|Y) := H(XY) - H(Y)$.

We will not define a conditional version of the Rényi entropy for arbitrary order α. Instead, we merely define conditional versions of the collision entropy and min-entropy. In order to do so, we first define the *conditional collision probability* and *conditional guessing probability*. The collision probability resp. guessing probability of X conditional on $\mathcal{A}$ are naturally defined as

$$p_{\mathsf{col}}(X|\mathcal{A}) = \sum_{x \in \mathcal{X}} P_{X|\mathcal{A}}(x)^2, \quad \text{and} \quad p_{\mathsf{guess}}(X|\mathcal{A}) = \max_{x \in \mathcal{X}} P_{X|\mathcal{A}}(x).$$

For a random variable Y, the conditional collision probability resp. conditional guessing probability of X given Y are defined as

$$\begin{aligned} p_{\mathsf{col}}(X|Y) &:= \sum_{y \in \mathcal{Y}} P_Y(y)\, p_{\mathsf{col}}(X|Y = y), \\ p_{\mathsf{guess}}(X|Y) &:= \sum_{y \in \mathcal{Y}} P_Y(y)\, p_{\mathsf{guess}}(X|Y = y). \end{aligned} \tag{2.1}$$

The *collision entropy conditional on an event* $\mathcal{A}$ and the *min-entropy conditional on an event* $\mathcal{A}$ are then given by, respectively

$$H_2(X|\mathcal{A}) = -\log p_{\mathsf{col}}(X|\mathcal{A}) \quad \text{and} \quad H_{\min}(X|\mathcal{A}) = -\log p_{\mathsf{guess}}(X|\mathcal{A}).$$

Definition 2.19 For a random variable Y, the *conditional collision entropy* resp. *conditional min-entropy*[4] of X given Y are defined as

$$H_2(X|Y) := -\log p_{\mathsf{col}}(X|Y), \quad \text{and} \quad H_{\min}(X|Y) := -\log p_{\mathsf{guess}}(X|Y).$$

The following relation always holds between the conditional versions of Shannon, collision and min-entropy:

$$H(X|Y) \geq H_2(X|Y) \geq H_{\min}(X|Y),$$

for all random variables X and Y. The proof of the left inequality uses Jensen's inequality; the proof of the right inequality follows from elementary observations.

2.4 Privacy Amplification

Suppose that Alice has a random variable X about which Eve has *side information*, modeled by a random variable Y that depends on X. *Privacy amplification* [BBR88, BBM95, HILL99] provides a way to extract a random variable K that has a smaller range than X, such that K is very close (in statistical distance) to a random variable U (having the same range as K) that is uniformly distributed and independent from Y. The random variable K will typically serve as a secret key.

First, we will explain privacy amplification in its original form, i.e., in the language of universal hashing. Subsequently, we will give an more general description in terms of extractors.

2.4.1 Universal Hashing

Definition 2.20 Let $\mathcal{G} := \{g_i\}_{i \in \mathcal{I}}$ be a family of functions $g_i : \mathcal{X} \to \mathcal{R}$, where $\mathcal{I}$, $\mathcal{X}$ and $\mathcal{R}$ are finite and non-empty sets. Let I be a random variable that is uniformly distributed over $\mathcal{I}$. The family $\mathcal{G}$ is called *universal* [CW77] if for all $x, x' \in \mathcal{X}$ such that $x \neq x'$ it holds that

$$\Pr[g_I(x) = g_I(x')] \leq \frac{1}{|\mathcal{R}|}.$$

There exist several constructions of universal families [CW77]. We give two constructions that will be used in later chapters. Let $n, r \in \mathbb{N}$ be such that $r \leq n$. The

[4] The definition for conditional min-entropy that we use is sometimes called *average* conditional min-entropy in the literature.

first family, $\mathcal{G}_1$, is well known:

$$\mathcal{G}_1 := \{f_A : A \in \mathbb{F}_2^{r \times n}\}$$

with

$$\begin{aligned} f_A : \mathbb{F}_2^n &\to \mathbb{F}_2^r \\ x &\mapsto Ax. \end{aligned}$$

To see that it is universal, note that for any $x, x' \in \mathbb{F}_2^n$ such that $x \neq x'$, the expression $Ax \oplus Ax' = A(x \oplus x')$, when viewed as the function $\mathbb{F}_2^{r \times n} \to \mathbb{F}_2^r$, $A \mapsto A(x \oplus x')$ is linear and surjective, hence the cardinality of $\{A \in \mathbb{F}_2^{n \times r} : A(x \oplus x') = y\}$ is the same for every $y \in \mathbb{F}_2^r$. Therefore, for A chosen uniformly at random, $\Pr[Ax = Ax'] = 2^{-r}$.

The second universal family, $\mathcal{G}_2$, is a variant on existing constructions of universal families, and will be of use in the proof of Theorem 4.6. In order to define it, we first introduce some more notation. For all $n \in \mathbb{N}$ where $n \geq 1$ and all $t \in [n]$ we let

$$[\cdot]_t : \mathbb{F}_{2^n} \to \mathbb{F}_{2^t}$$

be an arbitrary but fixed $\mathbb{F}_2$-linear surjective function. For example, when fixing a basis for $\mathbb{F}_{2^n}$, $[\cdot]_t$ can be defined as the projection on the subspace spanned by the first t basis vectors. Let $m, r \in \mathbb{N}$ be such that $r \leq m$. The second family is given by

$$\mathcal{G}_2 := \{h_{a,b} : a \in \mathbb{F}_{2^m}, b \in \mathbb{F}_{2^r}\}$$

with

$$\begin{aligned} h_{a,b} : \mathbb{F}_{2^m} \times \mathbb{F}_{2^r} &\to \mathbb{F}_{2^r} \\ (x, y) &\mapsto [a \cdot x]_r \oplus b \cdot y. \end{aligned}$$

Proposition 2.21 *The family $\mathcal{G}_2$ is universal.*

Proof. Let A and B be uniformly random over $\mathbb{F}_{2^m}$ and $\mathbb{F}_{2^r}$ respectively. Let (x, y), $(x', y') \in \mathbb{F}_{2^m} \times \mathbb{F}_{2^r}$ such that $(x, y) \neq (x', y')$ and consider

$$([A \cdot x]_r \oplus B \cdot y) \oplus ([A \cdot x']_r \oplus B \cdot y') = [A \cdot (x \oplus x')]_r \oplus B \cdot (y \oplus y').$$

Let us find the probability that this expression vanishes. In the following, we analyze the expression at the RHS of the equation above. In case $x = x' \wedge y \neq y'$, the left term vanishes and the right part is the bijective mapping $B \mapsto B \cdot (y \oplus y')$, and vanishes with probability 2^{-r}. In case $x \neq x' \wedge y = y'$, the right term vanishes and the left part inside the brackets, $A \mapsto A \cdot (x \oplus x')$, is a bijection and hence results in a uniformly distributed random variable over $\mathbb{F}_{2^m}$. Because the map $[\cdot]_r$ is linear and surjective, applying this map to the latter random variable results in a uniformly distributed random variable over $\mathbb{F}_{2^r}$, which equals zero with probability 2^{-r}. In case both $x \neq x'$ and $y \neq y'$, condition on $A = a$. For any a, by the bijective mapping $B \mapsto B \cdot (y \oplus y')$ the probability that the entire expression vanishes is 2^{-r}. In all cases, the probability that the expression vanishes equals 2^{-r}, hence the assertion follows. □

We want to be able to apply the functions of family $\mathcal{G}_2$ to vectors from an $\mathbb{F}_2$ vector space (bit strings). Hence, for every $n \in \mathbb{N}$ we fix a basis of $\mathbb{F}_{2^n}$, by which we can associate every vector in $\mathbb{F}_2^n$ with a unique field element in $\mathbb{F}_{2^n}$, and *vice versa*. We stress that the induced vector space isomorphism $\mathbb{F}_2^n \to \mathbb{F}_{2^n}$ is not a natural one; it depends on the chosen basis.

The Privacy Amplification Theorem

The following theorem is a modern variant of the original treatment of privacy amplification as found in [BBR88, BBM95, HILL99].

Theorem 2.22 (Privacy Amplification, Case of Classical Side Information) *Let $\mathcal{G} := \{g_i\}_{i \in \mathcal{I}}$ be a universal family of hash functions $g_i : \mathcal{X} \to \{0,1\}^r$, where $\mathcal{X}$ and $\mathcal{I}$ are finite and non-empty sets. Let X and Y be random variables over $\mathcal{X}$ and $\mathcal{Y}$ respectively, and let I and U be uniformly distributed over $\mathcal{I}$ and $\{0,1\}^r$ respectively, and such that I, U and XY are mutually independent. Let $K := g_I(X)$. Then*

$$\mathrm{SD}(KYI; UYI) \leq \tfrac{1}{2} \cdot 2^{-\frac{1}{2}(H_2(X|Y)-r)}$$

The proof can be found in [CF11]. The proof of a *quantum version*[5] of privacy amplification, which implies the statement above, can be found in [Ren05] (Theorem 5.5.1).

Note that the privacy amplification theorem is a powerful tool; it guarantees that the security of the extracted key increases *exponentially* in the gap $H_2(X|Y) - r$.

Also note that the functions from the family $\mathcal{G}_1$ introduced earlier each have range $\mathbb{F}_2^r$, while the privacy amplification theorem is stated in terms of hash functions

[5] In that version, Y is a quantum system that holds quantum information about X.

having range $\{0,1\}^r$. This should however not cause any confusion because of the elements of $\mathbb{F}_2$ are 0 and 1.

Let us now discuss the amount of randomness that is consumed by privacy amplification. Beyond the primary source of randomness, X, some additional randomness, sometimes called *catalyst randomness* or *seed*, is needed for the random variable I, i.e., to sample a function $g_i \in \mathcal{G}$ uniformly at random. Hence, for the present construction the amount of catalyst random bits needed is logarithmic in $|\mathcal{G}|$. Note that the size of the seed is an important parameter, not only because randomness should generally be regarded as a scarce resource, but also since the seed needs to be communicated.

To decrease the required number of random bits (and thereby to decrease the communication complexity of the privacy-amplification protocol), one could make use of a "δ-almost universal" family [Sti94]. For an appropriately chosen construction, see, e.g., the proof of Theorem 10 in [TSSR10], the amount of catalyst randomness becomes linear in the output length r, at the cost of slightly increasing $\mathrm{SD}(KYI; UYI)$.

Strongly Universal Families

Here, we define the related notion of a strongly universal family [CW81], which we will need later in the context of identification.

Definition 2.23 Let $\mathcal{G} := \{g_i\}_{i\in\mathcal{I}}$ be a family of functions $g_i : \mathcal{X} \to \mathcal{R}$, where $\mathcal{I}$, $\mathcal{X}$ and $\mathcal{R}$ are finite and non-empty sets. Let I be a random variable that is uniformly distributed over $\mathcal{I}$. The family $\mathcal{G}$ is called *strongly universal* if for all $x, x' \in \mathcal{X}$ such that $x \neq x'$ and for all $a, b \in \mathcal{R}$ it holds that

$$\Pr[g_I(x) = a \wedge g_I(x') = b] \leq \frac{1}{|\mathcal{R}|^2}.$$

2.4.2 Extractors

A more general approach to studying the (non-interactive) privacy amplification problem is in the language of *randomness extractors*. A main advantage of extractors over universal hashing is that there exist randomness extractors that require much shorter seed lengths.

Definition 2.24 Let $n, d, m \in \mathbb{N}$ and let $k, \varepsilon \in \mathbb{R}$ such that $k \geq 0$ and $\varepsilon \geq 0$. A function $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a (k,ε) *strong extractor* if for any pair of random variables (X,Y) such that X has range $\mathcal{X} \subseteq \{0,1\}^n$ and

$H_{\min}(X|Y) \geq k$, it holds that

$$\mathrm{SD}(\mathsf{Ext}(X;S)YS, UYS) \leq \varepsilon,$$

where S (the seed) is a uniformly distributed random variable over $\{0,1\}^d$ that is independent of XY, and U is a uniformly distributed random variable over $\{0,1\}^m$ that is independent of YS.

Remark 2.25 In Definition 2.24, the word *strong* indicates that $\mathsf{Ext}(X;S)$ must be close to uniform *even when given the seed S*. The definition for an *extractor* (without the adjective *strong*) is similar to Definition 2.24 but merely requires that

$$\mathrm{SD}(\mathsf{Ext}(X;S)Y, UY) \leq \varepsilon.$$

Remark 2.26 In the theoretical-computer-science literature, it is common to define extractors and strong extractors with respect to the *unconditional* min-entropy. The formal definition for strong extractors with respect to unconditional min-entropy is obtained from Definition 2.24 by removing every occurrence of Y. In [DORS08], extractors under this modified definition are called *worst-case* strong extractors, and strong extractors under Definition 2.24 are called *average-case* strong extractors.

The following theorem due to Vadhan guarantees that we can always turn any worst-case strong extractor into an average-case strong extractor, with only a slight loss in parameters.

Theorem 2.27 ([Vad12]) *Suppose that* Ext *is a worst-case (k,ε) strong extractor. Then,* Ext *is also an average-case $(k,3\varepsilon)$ strong extractor.*

A proof can be found in [CP11].

The following theorem bounds the parameters of worst-case extractors.

Theorem 2.28 ([RTS00]) *Suppose that $f : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a worst-case (k,ε) extractor. Then, the following bounds hold for d (the seed length) and m (the output length):*

$$\begin{aligned} d &= \log(n-k) + 2\log(1/\varepsilon) - O(1), \\ m &= d + k - 2\log(1/\varepsilon) + O(1). \end{aligned}$$

Moreover, [RTS00] show via the probabilistic method [AS00] that the bounds from Theorem 2.28 are tight up to constant factors.

Theorem 2.29 (Theorem 1.10 in [RTS00]) *For every $k, n, m \in \mathbb{N}$ such that $k \leq n$ and $\varepsilon \in \mathbb{R}$ such that $\varepsilon > 0$ there exists a worst-case (k, ε) extractor* $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *such that*

$$d = \left\lceil \max\left\{ \log\left(\frac{n-k}{\log e} + 1\right) + 2\log(\tfrac{1}{\varepsilon}), m - k + 2\log(\tfrac{1}{\varepsilon}) - \log(\log e) \right\} \right\rceil .$$

(Recall that e denotes the base of the natural logarithm.)

It is often not sufficient to merely know that extractors exist. Most applications require an *explicit* construction of an extractor. Informally, this means that the extractor can be efficiently constructed, i.e., in time polynomial in n. For a formal definition, we refer to [Sha02].

There exist several explicit constructions for (strong) extractors. In particular, a universal family of hash functions is an instance of an average-case[6] strong extractor and comes with explicit constructions [CW77]. When short seed length is important, universal hashing is not a good choice. An example of a strong extractor with much shorter seed length is the following.

Theorem 2.30 ([GUV09]) *For all $\alpha, \varepsilon \in \mathbb{R}$ such that $\alpha > 0$ and $\varepsilon > 0$ and all positive integers n, k, there is an explicit construction of a worst-case (k, ε) strong extractor* $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *with $d = O(\log n + \log(1/\varepsilon))$ and $m \geq (1-\alpha)k$.*

By Theorem 2.27, the extractor from [GUV09] is also an average-case strong extractor with the same parameters.

In this thesis, we will solely deal with average-case strong extractors.[7] Hence, we will use the following convention.

Convention 2.31 *From here, whenever we use the word* extractor, *we always mean an average-case strong extractor.*

[6] For a short proof of the fact that a universal family of hash functions is an average-case strong extractor, see [DORS08].

[7] Including average-case strong extractors with additional properties, see Definition 2.64.

2.5 Statistically Secure Encryption and Authentication

2.5.1 The One-Time Pad

Theorem 2.32 *Let M, K be independent random variables over $\mathbb{F}_2^n$, where K is uniformly distributed, and let $C := M \oplus K$. Then, for any $m, c \in \mathbb{F}_2^n$ it holds that*

$$\Pr[M = m|C = c] = \Pr[M = m].$$

By viewing M as a message, and K as a secret key, then C can be understood as a *perfect encryption* of M under K: the ciphertext C does not provide any information about the message M, whereas C and K together determine M (by definition of C). This (symmetric) encryption method is called the *one-time pad*. Shannon [Sha49] was the first to formally state and prove a statement equivalent to Theorem 2.32. In Section 2.8 (Proposition 2.53) we will give a more general version of Theorem 2.32.

The one-time pad gets its name from the fact that a key may only be used *once*: if a key is reused, the unconditional security property does no longer hold.

The one-time pad is actually not used in practice to encrypt large messages (except in ultra-high-security settings), because of the equal amount of key material needed. Nonetheless, in this thesis the one-time pad is an important building block.

2.5.2 Message Authentication

Consider the setting where a sender transmits a message to a receiver. The goal of message authentication is to convince the receiver that the received message is identical to the transmitted message, i.e., that it has not been modified by an adversary during transmission. A related problem that can also be solved by message authentication is where the sender has not transmitted any message, but the adversary injects a message into the channel instead.

Definition 2.33 Let $n, t, \ell \in \mathbb{N}$ and let $\delta \in \mathbb{R}$ such that $\delta \geq 0$. A family of functions

$$\{\mathsf{MAC}_k : \{0,1\}^n \to \{0,1\}^t\}$$

indexed by keys $k \in \{0,1\}^\ell$ is a δ-secure *message authentication code* if for any pair of fixed distinct messages, $m, m' \in \{0,1\}^n$ such that $m \neq m'$ and random variable K uniformly distributed over $\{0,1\}^\ell$ it holds that

$$p_{\mathsf{guess}}(\mathsf{MAC}_K(m') \,|\, \mathsf{MAC}_K(m)) \leq \delta.$$

A message authentication code can be used to authenticate a message in the following way. We require that the sender and receiver share a common secret key $K \in \{0,1\}^\ell$. Then, to authenticate a message a message m, the sender computes the *tag* $T := \mathsf{MAC}_K(m) \in \{0,1\}^t$ and transmits it along with the message. The receiver will compute $\mathsf{MAC}_K(m')$ where m' denotes the *received* message, and accepts the message if $\mathsf{MAC}_K(m') = T$.

The security property of the message authentication code guarantees that an adversary, who obviously knows m and T but does not know K, cannot select a message $m' \neq m$ of his choice and at the same time producing a valid tag $\widetilde{T}$ (such that $\widetilde{T} = \mathsf{MAC}_K(m')$, except with probability $\leq \delta$, where δ can be made arbitrarily small.

Carter and Wegman [CW81] showed that an *almost* strongly universal family of functions can be used as a secure message authentication code. An almost strongly universal family is a relaxed version of Definition 2.23, in that the upper bound $1/|\mathcal{R}|^2$ is slightly increased. The main benefit of this relaxed notion is that it allows for smaller families of functions, hence the required authentication-key size decreases, while incurring only a small increase in the security parameter δ.

Similar to the key used in the one-time pad, an authentication key may generally be used only a limited number of times, and, in case of using an almost strongly universal family, just once.

2.6 Hilbert Spaces

Here, we recall some basic facts from functional analysis. For an in-depth introduction, as well as the proofs of the statements that we present as facts, we refer to Kreyszig's book [Kre78].

A complex *inner product space* is a vector space V over the complex numbers that is equipped with an *inner product*. The latter is a map $(\cdot,\cdot) : V \times V \to \mathbb{C}$ such that

1. $(x, y+z) = (x,y) + (x,z)$
2. $(x, \alpha y) = \alpha(x,y)$
3. $(x,y) = \overline{(y,x)}$
4. $(x,x) \geq 0, \quad \text{and} \quad (x,x) = 0 \iff x = 0,$

for all vectors $x, y, z \in V$ and scalars $\alpha \in \mathbb{C}$. Note that the inner product defined here is linear in its second argument (and conjugate-linear in its first argument), as is common in physics and quantum information theory. On the other hand, in mathematical texts it is more common to define the inner product to be linear in its first argument, see, e.g., [Kre78]. Nonetheless, all statements in [Kre78] still hold with respect to our definition of the inner product, since both ways of defining the inner product are the same up to permutation of the arguments.

A complex *Hilbert space* is a complete[8] complex inner product space. In this thesis we will only deal with *finite-dimensional* complex Hilbert spaces, for which completeness is automatically guaranteed; this holds because every finite-dimensional normed space is complete, see, e.g. Theorem 2.4-2 in [Kre78].

Convention 2.34 *When we speak about a "Hilbert space" in this thesis, we always mean a* complex *Hilbert space of* finite dimension.

Convention 2.35 *As we will never need the notion of K-linearity for a (sub-)field K other than $\mathbb{C}$ itself, "linear" will always mean "$\mathbb{C}$-linear" in this section.*

2.6.1 Dirac's Braket Notation

We will use Dirac's "braket" notation, which is common in quantum mechanics. In this notation, a vector in a Hilbert space $\mathcal{H}$ is denoted as

$$|\psi\rangle$$

and is called a *ket* vector.

Let $\mathcal{H}^*$ denote the dual vector space of $\mathcal{H}$. For any vector $|\psi\rangle \in \mathcal{H}$, $\langle\psi| \in \mathcal{H}^*$ is the linear functional:

$$\begin{array}{rcccc} \langle\psi| & : & \mathcal{H} & \to & \mathbb{C}, \\ & & |\varphi\rangle & \mapsto & \langle\psi|\varphi\rangle, \end{array}$$

where $\langle\psi|$ is called a *bra* vector, and $\langle\psi|\varphi\rangle$ is the inner product $(|\psi\rangle, |\varphi\rangle)$ written in braket notation (which we will mainly use from now). Hence, by definition it holds that $\langle\psi||\varphi\rangle = \langle\psi|\varphi\rangle$.

Furthermore, we define the *outer product* $|\varphi\rangle\langle\psi|$ for any $|\psi\rangle, |\varphi\rangle \in \mathcal{H}$ as the linear operator

$$\begin{array}{rcccc} |\varphi\rangle\langle\psi| & : & \mathcal{H} & \to & \mathcal{H}, \\ & & |\chi\rangle & \mapsto & \langle\psi|\chi\rangle|\varphi\rangle. \end{array}$$

Hence, it holds that $|\varphi\rangle\langle\psi||\chi\rangle = |\varphi\rangle\langle\psi|\chi\rangle$.

[8] complete in the metric that is naturally induced by the inner product.

2.6.2 Operators

For arbitrary Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$, let $\mathrm{Hom}(\mathcal{H}, \mathcal{H}')$ denote the complex vector space of linear maps $\mathcal{H} \to \mathcal{H}'$ and let $\mathrm{End}(\mathcal{H})$ denote the complex algebra of linear operators $\mathcal{H} \to \mathcal{H}$. The elements in $\mathrm{End}(\mathcal{H})$ are also called endomorphisms.[9] From here, whenever we write *operator* in this thesis, we always mean *linear* operator.

The *identity operator*, which we denote by $\mathbb{I}_{\mathcal{H}}$, is the unique operator on $\mathcal{H}$ such that for any $|\varphi\rangle \in \mathcal{H}$, it holds that $\mathbb{I}_{\mathcal{H}}|\varphi\rangle = |\varphi\rangle$. Note that we omit the subscript of $\mathbb{I}_{\mathcal{H}}$ if the Hilbert space on which the identity operator acts is clear from its context. And, for a Hilbert space named $\mathcal{H}_A$, we will often denote the identity operator as $\mathbb{I}_A$.

The *adjoint* of an operator $T \in \mathrm{End}(\mathcal{H})$ is the unique operator

$$T^\dagger : \mathcal{H} \to \mathcal{H}$$

such that

$$(|\varphi\rangle, T|\psi\rangle) = (T^\dagger|\varphi\rangle, |\psi\rangle)$$

holds for all $|\varphi\rangle, |\psi\rangle \in \mathcal{H}$.

The bra of the vector $T|\varphi\rangle$ coincides with $\langle\varphi|T^\dagger$, as can easily be verified. As a consequence, the expression $\langle\varphi|T|\psi\rangle$, which can be interpreted as $\langle\varphi|T \in \mathcal{H}^*$ applied to $|\psi\rangle \in \mathcal{H}$ as well as $\langle\varphi| \in \mathcal{H}^*$ applied to $T|\psi\rangle \in \mathcal{H}$, causes no confusion since both interpretations give rise to the same value by definition of the adjoint.

We recall some special classes of operators. An operator $T \in \mathrm{End}(\mathcal{H})$ is

1. *normal* if $T^\dagger T = TT^\dagger$;
2. *unitary* if $T^\dagger T = TT^\dagger = \mathbb{I}$ (implies that T is normal);
3. *Hermitian* if $T^\dagger = T$ (implies that T is normal);
4. *positive semi-definite* if $\langle\psi|T|\psi\rangle \geq 0$ for all $|\psi\rangle \in \mathcal{H}$ (implies that T is Hermitian). We also write $T \geq 0$;
5. an *orthogonal projector* if $TT = T$ and $T^\dagger = T$ (implies that $T \geq 0$).

[9] In [Kre78], most statements about operators on normed spaces also cover the infinite-dimensional case, and as a consequence many statements require an operator to be *bounded*. Let $\mathcal{H}$ be a Hilbert space (recall that we always mean a *finite-dimensional* complex Hilbert space). A basic fact is that every operator $T \in \mathrm{End}(\mathcal{H})$ is bounded (see also Theorem 2.7-8 in [Kre78]), in that there exists a real number c such that for every $|x\rangle \in \mathcal{H}$ it holds that $\||T|x\rangle\|| \leq c\||x\rangle\||$, where $\|\cdot\|$ is the norm induced by the inner product.

Whenever we speak of a projector, we always mean an orthogonal projector. For all $|\psi\rangle \in \mathcal{H}$ that have norm 1, the operator $|\psi\rangle\langle\psi|$ is the orthogonal projector on the one-dimensional subspace spanned by $|\psi\rangle$ and has rank 1.

Every normal operator has a *spectral decomposition*.

Theorem 2.36 *Let $\mathcal{H}$ be an arbitrary Hilbert space and let d denote its dimension. For every normal operator $T \in \mathrm{End}(\mathcal{H})$ there exists an orthonormal basis of $\mathcal{H}$, $\{|\psi_i\rangle\}_{i\in[d]}$, such that*

$$T = \sum_{i\in[d]} \lambda_i |\psi_i\rangle\langle\psi_i|,$$

where the $\lambda_i \in \mathbb{C}$ for $i \in [d]$ are the eigenvalues of T. If T is Hermitian, then all eigenvalues are real.

In general, the spectral decomposition is not unique. In fact, it is unique (and the eigenvectors are unique up to multiplication by a complex scalar with norm 1) if and only if all eigenvalues are distinct. Note that if T is positive semi-definite, all eigenvalues are real and non-negative.

2.6.3 Tensor Products

To be able to compose quantum systems (see next section), we need the notion of a tensor product.

Definition 2.37 For finite-dimensional complex vector spaces V and W, a *tensor product* of V and W is a pair (U, ι), such that

- U is a finite-dimensional complex vector space and $\iota : V \times W \to U$ is a $\mathbb{C}$-bilinear map.
- For each pair (Z, f) where Z is a finite-dimensional complex vector space and $f : V \times W \to Z$ is a $\mathbb{C}$-bilinear map, there is a *unique* $\mathbb{C}$-linear map $f_* : U \to Z$ such that $f = f_* \circ \iota$.[10]

The latter property is called the *universal property* of a tensor product. Hence, the map f_*, which uniquely corresponds to f and is induced by the universal property,

[10] The "$\circ$" symbol denotes composition of maps.

makes the following diagram commutative:

$$\begin{tikzcd}
V \times W \arrow[r, "\iota"] \arrow[dr, "f"'] & U \arrow[d, "f_*"] \\
 & Z
\end{tikzcd}$$

Remark 2.38 The notion of a tensor product can be defined more generally (e.g., over an arbitrary field, or for infinite-dimensional spaces) in the same way as in Definition 2.37. However, as we do not need this generality here, we only define tensor products for finite-dimensional complex vector spaces.

In the remainder of this section, every occurrence of "linear combination" or "linear extension" should be read as a $\mathbb{C}$-linear combination or extension.

We will now show that a tensor product exists, by giving a natural construction. From the universal property, it follows that the tensor product is unique, up to a unique isomorphism. The particular tensor product that we construct below will be defined as *the* tensor product of V and W.

For finite-dimensional complex vector spaces V and W with respective dimensions d_V and d_W, let $V \otimes W$ denote the quotient space $F(V \times W)/R$, where $F(V \times W)$ is the free $\mathbb{C}$-vector space on $V \times W$, i.e., the vector space obtained by taking formal finite linear combinations of elements $(v, w) \in V \times W$, and R is the subspace of $F(V \times W)$ spanned by all elements of the form

$$\begin{aligned}
&(v_1 + v_2, w) - (v_1, w) - (v_2, w)\\
&(v, w_1 + w_2) - (v, w_1) - (v, w_2)\\
&\alpha(v, w) - (\alpha v, w)\\
&\alpha(v, w) - (v, \alpha w)
\end{aligned}$$

with $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$, and $\alpha \in \mathbb{C}$. We will call $V \otimes W$ the *tensor-product space* of V and W. The dimension of $V \otimes W$ is equal to $d_V \cdot d_W$.

For $v \in V$ and $w \in W$, the residue class $(v, w) + R$ in $V \otimes W$ is denoted by $v \otimes w$. The following properties hold by construction:

1. $(v_1 + v_2) \otimes w = (v_1 \otimes w) + (v_2 \otimes w)$
2. $v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2$
3. $\alpha(v \otimes w) = (\alpha v) \otimes w = v \otimes (\alpha w)$

for all $v, v_1, v_2 \in V$, for all $w, w_1, w_2 \in W$, and for all $\alpha \in \mathbb{C}$.

Define

$$\begin{aligned} \varphi : V \times W &\to V \otimes W \\ (v, w) &\mapsto v \otimes w. \end{aligned}$$

Definition 2.39 $(V \otimes W, \varphi)$ is *the* tensor product of V and W.

We will make use of the natural isomorphism

$$\phi : \mathrm{Hom}(V, V') \otimes \mathrm{Hom}(W, W') \to \mathrm{Hom}(V \otimes W, V' \otimes W') \tag{2.2}$$

for all finite-dimensional complex vector spaces V, V', W, W', given by the linear extension of

$$\phi(A \otimes B)(v \otimes w) = Av \otimes Bw$$

for all $A \in \mathrm{Hom}(V, V')$, for all $B \in \mathrm{Hom}(W, W')$, for all $v \in V$ and for all $w \in W$.

The tensor product $\mathcal{H} \otimes \mathcal{H}'$ of two *Hilbert spaces* $\mathcal{H}$ and $\mathcal{H}'$ is a finite-dimensional complex vector space, and we can turn $\mathcal{H} \otimes \mathcal{H}'$ into a Hilbert space by equipping it with an inner product.[11] A natural choice for this inner product is given by

$$(|\varphi\rangle \otimes |\psi\rangle, |\chi\rangle \otimes |\omega\rangle) := \langle\varphi|\chi\rangle \cdot \langle\psi|\omega\rangle, \tag{2.3}$$

for all $|\varphi\rangle, |\chi\rangle \in \mathcal{H}$ and for all $|\psi\rangle, |\omega\rangle \in \mathcal{H}'$. By linearity, this extends to all elements of $\mathcal{H} \otimes \mathcal{H}'$, i.e., elements of the form $\sum_i \alpha_i |\varphi_i\rangle \otimes |\psi_i\rangle$ for $\alpha_i \in \mathbb{C}$, $|\varphi_i\rangle \in \mathcal{H}$ and $|\psi_i\rangle \in \mathcal{H}'$ for all i. This inner product has the property that if $\{|\varphi_i\rangle\}_i$ is an orthonormal basis for $\mathcal{H}$ and $\{|\psi_j\rangle\}_j$ is an orthonormal basis for $\mathcal{H}'$, then $\{|\varphi_i\rangle \otimes |\psi_j\rangle\}_{i,j}$ is an orthonormal basis for $\mathcal{H} \otimes \mathcal{H}'$.

We will sometimes omit the tensor-product symbol in tensor products between Hilbert-space elements, i.e., $|\varphi\rangle|\psi\rangle$ for $|\varphi\rangle \in \mathcal{H}$ and $|\psi\rangle \in \mathcal{H}'$ should be read as the tensor product $|\varphi\rangle \otimes |\psi\rangle$.

By the natural isomorphism (2.2) and by the natural isomorphism $\mathbb{C} \otimes \mathbb{C} \cong \mathbb{C}$ (which follows from a simple argument involving the universal property) it follows that $\mathcal{H}^* \otimes \mathcal{H}'^* \cong (\mathcal{H} \otimes \mathcal{H}')^*$ and that the bra vector of an element $|\varphi\rangle \otimes |\psi\rangle \in \mathcal{H} \otimes \mathcal{H}'$ is naturally identified with

$$\langle\varphi| \otimes \langle\psi| \in \mathcal{H}^* \otimes \mathcal{H}'^*.$$

[11] Recall that completeness is guaranteed since we are in the finite-dimensional case.

Indeed, the bra vector $\langle\varphi|\otimes\langle\psi|$ acts on a ket vector $|\chi\rangle\otimes|\omega\rangle\in\mathcal{H}\otimes\mathcal{H}'$ as

$$(\langle\varphi|\otimes\langle\psi|)(|\chi\rangle\otimes|\omega\rangle)=\langle\varphi||\chi\rangle\otimes\langle\psi||\omega\rangle=\langle\varphi|\chi\rangle\langle\psi|\omega\rangle.$$

Similarly, it follows from (2.2) that $\mathrm{End}(\mathcal{H})\otimes\mathrm{End}(\mathcal{H}')\cong\mathrm{End}(\mathcal{H}\otimes\mathcal{H}')$, and that for any $A\in\mathrm{End}(\mathcal{H})$ and $B\in\mathrm{End}(\mathcal{H}')$, $A\otimes B$ acts on elements from $\mathcal{H}\otimes\mathcal{H}'$ as

$$(A\otimes B)(|\varphi\rangle\otimes|\psi\rangle)=A|\varphi\rangle\otimes B|\psi\rangle$$

for any $|\varphi\rangle\in\mathcal{H}$ and for any $|\psi\rangle\in\mathcal{H}'$. By linearity, this extends to all elements of $\mathcal{H}\otimes\mathcal{H}'$.

Finally, note that the natural isomorphism $\mathrm{End}(\mathcal{H}\otimes\mathcal{H}')\cong\mathrm{End}(\mathcal{H})\otimes\mathrm{End}(\mathcal{H}')$ identifies the outer product $(|\chi\rangle\otimes|\omega\rangle)(\langle\varphi|\otimes\langle\psi|)\in\mathrm{End}(\mathcal{H}\otimes\mathcal{H}')$ with $|\chi\rangle\langle\varphi|\otimes|\omega\rangle\langle\psi|\in\mathrm{End}(\mathcal{H})\otimes\mathrm{End}(\mathcal{H}')$.

2.6.4 Vector and Matrix Representations

For every $d\in\mathbb{N}$, let the complex vector space $\mathbb{C}^d$ be equipped with the standard inner product

$$\langle x|y\rangle=\bar{x}_1y_1+\cdots+\bar{x}_dy_d$$

for any $|x\rangle=[x_1\ \cdots\ x_d]^\mathsf{T}\in\mathbb{C}^d$ and $|y\rangle=[y_1\ \cdots\ y_d]^\mathsf{T}\in\mathbb{C}^d$. From here, we will view $\mathbb{C}^d$ as a Hilbert space.

For the sake of clarity of this section, we want to be able to indicate whether elements from $\mathbb{C}^d$ should be understood as row vectors or column vectors. Hence, we will write $\mathbb{C}^{d\times 1}$ when the elements of $\mathbb{C}^d$ should be understood as *column vectors*, and likewise $\mathbb{C}^{1\times d}$ when the elements of $\mathbb{C}^d$ should be understood as *row vectors*.

Let $\mathcal{H}$ be a Hilbert space of dimension d. Every choice of an orthonormal basis $\{|i\rangle\}_i$ of $\mathcal{H}$ induces the following vector-space isomorphisms

$$\begin{aligned}
&\mathcal{H}\to\mathbb{C}^{d\times 1} && \mathcal{H}^*\to\mathbb{C}^{1\times d} && \mathrm{End}(\mathcal{H})\to\mathbb{C}^{d\times d}\\
&|\varphi\rangle\mapsto\begin{bmatrix}\varphi_1\\ \vdots\\ \varphi_d\end{bmatrix} && \langle\varphi|\mapsto\begin{bmatrix}\bar{\varphi}_1 & \cdots & \bar{\varphi}_d\end{bmatrix} && T\mapsto\begin{bmatrix}t_{11} & \cdots & t_{1d}\\ \vdots & \ddots & \vdots\\ t_{d1} & \cdots & t_{dd}\end{bmatrix}
\end{aligned}\tag{2.4}$$

with $\varphi_i:=\langle i|\varphi\rangle$ for all $i\in[d]$ and with $t_{ij}:=\langle i|T|j\rangle$ for all $i,j\in[d]$. The matrix $[t_{ij}]_{i,j\in[d]}$ is the *matrix representation* of the operator T in the basis $\{|i\rangle\}_i$. Note that the leftmost isomorphism is a *Hilbert-space isomorphism*.[12]

[12] A Hilbert-space isomorphism is a vector-space isomorphism that preserves the inner product, i.e., for a Hilbert-space isomorphism $\tau:\mathcal{H}\to\mathcal{H}'$ it holds that $\langle y|\tau^\dagger\tau|x\rangle=\langle y|x\rangle$ for all $|x\rangle\in\mathcal{H}$ and all $|y\rangle\in\mathcal{H}'$.

Under these isomorphisms, the bra vector of a ket vector is given by the conjugate transpose. Furthermore, the respective actions of $\langle\psi| \in \mathcal{H}^*$ and $T \in \mathrm{End}(\mathcal{H})$ on a ket vector $|\varphi\rangle \in \mathcal{H}$ correspond to standard matrix multiplication, i.e., both diagrams below commute:

$$\begin{tikzcd}
\mathcal{H}^* \times \mathcal{H} \arrow[r, "\sim"] \arrow[rd] & \mathbb{C}^{1\times d} \times \mathbb{C}^{d\times 1} \arrow[d] \\
 & \mathbb{C}
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathrm{End}(\mathcal{H}) \times \mathcal{H} \arrow[r, "\sim"] \arrow[d] & \mathbb{C}^{d\times d} \times \mathbb{C}^{d\times 1} \arrow[d] \\
\mathcal{H} \arrow[r, "\sim"] & \mathbb{C}^{d\times 1}
\end{tikzcd}$$

where the $\xrightarrow{\sim}$-arrows are the vector-space isomorphisms given by (2.4), and the down arrows are given by the actions of $\mathcal{H}^*$ and $\mathrm{End}(\mathcal{H})$ on $\mathcal{H}$, and by standard matrix multiplication, respectively. Additionally, the outer product $|\varphi\rangle\langle\psi|$ can be understood as the matrix product between $|\varphi\rangle$ and $\langle\psi|$.

Let us recall the *Kronecker product*. For matrices

$$A = \begin{bmatrix} a_{11} & \dots & a_{1\ell} \\ \vdots & \ddots & \vdots \\ a_{k1} & \dots & a_{k\ell} \end{bmatrix} \in \mathbb{C}^{k\times\ell} \qquad \text{and} \qquad B = \begin{bmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \dots & b_{mn} \end{bmatrix} \in \mathbb{C}^{m\times n}$$

the Kronecker product $A \otimes B$ is the matrix given by

$$A \otimes B = \begin{bmatrix} a_{11}B & \dots & a_{1\ell}B \\ \vdots & \ddots & \vdots \\ a_{k1}B & \dots & a_{k\ell}B \end{bmatrix} \in \mathbb{C}^{km\times\ell n}$$

Note that the Kronecker product between two vectors follows as a special case.

Let $\mathcal{H}'$ be a Hilbert space of dimension d'. Under the isomorphisms (2.4) and under similar isomorphisms induced when fixing a basis for $\mathcal{H}'$, the tensor product can be identified with the Kronecker product, in the sense that each of the following diagrams commutes.

$$\begin{tikzcd}
\mathcal{H} \times \mathcal{H}' \arrow[r, "\sim"] \arrow[d, "\otimes"'] & \mathbb{C}^{d\times 1} \times \mathbb{C}^{d'\times 1} \arrow[d, "\otimes\ \text{Kronecker product}"] \\
\mathcal{H} \otimes \mathcal{H}' \arrow[r, "\sim^\star"] & \mathbb{C}^{dd'\times 1}
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathcal{H}^* \times \mathcal{H}'^* \arrow[r, "\sim"] \arrow[d, "\otimes"'] & \mathbb{C}^{1\times d} \times \mathbb{C}^{1\times d'} \arrow[dd, "\otimes\ \text{Kronecker product}"] \\
\mathcal{H}^* \otimes \mathcal{H}'^* \arrow[d, "\wr"', "\text{isomorph. (2.2)}"] & \\
(\mathcal{H} \otimes \mathcal{H}')^* \arrow[r, "\sim^\star"] & \mathbb{C}^{1\times dd'}
\end{tikzcd}$$

$$\begin{tikzcd}
\mathrm{End}(\mathcal{H}) \times \mathrm{End}(\mathcal{H}') \arrow[r, "\sim"] \arrow[d, "\otimes"'] & \mathbb{C}^{d\times d} \times \mathbb{C}^{d'\times d'} \arrow[dd, "\otimes"', "\substack{\text{Kronecker}\\ \text{product}}"] \\
\mathrm{End}(\mathcal{H}) \otimes \mathrm{End}(\mathcal{H}') \arrow[d, "\wr"', "\text{isomorph. (2.2)}"] & \\
\mathrm{End}(\mathcal{H} \otimes \mathcal{H}') \arrow[r, "\sim^{\star}"] & \mathbb{C}^{dd'\times dd'}
\end{tikzcd}$$

In the diagrams above, the isomorphisms marked with a star ($\star$) are induced by fixing orthonormal bases for $\mathcal{H}$ and $\mathcal{H}'$. To see this, note that by the natural choice of the inner product (2.3), choosing orthonormal bases for $\mathcal{H}$ and $\mathcal{H}'$ immediately fixes an orthonormal basis for $\mathcal{H} \otimes \mathcal{H}'$. The latter basis, in turn, induces the starred isomorphisms.

Whenever we fix a Hilbert space $\mathcal{H}$, whose dimension will be denoted as d, the isomorphisms (2.4) and the commuting diagrams from this section allow us to implicitly assume without loss of generality that $\mathcal{H} = \mathbb{C}^{d\times 1}$, that $\mathcal{H}^* = \mathbb{C}^{1\times d}$, and that $\mathrm{End}(\mathcal{H}) = \mathbb{C}^{d\times d}$.

For a more in-depth treatment of the tensor product, see [Lan05].

2.7 Quantum Systems and Operations

2.7.1 Postulates of Quantum Mechanics

Quantum mechanics is a mathematical model of a physical system. The theory of quantum mechanics is based on four postulates, from which the rest of the theory can be derived.

We will state the set of postulates of quantum mechanics in the language of density matrices.

Definition 2.40 A *density matrix* $\rho \in \mathrm{End}(\mathcal{H})$ on a Hilbert space $\mathcal{H}$ is a positive-semidefinite matrix having trace equal to 1, i.e., $\rho \geq 0$ and $\mathrm{tr}(\rho) = 1$. The set of density matrices on $\mathcal{H}$ is denoted as $\mathcal{D}(\mathcal{H})$.[13] If a density matrix has rank equal to 1 it is said to be *pure*, and in this case it can be written as $\rho = |\varphi\rangle\langle\varphi|$, where $|\varphi\rangle \in \mathcal{H}$ has norm 1.

Since positive-semidefiniteness implies normality, any density matrix has a spectral decomposition. Since by definition the trace of any density matrix equals 1, the eigenvalues of a density matrix always sum to one.

[13] It is straightforward to verify that this set is convex.

We will now state the four postulates. We basically follow [NC00], but we have rephrased the postulates in our language and notation.

1. *State:* To every quantum system we can associate a Hilbert space $\mathcal{H}$ (possibly *infinite*-dimensional), the *state space*, such that there is a one-to-one correspondence between the set of all possible states of this quantum system and $\mathcal{D}(\mathcal{H})$.

In this thesis (as is common in quantum information theory) we restrict to *finite-dimensional* quantum systems, i.e., systems with a finite-dimensional state space.

2. *Composition:* Let A and B be two finite-dimensional quantum systems, with state spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively. The state space of the *composite system AB* (also called the *joint system*) is given by the tensor product of the individual state spaces, i.e., $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$.

We say that two quantum systems A and B are *independent*, if the density matrix ρ_{AB} of the joint system AB can be decomposed as $\rho_{AB} = \rho_A \otimes \rho_B$.

3. *Evolution:* Any evolution of a finite-dimensional quantum system over a time-interval $[t_0, t_1]$ is described by a unitary transformation on the state space of that system that, with respect to time-dependence, solely depends on the boundary points t_0 and t_1. For a system with state space $\mathcal{H}$ and unitary $U \in \mathrm{End}(\mathcal{H})$, let $\rho \in \mathcal{D}(\mathcal{H})$ be the state at time t_0. The state at time t_1, ρ', is given by
$$\rho' = U\rho U^\dagger \in \mathcal{D}(\mathcal{H}).$$

4. *Measurement:* Any measurement on a finite-dimensional quantum system with state space $\mathcal{H}$ can be described by a collection $\{M_x\}_{x\in\mathcal{X}}$ of operators $M_x \in \mathrm{End}(\mathcal{H})$ that satisfy the *completeness condition*: $\sum_{x\in\mathcal{X}} M_x^\dagger M_x = \mathbb{I}$, for some finite and non-empty set $\mathcal{X}$. The operators M_x are called *measurement operators*. The index x refers to the possible outcomes of the measurement. When applying a measurement $\{M_x\}_{x\in\mathcal{X}}$ to a system with state space $\mathcal{H}$ that is in state ρ, the probability of obtaining outcome $x \in \mathcal{X}$ is given by
$$P_X(x) = \mathrm{tr}(M_x^\dagger M_x \rho) \quad \text{for all } x \in \mathcal{X}, \tag{2.5}$$
and the state of the system conditioned on having obtained outcome $x \in \mathcal{X}$ is
$$\rho^x = \frac{1}{P_X(x)} M_x \rho M_x^\dagger \in \mathcal{D}(\mathcal{H}) \quad \text{for all } x \in \mathcal{X}. \tag{2.6}$$

When we say that we apply a unitary $U_A \in \mathrm{End}(\mathcal{H}_A)$ to system A (having state space $\mathcal{H}_A$) of some joint quantum system AB (with state space $\mathcal{H}_A \otimes \mathcal{H}_B$), we mean applying the unitary $U_{AB} := U_A \otimes \mathbb{I}_B$ to the joint system. Similarly, when we say that we *measure* system A using a (complete) collection of measurement operators $\{M_{A,x}\}_{x\in\mathcal{X}}$ where $\mathcal{X}$ is a finite and non-empty set and where $M_{A,x} \in \mathrm{End}(\mathcal{H}_A)$ for every $x \in \mathcal{X}$, we mean that we measure the joint system using the collection of measurement operators $\{M_{AB,x}\}_{x\in\mathcal{X}}$, where $M_{AB,x} = M_{A,x} \otimes \mathbb{I}_B$ for every $x \in \mathcal{X}$.

2.7.2 States

In general, a state that is described by a density matrix is called a *mixture* (another way of saying this is to say that the state is *mixed*). In case this density matrix is pure, then we will also call the corresponding state (described by that density matrix) *pure*. Any pure state $\rho = |\varphi\rangle\langle\varphi| \in \mathcal{D}(\mathcal{H})$ may equivalently be described by means of the ket vector $|\varphi\rangle \in \mathcal{H}$ (which has norm 1). *Vice versa*, any ket vector $|\varphi\rangle \in \mathcal{H}$ with norm 1 describes a pure state with density matrix $\rho = |\varphi\rangle\langle\varphi| \in \mathcal{D}(\mathcal{H})$. We will also call a ket vector with norm 1 a *state vector*. When dealing with pure states, we sometimes find it more convenient to work with state vectors than with density matrices.

As a disclaimer, we are sometimes a bit sloppy with the terminology and use the term "state" not only for the state of a quantum system, but also for the quantum system itself, as well as for the density matrix representing that state (or for the state vector, in case the state is pure). This should, however, cause no confusion.

An important density matrix in $\mathcal{D}(\mathcal{H})$ is the *fully mixed state* $\frac{1}{d}\mathbb{I}$ with $d = \dim(\mathcal{H})$.

Let $\mathcal{H}$ be of arbitrary dimension d, and let $\{|i\rangle\}_{i\in[d]}$ be an orthonormal basis of $\mathcal{H}$. A *superposition* $|\psi\rangle \in \mathcal{H}$ is a pure state of the form

$$|\psi\rangle = \sum_{i\in[d]} \alpha_i |i\rangle$$

with $\alpha_i \in \mathbb{C}$ for every $i \in [d]$ and such that $\sum_{i\in[d]} |\alpha_i|^2 = 1$. The coefficients α_i are called *amplitudes*.

Let $|\psi\rangle \in \mathcal{H}$ be a state vector. The state vector $e^{\mathrm{i}\theta}|\psi\rangle$ for any $\theta \in \mathbb{R}$ represents the *same* state as $|\psi\rangle$. Indeed, the corresponding density matrices coincide:

$$(e^{\mathrm{i}\theta}|\psi\rangle)(e^{-\mathrm{i}\theta}\langle\psi|) = |\psi\rangle\langle\psi|.$$

2.7.3 Computational and Hadamard Basis

In the context of quantum information theory, the standard basis of $\mathcal{H} = \mathbb{C}^d$ is called the *computational basis*. In $\mathbb{C}^2$, we denote the basis vectors of the computational basis as

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

A state on $\mathcal{H} = \mathbb{C}^2$ is called a *qubit*.

When we write $|b\rangle$ for $b = (b_1, \cdots, b_n) \in \{0,1\}^n$, we mean the tensor product $|b_1\rangle \otimes \cdots \otimes |b_n\rangle \in \mathbb{C}^{2n}$, which is a member of the computational basis of $\mathbb{C}^{2n}$.

The *Hadamard matrix* on $\mathbb{C}^2$ is defined as follows:

$$H := \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \in \mathrm{End}(\mathbb{C}^2).$$

The Hadamard matrix is a *unitary* matrix. For $\mathbb{C}^d$ where $d = 2^n$, the Hadamard matrix on $\mathbb{C}^d$ is defined as the n-fold tensor product of H, i.e., $H^{\otimes n}$, which is a shorthand for $\underbrace{H \otimes \cdots \otimes H}_{n \text{ times}}$.

By applying the Hadamard matrix (of appropriate dimension) to the basis vectors of the computational basis, we obtain the *Hadamard basis* (also called *diagonal basis*), which is often used in this thesis. On $\mathbb{C}^2$, we define

$$|+\rangle := H|0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \quad |-\rangle := H|1\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

2.7.4 Partial Trace and Purification

Consider a composite system AB with state space $\mathcal{H}_A \otimes \mathcal{H}_B$. When given the density matrix ρ_{AB} for this system, we can obtain the density matrix of system A alone, also called the *reduced density matrix* ρ_A, by applying the *partial trace*, i.e., $\rho_A = \mathrm{tr}_B(\rho_{AB})$. We also say that we obtain ρ_A from ρ_{AB} by "tracing out" system B.

Definition 2.41 The partial trace $\mathrm{tr}_B : \mathrm{End}(\mathcal{H}_A \otimes \mathcal{H}_B) \to \mathrm{End}(\mathcal{H}_A)$ is the unique linear functional that satisfies

$$\mathrm{tr}_B(|\varphi_A\rangle|\varphi_B\rangle\langle\psi_A|\langle\psi_B|) = \mathrm{tr}_B(|\varphi_A\rangle\langle\psi_A| \otimes |\varphi_B\rangle\langle\psi_B|) := \langle\psi_B|\varphi_B\rangle|\varphi_A\rangle\langle\psi_A|$$

for any $|\varphi_A\rangle, |\psi_A\rangle \in \mathcal{H}_A$ and $|\varphi_B\rangle, |\psi_B\rangle \in \mathcal{H}_B$.

Definition 2.42 Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be Hilbert spaces, and let $\rho_A \in \mathcal{D}(\mathcal{H}_A)$. A pure state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ is a *purification* of ρ_A if

$$\rho_A = \operatorname{tr}_B(|\psi\rangle\langle\psi|).$$

Proposition 2.43 *Every mixed state $\rho_A \in \mathcal{D}(\mathcal{H}_A)$ has a purification $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ with $\dim(\mathcal{H}_B) \leq \dim(\mathcal{H}_A)$.*

The proof is by construction.

Proof. Let $\rho_A = \sum_i \lambda_i |a_i\rangle\langle a_i|$ be a spectral decomposition of ρ_A. Let $\mathcal{H}_B$ be a Hilbert space with $\dim(\mathcal{H}_B) = \dim(\mathcal{H}_A)$ and let $\{|b_i\rangle\}_i$ be an orthonormal basis for $\mathcal{H}_B$. We claim that

$$|\psi\rangle := \sum_i \sqrt{\lambda_i}|a_i\rangle|b_i\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B.$$

is a purification of ρ_A. To show this, we trace out B again:

$$\begin{aligned}\operatorname{tr}_B(|\psi\rangle\langle\psi|) &= \sum_i \sqrt{\lambda_i}|a_i\rangle|b_i\rangle \sum_j \sqrt{\lambda_j}\langle a_j|\langle b_j| = \sum_{ij} \sqrt{\lambda_i\lambda_j}|a_i\rangle|b_i\rangle\langle a_j|\langle b_j| \\ &= \sum_{ij} \sqrt{\lambda_i\lambda_j}\delta_{ij}|a_i\rangle\langle a_j| = \rho_A,\end{aligned}$$

thus $|\psi\rangle$ is indeed a purification of ρ_A. □

Note that the purification $|\psi\rangle$ is unique up to unitary equivalence: applying an arbitrary unitary $U \in \operatorname{End}(\mathcal{H}_B)$ to system B cannot change the state ρ_A obtained by tracing out B.

2.7.5 The Schmidt Decomposition

Every pure state of a *bipartite*[14] system can be decomposed using the *Schmidt decomposition.*

Proposition 2.44 *Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be arbitrary Hilbert spaces, and write d_A and d_B for their respective dimensions. Let $|\psi\rangle$ be an arbitrary pure state on $\mathcal{H}_A \otimes \mathcal{H}_B$. Then, there exist orthonormal bases for $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively $\{|a_i\rangle\}_{i\in[d_A]}$ and $\{|b_j\rangle\}_{j\in[d_B]}$, an integer $r \in \mathbb{N}$ such that $r \leq \min(d_A, d_B)$ and positive real numbers $\sigma_1, \ldots, \sigma_r$ satisfying $\sum_{i\in[r]} \sigma_i^2 = 1$ such that*

$$|\psi\rangle = \sum_{i\in[r]} \sigma_i |a_i\rangle|b_i\rangle.$$

[14] composed of *two* subsystems.

A consequence of Proposition 2.44 is the following. For any state $|\psi\rangle \in \mathcal{H}_{AB}$, the reduced density matrices ρ_A and ρ_B have exactly the same non-zero eigenvalues:

$$\rho_A = \mathrm{tr}_B(|\psi\rangle\langle\psi|) = \sum_{i\in[r]} \sigma_i^2 |a_i\rangle\langle a_i| \quad \text{and} \quad \rho_B = \mathrm{tr}_A(|\psi\rangle\langle\psi|) = \sum_{i\in[r]} \sigma_i^2 |b_i\rangle\langle b_i|.$$

2.7.6 Quantum Operations

The postulates of quantum mechanics give us all the possible actions that can be performed on a quantum state. We can also take an alternative viewpoint and study maps between operators (sometimes called superoperators).

Let $\mathcal{H}, \mathcal{H}'$ be Hilbert spaces. A map $\mathcal{E} : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{H}')$ is called *positive* if $\mathcal{E}(P) \geq 0$ for any $P \in \mathrm{End}(\mathcal{H})$ for which $P \geq 0$.

For a Hilbert space $\mathcal{H}_A$, let $\mathcal{I}_A : \mathrm{End}(\mathcal{H}_A) \to \mathrm{End}(\mathcal{H}_A)$ denote the identity superoperator.

Definition 2.45 A *completely positive trace-preserving map* (CPTP map) is a map $\mathcal{E} : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{H}')$ with the following properties:

1. *Complete positivity:* the map $\mathcal{E} \otimes \mathcal{I}_R$ is positive for any Hilbert space $\mathcal{H}_R$;
2. *Trace-preserving:* $\mathrm{tr}(\mathcal{E}(T)) = \mathrm{tr}(T)$ for any $T \in \mathrm{End}(\mathcal{H})$

By *Stinespring's dilation theorem*, any CPTP map can be represented as a unitary transformation on a larger system, followed by a partial trace. The unitary is applied to a composition of the input state and an auxiliary, fixed state (usually called *ancilla state*) of appropriate dimension.

Theorem 2.46 (Stinespring Dilation, see, e.g., [Pau02]) *Let $\mathcal{E} : \mathrm{End}(\mathcal{H}_A) \to \mathrm{End}(\mathcal{H}_B)$ be a CPTP map. Then, there exists a Hilbert space $\mathcal{H}_R$ with $\dim(\mathcal{H}_R) = \dim(\mathcal{H}_A)$ and a unitary $U \in \mathrm{End}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_R)$ such that for any $T \in \mathrm{End}(\mathcal{H}_A)$*

$$\mathcal{E}(T) = \mathrm{tr}_{AR}(U(T \otimes |\omega_\circ\rangle\langle\omega_\circ|)U^\dagger),$$

where $|\omega_\circ\rangle = |0\ldots0\rangle \in \mathcal{H}_B \otimes \mathcal{H}_R$.

Hence, any CPTP map can be constructed by combining actions that are provided by the postulates, and, *vice versa*, any possible combination of those actions corresponds to a particular CPTP map.

Furthermore, the Stinespring dilation theorem implies that a CPTP map is the most general physically realizable transformation between density matrices. We will also call a CPTP map a *quantum operation*.

No Cloning

The *no-cloning theorem* is a fundamental result in quantum information theory. It formally expresses the impossibility of copying an unknown quantum state.

Theorem 2.47 *Let $\mathcal{H}$ be an arbitrary Hilbert space. There is no quantum operation $\mathcal{E} : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{H} \otimes \mathcal{H})$ such that*

$$\mathcal{E}(|\varphi\rangle\langle\varphi|) = |\varphi\rangle|\varphi\rangle\langle\varphi|\langle\varphi| \tag{2.7}$$

holds for every $|\varphi\rangle \in \mathcal{H}$.

A state $|\varphi\rangle \in \mathcal{H}$ can only be cloned in the sense of (2.7) if it is selected from a known set of orthogonal states on $\mathcal{H}$.

The following proof is due to Yuen [Yue86].

Proof. Let $\mathcal{E} : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{H} \otimes \mathcal{H})$ and $|\varphi\rangle, |\psi\rangle \in \mathcal{H}$ be such that

$$\mathcal{E}(|\varphi\rangle\langle\varphi|) = |\varphi\rangle|\varphi\rangle\langle\varphi|\langle\varphi| \quad \text{and} \quad \mathcal{E}(|\psi\rangle\langle\psi|) = |\psi\rangle|\psi\rangle\langle\psi|\langle\psi|.$$

In the remainder, we will show that this implies that $|\varphi\rangle$ and $|\psi\rangle$ either represent the same state or are orthogonal states.

Let $\mathcal{H}_R = \mathcal{H}$. Using Stinespring's dilation theorem, we can represent $\mathcal{E}(|\varphi\rangle\langle\varphi|)$ as

$$\begin{aligned}\mathcal{E}(|\varphi\rangle\langle\varphi|) &= \mathrm{tr}_R(U|\varphi\rangle|\omega_\circ\rangle|\omega_\circ\rangle\langle\varphi|\langle\omega_\circ|\langle\omega_\circ|U^\dagger) \\ &= \mathrm{tr}_R(|\varphi\rangle|\varphi\rangle|\chi\rangle\langle\varphi|\langle\varphi|\langle\chi|) \\ &= |\varphi\rangle|\varphi\rangle\langle\varphi|\langle\varphi|\end{aligned}$$

where $U \in \mathrm{End}(\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H}_R)$ is unitary, $|\omega_\circ\rangle = |0\cdots 0\rangle \in \mathcal{H}$ and $|\chi\rangle \in \mathcal{H}_R$. Similarly, we can represent $\mathcal{E}(|\psi\rangle\langle\psi|)$ in this form, where we let $|\chi'\rangle \in \mathcal{H}_R$ be the ancilla state after applying U (instead of $|\chi\rangle$).

To shorten notation, we proceed our analysis using state-vector notation, which we may use since our analysis is based on pure states. By focusing on the expressions inside the partial trace over R, we have the following two equations

$$U(|\varphi\rangle|\omega_\circ\rangle|\omega_\circ\rangle) = |\varphi\rangle|\varphi\rangle|\chi\rangle, \tag{2.8}$$
$$U(|\psi\rangle|\omega_\circ\rangle|\omega_\circ\rangle) = |\psi\rangle|\psi\rangle|\chi'\rangle, \tag{2.9}$$

We now take the inner product of both equations, i.e., we first write (2.9) as

$$(\langle\psi|\langle\omega_\circ|\langle\omega_\circ|)U^\dagger = \langle\psi|\langle\psi|\langle\chi'|,$$

and subsequently left-multiply the left and right hand side of (2.8) by respectively the left and right hand side of the expression above, to obtain

$$\langle\psi|\varphi\rangle = (\langle\psi|\varphi\rangle)^2\langle\chi'|\chi\rangle.$$

This equation is satisfied if $\langle\psi|\varphi\rangle = 0$, i.e., when $|\psi\rangle$ and $|\varphi\rangle$ are orthogonal. If they are not, then we can divide at both sides by $\langle\psi|\varphi\rangle$, to obtain $1 = \langle\psi|\varphi\rangle\langle\chi'|\chi\rangle$, which is only satisfied when $\langle\psi|\varphi\rangle\langle\chi|\chi'\rangle = \exp(\mathrm{i}\theta)$ for $\theta \in \mathbb{R}$, which means that $|\varphi\rangle$ and $|\psi\rangle$ represent the same state.

We conclude from this reasoning that only orthogonal states can be cloned, hence a general cloning operation is impossible. □

2.7.7 Alternative Descriptions of Measurements

There are several common ways of expressing measurements in quantum mechanics. We have chosen to express the fourth postulate in terms of "general measurements," where the latter is jargon from [NC00]. In this section, we discuss some of the alternative descriptions.

A *positive-operator-valued measure* (POVM) is obtained by setting $E_x := M_x^\dagger M_x$ for all x, where $\{M_x\}_x$ is a complete collection of measurement operators. The operators E_x, which are positive semi-definite, are called the POVM elements and the collection $\{E_x\}_x$, which obviously satisfies the completeness condition (i.e., $\sum_x E_x = \mathbb{I}$), is called "the POVM." *Vice versa*, every family of positive-semidefinite matrices that add up to the identity is a POVM, since any positive-semidefinite matrix E_x can always be decomposed as $E_x = M_x^\dagger M_x$. Note however that this decomposition is not unique. By substituting $E_x := M_x^\dagger M_x$ in (2.5) we see that the POVM elements determine the probabilities of the outcomes of the measurement. A POVM does not uniquely specify the post-measurement state, by the non-uniqueness of the decomposition above.

A special case of the general-measurement formalism is the class of *projective measurements*, where each measurement matrix is a projector. The completeness condition implies that these projectors project to mutually orthogonal subspaces, i.e., for the projective measurement $\{P_x\}_x$, it holds that $P_x P_{x'} = \delta_{xx'} P_x \quad \forall x, x'$.

If $x \in \mathbb{R}$ holds for every outcome x, then a compact way of representing a projective measurement is in the form of a Hermitian matrix called *observable* O, which has the following spectral decomposition

$$O = \sum_x x P_x.$$

A *complete projective measurement* is a projective measurement where all projectors have rank 1. Let A be a subsystem of an arbitrary composite system, and denote the state space of a by $\mathcal{H}_A$ and its dimension by d. Let $\mathcal{B} := \{|\psi_i\rangle\}_{i\in[d]}$ be an orthonormal basis for $\mathcal{H}_A$. When we say "we measure A in a basis $\mathcal{B}$," we mean a complete projective measurement, where the projectors are given by $P_i := |\psi_i\rangle\langle\psi_i|$ for all $i \in [d]$.

In particular, whenever we say that we measure an n-qubit state in the basis b, where $b = (b_1, \ldots, b_n) \in \{0,1\}^n$ is a bit string, we always mean that we measure the state qubit-wise as follows: for every $i \in [n]$, we measure the ith qubit in the computational basis (on $\mathbb{C}^2$) if $b_i = 0$, and in the Hadamard basis (of $\mathbb{C}^2$) if $b_i = 1$.

In case of a pure state, measuring-in-a-basis comes down to representing the state in the measurement basis. Suppose that we want to measure a state $|\psi\rangle \in \mathbb{C}^d$ in the computational basis. Let the representation of $|\psi\rangle$ in the computational basis be given by

$$|\psi\rangle = \sum_{x\in[d]} \alpha_x |x\rangle,$$

where $\alpha_x \in \mathbb{C}$ for all $x \in [d]$. Note that we call these coefficients α_x *amplitudes*. The probabilities of the possible measurement outcomes are given by the (absolute) squares of the amplitudes, i.e., $p_x = |\alpha_x|^2$. The post-measurement state when obtaining outcome x is $|x\rangle$.

2.7.8 Entanglement

From the postulates we know that if a density matrix $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ can be written as $\rho_{AB} = \rho_A \otimes \rho_B$, then the subsystems A and B are independent. In this case, we say that the state ρ_{AB} is a *product state*. More generally, a state is called *separable* if it can be written as a convex combination of product states,

$$\rho_{AB} = \sum_i \xi_i \, \sigma_{A,i} \otimes \sigma_{B,i},$$

where $\sigma_{A,i} \in \mathcal{H}_A$ and $\sigma_{B,i} \in \mathcal{H}_B$ for all i and where $\{\xi_i\}_i$ is a distribution.[15]

If a state is not separable, it is called *entangled*. One of the simplest examples of an entangled state is the *EPR pair* $|\Phi_{AB}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$, where $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^2$,

[15] Although we defined a distribution to be a *function*, here we of course mean that the numbers ξ_i are non-negative and sum up to one.

which we define to be[16]

$$|\Phi_{AB}\rangle := \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle).$$

When measuring the state, the entanglement exhibits itself in the form of a peculiar phenomenon, which Einstein called *"spukhaftige Fernwirkung"* ("spooky action at a distance"). To see this, suppose that Alice measures her subsystem A first, in the computational basis. Her outcome will be a random bit. Let us suppose Alice obtained the outcome 0, which means that the joint state has collapsed to $|00\rangle$. Now, if Bob measures in the *same basis as Alice* (i.e., the computational basis), he will get *the same outcome*. This effect is not limited to the computational basis; it also occurs for the Hadamard basis,[17] which becomes clear if we represent $|\Phi_{AB}\rangle$ in the Hadamard basis:

$$\begin{aligned}
|\Phi_{AB}\rangle &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \\
&= \frac{1}{\sqrt{2}}\Big(\frac{1}{2}(|+\rangle + |-\rangle)(|+\rangle + |-\rangle) + \frac{1}{2}(|+\rangle - |-\rangle)(|+\rangle - |-\rangle)\Big) \\
&= \frac{1}{\sqrt{2}}(|++\rangle + |--\rangle).
\end{aligned}$$

Note that because of symmetry, we may interchange the roles of Alice and Bob in the discussion above.

A well known application of EPR pairs (or, in general, entangled states) is *quantum key distribution* (QKD), which is a protocol for establishing a secret key between two parties. See Chapter 1 for a high-level explanation of QKD, and Chapter 3 for a technical treatment and proof.

2.7.9 Hybrid Systems

A special case of composite systems are *hybrid systems*, which are systems composed of both quantum and non-quantum (classical) subsystems. A state of such a hybrid system will be called a *cq-state*. We often encounter hybrid systems in quantum cryptography. For example, cryptographic keys are typically represented as classical subsystems, whereas the adversary's (quantum) information is described by a quantum subsystem.

[16] In the physics literature, an EPR pair is sometimes defined differently, i.e., as $\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$, which is also called the *singlet state*.

[17] When Alice and Bob perform their measurements on a singlet state, they will get opposite (binary) outcomes for *any* basis they use, as long as they use the same basis [NC00].

Although the behavior of the classical subsystems can be fully described in the language of probability theory, it is possible and usually more convenient to also use the density matrix formalism for representing the classical subsystems. Let X be a random variable over $\mathcal{X}$ with distribution P_X. The *density-matrix representation of* P_X with respect to an orthonormal basis $\{|x\rangle\}_{x\in\mathcal{X}}$ for $\mathbb{C}^{|\mathcal{X}|}$ is given by

$$\rho_X = \sum_{x\in\mathcal{X}} P_X(x)|x\rangle\langle x|.$$

From Theorem 2.36 it follows that the probabilities $P_X(x)$ coincide with the eigenvalues of ρ_X.

Following [Ren05], we say that a composite system ρ_{XE} is *classical with respect to* $\{|x\rangle\}_{x\in\mathcal{X}}$ if there exists a collection $\{\rho_E^x\}_{x\in\mathcal{X}}$ of density matrices on $\mathcal{H}_E$ such that ρ_{XE} can be written as

$$\rho_{XE} = \sum_{x\in\mathcal{X}} P_X(x)|x\rangle\langle x| \otimes \rho_E^x.$$

Moreover, by saying that ρ_{XE} *is classical on* $\mathcal{H}_X$ (or simply: *classical on* X) we mean that there exists a basis $\{|x\rangle\}_{x\in\mathcal{X}}$ of $\mathcal{H}_X$ such that ρ_{XE} is classical with respect to $\{|x\rangle\}_{x\in\mathcal{X}}$.

Tracing out X from ρ_{XE} yields

$$\rho_E = \operatorname{tr}_X(\rho_{XE}) = \sum_{x\in\mathcal{X}} P_X(x)\rho_E^x,$$

and the partial trace over a classical system coincides with marginalizing over that random variable. Furthermore, the fully mixed state on a classical subsystem coincides with the density-matrix representation of the uniform probability distribution over the corresponding random variable.

For a state ρ_{XE} that is classical on X we say that X *is random and independent from* E if

$$\rho_{XE} = \frac{1}{|\mathcal{X}|}\mathbb{I}_X \otimes \rho_E.$$

For example, in a cryptographic setting where X represents a classical key and E the quantum system held by an adversary, the above would mean that the key is perfectly secret with respect to the adversary.

For an arbitrary state ρ_{XE} that is classical on X, we may condition on any event $\mathcal{A}$ that is defined by $\Pr[\mathcal{A}|X = x]$ for all $x \in \mathcal{X}$ with $P_X(x) > 0$, we write this as

$$\rho_{XE|\mathcal{A}} = \sum_{x\in\mathcal{X}} P_{X|\mathcal{A}}(x)|x\rangle\langle x| \otimes \rho_E^x.$$

Tracing out X from this state gives

$$\rho_{E|\mathcal{A}} = \mathrm{tr}_X(\rho_{XE|\mathcal{A}}) = \sum_{x \in \mathcal{X}} P_{X|\mathcal{A}}(x) \rho_E^x.$$

If $\mathcal{A}$ is defined as the event $X = x$, then we get that $\rho_{E|\mathcal{A}} = \rho_{E|X=x} = \rho_E^x$.

Let ρ_{XYE} be cq-state of hybrid system XYE with classical X and Y. To express that the random variable X is independent of the quantum subsystem E *when given the random variable* Y, we say that ρ_{XYE} equals $\rho_{X \leftrightarrow Y \leftrightarrow E}$, where

$$\rho_{X \leftrightarrow Y \leftrightarrow E} := \sum_{x,y} P_{XY}(x,y) |x\rangle\langle x| \otimes |y\rangle\langle y| \otimes \rho_E^y. \tag{2.10}$$

This notion is called *conditional independence* and the quantum version above was introduced in [DFSS07].

2.7.10 Distance between States

The *trace norm* of a matrix A is defined as $\|A\|_1 := \mathrm{tr}\sqrt{A^\dagger A}$, where $\sqrt{A^\dagger A}$ is the positive semi-definite square root[18] of $A^\dagger A$. In case A is Hermitian, then the trace norm of A simplifies to $\sum_i |\lambda_i|$, where λ_i are the eigenvalues of A. The *trace distance* (between states) is the quantum analogue of the statistical distance (between distributions).

Definition 2.48 The *trace distance* between two density matrices $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ is defined as $\delta(\rho, \sigma) := \frac{1}{2}\|\rho - \sigma\|_1$.

Below, we list a number of important properties, for any $\rho, \sigma \in \mathcal{D}(\mathcal{H})$.

1. *Non-negativity:* $0 \leq \delta(\rho, \sigma)$;
2. *Identity of indiscernibles:* $\delta(\rho, \sigma) = 0$ if and only if $\rho = \sigma$;
3. *Symmetry:* $\delta(\rho, \sigma) = \delta(\sigma, \rho)$;
4. *Subadditivity / triangle inequality:* $\delta(\rho, \sigma) \leq \delta(\rho, \tau) + \delta(\tau, \sigma)$ for any $\tau \in \mathcal{D}(\mathcal{H})$;
5. *Bounded from above:* $\delta(\rho, \sigma) \leq 1$, with $\delta(\rho, \sigma) = 1$ if and only if $\mathrm{tr}(\rho\sigma) = 0$ (i.e., when ρ and σ are orthogonal);
6. *Unitary invariance:* for any unitary U, $\delta(U\rho U^\dagger, U\sigma U^\dagger) = \delta(\rho, \sigma)$;
7. *Subadditivity w.r.t. tensor products:* $\delta(\rho \otimes \rho', \sigma \otimes \sigma') \leq \delta(\rho, \sigma) + \delta(\rho', \sigma')$ for any $\rho', \sigma' \in \mathcal{D}(\mathcal{H}')$, with "=" if and only if $\rho' = \sigma'$;

[18] For a definition of the positive semi-definite square root, see Def. 9.4-1 in [Kre78].

8. *Contractive for partial trace:* for bipartite density operators $\rho_{AB}, \sigma_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, $\delta(\mathrm{tr}_B(\rho_{AB}), \mathrm{tr}_B(\sigma_{AB})) \leq \delta(\rho_{AB}, \sigma_{AB})$

Note that properties 1–4 imply that the trace distance is a metric.

Similar to the statistical distance, the trace distance can be interpreted as the maximal distinguishing probability.

Proposition 2.49 *[NC00] Let $\rho, \sigma \in \mathcal{D}(\mathcal{H})$. Then,*

$$\delta(\rho, \sigma) = \max_{\{E_m\}} \mathrm{SD}(p, q)$$

where the maximum is over all POVMs on $\mathcal{H}$, and p and q are the distributions of the outcomes of measuring ρ and σ respectively using $\{E_m\}$.

For two *pure* states $|\psi\rangle$ and $|\varphi\rangle$, the trace distance simplifies to

$$\delta(|\psi\rangle\langle\psi|, |\varphi\rangle\langle\varphi|) = \sqrt{1 - |\langle\psi|\varphi\rangle|^2}.$$

For two cq-states $\rho_{XE}, \sigma_{XE} \in \mathcal{D}(\mathcal{H}_X \otimes \mathcal{H}_E)$ with classical X and $\mathrm{tr}_E(\rho_{XE}) = \mathrm{tr}_E(\sigma_{XE})$, it holds that

$$\delta(\rho_{XE}, \sigma_{XE}) = \sum_x P_X(x)\, \delta(\rho_E^x, \sigma_E^x).$$

An important property of the trace distance is that it cannot increase when applying an arbitrary CPTP map to the states.

Theorem 2.50 *Let $\mathcal{H}$ be an arbitrary Hilbert space and let $\rho, \sigma \in \mathcal{D}(\mathcal{H})$. Let $\mathcal{E}$ be an arbitrary CPTP map on $\mathcal{H}$. Then,*

$$\delta(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \leq \delta(\rho, \sigma)$$

Proof. By Stinespring's dilation theorem (Theorem 2.46), any quantum operation $\mathcal{E}$ can be represented as the sequence: (1) composition with an ancilla, (2) unitary transformation, and (3) partial trace. By subadditivity with respect to tensor products, and the fact that for ρ and σ the same ancilla is added, step (1) does not change the trace distance. By unitary invariance, step (2) does also not change the trace distance. By contractivity for the partial trace, step (3) cannot increase the trace distance. Hence the assertion follows. □

In quantum cryptography, we often want to express the statistical distance between some state and another state that models an "ideal" or "desired situation," for example a state of which a subsystem is the fully mixed state and independent from the other subsystems. For this purpose, we introduce a compact notation.

Definition 2.51 For a density matrix $\rho_{XE} \in \mathcal{D}(\mathcal{H}_X \otimes \mathcal{H}_E)$ with classical X, the *distance to uniform* of X given E is defined as

$$d_{\mathsf{unif}}(X|E) := \tfrac{1}{2}\|\rho_{XE} - \rho_U \otimes \rho_E\|_1,$$

where $\rho_U := \frac{1}{\dim(\mathcal{H}_X)}\mathbb{I}_X$.

If also E is classical, then $d_{\mathsf{unif}}(X|E)$ simplifies to

$$\begin{aligned} d_{\mathsf{unif}}(X|E) &= \tfrac{1}{2}\sum_{x,e}|P_{XE}(x,e) - P_U(x)P_E(e)| \\ &= \sum_e P_E(e)\,\tfrac{1}{2}\sum_x |P_{X|E}(x|e) - P_U(x)|. \end{aligned}$$

It is not too hard to show that for a tri-partite system XYE with classical X and Y

$$d_{\mathsf{unif}}(X|YE) = \sum_{y\in\mathcal{Y}} P_Y(y)\, d_{\mathsf{unif}}(X|E, Y{=}y).$$

From this, the following lemma follows immediately.

Lemma 2.52 *For any y: $d_{\mathsf{unif}}(X|E, Y{=}y) \leq d_{\mathsf{unif}}(X|YE)/\Pr[Y = y]$.*

2.8 The One-Time Pad in a Quantum Setting

Proposition 2.53 *Let $\ell \in \mathbb{N}$. For any classical random variables M, K over $\mathbb{F}_2^\ell$ and arbitrary quantum system E, let $\frac{1}{2}\|\rho_{MKE} - \rho_M \otimes \rho_U \otimes \rho_E\|_1 \leq \varepsilon$, where $\rho_{MKE} \in \mathcal{D}(\mathcal{H}_M \otimes \mathcal{H}_K \otimes \mathcal{H}_E)$ and ρ_U is the fully mixed state on $\mathcal{H}_K$. Then, it holds that*

$$\tfrac{1}{2}\|\rho_{MCE} - \rho_M \otimes \rho_U \otimes \rho_E\|_1 \leq \varepsilon$$

where $C := M \oplus K$.

Proof. Follows immediately from Theorem 2.50. □

2.9 Measures of Uncertainty for Density Matrices

Definition 2.54 Let $\mathcal{H}$ be some Hilbert space. For a density matrix $\rho \in \mathcal{D}(\mathcal{H})$ the *von Neumann entropy* is given by

$$H(\rho) := -\mathrm{tr}(\rho \log \rho),$$

Equivalently, $H(\rho) = -\sum_i \lambda_i \log \lambda_i$, where λ_i are the eigenvalues of ρ. When ρ is classical, then the von Neumann entropy coincides with the Shannon entropy.

The following definition of conditional min-entropy is due to [Ren05].

Definition 2.55 Let $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and $\sigma_B \in \mathcal{D}(\mathcal{H}_B)$. The *min-entropy of ρ_{AB} relative to σ_B* is defined[19] as

$$H_{\min}(\rho_{AB}|\sigma_B) := -\log \inf\{\lambda \in \mathbb{R} : \lambda \cdot \mathbb{I}_A \otimes \sigma_B - \rho_{AB} \geq 0\}.$$

The *min-entropy of ρ_{AB} when given $\mathcal{H}_B$* is defined[20] as

$$H_{\min}(\rho_{AB}|B) := \sup_{\sigma_B} H_{\min}(\rho_{AB}|\sigma_B).$$

When the state ρ_{AB} is clear from the context, we prefer writing $H_{\min}(A|B)$ for $H_{\min}(\rho_{AB}|B)$ and say "the min-entropy of A given B." If $\mathcal{H}_B$ is the trivial space $\mathbb{C}$, we obtain the unconditional min-entropy of ρ_A, denoted as $H_{\min}(\rho_A)$, or $H_{\min}(A)$ if ρ_A is clear from its context, which simplifies to

$$H_{\min}(\rho_A) = -\log \lambda_{\max}(\rho_A),$$

where $\lambda_{\max}(\rho_A)$ is the largest eigenvalue of ρ_A.

For the special case of a hybrid state $\rho_{XE} \in \mathcal{D}(\mathcal{H}_X \otimes \mathcal{H}_E)$ with classical X, it is shown in [KRS09] that the conditional min-entropy of a quantum state coincides with the negative logarithm of the *guessing probability conditional on quantum side information*

$$p_{\mathsf{guess}}(X|E) := \max_{\{M_x\}} \sum_x P_X(x) \operatorname{tr}(M_x \rho_E^x),$$

where the latter is the probability that the party holding $\mathcal{H}_E$ guesses X correctly using the POVM $\{M_x\}_x$ on $\mathcal{H}_E$ that maximizes p_{guess}. Thus,

$$H_{\min}(X|E) = -\log p_{\mathsf{guess}}(X|E). \tag{2.11}$$

The following proposition guarantees that the "averaging property" of the guessing probability (which holds by definition in the classical case, see (2.1) in Section 2.3.2) still holds when additionally conditioning on a quantum system.

[19] There exist choices for σ_B for which the set $\{\lambda \in \mathbb{R} : \lambda \cdot \mathbb{I}_A \otimes \sigma_B - \rho_{AB} \geq 0\}$ is empty. For this reason, we define $\inf \emptyset := \infty$.

[20] If $\mathcal{H}_B$ has finite dimension (in this thesis, we anyway solely deal with finite-dimensional Hilbert spaces), then the set $\mathcal{D}(\mathcal{H}_B)$ is compact and hence the supremum can be replaced by a maximum [Ren05].

Proposition 2.56 *For any state $\rho_{XYE} \in \mathcal{D}(\mathcal{H}_X \otimes \mathcal{H}_Y \otimes \mathcal{H}_E)$ that is classical on X and Y it holds that*

$$p_{\mathsf{guess}}(X|YE) = \sum_y P_Y(y)\, p_{\mathsf{guess}}(X|E, Y = y).$$

Proof. First, note that for any matrix M_x acting on $\mathcal{H}_Y \otimes \mathcal{H}_E$, we can always write $M_x = \sum_{y,y'} |y\rangle\langle y'| \otimes M_x^{y,y'}$, where $M_x^{y,y'}$ acts on $\mathcal{H}_E$ for every x, y, y'. Now, we write

$$\begin{aligned}
p_{\mathsf{guess}}(X|YE) &= \max_{\{M_x\}} \sum_x P_X(x) \mathrm{tr}(M_x \rho_{YE}^x) \\
&= \max_{\{M_x\}} \sum_x P_X(x) \mathrm{tr}(M_x \sum_y P_{Y|X}(y|x)\, |y\rangle\langle y| \otimes \rho_E^{x,y}) \\
&= \max_{\{M_x\}} \sum_{x,y} P_{XY}(x,y) \mathrm{tr}((\sum_{v,w} |v\rangle\langle w| \otimes M_x^{v,w})(|y\rangle\langle y| \otimes \rho_E^{x,y})) \\
&= \max_{\{M_x\}} \sum_{x,y} P_{XY}(x,y) \sum_v \langle v|y\rangle \mathrm{tr}(M_x^{v,y} \rho_E^{x,y}) \\
&= \max_{\{M_x\}} \sum_{x,y} P_{XY}(x,y) \mathrm{tr}(M_x^{y,y} \rho_E^{x,y}) \\
&= \sum_y P_Y(y) \max_{\{M_x^{y,y}\}} \sum_x P_{X|Y}(x|y) \mathrm{tr}(M_x^{y,y} \rho_E^{x,y}) \\
&= \sum_y P_Y(y)\, p_{\mathsf{guess}}(X|E, Y = y).
\end{aligned}$$

□

We also define the (unconditional) *max-entropy* of a density matrix. In the literature, one typically finds two different definitions for max-entropy (i.e., [Ren05] versus [KRS09]). The following definition satisfies our needs.

Definition 2.57 Let $\rho \in \mathcal{D}(\mathcal{H})$. The *max-entropy* of ρ is defined as

$$H_{\max}(\rho) := \log \operatorname{rank}(\rho).$$

Proposition 2.58 *The conditional min-entropy is invariant under local unitaries, that is, for $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and an arbitrary unitary U with product structure on $\mathcal{H}_A \otimes \mathcal{H}_B$, i.e., $U := U_A \otimes U_B$, it holds that*

$$H_{\min}(U \rho_{AB} U^\dagger | B) = H_{\min}(\rho_{AB}|B). \tag{2.12}$$

Furthermore, the unconditional max-entropy is unitarily invariant,

$$H_{\max}(U_A \rho_A U_A^\dagger) = H_{\max}(\rho_A). \tag{2.13}$$

Proof. To prove (2.12) it suffices to show that

$$\lambda \cdot \mathbb{I}_A \otimes \sigma_B - \rho_{AB} \geq 0 \implies \exists\, \sigma'_B \text{ such that } \lambda \cdot \mathbb{I}_A \otimes \sigma'_B - U\rho_{AB}U^\dagger \geq 0$$

Because U is unitary, it holds that

$$\lambda \cdot \mathbb{I}_A \otimes \sigma_B - \rho_{AB} \geq 0 \implies U(\lambda \cdot \mathbb{I}_A \otimes \sigma_B - \rho_{AB})U^\dagger \geq 0$$

Finally,

$$\begin{aligned} U(\lambda \cdot \mathbb{I}_A \otimes \sigma_B - \rho_{AB})U^\dagger &= \lambda \cdot U(\mathbb{I}_A \otimes \sigma_B)U^\dagger - U\rho_{AB}U^\dagger \\ &= \lambda \cdot U_A \mathbb{I}_A U_A^\dagger \otimes U_B \sigma_B U_B^\dagger - U\rho_{AB}U^\dagger \\ &= \lambda \cdot \mathbb{I}_A \otimes \sigma'_B - U\rho_{AB}U^\dagger, \end{aligned}$$

where $\sigma'_B = U_B \sigma_B U_B^\dagger$.

For (2.13), the claim immediately follows from the fact that a unitary transformation leaves the eigenvalues of the operator to which it is applied unchanged. □

The following proposition is known as the chain rule for min-entropy.

Proposition 2.59 ([Ren05]) *The following holds for all $\rho_{ABC} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$,*

$$H_{\min}(A|BC) \geq H_{\min}(AB|C) - H_{\max}(B).$$

The following shows that removing a classical subsystem only reduces the min-entropy.

Proposition 2.60 ([Ren05]) *The following holds for all $\rho_{AXC} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_X \otimes \mathcal{H}_C)$ with classical X,*

$$H_{\min}(AX|C) \geq H_{\min}(A|C).$$

As a corollary of Proposition 2.59 and 2.60, we obtain a chain rule that we will often use.

Corollary 2.61 (Chain Rule for Removing Classical Subsystems) *The following holds for all $\rho_{AXC} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_X \otimes \mathcal{H}_C)$ with classical X,*

$$H_{\min}(A|XC) \geq H_{\min}(A|C) - H_{\max}(X).$$

The following chain rule is particularly useful to prove security in the bounded-quantum-storage model.

Proposition 2.62 *For any $\rho \in \mathcal{D}(\mathcal{H}_{XYE})$ with classical X and Y it holds that*

$$H_{\min}(X|YE) \geq H_{\min}(X|Y) - H_{\max}(E).$$

To prove Proposition 2.62, we will use the following lemma.

Lemma 2.63 *For any state $\rho_{XYE} \in \mathcal{D}(\mathcal{H}_X \otimes \mathcal{H}_Y \otimes \mathcal{H}_E)$ that is classical on X and Y it holds that*

$$H_{\min}(XE|Y=y) \geq H_{\min}(X|Y=y) \tag{2.14}$$

for every $y \in \mathcal{Y}$.

Proof. Note that it suffices to show that $\lambda_{\max}(\rho^y_{XE}) \leq \lambda_{\max}(\rho^y_X)$ holds for every $y \in \mathcal{Y}$. Because ρ^y_{XE} is classical on X, there exists a unitary U acting on $\mathcal{H}_X$ such that $\widetilde{\rho}^y_{XE} := (U \otimes \mathbb{I}_E)\rho^y_{XE}(U^\dagger \otimes \mathbb{I}_E)$ is classical with respect to the computational basis $\{|x\rangle\}_{x\in\mathcal{X}}$ on $\mathcal{H}_X$ with $\mathcal{X} := [d]$. In particular, this means that $\widetilde{\rho}^y_{XE}$ has block-diagonal structure:

$$\widetilde{\rho}^y_{XE} = \sum_{x\in[d]} P_{X|Y}(x|y)|x\rangle\langle x| \otimes \rho^{x,y}_E = \begin{bmatrix} P_{X|Y}(1|y)\,\rho^{1,y}_E & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & P_{X|Y}(d|y)\,\rho^{d,y}_E \end{bmatrix}.$$

Note that because U is unitary, $\widetilde{\rho}^y_{XE}$ has the same eigenvalues as ρ^y_{XE}, where these eigenvalues are given by the union of the eigenvalues of the blocks on the diagonal of $\widetilde{\rho}^y_{XE}$. From this we see that the largest eigenvalue of $\widetilde{\rho}^y_{XE}$ (and thus of ρ^y_{XE}) cannot be larger than the largest eigenvalue of $\widetilde{\rho}^y_X := \operatorname{tr}_E(\widetilde{\rho}^y_{XE})$ (and thus of ρ^y_X). □

Proof of Proposition 2.62 . By (2.11) it is equivalent to show that

$$p_{\mathsf{guess}}(X|YE) \leq p_{\mathsf{guess}}(X|Y)\, 2^{H_{\max}(E)}.$$

Using Proposition 2.56, we write

$$\begin{aligned}
p_{\mathsf{guess}}(X|EY) &= \sum_y P_Y(y)\, p_{\mathsf{guess}}(X|E, Y=y) = \sum_y P_Y(y)\, 2^{-H_{\min}(X|E,Y=y)} \\
&\leq \sum_y P_Y(y)\, 2^{-(H_{\min}(XE|Y=y)-H_{\max}(E))} \\
&\leq 2^{H_{\max}(E)} \sum_y P_Y(y) 2^{-H_{\min}(X|Y=y)} = 2^{H_{\max}(E)}\, p_{\mathsf{guess}}(X|Y),
\end{aligned}$$

where the first inequality is Proposition 2.59, and the second inequality follows by Lemma 2.63. Hence, the claim follows. □

2.10 Extractors against Quantum Side Information

A natural generalization of the randomness extraction problem is to allow the side information, i.e., the random variable Y in Definition 2.24, to be a quantum state.

Definition 2.64 A function Ext : $\{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a (k,ε) *strong extractor against quantum side information*, if for any bipartite quantum system XE with classical X and with $H_{\min}(X|E) \geq k$, and for a uniform and independent seed S, we have

$$d_{\mathsf{unif}}\big(\mathsf{Ext}(X,S)\big|SE\big) \leq \varepsilon\,.$$

Note that we find "extractor against quantum side information" a too cumbersome terminology; thus we just call Ext a (strong) extractor, even though it is a stronger notion than the standard notion of a (strong) extractor. When necessary, we distinguish between the two notions by saying that an extractor is or is not *secure against quantum side information.*

A well-known example of a strong extractor (that is secure against quantum side information) is a two-universal hash function. The parameters of this extractor are given by the privacy amplification theorem for quantum adversaries due to Renner and König [RK05].

Theorem 2.65 (Privacy Amplification) *Let ρ_{XE} be a hybrid state with classical X. Let h : $\{0,1\}^n \times \{0,1\}^d \to \{0,1\}^q$ be a universal hash function, and let S be uniformly distributed over $\{0,1\}^d$, independent of X and E. Then it holds that*

$$d_{\mathsf{unif}}(h(X,S)|SE) \leq \frac{1}{2}\sqrt{2^{q-H_{\min}(X|SE)}} = \frac{1}{2}\sqrt{2^q\, p_{\mathsf{guess}}(X|SE)}.$$

At the time of this writing, the state of the art is described in [DPVR09], which presents an extractor secure against quantum side information with a seed length that is polylogarithmic in the input length ($d = O(\log^3 n)$).

2.11 Quantum Identification

In Chapters 4 and 5 we present contributions to different aspects of the *quantum identification protocol* (QID protocol) from [DFSS07]. In this section, we give an overview of this protocol.

First of all, let us say in some more detail what we mean by the term "protocol." A *protocol* is a specification of a sequence of operations (steps), to be performed by two or more parties. At the start, the protocol may take (classical or quantum) inputs, which may be specific for each party. In the following steps, the parties perform local computations and exchange (classical or quantum) messages. At the end, the parties produce their outputs. Whenever a party expects a message that either never arrives or arrives in the wrong format, that party will use a default message instead and will then continue executing the protocol.

The goal of *password-based identification* is to "prove" knowledge of a password w (or some other low-entropy key, like a PIN) without giving w away. More formally, given a user U and a server S that hold a pre-agreed password $w \in \mathcal{W}$, U wants to convince S that he indeed knows w, but in such a way that he gives away as little information on w as possible in case he is actually interacting with a dishonest server S* (who does not know w).

An informal behavioral description of a cryptographic task, such as the one above, can be formalized as an *ideal functionality*. In the case of password-based identification, the ideal functionality $\mathcal{F}$ computes the equality function; it takes as inputs the passwords of U and S, and outputs to S a single bit that tells whether the passwords are equal or not.

A contribution from [DFSS07] beyond the QID protocol itself are the formal security definitions for quantum password-based identification. As shown in [FS09], these definitions are uniquely determined by the ideal functionality and guarantee a special form of sequential composability: if QID protocol π securely implements ideal functionality $\mathcal{F}$ (according to the definitions given below), then any classical two-party protocol that makes sequential calls to $\mathcal{F}$ remains secure when the calls to $\mathcal{F}$ are replaced by invocations of the QID protocol π.

Definition 2.66 (Correctness) An identification protocol is said to be ε-*correct* if, after an execution by honest U and honest S, S accepts with probability $1 - \varepsilon$.

The following two definitions are a bit less obvious. The intuition behind them is as follows.

In the ideal world,[21] a dishonest party (either U^* or S^*, depending on which of the two definitions you consider) whose view is independent of the password W cannot do better than to guess this password, and learn whether this guess was correct.[22] Note that if the guess was wrong, then the dishonest party can discard this candidate password; if it was right, then there is no security left. It is crucial that the security definition allows this guessing strategy; otherwise the definition can never be achieved by any protocol. Let the adversary's guess be modeled as the random variable W'. Formally, we can express the final state in the ideal world (conditioned on that the guess was wrong) as $\rho_{W \leftrightarrow W' \leftrightarrow E|W \neq W'}$.

In the real world, the parties execute a protocol to emulate the ideal functionality. We want to show that no matter which strategy the dishonest party has, and when given all messages exchanged during the execution of the protocol, the dishonest party can essentially not learn more information about W than the dishonest party in the ideal world. Formally, this is done by showing that, after the execution of the protocol, there exists a random variable W' in the real world for which the joint state $\rho_{WW'E}$, conditioned on the event where $W' \neq W$, is close to the final state in the ideal world.

Definition 2.67 (User Security) An identification protocol for two parties U, S is ε-secure for the user U against (dishonest) server S^* if the following holds: If the initial state of S^* is independent of W, then its state E after execution of the protocol is such that there exists a random variable W' that is independent of W and such that

$$\delta(\rho_{WW'E|W \neq W'}, \rho_{W \leftrightarrow W' \leftrightarrow E|W \neq W'}) \leq \varepsilon.$$

Definition 2.68 (Server Security) An identification protocol for two parties U, S is ε-secure for the server S against (dishonest) user U^* if the following holds: whenever the initial state of U^* is independent of W, then there exists a random variable W' (possibly $\perp$) that is independent of W such that if $W \neq W'$ then S

[21] That is, the ideal setting where the ideal functionality is used.

[22] Note that in case the user is the dishonest party, he does not learn the correctness of his guess directly from the ideal functionality, since the latter only outputs a bit to the server. However, the dishonest user can typically deduce the correctness of his guess from subsequent behavior of the server.

accepts with probability at most ε, and if $W = W'$ then S accepts with certainty.[23] Furthermore, the common state ρ_{WE} after execution of the protocol (including S's announcement to accept or reject) satisfies

$$\delta(\rho_{WW'E|W\neq W'}, \rho_{W\leftrightarrow W'\leftrightarrow E|W\neq W'}) \leq \varepsilon.$$

It is well-known that these definitions cannot be achieved with suitable parameters (i.e., exponentially small ε) without making additional assumptions (beyond trusting the laws of quantum mechanics).

In [DFSS07], Damgård *et al.* show the existence of a secure identification protocol in the bounded-quantum-storage model. The protocol involves the communication of qubits, and is secure against an arbitrary dishonest server S^* that has limited quantum-storage capabilities and can only store a certain fraction of the communicated qubits, whereas the security against a dishonest user U^* holds unconditionally.

In fact, two QID protocols are proposed in [DFSS07], `QID` and `QID`$^+$. The former is truly password-based but does not protect against a man-in-the-middle attack, whereas the latter is secure against a man-in-the-middle attack but is not truly password-based, because U and S need to additionally share a secret high-entropy key.[24]

2.11.1 The Basic QID Protocol

Let $\mathcal{C} \subset \{0,1\}^n$ be a binary code with minimum distance d, and let $\mathfrak{c} : \mathcal{W} \to \mathcal{C}$ be its encoding function. Let $m := |\mathcal{W}|$, and typically, $m < 2^n$. Let $\mathcal{F}$ be the class of all linear functions from $\mathbb{F}_2^n$ to $\mathbb{F}_2^\ell$, where $\ell < n$, represented as $\ell \times n$ matrices over $\mathbb{F}_2$. Note that $\mathcal{F} = \mathcal{G}_1$ with $r = \ell$, where $\mathcal{G}_1$ is defined and shown to be universal in Section 2.4.1. Furthermore, let $\mathcal{G}$ be a strongly two-universal class of hash functions from $\mathcal{W}$ to $\mathbb{F}_2^\ell$. When we write H^v for any $v = (v_1, \ldots, v_n) \in \{0,1\}^n$ (recall that H is the Hadamard matrix on $\mathbb{C}^2$), we mean $H^{v_1} \otimes \cdots \otimes H^{v_n}$. `QID` is shown as Protocol 2.1.

We find it convenient to specify protocols in terms of fixed values (instead of random variables). In the proofs, we then usually switch to random-variable notation. We will not give the security proofs of `QID` here, they can be found in [DFSS07]. Instead, we describe at a high level how the protocol works and why it provides security.

[23] The latter clause is not present in [DFSS07], but achieves a more natural ideal functionality, as pointed out in [FS09].

[24] The high-entropy key is only needed to protect against a man-in-the-middle attack, security against dishonest U and S only relies on the password and holds even if the dishonest party knows the high-entropy key.

1. U selects $x \overset{\mathtt{r}}{\leftarrow} \{0,1\}^n$ and $\theta \overset{\mathtt{r}}{\leftarrow} \{0,1\}^n$ and sends $H^\theta|x\rangle$ to S.
2. S measures $H^\theta|x\rangle$ in basis $c = \mathfrak{c}(w)$. Let x' be the outcome.
3. U selects $f \overset{\mathtt{r}}{\leftarrow} \mathcal{F}$ and sends θ and f to S. Both compute $\mathcal{I}_w := \{i : \theta_i = \mathfrak{c}(w)_i\}$.
4. S selects $g \overset{\mathtt{r}}{\leftarrow} \mathcal{G}$ and sends g to U.
5. U computes and sends $z := f(x_{\mathcal{I}_w}) \oplus g(w)$ to S.
6. S accepts if and only if $z = z'$ where $z' := f(x'_{\mathcal{I}_w}) \oplus g(w)$.

Protocol 2.1: The quantum password-based-identification protocol `QID` from [DFSS07].

The protocol starts with a qubit-communication phase, in which the user sends BB84-encoded qubits (i.e., random basis states from either the computational or Hadamard basis) to the server. The server measures the ith qubit for all $i \in [n]$ in the computational or the Hadamard basis, depending on the ith position in a length-n binary codeword c, where this codeword is determined by the password w.

After this communication phase, where we assume that the server measured (most of) the qubits upon reception, the user announces the bases in which the qubits were encoded. At this point, the parties can derive a common raw key by selecting those positions where their bases coincide, i.e., $x_{\mathcal{I}_w}$. For the security of the user, it will be important that the server can only derive one raw key belonging to a single choice of w. Note that the bounded-quantum-storage assumption about the server is crucial here; if the server can delay all measurements beyond the point at which the user announces the bases, he can measure each qubit in the correct basis, and subsequently compute a separate raw key for each choice of w.

Next, the user sends a function f randomly chosen from a universal family to the server, and the server sends a function g randomly chosen from a strongly universal family to the user. Then, the user sends $z := f(x_{\mathcal{I}_w}) \oplus g(w)$ to the server. At the core of the user-security proof is a lower bound on the min-entropy of $x_{\mathcal{I}_w}$ from the dishonest server's point of view, which follows from the uncertainty relation from [DFR$^+$07]. The function f performs privacy amplification to this raw key, resulting in a shorter but almost uniform key $f(x_{\mathcal{I}_w})$. The latter key is used as a one-time pad such that z will be close to independent from w, regardless of g, and protects the user against a dishonest server.

The purpose of g is to protect the server against a dishonest user. By the strongly universal property, the value of $g(w)$ is different for each w with overwhelming probability, which makes it extremely unlikely that a dishonest user can produce a valid z for incorrect guesses of w. Furthermore, it guarantees that the accept/reject decision of the server cannot be exploited by the dishonest user to learn anything beyond the correctness of his guess for w.

2.11.2 Security against Man-in-the-Middle Attacks

As mentioned above, protocol `QID` is only proven secure against impersonation attacks. Moreover, `QID` is actually insecure in case of a man-in-the-middle attack. For example, the attacker can measure the first qubit in a fixed basis and forward the collapsed qubit. If the server subsequently rejects, then the attacker knows that he inserted an error; hence the first qubit must have been encoded in a basis opposite to the attacker's measurement basis. This in turn gives the attacker one bit of information on w.

The QID protocol that is in addition secure against man-in-the-middle attacks is called $\texttt{QID}^+$. It is obtained from a noise-tolerant version of `QID`, by introducing consistency checks in the qubit communication phase and to additionally authenticate all classical communication. For details about the noise-tolerant version of protocol `QID`, see [DFSS07]. For the high-level discussion here, it suffices to know that, informally speaking, $\{\mathrm{syn}_j\}_{j\in\mathcal{J}}$ for some non-empty set $\mathcal{J}$ is a special family of syndrome functions that also acts as an extractor: if a randomly selected syndrome function is applied to a random variable with large enough min-entropy, then the output will be close to the uniform distribution. This property prevents leakage of information about w.

The task of authenticating all classical messages can be performed using a standard information-theoretic authentication code, which requires an authentication key, which may only be re-used a limited number of times. Hence, when using standard authentication, the parties need to refresh the authentication key after a fixed number of protocol executions, e.g., using QKD. The main problem of this approach is that an attacker can repeatedly enforce the QID and QKD protocols to abort, in order to let the parties run out of key material. Damgård *et al.* [DFSS07] circumvent this problem by performing the authentication in the following special way such that the authentication key can be re-used. In Protocol 2.2, MAC^*_k is an *extractor-MAC*, which has the following property. If the message for which the tag is computed contains sufficient min-entropy conditioned on the adversary's view, then the tag is close to uniform when given the key k and the adversary's view. Note that the message is guaranteed to have sufficient min-entropy by including $x_{\mathcal{I}}$. Hence, the adversary only learns a very small amount of information about the key from observing the tag, which allows the honest parties to reuse the key.

1. U selects $x \overset{\mathtt{r}}{\leftarrow} \{0,1\}^n$ and $\theta \overset{\mathtt{r}}{\leftarrow} \{0,1\}^n$ and sends $H^\theta|x\rangle$ to S.
2. S selects a test set $\mathcal{T} \subset [n]$ of size ℓ at random, computes $c = \mathfrak{c}(w)$ and replaces c_i for all $i \in \mathcal{T}$ by random bits, and then measures $H^\theta|x\rangle$ in basis c. Let x' be the outcome, and let $test' := x'_{\mathcal{T}}$.
3. Let $\mathcal{I} := \{i : \theta_i = \mathfrak{c}(w)_i\}$. U selects $f \overset{\mathtt{r}}{\leftarrow} \mathcal{F}$ and $j \overset{\mathtt{r}}{\leftarrow} \mathcal{J}$ and sends θ, j, $s := \mathrm{syn}_j(x_{\mathcal{I}})$ and f to S.
4. S selects $g \overset{\mathtt{r}}{\leftarrow} \mathcal{G}$ and sends g and $\mathcal{T}$ to U.
5. U computes and sends $test := x_{\mathcal{T}}$, $z := f(x_{\mathcal{I}_w}) \oplus g(w)$ and $tag^* := \mathsf{MAC}^*_k(\theta, j, s, f, g, \mathcal{T}, test, z, x_{\mathcal{I}})$ to S.
6. S recovers $x_{\mathcal{I}}$ from $x'_{\mathcal{I}}$ using $test$ and s, and accepts if and only if (1) tag^* verifies correctly, (2) $test$ coincides with $test'$ at the positions where the bases coincide, and (3) $z = f(x'_{\mathcal{I}_w}) \oplus g(w)$.

Protocol 2.2: The quantum identification protocol $\mathtt{QID}^+$, which is also secure against man-in-the-middle attacks. To achieve the latter, it requires an additional high-entropy key, which is called k here.

3

Random Sampling from a Quantum Population

The content of this chapter is based on joint work with Serge Fehr [BF10].

Chapter Contents

3.1 Introduction . . . 96

3.1.1 Proposed Framework . . . 97

3.1.2 Applications . . . 97

3.1.3 Notation . . . 99

3.2 Sampling from a Classical Population . . . 99

3.2.1 Sampling Strategies . . . 99

3.2.2 The Error Probability . . . 102

3.2.3 Error Probabilities of the Example Sampling Strategies 103

3.3 Sampling from a Quantum Population . . . 108

3.3.1 Analyzing Sampling Strategies in the Quantum Setting 109

3.3.2 The Quantum vs. the Classical Error Probability . . . 111

3.3.3 Superpositions with a Small Number of Terms . . . 115

3.4 A Security Proof for Quantum Oblivious Transfer . . . 118

3.4.1 Security against Dishonest Alice . . . 120

3.4.2 The Harder Case: Security against Dishonest Bob . . . 122

3.5 An Accessible Proof for Quantum Key Distribution . . . 125

3.5.1 Survey of Existing QKD Proofs . . . 126

3.5.2 Our Proof . . . 127

3.6 Conclusion . . . 132

3.1 Introduction

Sampling allows to learn some information on a large population by merely looking at a relatively small number of individuals. For instance it is possible to predict the outcome of an election with very good accuracy by analyzing a relatively small subset of all the votes. In this chapter, we study the act of sampling from a *quantum* population, where we want to be able to learn information on a large quantum state by measuring only a small part. Specifically, we investigate the quantum version of the following classical sampling problem (and of variants thereof). Given a bit string $q = (q_1, \ldots, q_n) \in \{0,1\}^n$ of length n, the task is to estimate the Hamming weight of q by sampling and looking at only a few positions within q. This classical sampling problem is well understood. For instance the following particular *sampling strategy* works well: sample (with or without replacement) a linear number of positions uniformly at random, and compute an estimate for the Hamming weight of q by scaling the Hamming weight of the sample accordingly; Hoeffding's inequality (Theorem 2.11) guarantees that the estimate is close to the real Hamming weight except with small probability. Such a sampling strategy in particular allows to *test* whether q is close to the all-zero string $(0, \ldots, 0)$ by looking only at a relatively small number of positions, where the test is accepted if and only if all the sample positions are zero, i.e., the estimated Hamming weight vanishes.

In the quantum version of the above sampling problem, the string q is replaced by a n-qubit quantum system A. It is obvious that a sampling strategy from the classical setting can be *applied* to the quantum setting as well: pick a sample of qubit positions within A, measure (in the computational basis) these sample positions, and compute the estimate as dictated by the sampling strategy from the observed values (i.e., typically, scale the Hamming weight of the measured sample appropriately). However, what is not clear *a priori*, is how to formally *interpret* the computed estimate. In the special case of testing closeness to the all-zero string, one expects that if the measurement of a random sample only produces zeros then the initial state of A must have been close to the all-zero state $|0\rangle \cdots |0\rangle$. But what is the right way to measure closeness here? For instance it must allow for states of the form $|q\rangle$ where $q \in \{0,1\}^n$ has small Hamming weight, but it must also allow for superpositions with arbitrary states that come with a very small amplitude. In the general case of a sampling strategy that, in its classical usage, aims at estimating the Hamming weight (rather than at testing closeness to the all-zero string), it is not even clear what the estimate actually estimates when the sampling strategy is applied to an n-qubit quantum system, since we cannot speak of the Hamming weight of a quantum state. Furthermore, when applying a sampling strategy to a quantum population, how should we quantify its accuracy? And, when a definition for this accuracy has been established, is it actually feasible to compute (good bounds on) this accuracy? Finally, a last subtlety that is inherent to the quantum setting is that the execution of a sampling strategy actually changes the state of A due to the measurements.

3.1.1 Proposed Framework

In this chapter, we present a framework that answers the above questions and allows us to fully understand how a classical sampling strategy behaves when applied to a quantum population, i.e., to an n-qubit system or, more generally, to n copies of an arbitrary "atomic" system. Our framework incorporates the following. First, we specify an abstract property on the state of A (after the measurements done by the sampling strategy), with the intended meaning that this is the property one should conclude from the outcome of the sampling strategy when applied to A. We also demonstrate that this property has useful consequences: specifically, that a suitable measurement will lead to a high-entropy outcome; this is useful in particular for quantum-cryptographic purposes. Then, we define a meaningful measure, sort of a "quantum error probability" (although technically speaking it is not a probability), that tells how reliable it is to conclude the specified property from the outcome of the sampling strategy. Finally, we show that for *any* sampling strategy, the quantum error probability of the strategy, as we define it, is bounded by the square root of its classical error probability. This means that in order to understand how well a sampling strategy performs in the quantum setting, it suffices to analyze it in the classical setting. For typical sampling strategies, such as picking the sample uniformly at random, there are well-known good bounds on the classical error probability.

3.1.2 Applications

We demonstrate the usefulness of our framework by proposing new and simple(r) proofs for existing quantum-cryptographic protocols. Furthermore, we think that our framework can be valuable in other applications as well.

Simple Proof for Quantum Oblivious Transfer from Bit Commitment

The first application is to *quantum oblivious transfer* (QOT). It is well known that QOT is not possible from scratch; however, one can build a secure QOT scheme when given a *bit commitment* (BC) primitive "for free."[1] Like QOT, also QBC is impossible from scratch; nevertheless, the implication from BC to QOT is interesting from a theoretical point of view, since the corresponding implication does not hold in the classical setting. The existence of a QOT scheme based on a BC was suggested by Bennett *et al.* in 1991 [BBCS91];[2] however, no security proof was provided. Mayers and Salvail proved security of the QOT scheme against a restricted adversary that only performs *individual* measurements [MS94], and finally, in 1995, Yao gave a security proof against a general adversary, which is allowed to do fully *coherent* measurements [Yao95]. However, from today's perspective, Yao's proof is still not fully satisfactory: it is very technical, without intuition and hard to follow, and it measures the adversary's information in terms of "accessible information," which has proven to be a too weak information measure [BOHL$^+$05, RK05, KRBM07].

In Section 3.4, we show how our framework for analyzing sampling strategies in the quantum setting leads to a conceptually very simple and easy-to-understand security proof for QOT from BC. The proof essentially works as follows: When considering a purified version of the QOT scheme, the commit-and-open phase of the QOT scheme can be viewed as executing a specific sampling strategy. From the framework, it then follows that some crucial piece of information has high entropy from the adversary's point of view. The proof is then concluded by applying the privacy amplification theorem. Note that in [DFL$^+$09], it is shown that the same kind of analysis is not restricted to QOT but actually applies to a large class of two-party quantum-cryptographic schemes which are based on a commit-and-open phase.

Simple Proof for Quantum Key Distribution

In Section 3.5 we discuss our second application, being quantum key distribution (QKD). Also here, our framework allows for a simple and easy-to-understand security proof, namely for the BB84 QKD scheme.[3] Similar to our proof for QOT, we can view the checking phase of the BB84 scheme as executing a specific sampling strategy (although here some additional non-trivial observation needs to be made).

[1] We use BC and OT as short-hands of the respective abstract primitives, bit commitment and oblivious transfer, and we write QBC and QOT for potential schemes implementing the respective primitives in the quantum setting.

[2] At that time, QBC was thought to be possible, and thus the QOT scheme was claimed to be implementable from scratch.

[3] Actually, we prove security for an entanglement-based version of BB84, which was first proposed by Ekert, and which implies security for the original BB84 scheme.

From the framework, we can then conclude that the raw key has high entropy from the adversary's point of view, and again privacy amplification finishes the job.

As for QOT, also QKD schemes initially came without security proofs, and proving QKD schemes rigorously secure turned out to be an extremely challenging and subtle task. Nowadays, though, the security of QKD schemes is better understood, and we know of various ways of proving, say, BB84 secure, ranging from Shor and Preskill's proof based on quantum error-correcting codes to Renner's approach using a quantum de Finetti theorem which allows to reduce security against general attacks to security against the much weaker class of so-called collective attacks. As such, our proof may safely be viewed as "yet another BB84 QKD proof." Nevertheless, it has some nice features: it provides an explicit and easy-to-compute expression for the security of the scheme (in contrast to most proofs in the literature which merely provide an asymptotic analysis), it does not require any "symmetrization of the qubits" (e.g., by applying a random permutation) from the protocol, and it is technically not very involved (e.g., compared to the proofs involving Renner's quantum de Finetti theorem). Furthermore, it gives immediately a *direct* security proof, rather than a reduction to the security against collective attacks.

3.1.3 Notation

Throughout this chapter, $\mathcal{A}$ denotes some fixed finite alphabet with $0 \in \mathcal{A}$. It is safe to think of $\mathcal{A}$ as $\{0,1\}$, but our claims also hold for larger alphabets. For a string $q = (q_1, \ldots, q_n) \in \mathcal{A}^n$ of arbitrary length $n \geq 0$, the *Hamming weight* of q is defined as the number of non-zero entries in q: $\mathrm{wt}(q) := \big|\{i \in [n] : q_i \neq 0\}\big|$. We also use the notion of the *relative* Hamming weight of q, defined as $\eta(q) := \mathrm{wt}(q)/n$. By convention, the relative Hamming weight of the empty string $\perp$ is set to $\eta(\perp) := 0$. For a string $q = (q_1, \ldots, q_n) \in \mathcal{A}^n$ and a subset $J \subset [n]$, we write $q_J := (q_i)_{i \in J}$ for the restriction of q to the positions $i \in J$.

3.2 Sampling from a Classical Population

As a warm-up, and in order to study some useful examples and introduce some convenient notation, we start with the classical sampling problem, which is rather well-understood.

3.2.1 Sampling Strategies

Let $q = (q_1, \ldots, q_n) \in \mathcal{A}^n$ be a string of given length n. We consider the problem of estimating the relative Hamming weight $\eta(q)$ by only looking at a substring q_t of

q, for a small subset $t \subset [n]$.[4] Actually, we are interested in the equivalent problem of estimating the relative Hamming weight $\eta(q_{\bar{t}})$ of the *remaining* string $q_{\bar{t}}$, where $\bar{t}$ is the complement $\bar{t} = [n] \setminus t$ of t.[5] A canonical way to do so would be to sample a uniformly random subset (say, of a certain small size) of positions, and compute the relative Hamming weight of the sample as estimate. Very generally, we allow any strategy that picks a subset $t \subset [n]$ according to some probability distribution and computes the estimate for $\eta(q_{\bar{t}})$ as some (possibly randomized) function of t and q_t, i.e., as $f(t, q_t, s)$ for a *seed* s that is sampled according to some probability distribution. This motivates the following formal definition.

Definition 3.1 (Sampling Strategy) A *sampling strategy* Ψ is defined by the triple (P_T, P_S, f), where P_T is a distribution over the subsets of $[n]$, P_S is a (independent) distribution over a finite set $\mathcal{S}$, and f is a function

$$f : \{(t, v) : t \subset [n], v \in \mathcal{A}^{|t|}\} \times \mathcal{S} \to \mathbb{R}.$$

We stress that a sampling strategy Ψ, as defined here, specifies how to choose the sample subset as well as how to compute the estimate from the sample (thus a more appropriate but lengthy name would be a "sample-and-estimate strategy").

Remark 3.2 By definition, the choice of the seed s is specified to be independent of t, i.e., $P_{TS} = P_T P_S$. Sometimes, however, it is convenient to allow s to depend on t. We can actually do so without contradicting Definition 3.1. Namely, to comply with the independence requirement, we would simply choose a (typically huge) "container" seed that contains a seed for every possible choice of t, each one chosen with the corresponding distribution, and it is then part of f's task, when given t, to select the seed that is actually needed from the container seed.[6]

A sampling strategy Ψ can obviously also be used to *test* if q (or actually $q_{\bar{t}}$) is close to the all-zero string $0 \cdots 0$: compute the estimate for $\eta(q_{\bar{t}})$ as dictated by Ψ, and *accept* if the estimate vanishes and else *reject*.

We briefly discuss five example sampling strategies. The examples should illustrate the generality of the definition, and some of the examples will be used later on; however, the reader is free to skip (some of) them. We start with the canonical example mentioned in the beginning.

[4] More generally, we may consider the problem of estimating the Hamming *distance* of q to some arbitrary *reference string* $q_\circ$; but this can obviously be done simply by estimating the Hamming weight of $q' = q - q_\circ$.

[5] The reason for this, as will become clear later, is that in our applications, the sampled positions within q will be *discarded*, and thus we will be interested merely in the remaining positions.

[6] Alternatively, we could simply drop the independence requirement in Definition 3.1; however, we feel it is conceptually easier to think of the seed as being independently chosen.

Example 3.3 (Random Sampling *Without* Replacement) In random sampling without replacement, k *distinct* indices $i_1, \ldots, i_k$ within $[n]$ are chosen uniformly at random, where k is some parameter, and the relative Hamming weight of $q_{\{i_1,\ldots,i_k\}}$ is used as estimate for $\eta(q_{\bar{t}})$. Formally, this sampling strategy is given by $\Psi = (P_T, P_S, f)$ where $P_T(t) = 1/\binom{n}{k}$ if $|t| = k$ and else $P_T(t) = 0$, $\mathcal{S} = \{\perp\}$ and thus $P_S(\perp) = 1$, and $f(t, q_t, \perp) = f(t, q_t) = \eta(q_t)$.

With the second example, we show that also sampling with replacement is captured by our definition.

Example 3.4 (Random Sampling *With* Replacement) In random sampling with replacement, k indices $i_1, \ldots, i_k$ are chosen independently uniformly at random within $[n]$, where k is some parameter, and the relative Hamming weight of the string $(q_{i_1}, \ldots, q_{i_k})$ is used as estimate for $\eta(q_{\bar{t}})$. Note that here i_ℓ may coincide with $i_{\ell'}$ for $\ell \neq \ell'$, in which case $(q_{i_1}, \ldots, q_{i_k})$ is not equal to $q_{\{i_1,\ldots,i_k\}}$. To make this fit into Definition 3.1, we set t to be $\{i_1, \ldots, i_k\}$, and we let $f(t, q_t, s)$ be given by $\eta(q_{j_1}, \ldots, q_{j_k})$, where $j_1, \ldots, j_k$ is determined by the seed s among all possibilities with $\{j_1, \ldots, j_k\} = t$. It is cumbersome and of no importance to us to determine the correct distributions P_T and P_S for t and s, respectively; it is sufficient to realize that random sampling with replacement *is* captured by Definition 3.1.

Next, we sample by picking a uniformly random subset (without restricting its size).

Example 3.5 (Uniformly Random Subset Sampling) The sample set t is chosen as a uniformly random subset of $[n]$, and the estimate is computed as the relative Hamming weight of the sample q_t. Formally, $P_T(t) = 1/2^n$ for any $t \subseteq [n]$, and $\mathcal{S} = \{\perp\}$ and $f(t, q_t, \perp) = f(t, q_t) = \eta(q_t)$.

As a fourth example, we consider a somewhat unnatural and in some sense non-optimal sampling strategy. This example, though, will be of use in our analysis of quantum oblivious transfer in Section 3.4.

Example 3.6 (Random Sampling Without Replacement, Using Part of the Sample) This example can be viewed as a composition of Example 3.3 and 3.5. Namely, t is chosen as a random subset of fixed size k, as in Example 3.3, so that $P_T(t) = 1/\binom{n}{k}$ for $t \subset [n]$ with $|t| = k$. But now, only part of the sample q_t is used to compute the estimate. Namely, the estimate is computed as

$$f(t, q_t, s) = \eta(q_s),$$

where the seed s is chosen as a uniformly random subset s of t; i.e., $P_S(s) = 1/2^t$ for any $s \subseteq t$. Recall from Remark 3.2 that the choice of s is allowed to depend on t. We would like to point out that when we use Example 3.6 in Section 3.4, it is useful that the restriction to the subset s is part of the evaluation of f, rather than part of the selection of the sample subset t.

In the fifth example we consider another somewhat unnatural sampling strategy, which though will be useful for the QKD proof in Section 3.5.

Example 3.7 (Pairwise One-Out-of-Two Sampling, Using Part of the Sample) For this example, it is convenient to consider the index set from which the subset t is chosen, to be of the form $[n] \times \{0,1\}$. Namely, we consider the string $q \in \mathcal{A}^{2n}$ to be indexed by *pairs* of indices, $q = (q_{ij})$, where $i \in [n]$ and $j \in \{0,1\}$; in other words, we consider q to consist of n pairs (q_{i0}, q_{i1}). The subset $t \subset [n] \times \{0,1\}$ is chosen as $t = \{(1, j_1), \ldots, (n, j_n)\}$ where every j_k is picked independently at random in $\{0,1\}$. In other words, t selects one element from each pair (q_{i0}, q_{i1}). Furthermore, the estimate for $\eta(q_{\bar{t}})$ is computed from q_t as $f(t, q_t, s) = \eta(q_s)$ where the seed s is a random subset $s \subset t$ of size k.

Example 3.8 (Pairwise *Biased* One-Out-of-Two Sampling, Using Part of the Sample) In this example we consider a similar situation as in Example 3.7, except that we now construct t by sampling every j_k according to the *Bernoulli* distribution $(p, 1-p)$. Consequently, we compute the estimate for $\eta(q_{\bar{t}})$ slightly differently, but we will make this clear in Section 3.2.3.

3.2.2 The Error Probability

After having introduced the general notion of a sampling strategy, we will define a measure that captures for a given sampling strategy how well it performs. More precisely—but still informally—this measure should be the probability that the difference between the estimate $f(t, q_t, s)$ and the real value, $\eta(q_{\bar{t}})$, is smaller than some given number. For the definition, it will be convenient to introduce the following notation. For a given sampling strategy $\Psi = (P_T, P_S, f)$, consider arbitrary but fixed choices for the subset $t \subset [n]$ and the seed $s \in \mathcal{S}$ with $P_T(t) > 0$ and $P_S(s) > 0$. Furthermore, fix an arbitrary $\delta > 0$. Define $B^\delta_{t,s}(\Psi) \subseteq \mathcal{A}^n$ as

$$B^\delta_{t,s}(\Psi) := \{b \in \mathcal{A}^n : |\eta(b_{\bar{t}}) - f(t, b_t, s)| < \delta\},$$

i.e., as the set of all strings q for which the estimate is δ-close to the real value, assuming that subset t and seed s have been used. To simplify notation, if Ψ is clear from the context, we simply write $B^\delta_{t,s}$ instead of $B^\delta_{t,s}(\Psi)$. By replacing the specific values t and s by the corresponding (independent) random variables T and S, with distributions P_T and P_S, respectively, we obtain the *random variable* $B^\delta_{T,S}$, whose range consists of subsets of $\mathcal{A}^n$. By means of this random variable, we now define the *error probability* of a sampling strategy as follows.

Definition 3.9 (Error Probability) The *classical error probability* of a sampling strategy $\Psi = (P_T, P_S, f)$ is defined as the following value, parameterized by $0 < \delta < 1$:

$$\varepsilon^{\delta}_{\text{class}}(\Psi) := \max_{q \in \mathcal{A}^n} \Pr\Big[q \notin B^{\delta}_{T,S}(\Psi)\Big] \,.$$

By definition of the error probability, it is guaranteed that for any string $q \in \mathcal{A}^n$, the estimated value is δ-close to the real value except with probability at most $\varepsilon^{\delta}_{\text{class}}(\Psi)$. When used as a sampling strategy to test closeness to the all-zero string, $\varepsilon^{\delta}_{\text{class}}(\Psi)$ determines the probability of accepting even though $q_{\bar{t}}$ is "not close" to the all-zero string, in the sense that its relative Hamming weight exceeds δ. Whenever Ψ is clear from the context, we will write $\varepsilon^{\delta}_{\text{class}}$ instead of $\varepsilon^{\delta}_{\text{class}}(\Psi)$.

3.2.3 Error Probabilities of the Example Sampling Strategies

We will now analyze the error probabilities for the sampling strategies considered in Examples 3.3 to 3.8 (excluding Example 3.4) and we show them all to be exponentially small by applying Hoeffding's inequality in a suitable way.

Ex. 3.3: Random Sampling *Without* Replacement

It follows immediately from Theorem 2.13 that the estimate is δ-close to the relative Hamming weight $\eta(q)$ of q except with probability at most $2\exp(-2\delta^2 k)$. However, we want to analyze closeness of the estimate to $\eta(q_{\bar{T}})$ (still treating T as a random variable). This can be derived easily as follows. We can write $\eta(q) = \alpha\eta(q_T) + (1-\alpha)\eta(q_{\bar{T}})$, where $\alpha := k/n$, and thus can see that

$$\eta(q_{\bar{T}}) - \eta(q_T) = \frac{1}{1-\alpha}\Big(\eta(q) - \alpha\eta(q_T)\Big) - \eta(q_T) = \frac{1}{1-\alpha}\Big(\eta(q) - \eta(q_T)\Big)$$

so that

$$\begin{aligned}\varepsilon^{\delta}_{\text{class}} &= \max_q \Pr\Big[q \notin B^{\delta}_{T,S}\Big] = \max_q \Pr\big[|\eta(q_{\bar{T}}) - \eta(q_T)| \geq \delta\big] \\ &= \max_q \Pr\big[|\eta(q) - \eta(q_T)| \geq (1-\alpha)\delta\big] \leq 2\exp(-2(1-\alpha)^2\delta^2 k). \qquad (3.1)\end{aligned}$$

Under assumption of $k \leq n/2$, we obtain a simple bound for the latter expression,

$$\varepsilon^{\delta}_{\text{class}} \leq 2\exp(-\tfrac{1}{2}\delta^2 k). \qquad (3.2)$$

We obtain the following bound if we use the bound from [Ser74]:

$$\begin{aligned}\varepsilon^{\delta}_{\text{class}} &= \max_q \Pr\big[|\eta(q) - \eta(q_T)| \geq (1-\alpha)\delta\big] \\ &\leq 2\exp(-\tfrac{2(1-\alpha)^2\delta^2 kn}{n-k+1}) = 2\exp(-\tfrac{2k(n-k)^2\delta^2}{n(n-k+1)}) \leq 2\exp(-\tfrac{\delta^2 kn}{n+2}).\end{aligned}$$

for $k \leq n/2$, because $-\frac{2k(n-k)^2\delta^2}{n(n-k+1)}$ is convex in k, and $-\frac{\delta^2 kn}{2+n}$ is linear in k and equality holds at $k = 0$ and $k = n/2$, hence it is a tight linear upper bound.

Ex. 3.4: Random Sampling *With* Replacement

Computing the error probability for Example 3.4 actually turns out to be tricky. Although Theorem 2.11 applies and guarantees that the estimate is likely to be close to $\eta(q)$, showing that the estimate is likely to be close to $\eta(q_{\bar{T}})$ seems to be nontrivial here. Since we make no further use of this example sampling strategy, we refrain from analyzing its error probability.

Ex. 3.5: Uniformly Random Subset Sampling

Note that for any fixed choice $k = |t|$, t is obtained as in random sampling without replacement. Because t is sampled uniformly at random, the expectation of k is given by $\mathbb{E}[k] = n/2$. Hence, by making use of Hoeffding's inequality (Theorem 2.13), we can say that for $0 < \beta < \frac{1}{2}$, $\Pr[|\frac{k}{n} - \frac{1}{2}| \geq \beta] \leq 2\exp(-2\beta^2 n)$.

Informally, the idea is to start off with an upper bound on $\varepsilon^\delta_{\text{class}}$ obtained for Example 3.3 (the case of sampling without replacement), and transform it into an upper bound that holds under the assumption that $k \in [(\frac{1}{2} - \beta)n, (\frac{1}{2} + \beta)n]$. Note that we cannot use the simple bound (3.2) from Example 3.3, because that result was obtained under the assumption that $k \leq n/2$, and here this assumption does not hold. Instead, we use bound (3.1) from Example 3.3,

$$\varepsilon^\delta_{\text{class}} \leq 2\exp\Big(-2(1-\tfrac{k}{n})^2\delta^2 k\Big) \tag{3.3}$$

which *does* hold for all $k \in \{0, \ldots, n\}$.

To get an upper bound for (3.3), we replace the occurrences of k by the appropriate boundary points of the interval $[(\frac{1}{2} - \beta)n, (\frac{1}{2} + \beta)n]$. I.e.,

$$2\exp\Big(-2\Big(1 - \frac{(\frac{1}{2}+\beta)n}{n}\Big)^2 \delta^2(\tfrac{1}{2} - \beta)n\Big) = 2\exp\big(-2n\delta^2(\tfrac{1}{2}-\beta)^3\big).$$

To compute $\varepsilon^\delta_{\text{class}}$, we use a union bound to combine the upper bound above, which holds under the assumption that k lies inside the previously defined interval, with the upper bound on the probability that k does *not* lie in this interval,

$$\varepsilon^\delta_{\text{class}} \leq 2\exp\Big(-2n\delta^2(\tfrac{1}{2}-\beta)^3\Big) + 2\exp(-2\beta^2 n).$$

Setting $\beta = \delta/4$ in the expression above yields $-n\delta^2(2-\delta)^3/32$ for the exponent of the first summand, and $-n\delta^2/8$ for the exponent of the second summand. Because $0 < \delta < 1$ (Definition 3.9), a suitable upper bound for both exponents is $-n\delta^2/32$.[7] This gives the following simpler bound,

$$\varepsilon^{\delta}_{\text{class}} \leq 4\exp(-n\delta^2/32).$$

Ex. 3.6: Random Sampling Without Replacement, Using Part of the Sample

From Example 3.3 we know that $\Pr\big[|\eta(q_{\bar{T}}) - \eta(q_T)| \geq \xi\big] \leq 2\exp(-\frac{1}{2}\xi^2 k)$, for $k < n/2$. Additionally, the selection of the seed s and the computation of $f(t, q_t, s)$ can be viewed as applying uniformly random subset sampling to q_t. Hence, it follows from Example 3.5 that $\max_q \Pr\big[|\eta(q_T) - \eta(q_S)| \geq \gamma\big] \leq 4\exp(-k\gamma^2/32)$. Setting $\delta = \xi + \gamma$, and using triangle inequality and union bound, we obtain

$$\begin{aligned}\varepsilon^{\delta}_{\text{class}} &= \max_q \Pr\big[|\eta(q_S) - \eta(q_{\bar{T}})| \geq \delta\big] \\ &\leq \min_{0<\xi<\delta} \Big[2\exp(-\tfrac{1}{2}\xi^2 k) + 4\exp(-k(\delta-\xi)^2/32)\Big] \\ &\leq 6\exp(-k\delta^2/50),\end{aligned}$$

where the last inequality follows from setting $\xi = \delta/5$ such that the two exponents coincide.

Ex. 3.7: Pairwise One-Out-of-Two Sampling, Using Part of the Sample

For $\mathcal{A} = \{0,1\}$, a bound on the error probability $\varepsilon^{\delta}_{\text{class}}$ is obtained as follows. Let q be arbitrary, indexed as discussed earlier. First, we show that $\eta(q_{\bar{T}})$ is likely to be close to $\eta(q_T)$. For this, consider the pairs (q_{i0}, q_{i1}) for which $q_{i0} \neq q_{i1}$. Let there be ℓ such pairs (where obviously $\ell \leq n$.) We denote the restrictions of q_T and $q_{\bar{T}}$ to these indices i with $q_{i0} \neq q_{i1}$ by $\tilde{q}_T$ and $\tilde{q}_{\bar{T}}$, respectively. It is easy to see that $\mathrm{wt}(\tilde{q}_T) + \mathrm{wt}(\tilde{q}_{\bar{T}}) = \ell$. It follows that for any $\epsilon > 0$ we have

$$\begin{aligned}\Pr\big[|\eta(q_{\bar{T}}) - \eta(q_T)| \geq \epsilon\big] &= \Pr\big[|\mathrm{wt}(q_T) - \mathrm{wt}(q_{\bar{T}})| \geq n\epsilon\big] \\ = \Pr\big[|\mathrm{wt}(\tilde{q}_T) - \mathrm{wt}(\tilde{q}_{\bar{T}})| \geq n\epsilon\big] &= \Pr\big[|2\mathrm{wt}(\tilde{q}_T) - \ell| \geq n\epsilon\big] \\ \leq 2\exp\Big(-2\left(\tfrac{n\epsilon}{2\ell}\right)^2 \ell\Big) &= 2\exp\Big(-\tfrac{n\epsilon^2}{2}\cdot\tfrac{n}{\ell}\Big) \leq 2\exp\Big(-\tfrac{1}{2}\epsilon^2 n\Big),\end{aligned}$$

[7] Note that our goal is to find a short and simple expression, rather than finding the tightest bound.

where the third equality follows from replacing $\mathrm{wt}(\widetilde{q}_{\bar{T}})$ by $\ell - \mathrm{wt}(\widetilde{q}_T)$, and the first inequality follows from Hoeffding's inequality (as each entry of $\mathrm{wt}(\widetilde{q}_T)$ is 0 with independent probability $\frac{1}{2}$).

Furthermore, for any $\gamma > 0$ we have the following relation involving q_S:

$$\Pr\big[|\eta(q_T) - \eta(q_S)| \geq \gamma\big] \leq 2\exp\left(-2k\gamma^2\right),$$

which follows from directly applying Hoeffding's inequality. Applying the union bound and letting $\delta = \epsilon + \gamma$, we obtain

$$\begin{aligned}
\varepsilon^{\delta}_{\text{class}} &= \max_q \Pr\big[|\eta(q_{\bar{T}}) - \eta(q_S)| \geq \delta\big] \\
&< 2 \min_{0<\epsilon<\delta} \left[\exp\left(-\tfrac{1}{2}\epsilon^2 n\right) + \exp\left(-2k(\delta-\epsilon)^2\right)\right] \\
&\leq 4\exp\left(-\tfrac{2kn\delta^2}{(2\sqrt{k}+\sqrt{n})^2}\right) \leq 4\exp\left(-\tfrac{1}{3}\delta^2 k\right) \ ,
\end{aligned}$$

where the last line follows from choosing ϵ such that the two exponents coincide, and from doing some simplifications while assuming $k \leq n/2$.

Ex. 3.8: Pairwise *Biased* One-Out-of-Two Sampling, Using Part of the Sample

It will be convenient to define the index set t as the union of two subsets, $t_0 \subset [n] \times \{0\}$ and $t_1 \subset [n] \times \{1\}$. Note that the complements of these subsets should now be understood as $\bar{t}_0 = ([n] \times \{0\}) \setminus t_0$ and $\bar{t}_1 = ([n] \times \{1\}) \setminus t_1$. Let t_0 and t_1 be constructed as follows. We first sample a set $\widetilde{t} \subset [n]$; for each element of $[n]$, we include it in $\widetilde{t}$ with probability p. Then, $t_0 := \widetilde{t} \times \{0\}$ and $t_1 := ([n] \setminus \widetilde{t}) \times \{1\}$. Like t, the seed s is also defined as the union of two randomly chosen sets, $s = s_0 \cup s_1$, where $s_0 \subset t_0$ and $s_1 \subset t_1$.[8] These sets have fixed size; for a parameter $k \in \mathbb{N}$, $|s_0| = \frac{k}{2}$ and $|s_1| = \frac{k}{2}$. Now, the estimate for $\eta(q_{\bar{t}})$ is computed as $f(t, q_t, s) = \frac{1}{n}(|\bar{t}_0|\, \eta(q_{s_0}) + |\bar{t}_1|\, \eta(q_{s_1}))$.

We need to show that $\eta(q_{\bar{T}})$ is likely to be close to $\eta(q_S)$. Because we compute an estimate for $\eta(q_{\bar{T}})$ as a function of $\eta(q_{S_0})$ and $\eta(q_{S_1})$, we will first show that (with high probability) $\eta(q_{T_0}) \approx \eta(q_{S_0})$ and $\eta(q_{T_1}) \approx \eta(q_{S_1})$. Then, we argue that $\eta(q_{\bar{T}_0}) \approx \eta(q_{T_0})$ and $\eta(q_{\bar{T}_1}) \approx \eta(q_{T_1})$, from which we can also conclude (using the union bound) that $\eta(q_{\bar{T}_0}) \approx \eta(q_{S_0})$ and $\eta(q_{\bar{T}_1}) \approx \eta(q_{S_1})$. Finally, we apply the union bound again and combine the two bounds to obtain an upper bound for $\Pr\big[|\eta(q_{\bar{T}}) - \frac{1}{n}(|\bar{T}_0|\, \eta(q_{S_0}) + |\bar{T}_1|\, \eta(q_{S_1}))| \geq \delta\big]$.

[8] Again, Remark 3.2 applies.

The first step in the proof follows directly from Hoeffding's inequality,

$$\Pr\left[|\eta(q_{T_0}) - \eta(q_{S_0})| \geq \gamma\right] \leq 2\exp\left(-2|S_0|\gamma^2\right) = 2\exp(-k\gamma^2), \quad \forall\gamma > 0.$$

Trivially, this bound also applies to the relation between $\eta(q_{T_1})$ and $\eta(q_{S_1})$, if we substitute appropriately. The second step, showing that $\eta(\bar{T}_0)$ (respectively $\eta(\bar{T}_1)$) is likely to be close to $\eta(T_0)$ (resp. $\eta(T_1)$), is slightly more involved. Namely, although the sum of the sizes of T_0 and T_1 is constant (to be precise, $|T_0| + |T_1| = n$), their individual sizes are random. In Example 3.5 we have already encountered a similar, though not identical, situation: Example 3.5 considers uniformly random one-out-of-two sampling whereas here we analyze one-out-of-two sampling according to a Bernoulli $(p, 1-p)$ distribution. Nonetheless, it is straightforward to generalize the error-probability analysis for Example 3.5 to this (more general) case.

Let $X := |T_0|$. The expectation of X is given by $\mathbb{E}[X] = np$. Let $\mathcal{E}$ be the event that $X \in [(p-\beta)n, (p+\beta)n]$, for $\beta > 0$. From Hoeffding's inequality, we known that $\Pr[\bar{\mathcal{E}}] = \Pr[|\frac{X}{n} - p| \geq \beta] \leq 2\exp(-2\beta^2 n)$. Like in Section 3.2.3, we find an upper bound that holds conditioned on the event $\mathcal{E}$, by substituting the boundary points of the interval used to define $\mathcal{E}$ in (3.3),

$$\begin{aligned}\Pr\left[|\eta(q_{T_0}) - \eta(q_{\bar{T}_0})| \geq \delta \mid \mathcal{E}\right] &\leq -2(p-\beta)n\left(1 - \frac{(p+\beta)n}{n}\right)^2\\ &= 2\exp(-2n\delta^2(1-p-\beta)^2(p-\beta)).\end{aligned}$$

Next, we apply the union bound to show that for $0 < \epsilon < \gamma$

$$\begin{aligned}&\Pr\left[\left|\eta(q_{\bar{T}_0}) - \eta(q_{S_0})\right| \geq \gamma \mid \mathcal{E}\right]\\ &\quad\leq 2\exp(-2n\epsilon^2(1-p-\beta)^2(p-\beta)) + 2\exp\left(-k(\gamma-\epsilon)^2\right).\end{aligned}$$

By substituting p by $1-p$ in the expression above, we also obtain

$$\begin{aligned}&\Pr\left[\left|\eta(q_{\bar{T}_1}) - \eta(q_{S_1})\right| \geq \gamma \mid \mathcal{E}\right]\\ &\quad\leq 2\exp(-2n\epsilon^2(p-\beta)^2(1-p-\beta)) + 2\exp\left(-k(\gamma-\epsilon)^2\right).\end{aligned}$$

Finally, we combine the two bounds and we get rid of the conditioning on $\mathcal{E}$ by adding $\Pr[\bar{\mathcal{E}}]$. For any $\delta > 0$ and $0 < \epsilon < \delta$, we may write

$$\begin{aligned}
&\varepsilon^{\delta}_{\text{class}} = \max_q \Pr\big[|\eta(q_{\bar{T}}) - \frac{1}{n}(|\bar{T}_0|\,\eta(q_{S_0}) + |\bar{T}_1|\,\eta(q_{S_1}))| \geq \delta\big]\\
&= \max_q \Pr\big[|\mathrm{wt}(q_{\bar{T}}) - |\bar{T}_0|\,\eta(q_{S_0}) + |\bar{T}_1|\,\eta(q_{S_1})| \geq n\delta\big]\\
&= \max_q \Pr\big[|\mathrm{wt}(q_{\bar{T}}) - |\bar{T}_0|\,\eta(q_{S_0}) + |\bar{T}_1|\,\eta(q_{S_1})| \geq (|\bar{T}_0|\delta + |\bar{T}_1|\delta)\big]\\
&\leq \max_q \Pr\Big[\Big|\eta(q_{\bar{T}_0}) - \eta(q_{S_0})\Big| \geq \delta\Big] + \Pr\Big[\Big|\eta(q_{\bar{T}_1}) - \eta(q_{S_1})\Big| \geq \delta\Big]\\
&\leq 2\exp(-2n\epsilon^2(1-p-\beta)^2(p-\beta)) + 2\exp(-2n\epsilon^2(p-\beta)^2(1-p-\beta))\\
&\quad + 4\exp\Big(-k(\delta-\epsilon)^2\Big) + 2\exp(-2\beta^2 n).
\end{aligned}$$

3.3 Sampling from a Quantum Population

In this section, we apply a sampling strategy to a *quantum population* and study its behavior. More specifically, let $A = A_1 \cdots A_n$ be an n-partite quantum system, where the state space of each system A_i equals $\mathcal{H}_{A_i} = \mathbb{C}^d$ with $d = |\mathcal{A}|$, and let $\{|a\rangle\}_{a\in\mathcal{A}}$ be a fixed orthonormal basis of $\mathbb{C}^d$. We allow A to be entangled with some additional system E with arbitrary finite-dimensional state space $\mathcal{H}_E$. We may assume the joint state of AE to be pure, and as such be given by a state vector $|\varphi_{AE}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_E$; if not, then it can be purified by increasing the dimension of $\mathcal{H}_E$.

Similar to the classical sampling problem of testing closeness to the all-zero string, we can consider here the problem of testing if the state of A is close to the all-zero *reference state* $|\varphi^{\circ}_A\rangle = |0\rangle \cdots |0\rangle$ by looking at, which here means *measuring*, only a few of the subsystems of A. More generally, we will be interested in the sampling problem of estimating the "Hamming weight of the state of A," although it is not clear at the moment what this should mean. Actually, like in the classical case, we are interested in testing closeness to the all-zero state, respectively estimating the Hamming weight, of the *remaining subsystems* of A.

It is obvious that a sampling strategy $\Psi = (P_T, P_S, f)$ can be applied in a straightforward way to the setting at hand: sample t according to P_T, measure the subsystems A_i with $i \in t$ in basis $\{|a\rangle\}_{a\in\mathcal{A}}$ to observe $q_t \in \mathcal{A}^{|t|}$, and compute the estimate as $f(t, q_t, s)$ for s chosen according to P_S (respectively, for testing closeness to the all-zero state, accept or reject depending on the value of the estimate). However, it is a-priori *not* clear, how to interpret the outcome. Measuring a random subset of the subsystems of A and observing 0 all the time indeed seems to suggest that the original state of A, and thus the remaining subsystems, must be in some sense close to the all-zero state; but what is the right way to formalize this? In the case of a general sampling strategy for estimating the (relative) Hamming weight, what does the estimate actually estimate? And, do all strategies that perform well in the classical setting also perform well in the quantum setting?

We will give a rigorous analysis of sampling strategies when applied to an n-partite quantum system A, which will in particular answer the questions raised above. Later in the chapter, we demonstrate the usefulness of our analysis of sampling strategies for studying and analyzing quantum-cryptographic schemes.

3.3.1 Analyzing Sampling Strategies in the Quantum Setting

We start by suggesting the property on the remaining subsystems of A that one should expect to be able to conclude from the outcome of a sampling strategy. A somewhat natural approach is as follows.

Definition 3.10 For system AE, and similarly for any subsystem of A, we say that the state $|\varphi_{AE}\rangle$ of AE has *relative Hamming weight β within A* if it is of the form $|\varphi_{AE}\rangle = |b\rangle|\varphi_E\rangle$ with $b \in \mathcal{A}^n$ and $\eta(b) = \beta$.

Now, given the outcome $f(t, q_t, s)$ of a sampling strategy when applied to A, we want to be able to conclude that, up to a small error, the state of the remaining subsystem $A_{\bar{t}}E$ is a *superposition* of states with relative Hamming weight close to $f(t, q_t, s)$ within $A_{\bar{t}}$. To analyze this, we extend some of the notions introduced in the classical setting. Recall the definition of $B^{\delta}_{t,s}$, consisting of all strings $b \in \mathcal{A}^n$ with $|\eta(b_{\bar{t}}) - f(t, b_t, s)| < \delta$. By slightly abusing notation, we extend this notion to the quantum setting and write

$$\mathrm{span}(B^{\delta}_{t,s}) := \mathrm{span}(\{|b\rangle : b \in B^{\delta}_{t,s}\}) = \mathrm{span}(\{|b\rangle : |\eta(b_{\bar{t}}) - f(t, b_t, s)| < \delta\}) \, .$$

Note that if the state $|\varphi_{AE}\rangle$ of AE happens to be in $\mathrm{span}(B^{\delta}_{t,s}) \otimes \mathcal{H}_E$ for some t and s, and if exactly these t and s are chosen when applying the sampling strategy to A, then *with certainty* the state of $A_{\bar{t}}E$ (after the measurement) is in a superposition of states with relative Hamming weight δ-close to $f(t, q_t, s)$ within $A_{\bar{t}}$, regardless of the measurement outcome q_t.

Next, we want to extend the notion of error probability (Definition 3.9) to the quantum setting. The following approach turns out to be fruitful. We consider the *hybrid* system $TSAE$, consisting of the classical random variables T and S with distribution $P_{TS} = P_T P_S$, describing the choices of t and s, respectively, and of the actual quantum systems A and E. The state of $TSAE$ is given by

$$\rho_{TSAE} = \sum_{t,s} P_{TS}(t,s)|t,s\rangle\langle t,s| \otimes |\varphi_{AE}\rangle\langle\varphi_{AE}|\,.$$

Note that TS is independent of AE: $\rho_{TSAE} = \rho_{TS} \otimes \rho_{AE}$; indeed, in a sampling strategy t and s are chosen independently of the state of AE. We compare this *real* state of $TSAE$ with an *ideal* state which is of the form

$$\widetilde{\rho}_{TSAE} = \sum_{t,s} P_{TS}(t,s)|t,s\rangle\langle t,s| \otimes |\widetilde{\varphi}^{ts}_{AE}\rangle\langle\widetilde{\varphi}^{ts}_{AE}| \tag{3.4}$$

where $|\widetilde{\varphi}^{ts}_{AE}\rangle \in \mathrm{span}(B^{\delta}_{t,s}) \otimes \mathcal{H}_E$ for all (t,s) and for some given $\delta > 0$. Thus, T and S have the same distribution as in the real state, but here we allow AE to depend on T and S, and for each particular choice t and s for T and S, respectively, we require the state of AE to be in $\mathrm{span}(B^{\delta}_{t,s}) \otimes \mathcal{H}_E$. Hence, in an "ideal world" where the state of the hybrid system $TSAE$ is given by $\widetilde{\rho}_{TSAE}$, it holds *with certainty* that the state $|\psi_{A_{\bar{t}}E}\rangle$ of $A_{\bar{t}}E$, after having measured A_t and having observed q_t, is in a superposition of states with relative Hamming weight δ-close to $\beta := f(t, q_t, s)$ within $A_{\bar{t}}$. We now define the *quantum error probability* of a sampling strategy by looking at how far away the closest ideal state $\widetilde{\rho}_{TSAE}$ is from the real state ρ_{TSAE}.

Definition 3.11 (Quantum Error Probability) The *quantum error probability* of a sampling strategy $\Psi = (P_T, P_S, f)$ is defined as the following value, parameterized by $0 < \delta < 1$:

$$\varepsilon^{\delta}_{\text{quant}}(\Psi) = \max_{\mathcal{H}_E}\, \max_{|\varphi_{AE}\rangle}\, \min_{\widetilde{\rho}_{TSAE}}\, \tfrac{1}{2}\|\rho_{TSAE} - \widetilde{\rho}_{TSAE}\|_1,$$

where the first $\max$ is over all finite-dimensional state spaces $\mathcal{H}_E$, the second $\max$ is over all state vectors $|\varphi_{AE}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_E$, and the $\min$ is over all ideal states $\widetilde{\rho}_{TSAE}$ as in (3.4).[9]

As with $B^{\delta}_{t,s}$ and $\varepsilon^{\delta}_{\text{class}}$, we simply write $\varepsilon^{\delta}_{\text{quant}}$ when Ψ is clear from the context. We stress the meaningfulness of the definition: it guarantees that on average over the choice of t and s, the state of $A_{\bar{t}}E$ is $\varepsilon^{\delta}_{\text{quant}}$-close to a superposition of states with Hamming weight δ-close to $f(t, q_t, s)$ within $A_{\bar{t}}$, and as such it *behaves* like a superposition of such states, except with probability $\varepsilon^{\delta}_{\text{quant}}$. We will argue below and demonstrate in the subsequent sections that being close to a superposition of states with given approximate (relative) Hamming weight has some useful consequences.

[9] It is not too hard to see, in particular after having gained some more insight via the proof of Theorem 3.13 below, that the minimum and maxima exist.

Remark 3.12 Similarly to footnote 4, also here the results of the section immediately generalize from the all-zero reference state $|0\rangle\cdots|0\rangle$ to an arbitrary reference state $|\varphi^\circ_A\rangle$ of the form $|\varphi^\circ_A\rangle = U_1|0\rangle\otimes\cdots\otimes U_n|0\rangle$ for unitary operators U_i acting on $\mathbb{C}^d$. Indeed, the generalization follows simply by a suitable change of basis, defined by the U_i's. Or, in the special case where $\mathcal{A} = \{0,1\}$ and

$$|\varphi^\circ_A\rangle = H^{\hat\theta}|\hat x\rangle = H^{\hat\theta_1}|\hat x_1\rangle \otimes\cdots\otimes H^{\hat\theta_n}|\hat x_n\rangle$$

for a fixed reference basis $\hat\theta \in \{0,1\}^n$ and a fixed reference string $\hat x \in \{0,1\}^n$, we can, alternatively, replace in the definitions and results the computational by the Hadamard basis whenever $\hat\theta_i = 1$, and speak of the (relative) Hamming distance to $\hat x$ rather than of the (relative) Hamming weight.

3.3.2 The Quantum vs. the Classical Error Probability

It remains to discuss how difficult it is to actually *compute* the quantum error probability for given sampling strategies, and how the *quantum* error probability $\varepsilon^\delta_{\rm quant}$ relates to the corresponding *classical* error probability $\varepsilon^\delta_{\rm class}$. To this end, we show the following simple relationship between $\varepsilon^\delta_{\rm quant}$ and $\varepsilon^\delta_{\rm class}$.

Theorem 3.13 *For any sampling strategy Ψ and for any $\delta > 0$:*

$$\varepsilon^\delta_{\rm quant}(\Psi) \le \sqrt{\varepsilon^\delta_{\rm class}(\Psi)}.$$

As a consequence of this theorem, it suffices to analyze a sampling strategy in the classical setting, which is much easier, in order to understand how it behaves in the quantum setting. In particular, sampling strategies that are known to behave well in the classical setting, like examples 3.3 to 3.7, are also automatically guaranteed to behave well in the quantum setting. We will use this in the application sections.

Our bound on $\varepsilon^\delta_{\rm quant}$ is in general tight in the following sense.

Proposition 3.14 *There exist natural sampling strategies for which equality holds in Theorem 3.13.*

Later in this section, we will characterize this class of sampling strategies (which turns out to contain Example 3.3 and Example 3.7) and prove the proposition.

Proof of Theorem 3.13. We need to show that for any $|\varphi_{AE}\rangle \in \mathcal{H}_A\otimes\mathcal{H}_E$, with arbitrary $\mathcal{H}_E$, there exists a suitable ideal state $\widetilde\rho_{TSAE}$ such that $\frac{1}{2}\|\rho_{TSAE}-\widetilde\rho_{TSAE}\|_1 \le$

$\sqrt{\varepsilon^\delta_{\text{class}}}$. We construct $\tilde{\rho}_{TSAE}$ as in (3.4), where the $|\tilde{\varphi}^{ts}_{AE}\rangle$'s are defined by the following decomposition.

$$|\varphi_{AE}\rangle = \langle\tilde{\varphi}^{ts}_{AE}|\varphi_{AE}\rangle|\tilde{\varphi}^{ts}_{AE}\rangle + \langle\tilde{\varphi}^{ts\perp}_{AE}|\varphi_{AE}\rangle|\tilde{\varphi}^{ts\perp}_{AE}\rangle,$$

with $|\tilde{\varphi}^{ts}_{AE}\rangle \in \text{span}(B^\delta_{t,s})\otimes\mathcal{H}_E$, $|\tilde{\varphi}^{ts\perp}_{AE}\rangle \in \text{span}(B^\delta_{t,s})^\perp\otimes\mathcal{H}_E$ and $|\langle\tilde{\varphi}^{ts}_{AE}|\varphi_{AE}\rangle|^2 + |\langle\tilde{\varphi}^{ts\perp}_{AE}|\varphi_{AE}\rangle|^2 = 1$. In other words, $|\tilde{\varphi}^{ts}_{AE}\rangle$ is obtained as the re-normalized projection of $|\varphi_{AE}\rangle$ into $\text{span}(B^\delta_{t,s})\otimes\mathcal{H}_E$. Note that $|\langle\tilde{\varphi}^{ts\perp}_{AE}|\varphi_{AE}\rangle|^2$ equals the probability $\Pr[Q \notin B^\delta_{t,s}]$, where the random variable Q is obtained by measuring subsystem A of $|\varphi_{AE}\rangle$ in basis $\{|a\rangle\}^{\otimes n}_{a\in\mathcal{A}}$. Furthermore,

$$\begin{aligned}\sum_{t,s} P_{TS}(t,s)\,|\langle\tilde{\varphi}^{ts\perp}_{AE}|\varphi_{AE}\rangle|^2 &= \sum_{t,s} P_{TS}(t,s)\,\Pr[Q\notin B^\delta_{t,s}] = \Pr[Q\notin B^\delta_{T,S}]\\ &= \sum_q P_Q(q)\,\Pr[q\notin B^\delta_{T,S}],\end{aligned}$$

where by definition of $\varepsilon^\delta_{\text{class}}$, the latter is bounded above by $\varepsilon^\delta_{\text{class}}$. From elementary properties of the trace distance, and using Jensen's inequality, we can now conclude that

$$\begin{aligned}\tfrac{1}{2}\|\rho_{TSAE} - \tilde{\rho}_{TSAE}\|_1 &= \sum_{t,s} P_{TS}(t,s)\tfrac{1}{2}\Big\||\varphi_{AE}\rangle\langle\varphi_{AE}| - |\tilde{\varphi}^{ts}_{AE}\rangle\langle\tilde{\varphi}^{ts}_{AE}|\Big\|_1\\ &= \sum_{t,s} P_{TS}(t,s)\sqrt{1-|\langle\tilde{\varphi}^{ts}_{AE}|\varphi_{AE}\rangle|^2} = \sum_{t,s} P_{TS}(t,s)|\langle\tilde{\varphi}^{ts\perp}_{AE}|\varphi_{AE}\rangle|\\ &\le \sqrt{\sum_{t,s} P_{TS}(t,s)|\langle\tilde{\varphi}^{ts\perp}_{AE}|\varphi_{AE}\rangle|^2} \le \sqrt{\varepsilon^\delta_{\text{class}}},\end{aligned}$$

which was to be shown. □

As a side remark, we point out that the particular ideal state $\tilde{\rho}_{TSAE}$ constructed in the proof minimizes the distance to ρ_{TSAE}; this follows from the so-called Hilbert projection theorem.

The Tightness of Theorem 3.13

We show here that in general the inequality from Theorem 3.13 is tight. Specifically, we specify a natural class of sampling strategies for which Theorem 3.13 is an equality. Informally, this class consists of sampling strategies that behave in exactly the same way if the randomized choices T and S are replaced by *fixed* choices $t_\circ$ and $s_\circ$, and instead the coordinates of q are shuffled by means of a uniformly random permutation (chosen from a subgroup of all permutations). The formal definition is given below, but let us point out already here that Example 3.3 as well as the QKD sampling strategy discussed in Example 3.7 belong to this class. Indeed, for Example 3.3, instead of choosing a random subset T of size k one can equivalently choose a fixed subset and randomly permute the positions of q. And, similarly for Example 3.7, instead of choosing left or right from each pair (q_{i0}, q_{i1}) at random and then choosing a random subset of size k of the selected q_{ij}'s, one can equivalently fix these choices and swap each pair (q_{i0}, q_{i1}) with probability $\frac{1}{2}$ and apply a random permutation to the first index.

Let S_n denote the symmetric group of degree n, i.e., the group of permutations on $[n]$. For any $\pi \in S_n$ and $q = (q_1, \dots, q_n) \in \mathcal{A}^n$, we write πq to express that π permutes the *positions* of the elements of q, i.e., $\pi q = (q_{\pi^{-1}(1)}, \dots, q_{\pi^{-1}(n)})$. If $\mathcal{V}$ is a set of strings $q \in \mathcal{A}^n$, then $\pi\mathcal{V}$ means that the permutation π acts element-wise on $\mathcal{V}$.

Definition 3.15 (G-Symmetry of a Sampling Strategy) Let Ψ be a sampling strategy, let G be a subgroup of S_n, where n is the size of the population to which Ψ is applied, and let Π be a random permutation, uniformly distributed over G. We call Ψ *G-symmetric*, if there exist $t_\circ \subset [n]$ and $s_\circ \in \mathcal{S}$ such that

$$(\eta(q_{\bar{T}}), f(T, q_T, S)) \sim (\eta((\Pi q)_{\bar{t}_\circ}), f(t_\circ, (\Pi q)_{t_\circ}, s_\circ))$$

where "$\sim$" means that the pairs have the same probability distribution.

A direct consequence of this definition is the following relation, which we will apply later in this section.

$$\begin{aligned} B^\delta_{T,S} &= \{q \in \{0,1\}^n : |\eta(q_{\bar{T}}) - f(T, q_T, S)| < \delta\} \\ &\sim \{q \in \{0,1\}^n : |\eta((\Pi q)_{\bar{t}_\circ}) - f(t_\circ, (\Pi q)_{t_\circ}, s_\circ)| < \delta\} = \Pi^{-1} B^\delta_{t_\circ, s_\circ}. \end{aligned}$$

We can now rephrase Proposition 3.14 and prove it.

Proposition 3.16 *For any G-*symmetric *sampling strategy Ψ^{sym}_G and any $\delta > 0$:*

$$\varepsilon^\delta_{\text{quant}}(\Psi^{sym}_G) = \sqrt{\varepsilon^\delta_{\text{class}}(\Psi^{sym}_G)}$$

Proof. We need to show that there exists a system E and a state $|\varphi_{AE}\rangle$ such that $\frac{1}{2}\|\rho_{TSAE} - \tilde{\rho}_{TSAE}\|_1^2 = \varepsilon^\delta_{\text{class}}$ for $\tilde{\rho}_{TSAE}$ that minimizes the left hand side. As pointed out after the proof of Theorem 3.13, the particular construction of $\tilde{\rho}_{TSAE}$ used in the proof of Theorem 3.13 does minimize $\frac{1}{2}\|\rho_{TSAE} - \tilde{\rho}_{TSAE}\|_1$. Hence, it suffices to show that there exists a system E and a state $|\varphi_{AE}\rangle$ (that depends on G) such that

$$\tfrac{1}{2}\|\rho_{TSAE} - \widetilde{\rho}_{TSAE}\|_1^2 \overset{(3.5)}{=} \left[\sum_{t,s} P_{TS}(t,s)|\langle\varphi_{AE}|\widetilde{\varphi}_{AE}^{ts\perp}\rangle|\right]^2$$
$$\overset{(3.6)}{=} \sum_{t,s} P_{TS}(t,s)|\langle\varphi_{AE}|\widetilde{\varphi}_{AE}^{ts\perp}\rangle|^2 \overset{(3.7)}{=} \varepsilon_{\text{class}}^{\delta}.$$

where $\widetilde{\rho}_{TSAE}$ and $|\widetilde{\varphi}_{AE}^{ts\perp}\rangle$ are constructed as in the proof of Theorem 3.13. The derivation of equality (3.5) can be found in the proof of Theorem 3.13. The outline of the remaining part of the proof is as follows; we first present a candidate for $|\varphi_{AE}\rangle$ and then we show that equalities (3.6) and (3.7) do indeed hold for this state.

We choose E to be empty. Furthermore, we define

$$|\varphi_{AE}\rangle := \frac{1}{\sqrt{|G|}} \sum_{\pi\in G} |\pi q^*\rangle.$$

where q^* is such that $\Pr[q^* \notin B_{T,S}^{\delta}] = \varepsilon_{\text{class}}^{\delta}$. It follows from the projection construction for $\widetilde{\rho}_{TSAE}$ that

$$|\widetilde{\varphi}_{AE}^{ts\perp}\rangle = \frac{1}{\sqrt{|H_{t,s}|}} \sum_{\pi\in H_{t,s}} |\pi q^*\rangle,$$

where $H_{t,s} \subseteq G$, i.e., $H_{t,s} := \{\pi \in G : \pi q^* \notin B_{t,s}^{\delta}\}$.

To prove equality (3.6), we need to show that the inner product $|\langle\varphi_{AE}|\widetilde{\varphi}_{AE}^{ts\perp}\rangle|$ is independent of t and s. Because $|\varphi_{AE}\rangle$ is a uniform superposition over permutations of q^* and $|\widetilde{\varphi}_{AE}^{ts\perp}\rangle$ is a renormalized projection of $|\varphi_{AE}\rangle$, we can easily compute this inner product,

$$|\langle\varphi_{AE}|\widetilde{\varphi}_{AE}^{ts\perp}\rangle| = |H_{t,s}|/\sqrt{|G|\cdot|H_{t,s}|} = \sqrt{|H_{t,s}|/|G|}.$$

It suffices to show that $|H_{t,s}|$ is independent of (t,s). It follows from the G-symmetry that there exists a π such that $B_{t,s}^{\delta} = \pi B_{t_\circ,s_\circ}^{\delta}$. Furthermore, let Π be a random permutation, uniformly distributed over G. By definition of $H_{t,s}$ and because Π is *uniformly* distributed over G, we may write

$$\begin{aligned}|H_{t,s}| &= |G|\cdot\Pr[\Pi\, q^* \notin B_{t,s}^{\delta}]\\ &= |G|\cdot\Pr[q^* \notin \Pi^{-1}\pi B_{t_\circ,s_\circ}^{\delta}] = |G|\cdot\Pr[q^* \notin \Pi^{-1} B_{t_\circ,s_\circ}^{\delta}],\end{aligned} \tag{3.8}$$

where the last expression is clearly independent of (t,s).

Now, let us focus on equality (3.7). We derived in the proof of Theorem 3.13 that $\sum_{t,s} P_{TS}(t,s)\,|\langle\varphi_{AE}|\tilde{\varphi}^{ts\perp}_{AE}\rangle|^2 = \sum_q P_Q(q)\ \Pr[q\,{\notin}\,B^\delta_{T,S}]$, where the random variable Q is obtained by measuring subsystem A of $|\varphi_{AE}\rangle$. By definition of $|\varphi_{AE}\rangle$, $P_Q(q) > 0$ only for q of the form πq^* for some $\pi \in G$. Hence, to prove equality (3.7), we have to show that for any $\pi \in G$, $\Pr[\pi q^* \notin B^\delta_{T,S}] = \varepsilon^\delta_{\text{class}}$. This follows directly from the G-symmetry,

$$\begin{aligned}\Pr[\pi q^* \,{\notin}\, B^\delta_{T,S}] &= \Pr[\pi q^* \,{\notin}\, \Pi^{-1} B^\delta_{t_\circ,s_\circ}] = \Pr[q^* \,{\notin}\, \pi^{-1}\Pi^{-1} B^\delta_{t_\circ,s_\circ}]\\ &= \Pr[q^* \,{\notin}\, \Pi^{-1} B^\delta_{t_\circ,s_\circ}] = \Pr[q^* \,{\notin}\, B^\delta_{T,S}]. \qquad (3.9)\end{aligned}$$

Finally, note that (3.8) and (3.9) rely on the group structure of G. □

3.3.3 Superpositions with a Small Number of Terms

We give here an argument why being close to a superposition of states with a given approximate Hamming weight may be a useful property in the analyses of quantum-cryptographic schemes. For simplicity, and since this will be the case in our applications, we now restrict to the binary case where $\mathcal{A} = \{0,1\}$. Our argument is based on the following lemma, which follows from Lemma 3.1.13 in [Ren05]; for completeness, we give a direct proof of Lemma 3.17 below as well. Informally, the lemma states that measuring (part of) a *superposition* of a small number of orthogonal states produces a similar amount of uncertainty as when measuring the *mixture* of these orthogonal states.

Lemma 3.17 *Let A and E be arbitrary quantum systems, let $\{|i\rangle\}_{i\in I}$ and $\{|w\rangle\}_{w\in\mathcal{W}}$ be orthonormal bases of $\mathcal{H}_A$, and let $|\varphi_{AE}\rangle$ and ρ^{mix}_{AE} be of the form*

$$|\varphi_{AE}\rangle = \sum_{i\in J}\alpha_i|i\rangle|\varphi^i_E\rangle \in \mathcal{H}_A\otimes\mathcal{H}_E \quad \textit{and} \quad \rho^{\text{mix}}_{AE} = \sum_{i\in J}|\alpha_i|^2|i\rangle\langle i|\otimes|\varphi^i_E\rangle\langle\varphi^i_E|$$

for some subset $J \subseteq I$. Furthermore, let ρ_{WE} and ρ^{mix}_{WE} describe the hybrid systems obtained by measuring subsystem A of $|\varphi_{AE}\rangle$ and ρ^{mix}_{AE}, respectively, in basis $\{|w\rangle\}_{w\in\mathcal{W}}$ to observe outcome W. Then,

$$H_{\min}(\rho_{WE}|E) \geq H_{\min}(\rho^{\text{mix}}_{WE}|E) - \log|J|\,.$$

The main tool to prove this lemma is the Cauchy-Schwarz inequality.

Theorem 3.18 (Cauchy-Schwarz Inequality) *Let $\mathcal{H}$ be a Hilbert space. Then,*

$$|\langle\varphi|\psi\rangle|^2 \leq \langle\varphi|\varphi\rangle\langle\psi|\psi\rangle \qquad \forall\,|\varphi\rangle,|\psi\rangle\in\mathcal{H}.$$

A proof can be found in any standard textbook on functional analysis. For a proof written in Dirac's braket notation, see [NC00, Box 2.1].

Proof of Lemma 3.17. We will show that $|J|\rho^{\text{mix}}_{WE} \geq \rho_{WE}$. It then follows that for any density matrix σ_E and for any non-negative $h \in \mathbb{R}$

$$\begin{aligned} 2^{-(h-\log|J|)} \cdot \mathbb{I}_W \otimes \sigma_E - \rho_{WE} &\geq 2^{-h}|J| \cdot \mathbb{I}_W \otimes \sigma_E - |J|\rho^{\text{mix}}_{WE} \\ &= |J|(2^{-h} \cdot \mathbb{I}_W \otimes \sigma_E - \rho^{\text{mix}}_{WE}) \end{aligned}$$

so that if the right-hand side is positive semidefinite then so is the left-hand side. The claimed bound $H_{\min}(\rho_{WE}|E) \geq H_{\min}(\rho^{\text{mix}}_{WE}|E) - \log|J|$ then follows by the definition of the min-entropy.

Writing out the measurements explicitly yields

$$\begin{aligned} \rho_{WE} &= \sum_{w\in\mathcal{W}} (|w\rangle\langle w| \otimes \mathbb{I}_E)|\varphi_{AE}\rangle\langle\varphi_{AE}|(|w\rangle\langle w| \otimes \mathbb{I}_E) \\ &= \sum_{w\in\mathcal{W}} \sum_{i,j\in J} \alpha_i\bar{\alpha}_j |w\rangle\langle w|i\rangle\langle j|w\rangle\langle w| \otimes |\varphi^i_E\rangle\langle\varphi^j_E| \end{aligned}$$

and

$$\rho^{\text{mix}}_{WE} = \sum_{i\in J} |\alpha_i|^2 \sum_{w\in\mathcal{W}} |\langle w|i\rangle|^2 |w\rangle\langle w| \otimes |\varphi^i_E\rangle\langle\varphi^i_E|.$$

We want to show that $\langle\xi|((|J|\rho^{\text{mix}}_{WE} - \rho_{WE})|\xi\rangle \geq 0$ for all $|\xi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_E$. By the Schmidt decomposition, we may write $|\xi\rangle = \sum_{w\in\mathcal{W}} \beta_w|w\rangle|\psi^w_E\rangle$, where $|\psi^w_E\rangle \in \mathcal{H}_E$ for all $w \in \mathcal{W}$.

$$\begin{aligned} \langle\xi|\rho_{WE}|\xi\rangle &= \sum_{v,w,x\in\mathcal{W}} \bar{\beta}_v\beta_x \sum_{i,j\in J} \alpha_i\bar{\alpha}_j \langle v|w\rangle\langle w|i\rangle\langle j|w\rangle\langle w|x\rangle \otimes \langle\psi^v_E|\varphi^i_E\rangle\langle\varphi^j_E|\psi^x_E\rangle \\ &= \sum_{w\in\mathcal{W}} |\beta_w|^2 \sum_{i,j\in J} \alpha_i\bar{\alpha}_j \langle w|i\rangle\langle j|w\rangle \otimes \langle\psi^w_E|\varphi^i_E\rangle\langle\varphi^j_E|\psi^w_E\rangle \\ &= \sum_{w\in\mathcal{W}} |\beta_w|^2 \Big(\sum_{i\in J} \alpha_i\langle w|i\rangle\langle\psi^w_E|\varphi^i_E\rangle\Big)\Big(\sum_{j\in J} \bar{\alpha}_j\langle j|w\rangle\langle\varphi^j_E|\psi^w_E\rangle\Big) \\ &= \sum_{w\in\mathcal{W}} |\beta_w|^2 \Big|\sum_{i\in J} \alpha_i\langle w|i\rangle\langle\psi^w_E|\varphi^i_E\rangle\Big|^2, \end{aligned}$$

and

$$\begin{aligned}\langle\xi|\rho_{WE}^{\text{mix}}|\xi\rangle &= \sum_{v,w,x\in\mathcal{W}} \bar{\beta}_v\beta_x \sum_{i\in J} |\alpha_i|^2 |\langle w|i\rangle|^2 \langle v|w\rangle\langle w|x\rangle\langle\psi_E^v|\varphi_E^i\rangle\langle\varphi_E^i|\psi_E^x\rangle \\ &= \sum_{w\in\mathcal{W}} |\beta_w|^2 \sum_{i\in J} |\alpha_i|^2 |\langle w|i\rangle|^2 \langle\psi_E^w|\varphi_E^i\rangle\langle\varphi_E^i|\psi_E^w\rangle \\ &\geq \frac{1}{|J|}\sum_{w\in\mathcal{W}} |\beta_w|^2 \Big|\sum_{i\in J}\alpha_i\langle w|i\rangle\langle\psi_E^w|\varphi_E^i\rangle\Big|^2 = \frac{1}{|J|}\langle\xi|\rho_{WE}|\xi\rangle,\end{aligned}$$

where the inequality follows from Cauchy-Schwarz inequality (Theorem 3.18). Hence, the operator $|J|\rho_{WE}^{\text{mix}} - \rho_{WE}$ is positive semidefinite and the claim follows. □

We apply Lemma 3.17 to an n-qubit system A where $|\varphi_{AE}\rangle$ is a superposition of states with relative Hamming weight δ-close to β within A:[10]

$$|\varphi_{AE}\rangle = \sum_{\substack{b\in\{0,1\}^n \\ |\eta(b)-\beta|\leq\delta}} |b\rangle|\varphi_E^b\rangle\,.$$

It is well known that $\big|\{b\in\{0,1\}^n : |\eta(b)-\beta|\leq\delta\}\big| \leq \big|\{b\in\{0,1\}^n : \eta(b)\leq\beta+\delta\}\big| \leq 2^{nh(\beta+\delta)}$ for $\beta+\delta\leq\frac{1}{2}$, where the function h is the binary entropy function.[11]

Since measuring qubits within a state $|b\rangle$ in the *Hadamard* basis produces uniformly random bits, we can conclude the following.

Corollary 3.19 *Let A be an n-qubit system, let the state $|\varphi_{AE}\rangle$ of AE be a superposition of states with relative Hamming weight δ-close to β within A, where $\beta+\delta\leq\frac{1}{2}$, and let the random variable X be obtained by measuring A in basis $H^\theta\{|0\rangle,|1\rangle\}^{\otimes n}$ for $\theta\in\{0,1\}^n$. Then*

$$H_{\min}(X|E) \geq \mathrm{wt}(\theta) - nh(\beta+\delta)\,.$$

Consider now the following quantum-cryptographic setting. Bob prepares and hands over to Alice an n-qubit quantum system A, which ought to be in state $|\varphi_A^\circ\rangle = |0\rangle\cdots|0\rangle$. However, since Bob might be dishonest, the state of A could

[10] System A considered here corresponds to the subsystem $A_{\bar{t}}$ in the previous section, after having measured A_t of the ideal state.

[11] There exists a corresponding upper bound for the cardinality of a q-ary Hamming ball (with arbitrary q), expressed in terms of the so-called q-ary entropy function; we do not elaborate on this here, since we now focus on the binary case.

be anything, even entangled with some system E controlled by Bob. Our results now imply the following: Alice can apply a suitable sampling strategy to convince herself that the joint state of the remaining subsystem of A and of E is (close to) a superposition of states with bounded relative Hamming weight. From Corollary 3.19, we can then conclude that with respect to the min-entropy of the measurement outcome, the state of A behaves similarly to the case where Bob honestly prepares A to be in state $|\varphi_A^\circ\rangle$. By Remark 3.12, i.e., by doing a suitable change of basis, the same holds if $|\varphi_A^\circ\rangle = H^{\hat{\theta}}|\hat{x}\rangle$ for arbitrary fixed $\hat{\theta}, \hat{x} \in \{0,1\}^n$, where $\mathrm{wt}(\theta)$ is replaced by the Hamming distance between θ and $\hat{\theta}$. We will make use of this in the applications in the upcoming sections.

3.4 A Security Proof for Quantum Oblivious Transfer

In a (one-out-of-two) *oblivious transfer* (OT) Alice sends two messages, $m_0, m_1 \in \{0,1\}^\ell$ to Bob. Bob may choose to receive one of the two messages, m_c. The security requirements demand that Bob learns no information on the other message, m_{1-c}, while at the same time Alice remains ignorant about Bob's choice bit c.

Back in 1991, Bennett *et al.* proposed a quantum scheme for OT, i.e., a QOT scheme [BBCS91]. The scheme makes use of a *bit commitment* (BC), which at that point in time was believed to be implementable with unconditional security by a quantum scheme. Bennett *et al.*, however, merely claimed security of their scheme without providing any proof. In 1994, Mayers and Salvail proved the QOT scheme secure against a limited class of attacks [MS94], and, subsequently, Yao presented a full security proof without limiting the adversary's capabilities [Yao95]. However, Yao's proof is lengthy and very technical, and thus hard to understand. Furthermore, security is phrased and proven in terms of *accessible information*, of which we now know that it is a too weak information measure to guarantee security as required.

Here we show how our sampling-strategy framework naturally leads to a new security proof for Bennett *et al.*'s QOT scheme. The new proof is simple and conceptually easy-to-understand, and security is expressed and proven by means of a security definition that is currently accepted to be "the right one." Furthermore, it allows for an explicit bound on the imperfection of the scheme for any set of parameters (number of transmitted qubits, length of messages etc.), rather than merely providing an asymptotic security claim. Nowadays, we of course know that BC (as well as QOT) cannot be implemented with unconditional security by means of a quantum scheme: QBC is impossible [May97, LC97]. As such QOT cannot be instantiated from scratch. Nevertheless, the existence of a QOT scheme based on a (hypothetical) BC is still an interesting result, since in the non-quantum world, a BC alone does *not* allow to implement OT.

Below, we describe Bennett *et al.*'s QOT scheme (with some minor modifications), which we denote as QOT. Actually, QOT corresponds to the *randomized oblivious transfer* used within Bennett *et al.*'s QOT scheme, where the messages m_0 and m_1, called k_0 and k_1 in QOT, are not *input* by Alice (her input is empty: $\perp$) but randomly produced during the course of the scheme and then *output* to Alice. The desired non-randomized OT is then obtained simply by one-time-pad encrypting Alice's input messages m_0 and m_1 with the keys k_0 and k_1, respectively. Security of the non-randomized OT follows immediately from the security of the randomized OT by the properties of the one-time pad (see Proposition 2.53).

QOT is parameterized by parameters $n, k, \ell \in \mathbb{N}$, where n is the number of qubits communicated, ℓ the bit-length of the messages/keys k_0, k_1, and k is the size of the "test set" t, which we require to be at most $n/2$. QOT makes use of a universal hash function $g : \mathcal{R} \times \{0,1\}^n \to \{0,1\}^\ell$. For $x' \in \{0,1\}^{n'}$ with $n' < n$, we define $g(r, x')$ as $g(r, x)$ where $x \in \{0,1\}^n$ is obtained from x' by padding it with sufficiently many 0's. Furthermore, the scheme makes use of a BC, which we model as an ideal BC functionality. One can think of this ideal BC as a trusted party. This party accepts an input from the sender in the commit phase, and forwards this input to the receiver in the opening phase, and neither the sender nor the receiver is able to cheat in any way. Alternatively, at the cost of losing unconditional security against dishonest Alice, we may use a BC implementation that is perfectly binding and computationally hiding.[12] Finally, for simplicity, we assume a *noise-free* quantum channel. For the more realistic setting of noisy quantum communication, an error-correcting code can be applied in a similar fashion as in the original scheme; this will not significantly affect our proof. In the upcoming protocol descriptions, we make use of our convention to speak about a basis θ (or $\hat{\theta}$) in $\{0,1\}^n$ when we actually mean $H^\theta\{|0\rangle, |1\rangle\}^{\otimes n}$ (respectively $H^{\hat{\theta}}\{|0\rangle, |1\rangle\}^{\otimes n}$). Also, please recall the general remarks made about protocols at the beginning of Section 2.11. Protocol 3.1 shows the description of QOT.

Note that our protocol, contrary to most QOT protocols given in the literature (including [BBCS91]), uses the same seed r to compute both keys (k_0 and k_1). Why we can do this will be made clear in the proof against dishonest Bob.

[12] Note that we do not claim any kind of composability for this computational setting. In case of a perfectly hiding and computationally binding BC scheme, our techniques do not apply directly. A specific variant of the latter case (in which the BC is required to have some additional properties) is handled in [DFL$^+$09].

1. *Preparation:* Alice chooses $x \overset{\mathrm{r}}{\leftarrow} \{0,1\}^n$ and $\theta \overset{\mathrm{r}}{\leftarrow} \{0,1\}^n$ and sends the n qubits $H^\theta|x\rangle$ to Bob. Bob selects $\hat{\theta} \overset{\mathrm{r}}{\leftarrow} \{0,1\}^n$ and measures the received qubits in basis $\hat{\theta}$, obtaining $\hat{x} \in \{0,1\}^n$.
2. *Commitment:* Bob commits bit-wise to $\hat{\theta}$ and $\hat{x}$. Alice samples a random subset $t \subset [n]$ of cardinality k and asks Bob to open the commitments to $\hat{\theta}_i$ and $\hat{x}_i$ for all $i \in t$. Alice verifies the opened commitments by checking that $\hat{x}_i = x_i$ whenever $\hat{\theta}_i = \theta_i$. She internally stores the outcome of this check, i.e., `accept` or `reject`, for later use in step 4.
3. *Set partitioning:* Alice sends θ to Bob. Bob partitions $\bar{t}$ into the subsets $I_c = \{i \in \bar{t} : \theta_i = \hat{\theta}_i\}$ and $I_{1-c} = \{i \in \bar{t} : \theta_i \neq \hat{\theta}_i\}$ and sends I_0 and I_1 to Alice.
4. *Key extraction:* Alice chooses $r \overset{\mathrm{r}}{\leftarrow} \mathcal{R}$ and sends it to Bob. Bob computes $\hat{k}_c = g(r, \hat{x}_{I_c})$. In case of `accept`, Alice computes k_0 and k_1 as $k_0 := g(r, x_{I_0})$ and $k_1 := g(r, x_{I_1})$. Otherwise, i.e., in case of `reject`, she sets k_0 and k_1 to random ℓ-bit strings.

Protocol 3.1: $\texttt{QOT}(\perp; c)$

It is trivial to see that for honest Alice and Bob: $\hat{k}_c = k_c$.

Furthermore, security against dishonest Alice, who is trying to learn information on c, is easy to see and not the issue here: in case of a perfect BC functionality, Alice learns no information on c no matter what she does; in case of a computationally hiding BC implementation, all information she obtains on c is "hidden within the commitments," and thus computational security follows from the computational hiding property.

Nevertheless, we will give a formal proof for the dishonest-Alice case in the section below. The security definition that we use is compatible with that of [FS09], meaning that—when using an ideal BC functionality—sequential composability is guaranteed when Alice is dishonest.

In Section 3.4.2, we deal with security against dishonest Bob.

3.4.1 Security against Dishonest Alice

Theorem 3.20 *Consider an execution of* $\texttt{QOT}$ *between dishonest Alice and honest Bob. Let* $C \in \{0,1\}$ *be Bob's input and let* E *be Alice's quantum system at the end of* *the protocol. For any dishonest Alice, there exist random variables K_0 and K_1 such that $\hat{K}_C = K_C$ and*

$$\rho_{CK_0K_1E} = \rho_C \otimes \rho_{K_0K_1E}.$$

Proof. To analyze the security against dishonest Alice, we slightly modify `QOT` into a protocol in which it is obvious that the claim holds. Nevertheless, the modified protocol remains equivalent to `QOT` in that both protocols produce exactly the same final state $\rho_{C\hat{K}_CE}$.

The first modification is that we change the specification of the ideal BC functionality such that it allows Bob to cheat. I.e., it provides the option for Bob to modify his commitment in the opening phase. Bob will not yet make use of this cheating possibility (in the context of this modification). Obviously, under this modification the protocol produces exactly the same state.

The second modification is that we let Bob postpone his measurements. In step 2, he still commits to $\hat{\theta}$ but commits to, say, the all-zero string in place of $\hat{x}$. Then, upon Alice's opening request, Bob only measures the qubits indexed by t in basis $\hat{\theta}_t$, and uses the cheating possibility of the commitment scheme to open the correct measurement outcomes to Alice. Bob postpones the measurements of the remaining qubits (indexed by $\bar{t}$) to step 3, after Alice announces her basis θ. It is clear that postponing these measurements does not change the final state.

As a third modification, we let Bob measure these remaining qubits (indexed by $\bar{t}$) *in Alice's basis θ*. This means that the qubits indexed by I_{1-C} are measured in a basis with a *different* distribution than in `QOT`. Nevertheless, Bob never uses the outcomes of these measurements, so the final state $\rho_{C\hat{K}_CE}$ is indeed exactly the same as in the original protocol.

Because Bob measures in Alice's basis, it holds that $\hat{X}_{\bar{t}} = X_{\bar{t}}$ and thus he can compute $K_0 := g(r, X_{I_0})$ as well as $K_1 := g(r, X_{I_1})$. This immediately proves the existence claim of these random variables.

Then, note that Bob's input C is only used in step 4 when Bob computes $\hat{K}_C$. I.e., the state $\rho_{K_0K_1E}$ is computed completely regardless of C and thus

$$\rho_{CK_0K_1E} = \rho_C \otimes \rho_{K_0K_1E}.$$

Finally, Bob sets $\hat{K}_C = K_C$. Hence this proves the claim. □

3.4.2 The Harder Case: Security against Dishonest Bob

Proving security against dishonest Bob is much more subtle, and is the goal of this section. Clearly, *if* Bob indeed measures the qubits in the preparation phase with respect to some choice $\hat{\theta}$, then security is easy to see: no matter how he partitions $\bar{t}$ into I_0 and I_1, on at least one of x_{I_0} and x_{I_1} he has some lower bounded uncertainty, and privacy amplification finishes the job. The intuition is now that the commitment phase forces Bob to essentially measure all qubits with respect to some choice $\hat{\theta}$, as otherwise he will get caught with overwhelming probability. However, proving this rigorously is non-trivial.

For our proof of security against dishonest Bob, we first introduce a slightly modified version of the protocol, QOT*, shown as Protocol 3.2. QOT* is only of proof-technical interest because it asks Alice to perform some actions that she could not do in practice. However, her actions are well-defined, and it follows from standard arguments that Bob's view of QOT is exactly the same as of QOT*. It thus suffices to prove security (against dishonest Bob) for QOT*.

QOT* is obtained from QOT by means of the following two modifications. First, for every $i \in [n]$, instead of sending $H^{\theta_i}|x_i\rangle$, Alice prepares an EPR pair A_iB_i of which she sends B_i to Bob and measures A_i, at some later point in the protocol, in basis θ_i to obtain x_i. By elementary properties of EPR pairs, and since actions on different subsystems commute, this does not affect Bob's view of the protocol. Second, Alice measures her qubits A_t within the test subset t in *Bob's basis* $\hat{\theta}_t$ (rather than in θ_t) to obtain x_t, but she still only verifies correctness of Bob's $\hat{x}_i$'s with $i \in t$ for which $\hat{\theta}_i = \theta_i$. Note that by assumption on the BC, the string $\hat{\theta}$ to which Bob can open his commitments is uniquely determined at this point, and thus Alice's action is well-defined, although not feasible in real life. This modification only influences Alice's bits x_i for which $i \in t$ and $\hat{\theta}_i \neq \theta_i$; however, since these bits are not used in the protocol, it has no effect on Bob's view.

Our proof for the security of QOT*, and thus of QOT, against dishonest Bob follows quite easily from our treatment of sampling strategies from Section 3.3. The proof is given below, after the formal security statement in Theorem 3.21. We would like to point out that our security guarantee implies the security definition proposed and studied in [FS09] for (randomized) OT, which in particular implies sequential *composability* when used as a sub-routine in a classical outer protocol.

Theorem 3.21 (Security of QOT) *Consider an execution of QOT (respectively QOT*) between honest Alice and dishonest Bob. Let K_0 and K_1 be the keys in $\{0,1\}^\ell$ output by Alice. Then, there exists a bit c so that K_{1-c} is close to random-and-independent*

1. *Preparation:* Alice prepares n EPR pairs of the form $(|0\rangle|0\rangle + |1\rangle|1\rangle)/\sqrt{2}$, and sends one qubit of each pair to Bob, who proceeds as in the original scheme QOT to obtain $\hat{\theta}$ and $\hat{x}$. Alice chooses $\theta \overset{\mathbf{r}}{\leftarrow} \{0,1\}^n$, but she does not measure her qubits yet.
2. *Commitment:* Bob commits to $\hat{\theta}$ and $\hat{x}$, and Alice chooses a random subset $t \subset [n]$ of cardinality k, as in `QOT`. Next, Alice measures her qubits that are indexed by t in *Bob's* basis $\hat{\theta}_t$ to obtain x_t. Then, Alice sends t to Bob and they proceed as in `QOT`, meaning that Bob opens these commitments and Alice verifies them.
3. *Set partitioning:* As in `QOT`. Additionally, Alice measures her qubits corresponding to I_0 in basis θ_{I_0} to obtain x_{I_0} and her qubits corresponding to I_1 in basis θ_{I_1} to obtain x_{I_1}.
4. *Key extraction:* Exactly as in the original scheme `QOT`.

Protocol 3.2: $\texttt{QOT}^*(\perp; c)$

of Bob's view (given K_c) in that for any $\epsilon, \delta > 0$:

$$\tfrac{1}{2}\|\rho_{K_{1-c}K_cE} - \tfrac{1}{2^\ell}\mathbb{I} \otimes \rho_{K_cE}\|_1$$
$$\leq \frac{1}{2} \cdot 2^{-\frac{1}{2}\left(\left(\frac{1}{4}-\frac{\epsilon}{2}-h(\delta)\right)(n-k)-\ell\right)} + \sqrt{6}\exp(-\delta^2 k/100) + 2\exp\!\big(-2\epsilon^2(n-k)\big),$$

where E denotes the quantum state output by Bob, and $\mathbb{I}$ the identity operator on $\mathbb{C}^{2^\ell}$.

On a high level, the proof is as follows. Alice's checking procedure can be understood as applying a sampling strategy to the qubits she holds. From this we obtain that (except with a small error) the joint state she shares with Bob is a superposition of states with small relative Hamming weight within her subsystem $A_{\bar{t}}$. This implies that the joint state is a superposition of states with small relative Hamming weight also within $A_{I_{1-c}}$, where $c \in \{0,1\}$ is chosen such that $\theta_i \neq \hat{\theta}_i$ for approximately half (or more) of the indices i in I_{1-c}. It then follows from Corollary 3.19 that $x_{I_{1-c}}$, obtained by measuring $A_{I_{1-c}}$ in basis $\theta_{I_{1-c}}$, has high min-entropy, so that privacy amplification concludes the proof. The formal proof, which takes care of the details and keeps track of the error term, is given below.

Proof. We consider the state

$$|\varphi_{AE_\circ}\rangle \in \mathcal{H}_{A_1} \otimes \cdots \otimes \mathcal{H}_{A_n} \otimes \mathcal{H}_{E_\circ},$$

shared between Alice and Bob, after Bob has committed to $\hat{\theta}$ and $\hat{x}$, but before Alice chooses the test subset t. $|\varphi_{AE_\circ}\rangle$ is obtained from the n EPR-pairs by an arbitrary quantum operation (possibly involving measurements), applied only to Bob's part. Without loss of generality, we may assume that, given the commitments, the joint state is indeed pure. Furthermore, we consider the strings $\hat{\theta}$ and $\hat{x}$, to which Bob has committed. By the perfectly binding property, these are uniquely determined. For concreteness, and in order to have the notation fit nicely with Section 3.3, we assume $\hat{\theta} = \hat{x} = (0, \ldots, 0) \in \{0,1\}^n$; however, by Remark 3.12, the very same reasoning works for any $\hat{\theta}$ and $\hat{x}$.

The crucial observation now is that Alice's checking procedure within the commitment phase of QOT* can be understood as applying a sampling strategy to the state $|\varphi_{AE_\circ}\rangle$ in order to test closeness of A to the all-zero state $|0\rangle \cdots |0\rangle$. Indeed, Alice chooses a random subset $t \subset [n]$ of cardinality k, measures A_t (in the computational basis) to obtain x_t, and decides whether to accept or reject based on x_t; specifically, she takes a random subset $s \subseteq t$, given by $s = \{i \in t : \theta_i = \hat{\theta}_i\}$, and accepts if and only $x_s = 0$ for all $i \in s$. This is precisely the sampling strategy Ψ studied in Example 3.6, adapted to test closeness to $|0\rangle \cdots |0\rangle$ by accepting if and only if $f(t, x_t, s) = 0$. Note that, by the random choices of the θ_i's, s is indeed a random subset of t.

Thus, we can conclude that at the end of the commitment phase, for any fixed $\delta > 0$, the joint state of $A_{\bar{t}}E_\circ$ has collapsed to a state $|\psi_{A_{\bar{t}}E_\circ}\rangle$ that is (on average over Alice's choice of t and s) $\varepsilon^\delta_{\text{quant}}$-close to being a superposition of states with relative Hamming weight at most δ within $A_{\bar{t}}$ (except when Alice rejects the test, but in that case she will output random and independent keys at the end of the protocol and the theorem trivially holds). We proceed by assuming that the state $|\psi_{A_{\bar{t}}E_\circ}\rangle$ *equals* a superposition of states with small relative Hamming weight, and we take the error $\varepsilon^\delta_{\text{quant}}$ into account at the end of the proof.[13] Recall that by Theorem 3.13 and Example 3.6 (and its analysis in Section 3.2.3),

$$\varepsilon^\delta_{\text{quant}} \leq \sqrt{\varepsilon^\delta_{\text{class}}} \leq \sqrt{6}\exp(-k\delta^2/100) \,.$$

By the random choices of the θ_i's, it follows from Hoeffding's inequality (Theorem 2.11) that the Hamming weight of $\theta_{\bar{t}}$ is lower bounded by $\text{wt}(\theta_{\bar{t}}) \geq (\frac{1}{2} - \epsilon)(n - k)$ except with probability at most $2\exp(-2\epsilon^2(n-k))$.[14] Below, we assume that

[13] It now follows immediately from Corollary 3.19 that $H_{\min}(X_0 X_1|E_\circ)$ is "large," where X_0 collects the bits obtained by measuring A_{I_0} in basis θ_{I_0}, and correspondingly for X_1. However, in the end we need that $H_{\min}(X_{1-c}|X_c E_\circ)$ is "large" for some c, which does *not* follow from the former. Because of that, we need to make a small detour.

[14] Actually, for the one-sided bound, we could save the factor two in front of the exp.

the bound holds; we take the error probability into account at the end of the proof. It follows that regardless of how Bob divides $\bar{t}$ into I_0 and I_1, there exists $c \in \{0,1\}$ such that $\mathrm{wt}(\theta_{I_{1-c}}) \geq \frac{1}{2}(\frac{1}{2}-\epsilon)(n-k)$ (if Bob is honest, then c coincides with his input bit).

By re-arranging Alice's qubits, we write the state $|\psi_{A_{\bar{t}}E_\circ}\rangle$ as $|\psi_{A^{1-c}A^cE_\circ}\rangle$, where $A^0 := A_{I_0}$ and $A^1 := A_{I_1}$. Since $|\psi_{A_{\bar{t}}E_\circ}\rangle$ is a superposition of states with Hamming weight at most $(n-k)\delta$ within $A_{\bar{t}}$, it is easy to see that $|\psi_{A^{1-c}A^cE_\circ}\rangle$ is a superposition of states with Hamming weight at most $(n-k)\delta$ within A^{1-c}. Let the random variables X_{1-c} and X_c describe the outcome of measuring A^{1-c} and A^c in bases $\theta_{I_{1-c}}$ and θ_{I_c}, respectively, and let $\rho_{X_{1-c}X_cE_\circ}$ be the corresponding hybrid state. We may think of $\rho_{X_{1-c}X_cE_\circ}$ being obtained by *first* measuring A^{1-c}, resulting in a hybrid state $\rho_{X_{1-c}A^cE_\circ}$, and *then* measuring A^c; indeed, the order in which these measurements take place have no effect on the final state.

We can now apply Corollary 3.19 to the hybrid state $\rho_{X_{1-c}A^cE_\circ}$ obtained from measuring subsystem A^{1-c} within $|\psi_{A^{1-c}A^cE_\circ}\rangle$ and conclude that

$$H_{\min}(X_{1-c}|A^cE_\circ) \geq \mathrm{wt}(\theta_{I_{1-c}}) - h(\delta)\cdot|I_{1-c}| \geq \Big(\frac{1}{4}-\frac{\epsilon}{2}-h(\delta)\Big)(n-k)\,.$$

By the fact that quantum operations cannot decrease min-entropy (Lemma 3.1.12 in [Ren05]) it follows that the same bound in particular holds for $H_{\min}(X_{1-c}|X_cE_\circ)$. Applying privacy amplification[15] (Theorem 2.65), incorporating the error probabilities (expressed in terms of trace distance) obtained along the proof, and noting that Bob's processing of his information to obtain his final quantum state E does not increase the trace distance, concludes the proof. □

3.5 An Accessible Proof for Quantum Key Distribution

Recall that in quantum key distribution (QKD), Alice and Bob want to agree on a secret key in the presence of an adversary Eve. Alice and Bob are assumed to be able to communicate over a quantum channel and over an authenticated classical channel.[16] Eve may eavesdrop the classical channel (but not insert or modify messages), and she has full control over the quantum channel. For a more detailed introduction, see Chapter 1.

[15] Because we show a lower bound on the min-entropy of X_{1-c} *when given the raw key* X_c, we simply need one hash function instead of two independently chosen ones as in, e.g., [DFSS07].

[16] If the classical channel between Alice and Bob is not authentic, then authenticity of the communication can still be achieved by information-theoretic authentication techniques, at the cost of requiring Alice and Bob to initially share a short secret key.

The first and still most prominent QKD scheme is the famous BB84 QKD scheme due to Bennett and Brassard [BB84]. In this section, we show how our sampling-strategy framework leads to a simple security proof for the BB84 QKD scheme.

3.5.1 Survey of Existing QKD Proofs

Before discussing our proof, we want to discuss two existing QKD proofs, i.e., the proof by Shor and Preskill [SP00], as well as a more modern proof by Renner [Ren05] based on the quantum de Finetti theorem. We will merely explain the ideas behind those proofs on a high level, instead of discussing the proofs in detail.

Shor and Preskill's Proof

Entanglement purification is a functionality that takes as input a state that is fairly close to a product state of n EPR pairs, and outputs a state consisting of m perfect EPR pairs, for some $m < n$. It is well known that we can obtain a "shared-key *generation* protocol" from an entanglement-purification protocol by appending a step in which Alice and Bob measure their halves of the EPR pairs in a common basis.

Shor and Preskill [SP00] show an entanglement-purification protocol based on Calderbank-Shor-Steane codes, which are error-correcting codes for quantum states that can be used without requiring a quantum computer. The security of this entanglement-purification protocol follows from the earlier work of Lo and Chau.

Subsequently, Shor and Preskill reduce this entanglement-purification protocol to a key *distribution* protocol, in which, unlike the shared-key generation protocol, Alice samples a classical key herself and encodes it in a quantum state that she sends to Bob. By some additional simplifications, the latter protocol is further reduced to BB84.

Renner's Proof using the Quantum de Finetti Theorem

One of the reasons why it is non-trivial to prove the security of QKD is that the proof should hold for *any* input state. Renner's approach [Ren05] to circumvent this difficulty is to show that by performing some simple operations, the input state can be transformed into a state that is very close to a convex combination of product states, for which it is much easier to show security for the different subprotocols. I.e., it suffices to show security against collective attacks.[17]

[17] A side result of this approach is that *coherent attacks* on QKD are not more powerful than *collective attacks*.

More precisely, the first step is to apply a random permutation to the qubits, which makes the input state symmetric (i.e., permutation-invariant). Moreover, it can be shown that there always exists a purification of this state which is symmetric as well. Then, by the finite quantum de Finetti theorem, it holds that when tracing out certain parts of this purification, one obtains a state that is close to a convex combination of product states.

3.5.2 Our Proof

We will now discuss our proof for the BB84 scheme based on the sampling-strategy framework. Beyond its simplicity, our proof has some other nice features. For instance, it allows us to explicitly state (a bound on) the error probability of the QKD scheme for any given choices of the parameters. Additionally, our proof does not seem to take unnecessary detours or to make use of "loose bounds," and therefore we feel that the bound on the error probability we obtain is rather tight (although we have no formal argument to support this).

Our proof strategy can also be applied to other QKD schemes that are based on the BB84 encoding. For example, Lo *et al.*'s QKD scheme[18] [LCA05] can be proven secure by following exactly our proof, except that one needs to analyze a slightly different sampling strategy, namely the one from Example 3.8. On the other hand, it is yet unknown whether our framework can be used to prove, e.g., the six-state QKD protocol [Bru98] secure.

Actually, the QKD scheme we analyze is the entanglement-based version of the BB84 scheme (as initially suggested by Ekert [Eke91]). However, it is very well known and not too hard to show that security of the entanglement-based version implies security of the original BB84 QKD scheme.

The entanglement-based QKD scheme, QKD, is parameterized by the total number n of qubits sent in the protocol and the number k of qubits used to estimate the error rate of the quantum channel (where we require $k \leq n/2$). Additional parameters, which are determined during the course of the protocol, are the observed error rate β and the number $\ell \in \mathbb{N} \cup \{0\}$ of extracted key bits. QKD makes use of a universal hash function $g : \mathcal{R} \times \{0,1\}^{n-k} \to \{0,1\}^{\ell}$ and a linear binary error correcting code of length $n-k$ that allows to correct up to a β'-fraction of errors (except maybe with negligible probability) for some $\beta' > \beta$. The choice of how much β' exceeds β is a trade-off between keeping the probability that Alice and Bob end

[18] In this scheme, Alice and Bob bias the choice of the bases so that they measure a bigger fraction of the qubits in the same basis.

up with different keys small and increasing the size of the extractable key. We will write m for the bit size of the syndrome of this error-correcting code. Protocol `QKD` can be found below.

1. *Qubit distribution:* Alice prepares n EPR pairs of the form $(|0\rangle|0\rangle + |1\rangle|1\rangle)/\sqrt{2}$, and sends one qubit of each pair to Bob, who confirms the receipt of the qubits. Then, Alice picks $\theta \overset{\text{r}}{\leftarrow} \{0,1\}^n$ and sends it to Bob, and Alice and Bob measure their respective qubits in basis θ to obtain $x \in \mathbb{F}_2^n$ on Alice's side respectively $y \in \mathbb{F}_2^n$ on Bob's side.
2. *Error estimation:* Alice chooses a random subset $s \subset [n]$ of size k and sends it to Bob. Then, Alice and Bob exchange x_s and y_s and compute $\beta := \eta(x_s \oplus y_s)$.
3. *Error correction:* Alice sends the syndrome syn of $x_{\bar{s}}$ to Bob with respect to a suitable linear error correcting code (as described above). Bob uses syn to correct the errors in $y_{\bar{s}}$ and obtains $\hat{x}_{\bar{s}}$. Let m be the bit-size of syn.
4. *Key distillation:* Alice chooses a random seed r for a universal hash function g with range $\{0,1\}^\ell$, where ℓ satisfies $\ell < (1-h(\beta))n - k - m$ (or $\ell = 0$ if the right-hand side is not positive), and sends it to Bob. Then, Alice and Bob compute $k_\text{A} := g(r, x_{\bar{s}})$ and $k_\text{B} := g(r, \hat{x}_{\bar{s}})$, respectively.

Protocol 3.3: `QKD`

It is not hard to see that $K_\text{A} = K_\text{B}$ except with negligible probability (in n). Furthermore, if no Eve interacts with the quantum communication in the qubit distribution phase then $x = y$ in case of a noise-free quantum channel, or more generally, $\eta(X - Y) \approx \phi$ in case the quantum channel is noisy and introduces an error probability $0 \leq \phi < \frac{1}{2}$. It follows that $\beta \approx \phi$, so that using an error correcting code that approaches the Shannon bound, Alice and Bob can extract close to $(1 - 2h(\phi))(n - k)$ bits of secret key, which is positive for ϕ smaller than approximately 11%. The difficult part is to prove security against an active adversary Eve. We first state the formal security claim.

Note that we cannot expect that Eve has (nearly) no information on K_A, i.e., that $\frac{1}{2}\|\rho_{K_\text{A}E} - \frac{1}{|\mathcal{K}|}\mathbb{I}_{K_\text{A}} \otimes \rho_E\|_1$ is small, since the bit-length ℓ of K_A is not fixed but depends on the course of the protocol, and Eve can influence and thus obtain information on ℓ (and thus on K_A). Theorem 3.22 though guarantees that the bit-length ℓ is the *only* information Eve learns on K_A, in other words, K_A is essentially random-and-independent of E when given ℓ.

Theorem 3.22 (Security of QKD) *Consider an execution of QKD in the presence of an adversary Eve. Let K_A be the key obtained by Alice, and let E be Eve's quantum system at the end of the protocol. Let $\widetilde{K}$ be chosen uniformly at random of the same bit-length as K_A. Then, for any δ with $\beta + \delta \leq \frac{1}{2}$:*

$$\tfrac{1}{2}\|\rho_{K_A E} - \rho_{\widetilde{K}E}\|_1 \leq \tfrac{1}{2} \cdot 2^{-\frac{1}{2}\left(n - nh(\beta+\delta) - k - m - \ell\right)} + 2\exp\left(-\tfrac{1}{6}\delta^2 k\right).$$

From an application point of view, the following question is of interest. Given the parameters n and k, and given a run of the protocol with observed error rate β and where an error-correcting code with syndrome length m was used, what is the maximal size ℓ of the extractable key K_A if we want $\frac{1}{2}\|\rho_{K_A E} - \rho_{\widetilde{K}E}\|_1 \leq \epsilon$ for a given ϵ? From the bound in Theorem 3.22, it follows that for every choice of δ (with $\beta + \delta \leq \frac{1}{2}$), one can easily compute a possible value for ℓ simply by solving for ℓ. In order to compute the optimal value, one needs to maximize ℓ over the choice of δ.

The formal proof of Theorem 3.22 is given below. Informally, the argument goes as follows. The error estimation phase can be understood as applying a sampling strategy. From this, we can conclude that the state from which the raw key, $x_{\bar{s}}$, is obtained, is a superposition of states with bounded Hamming weight, so that Corollary 3.19 guarantees a certain amount of min-entropy within $x_{\bar{s}}$. Privacy amplification then finishes the proof.

To model the error estimation procedure as a sampling strategy, we will need to consider a modified but *equivalent* way for Alice and Bob to jointly obtain $x_s \in \mathbb{F}_2^k$ and $y_s \in \mathbb{F}_2^k$ from the initial joint state, which will allow them to obtain their sum $x_s \oplus y_s$, and thus to compute β, *before* they measure the remaining part of the state, whose outcome then determines $x_{\bar{s}}$. This modification is based on the so-called CNOT operation, U_{CNOT}, acting on $\mathbb{C}^2 \otimes \mathbb{C}^2$, which has the following properties for all $b, c \in \mathbb{F}_2$:

$$U_{\text{CNOT}}(|b\rangle|c\rangle) = |b\rangle|b \oplus c\rangle \quad \text{and} \quad U_{\text{CNOT}}(H|b\rangle H|c\rangle) = H|b \oplus c\rangle H|c\rangle\,, \quad (3.10)$$

where the first holds by definition of U_{CNOT}, and the second is straightforward to verify.

Proof. Throughout the proof, we use capital letters, Θ, X etc. for the *random variables* representing the corresponding choices of θ, x etc. in protocol QKD. Let the state, shared by Alice, Bob and Eve right after the quantum communication in the qubit distribution phase, be denoted by $|\psi_{ABE_\circ}\rangle$;[19] without loss of generality, we may indeed assume the shared state to be pure. For every $i \in [n]$, Alice and Bob then measure the respective qubits A_i and B_i from $|\psi_{ABE_\circ}\rangle$ in basis Θ_i, obtaining X_i and Y_i. This results in the hybrid state $\rho_{\Theta XYE_\circ}$. For the proof, it will be convenient to introduce the additional random variables $W = (W_1, \ldots, W_n)$ and $Z = (Z_1, \ldots, Z_n)$, defined by

$$Z_i := X_i \oplus Y_i \qquad \text{and} \qquad W_i := \begin{cases} X_i & \text{if } \Theta_i = 0 \\ Y_i & \text{if } \Theta_i = 1 \end{cases} . \tag{3.11}$$

Note that, when given Θ, the random variables W and Z are uniquely determined by X and Y *and vice versa*, and thus we may equivalently analyze the hybrid state $\rho_{\Theta WZE_\circ}$.

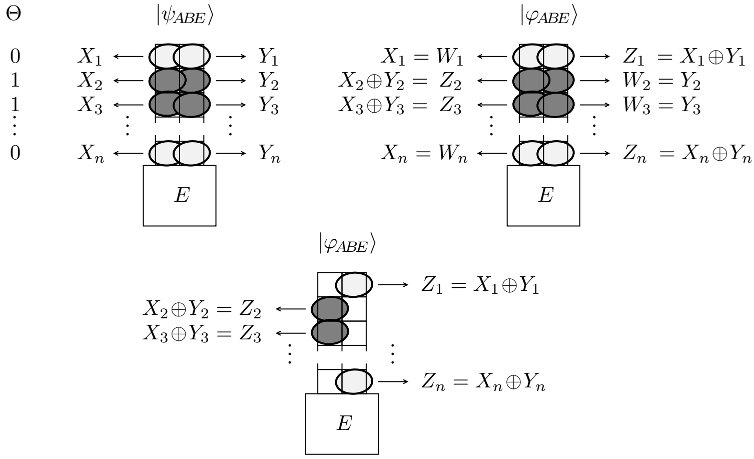

Figure 3.1: Original and modified experiments for obtaining the same state $\rho_{\Theta WZE_\circ}$.

For the analysis, we will consider a slightly *different* experiment for Alice and Bob to obtain the very *same* state $\rho_{\Theta WZE_\circ}$; the advantage of the modified experiment is that it can be understood as a sampling strategy. The modified experiment is as follows. First, the CNOT transformation is applied to every qubit pair A_iB_i within $|\psi_{ABE_\circ}\rangle$ for $i \in [n]$, such that the state $|\varphi_{ABE_\circ}\rangle = (U_{\text{CNOT}}^{\otimes n} \otimes \mathbb{I}_{E_\circ})|\psi_{ABE_\circ}\rangle$ is obtained. Next, Θ is chosen at random as in the original scheme, and for every

[19] Note that $E_\circ$ represents Eve's quantum state just after the quantum communication stage, whereas E represents Eve's entire state of knowledge at the end of the protocol (i.e., the quantum information and all classical information gathered during execution of QKD).

$i \in [n]$ the qubit pair A_iB_i of the transformed state is measured as in the original scheme depending on Θ_i; however, if $\Theta_i = 0$ then the resulting bits are denoted by W_i and Z_i, respectively, and if $\Theta_i = 1$ then they are denoted by Z_i and W_i, respectively, such that which bit is assigned to which variable depends on Θ_i. This is illustrated in Figure 3.1 (left and middle), where light and dark colored ovals represent measurements in the computational and Hadamard basis, respectively. It now follows immediately from the properties (3.10) of the CNOT transformation and from the relation (3.11) between X, Y and W, Z that the state $\rho_{\Theta W Z E_\circ}$ (or, equivalently, $\rho_{\Theta X Y E_\circ}$) obtained in this modified experiment is exactly the same as in the original.

An additional modification we may do without influencing the final state is to *delay* some of the measurements: we assume that first the qubits are measured that lead to the Z_i's, and only at some later point, namely after the *error estimation* phase, the qubits leading to the W_i's are measured (as illustrated in Figure 3.1, right). This can be done since the relative Hamming weight of $X_S \oplus Y_S$ for a random subset $S \subset [n]$ (of size k) can be computed given Z alone.

The crucial observation is now that this modified experiment can be viewed as a particular sampling strategy Ψ, as a matter of fact as the sampling strategy discussed in Example 3.7, being applied to systems A and B of the state $|\varphi_{ABE_\circ}\rangle$. Indeed: first, a subset of the $2n$ qubit positions is selected according to some probability distribution, namely of each pair A_iB_i one qubit is selected at random (determined by Θ_i). Then, the selected qubits are measured to obtain the bit string $Z = (Z_1, \ldots, Z_n)$. And, finally, a value β is computed as a (randomized) function of Z: $\beta = \eta(Z_S)$ for a random $S \subset [n]$ of size k. We point out that here the reference basis (as explained in Remark 3.12) is not the computational basis for all qubits, but the Hadamard basis on the qubits in system A and the computational basis in system B; however, as discussed in Remark 3.12, we may still apply the results from Section 3.3 (appropriately adapted).

It thus follows that for any fixed $\delta > 0$, the remaining state, from which W is then obtained, is (on average over Θ and S) $\varepsilon^{\delta}_{\text{quant}}$-close to a state which is (for any possible values for Θ, Z and S) a superposition of states with relative Hamming weight in a δ-neighborhood of β. Note that the latter has to be understood with respect to the fixed reference basis (i.e., the Hadamard basis on A and the computational basis on B). In the following, we assume that the remaining state *equals* such a superposition; we will take the error below into account at the end of the proof,

$$\varepsilon^{\delta}_{\text{quant}} \leq \sqrt{\varepsilon^{\delta}_{\text{class}}} \leq 2\exp(-\tfrac{1}{6}\delta^2 k)\,.$$

where the bound on $\varepsilon^{\delta}_{\text{class}}$ is derived in Section 3.2.3 (Example 3.7).

Recall that W is now obtained by measuring the remaining qubits; however, the basis used is opposite to the reference basis, namely the computational basis on the qubits A_i and the Hadamard basis on the qubits B_i. Hence, by Corollary 3.19 (and the subsequent discussion) we get a lower bound on the min-entropy of W:

$$H_{\min}(W|\Theta Z S E_\circ) \geq n - nh(\beta + \delta)\,.$$

Since W is uniquely determined by X (and vice versa) when given Θ and Z, the same lower bound also holds for $H_{\min}(X|\Theta Z S E_\circ)$. Note that in `QKD`, the k qubit-pairs that are used for estimating β are not used anymore in the key distillation phase, so we are actually interested in the min-entropy of $X_{\bar{S}}$. Additionally, we should take into account that Alice sends an m-bit syndrome SYN during the error correction phase. Hence, by using the chain rule, we obtain

$$H_{\min}(X_{\bar{S}}|\Theta Z X_S SYN E_\circ) \geq n - nh(\beta + \delta) - k - m.$$

Finally, we apply privacy amplification (Theorem 2.65) which concludes the proof. □

Probably, it is possible to prove the lower bound: $(1 - h(\beta + \delta))(n - k) - m$ using a different sampling strategy. However, for that case the error probability of the related classical sampling strategy becomes harder to analyze. We have chosen for the current proof strategy and bound for the sake of simplicity.

3.6 Conclusion

We have shown a framework for predicting some property (namely the approximate Hamming weight, appropriately defined) of a population of quantum states, by measuring a small sample subset. The framework allows for new and simple security proofs for important quantum cryptographic protocols: the Bennett *et al.* QOT and the BB84 QKD scheme.

We find it particularly interesting that with our framework, the protocols for QOT and QKD can be proven secure by means of very similar techniques, even though they implement fundamentally different cryptographic primitives, and are intuitively secure due to very different reasons (namely in QOT the commitments force Bob to measure the communicated qubits, whereas in QKD Eve disturbs the communicated qubits when trying to observe them).[20]

[20] As pointed out by Louis Salvail, a connection between the security of QOT and QKD has also been made by Mayers [May95, May96].

4

Authentication from a Weak Key with a Privacy Requirement

This chapter is based on joint work with Serge Fehr [BF11].

Chapter Contents

4.1 Introduction . . . 134

4.1.1 Related Work . . . 135

4.1.2 Motivation . . . 135

4.1.3 Contributions . . . 136

4.1.4 The Fuzzy Case . . . 136

4.2 Security Definition . . . 137

4.3 Dodis and Wichs' Authentication Protocol . . . 137

4.3.1 Towards Achieving Key-Privacy . . . 139

4.4 Our Construction . . . 141

4.5 Proofs of Security and Privacy . . . 143

4.6 Instantiating the Building Blocks . . . 148

4.6.1 Look-Ahead Extractors against Classical Side Information . . . 148

4.6.2 Look-Ahead Extractors and Quantum Side Information? . . . 155

4.6.3 Security and Instantiation of the MAC . . . 156

4.6.4 Instantiating Protocol AUTH . . . 156

4.7 The Fuzzy Case . . . 159

4.8 Application: Password-Based Identification . . . 160

4.9 Open Problem . . . 161

4.1 Introduction

In this chapter, we consider the problem of achieving authentic[1] communication based on a *weak key* over a public channel that might be under the control of an active adversary. A key is *weak* if its min-entropy is an arbitrarily small fraction of its bit length. We study this problem in the information-theoretic setting, i.e., we assume the adversary to be computationally unbounded.

First of all, note that because we are dealing with an *active* adversary, the standard approach of using an extractor to turn the weak key into a strong one (which can then be used to perform standard message authentication) will *not* work, since the adversary can tamper with the extractor's seed.

Specifically, we consider the following scenario. Alice and Bob share a *long-term* key W. When needed, Alice and Bob can extract a weak *session key* X_W from an auxiliary source of randomness with the help of W. It should be guaranteed by the property of the auxiliary source that a potential adversary Eve who does not know W has limited information on the weak session key X_W. This is formalized by requiring that $H_{\min}(X_W|WE) \geq k$ for some parameter k, where E denotes Eve's side information. This scenario occurs naturally in, e.g., Maurer's *bounded-storage model* [Mau90], where W determines which part of the huge string to read, as well as in the quantum setting, where W determines in which basis to measure some quantum state.

The goal is to authenticate a message μ from Alice to Bob with the help of the weak session key X_W, while guaranteeing *security*, in that if Eve tampers with μ then this will be detected, and *privacy*, in that Eve cannot learn information about the long-term key W. We stress that the privacy property is vital for Alice and Bob to be able to re-use W. Note that once Alice and Bob can do message authentication with a weak key, then they can also do key agreement, simply by doing standard randomness extraction where the seed for the extractor is communicated in an authentic way.

We want to emphasize that, by assumption, every new session key X_W for the same long-term key W contains fresh randomness, provided by the auxiliary source. Therefore, the goal above does not contradict the well-known impossibility result of re-using an authentication key without refreshing. Also note that we do not specify how exactly the auxiliary source of randomness produces X_W from W; on the contrary, we want security no matter how X_W is obtained, as long as X_W contains enough min-entropy (given the adversary's information and W).

[1] For an introduction to message authentication, see Section 2.5.2.

4.1.1 Related Work

With regard to the above security property, the problem of authentication from a weak key in the presence of an active adversary is a fairly well-studied problem. To the best of our knowledge, we are the first to study a special case of this problem where the weak key is obtained from a long-term key and where privacy of the long-term key needs to be guaranteed. In particular, the works that we will mention below do not address this special case, and moreover they all fail to satisfy the privacy property.

In the following discussion, let n be the bitsize of the key (in our case, the session key) and k its min-entropy (in bits). It was proved by Dodis and Wichs [DW09] that non-interactive authentication is impossible when $k \leq n/2$, even when the parties have access to local non-shared randomness, which we will assume. For a good overview of earlier work on the case $k > n/2$, we refer to [DW09].

The first protocol for interactive authentication from arbitrarily weak keys is due to Renner and Wolf [RW03]. It requires $\Theta(\ell)$ rounds of interaction to authenticate an ℓ-bit message. In [DW09], an authentication protocol from arbitrarily weak keys is described that only needs two rounds of interaction, which is optimal (in terms of the number of rounds). Chandran *et al.* [CKOR10] focus on minimizing entropy loss and describe a privacy amplification protocol that is optimal with respect to entropy loss (up to constant factors). Their construction needs a linear number of rounds (linear in the security parameter).

The case where Alice and Bob share highly-correlated, but possibly unequal keys—the "fuzzy" case—is addressed in [RW04] and improved upon by Kanukurthi and Reyzin [KR09], but also covered by [DW09] and [CKOR10].

4.1.2 Motivation

The main motivation for the work in this chapter comes from *password-based identification* in the bounded-quantum-storage model (BQSM). As already mentioned in Section 2.11, Damgård *et al.* [DFSS07] propose two identification protocols: QID, which is only secure against dishonest Alice or Bob, and QID$^+$, which is also secure against a man-in-the-middle (MITM) attack. However, only QID is truly password-based; in QID$^+$, Alice and Bob, in addition to the password, also need to share a high-entropy key.

Now, the observation is that with the help of an authentication protocol with long-term-key privacy, the protocol QID$^+$ can be turned into a truly password-based identification protocol in the BQSM with security against MITM attacks.

Based on $\mathtt{QID}^+$, Damgård *et al.* also propose an *authenticated* quantum key distribution protocol in the BQSM, which, in contrast to standard quantum key distribution protocols, does not require authenticated communication but has the authentication "built in." Furthermore, in contrast to using standard quantum key distribution in combination with standard authentication, in the authenticated quantum key distribution protocol the authentication keys can be re-used. By making $\mathtt{QID}^+$ truly password-based, Damgård *et al.*'s authenticated QKD protocol will become truly password-based as well.

4.1.3 Contributions

We propose a new four-round protocol for message authentication with a weak session key X_W. The protocol is an extension of the two-round protocol by Dodis and Wichs [DW09], which is based on *look-ahead extraction*. Given a secure look-ahead extractor, we prove that our protocol satisfies *security* and *long-term-key privacy*, meaning that the adversary Eve cannot tamper with the authenticated message without being detected, nor does she learn a non-negligible amount of information on the long-term key W.

For the case where Eve's side information about X_W is classical, we can use the construction for a look-ahead extractor that is given in [DW09]. Contrary to what we have claimed in [BF11] (see Section 4.6.2 for a more detailed explanation), it remains an open problem to construct a look-ahead extractor that is secure against quantum side information, or, to prove that the construction given in [DW09] (which is secure in the presence of classical side information) is also secure against quantum side information. Hence, we cannot yet construct an authentication protocol that is secure in the quantum setting, which would be needed for our envisioned application, i.e., truly password-based identification in the BQSM with security against MITM attacks.

4.1.4 The Fuzzy Case

We will also discuss the "fuzzy case," i.e., where there are some errors between Alice's and Bob's weak session key. If Eve's side information is classical, then our techniques are known to be secure in the fuzzy case; in the quantum setting, however, this remains to be shown. Precisely this latter case—the quantum setting—is relevant for our password-based-identification application.

4.2 Security Definition

In this chapter, an *authentication protocol* is understood as a classical protocol between two parties Alice and Bob. Alice inputs a message μ and a weak session key X_W, and Bob inputs a message μ' and the same session key X_W. At the end of the protocol, Bob announces a Boolean decision whether to "accept" or "reject." The weak session key X_W may depend arbitrarily on a long-term key W. During the execution of the protocol, an adversary Eve has full control over the communication between Alice and Bob.

We require the protocol to fulfill the following formal definition.

Definition 4.1 Let $E_\circ, E$ denote Eve's respective a priori and a posteriori quantum systems, where the latter includes Bob's decision on whether to accept or reject. An $(n, k, m, \delta, \varepsilon)$ message-authentication protocol with long-term-key privacy is defined to satisfy the following properties:

1. *Correctness:* If there is no adversary Eve present, then for any message $\mu \in \{0,1\}^m$ and $\mu' = \mu$, and for any (distribution of the) key $X_W \in \{0,1\}^n$, Bob accepts with certainty.
2. *Security:* If $H_{\min}(X_W|WE_\circ) \geq k$, then for any $\mu, \mu' \in \{0,1\}^m$ with $\mu \neq \mu'$, the probability that Bob accepts is at most δ.
3. *Long-Term-Key Privacy:* If $\rho_{WE_\circ} = \rho_W \otimes \rho_{E_\circ}$ and $H_{\min}(X_W|WE_\circ) \geq k$, then
$$\frac{1}{2}\|\rho_{WE} - \rho_W \otimes \rho_E\|_1 \leq \varepsilon\,.$$

4.3 Dodis and Wichs' Authentication Protocol

In this section, we describe a slightly modified version of the two-round message-authentication protocol due to Dodis and Wichs [DW09]. We will use this protocol later as a "starting point" to construct our message-authentication protocol. We start by giving a few definitions that are crucial for the understanding of the protocol by Dodis and Wichs.

Definition 4.2 (Epsilon Look-Aheadness) Let t, ℓ be positive integers. Let $A := (A_1, \ldots, A_t)$ and $B := (B_1, \ldots, B_t)$ be random variables over $(\{0,1\}^\ell)^t$, and let E be a quantum system. For all $i \in \{0, \ldots, t-1\}$ let ε_i be defined as

$$\varepsilon_i := d_{\mathsf{unif}}(A_{i+1} \ldots A_t | B_1 \ldots B_i E)\,.$$

The ordered pair (A, B) is *ε-look-ahead conditioned on E* if $\varepsilon \geq \max_i \varepsilon_i$.

Definition 4.3 (Look-Ahead Extractor) $\mathsf{laExt} : \{0,1\}^n \times \{0,1\}^d \to (\{0,1\}^\ell)^t$ is called a (k,ε)-*look-ahead extractor* if for any random variable $X \in \{0,1\}^n$ and quantum system E with $H_{\min}(X|E) \geq k$ the following holds. Let $S \in \{0,1\}^d$ be an independent and uniformly distributed seed, and let $\widetilde{S} \in \{0,1\}^d$ be adversarially chosen given S and E; this may involve a (partial) measurement of E, resulting in the new state E'. Then, the ordered pair $(R, \widetilde{R})$ where $R = (R_1, \ldots, R_t) := \mathsf{laExt}(X;S)$ and $\widetilde{R} = (\widetilde{R}_1, \ldots, \widetilde{R}_t) := \mathsf{laExt}(X;\widetilde{S})$ is ε-look-ahead conditioned on S, $\widetilde{S}$ and E'.

Informally, a look-ahead extractor has the property that even if the adversary is allowed to modify the seed, when given the first i blocks of the key that is extracted using the modified seed, the remaining blocks of the key that is extracted using the correct seed still look random.

Definition 4.4 (Look-Ahead-Secure MAC) A family of functions

$$\{\mathsf{MAC}_\kappa : \{0,1\}^m \to \{0,1\}^s\},$$

indexed by keys $\kappa \in (\{0,1\}^\ell)^t$ is an (ε,δ) *look-ahead-secure* MAC if for any pair of fixed and distinct messages $\mu_\mathrm{A}, \mu_\mathrm{B} \in \{0,1\}^m$, $\mu_\mathrm{A} \neq \mu_\mathrm{B}$, and any ordered pair of random variables $(K, K') \in (\{0,1\}^\ell)^{2t}$ satisfying the look-ahead property with parameter ε conditioned on quantum system E,

$$p_{\mathsf{guess}}(\mathsf{MAC}_K(\mu_\mathrm{B}) \,|\, \mathsf{MAC}_{K'}(\mu_\mathrm{A}) E) < \delta\,.$$

We are now ready to present the Dodis and Wichs message-authentication protocol DWMAC. The version that we present here, Protocol 4.1, is slightly modified in that we assume that Alice has already sent her message μ_A to Bob, who has received it as μ_B (possibly $\neq \mu_\mathrm{A}$). This modification is for simplicity, and because we do not aim at minimizing the number of rounds. X_W is the weak key, known to both Alice and Bob. The function $\mathsf{laExt} : \{0,1\}^n \times \{0,1\}^d \to (\{0,1\}^\ell)^t$ is a (k,ε)-look-ahead extractor and $\mathsf{MAC}_\kappa : \{0,1\}^m \to \mathbb{F}_{2^s}$ is a (ε,δ) look-ahead-secure MAC.

Security of DWMAC follows immediately from the definitions of the underlying building blocks: laExt ensures that Alice and Bob's versions of the key K satisfy the look-ahead property, and in this case it is guaranteed that MAC acts as a secure MAC, even when Alice's key was modified.

However, in our setting where we additionally want to maintain privacy of the long-term key W, which may arbitrarily depend on X_W, DWMAC does not seem to be good enough, unless Eve remains passive. Indeed, if Eve does not manipulate the communicated seed R, then by the assumed lower bound on $H_{\min}(X_W|WE)$ it

Alice(X_W, μ_A)		Bob(X_W, μ_B)
		$R \overset{\mathrm{r}}{\leftarrow} \{0,1\}^d$
	$\xleftarrow{R}$	
$K := \mathsf{laExt}(X_W; R)$		$K := \mathsf{laExt}(X_W; R)$
$T_\mathrm{A} := \mathsf{MAC}_K(\mu_\mathrm{A})$		$T_\mathrm{B} := \mathsf{MAC}_K(\mu_\mathrm{B})$
	$\xrightarrow{T_\mathrm{A}}$	
		accept if: $T_\mathrm{A} = T_\mathrm{B}$ else: abort

Protocol 4.1: Dodis and Wichs' two-round protocol DWMAC for message authentication from a weak key (X_W). When Alice wants to authenticate the message μ_A to Bob, then *Bob* first sends a random seed R to Alice, upon which Alice replies with the tag T_A.

follows that the extracted K on Bob's side is close to random and independent of W (and E), and thus T leaks no information on W. However, if Eve manipulates the seed R (for instance replaces it by a value of her choice), then there is no guarantee anymore that K, and thus T, does not leak information on W.

Another and more subtle way for Eve to (potentially) learn information on W is by not manipulating the message, i.e., have $\mu_\mathrm{A} = \mu_\mathrm{B}$, but manipulate the seed R and try to obtain information on W by observing if Bob accepts or not.

4.3.1 Towards Achieving Key-Privacy

We give here some intuition on how we overcome the above privacy issues of DWMAC with respect to the long-term key W. Similarly to our notation T_A and T_B to distinguish between the tag computed by Alice and by Bob, respectively, we write R_A and R_B etc. to distinguish between Alice and Bob's values of R etc., which may be different if Eve actively manipulates communicated messages.

A first approach to prevent leakage through T_A is to one-time-pad encrypt T_A. Let $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^k \to \mathbb{F}_{2^s}$ be a strong extractor (since we merely give a high-level explanation here, we do not specify all parameters of this extractor here). The key for the one-time pad is extracted from X_W by means of Ext, where Alice chooses the seed:

$$\begin{array}{lcl}
\text{Alice} & & \text{Bob} \\
 & \xleftarrow{\quad R \quad} & \\
S \overset{\mathtt{r}}{\leftarrow} \{0,1\}^k & & \\
Z := \mathsf{Ext}(X_W; S) & & \\
Q := T_{\mathrm{A}} \oplus Z & \xrightarrow{\quad S,Q \quad} & \\
 & & Z := \mathsf{Ext}(X_W; S) \\
 & & \text{accept if: } Q = T_{\mathrm{B}} \oplus Z
\end{array}$$

In the above protocol (and also below), we understand T_{A} and T_{B} to be computed as in `DWMAC`. Note that since it is Alice who chooses the seed S and because $H_{\min}(X_W|WE)$ is sufficiently large, Z_{A} is guaranteed to be (close to) random and independent of W (and E), and thus hides all information that T_{A} might leak on W. However, this modification renders the *security* of the protocol invalid. For instance, we cannot exclude that by modifying the seed S appropriately, Eve can enforce $Z_{\mathrm{B}} = T_{\mathrm{B}}$, so that she only needs to send $Q = 0$ to have Bob convinced.

In order to restore security while still preventing information to leak through T_{A}, we let Bob choose a random non-zero "multiplier" for the one-time pad key Z:

$$\begin{array}{lcl}
\text{Alice} & & \text{Bob} \\
 & \xleftarrow{\quad R \quad} & \\
S \overset{\mathtt{r}}{\leftarrow} \{0,1\}^k & & C \overset{\mathtt{r}}{\leftarrow} \mathbb{F}_{2^s}^* \\
Z := \mathsf{Ext}(X_W; S) & & \\
 & \xrightarrow{\quad S \quad} & \\
 & \xleftarrow{\quad C \quad} & \\
\text{abort if } C = 0 & & \\
Q := T \oplus C \cdot Z & \xrightarrow{\quad Q \quad} & \\
 & & Z := \mathsf{Ext}(X_W; S) \\
 & & \text{accept if: } Q = T_{\mathrm{B}} \oplus C \cdot Z
\end{array}$$

Leakage through T_{A} is still prevented since a non-zero multiple of a good one-time-pad key is still a good one-time-pad key. Furthermore, for security, we can intuitively argue as follows. Consider a snapshot of an execution of the protocol after S has been communicated. We give Eve the value T_{A} for free; this only makes her stronger. By the security of the underlying `DWMAC` protocol, we know that it is hard for Eve to guess T_{B}. Now, assuming that there exist two distinct values for C for which Eve can predict the corresponding value $Q_{\mathrm{B}} = T_{\mathrm{B}} \oplus C \cdot Z_{\mathrm{B}}$, it follows immediately that Eve can actually predict T_{B}; a contradiction. Hence, there can be at most one value for Bob's choice of C for which Eve can guess Q_{B} reasonably well.

We point out that the above intuitive reasoning involves *rewinding*; this is fine in the classical setting, but fails when quantum information is involved due to no-cloning (see, e.g., [VDG98]). Thus, in our formal security proof where we allow Eve to maintain a quantum state, we have to reason in a different way. As a consequence, in the actual protocol, Q is computed in a slightly different way.

One issue that we have not yet addressed is that Bob's decision to accept or reject may also leak information on W when $\mu_{\mathrm{A}} = \mu_{\mathrm{B}}$ and Eve modifies one (or both) of the seeds R and S. Note that this is not an issue if $\mu_{\mathrm{A}} \neq \mu_{\mathrm{B}}$ because then, by the security property, Bob rejects with (near) certainty. For instance it might be that changing the first bit of S changes Z or not, depending on what the first bit of X_W is. Thus, by changing the first bit of S and observing Bob's decision, Eve can learn the first bit of X_W, which may give one bit of information on W. The solution to overcome this problem is intuitively very simple: we use MAC not only to authenticate the actual message, but also to authenticate the two seeds R and S. Then, like in the case $\mu_{\mathrm{A}} \neq \mu_{\mathrm{B}}$, if Eve changes one of the seeds then Bob's will reject. Note that this modification introduces a circularity: the key K, which is used to authenticate the seed R (as well as the message and S) is extracted from X_W by means of the seed R. However, it turns out that we can deal with this.

4.4 Our Construction

We now turn to our construction for the message-authentication protocol with long-term-key privacy (Definition 4.1). Let $\mathsf{laExt} : \{0,1\}^n \times \{0,1\}^d \to (\{0,1\}^\ell)^t$ be a (k_K, ε_K) look-ahead extractor. Let $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^v \to \mathbb{F}_{2^q}$ be a (k_Z, ε_Z)-strong extractor. Let $\mathsf{MAC} : (\{0,1\}^\ell)^t \times (\{0,1\}^m \times \{0,1\}^d \times \{0,1\}^v) \to \mathbb{F}_{2^s}$ be an $(\epsilon, \lambda + \epsilon)$ look-ahead-secure MAC for any $\epsilon > 0$. Let X_W be the session key, shared among Alice and Bob. We require that $H_{\min}(X_W|WE_\circ) \geq \max(k_K + q, k_Z)$, and recall from Definition 4.1 that $E_\circ$ denotes Eve's a priori quantum system. Recall from Section 2.4.1 that for an element $x \in \mathbb{F}_{2^n}$ for arbitrary $n \in \mathbb{N}$, $[x]_q$ denotes an arbitrary linear surjective function $\mathbb{F}_{2^n} \to \mathbb{F}_{2^q}$. Protocol `AUTH` is as Protocol 4.2.

In Section 4.6, we show how to instantiate the building blocks to obtain a protocol with reasonable parameters that can be used in a scenario where Eve has classical side information. For the quantum setting, we cannot yet instantiate protocol `AUTH`: we currently do not have a construction for a look-ahead extractor that is provably secure against quantum side information.

Depending on the parameters of an instantiation of `AUTH` and on the bitsize of μ_{A}, it might be better (or even necessary) to authenticate a hash of the tuple (μ_{A}, R, S),

Alice(X_W, μ_{A})		Bob(X_W, μ_{B})
		$R \xleftarrow{\mathrm{r}} \{0,1\}^d$
	$\xleftarrow{R}$	
$K := \mathsf{laExt}(X_W; R)$		$K := \mathsf{laExt}(X_W; R)$
$S \xleftarrow{\mathrm{r}} \{0,1\}^v$		
	$\xrightarrow{S}$	
$Z := \mathsf{Ext}(X_W; S)$		$Z := \mathsf{Ext}(X_W; S)$
$T_{\mathrm{A}} := \mathsf{MAC}_K((\mu_{\mathrm{A}}, R, S))$		$T_{\mathrm{B}} := \mathsf{MAC}_K((\mu_{\mathrm{B}}, R, S))$
		$U \xleftarrow{\mathrm{r}} \mathbb{F}_{2^s}, V \xleftarrow{\mathrm{r}} \mathbb{F}_{2^q}^*$
	$\xleftarrow{U,V}$	
if $V = 0$: abort		
$Q := [U \cdot T_{\mathrm{A}}]_q \oplus V \cdot Z$		
	$\xrightarrow{Q}$	
		accept if: $Q = [U \cdot T_{\mathrm{B}}]_q \oplus V \cdot Z$
		else: abort

Protocol 4.2: Our new four-round message-authentication protocol AUTH.

instead of authenticating the tuple itself. In this case, we let Alice choose a small seed for an almost universal hash function and apply MAC_K to this seed and the hash of the the tuple (μ_{A}, R, S) (with respect to this seed). We will actually make use of this idea in Section 4.6.

Before going into the security proof for protocol AUTH, we resolve here the circularity issue obtained by authenticating the seed R that was used to extract the authentication key K.

Lemma 4.5 *Consider a family of functions* MAC_κ *(indexed by keys* $\kappa \in (\{0,1\}^\ell)^t$*) that is a* $(\xi, \lambda + \xi)$*-look-ahead-secure MAC for any* ξ*. Let* K*,* K'*,* M_{A} *and* M_{B} *be arbitrary random variables and* E *a quantum state, and let the ordered pair* $(K, K') \in (\{0,1\}^\ell)^{2t}$ *satisfy the look-ahead property with parameter* ε *conditioned on* M_{A}*,* M_{B}*,* E *and the event* $M_{\mathrm{A}} \neq M_{\mathrm{B}}$*. Then,*

$$p_{\mathsf{guess}}(\mathsf{MAC}_K(M_{\mathrm{B}}) \,|\, \mathsf{MAC}_{K'}(M_{\mathrm{A}}) M_{\mathrm{A}} M_{\mathrm{B}} E, M_{\mathrm{A}} \neq M_{\mathrm{B}}) < \lambda + t\varepsilon.$$

Note that in the lemma above the messages may depend on the keys, whereas Definition 4.4 considers *fixed* messages.

Proof. We condition on $M_{\mathrm{A}} = m_{\mathrm{A}}$ and $M_{\mathrm{B}} = m_{\mathrm{B}}$ where $m_{\mathrm{A}} \neq m_{\mathrm{B}}$. Because

(K, K') may depend on $(M_{\rm A}, M_{\rm B})$, conditioning on fixed values for the latter implies that (K, K') is not necessarily ε-look-ahead anymore. Let $\varepsilon_{m_{\rm A},m_{\rm B}}$ be the maximum over $i \in [t]$ of the following expression,

$$\varepsilon_{m_{\rm A},m_{\rm B},i} := d_{\textsf{unif}}(K_{i+1} \dots K_t | K'_1 \dots K'_i E, M_{\rm A} = m_{\rm A}, M_{\rm B} = m_{\rm B}).$$

Hence, by Definition 4.2, (K, K') is $\varepsilon_{m_{\rm A},m_{\rm B}}$-look-ahead conditioned on E and the events $M_{\rm A} = m_{\rm A}$ and $M_{\rm B} = m_{\rm B}$. Note that averaging $\varepsilon_{m_{\rm A},m_{\rm B},i}$ over $m_{\rm A}$ and $m_{\rm B}$ (conditioned on them being distinct) results in

$$\varepsilon_i = d_{\textsf{unif}}(K_{i+1} \dots K_t | K'_1 \dots K'_i M_{\rm A} M_{\rm B} E, M_{\rm A} \neq M_{\rm B}) \leq \varepsilon\,.$$

Furthermore, note that by conditioning on fixed and distinct values for $M_{\rm A}$ and $M_{\rm B}$, we fulfill the requirements for MAC look-ahead security from Definition 4.4. I.e. we can conclude that

$$p_{\textsf{guess}}(\textsf{MAC}_K(M_{\rm B}) \,|\, \textsf{MAC}_{K'}(M_{\rm A}) E, M_{\rm A} = m_{\rm A}, M_{\rm B} = m_{\rm B}) < \lambda + \varepsilon_{m_{\rm A},m_{\rm B}}.$$

It now follows that

$$\begin{aligned}
&p_{\textsf{guess}}(\textsf{MAC}_K(M_{\rm B}) \,|\, \textsf{MAC}_{K'}(M_{\rm A}) M_{\rm A} M_{\rm B} E, M_{\rm A} \neq M_{\rm B}) \\
&= \sum_{m_{\rm A},m_{\rm B}} P_{M_{\rm A} M_{\rm B} | M_{\rm A} \neq M_{\rm B}}(m_{\rm A}, m_{\rm B}) \\
&\qquad \cdot p_{\textsf{guess}}(\textsf{MAC}_K(M_{\rm B}) \,|\, \textsf{MAC}_{K'}(M_{\rm A}) E, M_{\rm A} = m_{\rm A}, M_{\rm B} = m_{\rm B}) \\
&< \sum_{m_{\rm A},m_{\rm B}} P_{M_{\rm A} M_{\rm B} | M_{\rm A} \neq M_{\rm B}}(m_{\rm A}, m_{\rm B}) \, (\lambda + \max_{i \in [t]} \varepsilon_{m_{\rm A},m_{\rm B},i}) \\
&\leq \lambda + \sum_{m_{\rm A},m_{\rm B}} P_{M_{\rm A} M_{\rm B} | M_{\rm A} \neq M_{\rm B}}(m_{\rm A}, m_{\rm B}) \sum_{i \in [t]} \varepsilon_{m_{\rm A},m_{\rm B},i} \\
&= \lambda + \sum_{i \in [t]} \sum_{m_{\rm A},m_{\rm B}} P_{M_{\rm A} M_{\rm B} | M_{\rm A} \neq M_{\rm B}}(m_{\rm A}, m_{\rm B}) \, \varepsilon_{m_{\rm A},m_{\rm B},i} \\
&= \lambda + \sum_{i \in [t]} \varepsilon_i \leq \lambda + \sum_{i \in [t]} \varepsilon = \lambda + t\varepsilon.
\end{aligned}$$

This concludes the proof. □

4.5 Proofs of Security and Privacy

In this section we show that protocol `AUTH` fulfills the properties listed in Definition 4.1. First of all, note that it is easy to see from the protocol description that the correctness property is satisfied, we do not elaborate further on this here.

Throughout the proofs, let $E_\circ$ be Eve's quantum side information before executing AUTH. E_i, where $i \in \{1, \dots, 4\}$, represents Eve's (quantum) side information after the ith round of communication, and hence includes the communicated random variables up to this ith round. E represents Eve's side information after executing AUTH, including Bob's decision to accept or reject (E_4 does not include this decision). Furthermore, like in Section 4.3.1, we write R_A and R_B etc. for Alice and Bob's respective values for R etc.

Theorem 4.6 (Security) *If $H_{\min}(X_W|WE_\circ) \geq k_K + q$, then Protocol AUTH fulfills the security property defined in Definition 4.1 with*

$$\delta \leq 3 \cdot 2^{-q} + \frac{1}{2}\sqrt{2^q(\lambda + t\,\varepsilon_K)}.$$

In fact, we will prove a slightly stronger statement than the security statement, which will be of use also in the proof of the key privacy statement. Let $M_\mathrm{A} := (\mu_\mathrm{A}, R_\mathrm{A}, S_\mathrm{A})$ and $M_\mathrm{B} := (\mu_\mathrm{B}, R_\mathrm{B}, S_\mathrm{B})$. We will prove that in protocol AUTH, if $H_{\min}(X_W|WE_\circ) \geq k_K + q$, and conditioned on the event $M_\mathrm{A} \neq M_\mathrm{B}$, Bob rejects except with probability

$$\delta' \leq 3 \cdot 2^{-q} + \frac{1}{2}\sqrt{2^q(\lambda + t\,\varepsilon_K / \Pr[M_\mathrm{A} \neq M_\mathrm{B}])}.$$

Note that this expression reduces to the simpler expression of Theorem 4.6 when proving security, because in that case $\mu_\mathrm{A} \neq \mu_\mathrm{B}$ (by Definition 4.1) which implies that $\Pr[M_\mathrm{A} \neq M_\mathrm{B}] = 1$.

Proof. Consider the phase in protocol AUTH after the second round of communication. Assume that Z_A and T_A are given to the adversary (this will only make her stronger). Let $K_\mathrm{A} := \mathsf{laExt}(X_W; R_\mathrm{A})$ and $K_\mathrm{B} := \mathsf{laExt}(X_W; R_\mathrm{B})$. (Recall that laExt is a (k_K, ε_K) look-ahead extractor.)

From the chain rule, and by subsequently using that R_B and S_A are sampled independently, it follows that

$$H_{\min}(X_W|Z_\mathrm{A}WE_2) \geq H_{\min}(X_W|WE_2) - q \geq H_{\min}(X_W|WE_\circ) - q.$$

By assumption on the parameters, i.e., $H_{\min}(X_W|WE_\circ) \geq k_K + q$, it follows that $(K_\mathrm{B}, K_\mathrm{A})$ is ε_K-look-ahead conditioned on Z_A, W and E_2. In order to apply Lemma 4.5, we additionally condition on the event $M_\mathrm{A} \neq M_\mathrm{B}$. By Lemma 2.52, it is guaranteed that ε_K grows at most by a factor $1/\Pr[M_\mathrm{A} \neq M_\mathrm{B}]$ as a result of this conditioning. We now apply Lemma 4.5 and conclude that

$$p_{\mathsf{guess}}(T_\mathrm{B}|T_\mathrm{A}Z_\mathrm{A}WE_2, M_\mathrm{A} \neq M_\mathrm{B}) \leq \lambda + t\,\varepsilon_K / \Pr[M_\mathrm{A} \neq M_\mathrm{B}].$$

The next step is to view $Q_{\mathrm{B}} := [U_{\mathrm{B}} \cdot T_{\mathrm{B}}]_q \oplus V_{\mathrm{B}} \cdot Z_{\mathrm{B}}$ as the output of a strong extractor, with seed $(U_{\mathrm{B}}, V_{\mathrm{B}})$. Indeed, as guaranteed by Proposition 2.21,the function

$$\begin{aligned} h : \mathbb{F}_{2^s} \times \mathbb{F}_{2^q} \times \mathbb{F}_{2^s} \times \mathbb{F}_{2^q} &\to \mathbb{F}_{2^q} \\ (t, z, u, v) &\mapsto [u \cdot t]_q \oplus v \cdot z, \end{aligned}$$

is a universal hash function (with random seed (u, v)). Thus, we can apply privacy amplification. One subtlety is that in protocol AUTH, V_{B} is random in $\mathbb{F}^*_{2^q}$, rather than in $\mathbb{F}_{2^q}$. Nonetheless, the overall state will be 2^{-q}-close in trace distance to a state where V_{B} would be random over $\mathbb{F}_{2^q}$, and hence, by triangle inequality, the distance-to-uniform increases by an additive term of at most $2 \cdot 2^{-q}$:

$$\begin{aligned} & d_{\mathsf{unif}}(Q_{\mathrm{B}} | U_{\mathrm{B}} V_{\mathrm{B}} T_{\mathrm{A}} Z_{\mathrm{A}} W E_2, M_{\mathrm{A}} \neq M_{\mathrm{B}}) \\ &\qquad \leq \tfrac{1}{2}\sqrt{2^q p_{\mathsf{guess}}(T_{\mathrm{B}} Z_{\mathrm{B}} | T_{\mathrm{A}} Z_{\mathrm{A}} W E_2, M_{\mathrm{A}} \neq M_{\mathrm{B}})} + 2 \cdot 2^{-q} \\ &\qquad \leq \tfrac{1}{2}\sqrt{2^q p_{\mathsf{guess}}(T_{\mathrm{B}} | T_{\mathrm{A}} Z_{\mathrm{A}} W E_2, M_{\mathrm{A}} \neq M_{\mathrm{B}})} + 2 \cdot 2^{-q} \\ &\qquad \leq \tfrac{1}{2}\sqrt{2^q (\lambda + t\,\varepsilon_K / \Pr[M_{\mathrm{A}} \neq M_{\mathrm{B}}])} + 2 \cdot 2^{-q}. \end{aligned}$$

Finally, we have that

$$\begin{aligned} \delta' &= p_{\mathsf{guess}}(Q_{\mathrm{B}} | Q_{\mathrm{A}} W E_3, M_{\mathrm{A}} \neq M_{\mathrm{B}}) \\ &\leq p_{\mathsf{guess}}(Q_{\mathrm{B}} | U_{\mathrm{B}} V_{\mathrm{B}} T_{\mathrm{A}} Z_{\mathrm{A}} W E_2, M_{\mathrm{A}} \neq M_{\mathrm{B}}) \\ &\leq 2^{-q} + d_{\mathsf{unif}}(Q_{\mathrm{B}} | U_{\mathrm{B}} V_{\mathrm{B}} T_{\mathrm{A}} Z_{\mathrm{A}} W E_2, M_{\mathrm{A}} \neq M_{\mathrm{B}}) \\ &\leq 3 \cdot 2^{-q} + \tfrac{1}{2}\sqrt{2^q (\lambda + t\,\varepsilon_K / \Pr[M_{\mathrm{A}} \neq M_{\mathrm{B}}])}. \end{aligned}$$

□

Theorem 4.7 (Long-Term-Key Privacy) *If $H_{\min}(X_W | W E_\circ) \geq \max(q + k_K, k_Z)$, then Protocol AUTH fulfills the long-term-key privacy property defined in Definition 4.1 with*

$$\varepsilon \leq 6 \cdot 2^{-q} + \sqrt{2^q (\lambda + t\,\varepsilon_K)} + \varepsilon_K + 2\,\varepsilon_Z.$$

Proof. We first prove that none of the messages exchanged during the protocol leaks information about W. Then, we show that in our protocol Bob's decision on whether to accept or reject neither leaks information about W.

In the first three rounds of AUTH, Alice and Bob solely exchange independent randomness, so these rounds trivially leak no information about W. The aim in this part of the proof is to show that the fourth message, $Q = [U \cdot T_\text{A}]_q \oplus V \cdot Z$, where T_A could depend on W, indeed keeps W private.

Because R_B is sampled independently of X_W, and by the chain rule, it follows that $H_{\min}(X_W|WE_1[U_\text{A} \cdot T_\text{A}]_q) \geq H_{\min}(X_W|WE_\circ) - q$. By assumption on the parameters in the statement of the theorem, i.e., $H_{\min}(X_W|WE_\circ) \geq q + k_Z$, and by the properties of Ext it follows that

$$d_{\mathsf{unif}}(Z_\text{A}|WE_2[U_\text{A} \cdot T_\text{A}]_q) \leq d_{\mathsf{unif}}(Z_\text{A}|S_\text{A}WE_1[U_\text{A} \cdot T_\text{A}]_q) \leq \varepsilon_Z.$$

By the fact that U_B and V_B are sampled independently, the following also holds

$$d_{\mathsf{unif}}(Z_\text{A}|WE_3[U_\text{A} \cdot T_\text{A}]_q) \leq \varepsilon_Z.$$

Then, by security of the one-time pad (see Proposition 2.53), by the fact that Eve cannot gain information on W by computing Q_B, and by assumption that $\rho_{WE_\circ} = \rho_W \otimes \rho_{E_\circ}$,

$$\tfrac{1}{2}\|\rho_{WE_4} - \rho_W \otimes \rho_{E_4}\|_1 \leq \tfrac{1}{2}\|\rho_{WE_3Q_\text{A}} - \rho_W \otimes \rho_{E_3Q_\text{A}}\|_1 \leq \varepsilon_Z.$$

This completes the first part of the proof.

It remains to show that Bob's decision to accept or reject cannot leak (a substantial amount of) information about W. To show this, we make the following case distinction. In case $\mu_\text{A} \neq \mu_\text{B}$, the security proof applies and Bob rejects except with probability $\delta \leq 3 \cdot 2^{-q} + \frac{1}{2}\sqrt{2^q(\lambda + t\,\varepsilon_K)}$. It now immediately follows that

$$\tfrac{1}{2}\|\rho_{WE_4} - \rho_{WE}\|_1 \leq \delta, \quad \text{and} \quad \tfrac{1}{2}\|\rho_W \otimes \rho_{E_4} - \rho_W \otimes \rho_E\|_1 \leq \delta.$$

Hence, in case $\mu_\text{A} \neq \mu_\text{B}$ (by the triangle inequality),

$$\tfrac{1}{2}\|\rho_{WE} - \rho_W \otimes \rho_E\|_1 \leq \varepsilon_Z + 2\delta.$$

We now turn to the case $\mu_\text{A} = \mu_\text{B}$ and we analyze for two disjoint events. Conditioned on $M_\text{A} \neq M_\text{B}$, the strengthened version of the security statement applies, i.e.,

$$\delta' \leq 3 \cdot 2^{-q} + \tfrac{1}{2}\sqrt{2^q(\lambda + t\,\varepsilon_K/\Pr[M_\text{A} \neq M_\text{B}])},$$

and again by applying the triangle inequality, we obtain

$$\tfrac{1}{2}\|\rho_{WE|M_\text{A} \neq M_\text{B}} - \rho_W \otimes \rho_{E|M_\text{A} \neq M_\text{B}}\|_1 \leq \varepsilon_Z + 2\delta'.$$

Secondly, we analyze for the event $M_\text{A} = M_\text{B}$. Nevertheless, we start this analysis without conditioning on $M_\text{A} = M_\text{B}$. (We'll condition on this event later in the proof.) Since S_{A} is sampled at random and independently of X_W, and since $H_{\min}(X_W|WE_\circ) > k_Z$, it follows that

$$d_{\mathsf{unif}}(Z_{\mathrm{A}}|S_{\mathrm{A}}WE_\circ) < \varepsilon_Z \, .$$

By the chain rule (and the independent choice of S_{A}),

$$H_{\min}(X_W|Z_{\mathrm{A}}WE_2) \geq H_{\min}(X_W|WE_\circ) - q > k_K \, ,$$

and thus

$$d_{\mathsf{unif}}(K_{\mathrm{B}}|R_{\mathrm{B}}Z_{\mathrm{A}}S_{\mathrm{A}}WE_\circ) < \varepsilon_K \, .$$

From the above, and the independent choices of R_{B} and S_{A}, it follows that

$$\tfrac{1}{2}\|\rho_{K_{\mathrm{B}}Z_{\mathrm{A}}R_{\mathrm{B}}S_{\mathrm{A}}WE_\circ} - \rho_U \otimes \rho_{U'} \otimes \rho_{R_{\mathrm{B}}} \otimes \rho_{S_{\mathrm{A}}} \otimes \rho_W \otimes \rho_{E_\circ}\|_1 \leq \varepsilon_K + \varepsilon_Z .$$

where ρ_U is the fully mixed state on $\mathcal{H}_{K_{\mathrm{B}}}$ and $\rho_{U'}$ is the fully mixed state on $\mathcal{H}_{Z_{\mathrm{A}}}$, and therefore that

$$\tfrac{1}{2}\|\rho_{K_{\mathrm{B}}Z_{\mathrm{A}}WE_2} - \rho_U \otimes \rho_{U'} \otimes \rho_W \otimes \rho_{E_2}\|_1 \leq \varepsilon_K + \varepsilon_Z .$$

We now condition on $M_{\mathrm{A}} = M_{\mathrm{B}}$. Note that conditioned on this event, $K_{\mathrm{A}} = K_{\mathrm{B}}$ and $Z_{\mathrm{A}} = Z_{\mathrm{B}}$, and therefore, from here on, we omit the subscripts for these random variables and simply write K and Z. From Lemma 2.52 (noting that whether the event $M_{\mathrm{A}} = M_{\mathrm{B}}$ holds is determined by E_2), we get

$$\tfrac{1}{2}\|\rho_{KZWE_2|M_{\mathrm{A}}=M_{\mathrm{B}}} - \rho_U \otimes \rho_{U'} \otimes \rho_W \otimes \rho_{E_2|M_{\mathrm{A}}=M_{\mathrm{B}}}\|_1 \leq \frac{\varepsilon_K + \varepsilon_Z}{\Pr[M_{\mathrm{A}} = M_{\mathrm{B}}]} .$$

U_{B} and V_{B} are chosen uniformly at random and independent of the rest (and also independently of the event $M_{\mathrm{A}} = M_{\mathrm{B}}$). Furthermore, since E is computed from (KZE_4) alone, it follows that

$$\tfrac{1}{2}\|\rho_{WE|M_{\mathrm{A}}=M_{\mathrm{B}}} - \rho_W \otimes \rho_{E|M_{\mathrm{A}}=M_{\mathrm{B}}}\|_1 \leq \frac{\varepsilon_K + \varepsilon_Z}{\Pr[M_{\mathrm{A}} = M_{\mathrm{B}}]} .$$

We now combine the analyses for the two disjoint events, and conclude that in case $\mu_{\mathrm{A}} = \mu_{\mathrm{B}}$,

$$\begin{aligned}
&\tfrac{1}{2}\|\rho_{WE} - \rho_W \otimes \rho_E\|_1 \\
&\quad \leq \Pr[M_{\mathrm{A}} \neq M_{\mathrm{B}}]\,\tfrac{1}{2}\|\rho_{WE|M_{\mathrm{A}}\neq M_{\mathrm{B}}} - \rho_W \otimes \rho_{E|M_{\mathrm{A}}\neq M_{\mathrm{B}}}\|_1 \\
&\qquad\qquad + \Pr[M_{\mathrm{A}} = M_{\mathrm{B}}]\,\tfrac{1}{2}\|\rho_{WE|M_{\mathrm{A}}=M_{\mathrm{B}}} - \rho_W \otimes \rho_{E|M_{\mathrm{A}}=M_{\mathrm{B}}}\|_1 \\
&\quad = \Pr[M_{\mathrm{A}} \neq M_{\mathrm{B}}]\,(\varepsilon_Z + 2\delta') + \varepsilon_K + \varepsilon_Z \\
&\quad \leq \Pr[M_{\mathrm{A}} \neq M_{\mathrm{B}}]\left[\varepsilon_Z + 6\cdot 2^{-q} + \sqrt{2^q(\lambda + t\,\varepsilon_K/\Pr[M_{\mathrm{A}} \neq M_{\mathrm{B}}])}\right] + \varepsilon_K + \varepsilon_Z \\
&\quad \leq 6\cdot 2^{-q} + \sqrt{2^q(\lambda + t\,\varepsilon_K)} + \varepsilon_K + 2\,\varepsilon_Z .
\end{aligned}$$

Note that we have computed two upper bounds on $\frac{1}{2}\|\rho_{WE} - \rho_W \otimes \rho_E\|_1$, for two distinct cases: $\mu_{\mathrm{A}} \neq \mu_{\mathrm{B}}$ and $\mu_{\mathrm{A}} = \mu_{\mathrm{B}}$. Obviously, the weaker (larger) upper bound holds in both cases, and we finally conclude that

$$\tfrac{1}{2}\|\rho_{WE} - \rho_W \otimes \rho_E\|_1 \leq 6 \cdot 2^{-q} + \sqrt{2^q(\lambda + t\,\varepsilon_K)} + \varepsilon_K + 2\,\varepsilon_Z .$$

□

4.6 Instantiating the Building Blocks

4.6.1 Look-Ahead Extractors against Classical Side Information

Dodis and Wichs [DW09] propose a construction for look-ahead extractors based on *alternating extraction* [DP07]. The construction uses two strong extractors, which are applied in an alternating fashion (we will explain the construction in detail later in this section). The following theorem due to [DW09] states for this construction how the parameters of the two extractors lead to the parameters of the constructed look-ahead extractor.

The security definition of a look-ahead extractor, Definition 4.3, considers quantum side information, represented by E. In this section, we consider the case where the side information E is purely classical. To avoid confusion, we will throughout this section write Z (instead of E) for the adversary's *classical* side information. Note that Z has arbitrary range.

Theorem 4.8 (cf. Theorem 10 in [DW09]) *Given a* $(k_w - 2t\ell, \varepsilon_w)$*-extractor* $\mathsf{Ext}_w : \{0,1\}^{n_w} \times \{0,1\}^\ell \to \{0,1\}^\ell$ *and an* $(n_q - 2t\ell, \varepsilon_q)$*-extractor* $\mathsf{Ext}_q : \{0,1\}^{n_q} \times \{0,1\}^\ell \to \{0,1\}^\ell$*, the construction in [DW09] yields an* $(k_w, t^2(\varepsilon_w + \varepsilon_q))$*-look-ahead extractor*

$$\mathsf{laExt} : \{0,1\}^{n_w} \times \{0,1\}^{n_q+\ell} \to (\{0,1\}^\ell)^t$$

Recently, Reyzin [RWY11] and, independently, Fehr (private communication) discovered that the proof given in [DW09] of Theorem 4.8 is not fully correct, due to a problem with Lemma 1 in [DP07].[2] Fortunately, the proof (of Theorem 4.8) could be fixed as shown in lecture notes by Reyzin [RWY11]. In the remainder of this section, we will explain the alternating extraction construction and reprove it for the case in which the side information is classical. Our proof of Theorem 4.9 (which is then used to prove Theorem 4.8) is inspired by Reyzin's proof [RWY11], but is

[2] To be precise, Dodis and Wichs' proof of look-ahead extraction is also affected by this issue.

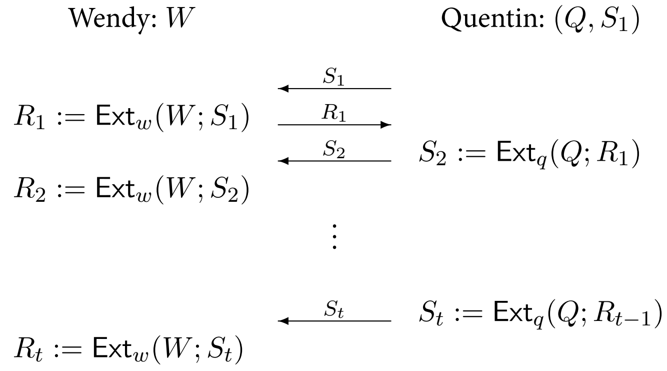
Wendy: W
Quentin: (Q, S_1)
S_1
$R_1 := \mathsf{Ext}_w(W; S_1)$
R_1
S_2
$S_2 := \mathsf{Ext}_q(Q; R_1)$
$R_2 := \mathsf{Ext}_w(W; S_2)$
$\vdots$
S_t
$S_t := \mathsf{Ext}_q(Q; R_{t-1})$
$R_t := \mathsf{Ext}_w(W; S_t)$

Figure 4.1: Alternating extraction explained.

more extensive (we consider auxiliary classical side information and we give formal min-entropy analyses, which are omitted in [RWY11]). Furthermore, our proof uses our Lemma 4.10, which we think is simpler than the corresponding Lemma 6 in [RWY11].

Look-Ahead Extractors from Alternating Extraction

The look-ahead extractor construction is easy to explain. Following [DP07], we identify two parties, Quentin and Wendy. With these parties, we associate the two extractors from Theorem 4.8, Ext_q and Ext_w, as well as two random variables, $Q \in \{0,1\}^{n_q}$ and $W \in \{0,1\}^{n_w}$, respectively. Quentin and Wendy perform *alternating extraction* as follows (see also Figure 4.1). Quentin begins by sending a string $S_1 \in \{0,1\}^\ell$ to Wendy. Wendy then uses S_1 as seed for her extractor: she computes $R_1 := \mathsf{Ext}_w(W; S_1)$ and sends R_1 back to Quentin. Quentin then uses R_1 as seed and computes $S_2 := \mathsf{Ext}_q(Q; R_1)$, and sends this to Wendy again, etc. The procedure stops after Wendy has computed R_t.

The alternating extraction procedure is a construction for a look-ahead extractor in the following way: W is the weakly random source, the tuple $S := (Q, S_1)$ acts as seed, and Wendy's output values $\{R_i\}_{i\in[t]}$ form the output, i.e., $(R_1, \dots, R_t) = \mathsf{laExt}(W; S)$.

Definition 4.3 considers *two* instances of a look-ahead extractor: the one at Bob's side,[3] which is provided with the original seed, and the one at Alice's side, which is

[3] We consider a setting where Alice wants to use the look-ahead extractor to authenticate a message to Bob. Recall that in such a setting Bob samples the seed.

provided with the adversarially modified seed. In terms of our alternating extraction explanation, Quentin and Wendy as described above reside on Bob's side. On Alice's side, we will call the corresponding parties $\widetilde{\text{Q}}$uentin and $\widetilde{\text{W}}$endy. $\widetilde{\text{Q}}$uentin's initial view consists of $(\widetilde{Q}, \widetilde{S}_1, Z)$ (where $(\widetilde{Q}, \widetilde{S}_1)$ equals $\widetilde{S}$ from Definition 4.3) and $\widetilde{\text{W}}$endy's initial view consists of (W, Z). $\widetilde{\text{Q}}$uentin and $\widetilde{\text{W}}$endy exchange ℓ-bit messages which we denote as $\widetilde{S}_i$ and $\widetilde{R}_i$ respectively. These messages are computed from their views in iteration i, which each consists of the party's initial view concatenated with the messages exchanged during alternating extraction.

To prove Theorem 4.8, we let Quentin and Wendy as well as $\widetilde{\text{Q}}$uentin and $\widetilde{\text{W}}$endy perform alternating extraction *synchronously*. In particular, we need Theorem 4.9 as an ingredient, which informally states that the ith message produced by Wendy looks random from the combined view of Quentin and $\widetilde{\text{Q}}$uentin, and *vice versa*. Note that the combined view of Quentin and $\widetilde{\text{Q}}$uentin equals the view of the (implicit) adversary in Definition 4.3.

We will use the following notation for collections of random variables S_i and R_i (as well as $\widetilde{S}_i$ and $\widetilde{R}_i$),

$$S_{[i]} := (S_1, \ldots, S_i) \quad \forall i \in \mathbb{N} \setminus \{0\},$$

and likewise for $R_{[i]}$, $\widetilde{S}_{[i]}$ and $\widetilde{R}_{[i]}$. Furthermore, $S_{[i]}$ for any $i < 1$ denotes the empty list, and likewise for $R_{[i]}$, etc.

Theorem 4.9 *Let ε_q and ε_w as in Theorem 4.8 and let W, Q, $\widetilde{Q}$, S_i, R_i, $\widetilde{S}_i$, $\widetilde{R}_i$ and Z be as described above. If $P_{S_1QWZ} = P_U P_{U'} P_{WZ}$, where P_U and $P_{U'}$ are uniform distributions on $\{0,1\}^\ell$ and $\{0,1\}^{n_q}$ respectively and if $H_{\min}(W|Z) \geq k_w$, then the following inequalities hold for all $i \in [t]$:*

$$d_{\mathsf{unif}}(S_i | W S_{[i-1]} R_{[i-1]} \widetilde{R}_{[i-1]} \widetilde{S}_{[i-1]} Z) \leq (\varepsilon_q + \varepsilon_w)(i-1) \tag{4.1}$$

$$d_{\mathsf{unif}}(R_i | Q R_{[i-1]} S_{[i]} \widetilde{R}_{[i-1]} \widetilde{S}_{[i]} \widetilde{Q} Z) \leq (\varepsilon_q + \varepsilon_w)(i-1) + \varepsilon_w, \tag{4.2}$$

Note that we require Q to be uniformly distributed; this stems from the parameters of Ext_q, which we adopt from Theorem 4.8. By adapting the parameters of Ext_q appropriately, alternating extraction also works when Q does not have full min-entropy (cf. [DW09, RWY11]). Nevertheless, since we anyway do not need this more general case, we find it simpler to state it as above.

As in [RWY11], the proof is based on the *conditional independence* of Q and W (when conditioned on the messages exchanged in the alternating-extraction protocol). This independence is crucial for inequalities (4.1) and (4.2) to hold because

S_i (R_i) is extracted from Q (W) via a seed that is computed from W (Q), and it is well known that for an extractor to work properly the seed must be (essentially) independent from the source.

Consider the general setting where two parties, holding independent random variables X and Y respectively, interact by exchanging messages, where each message is computed from the sender's random variable (i.e., *either* X or Y) and previously exchanged messages. Then, it is well known (and straightforward to prove) that $X \leftrightarrow M \leftrightarrow Y$ holds, where M represents the collection of the exchanged messages. Observe that alternating extraction (when viewing Wendy and $\widetilde{\text{W}}$endy as a single party and Quentin and $\widetilde{\text{Q}}$uentin as well) is a particular instance of the above general setting. Note that the (classical) side information Z should be treated as being part of M; it can be thought of an initial message that is sent from $\widetilde{\text{W}}$endy to $\widetilde{\text{Q}}$uentin.

To prove Theorem 4.9 we will use the following lemma, which is a corrected and extended version of Lemma 1 from [DP07].

Lemma 4.10 *Let A, B, C be arbitrary random variables over respectively $\mathcal{A}, \mathcal{B}, \mathcal{C}$ such that $A \leftrightarrow B \leftrightarrow C$. Then, for any function $f : \mathcal{A} \times \mathcal{C} \to \mathcal{Z}$ it holds that*

$$d_{\mathsf{unif}}(f(A,C)|BC) \leq d_{\mathsf{unif}}(f(A,U)|BU) + d_{\mathsf{unif}}(C|B)$$

where U is an independent random variable uniformly distributed over $\mathcal{C}$.

Proof.

$$\begin{aligned} d_{\mathsf{unif}}(C|B) &= \tfrac{1}{2}\|\rho_{CB} - \rho_U \otimes \rho_B\|_1 \\ &= \tfrac{1}{2}\|\rho_{CBA} - \rho_U \otimes \rho_{BA}\|_1 \\ &\geq \tfrac{1}{2}\|\rho_{f(A,C)BC} - \rho_{f(A,U)BU}\|_1 \end{aligned}$$

where the first equality is by definition of the trace distance to uniform, the second equality follows from the Markov property, and the inequality is by the fact that the trace distance cannot increase under quantum operations; see Theorem 2.50. Finally, the claim follows by applying triangle inequality. □

Proof of Theorem 4.9. We prove the statement by induction on i. Inequality (4.1) obviously holds for $i = 1$,

$$d_{\mathsf{unif}}(S_i|WS_{[i-1]}R_{[i-1]}\widetilde{S}_{[i-1]}\widetilde{R}_{[i-1]}Z)\big|_{i=1} = d_{\mathsf{unif}}(S_1|WZ) = 0.$$

The first half of the induction step is to show that, if (4.1) holds for i (the induction hypothesis), then (4.2) must hold for i, i.e.

$$d_{\mathsf{unif}}(R_i|QR_{[i-1]}S_{[i]}\widetilde{R}_{[i-1]}\widetilde{S}_{[i]}\widetilde{Q}Z) \leq (\varepsilon_q + \varepsilon_w)(i-1) + \varepsilon_w.$$

The (trace) distance to uniform cannot increase when applying the same operation to both states (in this case: removing W)

$$\begin{aligned} d_{\mathsf{unif}}(S_i|S_{[i-1]}R_{[i-1]}\widetilde{R}_{[i-1]}\widetilde{S}_{[i-1]}Z) &\leq d_{\mathsf{unif}}(S_i|WS_{[i-1]}R_{[i-1]}\widetilde{R}_{[i-1]}\widetilde{S}_{[i-1]}Z) \\ &\leq (\varepsilon_q+\varepsilon_w)(i-1). \end{aligned} \tag{4.3}$$

The following bound holds on the conditional min-entropy of W,

$$\begin{aligned} H_{\min}(W|S_{[i-1]}R_{[i-1]}\widetilde{R}_{[i-1]}\widetilde{S}_{[i-1]}Z) &\geq H_{\min}(W|S_{[i-1]}R_{[i-1]}\widetilde{R}_{[i-1]}\widetilde{S}_{[i-1]}QZ) \\ &= H_{\min}(W|S_1R_{[i-1]}\widetilde{R}_{[i-1]}QZ) \\ &\geq H_{\min}(W|S_1QZ) - H_{\max}(R_{[i-1]}\widetilde{R}_{[i-1]}) \\ &= H_{\min}(W|Z) - 2(i-1)\ell \\ &\geq k_w - 2(t-1)\ell, \end{aligned}$$

where the first inequality holds by strong subadditivity, the first equality holds because $W \leftrightarrow R_{[i-1]}\widetilde{R}_{[i-1]}S_1QZ \leftrightarrow \widetilde{S}_{[i-1]}S_{[i-1]} \setminus \{S_1\}$ (which holds because of the way the S_i and $\widetilde{S}_i$ are computed), the second inequality is the chain rule and the second equality holds because $P_{WZS_1Q} = P_{WZ}P_UP_Q$. The definition of Ext_w then guarantees that

$$d_{\mathsf{unif}}(\mathsf{Ext}_w(W;U)|US_{[i-1]}R_{[i-1]}\widetilde{S}_{[i-1]}\widetilde{R}_{[i-1]}Z) \leq \varepsilon_w, \tag{4.4}$$

for an independent and uniform seed U.

Given that $W \leftrightarrow S_{[i-1]}R_{[i-1]}\widetilde{S}_{[i-1]}\widetilde{R}_{[i-1]}Z \leftrightarrow Q$ is a Markov chain (as explained before Lemma 4.10), it follows that $W \leftrightarrow S_{[i-1]}R_{[i-1]}\widetilde{S}_{[i-1]}\widetilde{R}_{[i-1]}Z \leftrightarrow S_i$ holds as well, since S_i is a function of Q and R_{i-1}. Now, given the latter Markov chain and (4.3) and (4.4), we can apply Lemma 4.10 with $A = W$, $B = S_{[i-1]}R_{[i-1]}\widetilde{S}_{[i-1]}\widetilde{R}_{[i-1]}Z$, $C = S_i$ and $U = U$, which guarantees that

$$d_{\mathsf{unif}}(\mathsf{Ext}_w(W;S_i)|R_{[i-1]}S_{[i]}\widetilde{S}_{[i-1]}\widetilde{R}_{[i-1]}Z) \leq (\varepsilon_q+\varepsilon_w)(i-1)+\varepsilon_w.$$

Because it holds that $Q \leftrightarrow S_{[i]}R_{[i-1]}\widetilde{S}_{[i-1]}\widetilde{R}_{[i-1]}Z \leftrightarrow W$, we may additionally condition on Q in the expression above without increasing the trace distance to uniform. Furthermore, since both $\widetilde{Q}$ and $\widetilde{S}_i$ can be computed from the random variables that are already being conditioned on, we can also condition on them "for free." Since $R_i := \mathsf{Ext}_w(W;S_i)$ we obtain

$$d_{\mathsf{unif}}(R_i|QR_{[i-1]}S_{[i]}\widetilde{Q}\widetilde{S}_{[i]}\widetilde{R}_{[i-1]}Z) \leq (\varepsilon_q+\varepsilon_w)(i-1)+\varepsilon_w,$$

which is (4.2) for i and concludes the proof of the first half of the induction step.

The second half of the induction step is to take the expression above as the induction hypothesis and show that if hypothesis is true, then (4.1) must hold for $i+1$, i.e.

$$d_{\mathsf{unif}}(S_{i+1}|WS_{[i]}R_{[i]}\widetilde{S}_{[i]}\widetilde{R}_{[i]}Z) \leq (\varepsilon_q + \varepsilon_w)i.$$

This second part is essentially a "mirror image" of the above part.

By an elementary property of the trace distance, the distance to uniform cannot increase when applying a function to both states (it this case: removing systems Q and $\widetilde{Q}$):

$$\begin{aligned} d_{\mathsf{unif}}(R_i|R_{[i-1]}S_{[i]}\widetilde{R}_{[i-1]}\widetilde{S}_{[i]}Z) &\leq d_{\mathsf{unif}}(R_i|QR_{[i-1]}S_{[i]}\widetilde{Q}\widetilde{S}_{[i]}\widetilde{R}_{[i-1]}Z) \\ &\leq (\varepsilon_q + \varepsilon_w)(i-1) + \varepsilon_w. \end{aligned} \tag{4.5}$$

The following bound holds on the conditional min-entropy of Q,

$$\begin{aligned} H_{\min}(Q|R_{[i-1]}S_{[i]}\widetilde{R}_{[i-1]}\widetilde{S}_{[i]}Z) &\geq H_{\min}(Q|WR_{[i-1]}S_{[i]}\widetilde{R}_{[i-1]}\widetilde{S}_{[i]}Z) \\ &= H_{\min}(Q|WZS_{[i]}\widetilde{S}_{[i]}) \\ &\geq H_{\min}(Q|WZS_1) - H_{\max}(\widetilde{S}_{[i]}S_{[i]} \setminus \{S_1\}) \\ &= n_q - (2i-1)\ell \\ &\geq n_q - (2t-1)\ell, \end{aligned}$$

where the first inequality holds by strong subadditivity, the first equality holds because $Q \leftrightarrow WZS_{[i]}\widetilde{S}_{[i]} \leftrightarrow R_{[i-1]}\widetilde{R}_{[i-1]}$, the second inequality is the chain rule, the second equality holds because $P_{WZS_1Q} = P_{WZ}P_UP_{U'}$ and the last inequality follows because $i \leq t$. The definition of Ext_q then guarantees that

$$d_{\mathsf{unif}}(\mathsf{Ext}_q(Q;U)|UR_{[i-1]}S_{[i]}\widetilde{R}_{[i-1]}\widetilde{S}_{[i]}Z) \leq \varepsilon_q, \tag{4.6}$$

for an independent and uniform seed U.

Note that from the fact that $Q \leftrightarrow R_{[i-1]}S_{[i]}\widetilde{R}_{[i-1]}\widetilde{S}_{[i]}Z \leftrightarrow W$, it follows that $Q \leftrightarrow R_{[i-1]}S_{[i]}Z \leftrightarrow R_i$ since R_i is a function of W and S_i. Given this latter Markov chain and (4.5) and (4.6), we can apply Lemma 4.10 with $A = Q$, $B = S_{[i]}R_{[i-1]}\widetilde{S}_{[i]}\widetilde{R}_{[i-1]}Z$, $C = R_i$ and $U = U$, which guarantees that

$$d_{\mathsf{unif}}(\mathsf{Ext}_q(Q;R_i)|R_{[i]}S_{[i]}\widetilde{R}_{[i-1]}\widetilde{S}_{[i]}Z) \leq (\varepsilon_q + \varepsilon_w)i.$$

Because it holds that $Q \leftrightarrow S_{[i]}R_{[i]}\widetilde{S}_{[i]}\widetilde{R}_{[i-1]}Z \leftrightarrow W$, we may additionally condition on W without increasing the distance to uniform. Furthermore, we may condition on $\widetilde{R}_i$ as well since it is computed as a function of W and $\widetilde{S}_i$,

$$d_{\mathsf{unif}}(\mathsf{Ext}_q(Q; R_i)|WR_{[i]}S_{[i]}\widetilde{R}_{[i]}\widetilde{S}_{[i]}Z) \leq (\varepsilon_q + \varepsilon_w)i.$$

Finally, we obtain (4.1) for $i+1$ by noting that $S_{i+1} := \mathsf{Ext}_q(Q; R_i)$ and this proves the second half of the induction step. □

Finally, we prove the main claim. The proof below is essentially the same as the proof of Theorem 9 in [DW09], but then adapted to our notation.

Proof of Theorem 4.8. We need to prove that

$$d_{\mathsf{unif}}(R_{i+1}\dots R_t|\widetilde{R}_{[i]}S\widetilde{S}Z) \leq t^2(\varepsilon_q + \varepsilon_w).$$

Note that Definition 4.3 already requires that S_1 and Q are uniformly distributed and independent of W and Z and that $H_{\min}(W|Z) \geq k_w$, so Theorem 4.9 applies.

Consider (4.2) from Theorem 4.9, i.e.

$$d_{\mathsf{unif}}(R_i|QR_{[i-1]}S_{[i]}\widetilde{R}_{[i-1]}\widetilde{S}_{[i]}\widetilde{Q}Z) \leq (\varepsilon_q + \varepsilon_w)(i-1) + \varepsilon_w,$$

Let us remove the conditioning on $S_{[i]}$ and $\widetilde{S}_{[i]}$ except for S_1 and $\widetilde{S}_1$, by elementary properties of the trace distance this cannot increase the distance. As mentioned on page 149, $S := (Q, S_1)$ and similarly $\widetilde{S} := (\widetilde{Q}, \widetilde{S}_1)$, so we replace $(Q, S_1, \widetilde{Q}, \widetilde{S}_1)$ by $(S, \widetilde{S})$. Furthermore, we may obviously append independent uniform randomness without increasing the distance-to-uniform:

$$d_{\mathsf{unif}}(R_iU_{\ell(t-i)}|R_{[i-1]}\widetilde{R}_{[i-1]}S\widetilde{S}Z) \leq (\varepsilon_q + \varepsilon_w)(i-1) + \varepsilon_w, \tag{4.7}$$

We will evaluate (4.7) using the substitutions $i \to i+1$ up to $i \to t$:

$$\begin{aligned} d_{\mathsf{unif}}(R_{i+1}U_{\ell(t-i-1)}|R_{[i]}\widetilde{R}_{[i]}S\widetilde{S}Z) &\leq (\varepsilon_q + \varepsilon_w)i + \varepsilon_w, \\ d_{\mathsf{unif}}(R_{i+2}U_{\ell(t-i-2)}|R_{[i+1]}\widetilde{R}_{[i+1]}S\widetilde{S}Z) &\leq (\varepsilon_q + \varepsilon_w)(i+1) + \varepsilon_w, \\ &\vdots \\ d_{\mathsf{unif}}(R_t|R_{[t-1]}\widetilde{R}_{[t-1]}S\widetilde{S}Z) &\leq (\varepsilon_q + \varepsilon_w)(t-1) + \varepsilon_w. \end{aligned}$$

By recursively applying the triangle inequality to these expressions (a “hybrid argument”) we may conclude that

$$d_{\mathsf{unif}}(R_{i+1}\dots R_t|R_{[i]}\widetilde{R}_{[i]}S\widetilde{S}Z) \leq \tfrac{1}{2}t(t-1)(\varepsilon_q+\varepsilon_w) + (t-1)\varepsilon_w \leq t^2(\varepsilon_q+\varepsilon_w)$$

Finally, we obtain the claim simply by removing the conditioning on $\widetilde{R}_{[i]}$[4]. □

[4] we thus actually prove a slightly stronger statement, i.e., that the claim still holds when conditioning additionally on $R_{[i]}$

Parameters of an Explicit Look-Ahead Extractor

Dodis and Wichs use the explicit strong extractor from [GUV09] (see also Theorem 2.30) to instantiate the extractor in Theorem 4.8, and achieve the following parameters.

Theorem 4.11 (Theorem 11 in [DW09]) *For all integers $n \geq k$ and all $\varepsilon > 0$ there exist (k, ε)-look-ahead extractors* $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to (\{0,1\}^\ell)^t$ *as long as*

$$\begin{aligned} k &\geq 2(t+2)\max(\ell, O(\log(n) + \log(t) + \log(1/\varepsilon))) \\ &\geq O(t(\ell + \log(n) + \log(t) + \log(1/\varepsilon))), \end{aligned}$$

and $d \geq O(t(\ell + \log(n) + \log(t) + \log(1/\varepsilon)))$.

I.e., when neglecting logarithmic terms, k and d are both of order $t\ell$, the bit-size of the range of the extractor.

4.6.2 Look-Ahead Extractors and Quantum Side Information?

In the *Eurocrypt* paper [BF11] on which the present chapter is based, we claimed that one can obtain a look-ahead extractor that is secure against *quantum* side information simply by replacing the classical strong extractors in the original construction by extractors against quantum side information, and furthermore that the original proof strategy can also be used in the quantum setting. Unfortunately, during the preparation of this thesis we noticed that we have overlooked a subtle issue that renders the original proof strategy invalid for the quantum case. Although the alternating-extraction construction *could* still work in the quantum setting, we currently do not have a proof for it. We leave it as an important open problem.

Let us briefly explain here why the proof strategy for the classical setting does not apply in the quantum setting. Recall that according to Definition 4.3 the adversary creates $\widetilde{S} = (\widetilde{Q}, \widetilde{S}_1)$ given $S = (Q, S_1)$ and E, and that this process may involve a measurement on E, which then collapses to the state E'. This latter state E' may in particular include Q, and it typically depends on W as well. It is not clear how to generalize Lemma 4.10 to include this quantum side information. Moreover, the proof for the classical case makes statements about a probability space in which Z and $\widetilde{S}$, which is computed from Z, exist simultaneously. In the quantum setting, however, the *original* quantum state E does not exist anymore after it is measured (to produce $\widetilde{S}$); it collapses to the post-measurement state E', which is not guaranteed to have the necessary properties (like independence of Q).

4.6.3 Security and Instantiation of the MAC

To construct a MAC with look-ahead security, we adopt the construction given in [DW09]. Because our look-ahead security definition, Definition 4.4, is slightly weaker than the one given in [DW09] (in that both μ_{A} and μ_{B} are fixed), we obtain a better security parameter, as argued below.

With respect to a different aspect, the requirement on the MAC for constructing our protocol AUTH is somewhat stronger, because we need a "universal" MAC which is $(\epsilon, \lambda + \epsilon)$-look-ahead secure *for any* $\epsilon \geq 0$ (and some λ). (This requirement stems from the proof of Lemma 4.5.) It turns out that the construction from [DW09] satisfies this property.

Proposition 4.12 *For any positive integers m and ℓ, there exists a family of functions $\{\mathsf{MAC}_\kappa : \{0,1\}^m \to \{0,1\}^s\}$, indexed by keys $\kappa \in (\{0,1\}^\ell)^t$, that is $(\varepsilon, 2^{-\ell} + \varepsilon)$ look-ahead secure for any $\varepsilon > 0$, where $t = 4m$ and $s = 2m\ell$.*

For completeness, we very briefly describe the idea of the construction here. The function $\mathsf{MAC}_\kappa(\mu)$ outputs some of the blocks κ_i of the key $\kappa = (\kappa_1, \ldots, \kappa_t)$; where the choice of this subset is determined by μ. Furthermore, the construction guarantees that for any two distinct messages μ and μ', there exists an index $i_\circ < t$ such that $\mathsf{MAC}_\kappa(\mu)$ outputs more blocks κ_i with $i > i_\circ$ than $\mathsf{MAC}_\kappa(\mu')$ does. From the look-ahead property, it follows that given $\kappa'_1, \ldots, \kappa'_{i_\circ}$, the remaining blocks $\kappa_{i_\circ+1}, \ldots, \kappa_t$ are (ε-close to) random. Then, from the choice of $i_\circ$ and from the chain rule we conclude that when given $\mathsf{MAC}_{\kappa'}(\mu')$, the tag $\mathsf{MAC}_\kappa(\mu)$ still contains at least (nearly) ℓ bits of min-entropy.

Since the security of the MAC follows more or less directly from the look-ahead property (and an application of the chain rule), this construction is secure in the presence of quantum side information when the underlying look-ahead extractor is secure against quantum side information.

When comparing our Proposition 4.12 with Lemma 15 in Appendix E.3 of [DW09], our modification of fixing both μ_{A} and μ_{B} before executing DWMAC overcomes the need for a union bound over all possible messages μ_{B} and hence saves us a factor of 2^m.

4.6.4 Instantiating Protocol AUTH

We will instantiate protocol AUTH for the case of classical side information. Before doing so, we first need to slightly modify the protocol. Because the alternating-extraction construction that we use to instantiate laExt requires a relatively large seed, we cannot let Alice authenticate the tuple (μ_{A}, R, S) directly. Instead, Alice will sample a seed and for an almost universal hash function, and authenticates the seed and the hash of (μ_{A}, R, S). We will make use of the well-known polynomial construction for an almost universal hash function (see, e.g., [TSSR10]); for some field $\mathbb{F}$ and b a positive integer, let

$$\begin{array}{cccc} h: & \mathbb{F}^b \times \mathbb{F} & \to & \mathbb{F} \\ & (x_1, \ldots, x_b; \alpha) & \mapsto & \sum_{i=1}^{b} x_i \alpha^{b-i}. \end{array}$$

For α, the seed, randomly chosen from $\mathbb{F}$, the probability that two distinct inputs $x, x' \in \mathbb{F}^b$ collide is $p_{\mathsf{col}} := (b-1)/|\mathbb{F}|$.

This hashing-modification to AUTH will affect its security and privacy. We take care of this simply by adding p_{col} to the security and $2\, p_{\mathsf{col}}$ to the privacy upper bound. The latter factor of two comes from the triangle inequality, which appears because privacy (as defined in Definition 4.1) is a distance between two states.

We now combine Theorem 4.6, Theorem 4.7, Theorem 2.30, Theorem 4.11 and Proposition 4.12 and make use of the hashing modification explained above in order to obtain a lower bound on k, the min-entropy required by AUTH, in terms of desired security and privacy parameters and the bitsize of the message to be authenticated.

Corollary 4.13 *For any integers $n \geq k$, m and any $\varepsilon > 0$ and any $0 < \delta \leq \varepsilon/8$, we can construct an efficient four-round $(n, k, m, \delta, \varepsilon)$ message-authentication protocol with long-term-key privacy as long as (asymptotically)*

$$k = O\Big(\log(1/\varepsilon) + \big(\log(1/\delta) + \log(m')\big) \cdot \big(\log(1/\delta) + \log(m') + \log(n)\big)\Big),$$

where

$$m' = m + O\Big(\log(1/\varepsilon) + \big(\log(1/\delta) + \log(m')\big) \cdot \big(\log(1/\delta) + \log(m') + \log(n)\big)\Big).$$

Proof. We start by computing suitable parameters for the almost universal hash function. Let $\mathbb{F} := \mathbb{F}_{2^c}$ for a positive integer c, and let m' be the bitsize of the tuple (μ, R, S), i.e., $m' = m + d + v$. Hence, $b = m'/c$,[5] and $p_{\mathsf{col}} = 2^{-c}(m'/c - 1) \leq 2^{-c} m'$.

As required by the security and privacy proofs, $k > \max(q + k_K, k_Z)$. We first analyze k_K. Let $\delta' := 3 \cdot 2^{-q} + \frac{1}{2}\sqrt{2^q(2^{-\ell} + t\,\varepsilon_K)} + 2^{-c} m'$ (this expression

[5] Here, we assume that m' is an integer multiple of c. Note that this can always be achieved by zero-padding m'.

originates from combining Theorem 4.6, Proposition 4.12 and p_{col}). To simplify matters, we choose $q = \ell/2$, $c = \ell/2 + \log m'$ and $\varepsilon_K = 2^{-\ell}/t$ and we obtain

$$\begin{aligned}\delta' &= 3 \cdot 2^{-\ell/2} + \frac{1}{2}\sqrt{2^{\ell/2}(2 \cdot 2^{-\ell})} + 2^{-(\ell/2 + \log m')}\, m' \\ &= 3 \cdot 2^{-\ell/2} + 2^{-\frac{1}{2}-\frac{\ell}{4}} + 2^{-\ell/2} \lessapprox 2^{-\ell/4} \quad \text{(for large enough } \ell\text{)}.\end{aligned}$$

Because δ' is an upper bound for the security of AUTH, a sufficient condition to achieve the desired security level δ is when $\delta' \leq \delta$. Hence, we choose

$$\ell \geq 4\log(1/\delta).$$

The actual message to be authenticated consists of the seed and the hash value and therefore has bit-length $2c$. Then, by Proposition 4.12 we have that $t = 4(2c) = 4\ell + 8\log m' \geq 16\log(1/\delta) + 8\log m'$. We substitute this into the expression for ε_K:

$$\varepsilon_K \leq \delta^4/(16\log(1/\delta) + 8\log m').$$

Next, we plug this into the bound for k from Theorem 4.11. This yields

$$k_K = O\Big(\big(\log(1/\delta) + \log(m')\big) \cdot \big(\log(1/\delta) + \log(m') + \log(n)\big)\Big).$$

We now analyze k_Z. Let

$$\begin{aligned}\varepsilon' &:= 6 \cdot 2^{-q} + \sqrt{2^q(2^{-\ell} + t\,\varepsilon_K)} + \varepsilon_K + 2\,\varepsilon_Z + 2^{-c+1}m' \\ &= 2\delta' + \delta'^4/t + 2\varepsilon_Z + 2^{-c+1}m'\end{aligned}$$

be the upper bound on the privacy of AUTH (the expression follows from combining Theorem 4.7, Proposition 4.12 and p_{col}). To achieve the desired privacy ε, it suffices that $\varepsilon' \leq \varepsilon$. By substituting $\delta' = \delta$ and solving for ε_Z, we obtain $\varepsilon_Z \leq \frac{1}{2}\varepsilon - \delta - \frac{1}{2t}\delta^4 - 2^{-\ell/2} \leq \frac{1}{2}\varepsilon - \delta - \frac{1}{2t}\delta^4 - \delta^2$. From the latter expression, we see why we cannot choose δ arbitrarily large, compared to ε, because an upper bound for ε_Z should of course not be negative. Note that this parameter-dependency is not surprising; it stems from the fact that the privacy proof makes use of the security proof. Therefore, we choose $0 < \delta \leq \varepsilon/8$, such that $\varepsilon_Z \leq \frac{\varepsilon}{2} - \frac{\varepsilon}{8} - \frac{\varepsilon^4}{2^{13}t} - \frac{\varepsilon^2}{64}$. Lower bounding the RHS yields the simpler expression

$$\varepsilon_Z \leq \varepsilon/4.$$

Substituting this into the bound for k from Theorem 4.11 gives

$$k_Z = O\big(\log(1/\delta) + \log(n) + \log(1/\varepsilon)\big)$$

We upper-bound $\max(q + k_K, k_Z)$ by the sum $q + k_K + k_Z$:

$$\begin{aligned} k &\geq 2\log 1/\delta + k_K + k_Z \\ &= O\Big(\log(1/\varepsilon) + \big(\log(1/\delta) + \log(m')\big) \cdot \big(\log(1/\delta) + \log(m') + \log(n)\big)\Big). \end{aligned}$$

Remember that $m' = (m + d + v)$, where $v = O\big(\log(n) + \log(1/\varepsilon_Z)\big) = O\big(\log(n) + \log(1/\varepsilon)\big)$ and

$$d = O\Big(\big(\log(1/\delta) + \log(m')\big) \cdot \big(\log(1/\delta) + \log(m') + \log(n)\big)\Big).$$

□

4.7 The Fuzzy Case

Up to here, we assumed a scenario where Alice and Bob share *identical* copies of the session key X_W. Let us now consider the "fuzzy" case, where Alice and Bob hold keys that are only close in some sense, but not necessarily equal. This kind of scenario naturally arises when Alice and Bob obtain their session keys in the presence of noise. For simplicity and with our application (Section 4.8) in mind, we use the Hamming distance to measure closeness between keys.

Consider the following simple approach. Let Bob's key be called X_W. Before executing the authentication protocol, Bob sends some error-correcting information (like the syndrome of X_W with respect to some error-correcting code) to Alice, so that she can correct the errors in her key, X'_W. Since Eve has full control over the communication channel, she can also modify this error-correction information. In this case Alice might not correct X'_W successfully, in which case our protocol is not guaranteed to be secure. However, as stated in Theorem 22 in [DW09], this approach is secure (in the classical setting) if one uses alternating-extraction-based instantiations of look-ahead extractors. (Note that the parameters change slightly compared to the non-fuzzy case, to take into account the min-entropy loss due to the error correction information.) For this solution to work it is important that X_W has sufficient min-entropy when given Eve's (classical) side information, and that Bob sends the error-correcting information to Alice (i.e., the error-correction information must be sent in the same direction as the seed for the look-ahead extractor).

Because we currently do not have a provably secure construction for a look-ahead extractor against quantum side information, we cannot say whether the approach above also works in the setting where Eve is allowed to have quantum side information. This remains an open question that needs to be solved before protocol AUTH can be used to improve the quantum protocol $\mathtt{QID}^{+}$.

One subtlety is that the error-correcting information must not leak information about W, to preserve the privacy property. Exactly this problem is addressed in [DS05], and is generalized to the quantum setting in [FS09]. Note that it is straightforward to upper bound the min-entropy loss in X_W due to error correction: by the chain rule this is at most the bitsize of the error-correction information.

Finally, we want to make a remark about how this min-entropy loss (caused by sending the error-correction information) is incorporated in the parameters of Theorem 22 in [DW09]: Ext_q needs to be an $(n_q - (2\ell + \alpha)t, \varepsilon_q)$-extractor,[6] where α is the bitsize of the error-correction information. In words, there is a loss of αt in the first parameter, where one would expect only a loss of α. To us it seems that the factor t in front of α is not necessary; it is merely a consequence of the proof strategy of Theorem 22, which uses the alternating-extraction theorem (Theorem 9 in [DW09]) as a black box.

Furthermore, it seems that the requirement on the conditional min-entropy of W_A in Theorem 22 (from [DW09]) is not necessary; it is also not used in the proof.

4.8 Application: Password-Based Identification

We sketch here how an instantiation of protocol AUTH that a) is secure when Eve has quantum side information about the weak key, and b) is still secure in the fuzzy case, would lead to a truly password-based identification protocol in the bounded quantum storage model with security against man-in-the-middle attacks. We want to stress that to achieve a) and to be able to verify b), we still miss one building block, i.e., look-ahead extractors against quantum side information. Furthermore, recall that the protocols QID and $\mathtt{QID}^{+}$ are briefly explained in Section 2.11 (for details, the reader is referred to [DFSS07]). In particular, recall that the need for an additional high-entropy key in $\mathtt{QID}^{+}$ stems from the use of an extractor MAC, which is used to authenticate all classical communication.

Our idea of obtaining security against man-in-the-middle attacks without a high-entropy key is now simply to do the authentication of the classical communication by applying protocol AUTH when using $x_{\mathcal{I}_w}$ as weak session key. Our privacy property guarantees that the authentication does not leak information on the password w.

[6] For comparison: in Theorem 4.8, the non-fuzzy case, Ext_q is a $(n_q - 2\ell t, \varepsilon_q)$-extractor.

We stress that previous protocols for authentication based on weak keys would (potentially) leak here information on w.

As in Section 2.11, we abbreviate the user and server by U and S respectively. If the quantum communication is noisy (which it is in realistic scenarios) or if the man-in-the-middle attacker modifies some of the qubits (but few enough so that he is not detected) or θ, then U's and S's version of $x_{\mathcal{I}_w}$ are not identical. Thus, we indeed require that AUTH is secure in the fuzzy case.

If the analysis of the fuzzy case for the case of classical side information would more or less directly carry over to the quantum setting, then this would mean that we need a lower bound on the min-entropy of S's version of $x_{\mathcal{I}_w}$ (when given the adversary's side information). Although the analysis of Damgård *et al.* only guarantees min-entropy in U's version, we can slightly modify the protocol to also guarantee lower-bounded min-entropy on S's side. Instead of measuring the BB84 qubits in basis $c(w)$, S measures them in a *random* basis $\hat{\theta}$ and announces the difference $r = c(w) \oplus \hat{\theta}$. Then, U and S update the code c by shifting every code word by r, so that with respect to the updated code c', S has actually measured the BB84 qubits in basis $c'(w)$. This trick has also been used in [DFL$^+$09], though for a different reason, and has no real effect on the analysis of the protocol. However, since S now also measures in a random basis, we can apply the uncertainty relation of [DFR$^+$07] to get a lower bound on the min-entropy on S's side.

4.9 Open Problem

The main open problem of this chapter is showing the existence of (efficient) look-ahead extractors that are secure against quantum side information. It remains possible that the alternating-extraction construction also works against quantum side information, but it might also be the case that totally different techniques are needed.

5

Hybrid Security of Password-Based Identification

The content of this chapter is based on joint work with Serge Fehr, Carlos González-Guillén and Christian Schaffner [BFGS12].

Chapter Contents

5.1 Introduction 164

5.1.1 A New Uncertainty Relation 164

5.1.2 Related Work 168

5.2 An All-But-One Entropic Uncertainty Relation 169

5.2.1 Constructing Good Families of Bases 175

5.3 A New Quantum Identification Protocol 175

5.3.1 Description of Our New Protocol 177

5.4 (Unconditional) Server Security 178

5.5 User Security in the BQSM 179

5.6 User Security in the Single-Qubit-Operations Model 182

5.6.1 The Model 182

5.6.2 No Privacy Amplification 183

5.6.3 Single-Qubit Measurements 183

5.6.4 User Security of NEWQID 187

5.6.5 Attack against NEWQID using Operations on Pairs of Qubits 192

5.7 Conclusion 194

5.1 Introduction

In this chapter, we propose a new entropic uncertainty relation. Furthermore, we present a modified version of the quantum identification protocol QID introduced in Section 2.11, that we will refer to as NEWQID. We will show how the new uncertainty relation can be used to prove NEWQID secure in the bounded-quantum-storage model (BQSM). Moreover, we will introduce another security model in this chapter, which we call the *single-qubit-operations model* (SQOM), and show that NEWQID is also secure in this model.

5.1.1 A New Uncertainty Relation

Uncertainty relations are quantitative characterizations of the uncertainty principle of quantum mechanics, which expresses that for certain pairs of measurements, there exists no state for which the measurement outcome is determined for *both* measurements: at least one of the outcomes must be somewhat uncertain. *Entropic* uncertainty relations express this uncertainty in at least one of the measurement outcomes by means of an entropy measure, usually the Shannon entropy. Our new entropic uncertainty relation distinguishes itself from previously known uncertainty relations by the following collection of features:

1. It uses the *min-entropy* as entropy measure, rather than the Shannon entropy. Such an uncertainty relation is sometimes also called a *high-order* entropic uncertainty relation.[1] Since privacy amplification needs a lower bound on the min-entropy, high-order entropic uncertainty relations are useful tools in quantum cryptography.

2. It lower bounds the uncertainty in the measurement outcome for *all but one* measurements, chosen from an *arbitrary* (and arbitrarily large) family of possible measurements. This is clearly *stronger* than typical entropic uncertainty relations that lower bound the uncertainty on *average* (over the choice of the measurement).

3. The measurements can be chosen to be qubit-wise measurements, in the computational or Hadamard basis, and thus the uncertainty relation is applicable to settings that can be implemented using current technology.

[1] This is because the min-entropy coincides with the Rényi entropy H_α of high(est) order $\alpha = \infty$ (see Section 2.3.1).

To the best of our knowledge, no previous entropic uncertainty relation satisfies (1) and (2) simultaneously, let alone in combination with (3). Indeed, as pointed out in a recent overview article by Wehner and Winter [WW10], little is known about entropic uncertainty relations for more than two measurement outcomes, and even less when additionally considering min-entropy.

Explanation by means of a Simpler Entropic Uncertainty Relation

To explain our new uncertainty relation, we find it helpful to first discuss a simpler variant, which does not satisfy (1), and which follows trivially from known results. Fix an arbitrary family $\{\mathcal{B}_1, \ldots, \mathcal{B}_m\}$ of bases for a given quantum system, and let us denote the state space of this given system by $\mathcal{H}$. The *maximum overlap* of such a family is defined as the real number

$$c := \max\{|\langle\phi|\psi\rangle| : |\phi\rangle \in \mathcal{B}_j, |\psi\rangle \in \mathcal{B}_k, 1 \leq j < k \leq m\}.$$

Let $d := -\log(c^2)$. Furthermore, let $\rho \in \mathcal{D}(\mathcal{H})$ be an arbitrary quantum state, and let X denote the measurement outcome when ρ is measured in one of the bases. We model the choice of the basis by a random variable J, so that $H(X|J=j)$ denotes the Shannon entropy of the measurement outcome when ρ is measured in basis $\mathcal{B}_j$. It follows immediately from Maassen and Uffink's uncertainty relation [MU88] that

$$H(X|J=j) + H(X|J=k) \geq -\log(c^2) = d \quad \forall j \neq k.$$

As a direct consequence, there exists a choice j' for the measurement so that $H(X|J=j) \geq \frac{d}{2}$ for all $j \in \{1, \ldots, m\}$ with $j \neq j'$. In other words, for any state ρ there exists j' so that unless the choice for the measurement coincides with j', which happens with probability at most $\max_j P_J(j)$, there is at least $d/2$ bits of entropy in the outcome X.

Our new high-order entropic uncertainty relation shows that this very statement essentially still holds when we replace Shannon by min-entropy, except that j' becomes randomized: for any ρ, there exists a *random variable* J', independent of J, such that[2]

$$H_{\min}(X|J=j, J'=j') \gtrsim \frac{d}{2} \quad \forall\, j, j' \in [m] \text{ such that } j \neq j'$$

no matter what the distribution of J is. Thus, unless the measurement J coincides with J', there is roughly $d/2$ bits of min-entropy in the outcome X. Furthermore,

[2] The rigorous version of the approximate inequality $\gtrsim$ is stated in Theorem 5.3.

since J' is *independent* of J, the probability that J coincides with J' is at most $\max_j P_J(j)$, as is the case for a fixed J'.

Note that we have no control over (the distribution of) J'. We can merely guarantee that it exists and is independent of J. It may be insightful to interpret J' as a *virtual guess* for J, guessed by the party that prepares ρ, and whose goal is to have little uncertainty in the measurement outcome X. The reader may think of the following specific way of preparing ρ: sample j' according to some arbitrary distribution J', and then prepare the state as the, say, first basis vector of $\mathcal{B}_{j'}$. If the resulting mixture ρ is then measured in some basis $\mathcal{B}_j$, sampled according to an arbitrary (independent) distribution J, then unless $j = j'$ (i.e., our guess for j was correct), there is obviously lower bounded uncertainty in the measurement outcome X (assuming a non-trivial maximum overlap). Our uncertainty relation can be understood as saying that for *any* state ρ, no matter how it is prepared, there exists such a (virtual) guess J', which exhibits this very behavior: if it differs from the actual choice for the measurement then there is lower bounded uncertainty in the measurement outcome X. As an immediate consequence, we can for instance say that X has min-entropy at least $d/2$, except with a probability that is given by the probability of guessing J, e.g., except with probability $1/m$ if the measurement is chosen uniformly at random from the family. This is clearly the best we can hope for.

We stress that because the min-entropy is more conservative than the Shannon entropy, our high-order entropic uncertainty relation does not follow from its simpler Shannon-entropy version. Neither can it be deduced in an analogue way; the main reason being that for fixed pairs $j \neq k$, there is no strong lower bound on $H_{\min}(X|J{=}j) + H_{\min}(X|J{=}k)$, in contrast to the case of Shannon entropy. More precisely and more generally, the *average* uncertainty $\frac{1}{|J|}\sum_j H_{\min}(X|J{=}j)$ does not allow a lower bound higher than $\log|J|$. To see this, consider the following example for $|J| = 2$ (the example can easily be extended to arbitrary $|J|$). Suppose that ρ is the uniform mixture of two pure states, one giving no uncertainty when measured in basis j, and the other giving no uncertainty when measured in basis k. Then, $\frac{1}{2}H_{\min}(X|J{=}j) + \frac{1}{2}H_{\min}(X|J{=}k) = 1$. Because of a similar reason, we cannot hope to get a good bound for all but a *fixed* choice of j'; the probabilistic nature of J' is necessary (in general). Hence, compared to bounding the average uncertainty, the all-but-one form of our uncertainty relation not only makes our uncertainty relation stronger in that uncertainty for all-but-one implies uncertainty on average (yet not vice versa), but it also allows for *more* uncertainty.

By using asymptotically good error-correcting codes, one can construct families

$\{\mathcal{B}_1, \ldots, \mathcal{B}_m\}$ of bases that have a large value of d, and thus for which our uncertainty relation guarantees a large amount of min-entropy (we discuss this in more detail in Section 5.2.1). These families consist of qubit-wise measurements in the computational or the Hadamard basis, and thus are implementable with current technology.

The proof of our new uncertainty relation comprises a rather involved probability reasoning to prove the existence of the random variable J' and builds on earlier work presented in [Sch07].

Quantum Identification with Hybrid Security

As an application of our entropic uncertainty relation, we propose a new *quantum identification protocol* (for an introduction into quantum identification, see Section 2.11). Our uncertainty relation gives us the right tool to prove security of the new quantum identification protocol in the BQSM. The distinguishing feature of our new protocol is that it also offers some security in case the assumption underlying the BQSM fails to hold. Indeed, we additionally prove security of our new protocol against a dishonest server that has unbounded quantum storage capabilities and can reliably store all the qubits communicated during an execution of the protocol, but is restricted to non-adaptive single-qubit operations and measurements. [3] This is in sharp contrast to protocol QID by Damgård *et al.* (Section 2.11.1), which completely breaks down against a dishonest server that can store all the communicated qubits in a quantum memory and postpone the measurements until the user announces the correct measurement bases. On the downside, our protocol only offers security in case of a perfect single-qubit (e.g., single-photon) source, because multi-qubit emissions reveal information about w. Hence, given the immature state of single-qubit-source technology (as of 2012), our protocol is currently mainly of theoretical interest.

We want to stress that proving security of our protocol in this *single-qubit-operations model* (SQOM) is non-trivial. Indeed, as we will see, standard tools like privacy amplification are not applicable. Our proof involves certain properties of random linear codes and makes use of Diaconis and Shahshahani's XOR inequality (Theorem 2.8, see also [Dia88]).

[3] Because secure identification belongs to the class of secure 2PC functionalities, it is well known that *some* restriction is necessary (for references, see Section 1.3.2).

5.1.2 Related Work

The study of *entropic* uncertainty relations, whose origin dates back to 1957 with the work of Hirschman [Hir57], has received a lot of attention over the last decade due to their various applications in quantum information processing. We refer the reader to [WW10] for a recent overview on entropic uncertainty relations. Most of the known entropic uncertainty relations are of the form

$$\frac{1}{|J|}\sum_{j} H_\alpha(X|J{=}j) \geq h\,,$$

where H_α is the Rényi entropy.[4] I.e., most uncertainty relations only give a lower bound on the entropy of the measurement outcome X *on average* over the (random) choice of the measurement. As argued in Section 5.1.1, the bound h on the *min*-entropy can be at most $\log|J|$, no matter the range of X. Furthermore, an uncertainty relation of this form only guarantees that there is uncertainty in X for *some* measurement(s), but does not specify precisely for how many, and certainly it does not guarantee uncertainty for *all but one* measurements. The same holds for the high-order entropic uncertainty relation from [DFR$^+$07], which considers an exponential number of measurement settings and guarantees that except with negligible probability over the (random) choice of the measurement, there is lower-bounded min-entropy in the outcome. On the other hand, the high-order entropic uncertainty relation from [DFSS05] only considers *two* measurement settings and guarantees lower-bounded min-entropy with probability (close to) $\frac{1}{2}$.

The uncertainty relation we know of that comes closest to ours is Lemma 2.13 in [FHS11]. Using our notation, it shows that X is ϵ-close to having roughly $d/2$ bits of min-entropy (i.e., the same bound we get), but only for all but an ϵ-fraction of all the m possible choices for the measurement j, where ϵ is about $\sqrt{2/m}$.

With respect to our application, backing up the security of the identification protocol by Damgård *et al.* [DFSS07] against an adversary that can overcome the quantum-memory bound assumed by the BQSM was also the goal of [DFL$^+$09]. However, the solution proposed there relies on an unproven computational hardness assumption, and as such, strictly speaking, can be broken by an adversary in the SQOM, i.e., by storing qubits and measuring them later qubit-wise and performing (possibly infeasible) classical computations. On the other hand, by *assuming* a lower bound on the hardness of the underlying computational problem against quantum machines, the security of the protocol in [DFL$^+$09] holds against an adversary with much

[4] See Section 2.3.1 for the definition of H_α. Nevertheless, for most known uncertainty relations $\alpha = 1$, i.e., the Shannon entropy.

more quantum computing power than our protocol in the SQOM, which restricts the adversary to single-qubit operations.

We hope that with future research on this topic, new quantum identification (or other cryptographic) protocols will be developed with security in the same spirit as our protocol, but with a more relaxed restriction on the adversary's quantum computation capabilities, for instance that he can only perform a limited number of quantum computation steps, and in every step he can only act on a limited number of qubits coherently.

5.2 An All-But-One Entropic Uncertainty Relation

Throughout this section, $\{\mathcal{B}_1, \ldots, \mathcal{B}_m\}$ is an arbitrary but fixed family of bases for the state space $\mathcal{H}$ of a quantum system. For simplicity, we restrict our attention to an n-qubit system, such that $\mathcal{H} = (\mathbb{C}^2)^{\otimes n}$ for $n \in \mathbb{N}$, but our results immediately generalize to arbitrary quantum systems. We write the 2^n basis vectors of the j-th basis $\mathcal{B}_j$ as $\mathcal{B}_j = \{|x\rangle_j : x \in \{0,1\}^n\}$. Let c be the maximum overlap of $\{\mathcal{B}_1, \ldots, \mathcal{B}_m\}$ as defined in Section 5.1.1.

In order to obtain our entropic uncertainty relation that lower bounds the min-entropy of the measurement outcome for all but one measurement, we first state an uncertainty relation that expresses uncertainty by means of the probability measure of given sets.

Theorem 5.1 (cf. Thm. 4.18 in [Sch07]) *Let ρ be an arbitrary state of n qubits. For $j \in [m]$, let $Q^j(\cdot)$ be the distribution of the outcome when ρ is measured in the $\mathcal{B}_j$-basis, i.e., $Q^j(x) := \langle x|_j \, \rho \, |x\rangle_j$ for any $x \in \{0,1\}^n$. And for all subsets $\mathcal{X} \subset \{0,1\}^n$, let $Q^j(\mathcal{X}) := \sum_{x\in\mathcal{X}} Q^j(x)$. Then, for any family $\{\mathcal{L}^j\}_{j\in[m]}$ of subsets $\mathcal{L}^j \subset \{0,1\}^n$, it holds that*

$$\sum_{j\in[m]} Q^j(\mathcal{L}^j) \le 1 + c\,(m-1) \cdot \max_{\substack{j,k\in[m]\\ j\neq k}} \sqrt{|\mathcal{L}^j||\mathcal{L}^k|}.$$

A special case of Theorem 5.1, obtained by restricting the family of bases to the specific choice $\{\mathcal{B}_+, \mathcal{B}_\times\}$ with $\mathcal{B}_+ = \{|x\rangle : x \in \{0,1\}^n\}$ and $\mathcal{B}_\times = \{H^{\otimes n}|x\rangle : x \in \{0,1\}^n\}$ (i.e., either the computational or Hadamard basis for all qubits), is an uncertainty relation that was proven and used in the original paper about the BQSM [DFSS05]. The proof of Theorem 5.1 goes along similar lines as the proof in the journal version of [DFSS05] for the special case outlined above. The proof of Theorem 5.1 can be found in [Sch07], as well as in [BFGS12].

In the same spirit as Corollary 4.17 in [Sch07] (see also the full version of [DFSS05]), we reformulate above uncertainty relation in terms of a "good event" $\mathcal{E}$, which occurs with reasonable probability, and if it occurs, then the measurement outcomes have high min-entropy. The statement is obtained by choosing the sets $\mathcal{L}^j$ in Theorem 5.1 appropriately.

Because we now switch to entropy notation, it will be convenient to work with a measure of overlap between bases that is logarithmic in nature and *relative* to the number n of qubits. Hence, we define

$$\delta := -\frac{1}{n}\log c^2 .$$

We will later see that for "good" choices of bases, δ stays constant for growing n.

Corollary 5.2 *Let ρ be an arbitrary n-qubit state, let J be a random variable over $[m]$ (with arbitrary distribution P_J), and let X be the outcome when measuring ρ in basis $\mathcal{B}_J$.*[5] *Then, for any $\epsilon \in \mathbb{R}$ such that $0 < \epsilon < \delta/4$, there exists an event $\mathcal{E}$ such that*

$$\sum_{j\in[m]} \Pr[\mathcal{E}|J{=}j] \geq (m-1) - (2m-1)\cdot 2^{-\epsilon n}$$

and

$$H_{\min}(X|J{=}j,\mathcal{E}) \geq \Big(\frac{\delta}{2} - 2\epsilon\Big)n$$

for $j \in [m]$ with $P_{J|\mathcal{E}}(j) > 0$.

Proof. For $j \in [m]$ define

$$\mathcal{S}^j := \big\{x \in \{0,1\}^n : Q^j(x) \leq 2^{-(\delta/2-\epsilon)n}\big\}$$

to be the sets of strings with small probabilities and denote by $\mathcal{L}^j := \overline{\mathcal{S}}^j$ their complements.[6] Note that for all $x \in \mathcal{L}^j$, we have that $Q^j(x) > 2^{-(\delta/2-\epsilon)n}$ and therefore $|\mathcal{L}^j| < 2^{(\delta/2-\epsilon)n}$. It follows from Theorem 5.1 that

$$\begin{aligned}\sum_{j\in[m]} Q^j(\mathcal{S}^j) = \sum_{j\in[m]} (1-Q^j(\mathcal{L}^j)) &\geq m - (1+(m-1)\cdot 2^{-\epsilon n})\\ &= (m-1)-(m-1)2^{-\epsilon n}.\end{aligned}$$

We define $\mathcal{E} := \{X \in \mathcal{S}^J \,\wedge\, Q^J(\mathcal{S}^J) \geq 2^{-\epsilon n}\}$ to be the event that $X \in \mathcal{S}^J$ and at the same time the probability that this happens is not too small. Then $\Pr[\mathcal{E}|J{=}$

[5] I.e., $P_{X|J}(x|j) = Q^j(x)$, using the notation from Theorem 5.1.

[6] Here's the mnemonic: $\mathcal{S}$ for the strings with *s*mall probabilities, $\mathcal{L}$ for *l*arge.

$j] = \Pr[X \in \mathcal{S}^j \wedge Q^j(\mathcal{S}^j) \geq 2^{-\epsilon n} | J = j]$ either vanishes (if $Q^j(\mathcal{S}^j) < 2^{-\epsilon n}$) or else equals $Q^j(\mathcal{S}^j)$. In either case, $\Pr[\mathcal{E}|J = j] \geq Q^j(\mathcal{S}^j) - 2^{-\epsilon n}$ holds and thus the first claim follows by summing over $j \in [m]$ and using the derivation above. Furthermore, let $p = \max_j P_J(j)$, then $\Pr[\bar{\mathcal{E}}] = \sum_{j\in[m]} P_J(j) \Pr[\bar{\mathcal{E}}|J = j] \leq p \sum_{j\in[m]} \Pr[\bar{\mathcal{E}}|J = j] \leq p(m - (\sum_{j\in[m]} Q^j(\mathcal{S}^j) - 2^{-\epsilon n})) \leq p(1 + (2m-1) \cdot 2^{-\epsilon n})$, and $\Pr[\mathcal{E}] \geq (1-p) - p(2m-1) \cdot 2^{-\epsilon n}$

Regarding the second claim, in case $J = j$, we have

$$\begin{aligned} H_{\min}(X|J = j, \mathcal{E}) &= -\log\left(\max_{x \in \mathcal{S}^j} \frac{Q^j(x)}{Q^j(\mathcal{S}^j)}\right) \geq -\log\left(\frac{2^{-(\delta/2-\epsilon)n}}{Q^j(\mathcal{S}^j)}\right) \\ &= (\delta/2 - \epsilon)n + \log(Q^j(\mathcal{S}^j)). \end{aligned}$$

As $Q^j(\mathcal{S}^j) \geq 2^{-\epsilon n}$ by definition of $\mathcal{E}$, we have $H_{\min}(X|J = j, \mathcal{E}) \geq (\delta/2 - 2\epsilon)n$. □

We are now ready to state and prove our new all-but-one entropic uncertainty relation.

Theorem 5.3 *Let ρ be an arbitrary n-qubit state, let J be a random variable over $[m]$ (with arbitrary distribution P_J), and let X be the outcome when measuring ρ in basis $\mathcal{B}_J$. Then, for any $\epsilon \in \mathbb{R}$ such that $0 < \epsilon < \delta/4$, there exists a random variable J' with joint distribution $P_{JJ'X}$ such that (1) J and J' are independent and (2) there exists an event Ω with $\Pr[\Omega] \geq 1 - 2 \cdot 2^{-\epsilon n}$ such that*[7]

$$H_{\min}(X|J = j, J' = j', \Omega) \geq \Big(\frac{\delta}{2} - 2\epsilon\Big)n - 1$$

for all $j, j' \in [m]$ with $j \neq j'$ and $P_{JJ'|\Omega}(j, j') > 0$.

Note that, as phrased, Theorem 5.3 requires that J is fixed and known, and only then the existence of J' can be guaranteed. This is actually not necessary. By looking at the proof, we see that J' can be defined simultaneously in all m probability spaces $P_{X|J=j}$ with $j \in [m]$, without having assigned a probability distribution to J yet, so that the resulting random variable J' we obtain by assigning an *arbitrary* probability distribution P_J to J, satisfies the claimed properties. This in particular implies that the (marginal) distribution of J' is fully determined by ρ.

The idea of the proof of Theorem 5.3 is to (try to) define the random variable J' in such a way that the event $J \neq J'$ coincides with the "good event" $\mathcal{E}$ from

[7] Instead of introducing such an event Ω, we could also express the min-entropy bound by means of the *smooth* min-entropy of X given $J = j$ and $J' = j'$.

Corollary 5.2. It then follows immediately from Corollary 5.2 that $H_{\min}(X|J = j, J' \neq J) \geq (\delta/2 - 2\epsilon)n$, which is already close to the actual min-entropy bound we need to prove. This approach dictates that if the event $\mathcal{E}$ does not occur, then J' needs to *coincide* with J. Vice versa, if $\mathcal{E}$ does occur, then J' needs to be *different* to J. However, it is a priori unclear *how* to choose J' different from J in case $\mathcal{E}$ occurs. There is only one way to set J' to be equal to J, but there are many ways to set J' to be different from J (unless $m = 2$). It needs to be done in such a way that without conditioning on $\mathcal{E}$ or its complement, J and J' are independent.

Somewhat surprisingly, it turns out that the following does the job. To simplify this informal discussion, we assume that the sum of the m probabilities $\Pr[\mathcal{E}|J\!=\!j]$ from Corollary 5.2 equals $m - 1$ exactly. It then follows that the corresponding complementary probabilities, $\Pr[\bar{\mathcal{E}}|J\!=\!j]$ for the m different choices of $j \in [m]$, add up to 1 and thus form a probability distribution. J' is now chosen, in the above spirit depending on the event $\mathcal{E}$, so that its marginal distribution $P_{J'}$ coincides with this probability distribution: $P_{J'}(j') = \Pr[\bar{\mathcal{E}}|J\!=\!j']$ for all $j' \in [m]$. Thus, in case the event $\mathcal{E}$ occurs, J' is chosen according to this distribution but conditioned on being different from the value j, taken on by J. The technical details, and how to massage the argument in case the sum of the $\Pr[\mathcal{E}|J\!=\!j]$'s is not exactly $m - 1$, are worked out in the proof below.

Proof of Theorem 5.3. From Corollary 5.2 we know that for any $0 < \epsilon < \delta/4$, there exists an event $\mathcal{E}$ such that $\sum_{j\in[m]} \Pr[\mathcal{E}|J = j] = m - 1 - \alpha$, and thus $\sum_{j\in[m]} \Pr[\bar{\mathcal{E}}|J = j] = 1 + \alpha$, for $\alpha \in \mathbb{R}$ such that $-1 \leq \alpha \leq (2m - 1)2^{-\epsilon n}$. We make the case distinction between $\alpha = 0$, $\alpha > 0$ and $\alpha < 0$. We start with case $\alpha = 0$, we subsequently prove the other two cases by reducing them to the case $\alpha = 0$ by "inflating" and "deflating" the event $\mathcal{E}$ appropriately. The approach for the case $\alpha = 0$ is to define J' in such way that $\mathcal{E} \iff J \neq J'$, i.e., the event $J \neq J'$ coincides with the event $\mathcal{E}$. The min-entropy bound from Corollary 5.2 then immediately translates to $H_{\min}(X|J = j, J' \neq J) \geq (\delta/2 - 2\epsilon)n$, and to $H_{\min}(X|J = j, J' = j') \geq (\delta/2 - 2\epsilon)n$ for $j' \neq j$ with $P_{JJ'}(j, j') > 0$, as we will show. What is not obvious about the approach is how to define J' when it is supposed to be different from J, i.e., when the event $\mathcal{E}$ occurs, so that in the end J and J' are independent.

Formally, we define J' by means of the following conditional probability distributions:

$$P_{J'|JX\bar{\mathcal{E}}}(j'|j, x) := \begin{cases} 1 & \text{if } j = j' \\ 0 & \text{if } j \neq j' \end{cases}$$

and

$$P_{J'|JX\mathcal{E}}(j'|j,x) := \begin{cases} 0 & \text{if } j = j' \\ \dfrac{\Pr[\bar{\mathcal{E}}|J=j']}{\Pr[\mathcal{E}|J=j]} & \text{if } j \neq j' \end{cases}$$

We assume for the moment that the denominator in the latter expression does not vanish for any j; we take care of the case where it does later. Trivially, $P_{J'|JX\bar{\mathcal{E}}}$ is a proper distribution, with non-negative probabilities that add up to 1, and the same holds for $P_{J'|JX\mathcal{E}}$:

$$\sum_{j'\in[m]} P_{J'|JX\bar{\mathcal{E}}} = \sum_{j'\in[m]\setminus\{j\}} P_{J'|JX\bar{\mathcal{E}}} = \sum_{j'\in[m]\setminus\{j\}} \frac{\Pr[\bar{\mathcal{E}}|J=j']}{\Pr[\mathcal{E}|J=j]} = 1$$

where we used that $\sum_{j\in[m]} \Pr[\bar{\mathcal{E}}|J=j] = 1$ (because $\alpha = 0$) in the last equality. Furthermore, it follows immediately from the definition of J' that $\bar{\mathcal{E}} \implies J = J'$ and $\mathcal{E} \implies J \neq J'$. Hence, $\mathcal{E} \iff J \neq J'$, and thus the bound from Corollary 5.2 translates to $H_{\min}(X|J=j, J'\neq J) \geq (\delta/2 - 2\epsilon)n$. It remains to argue that J' is independent of J, and that the bound also holds for $H_{\min}(X|J = j, J' = j')$ whenever $j \neq j'$.

The latter follows immediately from the fact that conditioned on $J \neq J'$ (which is equivalent to $\mathcal{E}$), X, J and J' form a Markov chain $X \leftrightarrow J \leftrightarrow J'$, and thus, given $J = j$, additionally conditioning on $J' = j'$ does not change the distribution of X. For the independence of J and J', consider the joint probability distribution of J and J', given by

$$\begin{aligned} P_{JJ'}(j,j') &= P_{J'J\mathcal{E}}(j',j) + P_{J'J\bar{\mathcal{E}}}(j',j) \\ &= P_J(j)\Pr[\mathcal{E}|J=j]P_{J'|J\mathcal{E}}(j'|j) + P_J(j)\Pr[\bar{\mathcal{E}}|J=j]P_{J'|J\bar{\mathcal{E}}}(j'|j) \\ &= P_J(j)\Pr[\bar{\mathcal{E}}|J=j'], \end{aligned}$$

where the last equality follows by separately analyzing the cases $j = j'$ and $j \neq j'$. It follows immediately that the marginal distribution of J' is

$$P_{J'}(j') = \sum_j P_{JJ'}(j,j') = \Pr[\bar{\mathcal{E}}|J=j'],$$

and thus $P_{JJ'} = P_J \cdot P_{J'}$.

What is left to do for the case $\alpha = 0$ is to deal with the case where there exists j^* with $\Pr[\mathcal{E}|J=j^*] = 0$. Since $\sum_{j\in[m]} \Pr[\bar{\mathcal{E}}|J=j] = 1$, it holds that $\Pr[\bar{\mathcal{E}}|J=j] = 0$ for $j \neq j^*$. This motivates to define J' as $J' := j^*$ with probability 1. Note that this definition directly implies that J' is independent from J. Furthermore, by the above observations: $\mathcal{E} \iff J \neq J'$. This concludes the case $\alpha = 0$.

Next, we consider the case $\alpha > 0$. The idea is to "inflate" the event $\mathcal{E}$ so that α becomes 0, i.e., to define an event $\mathcal{E}'$ that contains $\mathcal{E}$ (meaning that $\mathcal{E} \implies \mathcal{E}'$) so that $\sum_{j\in[m]} \Pr[\mathcal{E}'|J = j] = m - 1$, and to define J' as in the case $\alpha = 0$ (but now using $\mathcal{E}'$). Formally, we define $\mathcal{E}'$ as the disjoint union $\mathcal{E}' = \mathcal{E} \vee \mathcal{E}_\circ$ of $\mathcal{E}$ and an event $\mathcal{E}_\circ$. The event $\mathcal{E}_\circ$ is defined by means of $\Pr[\mathcal{E}_\circ|\mathcal{E}, J = j, X = x] = 0$, so that $\mathcal{E}$ and $\mathcal{E}_\circ$ are indeed disjoint, and $\Pr[\mathcal{E}_\circ|J = j, X = x] = \alpha/m$, so that indeed

$$\begin{aligned}\sum_{j\in[m]} \Pr[\mathcal{E}'|J = j] &= \sum_{j\in[m]} (\Pr[\mathcal{E}|J = j] + \Pr[\mathcal{E}_\circ|J = j]) \\ &= (m - 1 - \alpha) + \alpha = m - 1\,.\end{aligned}$$

We can now apply the analysis of the case $\alpha = 0$ to conclude the existence of J', independent of J, such that $J \neq J' \iff \mathcal{E}'$ and thus $(J \neq J') \wedge \bar{\mathcal{E}}_\circ \iff \mathcal{E}' \wedge \bar{\mathcal{E}}_\circ \iff \mathcal{E}$. Setting $\Omega := \bar{\mathcal{E}}_\circ$, it follows that

$$H_{\min}(X|J = j, J \neq J', \Omega) = H_{\min}(X|J = j, \mathcal{E}) \geq (\delta/2 - 2\epsilon)n\,,$$

where $\Pr[\Omega] = 1 - \Pr[\mathcal{E}_\circ] = 1 - \alpha/m \geq 1 - (2m - 1)2^{-\epsilon n}/m \geq 1 - 2 \cdot 2^{-\epsilon n}$. Finally, using similar reasoning as in the case $\alpha = 0$, it follows that the same bound holds for $H_{\min}(X|J = j, J' = j', \Omega)$ whenever $j \neq j'$. This concludes the case $\alpha > 0$.

Finally, we consider the case $\alpha < 0$. The approach is the same as above, but now $\mathcal{E}'$ is obtained by "deflating" $\mathcal{E}$. Specifically, we define $\mathcal{E}'$ by means of $\Pr[\mathcal{E}'|\bar{\mathcal{E}}, J = j, X = x] = \Pr[\mathcal{E}'|\bar{\mathcal{E}}] = 0$, so that $\mathcal{E}'$ is contained in $\mathcal{E}$, and $\Pr[\mathcal{E}'|\mathcal{E}, J = j, X = x] = \Pr[\mathcal{E}'|\mathcal{E}] = \frac{m-1}{m-1-\alpha}$, so that

$$\sum_{j\in[m]} \Pr[\mathcal{E}'|J = j] = \sum_{j\in[m]} \Pr[\mathcal{E}'|\mathcal{E}] \cdot \Pr[\mathcal{E}|J = j] = m - 1\,.$$

Again, from the $\alpha = 0$ case we obtain J', independent of J, such that the event $J \neq J'$ is equivalent to the event $\mathcal{E}'$.

It follows that

$$\begin{aligned}H_{\min}(X|J = j, J \neq J') &= H_{\min}(X|J = j, \mathcal{E}') = H_{\min}(X|J = j, \mathcal{E}', \mathcal{E}) \\ &\geq H_{\min}(X|J = j, \mathcal{E}) - \log(P[\mathcal{E}'|\mathcal{E}, J = j]) \geq (\delta/2 - 2\epsilon)n - 1\,,\end{aligned}$$

where the second equality holds because $\mathcal{E}' \implies \mathcal{E}$, the first inequality holds because additionally conditioning on $\mathcal{E}'$ increases the probabilities of X conditioned on $J = j$ and $\mathcal{E}$ by at most a factor $1/P[\mathcal{E}'|\mathcal{E}, J = j]$), and the last inequality holds by Corollary 5.2) and because $P[\mathcal{E}'|\mathcal{E}, J = j]) = \frac{m-1}{m-1-\alpha} \geq \frac{1}{2}$, where the latter holds since $\alpha \geq -1$. Finally, using similar reasoning as in the previous cases, it follows that the same bound holds for $H_{\min}(X|J = j, J' = j')$ whenever $j \neq j'$. This concludes the proof. □

5.2.1 Constructing Good Families of Bases

Here, we discuss some interesting choices for the family $\{\mathcal{B}_1, \ldots, \mathcal{B}_m\}$ of bases. We say that such a family is "good" if $\delta = -\frac{1}{n}\log(c^2)$ converges to a strictly positive constant as n tends to infinity. There are various ways to construct such families. For example, a family obtained through sampling according to the Haar measure will be good with overwhelming probability (a precise statement, in which "good" means $\delta = 0.9$, can be found at the very end of the proof of Theorem 2.5 of [FHS11]). The best possible constant $\delta = 1$ is achieved for a family of *mutually unbiased bases*. However, for arbitrary quantum systems (i.e., not necessarily multi-qubit systems) it is not well understood how large such a family may be, beyond that its size cannot exceed the dimension plus 1.

In the upcoming section, we will use the following simple and well-known construction. For an arbitrary binary code $\mathcal{C} \subset \mathbb{F}_2^n$ of size m, minimum distance d and encoding function $\mathfrak{c} : [m] \to \mathcal{C}$, we can construct a family $\{\mathcal{B}_1, \ldots, \mathcal{B}_m\}$ of bases as follows. We identify the jth codeword, i.e., $\mathfrak{c}(j) = (c_1, \ldots, c_n)$ for $j \in [m]$, with the basis $\mathcal{B}_j = \{H^{\mathfrak{c}(j)}|x\rangle : x \in \mathbb{F}_2^n\} = \{(H^{c_1} \otimes \cdots \otimes H^{c_n})|x\rangle : x \in \mathbb{F}_2^n\}$. In other words, $\mathcal{B}_j$ measures qubit-wise in the computational or the Hadamard basis, depending on the corresponding coordinate of $\mathfrak{c}(j)$. It is easy to see that the maximum overlap c of the family obtained this way is directly related to the minimum distance of $\mathcal{C}$, namely $\delta = -\frac{1}{n}\log(c^2)$ coincides with the relative minimal distance d/n of $\mathcal{C}$. Hence, choosing an asymptotically good code immediately yields a good family of bases.

5.3 A New Quantum Identification Protocol

Our main application of the new uncertainty relation is in proving security of a new password-based identification protocol in the quantum setting. Recall that in password-based identification, a user U wants to convince a server S that he (U) knows a password w, in such a way that only a negligible amount of information is leaked about w in case U is interacting with a dishonest server S*. *Vice versa*, a dishonest user U^* (who does not know w) should not be able to gain information about w by interacting with S.

It is known that *without* any restriction on (one of) the dishonest participants, secure identification is impossible (even in the quantum setting). Indeed, if a quantum protocol is unconditionally secure against a dishonest user, then unavoidably it can be broken by a dishonest server with unbounded quantum storage and unbounded quantum computing power; this follows essentially from [Lo97] (see also [DFSS07]). Thus, the best one can hope for (for a protocol that is unconditionally secure against a dishonest user) is that in order to break it, unbounded quantum storage *and* unbounded quantum computing power are *necessary* for the dishonest server. Note that this is not the case for the existing quantum identification protocol QID, which we reviewed in Section 2.11.1: a dishonest server who can postpone the measurements of (most of) the qubits until the user announces the bases—by temporarily storing the qubits in a quantum memory—completely breaks the protocol. Thus, no quantum computing power at all is necessary to break QID, only sufficient quantum storage.

In this section, we propose a new identification protocol, NEWQID, which can be regarded as a first step towards closing the above gap. Like QID, our new protocol is secure against an unbounded dishonest user and against a dishonest server with limited quantum storage capabilities. Furthermore, and in contrast to QID, a minimal amount of quantum computation power is *necessary* to break the protocol, beyond sufficient quantum storage. Indeed, in addition to the security against a dishonest server with bounded quantum storage, we also prove security against a dishonest server that can store all the communicated qubits, but is restricted to measure them qubit-wise (in arbitrary qubit bases) at the end of the protocol execution. Thus, beyond sufficient quantum storage, quantum computation that involves *pairs* of qubits is necessary (and in fact sufficient) to break the new protocol.

Restricting the dishonest server to qubit-wise measurements may look restrictive; however, we stress that in order to break the protocol, the dishonest server needs to store many qubits *and* perform quantum operations on them that go beyond single-qubit operations; this may indeed be considerably more challenging than storing many qubits and measuring them qubit-wise. Furthermore, it turns out that proving security against such a dishonest server that is restricted to qubit-wise measurements is already challenging; indeed, standard techniques (e.g., privacy amplification) do not seem applicable here. Therefore, handling a dishonest server that can, say, act on *blocks* of qubits, must be left to future research.

The security properties that we want to achieve are given in Section 2.11. Similar to

QID, the new protocol will be shown to be unconditionally secure against dishonest users. The new uncertainty relation is the main ingredient for proving security against a dishonest server with bounded quantum storage. Our security proof against a dishonest server (having unbounded quantum storage) that is restricted to non-adaptive qubit-wise measurements uses very different techniques.

5.3.1 Description of Our New Protocol

Let $\mathcal{C} \subset \mathbb{F}_2^n$ be a binary code with minimum distance d, and let $\mathfrak{c} : \mathcal{W} \to \mathcal{C}$ be its encoding function. Let $m := |\mathcal{W}|$, and typically, $m < 2^n$. Let $\mathcal{F}$ be the class of all linear functions from $\{0,1\}^n$ to $\mathbb{F}_2^\ell$, where $\ell < n$, represented as $\ell \times n$ matrices over $\mathbb{F}_2$. Note that $\mathcal{F}$ is two-universal and coincides with the family $\mathcal{G}_1$ defined—and proved to be two-universal—in Section 2.4.1. Furthermore, let $\mathcal{G}$ be a strongly two-universal class of hash functions from $\mathcal{W}$ to $\mathbb{F}_2^\ell$. Protocol NEWQID is shown below.

1. U picks $x \overset{\mathtt{r}}{\leftarrow} \{0,1\}^n$ and sends $H^{\mathfrak{c}(w)}|x\rangle$ to S.
2. S measures in basis $\mathfrak{c}(w)$. Let x' be the outcome.
3. U picks $f \overset{\mathtt{r}}{\leftarrow} \mathcal{F}$ and sends it to S
4. S picks $g \overset{\mathtt{r}}{\leftarrow} \mathcal{G}$ and sends it to U
5. U computes and sends $z := f(x) \oplus g(w)$ to S
6. S accepts if and only if $z = z'$ where $z' := f(x') \oplus g(w)$

Protocol 5.1: Our new quantum password-based-identification protocol NEWQID. The difference between this new protocol and the existing protocol QID by Damgård *et al.* (see Protocol 2.1) is the way how the user prepares the state in step (1): in the new protocol the basis is chosen as a function of the password w, whereas in QID it is chosen at random and communicated in a later step in the protocol.

Note that our protocol is quite similar to QID (Section 2.11.1). The difference is that in our protocol, *both* parties, i.e., U and S, use $\mathfrak{c}(w)$ as basis for preparing/measuring the qubits in step (1) and (2), whereas in QID only S uses $\mathfrak{c}(w)$ and U uses a *random* basis $\theta \in \{0,1\}^n$ instead, and then U communicates θ to S and all the positions where θ and $\mathfrak{c}(w)$ differ are dismissed. Thus, in some sense, our new protocol is more natural since why should U use a random basis when he knows the right basis (i.e., the one that S uses)? In [DFSS07], using a random basis (for U) was crucial for their proof technique, which is based on an entropic uncertainty relation of a certain form, which asks for a random basis. However, using a random basis, which then needs to be announced, renders the protocol insecure against a dishonest server S^* that is capable of storing all the communicated qubits and then measure them in the right basis once it has been announced. Our new uncertainty relation applies to the case where an n-qubit state is measured in a basis that is sampled from a code $\mathcal{C}$, and thus is applicable to the new protocol where U uses basis $\mathfrak{c}(w) \in \mathcal{C}$. Since this basis is common knowledge (to the honest participants), it does not have to be communicated, and as such a straightforward store-and-then-measure attack as above does not apply.

A downside of our protocol is that security only holds in case of a perfect quantum source, which emits exactly one qubit when triggered. Indeed, a multi-photon emission enables a dishonest server S^* to learn information on the basis used, and thus gives away information on the password w in our protocol. As such, our protocol is currently mainly of theoretical interest.

It is straightforward to verify that (in the ideal setting with perfect sources, no noise, etc.) `NEWQID` satisfies the correctness property (Definition 2.66) perfectly. In the upcoming sections, we give proofs for server and user security.

5.4 (Unconditional) Server Security

First, we argue security of `NEWQID` against an arbitrary dishonest user U^* (that is merely restricted by the laws of quantum mechanics).

Theorem 5.4 *`NEWQID` is ε-secure for the server with $\varepsilon = \binom{m}{2}2^{-\ell}$.*

Proof. Clearly, from the steps (1) to (5) in the protocol `NEWQID`, U^* learns no information on W at all. The only information he may learn is by observing whether S accepts or not in step (6). Therefore, in order to prove server security, it suffices to show the existence of a random variable W', independent of W, with the property that S rejects whenever $W' \neq W$ (except with probability $\frac{1}{2}m(m-1)2^{-\ell}$) and that S accepts whenever $W' = W$.

We may assume that $\mathcal{W} = [m]$. Let $\rho_{WX'FGZE}$ be the state describing the password W, the variables X', F, G and Z occurring in the protocol from the server's point of view, and U^*'s quantum state E *before* observing S's decision to accept or reject. For any $w \in \mathcal{W}$, consider the state $\rho^w_{X'FGZE} := \rho_{X'FGZE|W=w}$. Note that the reduced state ρ^w_{FGZE} is the same for any $w \in \mathcal{W}$; this follows from the assumption that U^*'s initial state is independent of W and because F, G and Z are produced independently of W. We may thus write $\rho^w_{X'FGZE}$ as $\rho_{X'_wFGZE}$, and we can "glue together" the states $\rho_{X'_wFGZE}$ for all choices of w. This means, there exists a state

$\rho_{X'_1\cdots X'_m FGZE_1\cdots E_m}$ that correctly reduces to $\rho_{X'_w FGZE_w} = \rho_{X'_w FGZE}$ for any $w \in \mathcal{W}$, and conditioned on FGZ, we have that $X'_i E_i$ is independent of $X'_j E_j$ for any $i \neq j \in \mathcal{W}$. It is easy to see that for any $i \neq j \in \mathcal{W}$, G is independent of X'_i, X'_j and F. Therefore, by the strong two-universality of G, for any $i \neq j$ it holds that $Z'_i \neq Z'_j$ except with probability $2^{-\ell}$, where $Z'_w = F(X'_w) + G(w)$ for any w. Therefore, by the union bound, $Z'_1, \ldots, Z'_m$ are pairwise distinct and thus Z can coincide with at most one of the Z'_w's, except with probability $\varepsilon = \frac{1}{2}m(m-1)2^{-\ell}$. Let W' be defined such that $Z = Z'_{W'}$; if there is no such Z'_w then we let $W' = \perp$, and if there are more than one then we let it be the first. Recall, the latter can happen with probability at most ε. We now extend the state $\rho_{X'_1\cdots X'_m FGZW'E_1\cdots E_m}$ by W, chosen independently according to P_W. Clearly W' is independent of W. Furthermore, except with probability at most ε, if $W \neq W'$ then $Z \neq Z'_W$. Also note that $\rho_{X'_W FGZW'WE_W}$ is such that

$$\begin{aligned}\rho_{X'_W FGZWE_W} &= \sum_w P_W(w)\rho_{X'_w FGZE_w} \otimes |w\rangle\langle w| \\ &= \sum_w P_W(w)\rho^w_{X'FGZE} \otimes |w\rangle\langle w| = \rho_{X'FGZWE}.\end{aligned}$$

Thus, also with respect to the state $\rho_{X'FGZWE}$ there exist W', independent of W, such that if $W' \neq W$ then $Z \neq Z'$ except with probability at most ε. Finally, whenever $W = W'$ it follows by construction that $Z = Z'$ and S will always accept in this case. This was to be shown. □

5.5 User Security in the BQSM

Next, we consider a dishonest server S*, and first prove security of NEWQID in the *bounded-quantum-storage model*. In this model, as introduced in [DFSS05], it is assumed that the adversary (here S*) cannot store more than a fixed number of qubits, say q. The security proof of NEWQID in the bounded quantum storage model is very similar to the corresponding proof in [DFSS07] for their protocol, except that we use the new uncertainty relation from Section 5.2. Furthermore, since our uncertainty relation (Theorem 5.3) already guarantees the existence of the random variable W' as required by the security property, no *entropy-splitting* as in [DFSS07] is needed.

In the following, let $\delta := d/n$, i.e., the relative minimum distance of $\mathcal{C}$.

Theorem 5.5 *Let* S* *be a dishonest server whose quantum memory is at most* q *qubits at step 3 of NEWQID. Then, for any* $0 < \kappa < \delta/4$, *NEWQID is* ε*-secure for the* *user with*

$$\varepsilon = 2^{-\frac{1}{2}((\delta/2-2\kappa)n-1-q-\ell)} + 4 \cdot 2^{-\kappa n}.$$

Proof. We consider and analyze a purified version of NEWQID where in step (1) instead of sending $H^{\mathfrak{c}(W)}|X\rangle$ to S* for a uniformly distributed X, U prepares a fully entangled state $2^{-n/2}\sum_x |x\rangle|x\rangle$ and sends the second register to S* while keeping the first. Then, in step (3) when the memory bound has applied, U measures his register in the basis $\mathfrak{c}(W)$ in order to obtain X. Note that this procedure produces exactly the same common state as in the original (non-purified) version of NEWQID. Thus, we may just as well analyze this purified version.

The state of S* consists of his initial state and his part of the EPR pairs, and may include an additional ancilla register. Before the memory bound applies, S* may perform any unitary transformation on his composite system. When the memory bound is applied (just before step (3) is executed in NEWQID), S* has to measure all but q qubits of his system. Let the classical outcome of this measurement be denoted by y, and let E' be the remaining quantum state of at most q qubits. The common state has collapsed to a $(n+q)$-qubit state and depends on y; the analysis below holds for any y. Next, U measures his n-qubit part of the common state in basis $\mathfrak{c}(W)$; let X denote the classical outcome of this measurement. By our new uncertainty relation (Theorem 5.3) and subsequently applying the min-entropy chain rule that is given in Proposition 2.62 (to take the q stored qubits into account) it follows that there exists W', independent of W, and an event Ω that occurs at least with probability $1 - 2\cdot 2^{-\kappa n}$, such that

$$H_{\min}(X|E', W = w, W' = w', \Omega) \geq (\delta/2 - 2\kappa)n - 1 - q.$$

for any w, w' such that $w \neq w'$. Because U chooses F independently at random from a 2-universal family, privacy amplification guarantees that

$$d_{\mathsf{unif}}(F(X)|E'F, W = w, W' = w') \leq \varepsilon' := \frac{1}{2}\cdot 2^{-\frac{1}{2}((\delta/2-2\kappa)n-1-q-\ell)} + 2\cdot 2^{-\kappa n},$$

for any w, w' such that $w \neq w'$. Recall that $Z = F(X) \oplus G(W)$. By security of the one-time pad it follows that

$$d_{\mathsf{unif}}(Z|E'FG, W = w, W' = w') \leq \varepsilon', \tag{5.1}$$

for any w, w' such that $w \neq w'$. To prove the claim, we need to bound,

$$
\begin{aligned}
&\delta(\rho_{WW'E|W\neq W'}, \rho_{W\leftrightarrow W'\leftrightarrow E|W\neq W'}) \\
&\quad= \tfrac{1}{2}\|\rho_{WW'E'FGZ|W\neq W'} - \rho_{W\leftrightarrow W'\leftrightarrow E'FGZ|W\neq W'}\|_1 \\
&\quad\leq \tfrac{1}{2}\|\rho_{WW'E'FGZ|W\neq W'} - \rho_{WW'E'FG|W\neq W'} \otimes 2^{-\ell}\mathbb{I}\|_1 \\
&\qquad\quad + \tfrac{1}{2}\|\rho_{WW'E'FG|W\neq W'} \otimes 2^{-\ell}\mathbb{I} - \rho_{W\leftrightarrow W'\leftrightarrow E'FGZ|W\neq W'}\|_1
\end{aligned}
\tag{5.2}
$$

where the equality follows by definition of trace distance (Definition 2.48) and the fact that the output state E is obtained by applying a unitary transformation to the set of registers (E', F, G, W', Z). The inequality is the triangle inequality; in the remainder of the proof, we will show that both terms in (5.2) are upper bounded by ε'.

$$
\begin{aligned}
&\tfrac{1}{2}\|\rho_{WW'E'FGZ|W\neq W'} - \rho_{WW'E'FG|W\neq W'} \otimes 2^{-\ell}\mathbb{I}\|_1 \\
&\quad= \sum_{w\neq w'} P_{WW'|W\neq W'}(w, w')\, d_{\mathsf{unif}}(Z|E'FG, W = w, W' = w') \leq \varepsilon',
\end{aligned}
$$

where the latter inequality follows from (5.1).For the other term, we reason as follows:

$$
\begin{aligned}
&\tfrac{1}{2}\|\rho_{WW'E'FG|W\neq W'} \otimes 2^{-\ell}\mathbb{I} - \rho_{W\leftrightarrow W'\leftrightarrow E'FGZ|W\neq W'}\|_1 \\
&\quad= \tfrac{1}{2}\sum_{w\neq w'} P_{WW'|W\neq W'}(w, w')\, \|\rho^{w,w'}_{E'FG|W\neq W'} \otimes 2^{-\ell}\mathbb{I} - \rho^{w'}_{E'FGZ|W\neq W'}\|_1 \\
&\quad= \tfrac{1}{2}\sum_{w\neq w'} P_{WW'|W\neq W'}(w, w')\, \|\rho^{w,w'}_{E'FG|W\neq W'} \otimes 2^{-\ell}\mathbb{I} \\
&\qquad\quad - \sum_{\substack{w'' \\ \text{s.t. } w''\neq w'}} P_{W|W',W\neq W'}(w''|w')\rho^{w'',w'}_{E'FGZ|W\neq W'}\|_1 \\
&\quad= \tfrac{1}{2}\sum_{w'} P_{W'|W\neq W'}(w')\,\| \sum_{\substack{w \\ \text{s.t. } w\neq w'}} P_{W|W',W\neq W'}(w|w')\rho^{w,w'}_{E'FG|W\neq W'} \otimes 2^{-\ell}\mathbb{I} \\
&\qquad\quad - \sum_{\substack{w'' \\ \text{s.t. } w''\neq w'}} P_{W|W',W\neq W'}(w''|w')\rho^{w'',w'}_{E'FGZ|W\neq W'} \sum_{\substack{w \\ \text{s.t. } w\neq w'}} P_{W|W',W\neq W'}(w|w')\|_1 \\
&\quad= \tfrac{1}{2}\sum_{w\neq w'} P_{WW'|W\neq W'}(w, w')\, \|\rho^{w,w'}_{E'FG|W\neq W'} \otimes 2^{-\ell}\mathbb{I} - \rho^{w,w'}_{E'FGZ|W\neq W'}\|_1 \\
&\quad= \sum_{w\neq w'} P_{WW'|W\neq W'}(w, w')\, d_{\mathsf{unif}}(Z|E'FG, W = w, W' = w') \leq \varepsilon',
\end{aligned}
$$

where the first equality follows by definition of conditional independence (the quantum version, see (2.10) on page 81) and by a basic property of the trace distance; the third and fourth equality follow by linearity of the trace distance. The inequality on the last line follows from (5.1). This proves the claim. □

5.6 User Security in the Single-Qubit-Operations Model

We now consider a dishonest server S^* that can store an unbounded number of qubits. Clearly, against such a S^*, Theorem 5.5 provides no security guarantee anymore. We show here that there is still *some* level of security left. Specifically, we show that `NEWQID` is still secure against a dishonest server S^* that can reliably store all the communicated qubits and measure them qubit-wise and non-adaptively at the end of the protocol. This feature distinguishes our identification protocol from the protocol from [DFSS07], which completely breaks down against such an attack.

5.6.1 The Model

Formally, a dishonest server S^* in the SQOM is modeled as follows.

1. S^* may reliably store the n-qubit state $H^{\mathfrak{c}(w)}|x\rangle = H^{\mathfrak{c}(w)_1}|x_1\rangle \otimes \cdots \otimes H^{\mathfrak{c}(w)_n}|x_n\rangle$ received in step (1) of `NEWQID`.
2. At the end of the protocol, in step (5), S^* chooses an arbitrary sequence $\theta = (\theta_1, \ldots, \theta_n)$, where each θ_i describes an arbitrary orthonormal basis of $\mathbb{C}^2$, and measures each qubit $H^{\mathfrak{c}(w)_i}|x_i\rangle$ in basis θ_i to observe $Y_i \in \mathbb{F}_2$. Hence, we assume that S^* *measures all qubits at the end of the protocol.*
3. The choice of θ may depend on all the classical information gathered during the execution of the protocol, but we assume a *non-adaptive* setting where θ_i does not depend on Y_j for $i \neq j$, i.e., S^* has to choose θ entirely before performing any measurement.

Considering complete projective measurements acting on individual qubits, rather than general single-qubit POVMs, may be considered a restriction of our model. Nonetheless, general POVM measurements can always be described by projective measurements on a bigger system. In this sense, restricting to projective measurements is consistent with the requirement of single-qubit operations. It seems non-trivial to extend our security proof to general single-qubit POVMs.

The restriction to non-adaptive measurements (item 3) is rather strong, even though the protocol from [DFSS07] already breaks down in this non-adaptive setting. The restriction was introduced as a stepping stone towards proving the adaptive case. Up to now, we have unfortunately not yet succeeded in doing so, hence we leave the adaptive case for future research.

We also leave for future research the case of a less restricted dishonest server S^* that can do measurements on blocks that are less stringently bounded in size. Whereas the adaptive versus non-adaptive issue appears to be a proof-technical problem (`NEWQID` looks secure also against an adaptive S^*), allowing measurements on larger blocks will require a new protocol, since `NEWQID` becomes insecure when S^* can do measurements on blocks of size 2, as we show in Section 5.6.5.

5.6.2 No Privacy Amplification

One might expect that proving security of `NEWQID` in the SQOM, i.e., against a dishonest server S^* that is restricted to single-qubit operations should be straightforward, but actually the opposite is true, for the following reason. Even though it is not hard to show that after his measurements, S^* has lower bounded uncertainty in x (except if he was able to guess w), it is not clear how to conclude that $f(x)$ is close to random so that z does not reveal a significant amount of information about w. The reason is that standard privacy amplification fails to apply here. Indeed, the model allows S^* to postpone the measurement of all qubits to step (5) of the protocol. The hash function f, however, is chosen and sent already in step (3). This means that S^* can choose his measurements in step (5) depending on f. As a consequence, the distribution of x from the point of view of S^* may depend on the choice of the hash function f, in which case the privacy-amplification theorem does not give any guarantees.

5.6.3 Single-Qubit Measurements

Consider an arbitrary sequence $\theta = (\theta_1, \ldots, \theta_n)$ where each θ_i describes an orthonormal basis of $\mathbb{C}^2$. Let $|\psi\rangle$ be an n-qubit system of the form

$$|\psi\rangle = H^{b_1}|x_1\rangle \otimes \cdots \otimes H^{b_n}|x_n\rangle,$$

where x and b are arbitrary in $\mathbb{F}_2^n$. Measuring $|\psi\rangle$ qubit-wise in basis θ results in a measurement outcome $Y = (Y_1, \ldots, Y_n) \in \mathbb{F}_2^n$. Suppose that x, b and θ are in fact realizations of the random variables X, B and Θ respectively. It follows immediately from the product structure of the state $|\psi\rangle$ that

$$P_{Y|XB\Theta}(y|x,b,\theta) = \prod_{i=0}^{n} P_{Y_i|X_iB_i\Theta_i}(y_i|x_i,b_i,\theta_i),$$

i.e., the random variables Y_i are statistically independent conditioned on arbitrary fixed values for X_i, B_i and Θ_i but such that $P_{X_iB_i\Theta_i}(x_i,b_i,\theta_i) > 0$.

Lemma 5.6 *The distribution $P_{Y_i|X_iB_i\Theta_i}(y_i|x_i,b_i,\theta_i)$ exhibits the following symmetries:*

$$P_{Y_i|X_iB_i\Theta_i}(0|0,b_i,\theta_i) = P_{Y_i|X_iB_i\Theta_i}(1|1,b_i,\theta_i)$$

and

$$P_{Y_i|X_iB_i\Theta_i}(0|1,b_i,\theta_i) = P_{Y_i|X_iB_i\Theta_i}(1|0,b_i,\theta_i)$$

for all $i \in [n]$, for all b_i and θ_i with $P_{X_iB_i\Theta_i}(\xi,b_i,\theta_i) > 0$ for all $\xi \in \mathbb{F}_2$.

Proof. Let $\alpha, \beta \in \mathbb{C}$ be such that $\theta_i := \{\bar{\alpha}|0\rangle + \bar{\beta}|1\rangle, \bar{\beta}|0\rangle - \bar{\alpha}|1\rangle\}$. (We can always find such α and β.) Writing out the measurement explicitly gives

$$\begin{aligned} P_{Y_i|X_iB_i\Theta_i}(0|x_i,b_i,\theta_i) &= |(\alpha\langle 0| + \beta\langle 1|)H^{b_i}|x_i\rangle|^2 \qquad \text{and} \\ P_{Y_i|X_iB_i\Theta_i}(1|x_i,b_i,\theta_i) &= |(\beta\langle 0| - \alpha\langle 1|)H^{b_i}|x_i\rangle|^2. \end{aligned}$$

Hence, it suffices to prove that

$$|(\alpha\langle 0| + \beta\langle 1|)H^{b_i}|x_i\rangle|^2 = |(\beta\langle 0| - \alpha\langle 1|)H^{b_i}|x_i \oplus 1\rangle|^2 \tag{5.3}$$

for every $x_i, b_i \in \mathbb{F}_2$.

We first show (5.3) for $b_i = 0$. Let σ_1 be the first Pauli matrix defined by $\sigma_1|a\rangle = |a \oplus 1\rangle$ for every $a \in \mathbb{F}_2$. It follows immediately from the definition that σ_1 is a unitary matrix and it is easy to see that σ_1 is Hermitian. Then,

$$\begin{aligned} |(\alpha\langle 0| + \beta\langle 1|)|x_i\rangle|^2 &= |(\alpha\langle 0| + \beta\langle 1|)\sigma_1\sigma_1|x_i\rangle|^2 = |(\alpha\langle 1| + \beta\langle 0|)|x_i \oplus 1\rangle|^2 \\ &= |(\beta\langle 0| - \alpha\langle 1|)|x_i \oplus 1\rangle|^2 \end{aligned}$$

The last equation follows because the expression equals either $|\alpha|^2$ or $|\beta|^2$ (depending on $x_i \in \mathbb{F}_2$), hence we may freely change the sign of α. For $b_i = 1$, we have

$$|(\alpha\langle 0| + \beta\langle 1|)H|x_i\rangle|^2 = |(\alpha\langle 0| + \beta\langle 1|)(|0\rangle + (-1)^{x_i}|1\rangle)|^2 = |\alpha + (-1)^{x_i}\beta|^2$$

and

$$|(\beta\langle 0| - \alpha\langle 1|)H|x_i \oplus 1\rangle|^2 = |(\beta\langle 0| - \alpha\langle 1|)(|0\rangle - (-1)^{x_i}|1\rangle)|^2 = |\beta + (-1)^{x_i}\alpha|^2.$$

We see that those expressions are equal for every $x_i \in \mathbb{F}_2$. □

The symmetry characterized in Lemma 5.6 coincides with that of the *binary symmetric channel*, i.e., we can view Y as a "noisy version" of X, where this noise—produced by the measurement—is independent of X.

Formally, we can write Y as

$$Y = X \oplus \Delta, \tag{5.4}$$

where the random variable $\Delta = (\Delta_1, \ldots, \Delta_n) \in \mathbb{F}_2^n$ thus represents the error between the random variable $X \in \mathbb{F}_2^n$ that is "encoded" in the quantum state and the measurement outcome $Y \in \mathbb{F}_2^n$. By substituting (5.4) in Lemma 5.6, we get the following corollary.

Corollary 5.7 (Independence Between Δ and X) *For every $i \in [n]$ it holds that*

$$P_{\Delta_i|X_iB_i\Theta_i}(\delta_i|x_i, b_i, \theta_i) = P_{\Delta_i|B_i\Theta_i}(\delta_i|b_i, \theta_i)$$

for all $\delta_i \in \{0, 1\}$ and for all x_i, b_i and θ_i such that $P_{X_iB_i\Theta_i}(x_i, b_i, \theta_i) > 0$.

Furthermore, since the random variables Y_i are statistically independent conditioned on fixed values for X_i, B_i and Θ_i, it follows that the Δ_i are statistically independent conditioned on fixed values for B_i and Θ_i.

Definition 5.8 (Quantized Basis) For any orthonormal basis $\theta_i = \{|v_1\rangle, |v_2\rangle\}$ on $\mathbb{C}^2$, we define the *quantized basis* of θ_i as

$$\hat{\theta}_i := j^* \in \mathbb{F}_2, \quad \text{where } j^* \in \arg\max_{j \in \mathbb{F}_2} \max_{k \in \{1,2\}} |\langle v_k|H^j|0\rangle|.$$

If both $j \in \mathbb{F}_2$ attain the maximum, then j^* is chosen arbitrarily from $\mathbb{F}_2$. The quantized basis of the sequence $\theta = (\theta_1, \ldots, \theta_n)$ is naturally defined as the element-wise application of the above, resulting in $\hat{\theta} \in \mathbb{F}_2^n$.

We will use the bias (see Section 2.2.3) as a measure for the predictability of Δ_i.

Theorem 5.9 *When measuring the qubit $H^{b_i}|x_i\rangle$ for any $x_i, b_i \in \mathbb{F}_2$ in any orthonormal basis θ_i on $\mathbb{C}^2$ for which the quantized basis $\hat{\theta}_i$ is the complement of b_i, i.e., $\hat{\theta}_i = b_i \oplus 1$, then the bias of $\Delta_i \in \mathbb{F}_2$, where $\Delta_i = Y_i \oplus x_i$ and $Y_i \in \mathbb{F}_2$ is the measurement outcome, is upper bounded by*

$$\text{bias}(\Delta_i) \leq \frac{1}{\sqrt{2}}.$$

Since the theorem holds for any $x_i \in \mathbb{F}_2$ and since Corollary 5.7 guarantees that Δ_i is independent from an arbitrary random variable X_i, the theorem also applies when we replace x_i by the random variable X_i.

In order to prove Theorem 5.9, we need the following lemma.

Lemma 5.10 *If, for any orthonormal basis θ_i on $\mathbb{C}^2$, there exists a bit $b_i \in \mathbb{F}_2$ so that when measuring the qubit $H^{b_i}|x_i\rangle$ for any $x_i \in \mathbb{F}_2$ in the basis θ_i to obtain $Z_i \in \mathbb{F}_2$ it holds that*

$$\mathrm{bias}(Z_i) \geq 1/\sqrt{2},$$

then it holds that when measuring the qubit $H^{b_i \oplus 1}|x_i\rangle$ in the basis θ_i to obtain $Y_i \in \mathbb{F}_2$,

$$\mathrm{bias}(Y_i) \leq 1/\sqrt{2}.$$

Proof. First note that for any $x_i, b_i \in \mathbb{F}_2$ and any orthonormal basis θ_i on $\mathbb{C}^2$, measuring a state $H^{b_i}|x_i\rangle$ in $\theta_i = \{|v\rangle, |w\rangle\}$ where $|v\rangle = \alpha|0\rangle + \beta|1\rangle$ and $|w\rangle = \beta|0\rangle - \alpha|1\rangle$ gives the same outcome distribution (up to permutations) as when measuring one of the basis states of θ_i (when viewed as a quantum state), say $|w\rangle$, using the basis $\{H^{b_i}|x_i\rangle, H^{b_i}|x_i \oplus 1\rangle\}$. To see why this holds, note that it follows immediately that $|\langle w|H^{b_i}|x_i\rangle|^2 = |\langle x_i|H^{b_i}|w\rangle|^2$. Furthermore, we have already shown in the proof of Lemma 5.6 that

$$|\langle v|H^{b_i}|x_i\rangle|^2 = |\langle w|H^{b_i}|x_i \oplus 1\rangle|^2$$

holds.

Hence, we can apply Theorem 5.1 with $\rho = |w\rangle\langle w|$ (this implies that $n = 1$), $m = 2$ and $\mathcal{B}_0$ and $\mathcal{B}_1$ are the computational and Hadamard basis respectively. The maximum overlap between those bases is $c = 1/\sqrt{2}$. Theorem 5.1 gives us that

$$p_{\max}^{\{|0\rangle,|1\rangle\}} + p_{\max}^{\{|+\rangle,|-\rangle\}} \leq 1 + \frac{1}{\sqrt{2}},$$

where $p_{\max}^{\{|0\rangle,|1\rangle\}}$ and $p_{\max}^{\{|+\rangle,|-\rangle\}}$ respectively denote the maximum probability in the distribution obtained by measuring in the computational and Hadamard basis. By simple manipulations we can write this as a bound on the sum of the biases:

$$\begin{aligned} \frac{2}{\sqrt{2}} &\geq (2p_{\max}^{\{|0\rangle,|1\rangle\}} - 1) + (2p_{\max}^{\{|+\rangle,|-\rangle\}} - 1) \\ &= \mathrm{bias}(Y_i) + \mathrm{bias}(Z_i). \end{aligned} \tag{5.5}$$

From this relation, the claim follows immediately. □

Following [Sch07], we want to remark that both biases in (5.5) are equal to $1/\sqrt{2}$ when θ_i is the *Breidbart basis*, which is the basis that is precisely "in between" the computational and the Hadamard basis:[8]

$$|v\rangle = \cos(\tfrac{\pi}{8})|0\rangle + \sin(\tfrac{\pi}{8})|1\rangle \qquad \text{and} \qquad |w\rangle = \sin(\tfrac{\pi}{8})|0\rangle - \cos(\tfrac{\pi}{8})|1\rangle.$$

Proof of Theorem 5.9 . Let $\theta_i = \{|v_0\rangle, |v_1\rangle\}$. We will make a case distinction based on the value of

$$\mu := \max_{k \in \mathbb{F}_2} |\langle v_k | H^{\hat{\theta}_i} |0\rangle|. \tag{5.6}$$

If $\mu \leq \cos(\pi/8)$, then we also have that $\max_{k \in \mathbb{F}_2} |\langle v_k | H^{b_i} |x_i\rangle| \leq \cos(\pi/8)$ where $b_i = \hat{\theta}_i \oplus 1$, this holds by definition of the quantized basis (Definition 5.8). Then, the probability of obtaining outcome $Y_i = k^*$, where $k^* \in \mathbb{F}_2$ achieves the maximum in (5.6), is bounded by

$$P_{Y_i}(k^*) = |\langle v_{k^*} | H^{b_i} |x_i\rangle|^2 \leq \cos^2(\pi/8) = \tfrac{1}{2} + \tfrac{1}{2\sqrt{2}}.$$

Hence,

$$\text{bias}(\Delta_i) = \text{bias}(Y_i) = |P_{Y_i}(k^*) - (1 - P_{Y_i}(k^*))| = |2P_{Y_i}(k^*) - 1| \leq \tfrac{1}{\sqrt{2}}.$$

If $\mu > \cos(\pi/8)$, then when measuring the state $H^{\hat{\theta}_i}|x_i\rangle$ in θ_i to obtain $Z_i \in \mathbb{F}_2$, we have that $\text{bias}(Z_i) > 1/\sqrt{2}$ (this follows from similar computations as performed above). We now invoke Lemma 5.10 to conclude that when measuring the state $H^{b_i}|x_i\rangle$ in θ_i to obtain Y_i, $\text{bias}(\Delta_i) = \text{bias}(Y_i) < \frac{1}{\sqrt{2}}$. □

5.6.4 User Security of `NEWQID`

We are now ready to state and prove the security of `NEWQID` against a dishonest user in the SQOM.

Theorem 5.11 (User Security) *Let S^* be a dishonest server with unbounded quantum storage that is restricted to non-adaptive single-qubit operations, as specified in Section 5.6.1. Then, for any $\beta \in \mathbb{R}$ such that $0 < \beta < \frac{1}{4}$, user security (as defined in Definition 2.67) holds with*

$$\varepsilon \leq \tfrac{1}{2} 2^{\frac{1}{2}\ell - \frac{1}{4}(\frac{1}{4} - \beta)d} + \tbinom{m}{2} 2^{2\ell} \exp(-2d\beta^2)$$

Note that d is typically linear in n whereas ℓ is chosen independently of n, hence the expression above is negligible in d.

[8] In [Sch07], the corresponding state is called the "Hadamard-invariant state."

To prove Theorem 5.11 we need the following technical lemma and corollary. Recall that $\mathcal{F}$ denotes the class of all linear functions from $\mathbb{F}_2^n$ to $\mathbb{F}_2^\ell$, where $\ell < n$, represented as $\ell \times n$ matrices over $\mathbb{F}_2$. When $F \in \mathcal{F}$ acts on an n-bit vector $x \in \mathbb{F}_2^n$, we prefer the notation $F(x)$ over matrix-product notation Fx.[9] Furthermore, we write $\mathrm{span}(F)$ for the *row* span of F: the set of vectors obtained by making all possible $\mathbb{F}_2$ linear combinations of the rows of F, i.e., the set $\{sF : \forall s \in \mathbb{F}_2^\ell\}$, where s should be interpreted as a row vector and sF denotes a vector-matrix product. For two vectors $v, w \in \mathbb{F}_2^n$, the *Schur product* is defined as the element-wise product $v \odot w := (v_1w_1, v_2w_2, \dots, v_nw_n) \in \mathbb{F}_2^n$, and the *inner product* between v and w is given by $v \bullet w := v_1w_1 \oplus \cdots \oplus v_nw_n \in \mathbb{F}_2$. For an n-bit vector vector $v = (v_1, \dots, v_n)$ in $\mathbb{F}_2^n$, we write $|v|$ for its Hamming weight (as defined in Section 3.1.3), and, for any subset $\mathcal{I} \subseteq [n]$, we write $v_\mathcal{I}$ for the restricted vector $(v_i)_{i\in\mathcal{I}} \in \mathbb{F}_2^{|\mathcal{I}|}$.

Lemma 5.12 *Let n, k and ℓ be arbitrary positive integers, let $0 < \beta < \frac{1}{4}$ and let $\mathcal{I} \subset [n]$ such that $|\mathcal{I}| \geq k$, and let F be uniform over $\mathcal{F} = \mathbb{F}_2^{\ell\times n}$. Then, it holds except with probability $2^{2\ell}\exp(-2k\beta^2)$ (the probability is over the random matrix F) that*

$$|(f \odot g)_\mathcal{I}| > (\tfrac{1}{4} - \beta)k \qquad \forall f, g \in \mathrm{span}(F) \setminus \{\mathbf{0}\}$$

Proof. Without loss of generality, we will assume that $|\mathcal{I}| = k$. Now take arbitrary but non-zero vectors $r, s \in \mathbb{F}_2^\ell$ and let $V := rF$ and $W := sF$. We will analyze the case $r \neq s$; the case $r = s$ is similar but simpler. Because each element of F is an independent random bit, and r and s are non-zero and $r \neq s$, V and W are independent and uniformly distributed n-bit vectors with expected relative Hamming weight $1/2$. Hence, on average $|(V \odot W)_\mathcal{I}|$ equals $k/4$. Furthermore, using Hoeffding's inequality (Theorem 2.11), we may conclude that

$$\Pr\left[\frac{k}{4} - |(V \odot W)_\mathcal{I}| > \beta k\right] = \Pr\left[|(V \odot W)_\mathcal{I}| < (\tfrac{1}{4} - \beta)k\right] \leq \exp(-2k\beta^2)\,.$$

Finally, the claim follows by applying the union bound over the choice of r and s (each 2^ℓ possibilities). □

Recall that $\mathcal{C} \subset \mathbb{F}_2^n$ is a binary code with minimum distance d, $\mathfrak{c}(\cdot)$ its encoding function, and that $m := |\mathcal{W}|$.

[9] When using matrix-product notation ambiguities could arise, e.g., in subscripts of probability distributions like P_{FX}: then it is not clear whether this means the joint distribution of F and X or the distribution of F acting on X?

Corollary 5.13 *Let $0 < \beta < \frac{1}{4}$, and let F be uniformly distributed over $\mathcal{F}$. Then, F has the following property except with probability $\binom{m}{2}2^{2\ell}\exp(-2d\beta^2)$: for any string $s \in \mathbb{F}_2^n$ (possibly depending on the choice of F), there exists at most one $\widetilde{c} \in \mathcal{C}$ such that for any code word $c \in \mathcal{C}$ different from $\widetilde{c}$, it holds that*

$$|f \odot (c \oplus s)| \geq \tfrac{1}{2}(\tfrac{1}{4} - \beta)d \qquad \forall f \in \mathrm{span}(F) \setminus \{\mathbf{0}\}$$

We prove the statement by arguing for two $\widetilde{c}$'s and showing that they must be identical. In the proof, we will make use of elementary properties of the Schur product and the Hamming weight:

1. $|a| \geq |a \odot b|$ for all $a, b \in \mathbb{F}_2^n$. (Follows immediately.)
2. $|a \odot b| + |a \odot c| \geq |a \odot (b \oplus c)|$ for all $a, b, c \in \mathbb{F}_2^n$.
   Proof. $|a \odot (b \oplus c)| = |a \odot b \oplus a \odot c| \leq |a \odot b| + |a \odot c|$, where the equality is the distributivity of the Schur product, and the inequality is the triangle inequality for the Hamming weight. □

Proof. By Lemma 5.12 with $\mathcal{I} := \{i \in [n] : c_i \neq c'_i\}$ for $c, c' \in \mathcal{C}$, and by applying the union bound over all possible pairs (c, c'), we obtain that except with probability $\binom{m}{2}2^{2\ell}\exp(-2d\beta^2)$ (over the choice of F), it holds that

$$|f \odot g \odot (c \oplus c')| > (\tfrac{1}{4} - \beta)d \tag{5.7}$$

for all $f, g \in \mathrm{span}(F) \setminus \{\mathbf{0}\}$ and all $c, c' \in \mathcal{C}$ with $c \neq c'$.

Now, for such an F, and for every choice of $s \in \mathbb{F}_2^n$, consider $\widetilde{c}_1, \widetilde{c}_2 \in \mathcal{C}$ and $f_1, f_2 \in \mathrm{span}(F) \setminus \{\mathbf{0}\}$ such that

$$|f_1 \odot (\widetilde{c}_1 \oplus s)| < \tfrac{1}{2}(\tfrac{1}{4} - \beta)d \quad \text{and} \quad |f_2 \odot (\widetilde{c}_2 \oplus s)| < \tfrac{1}{2}(\tfrac{1}{4} - \beta)d.$$

We will show that this implies $\widetilde{c}_1 = \widetilde{c}_2$, which proves the claim. Indeed, we can write

$$\begin{aligned}(\tfrac{1}{4} - \beta)d &> |f_1 \odot (\widetilde{c}_1 \oplus s)| + |f_2 \odot (\widetilde{c}_2 \oplus s)| \\ &\geq |f_1 \odot f_2 \odot (\widetilde{c}_1 \oplus s)| + |f_1 \odot f_2 \odot (\widetilde{c}_2 \oplus s)| \geq |f_1 \odot f_2 \odot (\widetilde{c}_1 \oplus \widetilde{c}_2)|\end{aligned}$$

where the second inequality is property (1) from above applied twice and the third inequality is property (2). This contradicts (5.7) unless $\widetilde{c}_1 = \widetilde{c}_2$. □

Proof of Theorem 5.11. Consider an execution of NEWQID, with a dishonest server S^* as described in Section 5.6.1. We let W, X and Z be the random variables that describe the values w, x and z occurring in the protocol.

From NEWQID's description, we see that F is uniform over $\mathcal{F}$. Hence, by Corollary 5.13 it will be "good" (in the sense that the bound from Corollary 5.13 holds) except with probability $\binom{m}{2}2^{2\ell}\exp(-2d\beta^2)$. From here, we consider a fixed choice for F and condition on the event that it is "good," we will take the probability that F is "bad" into account at the end of the analysis. Although we have fixed F, we will keep using capital notation for it, to emphasize that F is a matrix. We also fix $G = g$ for an arbitrary g; the analysis below holds for any such choice.

Let Θ describe the qubit-wise measurement performed by S^* at the end of the execution, and Y the corresponding measurement outcome. By the non-adaptivity restriction and by the requirement in Definition 2.67 that S^* is initially independent of W, we may conclude that, once G and F are fixed, Θ is a function of Z. (Recall that $Z = F(X) \oplus g(W)$.)

We will define W' with the help of Corollary 5.13. Let $\hat{\Theta}$ be the quantized basis of Θ, as defined in Definition 5.8. Given a fixed value θ for Θ, and thus a fixed value $\hat{\theta}$ for $\hat{\Theta}$, we set s, which is a variable that occurs in Corollary 5.13, to $s = \hat{\theta}$. Corollary 5.13 now guarantees that there exists *at most one* $\tilde{c}$. If $\tilde{c}$ indeed exists, then we choose w' such that $\mathfrak{c}(w') = \tilde{c}$. Otherwise, we pick $w' \in \mathcal{W}$ arbitrarily (any choice will do). Note that this defines the random variable W', and furthermore note that $Z \to \Theta \to \hat{\Theta} \to W'$ forms a Markov chain. Moreover, by the choice of w' it immediately follows from Corollary 5.13 that for all $w \neq w'$ and for all $f \in \operatorname{span}(F) \setminus \{\mathbf{0}\}$ it holds that

$$|f \odot (\mathfrak{c}(w) \oplus \hat{\theta})| \geq \tfrac{1}{2}(\tfrac{1}{4} - \beta)d. \tag{5.8}$$

We will make use of this bound later in the proof.

Since the model (Section 5.6.1) enforces the dishonest server to measure all qubits at the end of the protocol, the system $E = (Y, Z, \Theta)$ is classical and hence the trace-distance-based user-security definition (Definition 2.67) simplifies to a bound on the statistical distance between distributions. I.e., it is sufficient to prove that

$$\mathrm{SD}(P_{EW|W'=w',W'\neq W}, P_{W|W'=w',W\neq W'}P_{E|W'=w',W\neq W'}) \leq \varepsilon$$

holds for any w'. Consider the distribution that appears above as the first argument to the statistical distance, i.e., $P_{EW|W'=w',W'\neq W}$. By substituting $E = (Y, Z, \Theta)$, it factors as follows[10]

$$\begin{aligned} P_{YZ\Theta W|W',W\neq W'} &= P_{W|W',W\neq W'}\, P_{Z\Theta|WW',W\neq W'}\, P_{Y|Z\Theta WW',W\neq W'} \\ &= P_{W|W',W\neq W'}\, P_{Z\Theta|W',W\neq W'}\, P_{Y|F(X)\Theta WW',W\neq W'}, \end{aligned} \tag{5.9}$$

where the equality $P_{Z\Theta|WW',W\neq W'} = P_{Z\Theta|W',W\neq W'}$ holds by the following argument: Z is independent of W (since $F(X)$ acts as one-time pad) and $Z \to \Theta \to W'$ is a Markov chain, and S^* (who computes Θ from Z) is initially independent of W by Definition 2.67, hence W is independent of Z, Θ and W', which implies the above equality. The equality $P_{Y|Z\Theta WW',W\neq W'} = P_{Y|F(X)\Theta WW',W\neq W'}$ holds by the observation that given W, Z is uniquely determined by $F(X)$ and vice versa.

In the remainder of this proof we will show that

$$d_{\mathsf{unif}}(Y|F(X)=u,\Theta=v,W=w,W'=w') \leq \tfrac{1}{2}2^{\frac{\ell}{2}-\frac{1}{4}(\frac{1}{4}-\beta)d},$$

for all u, v, w such that $w \neq w'$, where w' is determined by v. This then implies that the rightmost factor in (5.9) is essentially independent of W, and concludes the proof.

To simplify notation, we define $\mathcal{E}$ to be the event

$$\mathcal{E} := \{F(X)=u,\Theta=v,W=w,W'=w'\}$$

for fixed but arbitrary choices u, v and w such that $w \neq w'$, where w' is determined by v. We show closeness to the uniform distribution by using the XOR inequality from Diaconis *et al.* (Theorem 2.8), i.e., we use the inequality

$$d_{\mathsf{unif}}(Y|\mathcal{E}) \leq \tfrac{1}{2}\Big[\sum_{\alpha} \mathrm{bias}(\alpha\bullet Y|\mathcal{E})^2\Big]^{\frac{1}{2}},$$

where the sum is over all α in $\mathbb{F}_2^n \setminus \{\mathbf{0}\}$. We split this sum into two parts, one for $\alpha \in \mathrm{span}(F)$ and one for α not in $\mathrm{span}(F)$, and analyze the two parts separately.

Since X is uniformly distributed, it follows that for any $\alpha \notin \mathrm{span}(F)$, it holds that $P_{\alpha\bullet X|F(X)}(\cdot|u) = \frac{1}{2}$ (for any u). We conclude that

$$\begin{aligned} \tfrac{1}{2} = P_{\alpha\bullet X|F(X)} = P_{\alpha\bullet X|F(X)W} = P_{\alpha\bullet X|F(X)\Theta WW'} \\ = P_{\alpha\bullet Y|F(X)\Theta WW'} = P_{\alpha\bullet Y|\mathcal{E}} \quad \forall\alpha \notin \mathrm{span}(F). \end{aligned}$$

[10] Note that Convention 2.2 applies here.

The second equality follows since W is independent of X. The third equality holds by the fact that Θ is computed from $F(X) \oplus g(W)$ and W' is determined by Θ. The fourth equality follows by the security of the one-time pad, i.e., recall that $Y = X \oplus \Delta$, where by Corollary 5.7 it holds that $\Delta \in \mathbb{F}_2^n$ is independent of X when conditioned on fixed values for $B = \mathfrak{c}(W)$ and Θ. Hence, it follows that $\text{bias}(\alpha \bullet Y|\mathcal{E}) = 0$ for $\alpha \notin \text{span}(F)$.

For any non-zero $\alpha \in \text{span}(F)$, we can write

$$\begin{aligned}
\text{bias}(\alpha \bullet Y|\mathcal{E}) &= \text{bias}(\alpha \bullet (X \oplus \Delta)|\mathcal{E}) \\
&= \text{bias}(\alpha \bullet X \oplus \alpha \bullet \Delta|\mathcal{E}) && \text{(distributivity of dot product)} \\
&= \text{bias}(\alpha \bullet X|\mathcal{E})\text{bias}(\alpha \bullet \Delta|\mathcal{E}) && \text{(Corollary 5.7)} \\
&\leq \text{bias}(\alpha \bullet \Delta|\mathcal{E}) && (\text{bias}(\alpha \bullet X) \leq 1) \\
&= \prod_{i \in [n]} \text{bias}(\alpha_i \cdot \Delta_i|\mathcal{E}) && (\Delta_i \text{ independent}) \\
&= \prod_{i \in [n]:\alpha_i = 1} \text{bias}(\Delta_i|\mathcal{E}) \\
&\leq \prod_{\substack{i \in [n]:\alpha_i = 1 \\ \hat{\theta}_i = \mathfrak{c}(w)_i \oplus 1}} 2^{-\frac{1}{2}} && \text{(Theorem 5.9)} \\
&= 2^{-\frac{1}{2}|\alpha \odot (\mathfrak{c}(w) \oplus \hat{\theta})|} \leq 2^{-\frac{1}{4}(\frac{1}{4} - \beta)d} && \text{(by (5.8))}
\end{aligned}$$

Combining the two parts, we get

$$\begin{aligned}
d_{\mathsf{unif}}(Y|\mathcal{E}) &\leq \tfrac{1}{2}\Big[\sum_{\alpha} \text{bias}(\alpha \bullet Y|\mathcal{E})^2\Big]^{\frac{1}{2}} \\
&= \tfrac{1}{2}\Big[\sum_{\alpha \in \text{span}(F) \setminus \{\mathbf{0}\}} \text{bias}(\alpha \bullet Y|\mathcal{E})^2 + 0\Big]^{\frac{1}{2}} \leq \tfrac{1}{2} 2^{\frac{\ell}{2} - \frac{1}{4}(\frac{1}{4} - \beta)d}\,.
\end{aligned}$$

Incorporating the error probability of having a "bad" F completes the proof. □

5.6.5 Attack against NEWQID using Operations on Pairs of Qubits

We present an attack with which the dishonest server S^* can discard two passwords in one execution of NEWQID using coherent operations on pairs of qubits.

Before discussing this attack, we first explain a straightforward strategy by which S^* can discard one password per execution: S^* chooses a candidate password $\hat{w} \in \mathcal{W}$

and measures the state $H^{\mathfrak{c}(W)}|X\rangle$ qubit-wise in the basis $H^{\mathfrak{c}(\hat{w})}$ to obtain $Y \in \mathbb{F}_2^n$. S^* then computes $F(Y) \oplus g(\hat{w})$ and compares this to $Z = F(X) \oplus g(W)$, which he received from the user. If indeed $Z = F(Y) \oplus g(\hat{w})$, then it is very likely that $W = \hat{w}$, i.e., that S^* guessed the password correctly.

Let us now explain the attack, which is obtained by modifying the above strategy. The attack is based on the following observation [DFSS05]: if S^* can perform Bell measurements on qubit pairs $H^a|x_1\rangle \otimes H^a|x_2\rangle$, for $a, x_1, x_2 \in \mathbb{F}_2$, then he can learn the parity of $x_1 \oplus x_2$ for both choices of a simultaneously. This strategy can also be adapted to determine both parities of a pair in which the first qubit is encoded in a basis that is opposite to that of the second qubit, i.e., by appropriately applying a Hadamard gate prior to applying the Bell measurement.

Let the first bit of Z be equal to $f \bullet X \oplus g(W)_1$,[11] where $f \in \mathrm{span}(F) \setminus \{\mathbf{0}\}$. Let $\hat{w}_1, \hat{w}_2 \in \mathcal{W}$ be two candidate passwords. With the trick from above, S^* can measure the positions in the set

$$\mathcal{P} := \{i \in [n] : f_i = 1, \mathfrak{c}(\hat{w}_1)_i = 1 \oplus \mathfrak{c}(\hat{w}_2)_i\}$$

pairwise (assuming $|\mathcal{P}|$ to be even) using Bell measurements, while measuring the positions where $\mathfrak{c}(\hat{w}_1)$ and $\mathfrak{c}(\hat{w}_2)$ coincide using ordinary single-qubit measurements. This allows him to compute both "check bits" corresponding to both passwords *simultaneously*, i.e., those check bits coincide with $f \bullet Y_1 \oplus g(\hat{w}_1)_1$ and $f \bullet Y_2 \oplus g(\hat{w}_2)_1$, where $Y_1 \in \mathbb{F}_2^n$ and $Y_2 \in \mathbb{F}_2^n$ are the outcomes that S^* would have obtained if he had measured all qubits qubit-wise in either $\mathfrak{c}(\hat{w}_1)$ or $\mathfrak{c}(\hat{w}_2)$, respectively. If both these check bits are different from the bit Z_1, then S^* can discard both w_1 and w_2.

We have seen that in the *worst case*, the attack is capable of discarding two passwords in one execution, and hence clearly violates the security definition. On *average*, however, the attack seems to discard just one password per execution, i.e., a candidate password cannot be discarded if its check bit is consistent with Z_1, which essentially happens with probability $1/2$. This raises the question whether the security definition is unnecessarily strong, because it seems that not being able to discard more than one password on average would be sufficient. Apart from this, it might be possible to improve the attack, e.g., by selecting the positions where to measure pairwise in a more clever way, as to obtain multiple check bits (corresponding to multiple fs in the span of F) per candidate password, thereby increasing the probability of discarding a wrong candidate password.

[11] By $g(W)_1$ we mean the first bit of $g(W)$.

5.7 Conclusion

We view our work related to NEWQID as a first step in a promising line of research, aimed at achieving security in multiple models simultaneously. The main open problem in the context of the SQOM is to reprove our results in a more general model in which the dishonest server S^* can choose his basis adaptively. Also, it would be interesting to see whether similar results can be obtained in a model where the adversary is restricted to performing quantum operations on blocks of several qubits.

A

Notation

Symbol	Description	Page
$\mathbb{R}$	field of the real numbers	
$\mathbb{C}$	field of the complex numbers	
$\mathbb{Z}$	ring of rational integers	
$\mathbb{N}$	set of strictly positive integers	
$\mathbb{F}_q$	finite field of order q	42
$\mathbb{F}_q^*$	multiplicative group of $\mathbb{F}_q$	42
$\mathbf{0}$	zero vector in an $\mathbb{F}_2$ vector space of arbitrary dimension	48
$O(\cdot)$	Bachmann–Landau Big-Oh notation	43
$\Theta(\cdot)$	Bachmann–Landau Big-Theta notation	43
$[n]$	set of integers $\{1, \ldots, n\}$	
$[a, b]$	interval $\{x \in \mathbb{R} : a \leq x \leq b\}$	
log	binary logarithm	
e	base of the natural logithm (2.718 . . .)	
$\bar{a}$	complex conjugate of $a \in \mathbb{C}$	
$\|a\|$	absolute value of $a \in \mathbb{C}$	
$\|\mathcal{X}\|$	cardinality of the set $\mathcal{X}$	
$\|v\|$	Hamming weight of $v \in \{0, 1\}^n$ (used exclusively in Chapter 5)	
$\mathrm{wt}(v)$	Hamming weight of $v \in \{0, 1\}^n$ (used exclusively in Chapter 3)	99
$\mathrm{span}(\mathcal{S})$	linear span of the elements in the set $\mathcal{S}$	
$\mathrm{span}(A)$	linear span of the rows of matrix A	
A^T	transpose of matrix A	
$A^\dagger$	Hermitian transpose of complex matrix A	
$A \geq 0$	matrix A is positive semi-definite	
δ_{ij}	Kronecker delta symbol	42

$\oplus$	addition operator in $\mathbb{F}_2^n$ and $\mathbb{F}_{2^n}$	
$\otimes$	tensor product / Kronecker matrix product	65
$A^{\otimes n}$	n-fold Kronecker product of matrix A	73
$\mathcal{H}^{\otimes n}$	n-fold tensor product of Hilbert space $\mathcal{H}$	
$\odot$	Schur product	188
$\bullet$	standard inner product on $\mathbb{F}_2^n$	188
$x \overset{\mathrm{r}}{\leftarrow} \mathcal{X}$	x is picked independently and uniformly at random from the set $\mathcal{X}$	
$\mathrm{supp}(P_X)$	support of the distribution P_X	44
$h(p)$	binary entropy function of p	52
$H_{\min}(X)$	min-entropy of X	52
$p_{\mathsf{guess}}(X\|E)$	conditional guessing probability of X given E	84
$H_{\min}(A\|B)$	conditional min-entropy of A given B	84
$H_{\max}(A)$	max-entropy of A	85
H	2×2 Hadamard matrix	73
$\dim(\mathcal{H})$	dimension of $\mathcal{H}$	
$\mathrm{rank}(\rho)$	rank of ρ	
$\mathrm{tr}(\rho)$	trace of ρ	
$\mathrm{tr}_B(\rho_{AB})$	partial trace over B	73
$\mathbb{I}$	identity operator	64
$\mathrm{Hom}(\mathcal{H}, \mathcal{H}')$	complex vector space of linear maps $\mathcal{H} \to \mathcal{H}'$	
$\mathrm{End}(\mathcal{H})$	complex algebra of operators $\mathcal{H} \to \mathcal{H}$	
$\mathcal{D}(\mathcal{H})$	set of density operators on $\mathcal{H}$	70
$\|\varphi\rangle$	ket vector	63
$\langle\varphi\|$	bra vector	63
$\langle\varphi\|\psi\rangle$	inner product between $\|\varphi\rangle$ and $\|\psi\rangle$	63
$\|\varphi\rangle\langle\psi\|$	outer product between $\|\varphi\rangle$ and $\|\psi\rangle$	63
$\|\varphi\rangle\langle\varphi\|$	rank-1 projector (if and only if $\|\varphi\rangle$ has norm 1)	65
$\\|\rho\\|_1$	trace norm of ρ	81
$\delta(\rho, \sigma)$	trace distance between ρ and σ	81
$\mathrm{SD}(p, q)$	statistical distance between p and q	47
$d_{\mathsf{unif}}(X\|E)$	distance-to-uniform of X when given E	83

Bibliography

[AD96] Miklós Ajtai and Cynthia Dwork. A public-key cryptosystem with worst-case/average-case equivalence. *Electronic Colloquium on Computational Complexity*, 3(65), 1996.

[AS00] Noga Alon and Joel Spencer. *The Probabilistic Method (Second Edition)*. Wiley, 2000.

[BB84] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing*, pages 175–179, 1984.

[BBB$^+$92] Charles Bennett, François Bessette, Gilles Brassard, Louis Salvail, and John Smolin. Experimental quantum cryptography. *J. Crypt.*, 5:3–28, 1992.

[BBB$^+$00] Eli Biham, Michel Boyer, P. Oscar Boykin, Tal Mor, and Vwani Roychowdhury. A proof of the security of quantum key distribution (extended abstract). In *STOC*, pages 715–724, New York, 2000. ACM.

[BBBW82] Charles H. Bennett, Gilles Brassard, Seth Breidbart, and Stephen Wiesner. Quantum cryptography, or unforgeable subway tokens. In *CRYPTO*, LNCS, pages 267–275, 1982.

[BBCS91] Charles H. Bennett, Gilles Brassard, Claude Crépeau, and Marie-Hélène Skubiszewska. Practical quantum oblivious transfer. In *CRYPTO*, LNCS, pages 351–366, 1991.

[BBM92] Charles H. Bennett, Gilles Brassard, and N. David Mermin. Quantum cryptography without Bell's theorem. *Phys. Rev. Lett.*, 68:557–559, Feb 1992.

[BBM95] Charles Bennett, Gilles Brassard, and Ueli M. Maurer. Generalized privacy amplification. *IEEE Tran. Inf. Th.*, 41:1915–1923, 1995.

[BBR88] Charles H. Bennett, Gilles Brassard, and Jean-Marc Robert. Privacy amplification by public discussion. *SIAM J. Comput.*, 17(2):210–229, April 1988.

[BC90] Gilles Brassard and Claude Crépeau. Quantum bit commitment and coin tossing protocols. In *CRYPTO*, LNCS, pages 49–61, 1990.

[BCF$^+$11] Harry Buhrman, Nishanth Chandran, Serge Fehr, Ran Gelles, Vipul Goyal, Rafail Ostrovsky, and Christian Schaffner. Position-based quantum cryptography: impossibility and constructions. In *CRYPTO*, LNCS, pages 429–446. Springer, 2011. arXiv:1009.2490.

[BCS12] H. Buhrman, M. Christandl, and C. Schaffner. Complete insecurity of quantum protocols for classical two-party computation. arXiv:1201.0849, January 2012.

[BF10] Niek J. Bouman and Serge Fehr. Sampling in a quantum population, and applications. In Tal Rabin, editor, *CRYPTO*, volume 6223 of *LNCS*, pages 724–741. Springer, 2010. arXiv:0907.4246.

[BF11] Niek J. Bouman and Serge Fehr. Secure authentication from a weak key, without leaking information. In Kenneth Paterson, editor, *Eurocrypt*, volume 6632 of *LNCS*, pages 246–265. Springer, 2011. ePrint:2011/034.

[BFGS12] Niek J. Bouman, Serge Fehr, Carlos González-Guillén, and Christian Schaffner. An all-but-one entropic uncertainty relation, and application to password-based identification. In K. Iwama, Y. Kawano, and M. Murao, editors, *Theory of Quantum Computation, Communication and Cryptography (TQC2012)*, volume 7582 of *LNCS*, pages 29–44. Springer, 2012. arXiv:1105.6212.

[BFSS11] Harry Buhrman, Serge Fehr, Christian Schaffner, and Florian Speelman. The garden-hose game: A new model of computation, and application to position-based quantum cryptography. arXiv:1109.2563, September 2011.

[BK11] Salman Beigi and Robert König. Simplified instantaneous non-local quantum computation with applications to position-based cryptography. *New Journal of Physics*, 13(9):093036, 2011. arXiv:1101.1065.

[BLMS00] Gilles Brassard, Norbert Lütkenhaus, Tal Mor, and Barry C. Sanders. Limitations on practical quantum cryptography. *Phys. Rev. Lett.*, 85(6):1330–1333, August 2000. arXiv:quant-ph/9911054.

[Blu81] Manuel Blum. Coin flipping by telephone. In *CRYPTO*, pages 11–15, 1981.

[BM10] D. Bruss and T. Meyer. Quantum cryptography. In Fabio Benatti, Mark Fannes, Roberto Floreanini, and Dimitri Petritis, editors, *Quantum Information, Computation and Cryptography*, volume 808 of *Lecture Notes in Physics*, pages 277–308. Springer, 2010.

[BOHL$^+$05] Michael Ben-Or, Michal Horodecki, Debbie W. Leung, Dominic Mayers, and Jonathan Oppenheim. The universal composable security of quantum key distribution. In *TCC*, pages 386–406, 2005. arXiv:quant-ph/0409078.

[Bru98] Dagmar Bruss. Optimal eavesdropping in quantum cryptography with six states. *Phys. Rev. Lett.*, 81:3018, 1998. arXiv:quant-ph/9805019.

[BS93] Gilles Brassard and Louis Salvail. Secret-key reconciliation by public discussion. In Eli Biham, editor, *Eurocrypt*, volume 2656 of *LNCS*, pages 410–423. Springer, 1993.

[Cac97] Christian Cachin. *Entropy Measures and Unconditional Security in Cryptography*. PhD thesis, ETH Zürich (Switzerland), 1997.

[CF11] Ronald Cramer and Serge Fehr. The mathematical theory of information, and applications (lecture notes, version 2.0), 2011. Mathematical Institute, Leiden University.

[CGMO09] Nishanth Chandran, Vipul Goyal, Ryan Moriarty, and Rafail Ostrovsky. Position based cryptography. In Shai Halevi, editor, *CRYPTO*, volume 5677 of *LNCS*, pages 391–407. Springer, 2009. ePrint:2009/364.

[CHZ$^+$09] Wei Chen, Zheng-Fu Han, Tao Zhang, et al. Field experiment on a star-type metropolitan quantum key distribution network. *Photonics Technology Letters, IEEE*, 21(9):575–577, may 2009. arXiv:0708.3546.

[CK09] A. Chailloux and I. Kerenidis. Optimal quantum strong coin flipping. In *FOCS*, pages 527–533, oct 2009.

[CKOR10] Nishanth Chandran, Bhavana Kanukurthi, Rafail Ostrovsky, and Leonid Reyzin. Privacy amplification with asymptotically optimal entropy loss. In *STOC*, pages 785–794. ACM, 2010.

[Col07] Roger Colbeck. Impossibility of secure two-party classical computation. *Phys. Rev. A*, 76:062308, Dec 2007. arXiv:0708.2843.

[CP11] Kai-Min Chung and Rafael Pass. The randomness complexity of parallel repetition. In *FOCS*, pages 658–667, 2011.

[CT06] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory 2nd Edition*. Wiley Series in Telecommunications and Signal Processing. Wiley Interscience, July 2006.

[CW77] J. Lawrence Carter and Mark N. Wegman. Universal classes of hash functions. In *STOC*, pages 106–112, New York, 1977. ACM.

[CW81] J. Lawrence Carter and Mark N. Wegman. New hash functions and their use in authentication and set equality. *Journal of Computer and System Sciences*, 22(3):265–279, jun 1981.

[DFL$^+$09] Ivan Damgård, Serge Fehr, Carolin Lunemann, Louis Salvail, and Christian Schaffner. Improving the security of quantum protocols via commit-and-open. In *CRYPTO*, LNCS, pages 408–427. Springer, 2009. arXiv:0902.3918.

[DFR$^+$07] Ivan B. Damgård, Serge Fehr, Renato Renner, Louis Salvail, and Christian Schaffner. A tight high-order entropic quantum uncertainty relation with applications. In *CRYPTO*, LNCS, pages 360–378. Springer, 2007. arXiv:quant-ph/0612014.

[DFSS05] Ivan Damgård, Serge Fehr, Louis Salvail, and Christian Schaffner. Cryptography in the bounded quantum-storage model. In *FOCS*, pages 449–458. IEEE, 2005. arXiv:quant-ph/0508222.

[DFSS07] Ivan Damgård, Serge Fehr, Louis Salvail, and Christian Schaffner. Secure identification and QKD in the bounded-quantum-storage model. In *CRYPTO*, volume 4622 of *LNCS*, pages 342–359. Springer, 2007. arXiv:0708.2557.

[DH76] Whitfield Diffie and Martin E. Hellman. New directions in cryptography. *IEEE Tran. Inf. Th.*, 22(6):644–654, 1976.

[DHL$^+$04] D. P. DiVincenzo, M. Horodecki, D. W. Leung, J. A. Smolin, and B. M. Terhal. Locking classical correlations in quantum states. *Phys. Rev. Lett.*, 92(6), February 2004. arXiv:quant-ph/0303088.

[Dia88] P. Diaconis. *Group Representations in Probability and Statistics*, volume 11 of *Lecture Notes — Monograph series*. Institute of Mathematical Statistics, Hayward, CA, 1988.

[DK02] C. Doescher and M. Keyl. An introduction to quantum coin-tossing. arXiv:quant-ph/0206088, June 2002.

[DORS08] Yevgeniy Dodis, Rafail Ostrovsky, Leonid Reyzin, and Adam Smith. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. *SIAM J. Comput.*, 38(1):97–139, 2008. ePrint:2003/235.

[DP07] Stefan Dziembowski and Krzysztof Pietrzak. Intrusion-resilient secret sharing. In *FOCS*, pages 227–237. IEEE Computer Society, 2007. ePrint:2007/359.

[DP08] Stefan Dziembowski and Krzysztof Pietrzak. Leakage-resilient cryptography. In *FOCS*, pages 293–302, 2008.

[DP09] Devdatt P. Dubhashi and Alessandro Panconesi. *Concentration of Measure for the Analysis of Randomized Algorithms*. Cambridge Univ. Press, 2009.

[DPVR09] Anindya De, Christopher Portmann, Thomas Vidick, and Renato Renner. Trevisan's extractor in the presence of quantum side information. *arXiv*, 2009. arXiv:0912.5514.

[DR00] Joan Daemen and Vincent Rijmen. Rijndael for AES. In *AES Candidate Conference*, pages 343–348, 2000.

[DS02] Yevgeniy Dodis and Joel Spencer. On the (non)universality of the one-time pad. In *FOCS*, pages 376–388. IEEE, 2002.

[DS05] Yevgeniy Dodis and Adam Smith. Correcting errors without leaking partial information. In *STOC*, pages 654–663. ACM, 2005.

[DW09] Yevgeniy Dodis and Daniel Wichs. Non-malleable extractors and symmetric key cryptography from weak secrets. In *STOC*, pages 601–610, 2009. ePrint:2008/503.

[ECP$^+$05] Chip Elliott, Alexander Colvin, David Pearson, et al. Current status of the DARPA quantum network (invited paper). In *Quantum Information and Computation III*, volume 5815, pages 138–149, May 2005. arXiv:quant-ph/0503058.

[EGL85] Shimon Even, Oded Goldreich, and Abraham Lempel. A randomized protocol for signing contracts. *Commun. ACM*, 28(6):637–647, June 1985.

[Eke91] Artur K. Ekert. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.*, 67(6):661–663, August 1991.

[ElG85] Taher ElGamal. A public-key cryptosystem and a signature scheme based on discrete logarithms. *IEEE Tran. Inf. Th.*, 31(4):469–472, 1985.

[Fey82] Richard Feynman. Simulating physics with computers. *Int. J. of Th. Phys.*, 21:467–488, 1982.

[FFB$^+$11] F. Furrer, T. Franz, M. Berta, A. Leverrier, V. B. Scholz, M. Tomamichel, and R. F. Werner. Continuous variable quantum key distribution: Finite-key analysis of composable security against coherent attacks. arXiv:1112.2179, December 2011.

[FHS11] Omar Fawzi, Patrick Hayden, and Pranab Sen. From low-distortion norm embeddings to explicit uncertainty relations and efficient information locking. In *STOC*, pages 773–782. ACM, 2011. arXiv:1010.3007.

[FS09] Serge Fehr and Christian Schaffner. Composing quantum protocols in a classical environment. In *Theory of Cryptography Conference - TCC 09*, volume 5444 of *Lecture Notes in Computer Science*, pages 350–367. Springer, 2009. arXiv:0804.1059.

[Gol06] Oded Goldreich. On post-modern cryptography. ePrint:2006/461, 2006.

[GUV09] Venkatesan Guruswami, Christopher Umans, and Salil P. Vadhan. Unbalanced expanders and randomness extractors from Parvaresh–Vardy codes. *J. ACM*, 56(4), 2009.

[Hän10] Esther Hänggi. *Device-independent quantum key distribution*. PhD thesis, ETH Zürich (Switzerland), 2010. arXiv:1012.3878.

[HILL99] Johan Håstad, Russell Impagliazzo, Leonid A. Levin, and Michael Luby. A pseudorandom generator from any one-way function. *SIAM J. Comput.*, 28(4):1364–1396, 1999.

[Hir57] Jr. Hirschman, I. I. A note on entropy. *Am. J. Math.*, 79(1):152–156, 01 1957.

[HNDP02] R. J. Hughes, J. E. Nordholt, D. Derkacs, and C. G. Peterson. Practical free-space quantum key distribution over 10 km in daylight and at night. *New J. of Phys.*, 4:43, July 2002. arXiv:quant-ph/0206092.

[Hoe63] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[Hug04] Richard Hughes. A quantum information science and technology roadmap, part 2: Quantum cryptography (ARDA report). `http://qist.lanl.gov/qcrypt_map.shtml`, 2004.

[Hwa03] Won-Young Hwang. Quantum key distribution with high loss: Toward global secure communication. *Phys. Rev. Lett.*, 91:057901, Aug 2003.

[Kil88] Joe Kilian. Founding crytpography on oblivious transfer. In *STOC*, pages 20–31, New York, NY, USA, 1988. ACM.

[KMS11] Adrian Kent, William J. Munro, and Timothy P. Spiller. Quantum tagging: Authenticating location via quantum information and relativistic signaling constraints. *Phys. Rev. A*, 84, Jul 2011.

[Kob87] Neal Koblitz. Elliptic curve cryptosystems. *Math. Comp.*, 48:203–209, 1987.

[KR09] Bhavana Kanukurthi and Leonid Reyzin. Key agreement from close secrets over unsecured channels. In Antoine Joux, editor, *Eurocrypt*, volume 5479 of *LNCS*, pages 206–223. Springer, 2009. ePrint:2008/494.

[KRBM07] Robert König, Renato Renner, Andor Bariska, and Ueli Maurer. Small accessible quantum information does not imply security. *Phys. Rev. Lett.*, 98, Apr 2007.

[Kre78] Erwin Kreyszig. *Introductory Functional Analysis with Application*. Wiley, 1978.

[KRS09] Robert König, Renato Renner, and Christian Schaffner. The operational meaning of min- and max-entropy. *IEEE Tran. Inf. Th.*, 55(9):4337–4347, 2009. arXiv:0807.1338.

[Lan05] Serge Lang. *Algebra*. Graduate Texts in Mathematics. Springer, 2005.

[LC97] H.-K. Lo and H. F. Chau. Is quantum bit commitment really possible? *Phys. Rev. Lett.*, 78:3410–3413, April 1997. arXiv:quant-ph/9603004.

[LCA05] H.-K. Lo, H. F. Chau, and M. Ardehali. Efficient quantum key distribution scheme and a proof of its unconditional security. *J. Crypt.*, 18(2):133–165, 2005. arXiv:quant-ph/0011056.

[Lev09] Anthony Leverrier. *Theoretical study of continuous-variable quantum key distribution*. PhD thesis, Télécom ParisTech, 2009.

[Li12] Xin Li. Non-malleable extractors, two-source extractors and privacy amplification. In *FOCS*. IEEE, 2012.

[Lin92] T.A. Lindvall. *Lectures on the Coupling Method*. Dover Books on Mathematics Series. Dover, 1992.

[Lo97] Hoi-Kwong Lo. Insecurity of quantum secure computations. *Phys. Rev. A*, 56:1154–1162, Aug 1997. arXiv:quant-ph/9611031.

[LWW$^+$10] Lars Lydersen, Carlos Wiechers, Christoffer Wittmann, Dominique Elser, Johannes Skaar, and Vadim Makarov. Hacking commercial quantum cryptography systems by tailored bright illumination. *Nat. Photon.*, 4, October 2010. arXiv:1008.4593.

[Mau90] Ueli M. Maurer. A provably-secure strongly-randomized cipher. In *Eurocrypt*, LNCS, pages 361–373. Springer, 1990.

[May95] Dominic Mayers. On the security of the quantum oblivious transfer and key distribution protocols. In *CRYPTO*, pages 124–135, 1995.

[May96] Dominic Mayers. Quantum key distribution and string oblivious transfer in noisy channels. In *CRYPTO*, pages 343–357, 1996. arXiv:quant-ph/9606003.

[May97] Dominic Mayers. Unconditionally secure quantum bit commitment is impossible. *Phys. Rev. Lett.*, 78(17):3414–3417, April 1997. arXiv:quant-ph/9605044.

[May01] Dominic Mayers. Unconditional security in quantum cryptography. *J. ACM*, 48(3):351–406, 2001. arXiv:quant-ph/9802025.

[Mer78] Ralph C. Merkle. Secure communications over insecure channels. *Commun. ACM*, 21(4):294–299, April 1978.

[Mil86] Victor S. Miller. Use of elliptic curves in cryptography. In *CRYPTO*, volume 218 of *LNCS*, pages 417–426. Springer, 1986.

[Moc07] C. Mochon. Quantum weak coin flipping with arbitrarily small bias. arXiv:0711.4114, November 2007.

[MR09] Daniele Micciancio and Oded Regev. Lattice-based cryptography. In Daniel J. Bernstein, Johannes Buchmann, and Erik Dahmen, editors, *Post-Quantum Cryptography*, pages 147–191. Springer, 2009.

[MS94] Dominic Mayers and Louis Salvail. Quantum oblivious transfer is secure against all individual measurements. In *PhysComp*, pages 69–77. IEEE Computer Society, 1994.

[MU88] Hans Maassen and Jos B. M. Uffink. Generalized entropic uncertainty relations. *Phys. Rev. Lett.*, 60(12), 03 1988.

[MW97] Ueli Maurer and Stefan Wolf. Privacy amplification secure against active adversaries. In Burton S. Kaliski Jr., editor, *CRYPTO*, volume 1294 of *LNCS*, pages 307–321. Springer, August 1997.

[NC00] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge Univ. Press, first edition, 2000.

[NN93] Joseph Naor and Moni Naor. Small-bias probability spaces: Efficient constructions and applications. *SIAM J. Comput*, 22:838–856, 1993.

[Pau02] V.I. Paulsen. *Completely bounded maps and operator algebras*. Cambridge studies in advanced mathematics. Cambridge Univ. Press, 2002.

[PPM08] Andreas Poppe, Momtchil Peev, and Oliver Maurhart. Outline of the SECOQC quantum-key-distribution network in Vienna. *Int. J. Quant. Inf.*, 6:209–218, April 2008. arXiv:0804.0122.

[Rab81] Michael O. Rabin. How to exchange secrets by oblivious transfer. Technical Report TR-81, Aiken Computation Laboratory, Harvard University, 1981.

[RC09] R. Renner and J. I. Cirac. De Finetti representation theorem for infinite-dimensional quantum systems and applications to quantum cryptography. *Phys. Rev. Lett.*, 102:110504, Mar 2009. arXiv:0809.2243.

[Rén61] A. Rényi. On measures of entropy and information. In *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 547–561, 1961.

[Ren05] Renato Renner. *Security of Quantum Key Distribution.* PhD thesis, ETH Zürich (Switzerland), September 2005. arXiv:quant-ph/0512258.

[RK05] Renato Renner and Robert König. Universally composable privacy amplification against quantum adversaries. In *TCC*, volume 3378 of *LNCS*, pages 407–425. Springer, 2005. arXiv:quant-ph/0403133.

[RSA78] R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Commun. ACM*, 21(2):120–126, February 1978.

[RTS00] Jaikumar Radhakrishnan and Amnon Ta-Shma. Bounds for dispersers, extractors, and depth-two superconcentrators. *SIAM Journal on Discrete Mathematics*, 13, 2000.

[RW03] Renato Renner and Stefan Wolf. Unconditional authenticity and privacy from an arbitrarily weak secret. In Dan Boneh, editor, *CRYPTO*, volume 2729 of *LNCS*, pages 78–95. Springer, August 2003.

[RW04] Renato Renner and Stefan Wolf. The exact price for unconditionally secure asymmetric cryptography. In Christian Cachin and Jan Camenisch, editors, *Eurocrypt*, volume 3027 of *LNCS*, pages 109–125. Springer, 2004.

[RWY11] Leonid Reyzin, Drew Wolpert, and Sophia Yakoubov. Alternating extractors and leakage-resilient stream ciphers. 6.889 New Developments in Cryptography (lecture notes), `http://www.cs.bu.edu/~reyzin/teaching/s11cs937/notes-leo-2.pdf`, 2011.

[Sch07] Christian Schaffner. *Cryptography in the Bounded-Quantum-Storage Model.* PhD thesis, University of Aarhus (Denmark), September 2007. arXiv:0709.0289.

[Ser74] R. J. Serfling. Probability inequalities for the sum in sampling without replacement. *The Annals of Statistics*, 2(1):39–48, 1974.

[SFI$^+$11] M. Sasaki, M. Fujiwara, H. Ishizuka, et al. Field test of quantum key distribution in the Tokyo QKD network. *Optics Express*, 191:10387–+, May 2011. arXiv:1103.3566.

[Sha48] Claude E. Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27:379–423, July 1948.

[Sha49] Claude E. Shannon. Communication theory of secrecy systems. *The Bell system technical journal*, 28(4):656–715, 1949.

[Sha02] Ronen Shaltiel. Recent developments in explicit constructions of extractors. *Bulletin of the EATCS*, 77:67–95, 2002.

[Sho97] Peter W. Shor. Polynomial time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Sci. Statist. Comput.*, 26:1484, 1997. arXiv:quant-ph/9508027.

[Sho05] Victor Shoup. *A computational introduction to number theory and algebra*. Cambridge Univ. Press, New York, 2005.

[Sin99] Simon Singh. *The Code Book: The Evolution of Secrecy from Mary, Queen of Scots, to Quantum Cryptography*. Doubleday, 1st edition, 1999.

[SP00] Peter W. Shor and John Preskill. Simple proof of security of the BB84 quantum key distribution protocol. *Phys. Rev. Lett.*, 85:441–444, Jul 2000. arXiv:quant-ph/0003004.

[SS98] V. Scarani and A. Suarez. Introducing quantum mechanics: One-particle interferences. *American J. of Phys.*, 66:718–721, August 1998.

[SSS09] L. Salvail, C. Schaffner, and M. Sotakova. On the power of two-party quantum cryptography. In *ASIACRYPT*, volume 5912 of *LNCS*, pages 70–87. Springer, 2009. arXiv:0902.4036.

[Sti94] D. R. Stinson. Universal hashing and authentication codes. *Des. Codes Cryptography*, 4:369–380, October 1994.

[TSSR10] Marco Tomamichel, Christian Schaffner, Adam Smith, and Renato Renner. Leftover hashing against quantum side information. *IEEE Tran. Inf. Th.*, 2010. arXiv:1002.2436.

[Vad12] Salil P. Vadhan. Pseudorandomness. draft survey/monograph, 2012.

[VDG98] Jeroen Van De Graaf. *Towards a formal definition of security for quantum protocols*. PhD thesis, Univ. de Montreal (Quebec, Canada), 1998.

[Wie83] Stephen Wiesner. Conjugate coding. *SIGACT News*, 15:78–88, January 1983.

[WPGP$^+$12] Christian Weedbrook, Stefano Pirandola, García-Patrón, et al. Gaussian quantum information. *Rev. Mod. Phys.*, 84:621–669, May 2012. arXiv:1110.3234.

[WST08] Stephanie Wehner, Christian Schaffner, and Barbara M. Terhal. Cryptography from noisy storage. *Phys. Rev. Lett.*, 100, Jun 2008. arXiv:0711.2895.

[WW10] Severin Winkler and Jürg Wullschleger. On the efficiency of classical and quantum oblivious transfer reductions. In Tal Rabin, editor, *CRYPTO*, volume 6223 of *LNCS*, pages 707–723. Springer, 2010. ePrint:2009/508.

[Yao82] Andrew Chi-Chih Yao. Protocols for secure computations (extended abstract). In *FOCS*, pages 160–164. IEEE, 1982.

[Yao95] Andrew Chi-chih Yao. Security of quantum protocols against coherent measurements. In *STOC*, pages 67–75, 1995.

[Yue86] Horace P. Yuen. Amplification of quantum states and noiseless photon amplifiers. *Physics Letters A*, 113(8):405–407, 1986.

Index

1-2 OT, *see* one-out-of-two oblivious transfer
2PC (secure –), *see* secure two-party computation

accessible information, 35, 118
adjoint (of an operator), 64
Advanced Encryption Standard, 14
adversary, 16
AES, *see* Advanced Encryption Standard
almost, 62
alternating extraction, 148, 149
amplitudes, 72, 78
ancilla state, 75
authentication protocol, 137

BB84 protocol, 27
BC, *see* bit commitment
Bernoulli distribution, 46
bias, 46
binary entropy function, 52
binary random variable, 46
binomial distribution, 46
bipartite (system), 74
bit commitment, 98, 118
blinding attacks, 31
bounded-quantum-storage model, 32, 36
bounded-storage model, 39, 134
BQSM, *see* bounded-quantum-storage model
bra (vector), 63
braket notation, 63
Breidbart basis, 186

chain rule (for min-entropy), 86
chain rule (for Shannon entropy), 54
classical error probability, 103
coherent attacks, 29, 126
collective attacks, 29, 126
collision entropy, 52
collision probability, 53
complete projective measurement, 78
completely positive trace-preserving map, 75
completeness condition, 71
composite system, 71
computational basis, 73
concave, 49
conditional collision entropy, 55
conditional collision probability, 54
conditional distribution, 45
conditional entropy, 54
conditional guessing probability, 54
conditional independence (quantum version), 81
conditional min-entropy, 84
conditional min-entropy (of a random variable), 55
continuous-variable QKD, 30
convex, 49
coupling, 48
CPTP map, *see* completely positive trace-preserving map
cq-state, 79

decoy states, 31

density matrix, 70
density-matrix representation (of a distribution), 80
device-independent QKD, 31
DI-QKD, *see* device-independent QKD
diagonal basis, 73
distance to uniform, 83
distinguishing advantage, 47
distribution, 44, 45

encryption, 14
entanglement, 78
entanglement purification, 126
EPR pair, 78
events, 43
expectation, 49
expected value, 49
extractor-MAC, 93

finite probability space, 43
finite-key bound, 29
flat distribution, 46
free-space QKD, 30
fully mixed state, 72

guessing probability, 53
guessing probability (conditional on quantum side information), 84

Hadamard basis, 73
Hadamard matrix, 73
Hamming weight, 99
Hermitian operator, 64
Hilbert space, 63
hybrid system, 79

ideal functionality, 89
identity operator, 64
independent (quantum systems), 71
independent (statistically –), 45
individual attacks, 29
inner product, 62, 188
inner product (on the tensor-product space), 67
inner product space, 62

joint distribution, 44
joint system, 71

ket (vector), 63
Kronecker delta symbol, 42
Kronecker product, 69

locking, 36
look-ahead extractor, 138

MAC, *see* message authentication code
marginal distribution, 44
marginalization, 44
Markov chain, 45
matrix representation (of an operator), 68
max-entropy (of a density matrix), 85
max-entropy (of a random variable), 53
measuring in a basis, 78
message authentication code, 61
Millionaires' problem, 16
min-entropy (of a random variable), 52
min-entropy rate, 38
mixed (state), 72
mixture, 72
multiplexing channel, 34
mutually independent, 45

no cloning, 27, 76
no-cloning theorem, 76
normal operator, 64

oblivious transfer, 17, 34, 118
observable, 77
one-out-of-two oblivious transfer, 17
one-time pad, 61, 83
OT, *see* oblivious transfer
outcome (of a random variable), 44
outer product, 63

partial trace, 73
passively-quenched avalanche photodiode, 31
password-based identification, 89, 135
photon, 19
photon-number splitting attack, 31
polarizing beam splitter, 23
position-based cryptography, 33
position-based quantum cryptography, 34
positive-operator-valued measure, 77
positive-semidefinite operator, 64
POVM, *see* positive-operator-valued measure
privacy amplification, 55
product rule (for events), 43
product state, 78
projective measurements, 77
projector, 64
protocol, 89
pure (density matrix), 70
pure (state), 72
purification, 74

QID protocol, *see* quantum identification protocol
QKD, *see* quantum key distribution
QOT, *see* quantum oblivious transfer
quantum cryptography, 25, 34
quantum dynamics, 19
quantum error probability, 110
quantum identification protocol, 89, 167
quantum key distribution, 26, 79
quantum oblivious transfer, 97
quantum operation, 75
quantum population, 108
quantum protocol, 25
quantum theory, 19
qubit, 30, 73

Rényi entropy of order α, 52
Rabin oblivious transfer, 32
random variable, 44
randomized function, 15
randomized oblivious transfer, 119
range (of a random variable), 44
raw key, 28
realization (of a random variable), 44
reduced density matrix, 73
relative Hamming weight, 99
reverse reconciliation, 28
rewinding (argument), 141
RSA, 14

sample space, 43
sampling strategy, 96, 100
Schmidt decomposition, 74
Schur product, 188
secure two-party computation, 35
seed, 58
separable, 78
Shannon entropy, 51
single-qubit-operations model, 164
singlet state, 79
six-state QKD protocol, 35
spectral decomposition, 65
SQOM, *see* single-qubit-operations model
state space, 71
state vector, 72

statistical distance, 47
Stinespring's dilation theorem, 75
strictly convex, 49
strong extractor, 58
strong extractor against quantum side information, 88
strongly universal (family of functions), 58
superposition, 19, 72
support (of a distribution), 44
symmetric-key encryption, 14

tag (authentication –), 62
tensor product, 65
tensor-product space, 66
trace distance, 81
trace norm, 81

uncertainty principle, 23
uniform distribution, 45
union bound, 43
unitary operator, 64
universal (family of functions), 55
universal property (of a tensor product), 65

von Neumann entropy, 83

weak key, 134

Nederlandse samenvatting

In de kwantumcryptografie wordt onderzoek gedaan naar het benutten van kwantummechanische effecten voor cryptografische toepassingen, alsmede naar de veiligheid van bestaande en nieuwe cryptografische protocollen wanneer een kwaadaardige partij ("de vijand") berekeningen kan uitvoeren op een kwantumcomputer, en/of kwantuminformatie bezit over stochastische variabelen (bijv. cryptografische sleutels) die in deze protocollen een rol spelen.

Een bekende toepassing uit de kwantumcryptografie is sleuteldistributie (Engelse afkorting: QKD). Met behulp van dit protocol kunnen twee samenwerkende partijen via een onveilige kwantumverbinding (bijv. een optische vezel die gemanipuleerd kan worden door de vijand) en een geauthenticeerd klassiek communicatiekanaal op afstand een gezamenlijke en zeer veilige cryptografische sleutel genereren.

Hoewel de werking van het QKD-protocol op een intuïtief niveau vrij eenvoudig te begrijpen is, is het verre van triviaal om formeel te bewijzen dat de door het protocol geproduceerde sleutel daadwerkelijk veilig is. In dit proefschrift wordt een nieuwe bewijsmethode geïntroduceerd, die vervolgens succesvol wordt toegepast op het zgn. BB84-QKD protocol. Het resulterende bewijs is eenvoudiger dan de meeste bestaande QKD-veiligheidsbewijzen, en geeft een inzichtelijke, niet-asymptotische uitdrukking voor het bereikte veiligheidsniveau als functie van de protocolparameters. De nieuwe methode blijkt ook toepasbaar om de veiligheid te bewijzen van een kwantumprotocol voor het reduceren van *oblivious transfer* naar *bit commitment*. Het is goed mogelijk dat de bewijsmethode nog meer toepassingen heeft.

In een ander deel van dit proefschrift wordt de taak van berichtenauthenticatie in een nieuw scenario onderzocht. In dit scenario wordt aangenomen dat de vijand een beperkte hoeveelheid kwantuminformatie heeft over de cryptografische sleutel die gebruikt wordt voor authenticatie. Bovendien—en hiermee onderscheidt het scenario zich van eerder werk op dit gebied—wordt aangenomen dat de authenticatiesleutel in feite een sessie-sleutel is, die steeds opnieuw wordt afgeleid van een bron van *randomness* met behulp van een herbruikbare sleutel.

Het is onvermijdelijk dat een significante hoeveelheid informatie over de sessiesleutel openbaar wordt tijdens het uitvoeren van het authenticatieprotocol (en dus tevens in handen komt van de vijand). Het doel van dit onderzoek is om een authenticatiemethode te vinden die voorkomt dat er ook een significante hoeveelheid informatie over de herbruikbare sleutel vrijkomt. In dit proefschrift wordt een oplossing gepresenteerd voor het geval waarin de vijand louter klassieke informatie over de sessie-sleutel bezit. Voor het generieke geval waarin de vijand ook kwantuminformatie bezit is het voorgestelde protocol niet compleet.

Entropische onzekerheidrelaties zijn formele uitdrukkingen van het onzekerheidsprincipe van Heisenberg, die gebruik maken van een entropiemaat om de onzekerheid te kwantificeren. In dit proefschrift wordt een nieuwe entropische onzekerheidsrelatie gepresenteerd en bewezen. Het is de eerste onzekerheidsrelatie die een ondergrens geeft voor de min-entropie in het meetresultaat,[1] waarbij deze ondergrens geldt voor *op één na alle* metingen, gekozen uit een willekeurige (en willekeurig grote) familie van mogelijke metingen. Het gebruik van de min-entropie als onzekerheidsmaat maakt de onzekerheidsrelatie bijzonder geschikt voor gebruik in de kwantumcryptografie.

Als toepassing wordt een nieuw kwantum-identificatieprotocol gepresenteerd in het *bounded-quantum-storage* model; de nieuwe onzekerheidsrelatie vormt de kern van het formele veiligheidsbewijs voor dit protocol. In tegenstelling tot het oorspronkelijke kwantum-identificatieprotocol van Damgård *et al.* biedt het nieuwe identificatieprotocol ook enige mate van bescherming in het geval dat de *bounded-quantum-storage*-aanname niet geldt. Het protocol is nl. bestand tegen een vijand die een ongelimiteerde kwantum-opslagcapaciteit heeft, maar enkel (niet-adaptieve) operaties en metingen op afzonderlijke qubits kan uitvoeren. Het protocol van Damgård *et al.* biedt geen enkele bescherming tegen een dergelijke vijand.

[1] Verkregen door het meten van een willekeurige kwantumtoestand van een bepaalde dimensie.

Curriculum Vitae

Niek Johannes Bouman was born on September 2, 1983 in Son en Breugel, the Netherlands. From 1995 to 2001, he followed his pre-university education at the Lorentz Casimir Lyceum in Eindhoven, the Netherlands. He received his bachelor's degree (2005) and master's degree (2007, *cum laude*) in electrical engineering from Universiteit Twente in Enschede, the Netherlands. He wrote his bachelor's thesis about a problem in fingerprint recognition, under supervision of dr. Raymond Veldhuis. During his master's program he did an internship at IBM Research in Rüschlikon, Switzerland, and worked there on tape storage, under supervision of dr. Thomas Mittelholzer. He wrote his master's thesis about modulation and coding for linear Gaussian channels, under supervision of dr. Harm Cronie. From 2008 to 2012, he was a PhD student in prof. Ronald Cramer's cryptology group at CWI, under supervision of dr. Serge Fehr.