Université de Montréal

Towards a formal definition of security for quantum protocols

par

Jeroen van de Graaf

Département d'Informatique et de Recherche Opérationnelle

Faculté des Arts et des Sciences

Thèse présentée à la Faculté des études supérieures

en vue de l'obtention du grade de

Philosophiæ Doctor (Ph.D.)

en Informatique

Décembre 1997

Université de Montréal

Faculté des études supérieures

Cette thèse intitulée

Towards a formal definition of security for quantum protocols

présenté par:

Jeroen van de Graaf

a été évalué par un jury composé des personnes suivantes:

Michel Boyer
(président rapporteur)

Claude Crépeau
(directeur de recherche)

Gilles Brassard
(co-directeur)

Geňa Hahn
(membre du jury)

Joe Kilian
(examinateur externe)

David London
(représentant du doyen)

Thèse acceptée le: 29 mai 1998.

Summary

This thesis presents a formal model for quantum protocols, that is, cryptographic protocols based on quantum mechanics. It also develops a definition of security (which includes notions such as privacy and correctness) for quantum protocols.

First, existing classical (non-quantum) definitions of security are studied. There are essentially two approaches: one based on Turing Machines and simulation, the other using information theory. Both approaches are explored in detail and compared, then a particular variation of the first approach is chosen. Security is defined by introducing an interface (also called a simulator) which is capable of translating an attack on protocol 1 into an attack on protocol 2. The profile of a protocol is a stochastic variable, defined as the output of the participants and the view of the adversary who performs the attack. If an interface exists such that the profile of protocol 1 is indistinguishable from the profile of protocol 2 with the help of this interface, then protocol 1 is at least as secure as protocol 2. An absolute notion of security is obtained by comparing the protocol implementation with its 'ideal' equivalent.

In this approach, and in classical cryptography in general, indistinguishability of probability distributions is a crucial tool. The meaning of this notion is explored in the context of density matrices, the mathematical formalism which is used to describe the state of a quantum system. Various measures of distinguishability of quantum states relevant from a cryptographic point of view are investigated: probability of error, trace norm (related to the Kolmogorov distance), overlap (related to the Bhattacharyya coefficient) and Shannon distinguishability (defined through the mutual information). It is shown that if we require that the distinguishability between two families of quantum states vanish exponentially fast, these four distinguishability measures are equivalent.

Along the way upper and lower bounds for the Shannon distinguishability are obtained. This is a very useful result, because explicit computation of the Shannon distinguishability between two quantum states is often difficult (it requires solving a transcendental equation), whereas the other distinguishability measures, in particular the trace norm, are usually much easier to compute.

In the last chapter the two previous topics are combined. A formal model for quantum protocols is formulated. It uses Quantum Turing Machines to represent the participants, who communicate through a quantum channel. The different nature of quantum information as opposed to classical information is explored in detail, including the implications this has for the definition. As in the classical case, we introduce an interface to compare the security of two protocols for the same task. Subsequently some consequences of the proposed definition are investigated. In particular, strong evidence is given that one classical technique often used in proofs of security for classical protocols, called rewinding, is impossible in the quantum setting.

Key words: cryptographic protocols, zero knowledge protocols, quantum information theory, quantum cryptography.

Résumé

La question générale posée dans cette thèse est la suivante: qu'est-ce qu'un bon modèle formel pour des protocoles cryptographiques basés sur la mécanique quantique, et qu'est-ce qu'une définition raisonnable de la sécurité pour de tels protocoles quantiques?

On s'intéresse d'abord à cette question pour des protocoles classiques, en particulier pour des protocoles de calcul privé multi-partie. Pour un tel protocole, n participants veulent calculer ensemble une fonction $f(x_1, x_2, ..., x_n)$, où x_i est une entrée pour chacun d'eux. Personne ne veut révéler d'information à propos de son entrée, sauf ce qui peut être logiquement déduit de la sortie. Par exemple, un vote de majorité effectué sur Internet lorsqu'aucun participant ne jouit de la confiance de tous les autres. Trouver une définition pour des protocoles de calcul privé multi-partie est étonnamment difficile, même pour des protocoles à deux parties relativement simples, comme le transfert inconscient, ou l'évaluation d'une porte ET. Nous n'envisagerons que les protocoles à deux participants.

Les modèles existants peuvent être divisés en deux groupes. La première approche est basée sur les machines de Turing, qui sont étendues à partir d'un ruban dédié à l'envoi et à la réception des messages. Ce ruban partagé sert à modéliser la communication entre les participants d'un protocole. La sécurité est définie par l'introduction d'une interface (aussi appelé simulateur) qui est capable de traduire une attaque sur le protocole réel en une attaque sur une version idéalisée du même protocole. On définit une variable stochastique, le profil d'un protocole, comme les sorties des

participants et la vue de l'adversaire qui effectue l'attaque. Si une interface existe telle que le profil du protocole réel est indistingable, comparé à celui du protocole idéalisé au travers de cette interface, alors le protocole est dit sécuritaire.

Une deuxième approche utilise aussi les machines de Turing, mais sa définition de la sécurité est basée sur la théorie de l'information. La définition proposée traite séparément la confidentialité (aucune information erronée à propos de l'entrée d'une autre partie est obtenue) et la correction (la sortie de la fonction ne peut être influencée par une partie malhonnête). Bien que valable dans quelques cas spéciaux, il existe des exemples dans lesquels cette séparation est fatale, dans le sens où des protocoles non-sécuritaires satisfont ces deux aspects. Nous modifions cette approche pour résoudre ce problème et montrons qu'elle contient la première approche comme cas particulier. Néanmoins, nous développons des critères pour comparer les deux approches et choisissons la première dans la suite de la thèse.

Dans la conception des protocoles cryptographiques on définit souvent une famille de protocoles, indexée par un paramètre de sécurité. Ce dernier peut indiquer par exemple la taille d'une chaîne, le nombre d'étapes dans un protocole, ou le nombre de photons. Une notion centrale en cryptographie classique est l'indistingabilité exponentielle de deux distributions de probabilité. Cette notion exprime que la distance entre deux distributions converge vers zéro de manière exponentiellement rapide quand le paramètre de sécurité augmente; cette distance est habituellement prise au sens de Kolmogorov.

Donc, si on veut modéliser et analyser la sécurité des protocoles quantiques on a besoin d'une notion de distance entre deux états quantiques. L'état d'un système quantique est décrit par une matrice de densité, mais plusieurs notions de distance sont possibles. Nous étudions quatre mesures de distingabilité d'états quantiques pertinentes pour la cryptographie: probabilité d'erreur, norme de la trace (reliée à la distance de Kolmogorov), "overlap" (relié au coefficient de Bhattacharyya), et distingabilité de Shannon (définie à travers l'information mutuelle).

Ces mesures ont leur origine dans les mesures classiques d'indistingabilité des distributions de probabilité, à partir desquelles un ensemble d'inégalités peut être obtenu. Nous démontrons que ces inégalités sont aussi valables dans le modèle quantique, et obtenons des bornes supérieures et inférieures pour la distingabilité de Shannon. Ceci est important, car un calcul explicite de la distingabilité de Shannon est souvent difficile (il faut résoudre une équation transcendantale), alors que les autres mesures de distingabilité sont plus faciles à évaluer. En particulier, nous dérivons une expression simple d'une borne supérieure de la distingabilité de Shannon en fonction de la norme de la trace.

Comme dans le cas quantique, l'indistingabilité exponentielle de deux familles d'états quantiques peut être définie : on demande que la distingabilité entre deux familles s'annule exponentiellement vite. Il résulte que l'indistingabilité exponentielle est indépendante de la mesure spécifique choisie dans la définition: les quatre différentes définitions apparaissent donc équivalentes.

Dans le dernier chapitre nous présentons notre modèle pour des protocoles quantiques à deux parties, et donnons une définition de la sécurité. Nous nous restreignons aux protocoles avec des entrées et sorties classiques uniquement. Nous expliquons en détail comment un ordinateur quantique peut émuler les observations d'un état quantique, le fait que l'information quantique ne peut être copiée, et décrivons le principe qui nous empêche d'extraire une quelconque information d'un système quantique sans perturber l'état de celui-ci. Mais malgré le fait que l'information classique peut être considérée comme un cas particulier de l'information quantique, nous montrons que pour une bonne définition de la sécurité, les deux types d'information doivent être considérées séparément.

Après une discussion informelle, nous présentons notre modèle de façon formelle. Nous utilisons les machines de Turing quantiques pour les parties A et B, qui communiquent à travers un canal quantique modélisé par un ruban de communication quantique. Comme dans l'approche classique, nous utilisons une interface pour définir la sécurité d'un protocole relativement à un autre: on exige que le profil du premier protocole soit exponentiellement indistingable comparé au profil obtenu par un adversaire qui attaque le deuxième protocole à travers cette interface. Une notion absolue de sécurité est obtenue en comparant un protocole réel à son équivalent idéalisé. Finalement nous présentons des arguments convaincants sur l'impossibilité de la technique classique de retour en arrière ("rewinding") d'une machine de Turing dans notre modèle quantique.

Contents

List of Tables

List of Figures

List of Acronyms

BC	Bit Commitment	see page 20.
BCJL	Reference [BCJL93]	
BMS	Reference [BMS96]	
EPR	Reference [EPR35]	
GOT	Generalized Oblivious Transfer	see page 43.
MPC	Multi-Party Computation	see page 2.
OT	Oblivious Transfer	usually $\binom{2}{1}$-OT is meant.
$\binom{2}{1}$-OT	One-out-of-Two Oblivious Transfer	see page 20.
POK	Proof of Knowledge	see page 21.
POVM	Positive Operator-Valued Measure	see page 57.
PVM	Projection Valued Measurement	see page 57.
QTM	Quantum Turing Machine	see page 103.
TM	Turing Machine	see page 103.
TPC	Two-Party Computation	see page 20.

Chapter 1

Introduction

Mathematics belongs to man, not to God. [...]
If God has mathematics of his own that needs to be done,
let him do it himself.
ERRETT BISHOP

1.1 Classical cryptography

Suppose the board of directors of IACR (the International Association of Cryptographic Research) suddenly wants to hold a confidential vote on some sensitive issue (something related to its former editor-in-chief, say). Since its members work and live all around the world, one solution could be to use a trusted party, to whom everybody sends his vote, who counts the tallies and broadcasts them.

However, in cryptography there is no such thing as mutual trust: if all the parties involved in a protocol trust each other there would be no need for cryptography, and we could all go to the beach. Another presupposition is that the parties in a protocol are spatially separated, but can communicate through some channel. The challenge of cryptography is to design protocols which reduce trust to the minimum. In this ballot example, the question is whether the same functionality can be obtained without a trusted party. That is, whether the members of the board are able to hold their ballot while they are all connected through the internet, say, running some very clever Java-application, without relying on a trusted party.

This question was addressed, and answered affirmatively (at least theoretically) in the years 1985, '86, '87; it was then among the most exciting topics in cryptographic research. Not only secret voting was studied, but also other cryptographic tasks: how to play poker (or any other card game), how to resolve the millionaires' problem (two millionaires want to find out who is richest without revealing their wealth), or the match-making problem (boy and girl want to find whether they would be interested in dating each other, but if one party is not interested, he or she cannot find out what the other wanted).

All these tasks can be regarded as instances of a more general cryptographic task known as Multi-Party Computation (MPC): n participants want to compute together a function $y = f(x_1, x_2, ..., x_n)$ with the x_i being their inputs. Nobody wants to reveal information about his input except what can be logically deduced from the output, y. Observe that the match-making problem is just a disguised form of an evaluation of an AND gate.

If the participants of a cryptographic protocol for MPC do not want to put their trust in one person (or organization), they must agree on (and put their trust in) some other assumption. For most cryptographic protocols, this falls in one of three categories:

Participants are computational restricted: it is assumed that all the participants have only restricted computational resources and that computationally hard problems exist. For instance, they can assume that factoring large integers is hard.

The majority of participants is honest: If we assume that a reliable broadcast channel exists, that each pair of participants has a private communication channel and that more than $\frac{2}{3}n$ (later improved to $\frac{1}{2}n$) participants are honest, then there exist protocols for multi-party computations [CCD88, BGW88, RB89, BE89, BE91B].

- Participants are connected by error-prone channels: Usually it is assumed that when two participants communicate, the transmission of bits is errorless: a bit sent by one party equals the bit received by the other. (In practice error-correcting codes are applied to guarantee this property.) However, we can abandon this assumption and use this error-property (the fact that a communication channel may lose or invert its bits) in favor of cryptography. For instance, in the erasure channel (also called Rabin Oblivious Transfer [RA81]) a positive fraction of the bits transmitted by the sender disappear; the receiver sees a "?" instead of 0 or 1 some of the time. Another example is the Noisy Channel, in which a bit gets inverted during transmission with a certain probability. Both Oblivious Transfer and Noisy Channel are strong enough primitives to construct a protocol for MPC. Until October 1995 the Quantum Channel was believed to fall in this same category, but this turned out to be false. More on this subject later.

Since 1985 several protocols for MPC have appeared in the literature, for an overview with references, see [GR95] or [FR93]. But even though many protocols have been presented in these papers, hardly any of these papers contains a rigorous proof of their security. Often these papers present a protocol, and with some handwaving and a sketchy proof the protocols are claimed to be secure.

But this situation is unsatisfactory. History has shown us too many examples where encryption schemes were claimed to be secure, until a cleverer person came along and broke it. It was only during the Second World War that many mathematicians got involved in breaking the enemy's cryptosystem, and started to wonder about a mathematical basis for designing cryptosystems. By now, any encryption scheme or protocol should (ideally) be accompanied by a mathematical proof of its security. Similarly, we would like to accompany each proposed protocol for MPC with a proof of its security.

However, before one can prove something, a definition is needed, and this is where the problem lies. Above, in the context of Multi-Party Computation, we used the phrasing that "no party can learn anything about the other inputs, except from what can be logically deduced from its own input and output". Although intuitively very clear as such, this phrasing is not sufficiently precise. Finding a suitable and pleasant definition for the security of MPC is surprisingly difficult, even for relatively simple primitives as the match-making protocol (i.e. the AND) or Oblivious Transfer.

This is where this thesis begins. In Chapter 2 we will see why defining security of MPC is difficult. Then we discuss and compare in more detail the approaches taken by Beaver, Micali-Rogaway, Canetti et al, and Brassard-Crépeau. The first three have many features in common, and some of the differences will be of no concern for this thesis. At the end of the first chapter we adopt one approach, after having discussed the criteria for making this choice.

1.2 Quantum cryptography

The world is quantum!

Even though it is not apparent from our daily perceptions, the laws that describe the phenomena (including particles and radiation) of which our universe is constructed, are the laws of quantum mechanics.

Computation and communication are inherently *physical* processes. The Church-Turing thesis is a physical assumption about what we can compute. Turing's formalization of a computer was based on what a clerk following simple instruction could compute. (Even though Turing knew enough about quantum mechanics, it apparently did not occur to him that it could be used for computation.) And Cover and Thomas say the following about communication [CT91]:

> What do we mean when we say that A communicates with B? We mean that the physical acts of A have induced a desired physical state in B. This transfer of information is a physical process and therefore is subject to the uncontrollable ambient noise and imperfections of the physical signaling process itself. The communication is successful if the receiver B and the transmitter A agree on what was sent.

Since cryptography is founded on computation and communication, which are both physical, and since physics is essentially quantum, it becomes natural to base cryptography on quantum physics. To put it more precisely: quantum cryptography is the endeavor to build secure cryptographic protocols whose security is based on quantum mechanics.

What does this mean? In the preceding subsection we remarked that any encryption scheme or protocol should (ideally) be accompanied by a mathematical proof of its security. For instance, for a scheme based on the assumption that factoring integers is hard, one wants to prove the following theorem: "if someone can break this protocol, it must imply s/he has found an efficient algorithm for factoring"; Since the latter is considered unlikely, at least using a classical algorithm, the protocol is considered secure.

In quantum cryptography we want to design protocols together with proofs of the following statement: "if someone can break this protocol, it must imply s/he has found a contradiction with the laws of quantum mechanics as we know them." This feat would be even more surprising than finding an efficient way to factor; indeed, it would be a scientific revolution.

Quantum cryptography has a short but turbulent history. Apart from some early work by Wiesner, it started to evolve when Bennett and Brassard [BB84] defined a coding scheme (that is, they associated specific quantum states with the bits 0 and 1), usually called BB84, and invented a protocol to do key distribution. Then, after they built a prototype and published these results in the Journal of Cryptology and the Scientific American [BBBSS92, BBE92] quantum cryptography attracted a lot of attention. It began to stimulate, and be stimulated by, its two neighbors: quantum information theory and quantum computing.

Though no person has serious doubts about the security of the quantum key exchange protocol, getting the formal proof right turns out to be fairly complicated. In the initial analysis of the Quantum Key Distribution protocol, in which Alice sends quantum particles to Bob, it was said that eavesdropping was impossible because of Heisenberg's Uncertainty Principle. Apart from the fact that some physicist dislike the words "uncertainty principle" (and for good reasons), the early proof considered only intercept/resend attacks, in which Eve does a complete (von Neumann) measurement on the particle and replaces it by a new one.

However, instead of fully destroying the state of a quantum system, nature allows much more subtle kinds of measurements: Positive Operator-Valued Measures, in which Eve disturbs the state just a bit (which decreases her chance of being detected) and in which she only gets partial information about the state of the particle sent from Alice to Bob. So qualitatively Eve is faced with the following choice: a high disturbance, which means she obtains much information and has a high chance of being caught, or low disturbance, which means little information but low chance of being caught, and all the gradations in between. Information versus disturbance is the physical principle on which Quantum Key Distribution is based, but it plays a role in other quantum protocols as well, as we will see.

Of course, people started to look for applications of quantum cryptography other than Key Distribution. Classical cryptography offers a wide range of tasks, so it is plausible to search for quantum variants for each of them. This seemed to work well: protocols were proposed for Bit Commitment (BC) but finding a protocol with a proof for Oblivious Transfer seemed rather difficult, and finding a quantum signature scheme almost impossible (signatures seem to make an inherent use of one-way functions).

Functionally (i.e. as a task), Bit Commitment is not different from writing a 0 or 1 on a piece of paper, putting the paper in an envelope, sealing the envelop and putting it on the table. Then, in a later stage, A says: "I was committed to 0 (1)," and she proves this by opening the envelope and showing its contents to B. Though this may not be obvious, BC is an extremely powerful primitive: by combining large quantities of BCs one can construct powerful and beautiful protocols [BCC88].

The sequence of events related to Quantum Bit Commitment was dramatic:

(a) In 1993 Brassard, Crépeau, Jozsa and Langlois proposed a quantum protocol for Quantum Bit Commitment (QBC) including a proof of its security [BCJL93].

(b) Attempts by Mayers and Salvail to find a quantum protocol for Oblivious Transfer without additional assumptions never succeeded.

(c) In September 1995, Yao announces he has found a protocol for Oblivious Transfer assuming error-less quantum transmission.

(d) In October 1995, Mayers decided to look critically at the proof of BCJL and discovered the wrong statement was proven (see Section 4.4.6 for details). In addition, he realized that not only the proof was wrong, but that the protocol could not work. Independently, Lo and Chau had come to the same conclusion.

(e) During that winter alternative protocols for quantum BC were proposed, which were subsequently broken.

(f) In March 1996 Mayers presented convincing arguments that no protocol for QBC can exist.

Let us look a little closer at the last point. Initially, the hope was that cryptographic protocols could be developed whose security would be based solely on the laws of quantum mechanics. This turns out to be impossible for BC, thus casting doubts on the existence of quantum protocols for other tasks. For instance, given a protocol for Oblivious Transfer we can make a protocol for BC, so it follows that no protocol for quantum Oblivious Transfer can exist. It must be realized that this is a purely theoretical result. The security requirements applied are very strict; in other words, theoretically dishonest parties are supposed to have all kind of resources and master technological skills that in reality may not (and may never) exist. Nevertheless, the result is disappointing, in the sense that the idea of having quantum analogues for all existing classical protocols, without any computational assumption, turned out to be futile.

Clearly, Mayers's arguments led to much discussion. Part of the problem was that no one ever defined a proper model for what a quantum protocol exactly is, making the details of the proof somewhat hard to verify. Even today there is in my opinion no paper that clearly states the model with all its underlying assumptions properly explained, followed by a simple stepwise proof why QBC is impossible. The fact that recently two people have tried to advocate that QBC *is* possible [Bo97, Ke97] seems to confirm this point.

The last chapter of this thesis develops a formal model for quantum protocols. It also presents a definition of security. However, from a quantum model to a definition of security there is a big step, and to ease the transition we have to deal with another subject first.

1.3 Classical and quantum indistinguishability

In classical cryptography, many definitions of protocol security use the notion of indistinguishability. In the design of cryptographic protocols, we often actually define a *family* of protocols parameterized by a *security parameter* n (for instance denoting the length of some string, the number of rounds, or the number of photons). Typically, we require that the probability of cheating for each participant vanishes exponentially fast, i.e. is of the order $O(\varepsilon^n)$, where ε is between 0 and 1.

A common technique to define protocol security is to compare the protocol implementation (the family of protocols) with the *ideal protocol specification* and to require that these two become exponentially indistinguishable. We want to use this same technique, but before being able to do so, we must know how this indistinguishability notion generalizes to quantum theory. This is what Chapter 3 is about.

Classically, the objects we want to compare are probability distributions; for which several distance measures are known. Quantumly, the objects we deal with are density matrices. A density matrix is the complete description of what we know about the state of a quantum system (observe the "we know": a density matrix is actually subjective, an issue we have to deal with). Mathematically, it is a matrix to which we apply another matrix, describing the measurement applied to the quantum system. The resulting mathematical object is a probability distribution, associating a probability to each possible measurement outcome.

When studying the literature one finds that there are actually various (more or less) natural measures of distance between quantum states, i.e. between density matrices. Fortunately, when we study these notions in the context of families of quantum states and require that the distance between pairs vanish exponentially fast, it turns out that they are equivalent. In other words, even though there are various notions of distance between quantum states, there is only one obvious notion of exponential quantum indistinguishability.

Somewhat coincidentally, the work in Chapter 3 has also found another application. Along the way inequalities on quantum distinguishability measures were obtained, which have not appeared in the literature before. Moreover, the distinguishability notion based on the Shannon information is notoriously difficult to compute, but we show that it can be bounded by the trace norm (the equivalent of the Kolmogorov distance or variational distance), which is much easier to handle. This bound plays an important role in an improved proof of the security of Quantum Key Distribution [BBBGM98]. This subject, however, is not discussed in this thesis, though QKD will be used as an example to illustrate some points in Chapter 4.

1.4 Security for quantum protocols

The main objective of this thesis is to present a formal model for quantum protocols, and to develop a definition of security (which includes privacy and correctness) for them. Most of the proofs for quantum protocols were somewhat ad hoc; having a more general framework for security is desirable.

In the classical case, many protocols are proven secure as follows: suppose that A (Alice) is convinced that B (Bob) is always able to generate (simulate) by himself, in probabilistic polynomial time, all the messages that he sees during the execution of the protocol, without access to A's secret input. Then this protocol cannot leak any knowledge (= polynomially computable information) to B, neither about the secret nor about anything else. Such a protocol is called Zero Knowledge [GMR85, GMR89]. The question is how this notion translates in the quantum setting.

Before we can answer this question, we first need a proper model for quantum protocols. We examine in detail the differences between classical and quantum information, and the consequences this has for adapting the classical definition of security to the quantum setting. We will see that a quantum computer can emulate measurements and (classical) probabilistic computations. But despite the fact that classical computation can be seen as a special case of quantum computation, we argue that for a proper definition of security the two types of information must be kept separate. This is related to the fact that generally quantum information cannot be copied, whereas classical information can.

As a model of computation we use Quantum Turing Machines for A and B, and describe their quantum channel by means of a quantum communication tape. We define the output, view and profile of a quantum protocol, and use a quantum interface (simulator) to compare the security of two protocols. It is at this point that we need the results of Chapter 3 about exponential indistinguishability of quantum states. After the formal definition of security, we present strong evidence that in our model the classical technique of rewinding is impossible.

1.5 Structure

The structure of the thesis is extremely simple. Chapter 2 describes definitions of security for classical protocols, which make inherent use of the notion of exponential indistinguishability for probability distributions. In Chapter 3 this notion is extended to exponential indistinguishability for density matrices. Chapter 4 follows the definition proposed in Chapter 2, extended to the quantum scenario. Consequently, this last chapter often refers to Chapter 2, whereas the first two main chapters are almost independent of each other.

The subject of this thesis lies on the boundary between cryptography and physics. I have approached the subject from a cryptographer's point of view, trying to provide ample explanation concerning the underlying physical principles. Although many notions are explained in the text, familiarity with the following topics is helpful: state vector, Hilbert space, pure state, mixed state, superposition, entanglement, EPR-pairs, von Neumann measurement, bra and ket notation, density matrix, tensor product, partial tracing (also tracing out). A very brief introduction to most of these notions can be found in the appendix of [BCJL93]; otherwise [Is95] is highly recommended. Other helpful references are [Su86, Pe93, Hu89].

As a final remark: this thesis extends various existing notions to a quantum setting, and the adjective 'classical' should be interpreted as 'non-quantum'.

Chapter 2

Classical definitions of protocol security

Reality resists imitation through a model.
ERWIN SCHRÖDINGER

In a protocol for Multi-Party Computation n participants want to compute together a function $f(x_1, x_2, ..., x_n)$ with the x_i being their respective inputs. Nobody wants to reveal information about his input except what can be logically deduced from the output. An example is, for instance, a majority vote held on the Internet, without a trusted party. Finding a definition of security for Multi-Party Computation is surprisingly difficult, even for relatively simple two-party protocols as $\binom{2}{1}$-OT, or evaluating an AND *gate.*

In this chapter we explore what exactly these difficulties are. We present a definition, which is a minor variation on Beaver's approach, and compare it with the works of Micali and Rogaway and Canetti, Feige, Goldreich and Naor, which have similarities in underlying philosophy.

We also compare with the approach of Brassard and Crépeau, whose definition is *based on information theory. In particular, we show that a protocol which is secure according to the newly proposed definition, is also secure under their definition (but not the other way around).*

The primary purpose of this chapter is not *to present a complete formal model of security for protocols for Multi-Party Computation, but to survey the current literature on the subject. The goal of this thesis is to study definitions of security for quantum protocols, and this chapter serves to prepare the stage for this definition in Chapter 4, and to justify the choices made there.*

2.1 Introduction

2.1.1 Brief historical overview

Traditionally, cryptology can be regarded as the sum of cryptanalysis (the breaking of cryptosystems) and cryptography, the art of secret writing. In the last 50 years, and especially in the last few decades, we have seen an effort to change cryptography from an art to a science: a proposal for a new cryptosystem or protocol should nowadays be accompanied by a mathematical proof of its security in order to be accepted in the scientific community. In practice, however, it is hardly true that every protocol in use today has been certified with such mathematical rigor, and (hopefully) sound engineering principles were applied.

It was before and during the Second World War that mathematicians in England and the United States became involved in information theory and cryptology. The first formal proof of security of a cryptosystem was published by Shannon [SH48, SH49]. In the first paper he defines the notion of (mutual) information, in the second paper he uses it to show that the one-time pad is perfectly secure.

At first glance information theory might seem the ultimate tool to defining security, but this is not always true. A major problem with the information-theoretic approach is that it cannot provide a meaningful definition of security in systems whose security rely on the hypothesis that some problems are computationally intractable. For instance, many pseudo-random number generators, public-key cryptosystems and protocols rely on the assumption that it is practically impossible to factor an integer N that is the product of two large (say 200 digits) prime numbers P and Q. When one party announces his integer N, then from an information-theoretic point of view he also announces the factors P and Q, and that's the end of the security discussion. To put it differently, information theory is only concerned with mathematical existence, not with computability, and in this sense an eavesdropper is considered to have unlimited computational resources. Although this thesis is not primarily concerned with computational cryptography, these issues have influenced the field enormously.

For example, a tremendous contribution to dealing with the issue of defining computational security was given by Goldwasser, Micali and Rackoff [GMR85, GMR89], who defined the notion of zero knowledge. The intuition is as follows: suppose that A (Alice) is convinced that B (Bob) is always able to generate (simulate) by himself, in probabilistic polynomial time, all the messages that he sees during the execution of the protocol, without access to A's secret input. Then this protocol cannot leak any knowledge (= polynomially computable information) to B, neither about the secret nor about anything else.

Their definition gave enormous impetus to the field. It was the first example of a definition of security that applied to a *class* of protocols, not just one specific protocol in particular. Apart from more technical (but nevertheless crucial) issues as rewinding/resetting the simulator and auxiliary input (which we discuss in more detail later on), it also led to a long discussion about what a definition of security should exactly include. This, however, also depends strongly on the context in which the protocol is used.

For instance, if a protocol for Proofs of Knowledge is used, in which A convinces another party B that she knows a witness of some statement in *NP*, then any amount of knowledge leaked about the witness is too much. But on the other hand, if the protocol is used as a signature scheme, leaking some bits of information is not necessarily disastrous, as long as the large majority of the bits cannot be forged. To deal with these issues the notions of witness hiding and witness indistinguishability were defined [FS90], but they are completely inadequate for our purposes.

Soon after the introduction of zero knowledge, people started to study the question whether Multi-Party Computation could be done. Suppose m participants want to compute together a function $f(x_1, x_2, ..., x_m)$ with the x_i being their respective inputs. Let their respective outputs be defined as f_i. Nobody wants to reveal more information than is strictly necessary. That is, nobody party wants to reveal any information about his or her input except what can logically be deduced from one's input and the output. The question was whether protocols exist that implement this task. This was an important question, because many protocols, including the millionaires problem [YA82A], digital voting [CH81], coin flipping [BL82], oblivious transfer [RA81] and mental poker [SRA81] are just special cases of this.

We now know that such protocols indeed exist; in fact there are various ways to implement MPC based on different assumptions. For a short overview see [GR95], but for a more thorough comparison, Chapter 2 of Franklin's thesis [FR93] is highly recommended. One common feature of these protocols is that most of them employ some sort of Bit Commitment or Secret Sharing schemes that each participant must use in order to provide his input, thus implicitly defining what the input of each participant is.

Another common feature is that many (if not all) approaches use the circuit representation of the function f, and show how individual gates (e.g. NOT and AND) can be implemented. The question: "Which are the minimum cryptographic primitives needed to implement Two-Party Computation?" was answered by Kilian [KI88]. He showed that if a protocol for Oblivious Transfer between each pair of participants exists, then they can evaluate any function f. The reduction from Multi-Party Computation to Oblivious Transfer was done by Beaver and Goldwasser [BG89].

A more efficient reduction from MPC to OT using coding theory can be found in [CGT95].

Nevertheless, despite the abundance of papers on the subject of Multi-Party Computation, and even though the security of all the proposed protocols is commonly believed to be true, most of them lack a rigorous mathematical proof. Even worse, they lack a proper definition. This is probably due to the fact that covering all the security aspects of MPC in a definition turns out to be much more complicated than one might expect.

2.1.2 Problems with defining security

Covering all the security aspects of a cryptographic task in a definition has turned out to be a quite difficult endeavor (many researchers speak about a project taking years). For instance, we have to consider the following security aspects:

Correctness: when all participants are honest, the output of the protocol is the same as the function it is emulating.

Privacy: no coalition of m' participants can learn information about the input of an honest participant if he does not cooperate, except from what can be deduced logically from the coalition's inputs and the output.

Honesty: no coalition of m'' participants can make an honest participant accept an output that is inconsistent with the inputs of all the other participants.

Fairness: each participant should learn its output, i.e. no participant should be allowed to learn its output and quit before other participants have learned their output.

Resilience: this reflects the protocol's ability to complete the computation of f if several participants stop cooperating during the protocol.

There are two major problems with this approach. In the first place, all these security aspects mentioned above seem to have a somewhat ad hoc character: how can we be sure whether we have covered all security aspects? Secondly, it can be shown that proving a protocol private and correct does not necessarily lead to a secure protocol. In other words, the simple conjunction of all security aspects is not always secure. The counterexample below (here adopted to a two-party setting) is due to Micali and Rogaway [MR91].

This counterexample uses the notion of a Bit Commitment. Intuitively this cryptographic task can be compared to A writing a 0 or 1 on a piece of paper, putting it in a strong-box, closing the strong-box and giving the strong-box to B, while keeping the key. This is the *commit* phase. In the *open* phase A gives the key to B, who opens the strong-box and sees the bit.

> COUNTEREXAMPLE: Let x and y be the inputs of A and B respectively. The function g they want to evaluate jointly is defined as
>
> $$g_{\mathsf{A}}(x,y) = g_{\mathsf{B}}(x,y) := \begin{cases} 0x & \text{if } \mathrm{lsb}(y) = 0 \\ 1y & \text{if } \mathrm{lsb}(y) = 1, \end{cases}$$
>
> where $\mathrm{lsb}(y)$ stands for the least significant bit of y, and where juxtaposition serves as the concatenation operator.
>
> Consider now the following protocol implementation:
>
> *(1)* Both A and B use some perfectly secure Bit Commitment scheme to encrypt their input bits.
>
> *(2)* B opens $\mathrm{lsb}(y)$.
>
> *(3)* On the value of the bit opened, either A or B completely opens her or his input string, which is output by both parties.
>
> *(4)* A sends x to B.

Clearly this protocol implementation leaks more information than it should because of step *(4)*: according to the specification B learns A's input x only when the $\mathrm{lsb}(y) = 0$, whereas in the implementation B learns the value of x always.

Nevertheless, this protocol is private and is correct, when these notions are defined independently. Intuitively, the argument is as follows.

To see that the protocol is correct for both A and B is easy: obviously both parties output what they are supposed to. The fact that the protocol does more in step *(4)* is true, but is irrelevant as far as correctness is concerned; B can just ignore this information.

To see why the protocol is private for B, one must realize that for a dishonest B, who wants to maximize his information about x, the protocol allows for a perfectly honest strategy in which he always learns x: choose $\text{lsb}(y) = 0$. Thus, according to the protocol, A opens x in step *(3)*, and when receiving x again in step *(4)* B doesn't learn anything he didn't know already. That the protocol is also private for A is obvious.

The function g is not as far-fetched as it seems; simple Boolean functions as AND and OR allow B a choice of input in which he *always* learns A's input, making any definition of security that defines privacy and correctness separately void. Observe that the counterexample also works if A does not receive the output of g (though we have to be more creative to find a protocol implementation of this task).

This point is crucial if we want to use sub-protocols as building blocks for designing larger protocols, in the same way as subroutines are used to write large computer programs. Suppose we design a cryptographic task, which consists of several sub-tasks. To obtain a protocol implementation we would like to take an implementation for each subtask and combine them. But this implies we need that this sub-protocol acts as a black box, implementing the task it was designed for and providing the same security aspects as a black box. Having such a notion of "black box security" seems highly desirable if one wants to obtain a modular approach to protocol design. In other words, using terminology from complexity theory, we want to reduce the security of one protocol to some other protocol (or primitive).

Micali and Rogaway [MR91] were among the first to present formal definitions of MPC, using simulators in a way analogous to the definition of zero knowledge. This work continues the approach taken by Crépeau in [CR90], which studies reductions between various flavors of Oblivious Transfer. Another approach was taken by Beaver [BE90, BE91B]. Though from a global point of view his approach seems similar, it differs significantly in some of the technical details. More recently Canetti, Feige, Goldreich and Naor [CA96, CFGN96] drawing inspiration from these results, have come up with another variation. In particular, their approach deals with MPC in a computational setting, which for a long time was considered to be difficult, but which makes it also less relevant here.

A completely different avenue was pursued by Crépeau and Brassard [CR94, BCS96, BC97], who give a definition of One-out-of-Two Oblivious Transfer purely in terms of information theory.

The goal of this chapter is *not* to present a complete formal model of security for protocols for MPC, with formal proofs of all composition theorems etc.; this could easily be the subject of a thesis in itself. Instead, we want to emphasize the *ideas* behind the various definitions that exist, in order to prepare the stage for definitions of security for quantum protocols, the subject of Chapter 4. We do this by proposing a definition which is a minor variation on Beaver's approach. Then we compare this approach with Micali and Rogaway's, and Brassard and Crépeau's, and argue why the choice made here is preferable.

So as not to complicate matters too much, we restrict ourselves to the task of Two-Party Computation. Consequently, many of the difficulties that arise when defining security of protocols for *Multi*-Party Computation are, interesting as they might be, eventually ignored. From now on, x and y will denote A's and B's input, while f_{A} and f_{B} are their respective outputs. As usual, A is feminine, B masculine, but E, the adversary, is referred to as a thing (contrary to the convention of calling it Eve).

2.2 The definition

2.2.1 Cryptographic tasks

It is important to make a distinction between a protocol specification and a protocol implementation, just as software engineering teaches us to do the same for programs. When talking about "a protocol for coin flipping" it is not immediately clear whether we refer to an abstract protocol for coin flipping (i.e. the protocol specification) or whether we refer to a specific implementation. In this thesis, and certainly in this chapter, we want to be precise about this distinction. Therefore we call a protocol specification a *cryptographic task* (or just *task*); for simple tasks (like BC or OT) also the word *primitive* is used. On the other hand, *protocol* actually means a protocol implementation.

There are many conceivable ways of specifying a task, for instance simply by giving a specific protocol implementation, by specifying what would happen if some (or all) of the participants were honest, or by referring to some everyday procedure, like flipping a coin. A more uniform way is obtained by introducing a neutral party, called the *trusted host*, who serves as an intermediary to accomplish the task. For instance, Figure 2.1 and 2.2 specify the tasks for Bit Commitment and for Two-Party Computation, which are denoted as BC and TPC_f respectively. It is sometimes argued that TPC, or rather its generalization MPC, is so general that it comprises every cryptographic task conceivable. In a computational setting this may be true, but in a setting without computational assumptions this is not obvious at all. For instance, given a primitive for MPC it is not clear how to build a signature scheme.

One-out-of-Two Oblivious Transfer, denoted $\binom{2}{1}$-OT, is a special case of TPC_f. Let x be a string consisting of two bits, so $x = x_0x_1$, let y consist of one bit, and let f_B be the selection function: $f_\mathsf{B}(x_0x_1, y) := x_y$. Furthermore, A's output is the empty string, i.e. $f_\mathsf{A} = \Lambda$.

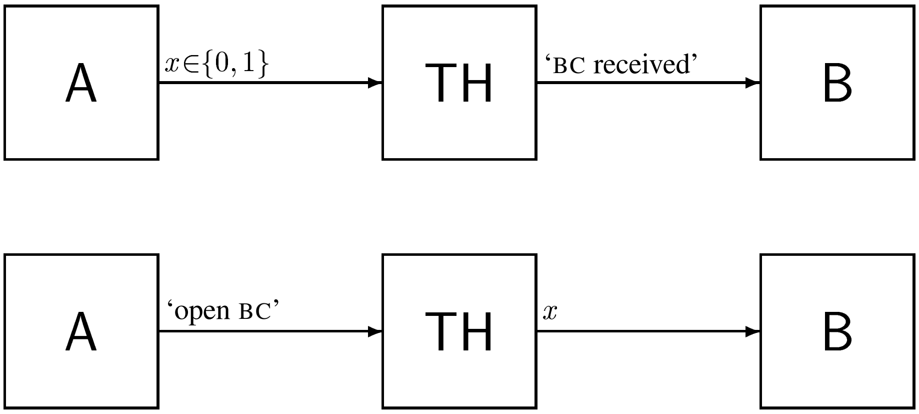

Figure 2.1: Task for Bit Commitment. First the COMMIT phase, possibly followed by the OPEN phase. The trusted host is the same.

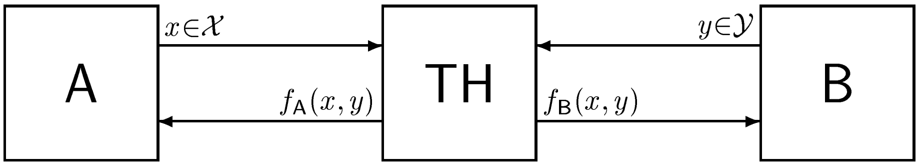

Figure 2.2: Task for Two-Party Computation of $f(x, y)$.

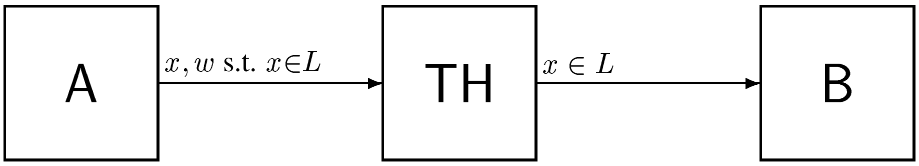

Figure 2.3: Task for Proof of Knowledge.

Figure 2.3 specifies the task for a Proof of Knowledge. Observe that, apart from sending x, the prover A also sends the witness w to the trusted host. This point is overlooked in [BE90]. In the first place it is logical to assume that the trusted party is restricted to probabilistic polynomial time only; otherwise the trusted host could be asked to perform tasks that are hard to compute or not computable at all (as deciding whether two graphs are isomorphic or deciding whether two programs always have the same output). Moreover, if A does not send w as well, then B could use the trusted host as an oracle [BE96].

2.2.2 Cryptographic protocols

In the context of this chapter an honest party (including a trusted host, to be introduced later) is supposed to have access to a probabilistic polynomial time program which is entirely dedicated to running a protocol. We model this program as a Turing Machine to have an explicit way to specify the parties' input and output. For a precise description of this model we make some minor modifications to the one presented by Crépeau [Cr90], in order to obtain a model that easily extends to quantum protocols, the subject of Chapter 4.

Apart from the input tape, work tape, random tape and output tape these Turing Machines, $\mathtt{A}$ and $\mathtt{B}$, are equipped with one shared tape, the *communication tape*. Auxiliary inputs are dealt with by quantifying over all possible stochastic variables (see the remark following Definition 2.1), hence no auxiliary input tape is required. The parties take turns using this tape to send each other messages, and do not have simultaneous access to it. This is accomplished as follows: When a party does not have access to the tape, her/his tape head is in the start position. When a party gets access to the tape, s/he reads the message until the end marker and returns the head to the initial position. At a later stage this party sends a message by writing it on the tape finishing with an end marker. Then s/he returns the head to the start position, and gives a signal to activate the other party.

Both machines use some pre-agreed alphabet (for instance ASCII) and message format (for instance assigning special characters to act as field delimiters or end-of-message markers). Various choices are possible and have been presented in various places in the literature (see [Be91b, MR91]), but supplying such details hardly contributes to clarity of exposition. The main point is that the reader be convinced that, given a specific choice of an alphabet and a message format, the stochastic variables introduced in the next paragraphs are well-defined. Without loss of generality we can assume they are defined over $\{0,1\}^*$.

A protocol proceeds in steps; we assume that A starts. The input of a party is the contents of her/his input tape. Observe that the definition ensures that the operation of each program depends only on its input, its random bits, and on the message received, so that any additional information an honest party may possess does not influence program execution. The messages sent include meta-messages such as ABORT. If neither party aborts, then they end the protocol by sending each other ACCEPT messages and entering a halting state.

Let $\mathtt{A}$ and $\mathtt{B}$ denote the programs that A and B run when they execute protocol $\Pi = [\mathtt{A}, \mathtt{B}]$. Let $[\mathtt{A}, \mathtt{B}](x)(y)$ denote the pair containing A's and B's *output* when they execute this protocol on input x and y. This is a stochastic variable because of the random tapes. By taking the first or second component we obtain a party's individual output, denoted as $[\mathtt{A}, \mathtt{B}]_{\mathsf{P}}(x)(y)$.

We assume that each party devotes one track of its work tape for logging purposes only: we assume s/he makes a copy of each symbol it reads from the input tape, the random tape and the communication tape. S/he also makes copies of symbols sent off through the communication tape, and puts special markers when a new step begins. Thus the final state of the work tape of P includes the finite part of P's random tape that s/he has read to obtain random bits, and the messages sent back and forth between A and B during the execution of $[\mathtt{A}, \mathtt{B}]$ on x and y. This defines another stochastic variable which is usually called P's *view* of protocol $[\mathtt{A}, \mathtt{B}]$, denoted as $[\mathtt{A}, \mathtt{B}]^*_{\mathsf{P}}(x)(y)$.

Attacks are modeled through an adversary, denoted E, who runs a Turing Machine $\mathtt{E}$. The adversary can be more powerful than A and B: it is allowed unlimited computational resources, and has the capacity to corrupt parties. The latter means that when a party has been corrupted, E completely dominates and dictates this party's behavior. In particular, E sees all the coin tosses generated by P, sees all the incoming messages received by P, defines the outgoing messages to be sent by P, etc. It is E's objective to disrupt the honest execution of Π in one way or another, for instance obtaining partial information about the input of an honest party, or influencing the function's output.

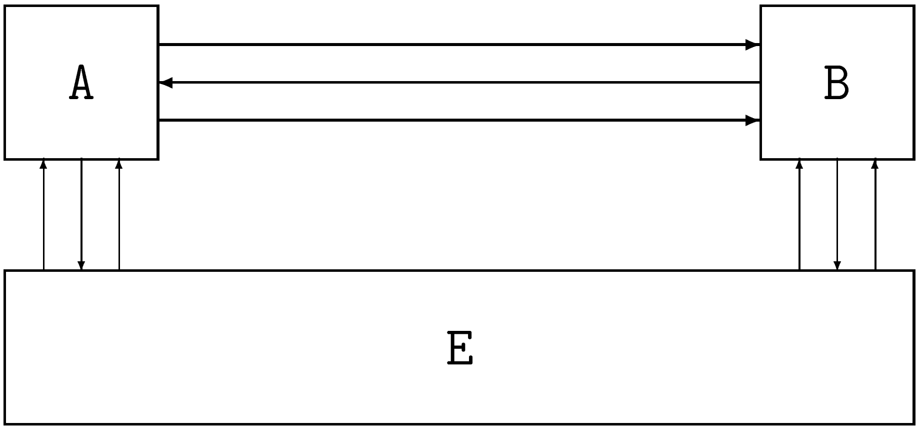

Figure 2.4: $\mathtt{E}$ attacks protocol $[\mathtt{A}, \mathtt{B}]$.

In the special case of two parties, E corrupts either A, either B, or no one (needed for correctness). The case that E corrupts both is less interesting, since no honest party is left whose security interests have to be protected (but in a multi-party setting with a dynamic adversary it *is* interesting [BE91B]). E cannot eavesdrop on the communication between A and B without corrupting either, so when corruption happens we may assume this takes place before the protocol started (conceptually in step 0), because a corruption during step r of the protocol is equivalent to the same party being corrupted from the beginning, but following the honest protocol up to step r.

Now we can extend the definition of protocol output and view to the situation where an adversary E running program $\mathtt{E}$ is present. Let $[\mathtt{A}, \mathtt{B}\colon \mathtt{E}]_{\mathsf{P}}(x)(y)$ denote the stochastic variable that describes the output obtained by P when executing the protocol $[\mathtt{A}, \mathtt{B}]$ in the presence of an adversary $\mathtt{E}$ who attacks either A or B. The view of a party is written as $[\mathtt{A}, \mathtt{B}\colon \mathtt{E}]^*_{\mathsf{P}}(x)(y)$, where P can be A, B or E.

In the definition in the next section we are interested in the triple consisting of the output of A, the output of B, and the view of E resulting from the same run of a protocol. We call this the *profile* of the protocol:

$$[\mathtt{A}, \mathtt{B}\colon \mathtt{E}]^{\ddagger}(x)(y) \stackrel{\text{def}}{=} \langle [\mathtt{A}, \mathtt{B}\colon \mathtt{E}]_{\mathsf{A}}(x)(y), [\mathtt{A}, \mathtt{B}\colon \mathtt{E}]_{\mathsf{B}}(x)(y), [\mathtt{A}, \mathtt{B}\colon \mathtt{E}]^*_{\mathsf{E}}(x)(y) \rangle \qquad (2.1)$$

Let X be a stochastic variable on the set $\mathcal{X}$, the set of A's allowed inputs, and let

Y be defined similarly for B. We assume that A, B and E know the joint stochastic variable $\langle X, Y\rangle$ on the set $\mathcal{X}\times\mathcal{Y}$. That is, for each run of the protocol we assume that a sample $\langle x, y\rangle$ is drawn from $\langle X, Y\rangle$ and that x and y are subsequently written on $\mathtt{A}$'s resp. $\mathtt{B}$'s input tape. Note that a corrupted player can completely ignore its input tape, $\mathtt{E}$ may supply an alternative input if it desires so. We can replace x and y (the instances) with X and Y (the stochastic variables) in the definitions of output and view in the previous paragraphs. The stochastic variables thus obtained, written as $[\mathtt{A}, \mathtt{B}]_{\mathsf{P}}(X)(Y)$ etc., are again well-defined.

In the design of cryptographic protocols, we often actually define a *family* of protocols parameterized by a *security parameter*, n (for instance denoting the length of some string, the number of rounds, or the number of photons used). So, instead of obtaining one stochastic variable, we obtain a family of stochastic variables, parameterized by a *security parameter*. We say that two families of stochastic variables $\{Z_1^{(n)}\}$ and $\{Z_2^{(n)}\}$ are *exponentially indistinguishable* (which is denoted as $\{Z_1^{(n)}\} \approx \{Z_2^{(n)}\}$) if they converge to each other exponentially fast:

$$\sum_{z\in\{0,1\}^*} |\mathrm{Prob}[Z_1^{(n)}{=}z] - \mathrm{Prob}[Z_2^{(n)}{=}z]| \leq \varepsilon^n \tag{2.2}$$

for some ε between 0 and 1. This notion is discussed in more detail in Section 3.7 of this thesis. In the remainder of this chapter we will not explicitly include the security parameter in the notation or the definitions; its presence is implicitly understood.

2.2.3 Comparing the security of protocols

For a detailed analysis of the security of cryptographic protocols we need a method to compare the security of two protocols. For instance, in a computational setting one could imagine some protocol Π_1 based on the assumption that the Schnorr identification protocol [Sc91] is secure. Formally this statement means that each attack on Π_1 implies that there exists an attack on Schnorr's scheme. In other words, we can say that Π_1 is at least as secure (as difficult to break) as Schnorr's scheme.

More generally we can define what it means for protocol Π_1 to be at least as secure as Π_2, when we assume that Π_1 and Π_2 implement the same task. (Later, in Subsection 2.2.7, we discuss the situation when Π_1 and Π_2 implement different tasks.) Suppose that an attack E on Π_1 exists. As in the previous paragraph, we require that any attack on Π_1 implies an attack on Π_2. This requirement can be enforced by demanding the existence of an algorithm, S, capable of translating any enemy $\mathtt{E}_1$ attacking Π_1 into an enemy $\mathtt{E}_2$ attacking Π_2. Let us define this more formally, following the terminology used in [BE91B] but the notation introduced in [BCS96].

Let $\mathcal{E}_1$ be the set of possible Turing Machines E representing valid attacks on protocol Π_1; define $\mathcal{E}_2$ similarly. We say that S is an *interface* between Π_1 and Π_2 if for all $\mathtt{E}_1 \in \mathcal{E}_1$ we have that $\mathtt{S}(\mathtt{E}_1) \in \mathcal{E}_2$. That is, S translates any valid attack on Π_1 into a valid attack on Π_2. The purpose of the interface is to make protocols comparable. On one side of the interface the adversary E is situated, who thinks it attacks protocol Π_1. On the other side we have Π_2. It is the interface's responsibility to let E talk 'smoothly' to Π_2, if necessary by modifying or creating messages.

As an elementary example, Π_1 and Π_2 could be two identical protocols, except that the message formats used are different. In this case the interface just has to translate the messages from one format to the other. The result is that an E attacking protocol Π_1 is translated into $\mathtt{S}(\mathtt{E}_1)$ attacking protocol Π_2, so $\mathtt{S}(\mathtt{E}_1)$ is an adversary in its own right.

From the interface's point of view, E is regarded as the code of some Turing Machine, and S can take "snapshots" of E's configuration. This mechanism is called *universal simulation*; even though we have modeled E so far as Turing Machine, the existence of a Universal Turing Machine makes this choice reasonable. The rationale for this choice is discussed later in Section 2.3, but the idea is to allow S capabilities similar to an operating system like suspend, dump, resume, etc. We also require that the time $\mathtt{S}(\mathtt{E}_1)$ needs to attack protocol Π_2 is bounded polynomially by the time it takes E to attack Π_1. Observe that [BE91B] uses black box access, and poses no time restriction on S.

Even though we assume here only two parties and corruptions right from the start, it is important to mention the following point: Another restriction we impose on the interface is that it translates corruption requests faithfully, i.e. instantaneously and to the same party. In particular, an interface is not allowed to issue a corruption request on its own (without a corresponding request from $\mathtt{E}$). If such interfaces were allowed there would be no guarantee that both protocols behave the same in the absence of an adversary, thus jeopardizing the correctness-part in the definitions that follows.

The following definition formalizes the intuition presented in the first two paragraphs of this section, and is very similar to the one in [BE91B]. Note that in the equation below, the third component of the profile $[\mathtt{A}_2, \mathtt{B}_2 \colon \mathtt{S}(\mathtt{E})]^{\ddagger}(X)(Y)$ stands for E's view in protocol Π_2, *not* $\mathtt{S}(\mathtt{E})$'s view. Recall that the first and the second component stand for A's resp. B's output. Consequently, the definition implies that the protocols Π_1 and Π_2 implement the same task.

2.1 Definition *Protocol* $\Pi_1 = [\mathtt{A}_1, \mathtt{B}_1]$ *is* <u>*at least as secure as*</u> $\Pi_2 = [\mathtt{A}_2, \mathtt{B}_2]$ *if there exists an interface* $\mathtt{S}$ *from* $[\mathtt{A}_1, \mathtt{B}_1]$ *to* $[\mathtt{A}_2, \mathtt{B}_2]$ *such that for all* $\mathtt{E} \in \mathcal{E}_1$ *and for all possible* $\langle X, Y \rangle$ *which are joint stochastic variables on* $\mathcal{X} \times \mathcal{Y}$*:*

$$[\mathtt{A}_1, \mathtt{B}_1 \colon \mathtt{E}]^{\ddagger}(X)(Y) \approx [\mathtt{A}_2, \mathtt{B}_2 \colon \mathtt{S}(\mathtt{E})]^{\ddagger}(X)(Y). \tag{2.3}$$

By choosing X and Y to range over all possible stochastic variables, we deal with the problem of auxiliary information, such as information obtained from previous protocols. This issue is crucial in order to prove that the concatenation of secure protocols remains secure. To see why the definition presented here is equivalent to one where the universal quantifier ranges over all possible input *and* auxiliary input (see for instance [GMR89, GO94, CR94, BE91A]), observe the following: from A's point of view, the variable X *given* an (arbitrary) auxiliary input can be regarded as another stochastic variable X'; its corresponding probability distribution is simply a readjustment of the probability distribution on X using the rules of conditional probabilities.

2.2.4 Sequential composition of protocols

One of the main objectives when defining security of protocols for Multi-Party Computation is to obtain a composition theorem: if we compose two arbitrary protocols which are secure, we want the sequential composition to be secure as well. (As an aside, note that the notion of security discussed in [BCC88], minimum disclosure, cannot meet this objective. A minimum disclosure protocol Π_1 does not leak any information about the secret, but may leak other information. This other information could be vital in the breaking of another minimum disclosure protocol Π_2.)

First we make a small modification of a Turing Machine's behavior at the end of the protocol: each symbol on the track of the work tape not used for logging purposes is moved (copied and erased) to the logging track, ended with a special protocol separator symbol. So all the information is moved to the logging track while the non-logging track is empty.

Let concatenation be denoted by juxtaposition. Then $[\mathtt{A}'\mathtt{A}'', \mathtt{B}'\mathtt{B}'']$ stands for the execution of protocol $[\mathtt{A}', \mathtt{B}']$ followed by $[\mathtt{A}'', \mathtt{B}'']$. A special symbol is used by A in order to separate the two inputs x' and x'' to both protocols. Note that it is permitted that x'' depends on A's output of $[\mathtt{A}', \mathtt{B}']$. Similar arrangements apply to B's input, both parties' output, and likewise for E. For simplicity we follow the convention that an abort in the first protocol means an abort of the whole protocol. The definition of the output of A and B, the view of E and therefore the profile of the composite protocol $[\mathtt{A}'\mathtt{A}'', \mathtt{B}'\mathtt{B}'']$ should now be clear. We can now state the following protocol composition theorem.

2.2 Theorem *Let* $[\mathtt{A}'_1, \mathtt{B}'_1]$ *be as secure as* $[\mathtt{A}'_2, \mathtt{B}'_2]$*, and let* $[\mathtt{A}''_1, \mathtt{B}''_1]$ *be as secure as* $[\mathtt{A}''_2, \mathtt{B}''_2]$*. Then* $\Pi_1 = [\mathtt{A}'_1\mathtt{A}''_1, \mathtt{B}'_1\mathtt{B}''_1]$ *is as secure as* $\Pi_2 = [\mathtt{A}'_2\mathtt{A}''_2, \mathtt{B}'_2\mathtt{B}''_2]$

PROOF: We have to show that there exists an interface $\mathtt{S}$ from Π_1 to Π_2 which satisfies equation (2.3) for any adversary $\mathtt{E}$.

Suppose $\mathtt{E}$ corrupts neither party. Then $\mathtt{S}$ corrupts neither party, and (2.3) is fulfilled almost trivially.

Suppose $\mathtt{E}$ corrupts one party. Then $\mathtt{S}$ runs $\mathtt{S}'$ until the end of Π' is encountered, which can be recognized because of the ACCEPT messages sent. Then $\mathtt{S}$ runs $\mathtt{S}''$ until the end. To see why $\mathtt{S}$ indeed satisfies equation (2.3), observe that $\mathtt{S}$ can simulate Π' correctly because $\mathtt{S}'$ can. Similarly, $\mathtt{S}$ can simulate Π' correctly; the auxiliary information possibly obtained in Π' cannot change this, because Definition 2.1 ranges over all possible joint stochastic variables $\langle X, Y \rangle$. Since for both Π' and Π'' a simulator must exist by hypothesis, $\mathtt{S}$ is an interface between Π_1 and Π_2. $\Box$

This proof can be extended to the sequential composition of $\mathrm{poly}(n)$ protocols, where n is the security parameter. Recall that we defined $\mathtt{S}$ to be restricted to polynomial time in $\mathtt{E}$, with the possibility to rewind. In sequential composition, if simulation of a sub-protocol fails, we can go back to the state before the sub-protocol had begun (because we had taken a snapshot), and try again. This technique is called *rewinding*, and we will discuss it later on. Even though it means that $\mathtt{S}$ needs more time the overall time only expands polynomially, as long as the probability of success is non-negligible. However, in parallel composition this does not work, because our probability of successful simulation is now exponentially small in n if rewinding is necessary.

2.2.5 Ideal protocols

An ideal protocol is the implementation of a task $\mathbf{T}$ in the presence of a trusted host. An ideal protocol is not the same as a task, though they look very much alike. A task is merely a specification of what a protocol is supposed to achieve. In contrast, in an ideal protocol A and B (when corrupted by the adversary E) can be dishonest and deviate from the protocol in any way they like. To put it differently, an ideal protocol is supposed to deal with all the eventualities of cheating or non-cooperating parties. For instance, in Bit Commitment a cheating A could send a bit string instead of a bit, or just have sent nothing after some timer has expired. When this happens the trusted party sends the message NOT-COOPERATING to B and aborts.

A trusted host is by definition incorruptible and always cooperating. Observe that an adversary attacking an ideal protocol, even though limited, has some power nevertheless: it can force a corrupted player to supply different inputs than s/he is supposed to, or refuse to cooperate in the middle of the protocol.

The way the trusted host handles the irregularities it detects is always specified and is known to both A and B. A party who sends an input value outside the domain ($\mathcal{X}$ or $\mathcal{Y}$) is cheating; when this occurs the trusted party could either report cheating and abort, or report and set the input to some default value known to A and B, or even set the input to some default value without reporting at all. The last two possibilities might be more relevant in a multi-party setting; in a two-party setting we assume that once cheating has been detected, any cooperation is broken off forever, so we choose the first possibility. We also suppose that in the ideal protocol the channel from an honest participant to a Trusted Party cannot be eavesdropped by E.

After this (intuitive) description of the trusted party's behavior, the ideal protocol for a particular task is essentially unique. Ideal protocols will be written with calligraphic letters. It seems not difficult to extend the tasks presented in Section 2.2.1, such that ideal protocols for the tasks of BC, $\binom{2}{1}$-OT, TPC and POK are obtained.

However there is a subtle problem here with BC: in the ideal protocol a trusted host can detect a cheating A who commits to rubbish already in the COMMIT phase, but in most (if not all) implementations this is discovered in the OPEN phase. This seems an inherent consequence of the fact that BC is a two-phase task, as opposed to the other examples. One (maybe inelegant) way to modify the definition of an ideal Bit Commitment, is to specify that cheating which occurs during the COMMIT phase is only reported during the OPEN phase. Though this can hardly be called 'ideal', it seems a better reflection of what protocol implementations can achieve in reality. The modified ideal task for BC, written $\mathcal{BC}$, is given in Figure 2.5.

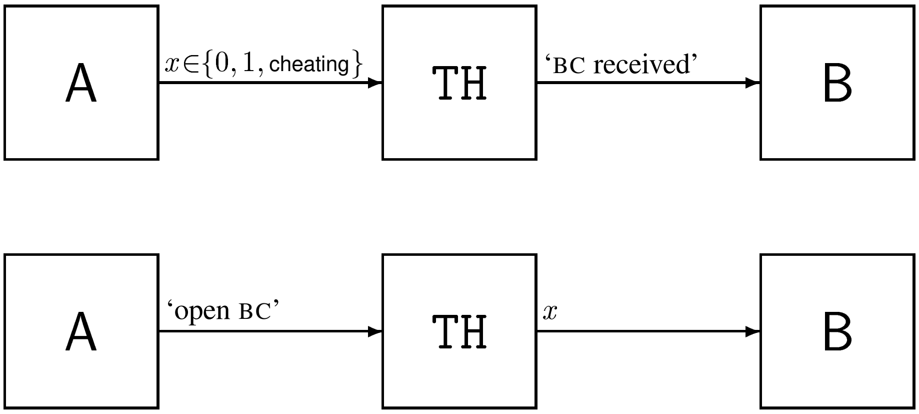

Figure 2.5: Modified ideal protocol for Bit Commitment. The trusted host does not report cheating during the COMMIT phase, but during the OPEN phase.

It can be argued that the distinction between a task and an ideal protocol is somewhat artificial. This is partially intentional; we defined a task to resemble an ideal protocol. The main issue is that the ideal protocol defines properly the trusted host's behavior when it detects cheating. And it seems convenient to be able to say that an ideal protocol, for instance $\mathcal{BC}$, is a particular implementation of the task BC.

2.2.6 Security

The ideal protocol provides us the perfect yardstick for expressing the security of a cryptographic protocol: we expect that a secure implementation of some task provide at least the same security as (is at least as difficult to break as) the ideal protocol for that task, thus comparing a situation without trust with one where there is trust. This is summarized in the following definition, which combines the definition of relative security with the notion of an ideal protocol.

2.3 Definition *Let* $\Pi_1 = [\mathtt{A}, \mathtt{B}]$ *be the protocol implementation of a cryptographic task* $\mathbf{T}$*, while* $\Pi_2 = [\mathcal{A}, \mathcal{B}]$ *is the ideal protocol for* $\mathbf{T}$*. Then* $[\mathtt{A}, \mathtt{B}]$ *is a* <u>*secure protocol for* $\mathbf{T}$</u> *if* $[\mathtt{A}, \mathtt{B}]$ *is as secure as* $[\mathcal{A}, \mathcal{B}]$.

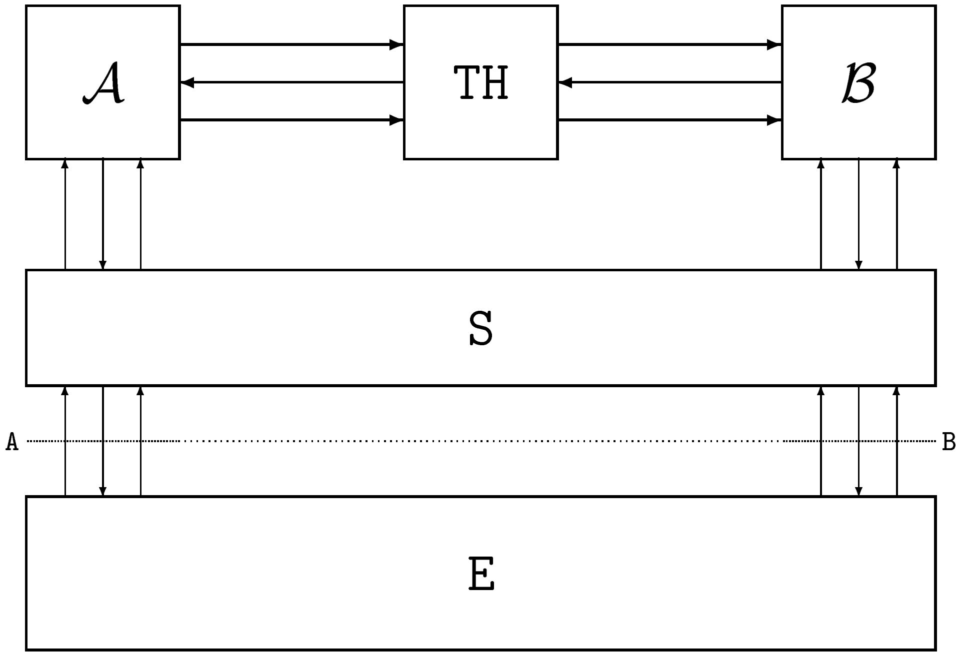

Figure 2.6: E attacks the ideal protocol $[\mathcal{A}, \mathcal{B}]$ through S. The dotted line indicates what 'image' S is supposed to show to E. Compare this with Figure 2.4.

Figure 2.6 depicts the intuition behind this definition.

Note that there exist various other ways to express this definition. Some people would say that the protocol $[\mathtt{A}, \mathtt{B}]$ is simulatable with the help of the ideal protocol $[\mathcal{A}, \mathcal{B}]$. For the special case of joint function evaluation, i.e. MPC, the ideal protocol is also called a function oracle. One could also say that the protocol $[\mathtt{A}, \mathtt{B}]$ reduces to the ideal protocol $[\mathcal{A}, \mathcal{B}]$.

Correctness and privacy can now be seen as special cases of this definition. Correctness (and honesty) is obtained by restricting the two profiles (triples) in Equation (2.3) to their first two components, A's and B's output. Privacy is obtained by restricting to the last component, E's view.

Let us now return to our Counterexample on page 17, and argue that the protocol presented there does not pass this new definition. Let $[\mathtt{A}, \mathtt{B}]$ be the protocol implementation using Bit Commitments and let $[\mathcal{A}, \mathcal{B}]$ be the ideal protocol for Two-Party Computation of the function g. Consider an adversary E that corrupts B but that is otherwise honest. Such an adversary, representing an honest-but-curious B, is perfectly well-defined, and we show that there cannot exist an interface which satisfies definition 2.1.

Recall that the interface S, with the help of $[\mathcal{A}, \mathcal{B}]$, must provide a profile indistinguishable from the profile of $[\mathtt{A}, \mathtt{B}]$, for any distributions on the inputs x and y. Suppose that $lsb(y) = 1$. If the interface S passes y to $[\mathcal{A}, \mathcal{B}]$ unmodified, it will not learn x. Therefore, except with negligible probability, S cannot simulate step *(4)* of $[\mathtt{A}, \mathtt{B}]$, and E's view in $[\mathtt{A}, \mathtt{B}]$ will almost always differ from its view in $[\mathcal{A}, \mathcal{B}]$. Alternatively, S could modify the least significant bit of y to 0, in order to learn x. But this will also modify A's output in $[\mathcal{A}, \mathcal{B}]$ to $0x$, whereas in $[\mathtt{A}, \mathtt{B}]$ A's output equals $1y$. We conclude that whichever value S chooses as the least significant bit of y as input to $[\mathcal{A}, \mathcal{B}]$, the profiles of $[\mathtt{A}, \mathtt{B}]$ and $[\mathcal{A}, \mathcal{B}]$ are always different.

2.2.7 Reductions

The composition of ideal protocols does not automatically yield an ideal protocol. The problem is that for each sub-protocol (primitive, black box) a different Trusted Party is used, but there is no "overall" trusted party who verifies that inputs and outputs from sub-protocols are indeed consistent. In other words, even though the protocols themselves are guaranteed to be honest, their input and output do not have to be. This can be solved by imposing extra conditions on the input of sub-protocols. For instance, in the reduction from $\mathcal{BC}$ to $\binom{2}{1}$-$\mathcal{OT}$, the protocol is performed n times in parallel [Ki90], or (more efficiently) an error correcting code is imposed on the inputs to $\binom{2}{1}$-$\mathcal{OT}$. Note that this is a one stage reduction: input are given to many copies of $\binom{2}{1}$-$\mathcal{OT}$, and many outputs result.

Things become even more complicated in multi-stage reduction, where the output of one ideal protocol is used as input for the next protocol. For instance, the ideal protocols could represent the AND, OR and NOT gates of a circuit to be evaluated. Only Beaver's methods [Be90, Be91a] seem sufficiently developed to give a formal proof that Two- and Multi-Party Computation can be built from such primitives. The details of these techniques are beyond the scope of this thesis.

2.3 Comparison with other definitions

The discussion so far has mostly followed the approach presented in [BE91B]. First we compare it briefly with similar approaches. Then we elaborate on the computational power of S and its access to E.

2.3.1 Oracles, simulators, etc.

Besides [BE91B], definitions of security have been proposed by Micali and Rogaway [MR91] and Canetti, Feige, Goldreich and Naor [CFGN96]. In spirit, all three approaches are similar. It is in technical details and in terminology that differences occur. The issues are the following:

(a) An interface in [BE91B] corresponds to a simulator in [MR91] and [CFGN96]. Similarly, an ideal protocol in [BE91B] corresponds to an oracle in [MR91] and [CFGN96].

Let us restate the intuition behind [BE91B]: protocol Π_1 is as secure as the ideal protocol Π_2 if each attack on Π_1 can be converted into an attack on Π_2: an attack on the whole implies an attack on its components. The intuition behind [MR91] and [CFGN96] is an extension of the intuition behind zero knowledge: protocol Π_1 is secure if the adversary can simulate its view of the protocol with the help of an oracle which represents the ideal protocol. In the original definition of zero knowledge no oracle was necessary, since the fact that x is in L is, by assumption, exactly the information conveyed by this protocol.

To a large extent both intuitions are valid and interchangeable, they just are different ways of seeing the same problem, the simulator paradigm probably preferred by people who are more familiar with the original definition of zero knowledge. Indeed, Figure 2.6 supports both point of views. But it seems that the approach of [BE91B] is somewhat crisper, making it easier to derive precise definitions, especially when the exact powers of the interface (simulator) and the ideal protocol (oracle) have to be specified. Moreover, in situations such as the example given at the beginning of Section 2.2.3 (a protocol as secure as Schnorr's identification protocol) the simulator paradigm seems less appropriate.

(b) In [BE91B] the notion of relative security and the notion of an ideal protocol are introduced separately, and absolute security is obtained by using the ideal protocol as just another protocol implementation. In [MR91] and [CFGN96] these two notions are intertwined in one definition, which makes it less general.

(c) In [BE91B] the protocol task could be anything, whereas [MR91] and [CFGN96] restrict themselves to Multi-Party Computation (i.e. evaluation of a function f). For instance, a definition of security for BC does not follow immediately; though not difficult, some minor adjustments seem necessary.

(d) In [BE91B] the profile is used to compare protocols, whereas the condition in [MR91] (definition (3.10)) gets rather complicated, reason why we prefer the first. We return to this point in Section 2.5.2. In [CFGN96] the approach of [BE91B] is adopted.

2.3.2 The relation between S and E

In this section we return to the discussion about the computational strength of the interface S, and its access to the adversary program E, as was mentioned in Section 2.2.3. Though at first glance this may seem a minor technical point, it turns out that the choices made here significantly influence the notion of security obtained. This is already true for the classical definition; but is also of crucial importance in the quantum setting, as we will see in Chapter 4.

Much of what is discussed in this section goes back to the definition of zero knowledge. As has been mentioned already, the intuition behind the definition of zero knowledge is as follows: suppose that a participant is always able to reproduce by himself all the messages that he sees during the execution of the protocol. Thus this protocol cannot leak any knowledge (= polynomially computable information) about the secret nor about anything else.

Even though the intuition is clear, there are various ways to formalize it. The most important issues are listed below; they are interrelated.

(a) The order of the quantifiers.

In the original definition of zero knowledge as presented in [GMR85], it is required that for *all* adversaries E (verifiers) there *exists* an interface (simulator) S that reproduces the conversation with (almost) the right probability distribution. However, in [BC89] it is argued that, certainly in the original context of computational protocols for proofs of knowledge, this seems unreasonable. For example, suppose that such an S exists, but that it is computationally infeasible to find. This implies that for a polynomial-time interface it would be easier to find the secret (the witness proving that $x \in L$) in order to provide a simulation, than to find the simulator. Therefore it seems impossible to maintain that E has gained nothing from the interaction.

A possible alternative, put forward in [BC89], is to reverse the quantifiers in the definition: there *exists* an interface (simulator) S such that for *all* adversaries E the conversation can be reproduced. This line of thinking is also adopted by [BE91B] and [MR91]. (The manuscript of [MR91] seems to contain reminders to the authors to devote a paragraph to the order of the quantifiers; unfortunately none have been written. In his thesis [RO91], Rogaway makes the same choice, but does not discuss the issue either.)

(b) The computational power of the interface. The interface should never be more powerful than the adversary, since it would imply that the combined adversary $\mathtt{S}(\mathtt{E})$ is much stronger than $\mathtt{E}$ alone. Consequently, the definition would proclaim Π_1 as secure as Π_2, even though the first one is attacked by a much weaker adversary.

Note that this can happen inadvertently. For instance, even if $\mathcal{E}_1$ includes unbounded adversaries programs $\mathtt{E}$, the quantifier *for all* includes also computationally limited $\mathtt{E}$. If $\mathtt{S}$ needs to work very hard (say exponential time) in order for $\mathtt{S}(\mathtt{E})$ to be an attack against Π_2, the same argument applies. As an example, take Π_1 to be a (non-cryptographic) protocol in which A sends a string to B; while in Π_2 the same string is 'encoded' as the solution of some *NP*-complete problem. We will return to this point below and justify our choices.

(c) The extent to which the interface can influence the adversary. In all definitions $\mathtt{S}$ takes $\mathtt{E}$ as an input. However, there are several conceivable ways this can be done. The following two ways are the most important ones.

(i) $\mathtt{S}$ gets the *code* of the algorithm $\mathtt{E}$, for instance as a program for a universal Turing Machine. This is called *universal simulation* by Oren [OR87], and is also preferred in [BC89]. Of course, $\mathtt{S}$ is not granted any powers that are considered infeasible, but it seems reasonable to allow $\mathtt{S}$ powers similar to an operating system running a specific program. In particular, we could allow $\mathtt{S}$ to make a complete 'snapshot' of the program $\mathtt{E}$, i.e. a dump of all its memory and program counters, to run several copies of this process maybe using different inputs, to suspend or to kill such processes, etc. Note that it explicitly includes the capacity to *rewind* $\mathtt{E}$, i.e. to make $\mathtt{E}$ go back to an earlier state. This can be done by making snapshots during the computation, and to erase any trace of a computation that went "too far".

(ii) S gets E as a resettable *black box*. This means that S has no knowledge about, nor control over what is going on inside the program E. However, S is allowed read/write access to all E's tapes, including its random tape. In other words, it is somewhat like universal simulation though without the ability to make snapshots, but including the ability to start all over again, using different bits on the input and random tapes.

Resetting gives the interface S the power to rewind E. The idea is that S can restart E from its initial configuration and halt it when it reaches the desired configuration. Note, however, that this implies that the randomness is provided to E by means of a random tape. As soon as E gets its randomness from an internal random source (for instance, by pushing a button or really tossing a coin), S's capability to rewind E is revoked.

The history of these notions is as follows. The original definition of zero knowledge [GMR85] contained a flaw: it lacked the desired property that the sequential composition of zero knowledge protocols was again zero knowledge. This was observed by several researchers independently, including the authors. It is resolved by requiring that, even when given arbitrary side information, usually called *auxiliary input*, the real and the simulated conversation are indistinguishable.

Oren [OR87] showed that black box zero knowledge implies universal simulation zero knowledge, which in its turn implies auxiliary input zero knowledge. Moreover, even though the black box definition seems restrictive, it turns out that all proofs in fact used a reasoning equivalent to black box simulation to show that such a simulator exist. To some extent this has led to black box zero knowledge having been accepted as the *de facto* standard; for example, in Oren's final paper, together with Goldreich [GO94], no mention of universal simulation is made. This is unfortunate, because it somehow precludes any discussion about the rationale behind the black box definition. It seems that the following fundamental question often has been side-stepped: *What is the significance of the existence of an adversary* E_1 *which attacks* Π_1*, in an attack against* Π_2*?* We believe that the following guidelines should be taken into account, though do not claim to have the ultimate answer to all the issues discussed in the previous subsections.

(a) *The interface* `S` *and the adversary program* `E` *are 'friends'.* They cooperate in attacking Π_2 (or in providing the optimal simulation, if you prefer this paradigm). There is no reason why `E` hides something from `S`, on the contrary.

(b) *The manner in which the attack against* Π_2 *is modeled given* `E` *should be 'constructive'.* Merely mathematical existence of `S` still does not mean that an entity which wants to construct `S` given `E` is capable of doing so. Observe that our 'universe' is full of mathematical notions which can be defined perfectly, but which are irrelevant since they are not computable. (Chaitin's number Ω, a real number that encodes the Halting probability, is an example; see for instance [LV93] for discussion.) This is why we prefer that there *exists* an interface `S` for *all* adversaries `E` which can reproduce the same distribution.

(c) *The computational effort it takes to attack* Π_2 *should bear some reasonable relation to the computational effort it takes* `E`. This means that `S` should not be given powers that are computationally infeasible. We believe that *universal simulation* is a good approximation of what in reality `S` can achieve, which includes the making of 'snapshots'.

In addition, we require that the total time it takes `S` to attack protocol Π_2, which includes the time it takes to run `E` as a procedure, is polynomially bounded by the time it takes `E` to attack Π_1.

The following table presents an overview of the choices made by several authors on the issues discussed in this section. However, the reader must be warned that more differences in the models and definitions exist, in particular with respect to the notion of indistinguishability used. The model referred to in the last line is quite different, and is the topic of Section 2.5.

paper	order quantifiers	computational power of E	computational power of S	simulation (S's access to E)
[GMR89]	∀E ∃S	*BPP*	*BPP*	auxiliary input
[OR87, GO94]	∀E ∃S	*BPP*	*BPP*	black box
[CFGN96]	∀E ∃S	*BPP*	*BPP*	black box
[BC89]	∃S ∀E	*BPP*	*BPP*	universal simulation
[MR91]	∃S ∀E	non-computable	unrestricted TM	black box
[BE91A]	∃S ∀E	unrestricted TM	unrestricted TM	black box
here	∃S ∀E	unrestricted TM	*BPP*	universal simulation
[BCS96]	∀E ∃S	(the equivalent of) E and S can be non-computable		

Table 2.1: A comparison of various definitions of security.

2.4 Simulatability and security

2.4.1 Physical assumptions

It must be emphasized that the preceding definitions hide some physical assumptions. For instance, the choice to extract the randomness of a Probabilistic Turing Machine by putting random bits on a dedicated tape and then using a Deterministic Turing Machine is one such assumption. If the randomness came from some internal, non-resettable source instead (or if we simply assumed that the tape is read-only and cannot be tampered with), then the simulation fails and many existing proofs of security for protocols completely collapse.

Next it must be realized that the view of a party is also something physical. In particular, the view is what a party sees on the tapes of his Turing Machine, to be compared with what he sees when using envelopes, playing cards, etc. Moreover, the fact that the interface S can create or modify the incoming messages of the adversary E is also a physical assumption. If A and B use pieces of paper to transmit their bits, with unforgeable signatures on the reverse side, the interface will almost always fail to satisfy E.

The previous paragraph, together with the definition of ideal Bit Commitment $\mathcal{BC}$ by means of a trusted host, implies that the interface can trivially open a bit commitment anyway it likes. It simply ignores the bit sent by $\mathcal{B}$ in the OPEN phase and replaces it by the bit it needs. In the terminology of [BCC88] the ideal $\mathcal{BC}$ protocol is chameleon, i.e. it can be opened by the interface in either way. Note that this simulation does not significantly differ from the way Kilian simulates the $\binom{2}{1}$-OT [KI90, p. 104]. The point is that in this model the view of a party is limited to what is written on his communication tape; who puts incoming messages on it is invisible to him. The tape acts as a barrier, beyond which any party, including the adversary, cannot penetrate.

Such a barrier is also present in what is probably the first indistinguishability test in computer science: the Turing test. In this well-known procedure, a human must distinguish another human from a computer generating human-like conversation. The interaction takes place through a keyboard and screen, which act as the barrier. In the every-day context of cryptography this barrier is not unreasonable: my physical view is limited to the wire that connects my computer to the socket in the wall, beyond it there is a network of which I know nothing, apart from the bits coming in and out.

2.4.2 Protocol security

Simulation, that is, an interface having to satisfy E, seems so far the only *general* method of defining security for protocols. However, it can be argued that for *individual* protocols, defined for a specific task, this need not be the case. For instance, the definition of a secure Bit Commitment protocol is possible without necessarily using simulatability. But this raises the following two questions:

(a) Do we believe that $\mathcal{BC}$, and a physical implementation of BC, e.g. using an envelope or a strong-box, provide exactly the same security?

(b) If the answer to *(a)* is yes, can we freely interchange one for the other in the context of a larger protocol, without affecting security?

The first question can, in my opinion, never be answered mathematically, only by common sense, since the two models are inherently incompatible. In my opinion, the answer to this question must be 'yes'. In the first place it seems easy to verify that whatever A and B can do in one protocol, it can also do in the other. Moreover, the definition of $\mathcal{BC}$ is completely based on our (classical) physical intuitions about what envelopes can achieve; if there was a mismatch it would mean that our definition of $\mathcal{BC}$ did not capture the right notion and we would have to adjust it. Answering 'no' would be to reject the definition. Indeed, the definition of many protocol tasks, and the kind of security it should provide, are based on the functionality of physical objects, such as envelopes, playing cards, signatures, unreliable mail service, elections, etc. Cryptographic protocols try to *model* this functionality.

With regard to the second question: if we believe that an envelope provides the same security as a strong-box, then using one as a primitive for a larger protocol or the other should not make any difference. However, if we extend this reasoning to $\mathcal{BC}$, we suddenly get a feature that envelopes and strong-boxes don't have: since E does not know where a bit sent to it comes from, S can open any $\mathcal{BC}$ either as a 0 or as a 1. This is a simple consequence of the different physical model used, but leads to some puzzling questions.

To conclude, my opinion is that simulatability is *not* a feature that reality gives us, but always depends on the way we model reality. If we change the model, we may be able to simulate more, or less. The latter seems to be the case for quantum protocols. On the other hand, no alternative 'paradigm' seems available, at least not at the moment. Information theory does not provide a way out, as we will see in the next section.

2.5 Another approach: using information theory

2.5.1 The definition

In [CR94] and [BCS96], Brassard and Crépeau follow a different approach to obtain a suitable definition of security for the particular task of One-out-of-Two Oblivious Transfer. This approach is extended in [BC97] to Generalized Oblivious Transfer (GOT). This protocol is an extension of of $\binom{2}{1}$-OT: B gets only one bit of information about the two bits put in by A. However, in GOT B's bit can be any Boolean function on the two input bits of A at B's choosing. Again, A receives no output. So A's input consists of two bits, x_0 and x_1, B's input consists of a predicate y from the set of all possible predicates (functions) from $\{0,1\}^2$ to $\{0,1\}$. Then B receives $y(x_0, x_1)$ while A receives nothing.

Below we have "naively" extended the definition of security for GOT from [BC97] to secure evaluation of an arbitrary function $f(x, y)$, that is, to TPC. So when we take the definition below and substitute f=GOT, $\mathcal{X}=\{0,1\}$ and $\mathcal{Y}=\{p : \{0,1\}^2 \to \{0,1\}\}$ we obtain their definition, the definition is downwards compatible. Then we compare this new definition with the one we presented earlier, Definition 2.3, and explain why we call the extension naive.

2.4 Definition (Naive extension of [BC97] – correctness part)
Protocol $[\mathtt{A},\mathtt{B}]$ is correct for Two Party Computation of the function $f(x, y)$ if

(a) for all $x \in \mathcal{X}, y \in \mathcal{Y}$

$$[\mathtt{A},\mathtt{B}](x)(y) \approx f(x, y); \tag{2.4}$$

(b) for any program $\tilde{\mathtt{A}}$ there exists a program $\tilde{\mathtt{A}}'$ such that for all $x \in \mathcal{X}, y \in \mathcal{Y}$

$$[\tilde{\mathtt{A}},\mathtt{B}]_{\mathsf{B}}(x)(y)\big|\,\mathtt{B}\text{ accepts} \approx [\mathtt{A},\mathtt{B}]_{\mathsf{B}}(\tilde{\mathtt{A}}'(x))(y)\big|\,\mathtt{B}\text{ accepts} \tag{2.5}$$

(c) for any program $\tilde{\mathtt{B}}$ there exists a program $\tilde{\mathtt{B}}'$ such that for all $x \in \mathcal{X}, y \in \mathcal{Y}$

$$[\mathtt{A},\tilde{\mathtt{B}}]_{\mathsf{A}}(x)(y)\big|\,\mathtt{A}\text{ accepts} \approx [\mathtt{A},\mathtt{B}]_{\mathsf{A}}(x)(\tilde{\mathtt{B}}'(y))\big|\,\mathtt{A}\text{ accepts} \tag{2.6}$$

Condition (2.4) expresses that, when both participants are honest, the protocol indeed performs the task it has been designed for. Condition (2.5) expresses that the best cheating A can do in order to influence B's outcome is always equivalent to her calculating an alternative input $\tilde{\mathtt{A}}'(x)$ and otherwise being honest during protocol execution, something which she can do anyhow. We assume that $\tilde{\mathtt{A}}'$ has access to $\tilde{\mathtt{A}}$. By the way, observe that the right hand side can also be replaced by $f_{\mathsf{B}}(\tilde{\mathtt{A}}'(x), y)$ because of the first condition. Condition (2.6) is similar to (2.5) with the roles of A and B reversed.

The preceding definition constituted the correctness part. For expressing privacy we use the Shannon *entropy* function, denoted $\mathbf{H}$ (see [CT91] for a definition, for instance).

2.5 Definition (Naive extension of [BC97] – privacy part)
Protocol $[\mathtt{A}, \mathtt{B}]$ *is private for Two Party Computation of the function* $f(x, y)$ *if*

(a) *for any joint stochastic variable* $\langle X, Y\rangle$ *over* $\mathcal{X}\times\mathcal{Y}$ *and for any program* $\tilde{\mathtt{A}}$ *there exists a stochastic variable* $\tilde{X}$ *over* $\mathcal{X}$ *such that*

$$\mathbf{H}\big(Y\big|\,[\tilde{\mathtt{A}}, \mathtt{B}]^*_{\mathsf{A}}(X)(Y), f_{\mathsf{A}}(\tilde{X}, Y)\big) \approx \mathbf{H}\big(Y\big|\, f_{\mathsf{A}}(\tilde{X}, Y)\big) \tag{2.7}$$

(b) *for any joint stochastic variable* $\langle X, Y\rangle$ *over* $\mathcal{X}\times\mathcal{Y}$ *and for any program* $\tilde{\mathtt{B}}$ *there exists a stochastic variable* $\tilde{Y}$ *over* $\mathcal{Y}$ *such that*

$$\mathbf{H}\big(X\big|\,[\mathtt{A}, \tilde{\mathtt{B}}]^*_{\mathsf{B}}(X)(Y), f_{\mathsf{B}}(X, \tilde{Y})\big) \approx \mathbf{H}\big(X\big|\, f_{\mathsf{B}}(X, \tilde{Y})\big) \tag{2.8}$$

The right hand side of condition (2.7) denotes A's uncertainty about Y, given one run of the ideal protocol to which she can supply an alternative input $\tilde{X}$. The left hand side of (2.7) reflects that in the real protocol A potentially gains additional information through her view of the protocol. Nevertheless, the equality expresses that the additional information A gains about B's input through her view is negligible, so the real protocol is as private as the ideal one. Using mutual information ($\mathbf{I}$) instead of entropy ($\mathbf{H}$) the last two conditions can be written more succinctly.

Expressed in this way, however, the comparison between the ideal protocol and the real implementation becomes more evident.

Typically the stochastic variable $\tilde{X}$ depends on $\tilde{\mathtt{A}}$ and on X (which includes dependence on auxiliary input). In this definition only the mathematical existence of $\tilde{X}$ is required, it need not be computable. So with this information-theoretic definition of privacy protocols can exist which are private against an all-knowing and infinitely powerful entity, but not against a party with limited resources.

In the model presented in Section 2.2 the role of $\tilde{X}$ (providing the ideal protocol with a suitable input) is taken care of by the interface. Its other responsibility, providing a proper simulation of the protocol $[\mathtt{A}, \mathtt{B}]$ to $\tilde{\mathtt{B}}$ (corresponding to the adversary), is implied by equation (2.7). The fact that A gains no information about Y from her view $f_{\mathsf{A}}(\tilde{X}, Y)$ means that in a way she 'knew' it already, so it can be written as a function: $[\tilde{\mathtt{A}}, \mathtt{B}]^*_{\mathsf{A}}(x)(y) = V(f_{\mathsf{A}}(\tilde{x}, y))$. As with $\tilde{X}$, the function V is not necessarily efficiently computable.

The most significant differences between Definitions 2.4 and 2.5 together, contrasted to Definition 2.3 presented earlier, are the following:

(a) Here correctness and privacy are defined separately, and the same problem with the Counterexample occurs. It is argued in [BCS96] that in the special case of (Generalized) Oblivious Transfer this is not a problem because A receives no output.

(b) The order of the quantifiers is reversed. This issue was already discussed in the previous section.

(c) The computational powers of the interfaces differ: here $\tilde{X}$ and V can be any function, while in Definition 2.3 they have been combined in one algorithm $\mathtt{S}$ which is restricted to probabilistic polynomial time.

2.5.2 Modifying the definition; awareness

Now we will modify Definitions 2.4 and 2.5, so that the result resembles Definition 2.3 more closely. The following definition reverses the order of the quantifiers, and requires that A' used for correctness and privacy be identical. But it does not resolve the issue about V being possibly non-computable, so the equivalence of the two approaches cannot be established.

2.6 Definition *Protocol $[\mathtt{A},\mathtt{B}]$ is* <u>*secure*</u> *for Two Party Computation of the function $f(x,y)$ if*

(a) for all $x \in \mathcal{X}$ and for all $y \in \mathcal{Y}$

$$[\mathtt{A},\mathtt{B}](x)(y) \approx f(x,y); \tag{2.9}$$

(b) there exists a program $\tilde{\mathtt{A}}'$ such that for any program $\tilde{\mathtt{A}}$, for all joint stochastic variable $\langle X,Y\rangle$ over $\mathcal{X}\times\mathcal{Y}$, and for all $x \in \mathcal{X}, y \in \mathcal{Y}$

$$[\tilde{\mathtt{A}},\mathtt{B}](x)(y)\big|\,\mathtt{B}\text{ accepts} \approx f(\tilde{\mathtt{A}}'(x),y) \tag{2.10}$$

and

$$\mathbf{I}\big(Y;[\tilde{\mathtt{A}},\mathtt{B}]^*_{\mathtt{A}}(X)(Y)\big|\,f(\tilde{\mathtt{A}}'(X),Y)\big) \approx 0 \tag{2.11}$$

(c) as (b), with the roles of A *and* B *reversed*

This definition resembles definition (3.10) of [MR91]. There Micali and Rogaway explicitly require that the cheating party can compute the alternative input; they call this *input awareness*. The function $\tilde{\mathtt{A}}'$ is the adversary input function, for which they insist that the same function be used in both the privacy and the correctness part. Indeed, it is easy to see that the Counterexample no longer passes the definition, because the $\tilde{\mathtt{B}}$ that would pass the privacy condition (the reverse of (2.11)) cannot simultaneously satisfy the correctness condition (the reverse of (2.10)).

The last definition also captures the notion of *output awareness*, through a small modification that could easily escape the reader's attention. The original correctness condition (2.5) expressed that only the honest party (B in this case) gets the same output (it is subscripted with B). The new correctness condition (2.10) requires however that all parties, *including* the cheating party, get the same output.

The crucial point, observed in [MR91], is that *even if a party is corrupted*, we have to worry about its output being well-defined as a result of the protocol. In the ideal protocol, a cheating party gets the output from the trusted host; this should also be reflected in the real protocol. Moreover, for a solitary protocol we may not have to care about the output of a cheating party. But if we want to prove a theorem about protocol composition, we need the outputs of all parties to be well-defined. Otherwise, in the case where the input of a protocols depends on the output of a preceding protocol, the adversary runs in trouble.

This additional requirement is captured by output awareness: the adversary should be able to compute a cheating party's output as a function of all the messages it sees. *Awareness* stands for both input awareness and output awareness. Definition 2.4 covers input awareness, but careful inspection reveals that a cheating $\tilde{\mathsf{B}}$, for instance, does not necessarily learn what output he obtains from the protocol. The same objection applies to the earlier definitions for security of (Generalized) Oblivious Transfer, so in this respect these definitions are inadequate.

2.6 Conclusion

In the last section we have seen how subtle and complicated the definition based on information theory has become. We had to realize that correctness and privacy must be coupled, and the notion of (input and output) awareness had to be introduced. On the other hand, observe that these two constraints are natural derivatives of Definition 2.3.

Note added in proof: *After this thesis was handed in, Joe Kilian pointed out that there is an additional complication with definition 2.4, which also applies to its* *earlier versions. Suppose that there exists an* $\tilde{\mathtt{A}}$ *who always tries to cheat, and that she succeeds doing so with exponential small probability. If she fails* B *always rejects. In this case equation 2.5 will not be fulfilled, even though the protocol 'should' pass the definition. It is very difficult to argue that such an* $\tilde{\mathtt{A}}$ *does* not *exist, so the definition rejects many protocols that may be considered secure.*

Indeed, the pleasing aspect of Beaver's approach is its intuitive appeal: we may have to worry about the exact relation between interface and adversary, but otherwise it seems quite convincing that all security aspects are incorporated. Moreover, composition theorems are relatively easy to prove. It is for this reason that we abandon the information-theoretic approach here, and adopt the interface (simulator) approach in this thesis.

Chapter 3

Distinguishability of quantum states

> [...] the 'imitation game.' It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman.
>
> ALAN TURING

This chapter surveys four measures of distinguishability of quantum states: probability of error, trace norm (related to the Kolmogorov distance), overlap (related to the Bhattacharyya coefficient) and Shannon distinguishability (defined through the mutual information).

These measures have their roots in measures of distinguishability of probability distributions, from which a collection of inequalities can be obtained. These inequalities can be shown to hold also in the quantum model. We obtain upper and lower bounds for the Shannon distinguishability. This is important because explicit computation of the Shannon distinguishability is often difficult, while the other distinguishability measures are usually easier to evaluate. We derive a very simple upper bound on the Shannon distinguishability as a function of the trace norm.

Moreover, cryptography often uses the notion of exponential indistinguishability of two probability distributions. This notion expresses that the Kolmogorov distance between two probability distributions vanishes exponentially fast when the security parameter increases. Similarly, exponential indistinguishability of two families of quantum states can be defined: we require that the distinguishability between two families of quantum states vanishes exponentially fast. It turns out that exponential indistinguishability is independent of the specific measure chosen in the definition; the four different definitions are equivalent.

3.1 Introduction

If one wants to model and analyze the cryptographic security of quantum protocols, one of the first questions to be answered is: What does it mean for two quantum states to be "close" to each other or to be "far" apart? That is, we need a notion of *distance* between two quantum states. The state of a quantum system is usually described by means of a *density matrix*, but various definitions of distance are available, and preferences may depend on the problem at hand and the point of view taken.

However, key properties that we would like them to satisfy are *(1)* that they are symmetric, *(2)* that they are invariant under unitary transformation (in other words, independent of the choice of the basis), and *(3)* that they somehow correspond to a conceivable physical experiment. Because of the latter constraint it is more appropriate to speak of measures of *distinguishability*.

In this chapter we will discuss four notions of distinguishability that are of particular interest to cryptography: probability of error, trace norm, overlap and Shannon distinguishability. Each of these four distinguishability measures is a generalization of a distinguishability measure between two (classical) probability distributions.

This also allows us to find the interrelations (inequalities) between these quantum distinguishability measures.

We establish a connection between the probability of error (or the Kolmogorov distance; the two are intimately connected) and the trace norm which apparently has not been noticed before. We derive a simple upper bound on the Shannon distinguishability as a function of the trace norm: $SD(\rho_0, \rho_1) \leq \frac{1}{2}\mathrm{Tr}|\rho_0 - \rho_1|$. Here ρ_0 and ρ_1 are density matrices, representing the states of two quantum systems. This bound may be very useful; for instance, in [BBBGM98] it is used to prove security of Quantum Key Distribution for a general class of attacks. Similarly, we can bound the Shannon distinguishability by the overlap between ρ_0 and ρ_1.

In the design of cryptographic protocols, we often actually define a *family* of protocols parameterized by a *security parameter*, n (denoting for example the length of some string, the number of rounds, or the number of photons). Typically we require that the probability of cheating for each participant vanishes exponentially fast, i.e. is of the order $O(\varepsilon^n)$, for ε between 0 and 1. In Definition 2.1 this plays an important role as well. It is therefore natural to consider two families of quantum states parameterized by n and to require that the distinguishability between the two families (defined by the pairwise distinguishability) vanishes exponentially fast. A priori this definition will depend on which distinguishability measure is chosen, but it is easy to show, as will be done, that exponential indistinguishability with respect to one distinguishability measure implies exponential indistinguishability with respect to each of the other distinguishability measures. In other words, these four notions are equivalent, and therefore it seems legitimate to talk of *the* notion of exponential indistinguishability for two families of quantum states.

This notion could be useful when defining the security of quantum protocols; indeed, this was initially the main motivation for this research. For instance, in [BMS96] it is proven that the Shannon distinguishability of the parity (i.e. the overall exclusive-or) of a quantum-bit-string decreases exponentially with the length of the string. This statement as such can be proven very easily using the techniques presented here, as will be shown in Section 8, without the need to keep track of the exact description of the density matrices involved (these tend to become quite complicated because of the tensor products involved).

The contribution of this chapter is three-fold. In the first place, even though some parts of this chapter are based on facts that have been known for some time, many of these facts are scattered throughout the literature, sometimes in papers that use an unusual terminology making it difficult to see that they are relevant. Elements of this work can already be found at various places in the work of Fuchs (see for instance [FU95]), but to our knowledge no paper presents a systematic overview of the distinguishability measures for quantum states from a cryptographer's point of view. In the second place, the bounds (inequalities) presented in Section 6 are useful, especially those that bound the Shannon distinguishability. The third contribution is the definition of *exponential indistinguishability* for quantum states, as explained two paragraphs earlier.

The structure of the chapter is as follows. In the following section we discuss probability distributions, mainly to introduce the setting and the notation to be used later. Section 3 discusses density matrices and measurements, showing how they lead to probability distributions as well. In Section 4 we discuss four measures of distinguishability, first classically, then quantumly. After a short summary in Section 5, we discuss several inequalities, again both classically and quantumly. In Section 7 these inequalities are applied to proving a theorem about exponential indistinguishability. Section 8 discusses an application.

3.2 Probability distributions

Let X_0 be a stochastic variable over the finite set $\mathcal{X}$, then we can define $p_0(x) \stackrel{\text{def}}{=} \text{Prob}[X_0 = x]$, so X_0 induces a probability distribution p_0 over $\mathcal{X}$. Let p_1 be defined likewise. Of course, $\sum_{x\in\mathcal{X}} p_t(x) = 1$ for t=$0, 1$. After relabeling the outcomes $x_1, x_2, x_3, \ldots x_m$ to $1, 2, 3, \ldots m$ we get:

		$x{=}1$	$x{=}2$	$x{=}3$	$\ldots$	$x{=}m$
X_0	$\pi_0 = \frac{1}{2}$	$p_0(1)$	$p_0(2)$	$p_0(3)$	$\ldots$	$p_0(m)$
X_1	$\pi_1 = \frac{1}{2}$	$p_1(1)$	$p_1(2)$	$p_1(3)$	$\ldots$	$p_1(m)$

Here π_0 and π_1 are the *a priori* probabilities of the two stochastic variables; they sum up to 1. Throughout this chapter $\pi_0 = \pi_1 = \frac{1}{2}$. Even though much of what will be said could be extended to the case $\pi_0 \neq \pi_1 \neq \frac{1}{2}$, it seems not relevant to the questions addressed here. Two distributions are *equivalent* (i.e. *perfectly indistinguishable*) if $p_0(x) = p_1(x)$ for all $x \in \mathcal{X}$, and they are *orthogonal* (i.e. *maximally distinguishable*) if there exists no x for which both $p_0(x)$ and $p_1(x)$ are non-zero.

Observe that $p_t(x)$ denotes the conditional probability that $X{=}x$ given that $T{=}t$, written as $\mathrm{Prob}[X{=}x|T{=}t]$. So the joint probability is half that value: $\mathrm{Prob}[X{=}x \wedge T{=}t] = \mathrm{Prob}[T{=}t]\ \mathrm{Prob}[X{=}x|T{=}t] = \pi_t p_t(x) = \frac{1}{2}p_t(x)$.

The following definitions and simple observations will be useful later on. We define the conditional probability $r_t(x) := \mathrm{Prob}[T{=}t|X{=}x]$, and the probability that $X{=}x$ *regardless* of t, that is, $p(x) := \mathrm{Prob}[X{=}x]$. Using Bayes' Theorem we get:

$$\begin{aligned} r_t(x) &= \mathrm{Prob}[T{=}t|X{=}x] \\ &= \mathrm{Prob}[T{=}t]\ \mathrm{Prob}[X{=}x|T{=}t]/\mathrm{Prob}[X{=}x] \\ &= \tfrac{1}{2}p_t(x)/p(x) \end{aligned}$$

Observe that $r_0(x) + r_1(x) = 1$ for all x. Using $p(x)$ and $r_t(x)$ we can represent the situation also in the following way:

		$x{=}1$	$x{=}2$	$x{=}3$	$\ldots$	$x{=}m$
X		$p(1)$	$p(2)$	$p(3)$	$\ldots$	$p(m)$
X_0	$\pi_0 = \frac{1}{2}$	$r_0(1)$	$r_0(2)$	$r_0(3)$	$\ldots$	$r_0(m)$
X_1	$\pi_1 = \frac{1}{2}$	$r_1(1)$	$r_1(2)$	$r_1(3)$	$\ldots$	$r_1(m)$

3.3 Density matrices and measurements

Recall that a quantum state is said to be in a *pure* state if there exists a measurement that can confirm this fact with probability 1. Pure states can be represented by vectors in an N-dimensional Hilbert space, i.e. a vector space with inner product. In this chapter N is finite. Now consider the following preparation: A flips a fair coin, and depending on the outcome sends one of two different pure states $|\psi_0\rangle$ or $|\psi_1\rangle$ to B. Then the "pureness" of the quantum states is "diluted" by the classical uncertainty about the resulting coin flip. In this case no deterministic measurement exists, and the quantum state is said to be in a *mixed* state.

How do we deal with mixed states mathematically? Instead of using a vector $|\psi\rangle$ to represent a pure state, we use the operator which projects onto the subspace spanned by $|\psi\rangle$. In bra-ket notation this operator is written as $|\psi\rangle\langle\psi|$. For a mixed state we construct a new operator by taking the sum of the projection operators, weighted according to their probabilities. The result is called a *density matrix* or *density operator*. It reflects B's knowledge about the state of the system, that is, his best prediction about any measurement results.

EXAMPLE: Consider the case where A prepares either a horizontally or a vertically polarized photon. We choose a basis such that $|\mathsf{H}\rangle = \binom{1}{0}$ and $|\mathsf{V}\rangle = \binom{0}{1}$. Then A's preparation is perceived by B as the mixed state

$$\tfrac{1}{2}|\mathsf{H}\rangle\langle\mathsf{H}| + \tfrac{1}{2}|\mathsf{V}\rangle\langle\mathsf{V}| = \tfrac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \tfrac{1}{2}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}, \quad (3.1)$$

which is the "completely mixed state".

Note that the same density matrix will be obtained if A prepares an equal mixture of left-polarized and right-polarized photons. In fact, in this two-dimensional example any equal mixture of two orthogonal pure states will yield the same density matrix.

Density matrices are always Hermitian i.e. $\rho = \rho^\dagger$. (The adjoint $\rho^\dagger$ of the matrix ρ is obtained by transposing ρ and taking the complex conjugate of all its entries.) This ensures that all the eigenvalues are real. To ensure that they are positive we require that ρ be positive, semi-definite (i.e. $\langle\psi|\rho|\psi\rangle \geq 0$ for all $\psi \in \mathcal{H}$). Finally we require that the eigenvalues (i.e. the trace of ρ) sum up to 1. These conditions make a probability interpretation of equation (3.3) below meaningful.

3.1 Definition (see for instance [SU86, IS95, PE93])
A density matrix ρ is an $N \times N$ matrix which has trace 1, which is Hermitian and which is positive, semi-definite.

Any source of quantum samples (that is, an imaginary Alice who indefinitely prepares quantum states in a prescribed way) is called an *ensemble*, which can be viewed as the quantum counterpart of a stochastic variable. A density matrix describes B's knowledge about the quantum state of a sample completely; two different ensembles (i.e. two different preparations) with the same density matrices are indistinguishable as far as B is concerned. There exists no measurement that will allow B to distinguish between these two ensembles.

However, the fact that a density matrix describes the a priori knowledge of B implies that additional classical information can change the density matrix, even when no measurement is performed and the state remains untouched. Two typical cases are *(1)* when A reveals information about the way she prepared the state(s) to B, or *(2)* when A and B share entanglement (EPR particles, for instance), and A sends the results of the measurements she performed to B. Observe that, consequently, a density matrix is subjective in the sense that it depends on what B knows.

EXAMPLE (continued): *(1)* Suppose that, after A has sent an equal mixture of $|\mathsf{H}\rangle$ and $|\mathsf{V}\rangle$, she reveals to B that for that particular sample she prepared $|\mathsf{V}\rangle$. Then B's density matrix suddenly changes from

$$\begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \quad \text{to} \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{3.2}$$

(2) An identical change happens in the following situation: A prepares an EPR pair whose combined state is written as

$$|\Psi^-\rangle = \tfrac{1}{\sqrt{2}}(|\mathsf{H}\rangle|\mathsf{V}\rangle - |\mathsf{V}\rangle|\mathsf{H}\rangle),$$

known as the singlet state, and sends one of the particles to B. If A and B measure both particles using the same basis, their measurement outcomes will always be anti-correlated. So if A after measuring finds that her particle is

> horizontally polarized ($|\mathsf{H}\rangle$) and if she tells this to B, then B's density matrix changes according to (3.2).

As an aside, it is interesting to remark that people disagree about whether the density matrix should be regarded as *the* state of the system or not, depending on the interpretation of quantum mechanics they adhere to. Consider for instance the situation before B has received the additional classical information to be sent by A. What is the state of his system? A pragmatist will answer that the state is simply described by B's density matrix, whereas a realist will argue that the state actually is different (maybe known by A or by God) and that B is merely ignorant of this. For discussion on this topic see for instance [Is95, Me96]. Here we will leave this philosophical question whether this is *really* the state or not unanswered, but adhere to the pragmatic approach, which seems more relevant from an information-theoretical point of view.

Now let us describe how to compute the probability of a certain measurement result from a density matrix. Mathematically speaking a density matrix ρ can be regarded as an object to which we can apply another operator E_x. Taking the trace of the product yields the probability that the measurement result is x given that the state was ρ, i.e. $\mathrm{Prob}[\mathrm{result}{=}x|\mathrm{state}{=}\rho] = \mathrm{Tr}(\rho E_x)$. Here the x serves as a label, connecting the operator E_x and the outcome x, but it has no specific, physical meaning.

This formula explains the name 'density matrix' or 'density operator': one has to apply another operator (corresponding to the measurement) and take the trace in order to obtain a density function (wave function), which is the equivalent of a probability distribution in the continuous case. Note that 'matrix' and 'operator' are freely interchanged, which is justified as long as it is understood that all operators are linear.

A *measurement* is the application of not just one operator, but of a collection (ordered set) of operators.

3.2 Definition (see [PE93])

*Let $\mathcal{E} = \langle E_1, \ldots, E_m \rangle$ be a collection (ordered set) of operators such that (1) all E_x are positive semi-definite operators, and (2) $\sum_x E_x = Id$. Such a collection specifies a Positive Operator-Valued Measure (*POVM*).*

Applying a POVM *to a system whose state is described by a density matrix ρ results in a probability distribution according to:*

$$\mathrm{Prob}[\mathrm{result} = x | \mathrm{state} = \rho] = \mathrm{Tr}(\rho E_x) \tag{3.3}$$

where x ranges from 1 to m.

As an alternative name for a POVM one also sees Probability Operator Measure. It is a postulate of quantum mechanics that any measurement can be described by a POVM, though it is not always clear that each POVM can indeed be experimentally realized. The set of all possible measurements, i.e. POVMs, is denoted by $\mathcal{M}$.

WARNING: The way measurements are defined in this chapter is not standard; indeed, many physicists are not familiar with this approach. In the usual approach, as developed by von Neumann, a measurement is defined as an Hermitian operator, and the eigenvalues of this operator correspond to the measurement results. However, both approaches can be proven equivalent using the spectral decomposition theorem [IS95].

A Projection Valued Measurement (PVM) (corresponding to what is usually called a von Neumann measurement) is a special case of a POVM: if we add the requirement that $E_x E_y = \delta(x, y) E_x$ (with $\delta(x, y)$=1 if x=y and 0 otherwise—i.e. the Kronecker-delta), then the operators E_x are necessarily projections. Note that for POVMs the E_x do not necessarily commute. Though PVM's are better known, they sometimes yield less information than POVMs. (On the other hand it has been shown that each POVM is equivalent to a PVM in a larger Hilbert space by adding a system in a known state, see [PE93] for discussion.)

EXAMPLE: Measuring whether a photon is polarized according to angle α or to $\alpha + \pi/2$ is done by the POVM

$$\left\{ \begin{pmatrix} c^2 & cs \\ cs & s^2 \end{pmatrix}, \begin{pmatrix} s^2 & -cs \\ -cs & c^2 \end{pmatrix} \right\},$$

where $c = \cos\alpha$ and $s = \sin\alpha$. This is a PVM. When applied to a photon known to be in state $|\mathsf{H}\rangle$, for instance, this results in the probability distribution $\langle c^2, s^2\rangle$, using equation (3.3).

An example of a POVM which is not a PVM is the symmetric three-outcome 'trine' POVM

$$\left\{ \tfrac{2}{3}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \tfrac{2}{3}\begin{pmatrix} \gamma^2 & \gamma\sigma \\ \gamma\sigma & \sigma^2 \end{pmatrix}, \tfrac{2}{3}\begin{pmatrix} \gamma^2 & -\gamma\sigma \\ -\gamma\sigma & \sigma^2 \end{pmatrix} \right\},$$

where $\gamma = \cos(\pi/3) = -\cos(2\pi/3) = \frac{1}{2}$ and $\sigma = \sin(\pi/3) = \sin(2\pi/3) = \frac{1}{2}\sqrt{3}$. This simplifies to

$$\left\{ \begin{pmatrix} \frac{2}{3} & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{6} & \frac{1}{6}\sqrt{3} \\ \frac{1}{6}\sqrt{3} & \frac{1}{2} \end{pmatrix}, \begin{pmatrix} \frac{1}{6} & -\frac{1}{6}\sqrt{3} \\ -\frac{1}{6}\sqrt{3} & \frac{1}{2} \end{pmatrix} \right\}.$$

Applying this POVM to the state $|\mathsf{V}\rangle$ results in the probability distribution $\langle 0, \frac{1}{2}, \frac{1}{2}\rangle$, again according to (3.3).

We have just seen that a measurement (a POVM) applied to a density matrix results in a probability distribution. Suppose now we have two density matrices defined over the same Hilbert space. Then we find ourselves back in the (classical) situation described in the previous section: comparing two probability distributions over the same outcome space X. In particular, let ρ_0, ρ_1 be two density matrices, and let $\mathcal{E} = \{E_1, \ldots, E_m\}$ again denote a POVM. Let $p_0(\mathcal{E})$ denote the probability distribution obtained by performing the POVM $\mathcal{E}$ on a system in state ρ_0 according to equation (3.3); let $p_1(\mathcal{E})$ be defined likewise. Then we have:

		$x{=}1$	$x{=}2$	$x{=}3$	$\ldots$	$x{=}m$
$p_0(\mathcal{E})$	π_0	$\mathrm{Tr}(\rho_0 E_1)$	$\mathrm{Tr}(\rho_0 E_2)$	$\mathrm{Tr}(\rho_0 E_3)$	$\ldots$	$\mathrm{Tr}(\rho_0 E_m)$
$p_1(\mathcal{E})$	π_1	$\mathrm{Tr}(\rho_1 E_1)$	$\mathrm{Tr}(\rho_1 E_2)$	$\mathrm{Tr}(\rho_1 E_3)$	$\ldots$	$\mathrm{Tr}(\rho_1 E_m)$

As before, π_0 and π_1 denote the a priori probabilities and are assumed to equal $\frac{1}{2}$.

3.4 Measures of distinguishability

This section discusses four different definitions of distinguishability, for probability distributions and for density matrices. Note that being able to distinguish between distributions, that is, between alternative hypotheses, is an important and well-studied problem with a vast literature. It goes under the name of classification, discrimination or feature evaluation, and has applications like speech recognition and radar detection. For a general overview, consult [BE82]. The problem studied here is a special case, in the sense that we want to distinguish between only two discrete probability distributions which have equal a priori probabilities. However, for the quantum case these probability distributions are determined by the choice of the measurement, thus optimization over all POVMs is required. In the following subsections the classical measure of distinguishability is discussed first, followed by its quantum counterpart.

3.4.1 Probability of error

Consider the following experimental situation in which B is asked to distinguish between two stochastic variables, as follows. First A provides him with one sample, x, chosen from either X_0 or X_1 with equal probability. Then B's task is to guess which of the two stochastic variables the sample came from, X_0 or X_1. Clearly, the average probability that B makes the right guess serves as a measure of distinguishability between the two probability distributions p_0 and p_1 induced by X_0 and X_1 respectively.

It is well known that B's optimal strategy is to look at the *a posteriori* probabilities: given the sample x his best choice is the t for which $r_t(x)$ is maximal (see the representation at the end of Section 3.2). This strategy is known as *Bayes' strategy*. So the average probability of successfully identifying the distribution now equals $\sum_{x\in X} p(x) \max\{r_0(x), r_1(x)\} = \frac{1}{2}\sum_{x\in X} \max\{p_0(x), p_1(x)\}$. Conversely, we can also express the probability that B fails.

3.3 Definition *The* <u>*probability of error between two probability distributions*</u> *is defined by*

$$\mathtt{PE}(p_0, p_1) \stackrel{\text{def}}{=} \tfrac{1}{2} \sum_{x \in X} \min\{p_0(x), p_1(x)\} \tag{3.4}$$

Two identical distributions have $\mathtt{PE} = \frac{1}{2}$*, and two orthogonal distributions have* $\mathtt{PE} = 0$.

WARNING: PE does not act as a distance: when two distributions are close to one another, PE is *not* close to 0, but close to $\frac{1}{2}$.

In the quantum model, the experimental set-up is almost identical. Again A has two ensembles, one according to ρ_0, the other one according to ρ_1. She provides B with a quantum sample chosen from one of the two ensembles with equal probability. Again B must guess from which ensemble the sample came: ρ_0 or ρ_1. And again, the Bayesian strategy of optimizing the a posteriori probabilities is optimal. However, in the quantum model B has an extra degree of freedom: he can choose the measurement he applies to his sample. Clearly, B's probability of successfully distinguishing ρ_0 from ρ_1 depends on the measurement he performs; he would like to choose the measurement that minimizes his probability of error. So we define:

3.4 Definition *The* <u>*probability of error between two density matrices*</u> ρ_0 *and* ρ_1 *is defined by*

$$PE(\rho_0, \rho_1) \stackrel{\text{def}}{=} \min_{\mathcal{E} \in \mathcal{M}} \mathtt{PE}(p_0(\mathcal{E}), p_1(\mathcal{E})), \tag{3.5}$$

where the POVM $\mathcal{E}$ *ranges over the set of all possible measurements* $\mathcal{M}$.

The question of finding an explicit formula for the optimal POVM in this definition was first studied by Helstrom [HE76, pp. 106–108]. He shows that the POVM $\mathcal{E}^*$ that *minimizes* $\mathtt{PE}(p_0(\mathcal{E}), p_1(\mathcal{E}))$ is actually a PVM. Knowing the optimal POVM, the probability of error can be expressed explicitly. The expression he gives is,

$$PE(\rho_0, \rho_1) = \tfrac{1}{2} + \tfrac{1}{2} \sum_{\lambda_j \leq 0} \lambda_j \ , \tag{3.6}$$

where the λ_j denote the eigenvalues of the matrix $\Gamma = \rho_0 - \rho_1$.

This expression can be cleaned up a little in the following way. Consider the function $f(x) = \frac{1}{2}(x - |x|)$. It vanishes when $x \geq 0$ and is the identity function otherwise. Thus, with its use, we can expand the summation in Eq. (3.6) to be over all the eigenvalues of Γ:

$$\begin{aligned} \mathtt{PE}(\rho_0, \rho_1) &= \tfrac{1}{2} + \tfrac{1}{2}\sum_{j=1}^{N} f(\lambda_j) \\ &= \tfrac{1}{2} + \tfrac{1}{4}\mathrm{Tr}\Gamma - \tfrac{1}{4}\sum_{j=1}^{N} |\lambda_j| \\ &= \tfrac{1}{2} - \tfrac{1}{4}\mathrm{Tr}|\Gamma| \ . \end{aligned}$$

Hence we have the following proposition:

3.5 Proposition *Given two arbitrary density matrices ρ_0 and ρ_1, the probability of error equals*

$$\mathtt{PE}(\rho_0, \rho_1) = \tfrac{1}{2} - \tfrac{1}{4}\sum_{j=1}^{N} |\lambda_j| = \tfrac{1}{2} - \tfrac{1}{4}\mathrm{Tr}|\rho_0 - \rho_1|, \tag{3.7}$$

where the λ_j are the eigenvalues of $\rho_0 - \rho_1$.

$\mathtt{PE}(\rho_0, \rho_1)$ is therefore just a simple function of the distance between ρ_0 and ρ_1, when measured as the trace norm of their difference.

3.4.2 Kolmogorov distance

Among cryptographers another measure of distinguishability between probability distributions is often used.

3.6 Definition *The Kolmogorov distance between two probability distributions is defined by*

$$\mathtt{K}(p_0, p_1) \stackrel{\mathrm{def}}{=} \tfrac{1}{2}\sum_{x \in X} |p_0(x) - p_1(x)|. \tag{3.8}$$

Two identical distributions have $\mathtt{K}=0$, and two orthogonal distributions have $\mathtt{K}=1$.

In some references the factor $\frac{1}{2}$ plays no role, and it is often left out, yet in other references an extra factor $\frac{1}{2}$ is included. Here we left it in because we wanted a value between 0 and 1.

Probability of error and Kolmogorov distance are closely related.

3.7 Proposition

$$\mathtt{PE}(p_0, p_1) = \tfrac{1}{2} - \tfrac{1}{2}\mathtt{K}(p_0, p_1) \tag{3.9}$$

This fact is not very difficult to prove, see [To72]. The most important step is to split the sum over X in two disjoint sub-sums, one for which $p_0(x) < p_1(x)$, and one for which $p_0(x) \geq p_1(x)$.

In the quantum scenario we will optimize again over all possible measurements. But here this means that we have to find the POVM that maximizes the Kolmogorov distance.

3.8 Definition *The Kolmogorov distance between two density matrices ρ_0 and ρ_1 is defined by*

$$K(\rho_0, \rho_1) \stackrel{\text{def}}{=} \max_{\mathcal{E}\in\mathcal{M}} \mathtt{K}(p_0(\mathcal{E}), p_1(\mathcal{E})), \tag{3.10}$$

where the POVM $\mathcal{E}$ *ranges over the set of all possible measurements* $\mathcal{M}$.

The relation between probability of error and Kolmogorov distance (eq. 3.9) shows that the two measurements that optimize $\mathtt{PE}$ and $\mathtt{K}$ are necessarily identical: the measurement $\mathcal{E}^*$ minimizes $\mathtt{PE}(p_0(\mathcal{E}), p_1(\mathcal{E})) = \frac{1}{2} - \frac{1}{2}\mathtt{K}(p_0(\mathcal{E}), p_1(\mathcal{E}))$ if and only if $\mathcal{E}^*$ maximizes $\mathtt{K}(p_0(\mathcal{E}), p_1(\mathcal{E}))$. See also the appendix of [Fu96]. Combining equations (3.7) and (3.9) we get:

3.9 Proposition *The Kolmogorov distance between two density matrices ρ_0 and ρ_1 equals*

$$K(\rho_0, \rho_1) = \tfrac{1}{2}\mathrm{Tr}|\rho_0 - \rho_1|, \tag{3.11}$$

where the λ_j are the eigenvalues of $\rho_0 - \rho_1$.

Observe that $\mathrm{Tr}|\rho_0 - \rho_1|$ is known as the *trace norm* on operators [RS72, OR96], so K has the additional property of satisfying the triangle inequality. The trace norm appears to be of unique significance within the class of all norms because of its connection to probability of error.

3.4.3 Bhattacharyya coefficient

Another distinguishability measure was defined in the 1940s because in some (continuous) cases the exact probability of error is difficult to evaluate exactly. See [KA67, TO71, TO72].

3.10 Definition *The Bhattacharyya coefficient between two probability distributions p_0 and p_1 is defined by*

$$\mathtt{B}(p_0, p_1) = \sum_{x \in X} \sqrt{p_0(x) p_1(x)} \tag{3.12}$$

Two identical distributions have $\mathtt{B} = 1$, whereas two orthogonal distributions have $\mathtt{B} = 0$.

WARNING: $\mathtt{B}$ does not act as a distance: when two distributions are close to one another, $\mathtt{B}$ is *not* close to 0, but close to 1. In some papers this is remedied by taking $\arccos(\mathtt{B})$ or $-\ln(\mathtt{B})$.

No simple operational definition for the Bhattacharyya coefficient seems to exist. However, geometrically it is related to the angle between p_0 and p_1, interpreted as vectors in the m-dimensional real space. A similar relation exists in the quantum model, as we will see shortly.

For the quantum model, we can now define a distinguishability measure by minimizing over all measurements possible.

3.11 Definition *The Bhattacharyya coefficient between two density matrices ρ_0 and ρ_1 is defined by*

$$B(\rho_0, \rho_1) \stackrel{\text{def}}{=} \min_{\mathcal{E} \in \mathcal{M}} \mathtt{B}(p_0(\mathcal{E}), p_1(\mathcal{E})), \tag{3.13}$$

where the POVM $\mathcal{E}$ *ranges over the set of all possible measurements* $\mathcal{M}$*.*

Surprisingly, it turns out that B is equivalent to another measure, studied by Wootters [Wo81] and Jozsa [Jo94], among others. Suppose $|\psi_0\rangle$ and $|\psi_1\rangle$ are pure states. When we think of these two state vectors geometrically, a natural notion of distinguishability is the angle between $|\psi_0\rangle$ and $|\psi_1\rangle$, or any simple function of this angle. In particular, using the formula $|\langle\psi_0|\psi_1\rangle| = |\psi_0||\psi_1|\cos\alpha$ and recalling that state vectors are usually normalized, we can define $overlap(|\psi_0\rangle, |\psi_1\rangle) := |\langle\psi_0|\psi_1\rangle|$ as a measure of distinguishability. The question is: what to do for mixed states?

The answer was given by Uhlmann [Uh76] and simplified by Jozsa [Jo94]. If ρ_0 is the density matrix of a mixed state in the Hilbert space $\mathcal{H}_1$, then we can always extend the Hilbert space such that ρ_0 becomes a pure state in the combined Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$. More precisely, we can always find an extension $\mathcal{H}_2$ of $\mathcal{H}_1$ and a pure state $|\varphi_0\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$, such that $\mathrm{Tr}_2(|\varphi_0\rangle\langle\varphi_0|) = \rho_0$. Here the operator Tr_2 means to trace out $\mathcal{H}_2$, the Hilbert space that was added, as explained in Subsection 4.3.3. When this condition holds, $|\varphi_0\rangle$ is said to be a *purification* of ρ_0. Similarly, if $|\varphi_1\rangle$ is the purification of ρ_1, we are back to a situation with two pure states, and we can apply the formula above, leading to the following definition.

3.12 Definition *The* <u>*(generalized) overlap*</u> *between two density matrices is defined by*

$$overlap(\rho_0, \rho_1) \stackrel{\text{def}}{=} \max|\langle\varphi_0|\varphi_1\rangle|, \tag{3.14}$$

where the maximum is taken over all purifications $|\varphi_0\rangle$ *and* $|\varphi_1\rangle$ *of* ρ_0 *and* ρ_1 *respectively.*

It can be shown [Uh76] [Jo94] that

$$overlap(\rho_0, \rho_1) = \mathrm{Tr}\left(\sqrt{\sqrt{\rho_0}\rho_1\sqrt{\rho_0}}\right), \tag{3.15}$$

but for dimensions higher than 2 this formula is difficult to evaluate, unless the eigenvalues of ρ_0 and ρ_1 are known. However, *overlap*(ρ_0, ρ_1) has the nice property that it is multiplicative over tensor products:

$$overlap(\rho_0 \otimes \rho_1, \rho_2 \otimes \rho_3) = overlap(\rho_0, \rho_2)\ overlap(\rho_1, \rho_3). \tag{3.16}$$

The following proposition establishes the connection between the Bhattacharyya coefficient and the overlap.

3.13 Proposition (Fuchs and Caves [FC95])
Overlap is equivalent to the quantum Bhattacharyya coefficient:

$$overlap(\rho_0, \rho_1) = \mathrm{B}(\rho_0, \rho_1). \tag{3.17}$$

Apart from the geometric interpretation of the overlap mentioned above, there also exists a physical interpretation for the *square* of this value. The following discussion is based on Jozsa's work.

Consider again the situation in which A provides B with a quantum state according to ρ_0 or ρ_1. Let us first suppose that both states are pure, so $\rho_0 = |\psi_0\rangle\langle\psi_0|$ and $\rho_1 = |\psi_1\rangle\langle\psi_1|$. Define $F(\rho_0, \rho_1) = (overlap(\rho_0, \rho_1))^2$. There is an obvious physical interpretation of this quantity: it is the probability that a particle in state $|\psi_1\rangle$ passes the test: "Are you yes/no in state $|\psi_0\rangle$?" Von Neumann called this the transition probability. In a communication model, if we see $|\psi_0\rangle$ as the input of a channel and $|\psi_1\rangle$ as the output, we could call this the fidelity.

This interpretation can be extended to the case where ρ_1 is mixed but ρ_0 is still pure: $F(\rho_0, \rho_1) = \langle\psi_0|\rho_1|\psi_0\rangle$ is the probability that a particle in state ρ_1 passes the test: "Are you yes/no in state ψ_0?" For the most general case, where ρ_0 and ρ_1 are both mixed, no physical interpretation seems to exist; nevertheless, F remains well-defined in this case and it is natural to continue to call it 'fidelity'.

3.4.4 Shannon distinguishability

Now we come to the last, and maybe most important notion of distinguishability. Mutual information, as defined by Shannon [SH48], can also be used as a distinguishability measure between probability distributions [Li56, Be82]. Recall that the (Shannon) entropy function $\mathbf{H}$ is defined as

$$\mathbf{H}(X) \stackrel{\text{def}}{=} -\sum_{x\in\mathcal{X}} \mathrm{Prob}[X{=}x] \lg \mathrm{Prob}[X{=}x], \tag{3.18}$$

where $\lg$ denotes the logarithm function with base 2. The argument of $\mathbf{H}$ can be either a stochastic variable or a probability distribution. A special case is the function $\mathbf{H}_2(x)$, which is the entropy of the distribution $\langle x, 1-x\rangle$.

Consider the following elementary example. Suppose we have two boxes, each containing colored balls. Let $t \in \mathcal{T}{=}\{0,1\}$ denote a particular identity of one of the boxes, whereas T denotes the stochastic variable representing the identity of the box. Then $\mathrm{Prob}[T = t]$ is just what was called the a priori probability π_t in Section 3.2. Recall that in our case $\pi_0 = \pi_1 = \frac{1}{2}$, so $\mathbf{H}(T) = 1$. Similarly, let X denote the stochastic variable corresponding to the color of the balls, regardless of the box it came from. Recall that $\mathrm{Prob}[X{=}x]$ was written as $p(x)$.

Considering the same experiment as with the probability of error, in which A picks a ball from one of the two boxes and gives it to B, one can ask now: How much information does X (the color of a picked ball) convey about T (the identity of the box it came from)?

Information is defined as the reduction of uncertainty, where uncertainty is quantified using the Shannon entropy. Consider two quantities: *(1)* the average uncertainty of B about T before he was handed a sample (or ball); and *(2)* his average uncertainty about T after he was handed a sample. This difference expresses the amount of information gained through the experiment, and can also be used as a measure of distinguishability between two distributions. Thus we obtain:

$$\begin{aligned}
\text{average information} &= (\text{average uncertainty about } T) - \\
&\qquad (\text{average uncertainty about } T \text{ when } x \text{ is given}) \\
&= \mathbf{H}(T) - \mathbf{H}(T|X) \\
&= \mathbf{H}_2(\tfrac{1}{2}) - \sum_{x\in\mathcal{X}} p(x)\ \mathbf{H}(T|X = x)
\end{aligned}$$

$$\begin{aligned} &= 1 - \sum_{x \in \mathcal{X}} p(x)\ \mathbf{H}_2(r_0(x)) \\ &= I(T;X). \end{aligned}$$

This leads to the following definition:

3.14 Definition *The Shannon distinguishability between two probability distributions p_0 and p_1 is defined by:*

$$\mathtt{SD}(p_0, p_1) \stackrel{\text{def}}{=} \mathbf{I}(T;X). \tag{3.19}$$

Shannon distinguishability can also be applied in the quantum scenario: find the measurement that optimizes the Shannon distinguishability of the two distributions obtained when applying the same measurement to two different quantum states.

3.15 Definition *The Shannon distinguishability between two density matrices ρ_0 and ρ_1 is defined as*

$$\mathit{SD}(\rho_0, \rho_1) \stackrel{\text{def}}{=} \max_{\mathcal{E} \in \mathcal{M}} \mathtt{SD}(p_0(\mathcal{E}), p_1(\mathcal{E})), \tag{3.20}$$

where the POVM $\mathcal{E}$ *ranges over the set of all possible measurements $\mathcal{M}$.*

There is an unfortunate problem for this measure of distinguishability: calculating the value $\mathit{SD}(\rho_0, \rho_1)$ is generally a difficult problem. Apart from a few special cases, no explicit formula for SD solely in terms of ρ_0 and ρ_1 is known. Even stronger than that: no such formula can exist in the general case [FC94]. This follows from the fact that optimizing the Shannon distinguishability requires the solution of a transcendental equation. (See also [PE93] and [MO97] for a discussion of other aspects of SD.)

3.5 Overview

The material presented in the previous four subsections can be summarized in the following table.

	classical definition	$p_0 = p_1$	$p_0 \perp p_1$	optimization criterion	quantum expression
PE	$\frac{1}{2}\sum \min\{p_0(x), p_1(x)\}$	1/2	0	min	$\frac{1}{2} - \frac{1}{4}\mathrm{Tr}\lvert\rho_0 - \rho_1\rvert$
K	$\frac{1}{2}\sum \lvert p_0(x) - p_1(x)\rvert$	0	1	max	$\frac{1}{2}\mathrm{Tr}\lvert\rho_0 - \rho_1\rvert$
B	$\sum(p_0(x)p_1(x))^{1/2}$	1	0	min	*overlap*(ρ_0, ρ_1)
SD	*via* $I(X;T)$	0	1	max	does not exist

Table 3.1: Classical and quantum notions of indistinguishability.

3.6 Inequalities

We have seen already (equation (3.9)) that probability of error and Kolmogorov distance are related through the equality:

$$\mathtt{PE}(p_0, p_1) = \tfrac{1}{2} - \tfrac{1}{2}\mathtt{K}(p_0, p_1) \tag{3.21}$$

The other pairs of distinguishability measures are related through inequalities which can be found in various places in the literature [KA67, HR70, TO71, TO72, BE82].

3.16 Proposition *Let p_0 and p_1 be probability distributions. The following relations hold:*

$$1 - \mathtt{B}(p_0, p_1) \leq \mathtt{K}(p_0, p_1) \leq \sqrt{1 - \mathtt{B}(p_0, p_1)^2} \tag{3.22}$$

$$1 - \mathbf{H}_2(\mathtt{PE}(p_0, p_1)) \leq \mathtt{SD}(p_0, p_1) \leq 1 - 2\mathtt{PE}(p_0, p_1). \tag{3.23}$$

$$1 - \mathtt{B}(p_0, p_1) \leq \mathtt{SD}(p_0, p_1) \leq 1 - \mathbf{H}_2(\tfrac{1}{2} - \tfrac{1}{2}\sqrt{1 - \mathtt{B}(p_0, p_1)}) \tag{3.24}$$

Before giving a proof of this proposition, we would like to state its quantum equivalent.

3.17 Theorem *Proposition 3.16 can be generalized to the quantum scenario: one can substitute PE, K, B and SD and use density matrices ρ_0 and ρ_1 as operands.*

Alternatively, using the quantum expressions, equations (3.22), (3.23) and (3.24) can be expressed in the following, equivalent form:

$$1 - \mathit{overlap}(\rho_0, \rho_1) \le \tfrac{1}{2}\mathrm{Tr}|\rho_0 - \rho_1| \le \sqrt{1 - \mathit{overlap}(\rho_0, \rho_1)^2} \tag{3.25}$$

$$1 - \mathbf{H}_2(\tfrac{1}{2} - \tfrac{1}{4}\mathrm{Tr}|\rho_0 - \rho_1|) \le \mathtt{SD}(\rho_0, \rho_1) \le \tfrac{1}{2}\mathrm{Tr}|\rho_0 - \rho_1|. \tag{3.26}$$

$$1 - \mathit{overlap}(\rho_0, \rho_1) \le \mathtt{SD}(\rho_0, \rho_1) \le 1 - \mathbf{H}_2(\tfrac{1}{2} - \tfrac{1}{2}\sqrt{1 - \mathit{overlap}(\rho_0, \rho_1)}) \tag{3.27}$$

The importance of this theorem is that, while the Shannon distinguishability itself may sometimes be difficult to calculate exactly, the inequalities provide a way to bound it. We will use this in the next section.

PROOF OF PROPOSITION 3.16:
We will start with proving (3.22). We start with the left inequality:

$$\begin{aligned}
1 - \mathtt{B}(p_0, p_1) &= \tfrac{1}{2}\left(\sum_{x\in\mathcal{X}} p_0(x) + \sum_{x\in\mathcal{X}} p_1(x) - 2\sum_{x\in\mathcal{X}} \sqrt{p_0(x)p_1(x)}\right) \\
&= \tfrac{1}{2}\sum_{x\in\mathcal{X}} \left(|\sqrt{p_0(x)} - \sqrt{p_1(x)}|\right)^2 \\
&\le \tfrac{1}{2}\sum_{x\in\mathcal{X}} |p_0(x) - p_1(x)| \\
&= \mathtt{K}(p_0, p_1)
\end{aligned}$$

The inequality in the penultimate transition is valid for each term individually: $|\sqrt{a} - \sqrt{b}| \le \sqrt{a} + \sqrt{b}$ because of the triangle inequality. Taking the absolute value on the right does not change the expression; now multiply both sides by $|\sqrt{a} - \sqrt{b}|$.

We need the Schwarz inequality, $(\sum_{x\in\mathcal{X}} c(x)d(x))^2 \le (\sum_{x\in\mathcal{X}} c(x)^2)(\sum_{x\in\mathcal{X}} d(x)^2)$, for the right inequality of (3.22), where c and d are vectors in the m-dimensional real vector space. Define the coordinates as $c(x) = |\sqrt{p_0(x)} - \sqrt{p_1(x)}|$ and $d(x) = |\sqrt{p_0(x)} + \sqrt{p_1(x)}|$. Then we find that

$$\begin{aligned}
\mathtt{K}(p_0, p_1)^2 &= \tfrac{1}{4}\left(\sum_{x\in\mathcal{X}} |p_0(x) - p_1(x)|\right)^2 \\
&= \tfrac{1}{4}\left(\sum_{x\in\mathcal{X}} |\sqrt{p_0(x)} - \sqrt{p_1(x)}||\sqrt{p_0(x)} + \sqrt{p_1(x)}|\right)^2
\end{aligned}$$

$$
\begin{aligned}
&\leq \tfrac{1}{4}\left(\sum_{x\in\mathcal{X}}|\sqrt{p_0(x)}-\sqrt{p_1(x)}|\right)^2\left(\sum_{x\in\mathcal{X}}|\sqrt{p_0(x)}+\sqrt{p_1(x)}|\right)^2 \\
&= \tfrac{1}{4}\left(\sum_{x\in\mathcal{X}}p_0(x)+p_1(x)-2\sqrt{p_0(x)p_1(x)}\right)\left(\sum_{x\in\mathcal{X}}p_0(x)+p_1(x)+2\sqrt{p_0(x)p_1(x)}\right) \\
&= \tfrac{1}{4}(2-2\mathtt{B}(p_0,p_1))(2+2\mathtt{B}(p_0,p_1)) \\
&= 1-\mathtt{B}(p_0,p_1)^2
\end{aligned}
$$

In order to prove the left inequality in (3.23) we observe that this is a special case of Fano's inequality (see for instance [CT91]):

$$\mathbf{H}(T|X) \leq \mathbf{H}_2(\mathtt{PE}(p_0,p_1)) + \mathtt{PE}(p_0,p_1)\ \log(\#\mathcal{T}-1).$$

Recall from Subsection 3.4.4 that the cardinality of $\mathcal{T}$, denoted as $\#\mathcal{T}$, equals 2. One could also use Jensen's inequality.

For the right inequality of (3.23) we expand $\mathbf{I}(X;T)$ as $\mathbf{H}(T)-\mathbf{H}(T|X)$ to obtain an inequality between SD and PE (see also [HR70]). Recalling the definitions of $r_t(x)$ and $p(x)$, observing that $r_1(x)=1-r_0(x)$ and that $2\min\{r,1-r\}\leq\mathbf{H}_2(r)$ for all r between 0 and 1 (see Figure 3.1), we obtain:

$$
\begin{aligned}
\mathtt{SD}(p_0,p_1) &= 1-\sum_{x\in\mathcal{X}}p(x)\ \mathbf{H}_2(r_0(x)) \\
&\leq 1-\sum_{x\in\mathcal{X}}p(x)\cdot 2\min\{r_0(x),1-r_0(x)\} \\
&\leq 1-2\mathtt{PE}(p_0,p_1).
\end{aligned}
$$

The left hand inequality of (3.24) is obtained in a very similar way, using that $\mathbf{H}_2(r)\leq 2\sqrt{r(1-r)}$ (see Figure 3.1):

$$
\begin{aligned}
\mathtt{SD}(p_0,p_1) &= 1-\sum_{x\in\mathcal{X}}p(x)\ \mathbf{H}_2(r_0(x)) \\
&\geq 1-\sum_{x\in\mathcal{X}}p(x)\cdot 2\sqrt{r_0(x)(1-r_0(x))} \\
&= 1-\sum_{x\in\mathcal{X}}\sqrt{p_0(x)p_1(x)} \\
&= 1-\mathtt{B}(p_0,p_1).
\end{aligned}
$$

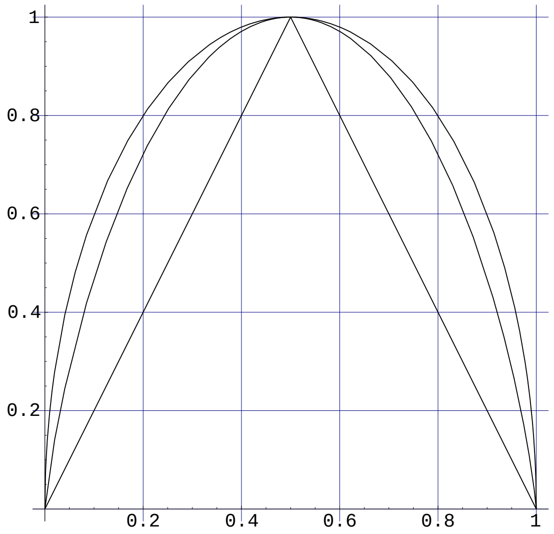

Figure 3.1: $2\min\{x, 1-x\} \leq \mathbf{H}_2(x) \leq 2(x(1-x))^{\frac{1}{2}}$.

(Formally this is proven by looking at the first and second derivatives.)

For the right hand side of (3.24) we define the function

$$g(r) := \tfrac{1}{2} - \tfrac{1}{2}\sqrt{1-r^2}. \tag{3.28}$$

If we let $h(r) = 2\sqrt{r(1-r)}$ for $0 \leq r \leq 1$, then we get that $h(r) = h(1-r)$, that $gh(r) = \min\{r, 1-r\}$, and that $hg(r) = r$. So $\mathbf{H}_2(r_0(x)) = \mathbf{H}_2(\min\{r_0(x), 1 - r_0(x)\}) = \mathbf{H}_2(gh(r_0(x)))$, and

$$\begin{aligned}
\mathtt{SD}(p_0, p_1) &= 1 - \sum_{x\in\mathcal{X}} p(x)\ \mathbf{H}_2(r_0(x)) \\
&= 1 - \sum_{x\in\mathcal{X}} p(x)\ \mathbf{H}_2(g(\ \ h(r_0(x))\ )) \\
&\leq 1 - \mathbf{H}_2(g(\sum_{x\in\mathcal{X}}\ p(x)h(r_0(x)))) \\
&= 1 - \mathbf{H}_2(g(\mathtt{B}(p_0, p_1))) \\
&= 1 - \mathbf{H}_2(\tfrac{1}{2} - \tfrac{1}{2}\sqrt{1 - (\mathtt{B}(p_0, p_1))^2})
\end{aligned}$$

Here we used Jensen's inequality on the (composite) function $\mathbf{H}_2(g(r))$; this function is convex. This concludes the proof of Proposition 3.16.$\square$

PROOF OF THEOREM 3.17:

First we prove equation (3.25). We start with the first inequality. Let $\mathcal{E}_B^*$ denote a

POVM which optimizes $\mathtt{B}$; define $\mathcal{E}_K^*$ likewise.

$$\begin{aligned} 1 - B(\rho_0, \rho_1) &= 1 - \mathtt{B}(p_0(\mathcal{E}_B^*), p_1(\mathcal{E}_B^*)) \\ &\leq \mathtt{K}(p_0(\mathcal{E}_B^*), p_1(\mathcal{E}_B^*)) \\ &\leq \mathtt{K}(p_0(\mathcal{E}_K^*), p_1(\mathcal{E}_K^*)) \\ &= K(\rho_0, \rho_1) \end{aligned}$$

The second inequality of (3.25) follows from

$$\begin{aligned} K(\rho_0, \rho_1) &= \mathtt{K}(p_0(\mathcal{E}_K^*), p_1(\mathcal{E}_K^*)) \\ &\leq \sqrt{1 - \mathtt{B}(p_0(\mathcal{E}_K^*), p_1(\mathcal{E}_K^*))^2} \\ &\leq \sqrt{1 - \mathtt{B}(p_0(\mathcal{E}_B^*), p_1(\mathcal{E}_B^*))^2} \\ &= \sqrt{1 - B(\rho_0, \rho_1)^2} \end{aligned}$$

Equations (3.26) and (3.27) are proven in a similar way. In particular, in (3.25) the function on the extreme left, $f(x) = 1 - x$, and on the extreme right, $f(x) = \sqrt{1 - x^2}$ are monotonously decreasing. In addition, $\mathtt{B}$ must be minimized whereas $\mathtt{K}$ must be maximized. The same is true for (3.26) and (3.27). □

3.7 Exponential indistinguishability

As already described in the Introduction, when designing cryptographic protocols, we often actually define a *family* of protocols parameterized by a *security parameter* n (for instance denoting the length of some string, the number of rounds, or the number of photons). Typically we require that the probability of cheating for each participant vanishes exponentially fast, i.e. is of the order $O(\varepsilon^{-n})$, with ε between 0 and 1. For instance, one technique is to compare the protocol implementation, which is usually a family of protocols parameterized by n, with an *ideal protocol specification*, and to prove that these two become exponentially indistinguishable [BE91B].

3.18 Definition *Let* $\{X_0\} = \langle X_0^{(1)}, X_0^{(2)}, X_0^{(3)} \ldots \rangle$ *denote a* *family* *of stochastic variables with corresponding distributions* $\langle p_0^{(1)}, p_0^{(2)}, p_0^{(3)} \ldots \rangle$ *. Let* $\{X_1\}$ *be defined similarly. Then* $\{X_0\}$ *and* $\{X_1\}$ *are* *exponentially indistinguishable* *if there exists an* n_0 *and an* ε *between* 0 *and* 1 *such that*

$$\forall n \geq n_0 : \mathtt{K}(p_0^{(n)}, p_1^{(n)}) \leq \varepsilon^n$$

Examples can be constructed easily. For instance, let X_0^n be uniformly distributed over $\{0,1\}^n$ (the set of strings of length n), i.e. for each $x \in \{0,1\}^n$ we have $p_0^{(n)}(x) = 2^{-n}$. This defines the family of uniform distributions $\{X_0\}$. We define $\{X_1\}$ identically, except that $p_1^{(n)}(0^n) = 0$, while $p_1^{(n)}(1^n) = 2^{-n+1}$. So 0^n, the word with all zeroes, has zero probability, and 1^n, the word with all ones, has double the probability. Clearly the two families $\{X_0\}$ and $\{X_1\}$ are exponentially indistinguishable.

The reader should be aware that in *computational* cryptography more refined notions of distinguishability have been defined [GMR89]. In *polynomial indistinguishability* it is only required that the families converge as fast as $1/n^k$, for any $k > 0$. Though we will not be arguing this formally, it is not hard to see that the proof of the Lemma below also applies for polynomially indistinguishable families. Yet another refinement is *computational indistinguishability*: the sample is given to a polynomial time Turing machine which outputs a 0 or 1. We look at the Kolmogorov distance of the possible outputs while maximizing over all polynomial time Turing Machines, and require that this converges as $1/n^k$ for any $k > 0$. Computational indistinguishability turned out to be extremely powerful for defining notions such as pseudo-random number generators [YA82B] and zero knowledge protocols [GMR89], but is of no concern here. All three notions have in common that a distinguisher who is restricted to polynomial time, and who is given a sample of either source, will not be able to tell the identity of the source.

Here we followed the computational-cryptographic tradition of defining exponential indistinguishability using the Kolmogorov distance. This choice, however, is in no way crucial, since the next lemma shows that we could have taken any of the four distinguishability measures. In other words $\mathtt{K}$, $\mathtt{PE}$, $\mathtt{B}$ and $\mathtt{SD}$ turn out to be equivalent when we require exponentially fast convergence. There is a small technicality here: indistinguishable distributions have $\mathtt{PE} = \frac{1}{2}$ and $\mathtt{B} = 1$, so exponential indistinguishability means convergence to those values, instead of convergence to 0, as is the case with $\mathtt{K}$ and $\mathtt{SD}$.

3.19 Lemma *Let $\{X_0\}$ and $\{X_1\}$ be two families of stochastic variables which are exponentially indistinguishable with respect to* one *of the distinguishability measures* $\mathtt{K}$, $\mathtt{PE}$, $\mathtt{B}$, $\mathtt{SD}$. *Then $\{X_0\}$ and $\{X_1\}$ are exponentially indistinguishable with respect to* each *of* $\mathtt{K}$, $\mathtt{PE}$, $\mathtt{B}$, $\mathtt{SD}$.

PROOF: The equivalence between exponential indistinguishability for $\mathtt{PE}$ and $\mathtt{K}$ follows from equation (3.9), whereas the other equivalences are derived using equations (3.22) through (3.24).

For instance, to prove that exponential indistinguishability for $\mathtt{K}$ implies exponential indistinguishability for $\mathtt{B}$ would go as follows. Suppose

$$[\exists n_0, \varepsilon][\forall n \geq n_0] \ : \ \mathtt{K}(p_0^{(n)}, p_1^{(n)}) \leq \varepsilon^n. \tag{3.29}$$

Using the left hand side of equation (3.22) it follows that $1 - \mathtt{B}(p_0^{(n)}, p_1^{(n)}) \leq \varepsilon^n$.

For the reverse direction: if $1 - \mathtt{B}(p_0^{(n)}, p_1^{(n)}) \leq \varepsilon^n$ then

$$1 - (\mathtt{B}(p_0^{(n)}, p_1^{(n)}))^2 \leq \varepsilon^n (1 + \mathtt{B}(p_0^{(n)}, p_1^{(n)})) \leq 2\varepsilon^n, \tag{3.30}$$

so $\mathtt{K}(p_0^{(n)}, p_1^{(n)}) \leq \sqrt{2}\sqrt{\varepsilon}^n$ using the right hand side of equation (3.22). If we choose $\bar{\varepsilon} < \sqrt{\varepsilon}$ and $\bar{n}_0 > n_0$ such that $\sqrt{2} < (\bar{\varepsilon}/\sqrt{\varepsilon})^{\bar{n}_0}$, then $\mathtt{K}(p_0^{(n)}, p_1^{(n)}) \leq \bar{\varepsilon}^n$ for $n \geq \bar{n}$.

The other implications are proven in a similar way. As far as $\mathbf{H}_2(x)$ is concerned, it is sufficient to recall (see fig. 3.1) that

$$2 \min\{x, 1-x\} \leq \mathbf{H}_2(x) \leq 2\sqrt{x(1-x)} \tag{3.31}$$

This concludes the proof. $\square$

The obvious next step is to define exponential indistinguishability for density matrices, and to show that the choice of the distinguishability measure is immaterial.

3.20 Definition *Let $\{\rho_0^{(n)}\} = \langle \rho_0^{(1)}, \rho_0^{(2)}, \rho_0^{(3)} \ldots \rangle$ denote a family of density matrices defined over the Hilbert spaces $\mathcal{H}^{(1)}, \mathcal{H}^{(2)}, \mathcal{H}^{(3)} \ldots$ Let $\{\rho_1^{(n)}\}$ be defined similarly. Then $\{\rho_0^{(n)}\}$ and $\{\rho_1^{(n)}\}$ are exponentially indistinguishable if there exists an n_0 and an ε between 0 and 1 such that*

$$\forall n \geq n_0 : K(\rho_0^{(n)}, \rho_1^{(n)}) \leq \varepsilon^n$$

An example will be presented in the next section. But first let us end this section with the quantum analogue of Lemma 3.19.

3.21 Theorem *Let $\{\rho_0^{(n)}\}$ and $\{\rho_1^{(n)}\}$ be two families of density matrices which are exponentially indistinguishable with respect to* one *of the distinguishability measures K, PE, B, SD. Then $\{\rho_0^{(n)}\}$ and $\{\rho_1^{(n)}\}$ are exponentially indistinguishable with respect to* each *of K, PE, B, SD.*

PROOF: This follows now immediately from the proof of Lemma 3.19 using Theorem 3.17. □

3.8 Applications

Let us now look at an application of the quantum-exponential indistinguishability idea. In particular, we look at the problem of the parity bit in quantum key distribution as studied in [BMS96] (henceforth called BMS). Let $|\psi_0\rangle = \binom{\cos\alpha}{\sin\alpha}$ and $|\psi_1\rangle = \binom{\cos\alpha}{-\sin\alpha}$, and let ρ_0 and ρ_1 be the corresponding density matrices. The bits 0 and 1 are represented by ρ_0 and ρ_1 Likewise, the bit string $z = z_1 z_2 \ldots z_n$ is represented by n different photons, the ith photon being in state ρ_i. Thus the combined state for the string z is given by

$$\rho_z = \rho_{z_1} \otimes \rho_{z_2} \otimes \cdots \otimes \rho_{z_n} \tag{3.32}$$

where $\otimes$ stands for the tensor product.

Now let $Z_0^{(n)}$ denote all the strings of length n with even parity (i.e. with overall exclusive-or equal to 0) and $Z_1^{(n)}$ all strings of length n with odd parity. Then define

$$\rho_j^{(n)} = \frac{1}{2^{n-1}} \sum_{z \in Z_j^{(n)}} \rho_z \tag{3.33}$$

for $j = 0, 1$. In BMS these two density matrices are explicitly calculated in order to compute their Shannon distinguishability as a function of n and α, for $0 \leq \alpha < \pi/4$. This is extremely important because the parity bit appears in the proof of security of the BB84 key exchange protocol [BB84].

Here we compute the distance between $\rho_0^{(n)}$ and $\rho_1^{(n)}$ in terms of Kolmogorov distance and overlap. And for the special case $n = 2$ we will study the inequalities obtained in Theorem 3.16, just as an illustration to see how good these bounds are. Observe that, at this point in time, the problem of the parity bit is one of the few non-trivial examples in which Shannon distinguishability, Kolmogorov distance (and related, probability of error) and overlap can be computed, yet the optimal POVMs are known to be different. For the next few paragraphs the reader is advised to consult BMS, or to take (3.34), (3.41) and (3.42) for granted.

First let us calculate the Kolmogorov distance $K(\rho_0^{(n)}, \rho_1^{(n)})$ as a function of n and α. In BMS it is shown that $\Delta^{(n)} \stackrel{\text{def}}{=} \frac{1}{2}(\rho_0^{(n)} - \rho_1^{(n)})$ has non-zero entries on the secondary diagonal only. Moreover, it is not difficult to see that all these entries equal $c^n s^n$, where $c = \cos\alpha, s = \sin\alpha$. Therefore $\Delta^{(n)}$ has 2^{n-1} eigenvalues equal to $-c^n s^n$, and 2^{n-1} eigenvalues equal to $+c^n s^n$, so

$$K(\rho_0^{(n)}, \rho_1^{(n)}) = \sum |\lambda_j| = (2cs)^n = (\sin 2\alpha)^n = S^n, \tag{3.34}$$

where $S = \sin 2\alpha$. Obviously, $\{\rho_0^{(n)}\}$ and $\{\rho_1^{(n)}\}$ are exponentially indistinguishable.

Computing the overlap between $\rho_0^{(n)}$ and $\rho_1^{(n)}$ is more elaborate. For the discussion in this section we will denote the overlap as B and define it through equation (3.15):

$$B(\rho_0, \rho_1) := \mathrm{Tr}\left(\sqrt{\sqrt{\rho_0}\rho_1\sqrt{\rho_0}}\right) \tag{3.35}$$

This is justified because of Proposition 3.13. In BMS equation (19) it is shown that $\rho_0^{(n)}$ and $\rho_1^{(n)}$ are block-diagonal, with 2×2 blocks. If σ is of the form

$$\begin{pmatrix} [\sigma^u] & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & [\sigma^l] \end{pmatrix} \tag{3.36}$$

we will write this as $\sigma = \sigma^u \oplus \sigma^l$. It is not difficult to see that the following three equalities now hold:

$$\begin{aligned} \mathrm{Tr}(\sigma^u \oplus \sigma^l) &= \mathrm{Tr}(\sigma^u) + \mathrm{Tr}(\sigma^l) \\ (\sigma_0^u \oplus \sigma_0^l)(\sigma_1^u \oplus \sigma_1^l) &= \sigma_0^u \sigma_1^u \oplus \sigma_0^u \sigma_1^u \\ \sqrt{(\sigma^u \oplus \sigma^l)} &= \sqrt{(\sigma^u)} \oplus \sqrt{(\sigma^l)} \end{aligned}$$

From this it follows that

$$\mathrm{Tr}\left(\sqrt{\sqrt{\sigma_0}\sigma_1\sqrt{\sigma_0}}\right) = \mathrm{Tr}\left(\sqrt{\sqrt{\sigma_0^u}\sigma_1^u\sqrt{\sigma_0^u}}\right) + \mathrm{Tr}\left(\sqrt{\sqrt{\sigma_0^l}\sigma_1^l\sqrt{\sigma_0^l}}\right) \tag{3.37}$$

which we rewrite as

$$B(\sigma_0, \sigma_1) = B(\sigma_0^u, \sigma_1^u) + B(\sigma_0^l, \sigma_1^l). \tag{3.38}$$

This means that we can evaluate B by evaluating each block individually according to (3.35) and sum up the results. Note that the operands on the right are not normalized, i.e. their trace does not equal 1.

Again according to BMS eq. (20) (but slightly changing notation) each block is of the form

$$\sigma_0^{(n,k)} = \begin{pmatrix} u^2 & uv \\ uv & v^2 \end{pmatrix} \text{ for even parity; } \sigma_1^{(n,k)} = \begin{pmatrix} u^2 & -uv \\ -uv & v^2 \end{pmatrix} \text{ for odd parity,} \tag{3.39}$$

with $u = c^{n-k}s^k, v = c^k s^{n-k}$. So in fact these density matrices represent pure states, $\binom{u}{v}$ and $\binom{u}{-v}$, and the formula for B reduces to taking the inner product. We find that

$$B(\sigma_0^{(n,k)}, \sigma_1^{(n,k)}) = c^{2(n-k)}s^{2k} - c^{2k}s^{2(n-k)}. \tag{3.40}$$

Summing up over all blocks of $\rho_i^{(n)}$ we get

$$\begin{aligned} B(\rho_0^{(n)}, \rho_1^{(n)}) &= \sum_{k=0}^{\lfloor (n-1)/2 \rfloor} \binom{n}{k} B(\sigma_0^{(n,k)}, \sigma_1^{(n,k)}) \\ &= \sum_{k=0}^{(n-1)/2} \binom{n}{k} (c^{2(n-k)} s^{2k} - c^{2k} s^{2(n-k)}), \end{aligned}$$

for n odd. For n even we find the same expression: even though for $k = n/2$ there are only $\frac{1}{2}\binom{n}{k}$ blocks, their contribution is zero since $B(\sigma_0^{(n,k)}, \sigma_1^{(n,k)}) = 0$ in this case. For the case $n = 2$ this expression reduces to

$$B(\rho_0^{(2)}, \rho_1^{(2)}) = c^4 - s^4 = (c^2 - s^2)(c^2 + s^2) = C, \tag{3.41}$$

where $C = \cos 2\alpha$.

For the Shannon distinguishability in the special case $n = 2$ BMS eq. (44) gives

$$SD(\rho_0^{(2)}, \rho_1^{(2)}) = \frac{1}{2}(1 + C^2)\mathbf{H}_2\left(1 - \frac{C^2}{1 + C^2}\right) + \frac{S^2}{2}. \tag{3.42}$$

Now we are in a position to substitute (3.34), (3.41) and (3.42) into equations (3.25), (3.26) and (3.27). Observe that (3.25) is easily seen to hold, in fact with equality holding on the right hand side. Equations (3.26) and (3.27) are illustrated in Figures 3.2 and 3.3. The horizontal axis represents the angle α between $|\psi_i\rangle$ and $\binom{1}{0}$, meaning that for $\pi/4$ (≈ 0.785) the states $|\psi_0\rangle$ and $|\psi_1\rangle$ are orthogonal. The fact that the bounds using overlap seem tighter might be explained by the fact that the function $2\sqrt{x(1-x)}$ resembles $\mathbf{H}_2(x)$ fairly well.

3.9 Conclusion

We have presented four measures of distinguishability for density matrices, shown their relation to classical measures of distinguishability for probability distributions, thus obtaining several inequalities and the generalized notion of exponential indistinguishability. The latter will be useful in the next chapter. The preceding presentation somewhat sidesteps the issue of the exact measurements involved, partially to make this work more accessible to non-physicists. Interesting questions are, for instance, under which conditions the measurements that optimize PE, B and SD are identical, or under what conditions equality holds in the inequalities presented. Such questions are still subject of ongoing research.

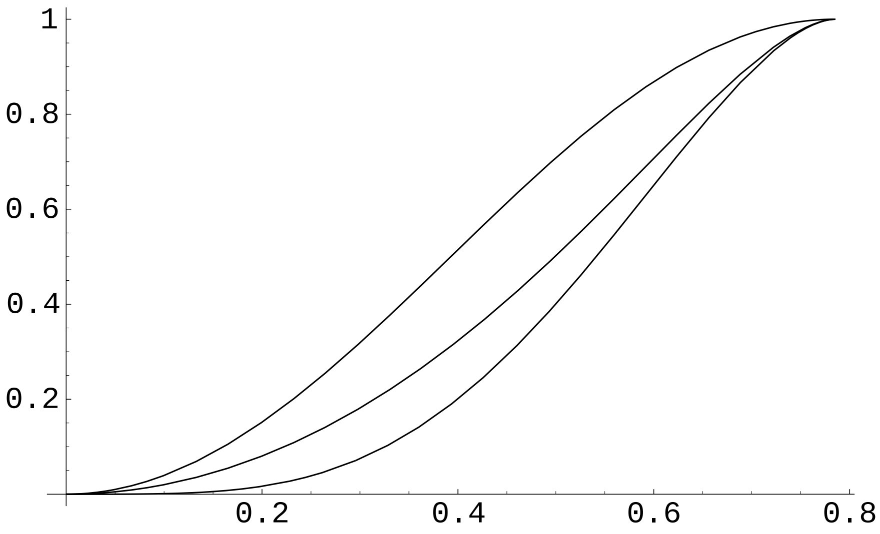

Figure 3.2: Equation 3.26 for the parity bit with $n = 2$ and with $\alpha \in [0, \pi/4]$ on the horizontal axis.

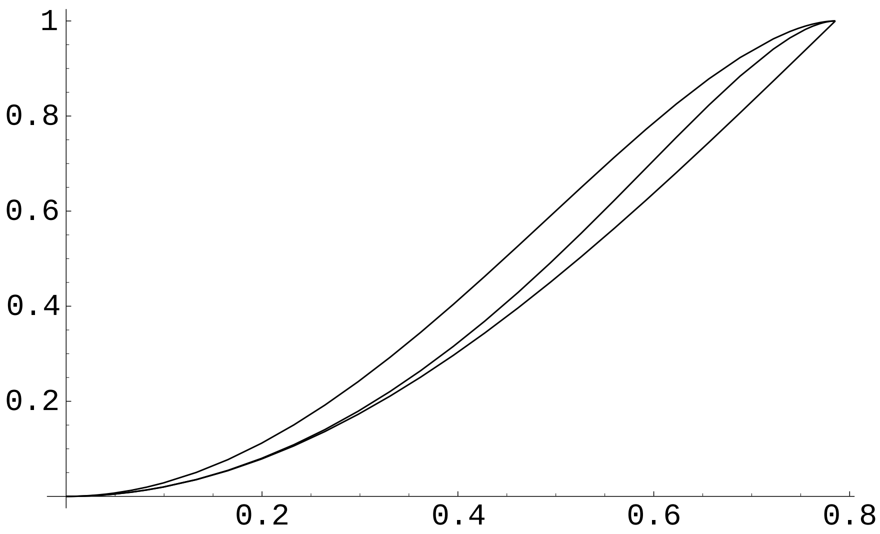

Figure 3.3: Equation 3.27 for the parity bit with $n = 2$ and with $\alpha \in [0, \pi/4]$ on the horizontal axis.

Chapter 4

A model for quantum protocols

Quantum particles: the dreams that stuff is made of.

DAVID MOSER

In this final chapter we present our model for two party quantum protocols, and a definition of security. We restrict ourselves to protocol tasks with classical input and output only, and describe how two quantum computers can execute a protocol using a quantum channel.

We discuss in detail the differences between classical and quantum information: the fact that quantum information cannot be copied and the physical principle that prevents us from extracting any classical information from a quantum system without disturbing its state. And despite the fact that classical information can be seen as a special case of quantum information, we show that from a cryptographic point of view the two types of information must be kept separated.

For the formal model we use Quantum Turing Machines for the parties and describe the quantum channel by means of a quantum communication tape. We define the output, view and profile of a quantum protocol, and use a quantum interface (simulator) to compare the security of two protocols.

After the formal definition of security we present strong evidence that the classical technique known as rewinding is impossible in quantum protocols.

4.1 Introduction

In Chapter 2 we have presented a classical definition of security, whereas in Chapter 3 we defined exponential indistinguishability for quantum states. In this chapter we extend the definition of security to quantum protocols, i.e. to protocols whose task specification remains classical, but in which the participants communicate through a quantum channel. The idea is simple; we take the definition as presented in Section 2.2 and see what happens when we replace the classical Turing Machines by two Quantum Turing Machines who communicate through a quantum tape. Chapter 3 provides us with the right notion of exponential indistinguishability.

Despite the fact that there are many papers on quantum cryptography, no formal definition has been presented so far. A protocol model was presented by Yao [YA95] and was adapted by Mayers [MA97] and Lo and Chau [LC97A, LC97B]. However, the explanations and justification for the underlying assumption was either brief or imprecise, and sometimes difficult to comprehend for someone without a proper physical background.

Here we examine the model from the point of view of a cryptographer with a strong interest, but not necessarily a strong background in quantum physics. In particular we give the intuition behind some of the quantum phenomena, and try to determine the physical principles responsible for them. Therefore, the first part of this chapter will be somewhat informal.

First we introduce briefly qubits and quantum circuits. In Section 3 we discuss in detail how a quantum computer can emulate classical computation, including measurements. Subsequently, in Section 4, we describe informally how two parties can communicate through a quantum channel. There we also discuss the fundamental difference between classical and quantum information, and the implication this has for the view of the adversary. In Section 5 we present the formal model, based on Quantum Turing Machines, and the definition of security. Section 6 discusses rewinding of the adversary.

4.2 Qubits and unitary transformations

The qubit is the basic unit of quantum information, as the bit is for classical information. In quantum mechanics the state of a (discrete) system is represented by a unit vector in an N-dimensional Hilbert space over the field of complex numbers, and a state thus representable is called a *pure* state. In this formalism it is natural to represent the mutually exclusive bits 0 and 1 as orthogonal vectors $|0\rangle$ and $|1\rangle$, which together span a 2-dimensional Hilbert space. The *superposition principle* states that any state $|\psi\rangle$ which is a linear superposition of $|0\rangle$ and $|1\rangle$ and which has norm 1 is itself also a pure state. Consequently, a *qubit* is a variable which can attain any value

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \tag{4.1}$$

with $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 = 1$. We say that a qubit is in state $|\psi\rangle$, while α and β are called the amplitudes.

For concreteness it may be useful to think of photon polarization as an implementation of a qubit, but most people do not think any more of bits in terms of On/Off, voltage levels or magnetization directions. Moreover, photons may not be the most convenient medium to store qubits; for this purpose ion traps are probably more practical. In any case, here we will not be concerned with implementation issues; nevertheless we will sometimes talk about a qubit as if it where a particle.

Any pure state of a collection of n qubits can be described by a vector in an N-dimensional Hilbert space $\mathcal{H}$, when $N = 2^n$. The N basis vectors can conveniently be labeled by the set of strings of length n; so each

$$|b_1 \ldots b_n\rangle = \bigotimes_{i=1}^{n} |b_i\rangle$$

represents a basis vector, where $b_i \in \{0,1\}$ and where $\otimes$ stands for tensor product. From the superposition principle it follows that a state like $|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ is also a pure state. There is no way we can write $|\Psi^-\rangle$ as a product of pure states of two individual qubits, i.e. as $(\alpha_1|0\rangle + \beta_1|1\rangle) \otimes (\alpha_2|0\rangle + \beta_2|1\rangle)$.

According to quantum mechanics any state evolution of a discrete system is described by a unitary transformation. A *unitary transformation* U is a linear operator on $\mathcal{H}$ which has an inverse and which preserves scalar products, i.e.

$$\langle U\psi | U\varphi\rangle = \langle \psi | \varphi\rangle \tag{4.2}$$

From these two conditions it can be shown that $UU^\dagger = U^\dagger U = Id$, where $U^\dagger$ is obtained by transposing the matrix representing U and taking the complex conjugate of all its entries. This condition can serve as an alternative characterization of a unitary transformation. Intuitively, a unitary operator is the analogue of a rotation in $\mathbb{R}^N$; in particular, it is equivalent to choosing another basis.

As an example, consider the following unitary transformation on two qubits, defined by the image of each basis vector under U:

$$\begin{aligned} |00\rangle &\to |00\rangle \\ |01\rangle &\to |01\rangle \\ |10\rangle &\to |11\rangle \\ |11\rangle &\to |10\rangle. \end{aligned}$$

This operation is called the XOR, because $|ab\rangle \longrightarrow |a\,(a \oplus b)\rangle$. Observe that U shows its 'xoring' behavior only if the input states correspond to the classical states $|0\rangle$ or $|1\rangle$, but otherwise any intuition is likely to fail. For instance, in the $\{|0\rangle, |1\rangle\}$ basis the role of $|a\rangle$ and $|b\rangle$ (i.e. the first and second qubits) is not symmetric: only $|b\rangle$ is affected whereas $|a\rangle$ is preserved. For this reason this operation is also called *controlled-not*: a controls whether b is inverted or not. But surprisingly, in the Hadamard basis $\{\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\}$ the roles of $|a\rangle$ and $|b\rangle$ are reversed. To see that XOR is indeed unitary, it is sufficient to write it as a matrix in the basis $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ and one obtains:

$$\mathsf{XOR} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{4.3}$$

The XOR is one of the simplest quantum operations, and is also referred to as an XOR *gate*; people use it as a building block to design quantum circuits. Several alternative gates have been defined; most notably the Toffoli gate T. This is 3-qubit gate, also called the *controlled-controlled-not* because

$$|abc\rangle \xrightarrow{\mathsf{T}} |ab\,((a \wedge b) \oplus c)\rangle.$$

So c is inverted only if a=b=1.

By using the 'constants' $|0\rangle$ and $|1\rangle$ as input values, a Toffoli gate T allows us to do various operations on one qubit, like Not, XOR or Fanout, and on two qubits, like And and Swap. We will give the Fanout explicitly, because it is of significant importance later on. Let d be either $|0\rangle$ or $|1\rangle$. Then $|d10\rangle \xrightarrow{\mathsf{T}} |d1d\rangle$. This Fanout works *only* when the input qubit $d \in \{|0\rangle, |1\rangle\}$, i.e. when the input qubit is known to represent a classical value. The Fanout does not work for an *arbitrary* qubit state $|\psi\rangle$, so there is no contradiction with the *no-cloning* theorem for quantum states, which we will explain in Section 4.4.3. For more details about gates, see for instance [St97].

4.3 Emulating classical computation and measurements by a quantum computer

Classical information and quantum information are quite different, and it seems that a general model of quantum computation has to deal with two types of information.

This, however, is not very convenient; it is much more desirable to have one type of information to deal with. In the next subsection we will show that any probabilistic computation can be emulated by a quantum computer.

In addition, we have to deal with measurements. It is a priori not clear what effects measurements during the computation have, whether a party gains or loses computational power by switching back and forth between classical and quantum information. Most textbooks about quantum mechanics present two possible kinds of transformations on the state of a quantum system: *(1)* unitary transformation, which are reversible, and *(2)* measurements (usually described as projection operators), which are not reversible. From this, one could easily come to the conclusion that measurements are irreversible, and therefore cannot be emulated by a (reversible) quantum computer. To show that a quantum computer *can* emulate any measurement is the subject of Subsection 4.3.2.

In this section we discuss emulating classical computation. The quantum computer is not yet explicitly defined, it can be thought of as a quantum circuit, a Quantum Turing Machine (to be introduced in Section 4.5.1), or any collection of unitary transformations on a register of qubits. We will return to emulating classical communication in a cryptographic protocol in Section 4.4.1. Most of the observations in this section are not new, but we believe it is worth discussing them thoroughly, in particular since we want to extend these type of arguments.

4.3.1 Deterministic and random bits

A party who wants to represent classical bits by means of qubits can simply use orthogonal states, say $|0\rangle$ and $|1\rangle$, to encode 0 and 1. We have seen this already before when we discussed quantum gates. Obviously, if a qubit is known to contain a classical bit and the basis is known, then the original value can always be retrieved by a measurement in this basis.

To obtain a random bit with a quantum computer a *quantum coin flip* is executed (also referred to as a role of a quantum die). One takes a qubit known to be in state $|0\rangle$ (or $|1\rangle$), and applies the unitary operator

$$\mathsf{QCF} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix}. \tag{4.4}$$

Then a measurement in the $\{|0\rangle, |1\rangle\}$ basis is performed, resulting in either 0 or 1 with equal probability.

In the next subsection we will see that one can emulate the measurement and postpone the actual measurement until the end. So the measurement result will be a qubit $|\mathsf{b}\rangle$ instead of a classical bit b. The quantum computer evaluates all possible execution paths in superposition until the end of the computation, when the measurement is performed. From all possible executions now one is 'chosen', according to the proper probability distribution.

Observe that this has the somewhat counterintuitive consequence that all randomness has disappeared from the computation itself: the state evolution is fully *deterministic*, the probabilistic aspect comes from the measurement at the end.

4.3.2 Measurements

Nobody seems to understand precisely what exactly happens during a measurement, and the subject is often debated upon. However, this is beside the point here. The crucial observation is that, in some sense, any measurement can be considered as being reversible (unitary). This is exactly the point of view we take here, because it shows us that a quantum computer can emulate any measurement.

In Chapter 3, where we used the POVM formalism to describe measurements, we put the boundary between the classical and the quantum world between the observed system and the measurement apparatus. In other words, the measurement apparatus was completely factored out in the mathematical formalism used. Here we do the opposite: we consider the measurement apparatus as part of the quantum system as well. Then any measurement can be reformulated as an interaction between

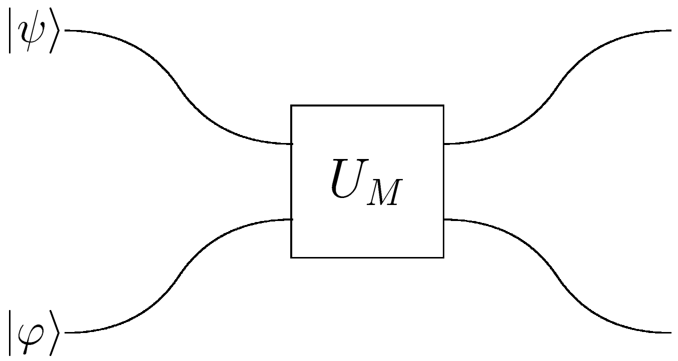

Figure 4.1: A measurement M of a system in state $|\psi\rangle$ interpreted as an interaction U_M with an auxiliary system in a known state $|\varphi\rangle$.

quantum system and measurement apparatus, described by a unitary transformation on both systems combined.

Let us give an example. Suppose that a qubit in state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ is measured according to the $\{|0\rangle, |1\rangle\}$ basis (see Figure 4.1). Let $|\varphi\rangle$ be the initial state of the measurement apparatus, and let U_M denote the unitary transformation corresponding to the measurement M, i.e. the unitary transformation that 'xors' information about the observed system (here the qubit) into the measurement apparatus. Then the state evolution of the 'measurement' can be described by

$$|\psi\rangle|\varphi\rangle \xrightarrow{U_M} \alpha|0\rangle|\varphi_0\rangle + \beta|1\rangle|\varphi_1\rangle. \tag{4.5}$$

How is the resulting state to be interpreted? Suppose that a 'real' measurement would occur, then we would argue as follows: if, after suitable amplification of its state, the measurement apparatus indicates that its state is $|\varphi_0\rangle$, then we would deduce that the qubit was in state $|0\rangle$. This happens with probability $|\alpha|^2$. And with probability $|\beta|^2$ the opposite happens: the measurement apparatus is found to be in state $|\varphi_1\rangle$, indicating that the qubit was in state $|1\rangle$. So we find that after the measurement interaction the states of the qubit and the measurement apparatus are correlated: $|\varphi'\rangle = |\varphi_b\rangle \Leftrightarrow |\psi'\rangle = |\mathsf{b}\rangle$. This correlation is called *entanglement*.

The previous description was just to explain how the state could be interpreted.

However, we are in no way forced to assume that a 'real' measurement actually has occurred. To put it differently, any quantum computer can emulate the evolution U_M described above: the measurement apparatus is simply regarded as a part of the quantum computer itself. And in this extended model the emulated measurement U_M can be undone (reversed): just apply U_M^{-1}. Note that in this section we are not interested in actually undoing the measurement, we only want to convince ourselves that the evolution is unitary. However, we will return to this point in Section 4.6 when discussing rewinding.

The preceding description of a measurement using entanglement was already given by Schrödinger in 1935 [Sc35, section 10], in the same paper in which he introduced his cat. This was several months after Einstein, Podolsky and Rosen's paper [EPR35]. See also [Sv97] for some helpful discussion.

Talking about EPR: it is interesting to elaborate on the previous discussion. The observation is the following: a quantum computer that emulates the measurement does not need the whole measurement apparatus to store $|\varphi_0\rangle$ or $|\varphi_1\rangle$. It suffices to take only one qubit since the subspace spanned by the (orthogonal) states $|\varphi_0\rangle$ and $|\varphi_1\rangle$ is of dimension two. So it becomes logical to write $|\varphi_b\rangle := |\mathsf{b}\rangle$. If we now set the initial state $|\varphi\rangle$ of this 'measurement' qubit to $|0\rangle$ we get

$$(\alpha|0\rangle + \beta|1\rangle)|0\rangle = \alpha|00\rangle + \beta|10\rangle \xrightarrow{U_M} \alpha|00\rangle + \beta|11\rangle, \tag{4.6}$$

which is exactly the evolution of the XOR. The resulting state is entangled, not unlike the state $|\Psi^-\rangle$ often used in EPR 'paradoxes'.

However, in the remainder of the computation both qubits play quite a different role. The 'measured' qubit continues to play a role in the quantum computation, and it can interact with other qubits. The 'measurement' qubit, on the other hand, is preciously kept isolated in memory until the computation has completed. Then it will be measured according to the same bases that would have been used if the measurement had been performed immediately, the $\{|0\rangle, |1\rangle\}$ basis. It is clear that this does *not* change the probability distribution of the measurement results. Consequently, during the emulation nothing is changed if parties do not perform their measurements immediately. Instead, they can keep the information in quantum memory and measure it (make it classical) at a later stage, without changing any probability distribution on the results. Note that a measurement in an arbitrary basis can be described by a unitary operation (implementing the change of basis) followed by a measurement in $\{|0\rangle, |1\rangle\}$ basis.

We have almost accomplished what we want: we can emulate deterministic and random bits and we can emulate each conceivable measurement. The only issue remaining is the if statement, where the condition is a measurement outcome, and the actions are unitary operators. Let $|\varphi'\rangle \in \mathcal{H}_1$ be the result of a measurement, and suppose we want to implement: if $|\varphi'\rangle = |0\rangle$ then U_0 else U_1, where U_0 and U_1 act on $\mathcal{H}_2$. It is not difficult to see that there exists a unitary transformation U_{if} on $\mathcal{H}_1 \otimes \mathcal{H}_2$ which has the same effect as this if statement.

The overall conclusion is that no computational power is lost or gained by emulating measurements. On the other hand, we obtain a uniform description using unitary transformations and quantum information (qubits) only, while all measurements are performed at the end.

4.3.3 Partial knowledge about entangled systems

The previous discussion about entanglement provides a good starting point to explain a mathematical technique, called tracing out.

Suppose A has a qubit in state $\alpha|0\rangle + \beta|1\rangle$, while B has a qubit in state $|0\rangle$. Suppose they jointly apply the XOR to their two qubits, then they separate again. The overall state of both qubits together evolves as in equation (4.6), resulting in $|\Phi\rangle = \alpha|00\rangle + \beta|11\rangle$. This describes perfectly the overall state of both qubits, but what do we know about the state of A's individual qubit? Or rather, what does A know about the state of her qubit?

It follows from the earlier discussion that with probability $|\alpha|^2$ A's qubit is in state

$|0\rangle$, and with probability $|\beta|^2$ in state $|1\rangle$; and the same is true for B. This was easy to see in this particular case. But in the general case, how do we calculate the partial knowledge a party has about an entangled system when not all the particles are in his hands?

The general strategy is to use density matrices (see Section 3.3) to describe the combined state, and then to *trace out*. The density matrix describing the state $|\Phi\rangle$ is

$$|\Phi\rangle\langle\Phi| = \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & b \end{pmatrix}, \tag{4.7}$$

where $a = |\alpha|^2$ and $b = |\beta|^2$. Labeling the rows and columns by the strings 00, 01, 10 and 11, each entry of this matrix can be represented by $d_{ij,kl}$, with $i, j, k, l \in \{0, 1\}$.

Let E be the matrix obtained by A after tracing out B's Hilbert space. The entries of E are defined by $e_{i,k} = \mathrm{Tr}(D_{i,k})$, where

$$D_{i,k} = \begin{pmatrix} d_{i0,k0} & d_{i1,k0} \\ d_{i0,k1} & d_{i1,k1} \end{pmatrix} \tag{4.8}$$

In the example

$$E = \mathrm{Tr}_{\mathsf{B}}|\Phi\rangle\langle\Phi| = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \tag{4.9}$$

exactly reflecting that with probability $|\alpha|^2$ A's qubit is in state $|0\rangle$, and with probability $|\beta|^2$ in state $|1\rangle$.

Similarly, we can trace out A's Hilbert space by interchanging the roles of i and k with the roles of j and l. In this particular example A and B obtain the same result after tracing out, but that need not be true generally. Note that some entries (like $d_{00,01}$) are lost in the process, so tracing out always implies a loss of information, unless the two qubits are not entangled. Tracing out is easily generalized to any tensor product of two vector spaces of finite dimension.

4.4 Quantum protocols informally

In this chapter our purpose is to describe a general model for quantum protocols and security. Before presenting a formal definition in the next section, we will discuss some important issues first in an informal way.

4.4.1 The quantum channel

We assume that A and B each have a quantum computer, and that they communicate through a quantum channel. What the latter precisely means is actually not very clear. For instance, we could say that they communicate by exchanging some quantum particles, but a particle can be in a superposition of being sent and not being sent. This is not so far-fetched as it may seem; it is possible to design quantum protocols using half-silvered mirrors, resulting in photons being in superposition of going through and being reflected. Nevertheless, it can be shown that this example as well can be modeled by using qubits.

We assume that A and B each have a private collection of qubits, on which they can do some unitary transformations, and that they share some quantum channel. This could be, for instance, a register of qubits which they modify in turns, or a collection of particles they send back and forth. In the QTM formalism it will be a communication tape they can both access.

Described more formally: A has her Hilbert space $\mathcal{H}^{\mathsf{A}}$, likewise B has his Hilbert space $\mathcal{H}^{\mathsf{B}}$, while their shared Hilbert space, the quantum channel, is denoted $\mathcal{H}^{\mathcal{C}}$. During step j, party P applies the transformation $U_j^{\mathsf{P}}|\psi\rangle$, with $|\psi\rangle \in \mathcal{H}^{\mathsf{P}} \otimes \mathcal{H}^{\mathcal{C}}$.

Suppose A wants to send quantum information to B. She applies the unitary transformation U on her qubits and the shared qubits, defined by $U^{\mathsf{A}}(|\psi^{\mathsf{A}}\rangle \otimes |\psi^{\mathcal{C}}\rangle)$. Generally this means that A's qubits will become entangled with those of the channel. Then it becomes B's turn: he applies a unitary transformation $U^{\mathsf{B}}(|\psi^{\mathcal{C}}\rangle \otimes |\psi^{\mathsf{B}}\rangle)$ to the shared qubits and his own, and the qubits of A, B and $\mathcal{C}$ end up being all entangled. Note that, in some intuitive sense, the unitary transformation applied by A and B correspond to reading *and* writing simultaneously.

This description seems sufficient to model any kind of quantum communication between A and B. However, is the model complete? One restriction is clear: we have restricted ourselves to *discrete* phenomena. This means that we cover only physical phenomena which, if measured, can give a finite number of (mutually exclusive) outcomes. Examples are: electron spin, photon polarization, energy levels of an atom. Continuous phenomena are excluded; an example would be electron position. Measurements here give a continuum of results, like "the electron is observed at position x". Defining a qubit in the context of a continuous phenomenon does not seem meaningful.

It is a postulate (i.e. an assumption about how reality can be described) of quantum mechanics that the pure state of every conceivable discrete quantum phenomenon can be described by a vector in an N-dimensional Hilbert space, and $\lceil \lg N \rceil$ qubits are sufficient to provide an equivalent description. In particular, this implies that any kind of (discrete) quantum phenomenon, even those that are yet to be discovered, can be described using this mathematical formalism, so in this respect the model is complete.

4.4.2 Dealing with classical information

In many proposed quantum protocols only part of the communication is sent through a quantum channel. Often a significant amount of communication is sent through a (reliable) classical channel. Thus it seems that we are dealing with two different types of information: classical and quantum. Measurements can be seen as a process that converts quantum information to classical information. A priori it seems necessary that any model for quantum protocols should describe classical and quantum information separately, and that A and B use two channels: one classical and one quantum.

On the other hand, we saw in the previous section that classical computation and measurements can be emulated by a quantum computer, so why not do the same for two quantum computers communicating through a quantum channel? Even though some physicists would not hesitate long and accept these statements without much doubt, it is not so clear from the onset, especially since security issues are involved. Cryptographers simply have to be extremely careful about the models they use for proving security. The next few subsections will show that this care is justified.

4.4.3 Information and disturbance

The most striking difference between classical and quantum information is that the former can be copied, whereas the latter cannot. Even though quantum information can be transferred onto another quantum system (through a unitary transformation or teleportation), it is impossible to gain *classical* information about the quantum state without changing the state. Let us examine this crucial difference in more detail.

Consider a qubit which is to be measured. The measurement apparatus can be interpreted as an auxiliary quantum system, and the measurement is a unitary transformation on the combined system. This situation is exactly as has been described before (Section 4.3.2), but the purpose is different: before we wanted to argue that measurements can be emulated by a quantum computer, whereas here we want to point out that any gain in classical information must imply a disturbance of the quantum state.

When subjected to a (complete) von Neumann measurement, the qubit is asked "Are you in state $|0\rangle$ or in $|1\rangle$?" This measurement yields one bit of classical information (0 or 1, with probability $|\alpha|^2$ resp. $|\beta|^2$) which is maximal. The state is now fully disturbed: the initial state is lost, and the system has changed ('collapsed') to $|0\rangle$ or $|1\rangle$, in agreement with the result.

In particular we know that we cannot simultaneously learn the answer to the question “Are you in state $|0'\rangle$ or in $|1'\rangle$”, when

$$\begin{aligned} |0'\rangle &= \tfrac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\ |1'\rangle &= \tfrac{1}{\sqrt{2}}(|0\rangle - |1\rangle), \end{aligned}$$

which is a conjugate basis. For instance, when this measurement is performed after the first one, its result is random. It seems obvious that we can extract only one classical bit of information from a qubit, and this fact is often attributed to the Heisenberg Uncertainty Principle, but the latter is not correct. In the first place, what is being referred to is not a new principle, but just a set of equations that follow from the theory. In addition, these relations do not say anything about making two conjugate measurements on the same particle at all: The Heisenberg Uncertainty Relations reflect a statistical statement about measurements on an ensemble of particles. For more discussion on this subject, see [PE93, page 93] [IS95, page 126] [FU96, sect. 3].

Nature allows for a much more subtle kind of measurement than the measurement described above. There exist partial measurements, in which the state can disturbed a bit and in which only partial information about the state is obtained. The idea is simply to use a different U_M than the one describing complete measurements. Obviously, one extreme case is when no measurement takes place ($U = Id$): no information about the quantum state is obtained, and the state is not disturbed. All the possibilities between these two extremes exist. In Chapter 3 we discussed an alternative formalism to describe such measurements without explicitly having to deal with the auxiliary system, called POVMs, but for the discussion here this formalism is less appropriate.

The fact that we *must* cause disturbance if we want to gain information, follows from a proposition, shown by Bennett, Brassard and Mermin.

4.1 Proposition ([BBM92])

Suppose M is a measurement that leaves two non-orthogonal states undisturbed. Then M cannot yield any information about which of these two states a given system is in.

PROOF: Let again U_M be the unitary operator corresponding to M. Let $|\psi_a\rangle$ and $|\psi_b\rangle$ be non-orthogonal, so their inner product is non-zero. Let $|\varphi\rangle$ be the state of the measuring system. A state evolution that disturbs neither $|\psi_a\rangle$ nor $|\psi_b\rangle$ would look like

$$\begin{aligned} |\psi_a\rangle|\varphi\rangle &\xrightarrow{U_M} |\psi_a\rangle|\varphi_a\rangle \\ |\psi_b\rangle|\varphi\rangle &\xrightarrow{U_M} |\psi_b\rangle|\varphi_b\rangle \end{aligned}$$

But since U_M is unitary,

$$\langle\psi_a|\psi_b\rangle\langle\varphi|\varphi\rangle = \langle\varphi\psi_a|\psi_b\varphi\rangle = \langle\varphi_a\psi_a|\psi_b\varphi_b\rangle = \langle\psi_a|\psi_b\rangle\langle\varphi_a|\varphi_b\rangle \tag{4.10}$$

The first and the last equality are true because the inner product is multiplicative over the tensor product, and the middle equality is true because unitary transformations are inner product preserving. Consider the left-most and the right-most expression. Since $\langle\varphi|\varphi\rangle = 1$ and $\langle\psi_a|\psi_b\rangle \neq 0$, we obtain that $\langle\varphi_a|\varphi_b\rangle = 1$. Since $|\varphi_a\rangle$ and $|\varphi_b\rangle$ are normalized, it means that $|\varphi_a\rangle = |\varphi_b\rangle$. So the final state after the measurement interaction is the same, regardless of the state the measured system had before. Consequently, M yields no information. □

It follows that cloning (i.e. creating *two* particles each in state $|\psi\rangle$ from one particle in state $|\psi\rangle$) of quantum states is impossible: if we could clone we could measure one particle to gain information, and keep the other. More direct proofs of the *no cloning theorem* show that the evolution $|\psi\rangle|\varphi_0\rangle \rightarrow |\psi\rangle|\psi\rangle$ (with $|\varphi_0\rangle$ some fixed initial state of the auxiliary system) cannot be unitary for arbitrary states $|\psi\rangle$ [WZ82, PE93].

This information disturbance principle is the physical principle on which the security of Quantum Key Distribution is based. We do not want to discuss this protocol in full detail here, but it provides a good intuition of what information-causes-disturbance means. Moreover, it shows how an E attacking a quantum protocol gains information. As is shown in Figure 4.2, A and B communicate through a quantum channel that is being eavesdropped upon by E. They also have access to an unjammable classical broadcast channel.

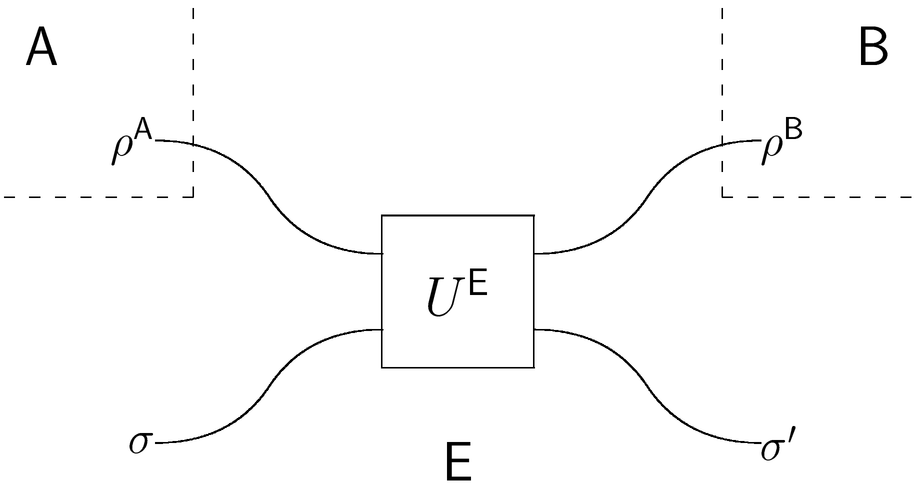

Figure 4.2: Quantum Key Distribution: E eavesdrops on the quantum channel between A and B. The final states are obtained after tracing out (see text).

E's attack is modeled through a unitary transformation U^{E} that it applies to the particles sent through the quantum channel. From E's point of view A's initial state is mixed, expressed by the density matrix ρ^{A}. The auxiliary system used by E has initial state σ, and final state σ', The latter is obtained by tracing out B's particles:

$$\sigma' = \mathrm{Tr}_{\mathsf{B}}(U^{\mathsf{E}}(\rho_A \otimes \sigma)(U^{\mathsf{E}})^\dagger).$$

The final state of B is obtained by tracing out E's system.

The distance between σ and σ' is a measure for the amount of 'information' E has gained. On the other hand, the distance between ρ^A and ρ^B is in indication for the amount of disturbance that U^{E} has caused; clearly, in the absence of E they are equal. Various reasonable notions of distance exist, this was exactly the subject of Chapter 3. The information disturbance principle now tells us that if E tries to gain information, this necessarily causes a disturbance in the transmission between A and B with overwhelming probability. This they will detect through their public discussion after the quantum transmission has taken place.

Quantifying the relation between the information E gains and the disturbance caused by the attack is generally not easy. In [FGGNP97] this problem is addressed for one qubit, using Shannon information (distinguishability) as the distance measure. A simple relation is derived for the information obtained by a partial measurement in the $\{|0\rangle, |1\rangle\}$ basis, as a function of the disturbance this measurement causes in the conjugate $\{|0'\rangle, |1'\rangle\}$ basis, and it is shown the sum of information gained and disturbance caused never exceeds one bit. A less rigorous derivation of this relation is found in [CG97], which also deals with the case for two qubits. For more than two qubits no general formulas concerning information and disturbance are known. For a more thorough exposition about the information disturbance principle the reader is referred to [FU96].

4.4.4 Redefining the view

The different nature of quantum information has significant consequences for the model, in particular for the definition of the view. Look for instance at Figure 2.4 on page 24, and assume that A, B and E are quantum computers, and suppose that A is corrupted. This means that A gives all the qubits she receives immediately to E, and sends only qubits that it gave her. However, unlike in the classical case, E's capability to look at the qubits it receives from A is severely restricted by the fact that if it looks, it will also cause a disturbance.

In general, from the discussion in the previous subsection it follows that *parties cannot monitor a quantum channel without disturbing it*. Copying the channel traffic and storing it for later reference is impossible. Consequently, the *quantum view*, i.e. the view that a party has of the quantum communication taking place in a protocol, is completely different from the classical view as defined in Section 2.2.2. To emphasize this point consider an elementary example: Suppose that A sends a qubit to B, and B does nothing to this qubit and sends it back. (In the context of a bigger protocol this could be one meaningful course of action among other possibilities). In this case no communication took place, and the view of B is empty (contains no information).

This may be somewhat puzzling, until we realize that A and B communicate only if "... physical acts of A have induced a desired physical state in B" [CT91]. Since no physical change in B's system takes place, no communication has occurred. Here we are reminded of the fact that, even though we are used to model communication in a mathematical way, communication is intrinsically a *physical* process.

The example of Quantum Key Distribution (see Figure 4.2) illustrates in which way an adversary E can extract information from quantum interactions: knowing its own initial state and its own unitary evolution, P's final state contains information about the states of the systems it interacted with. Therefore we define the quantum view of P as its final state. Recall that a quantum state is commonly represented as a density matrix, which can be regarded as a generalization of a stochastic variable (see Chapter 3).

It is very tempting to regard the classical view simply as a special case of the quantum view. This may seem evident from a physical point of view, but it is wrong from a cryptographic point of view. We cannot measure quantum information without disturbing the state, but for classical information this is not the case. If a party knows that some qubits represent classical bits which are encoded by sending orthogonal states, s/he can measure them and obtain the information therein contained without being caught, so it is logical to assume that s/he will do so. Therefore *it is necessary to separate the classical and the quantum communication taking place in a protocol, and to define the classical view and the quantum view separately.*

At this point there are two ways to proceed:

(a) We assume that A and B are connected through a *classical channel and a quantum channel*. The view of P (either A or B) is defined as the combination of the classical view (the concatenation of all classical bits sent and received by P) together with the final state ρ_F^{P}.

(b) We assume that A and B are connected through a *quantum channel only*. This quantum channel incorporates the classical channel, as follows:

(i) P sends a 0 or a 1 as $|0\rangle$ or $|1\rangle$.

(ii) For each classical bit that P receives s/he emulates a measurement according to the $\{|0\rangle, |1\rangle\}$ basis as described in Section 4.3.2.

(iii) For each qubit that P sends or receives and which is known to represent a classical bit, P makes a copy of it which s/he stores away in quantum memory until the end of the protocol. (This can been done using the Fanout, for instance, so this copying of classical bits is certainly unitary.)

Both descriptions are clearly equivalent: if we represent the classical view obtained in *(a)* using qubits in state $|0\rangle$ or $|1\rangle$, then the combined view obtained in *(a)* must be equivalent to the quantum view obtained in *(b)*. We choose option *(b)* because of its simplicity: we have only one type of information and measurements can be emulated as described before, so they can all be postponed until the end of the protocol. This is in agreement with the model used by Mayers [MA97]. The only difference in the two scenarios is that in *(a)* the measurements (and the 'collapses') occur during the protocol, whereas in *(b)* the measurements are emulated and the collapse occurs at the end. As has been said in Section 4.3.2: the point of time of a measurement does not changed the probability distributions on its results.

There is another way to phrase that security is not affected in the transition from *(a)* to *(b)* [MS97]. Intuitively, an honest party can always (and should always) measure those bits sent to him which correspond to classical bits. So even though E can fancy itself that it is doing a quantum attack on the classical bits, all these bits will collapse to either $|0\rangle$ or $|1\rangle$. All possible entangled states E might have created also collapse as a result of the honest party's measurement, and effectively E's attack is equivalent to a (probabilistic) classical attack. It is possible to show this formally with the help of Definition 4.2, but in the end the argument does not differ much from the one given in the previous paragraph: as long as an honest participant measures (forces a 'collapse') classical bits, s/he is protected.

4.4.5 Comparison with other work

The first model for quantum protocols was presented by Yao [YA95], who mentions that separating out the classical communication is useful. This model was adapted both by Mayers [MA97] and by Lo and Chau [LC97A] to prove that Quantum Bit Commitment is impossible.

The model introduced in [MA97] addresses the issue of classical information, but without giving any explanation, which was provided by personal communication [MS97]. This issue is nowhere addressed in the works of Lo and Chau, but this is not problematic when they prove that QBC is impossible, for the following reason. Intuitively, the proof shows that for every quantum computer B there exists a local unitary transformation that a cheating A can perform, allowing her to open as a 0 or as a 1. But since the quantifier ranges over all B, this also includes the B's that measure and copy all classical information as described in *(b)* in the previous subsection. In other words, Lo and Chau's model is good enough for proving that QBC is impossible. But the argument that "classical communication is a special case of quantum communication and needs no further consideration" [LC97B] seems incorrect for defining a general model of security for arbitrary quantum protocols.

4.4.6 Protocol tasks

In Chapter 2 we separated tasks (the ideal behavior of honest parties) from the ideal protocol (which included all possible dishonest behaviors). But we saw already that we can specify tasks which might not be implementable in practice; for instance, most implementations of bit commitment cannot prevent A from sending in a value that contains nonsense.

In the case of Quantum Bit Commitment things are even more dramatic. The flaw in the proof of [BCJL93] was the following. It was proven that A could not open both as a 0 and a 1 simultaneously. However, this overlooked a possibility that classically does not make sense, but quantumly does: what if A can open a 0 or a 1 *at her choosing*? Classically, if A opens and B sees a 0, then he can assume the 0 was there in the COMMIT phase. But quantumly this isn't true anymore: by clever use of entanglement A can create a state on her side which allows her to open as a 0 or a 1 as she pleases. Moreover, if she is never requested to open she is able to exploit this state, see [BCMS97]. Observe that this does not mean that our task specification for BC is flawed, but that in a purely quantum-mechanical model Bit Commitment does not exist.

Considering tasks which have qubits as input and output stretches the imagination even more. What does it mean to commit to a qubit? Does it mean that A gives a physical qubit to the trusted host, or a construction prescription? Certainly a trusted host cannot inspect a qubit given to him. And how to deal with entangled qubits?

For now we will leave these questions for what they are, and restrict ourselves to cryptographic tasks which are classical. That is, A's and B's inputs and outputs are classical, even though they use quantum computers and a quantum channel. Despite the negative results about QBC and its implications for other two-party quantum protocols [Lo97], our quest for a proper model and definition is not meaningless, since we can always make some extra assumption (see the end of Section 4.5.4).

4.5 Formal definitions

After all these preparatory discussions we are finally ready to present our definition of security for quantum protocols. We follow the same approach as in Section 2.2.2, making modifications when the different nature of quantum information forces us to. But first we present a formal model of computation.

4.5.1 The Quantum Turing Machine

In principle, quantum algorithms can conveniently be described using the formalism of quantum circuits, but no complexity measure seems well-established, at least at the time of writing. On the other hand, even though describing a quantum algorithm by means of a Quantum Turing Machine (QTM) can be tedious, there is a well-defined complexity class based on them. Moreover, since the classical definitions in Chapter 2 use probabilistic TMs, it will be easier to extend the definition using QTMs instead of quantum circuits.

Quantum Turing Machines were first considered by Deutsch [DE85], and considerably improved by Bernstein and Vazirani [BV97]. We will very briefly summarize those points from their paper that we need here.

A deterministic TM can be defined as a triplet $\langle \Gamma, Q, \delta \rangle$, where Γ is the alphabet with blank symbol $\top$, where Q is the set of states with initial state q_0 and final state q_f, and where δ is the transition function: $\langle q, s \rangle \xrightarrow{\delta} \langle q', s', d \rangle$. This means that when the machine is in state q and reads an s on the tape, it will write s', move the head in direction $d \in \{\mathsf{left}, \mathsf{right}\}$, and change to the state q'.

For a probabilistic TM one pair $\langle q, s \rangle$ can map to several triples $\langle q', s', d \rangle$; each triple is associated with a certain probability which all sum up to one. For a quantum TM one pair $\langle q, s \rangle$ can map to several triples $\langle q', s', d \rangle$; each such transition is associated with a certain amplitude. The sum of the squares of the norm of these amplitudes must equal 1.

The *configuration* (or instantaneous description) of a classical TM includes the current state, the contents of the tape, and the position of the tape head. Let $\{c_i\}_{i=1}^{\infty}$ be an enumeration of all possible configurations. In the section about quantum circuits we have seen that to define a unitary transformation, it suffices to define the transformation on the classical basis states $|b_1...b_n\rangle$ with $b_i \in \{0, 1\}$. For the QTM we do something similar, using the configurations $\{c_i\}_{i=1}^{\infty}$ as a basis.

We construct the (infinite-dimensional) matrix $U = U_{\langle\Gamma,Q,\delta\rangle}$ as follows: the matrix entry with column c and row c' contains the amplitude associated to go from configuration c to c' in one step of the QTM, as prescribed by δ. Note that the number of non-zero entries in each row, i.e. the number of configurations c' reachable from c, is certainly finite because δ has a finite description (in fact it is at most $2 \cdot \#Q \cdot \#\Gamma$). So each row of U has zero entries almost everywhere. In the special case of a reversible deterministic TM there is only one transition leading from c to c', and we obtain a permutation matrix.

We call U the *time evolution operator*, but in order to get a time evolution agreeing with quantum mechanics we must require that U be unitary (or, as can be shown equivalent, preserves Euclidean length). So if $\gamma := \sum_{j=1}^{\infty} \alpha_j c_j$ is a superposition of configurations with Euclidean length 1 (i.e. $\sum_{j=1}^{\infty} |\alpha_j|^2 = 1$), then one step of the QTM is defined by

$$\gamma \xrightarrow{U} U(\sum_{j=1}^{\infty} \alpha_j c_j) = \sum_{j=1}^{\infty} \alpha_j U(c_j). \tag{4.11}$$

It is crucial to keep in mind that the definition of the QTM is based on the operator U acting on the complex vector space spanned by $\{c_i\}_{i=1}^{\infty}$. Even though U is unambiguously defined by Q, Γ and δ, one easily loses intuition for what is going on. For concreteness, we can still think of each tape symbol as a qustring of $\lceil \lg \#\Gamma \rceil$ qubits, with the $\top$ represented by the qustring consisting of $|0\rangle$'s only. But imagining the tape head in many superpositions becomes rather mind-boggling.

It may not seem evident to verify that $U_{\langle\Gamma,Q,\delta\rangle}$ indeed is unitary, but Bernstein and Vazirani show how this requirement translates to a set of conditions on δ which are relatively easy to verify, preserving Euclidean length being one of them.

Note that the amplitudes associated to each image of $\delta(q, s)$ are restricted to the subset of complex numbers whose real and imaginary parts can be approximated with an accuracy of 2^{-n} in a time polynomial in n for a deterministic Turing Machine. Without such a restriction non-computable information could be encoded in the binary expansion of the amplitudes. Another point is that not all different execution paths, represented by different configurations in superposition, need to enter q_f in the same time step. This is resolved by requiring that all inputs of same length take the same amount of time.

A language $L \subset (\Gamma - \top)^*$ belongs to the class *BQP* if the following holds: there exists a QTM that needs polynomially many steps in the length of its input z, such that if $z{\in}L$ ($z{\notin}L$) then z is accepted (rejected) with probability at least $\frac{2}{3}$. The class *BQP* is considered the quantum analogue of *BPP*: it reflects the class of languages efficiently computable on a quantum computer.

One of the main results of [BV97] is the fact that a universal QTM exists with only a polynomial simulation overhead. This allows us to use universal simulation as the mechanism through which the interface S has access to the adversary E.

4.5.2 Protocol, output, view, profile, the adversary

In order to execute a quantum protocol we suppose that A has access to a Quantum Turing Machine, called $\mathtt{A}$. This QTM is equipped with an input tape, work tape and output tape. Before the protocol starts the (classical) input x is written on the input tape, whereas the work tape and output tape are in some well-defined pure state. The quantum state corresponding to $\mathtt{A}$'s overall initial configuration is described by $|\psi_0^{\mathsf{A}}(x)\rangle$. We define $|\psi_0^{\mathsf{B}}(y)\rangle$ in a similar way. $\mathtt{A}$ and $\mathtt{B}$ both have access to a communication tape, whose initial state is described by $|\psi_0^{\mathcal{C}}\rangle$.

We assume that A begins and that B ends, so k, the number of steps, is even. We will describe A's action during step j of the protocol for j odd; for j even A does nothing.

(a) A copies to one track of her work tape (reserved for logging purposes) all qubits received on the channel $\mathcal{C}$ representing classical bits.

(b) A applies a unitary operator U_j^{A} on the Hilbert space $\mathcal{H}^{\mathsf{A}} \otimes \mathcal{H}^{\mathcal{C}}$.

(c) A copies to her logging track all qubits on $\mathcal{C}$ representing classical bits before she sends them.

B does nothing during step j, so $U_j^{\mathsf{B}} = \mathit{Id}$. Note that anything he could do while it is A's turn is restricted to a unitary transformation on his own qubits, which can be included in the previous or next step.

The unitary transformation on the full Hilbert space $\mathcal{H}^{\mathsf{A}} \otimes \mathcal{H}^{\mathcal{C}} \otimes \mathcal{H}^{\mathsf{B}}$ in step j is

$$U_j^{\mathrm{tot}} = U_j^{\mathsf{A}} \otimes \mathit{Id}^B \qquad (j \text{ odd}) \tag{4.12}$$

For j even, it is B's turn and A does nothing. U^{B} acts on $\mathcal{H}^{\mathcal{C}} \otimes \mathcal{H}^{\mathsf{B}}$, and we find

$$U_j^{\mathrm{tot}} = \mathit{Id}^A \otimes U_j^{\mathsf{B}} \qquad (j \text{ even}) \tag{4.13}$$

We see that $\mathtt{A}$ and $\mathtt{B}$ can be described as a finite sequence of operators:

$$\mathtt{A} = [U_1^{\mathsf{A}}, \mathit{Id}^A, U_3^{\mathsf{A}}, \mathit{Id}^A, ..., U_{k-1}^{\mathsf{A}}, \mathit{Id}^A] \tag{4.14}$$

and

$$\mathtt{B} = [\mathit{Id}^B, U_2^{\mathsf{B}}, \mathit{Id}^B, U_4^{\mathsf{B}}, ..., \mathit{Id}^B, U_k^{\mathsf{B}}] \tag{4.15}$$

All the unitary operators U_j^{P} above are computable in polynomial time by a QTM.

Define

$$|\psi_0^{\mathrm{tot}}\rangle := |\psi_0^{\mathsf{A}}(x)\rangle \otimes |\psi_0^{\mathcal{C}}\rangle \otimes |\psi_0^{\mathsf{B}}(y)\rangle \tag{4.16}$$

The protocol $[\mathtt{A}, \mathtt{B}](x)(y)$ evolves as follows:

$$|\psi_j^{\mathrm{tot}}\rangle := U_j^{\mathrm{tot}}(|\psi_{j-1}^{\mathrm{tot}}\rangle), \tag{4.17}$$

where $j = 1...k$.

At the end of the protocol, B moves the contents of the communication tape to its work tape, returning the communication tape to its initial state. Then each party P makes a measurement of the output tape in the standard basis. The result of this measurement defines a stochastic variable, the party's *output* of the protocol, written as $[\mathtt{A}, \mathtt{B}]_{\mathsf{P}}(x, y)$.

The *view* of a party P is the final state *after* the measurement has been performed. This final state, written as $[A, B]^*_{\mathsf{P}}(x, y)$, is a density matrix. It is obtained by tracing out as explained in Subsection 4.3.3.

The adversary E is modeled like a player, but is more powerful. It is also a QTM, but we put no restrictions on the time it takes. E has an extra step 0, denoted by U_0^{E} in which it can issue a (classical) corruption request (in a two-party protocol it does not seem meaningful to have E being in a superposition of corrupting A and B, but in a multi-party setting it is). The operator U_j^{E} describes its behavior during round j. For j odd or even, depending whether A or B is corrupted, $U_j^{\mathsf{E}} = \mathit{Id}$. During the protocol E behaves as any party monitoring a combined channel: it measures the qubits known to be classical and stores them.

At the end E must provide the corrupted party with an output; this follows from the discussion about output awareness in Section 2.5.2 and our proposed definition of security. Hence, at the end of the protocol E performs a measurement M^{E}, and provides the result to the corrupted party, which defines this party's output. Its quantum state after the measurement is defined as the adversary's view of the protocol, and is written as $[A, B\colon E]^*_{\mathsf{E}}(x, y)$.

Just as in the classical case, we are interested in the *profile*, the triple consisting of the output of A, the output of B, and the view of E. Therefore we define the following variable:

$$[A, B\colon E]^{\ddagger}(x)(y) \stackrel{\text{def}}{=} \langle [A, B\colon E]_{\mathsf{A}}(x)(y), [A, B\colon E]_{\mathsf{B}}(x)(y), [A, B\colon E]^*_{\mathsf{E}}(x)(y) \rangle. \quad (4.18)$$

Even though strictly speaking the first and the second component are stochastic variables, we will regard them as density matrices.

Just as in the classical case, we can assume that the inputs x and y are samples from the stochastic variable X and Y, so $[A, B\colon E]^{\ddagger}(X)(Y)$ is well-defined.

Recall that cryptographic protocols are usually designed with a security parameter, n, which is not explicit in the notation. The notion of *exponential indistinguishability for quantum states* has been explored in detail in Chapter 3.

4.5.3 Comparing the security of protocols; interface

For comparing the security of two quantum protocols we follow the same course as for comparing the security of classical protocols, as explained in Section 2.2.3. Just as before we introduce an *interface* which can translate any attack $\mathtt{E}$ on Π_1 into an attack $\mathtt{S}(\mathtt{E})$ on Π_2, but now we assume that $\mathtt{S}$ is a QTM. Again we use universal simulation.

The following definition is the analogue of the classical definition of relative security, Definition 2.1. The difference is that everything, except the protocol output, has become quantum.

4.2 Definition *Protocol* $\Pi_1 = [\mathtt{A}_1, \mathtt{B}_1]$ *is at least as secure as* $\Pi_2 = [\mathtt{A}_2, \mathtt{B}_2]$ *if there exists an interface* $\mathtt{S}$ *from* $[\mathtt{A}_1, \mathtt{B}_1]$ *to* $[\mathtt{A}_2, \mathtt{B}_2]$ *such that for all* $\mathtt{E} \in \mathcal{E}_1$ *and for all possible* $\langle X, Y\rangle$ *which are joint stochastic variables on* $\mathcal{X} \times \mathcal{Y}$*:*

$$[\mathtt{A}_1, \mathtt{B}_1 \colon \mathtt{E}]^{\ddagger}(X)(Y) \approx [\mathtt{A}_2, \mathtt{B}_2 \colon \mathtt{S}(\mathtt{E})]^{\ddagger}(X)(Y). \tag{4.19}$$

By choosing X and Y to range over all possible stochastic variables, we deal with the problem of side information, classical and quantum. This follows from the fact that we consider *classical* tasks only: whatever quantum system E ends up with, it can never be entangled with the input or output of the honest party, since s/he has been honest and has 'collapsed' the adversary's state to something classical. Therefore, all that has been said about composition of protocols and about protocol reductions carries over to quantum protocols.

4.3 Theorem *Let* $[\mathtt{A}'_1, \mathtt{B}'_1]$ *be as secure as* $[\mathtt{A}'_2, \mathtt{B}'_2]$*, and let* $[\mathtt{A}''_1, \mathtt{B}''_1]$ *be as secure as* $[\mathtt{A}''_2, \mathtt{B}''_2]$*. Then* $\Pi_1 = [\mathtt{A}'_1\mathtt{A}''_1, \mathtt{B}'_1\mathtt{B}''_1]$ *is as secure as* $\Pi_2 = [\mathtt{A}'_2\mathtt{A}''_2, \mathtt{B}'_2\mathtt{B}''_2]$.

If we consider cryptographic tasks for qubits this reasoning is likely to fail. Entanglement allows for interdependencies between such protocols that classically do not exist and probably each new protocol would have to be proven from scratch, even if the building blocks are well-known sub-protocols.

4.5.4 Ideal protocols and absolute security

Since we model classical protocol tasks, we can use the same definitions of ideal protocols as in Section 2.2.1. To facilitate the translation between the quantum implementation and the classical ideal protocol, there are two possibilities:

(a) We assume that $\mathcal{A}$ and $\mathcal{B}$ come with an 'translator' which codes classical bits b to qubits $|\mathsf{b}\rangle$ and vice versa.

(b) We assume that $\mathcal{A}$ and $\mathcal{B}$ are also quantum computers, and furthermore that the trusted host 'collapses' each qubit it sends or receives by measuring it according to the $\{|0\rangle, |1\rangle\}$ basis.

Since an honest $\mathcal{A}$ or $\mathcal{B}$ communicates only classical information with the trusted host, the possibilities are equivalent. We choose possibility *(b)*.

The definition of absolute security of a quantum protocol follows the same procedure; we require that the protocol be at least as secure as the ideal task:

4.4 Definition *Let* $\Pi_1 = [A_1, B_1]$ *be the protocol implementation of a cryptographic task* $\mathbf{T}$*, while* $\Pi_2 = [\mathcal{A}, \mathcal{B}]$ *is the ideal protocol for* $\mathbf{T}$*. Then* $[A_1, B_1]$ *is a* <u>*secure protocol for* $\mathbf{T}$</u> *if* $[A_1, B_1]$ *is as secure as* $[\mathcal{A}, \mathcal{B}]$.

At the moment of writing we do not know a candidate protocol that could satisfy this definition, and an additional assumption of some sort is probably necessary. For instance, we can assume that decoherence (quantum noise) exists, that the number of particles which can be measured coherently is only a constant fraction of all the particles sent, or that honest parties are available to bootstrap the protocol. Applying the definition in this context will be the subject of future research.

4.6 Rewinding the adversary

In Section 2.3 we spent a considerable effort explaining the various choices in the way the interface `S` controls the adversary `E`. In light of the different nature of quantum information we have to reconsider this.

The points which were addressed in Section 2.3 were the following:

(a) *The order of the quantifiers.*

There is no reason to reopen this discussion in the quantum scenario; the same argumentation as in the classical case applies.

(b) *The computational power of the interface.* We have argued before that the interface should not be too powerful. Therefore we require that the time it takes the interface S to attack Π_2 is polynomial in the time it takes the adversary E to attack Π_1.

Observe that if we put no such limitation on the computational power of S, it could simulate E simply by working out all the mathematical equations. This follows from Feynman's observation that every quantum evolution can be simulated classically with exponential slowdown [FE82]. The exponent comes from the fact that the state of a system consisting of n particles is described by a vector in a 2^n-dimensional Hilbert space.

(c) *The extent to which the interface can influence the adversary.* This is an important issue. Classically, we discussed two choices:

(i) *universal simulation*, for which we reasoned that taking snapshots and resetting was possible.

(ii) *black-box simulation*, for which resetting was possible, provided the randomness of E was supplied via the random tape, not internally.

Quantumly we consider the same choices. However, due to the information disturbance principle discussed in Section 4.4.3 it turns out that the classical techniques to simulate a probability distribution fail, as we will see shortly.

What does rewinding exactly mean in the classical case? Suppose a computer is running some program. Let c_1 be the configuration of this program at some well-defined time t_1. When the program is implemented by a TM, the configuration includes the current state, the contents of all tapes, and the position of the tape heads. We would either like to remember c_1, or to be able to reconstruct it accurately. We proceed with the computation until some later time t_2. In one way or another we subject the current configuration c_2 to a test T. If the test succeeds then continue the computation, else return to the configuration c_1 in one way or another, but in polynomial time.

In case of black-box simulation, configuration c_1 can be reconstructed by resetting the black-box and re-running the whole program again. Even though this does not change the behavior of the simulated E, the interface S can remember the random choices that worked, and try different ones where the simulation failed. In some sense this can be regarded as backtracking for probabilistic algorithms: if the algorithm fails it can return to a previous configuration in which it was successful (at least up to that point) and try another branch. This kind of simulation is often used in proofs of security for sequential composition of protocols under computational assumptions. It is also used, for instance, in [NOVY92]. In universal simulation 'snapshots' are allowed, so resetting and re-computing is unnecessary.

Quantumly it is tempting to think that rewinding might be possible by reversing the computation, but this is false. Let $|\psi_i\rangle$ be the state at t_i, $i = 1, 2$. Let $U_{1,2}$ be the unitary evolution between time t_1 and t_2. Let T be the two-valued SUCCESS/FAIL measurement following $U_{1,2}$, and let U_T denote the unitary operator emulating T. As before, we put the measurement result in one qubit, and get the unitary evolution

$$|\psi_1\rangle \xrightarrow{U_{1,2}} |\psi_2\rangle \xrightarrow{U_T} \alpha|\mathsf{S}\rangle|\psi_S\rangle + \beta|\mathsf{F}\rangle|\psi_F\rangle \tag{4.20}$$

Now we would like to implement something like: if $|\mathsf{S}\rangle$ then continue with $|\psi_S\rangle$ else $U_T^{-1}(|\mathsf{F}\rangle|\psi_F\rangle)$. However, the 'else' case does not have the desired effect of returning to $|\psi_2\rangle$ because the information contained in $|\psi_S\rangle$ is lost. So the suggested evolution is almost certainly not unitary.

To deal with the question rigorously, we appeal to Proposition 4.1 once more: the qubit that contains the test result after the emulated measurement U_T is classical. So it can be measured without causing a further disturbance, and we can gain one bit of information with non-negligible probability. It follows that every state not orthogonal to $|\psi_F\rangle$ has been disturbed. So we conclude:

4.5 Proposition *Rewinding by reversing the unitary transformation induced by* $\mathtt{E}$*, or taking snapshots is impossible.*

The overall conclusion with regard to measurements is that a measurement can always be undone, see Section 4.3.2, provided that *all* its possible (classical) measurement results are 'forgotten'. We cannot accept some results and undo others. In Svozil's extension [Sv97, p. 12] of Schrödinger's words [Sc35]: "By no way can we open the prediction catalog at one page [measure a density matrix in one basis], read and (irreversibly) memorize the page, close it; then open it at another, complementary [conjugate], page."

But with regard to rewinding, showing that rewinding by reversing or by taking snapshots is impossible does not show that no other ways to rewind in polynomial time exists. Below we show that the classical technique of resetting the black box and computing 'forward', as has been described above, can only exist if quantum one-way functions do not exist. More precisely, suppose that $\mathtt{S}$ has a polynomial time algorithm R which takes as input $|\psi_1\rangle$ and a description of $U_{1,2}$ and T, and which outputs $|\psi_2\rangle$ such that $T(|\psi_2\rangle) = \textsc{success}$. Then R can be used to invert any one-way permutation.

The existence of one-way permutations, easy to compute in quantum polynomial time but hard to invert, is likely because of the results of Bennett, Bernstein, Brassard and Vazirani [BBBV97]. They show that, relative to a length preserving permutation oracle chosen at random, with probability 1 the class *NP*$\cap$*co-NP* cannot be solved in time $0(2^{n/3})$ on a QTM. It must be realized that if the class *NP*$\cap$*co-NP* can be solved in polynomial time, then one-way permutations cannot exist: for any permutation f the language $\{y \in \{0,1\}^n | \exists x \in \{0,1\}^{n-1}$ s.t. $f(x0){=}y\}$ is evidently in *NP*. It is also in *co-NP* since its complement can be described as the set of y with a pre-image ending in 1. So deciding whether the pre-image of y ends in a 0 or 1 can be decided efficiently; a similar technique applies to retrieve the other bits.

4.6 Proposition *Let $|\psi_2\rangle$ be an arbitrary quantum state, and let T be an arbitrary two outcome test on $|\psi_2\rangle$. Suppose that there exists a quantum algorithm R which takes as input $|\psi_1\rangle$, $U_{1,2}$ and T, and which yields a new state $|\psi_2\rangle$ such that $T(|\psi_2\rangle) =$ SUCCESS with overwhelming probability. Then there exists a polynomial time algorithm for inverting any one-way permutation.*

PROOF: Let $f : \{0,1\}^n \rightarrow \{0,1\}^n$ be a permutation. Let $\bar{x} \in \{0,1\}^n$ be arbitrary, and let $\bar{y} := f(\bar{x})$. We construct a polynomial time algorithm which inverts f at $\bar{y}$. First we create the state

$$|\psi_0\rangle = \sum_{x \in \{0,1\}^n} \frac{1}{2^{n/2}} |x; 0^n\rangle.$$

Then we evaluate the image of this state under f:

$$|\psi_1\rangle := f(|\psi_0\rangle) = \sum_{x \in \{0,1\}^n} \frac{1}{2^{n/2}} |x; f(x)\rangle$$

Let $\bar{y}_i$ be the ith bit of $\bar{y}$ and define the test $T(|\psi_2\rangle) =$ "the last bit of $|\psi_2\rangle$ equals $\bar{y}_n$".

By hypothesis, $R = R(|\psi_1\rangle, \mathit{Id}, T)$ computes in polynomial time a new state $|\psi_2\rangle$ such that the last bit of $|\psi_2\rangle$ equals $\bar{y}_n$. We proceed in this fashion for all the bits of $\bar{y}$, using the output of each procedure call to R as input to the next. After n steps the final state is $|\bar{x}; \bar{y}\rangle$, so f has been inverted. $\Box$.

So we have proven that rewinding by reversible computation, by taking snapshots, and by resetting and roll forwards from the beginning is impossible. Still, this does not prove that rewinding is completely ruled out. On the other hand, the kind of rewinding used here is in fact very realistic. For instance, in [NOVY92], Naor, Ostrowsky, Venkatesan and Yung give a Bit Commitment scheme which is computationally binding. That is, if an algorithm E exists that can open the Bit Commitment as 0 or 1 with non-negligible probability, they show that there exists a (classical) polynomial time algorithm S which inverts any one-way permutation f.

This algorithm S makes use of classical rewinding. Suppose we replace the classical algorithms E and S in this reduction by quantum algorithms E and S, and we allow quantum rewinding. Then, using the techniques developed in [BCMS97], we can obtain a quantum polynomial time algorithm that inverts any one-way permutation. Since rewinding is the only classical assumption the reduction (i.e. the description of S) uses, this gives us yet more evidence that quantum rewinding is generally impossible.

The fact that in the quantum model rewinding is probably impossible is worrying. For instance, if we take any Zero Knowledge protocol and run it on Quantum Turing Machines, then suddenly the proof of its security does not hold any longer. The general conclusion of this section is that proving the security of a quantum protocol by using simulatability is much harder, but at this point in time no promising alternative to proving security seems available.

4.7 Extensions and open problems

So far, we have assumed that the channel that connects A and B is perfect, and that it cannot be eavesdropped on. Observe in the first place that such assumptions are justified because of the existence of quantum error correcting codes (see for instance [SH96, PR97] and the references therein). Intuitively, the idea is to embed each 2-dimensional qubit in a much larger Hilbert space, of 2^7 dimensions, for instance. When an error in one of the seven qubits occurs, the state leaves its allotted subspace. We can always correct this error and return the state to the subspace it should be in, surprisingly *without* disturbing the state.

To deal with an eavesdropper E, we only have to remember Figure 4.2, in which U^{E} is a transformation that it can perform on the quantum channel after A has sent but before B has received. It might be interesting to prove the security of the Quantum Key Distribution protocol by using simulation. In the ideal task a trusted host sends the same random string to both A and B, and E can corrupt nobody. So the question becomes whether an interface S, without talking to anybody, can produce a conversation which is indistinguishable from the real interaction between A and B.

Another extension is to include decoherence. Decoherence is quantum noise: in a non-perfect transmission we must consider the possibility that a quantum particle gets entangled with a particle which is not part of our system, usually called the environment. This is again a unitary transformation, just as in eavesdropping, but since we do not control the 'strange' particle the effect is an uncontrolled disturbance of the state, i.e. noise. Decoherence is exactly the phenomenon we are trying to fight with quantum error correction. But instead of regarding it as nuisance, it might also be used to our advantage. If decoherence can be proven to exist and can be quantified properly, it might very well be an additional hypothesis to base quantum protocols on. It would basically exploit the same idea as has been done classically by Crépeau and Kilian [CK88, CR97], in which it is shown how to build Oblivious Transfer from a Noisy Channel.

A somewhat different question concerns transferability of proofs. In a classical zero knowledge protocol, if A provides B with a proof, this proof is certainly not transferable to a third party D: by definition B could have created this proof all by himself. The question is whether the same argument still applies in quantum protocols, in particular when D first supplies B with a quantum system in a state unknown to B. In addition, B is supposed to perform measurements on this system depending on messages received from A, and to use the measurement results as a source of randomness in his communication with A. After B has finished interacting with A he sends the transcript of the conversation, together with the modified quantum system back to D. The question is whether B is able to satisfy D without ever talking to A, or whether this serves as a transferable proof of knowledge.

Chapter 5

Conclusion

> We must never forget that "reality" too is a human word just like "wave" or "consciousness". Our task is to use these words correctly—that is, unambiguously and consistently.
>
> NIELS BOHR (as quoted by Kalckar)

The main objective of this thesis, presenting a formal model for quantum protocols and developing a definition of security, has been achieved. However, it is still too early to judge precisely the merits of this definition. A part of the problem is that no two-party protocol task fulfills the definition without an additional assumption, for instance the existence of noisy quantum channels (decoherence). On the other hand, the notions of exponential indistinguishability for quantum systems and of reduction between protocols ("an attack on protocol Π_1 must imply an attack on Π_2"), as a fundamental method to compare the security of quantum protocols, have been defined properly.

We have presented good reasons to believe that the classical technique of rewinding is quantumly impossible. But in a quantum world, where everything is reversible and no information gets lost, the fact that rewinding is impossible looks actually very reasonable. The initial objection when this research was started, that Quantum

Bit Commitment was secure yet not rewindable or simulatable, has disappeared. It will be interesting so see how these notions develop when we make the additional assumption that decoherence exists.

A practical problem encountered when trying to prove security is the fact that many quantum protocols use error-correcting codes at some stage. Maybe the techniques developed in Chapter 3 can be extended to deal with this problem.

From Chapter 2 it is obvious that, after 20 years of (public) research in cryptography, the field is still in a precarious state. Even just classically we talk about protocol security all the time, but we still have problems defining it properly, or coming up with rigorous proofs. Yet, I am not aware of more than 5 papers continuing the approaches of Beaver and/or Micali-Rogaway presented 6 years ago. In my opinion there is a serious problem here, and it deserves much wider attention.

Even if practical implementations of quantum protocols remain restricted to quantum key distribution, I believe quantum cryptography has taught cryptography an important lesson. It has shown us that our intuitions about security and our definitions are based on *classical* assumptions. Since the overwhelming majority of protocols in use today is classical anyway, this is perfectly all right. Nevertheless, from a purely theoretical perspective things look different. Physical experiments show us that the world is intrinsically quantum. Therefore we should also accept the quantum-version of the Church-Turing thesis: that quantum computation is *the* fundamental notion of computability. But if we accept that, we must certainly investigate what security defined in terms of quantum mechanics means. This has led already to some surprises, new insights and puzzling questions about what security really means. In this respect we only begin to understand what's going on.

The title of my thesis is tentative. I hope that this work contributes to a proper understanding of the topics mentioned above.

Acknowledgments

Regardez le ciel. Demandez-vous : Le mouton oui ou non
a-t-il mangé la fleur? Et vous verrez comme tout change...
Et aucune grande personne ne comprendra jamais que ça a
tellement d'importance!

LE PETIT PRINCE

In the first place I would like to thank my two supervisors, Gilles Brassard and Claude Crépeau. I can never decide who I should thank more, therefore I won't. For neither of them I have ever been able to figure out how exactly they do research or how they think, but both have mathematical, physical and cryptographic intuitions that would always baffle me ("Where did he get that from??"). Working with them in the vanguard of science is extremely stimulating and exciting, and I have learned enormously from them. They also kept their confidence in me when progress was not as expected; for this I thank them most. The Chapters 2 and 4 grew out of many long discussions with them.

The money for my PhD came indirectly from the research councils of Canada and Québec. I want to thank NSERC and FCAR for making this research possible.

Intellectually the most shocking experience of the past five years was to be confronted with quantum mechanics (maybe apart from, on the sly, Richard Dawkins' books on evolution). I had always been interested in the subject, but probably got a little more than I was prepared for. Actually, during the first months in Montreal

I thought that the whole subject of quantum computation was ridiculous, and I was shocked when I began to understand that the Church-Turing thesis is indeed a physical assumption. I would like to thank Richard Jozsa, Bill Wootters, Chris Fuchs and Tal Mor for their endless efforts trying to explain quantum physics to me. I think their time and energy have not been completely in vain: I often had (and still have) this ”ooh, so *that* is what he meant” experience, usually several years later.

Chapter 3 is the result of work done together with Chris. He deserves an extra thanks for explaining to me that seeing a cat in superposition of being alive or dead is not more improbable than seeing tomorrow’s weather with 50 % chance of sunshine and 50 % chance of snow. From Tal I learnt another important thing: to be happy with small results.

Furthermore, I have had many good discussions with the people from the lab, Alain Tapp, Dominic Mayers and Michel Boyer in particular.

For supporting me in one way or another I would like to thank Eliana, Frédéric & Ordia, Hassan & Mounia and Jean-François in Montreal, and my friends and family in the Netherlands, especially Alexander, Frank, Heleen, Liesbeth, Teake, and my parents.

Finally, I would like to thank Ana Paola for two and a half years of e-mail support and patience.

Bibliography

[BE89] BEAVER, D., "Multiparty protocols tolerating half faulty processors", In *Proc. CRYPTO 89* (1989), G. Brassard, Ed., Springer-Verlag, pp. 560–572, Lecture Notes in Computer Science No. 435.

[BG89] BEAVER, D. AND S. GOLDWASSER, "Multiparty computation with faulty majority", In *Proc. CRYPTO 89* (1989), G. Brassard, Ed., Springer-Verlag, pp. 589–590, Lecture Notes in Computer Science No. 435.

[BE90] BEAVER, D., *Security, Fault Tolerance, and Communication Complexity in Distributed Systems*, PhD thesis, Harvard University, 1990.

[BE91A] BEAVER, D., "Foundations of secure interactive computing", In *Proc. CRYPTO 91* (1991), J. Feigenbaum, Ed., Springer, pp. 377–391, Lecture Notes in Computer Science No. 576.

[BE91B] BEAVER, D., "Secure multiparty protocols and zero-knowledge proof systems tolerating a faulty minority", *Journal of Cryptology* **4**, 2 (1991), pp. 75–122.

[BE96] BEAVER, D., *personal communication*, June 1996.

[BE82] BEN-BASSAT, M., "Use of distance measures, information measures and error bounds in feature evaluation", In *Handbook of Statistics, Volume 2 — Classification, Pattern Recognition and Reduction of Di-*

mensionality (1982), P. R. Krishnaiah and L. N. Kanal, Eds., North-Holland, pp. 773–791.

[BB84] BENNETT, C. H. AND G. BRASSARD, "Quantum cryptography: Public key distribution and coin tossing", In *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing* (Bangalore, India, 1984), pp. 175–179.

[BBBSS92] BENNETT, C. H., F. BESSETTE, G. BRASSARD, L. SALVAIL AND J. SMOLIN, "Experimental quantum cryptography", *Journal of Cryptology* **5**, 1 (1992), pp. 3–28.

[BBE92] BENNETT, C. H., G. BRASSARD AND A. EKERT, "Quantum cryptography", *Scientific American*, October (1992), pp. 50–57.

[BBM92] BENNETT, C. H., G. BRASSARD AND N. MERMIN, "Quantum cryptography without Bell's theorem", *Physical Review Letters* **68**, 5 (1992), pp. 557–559.

[BMS96] BENNETT, C. H., T. MOR AND J. SMOLIN, "The parity bit in quantum cryptography", *Physical Review A* **54**, 3 (1996), p. 2675, also available at http://xxx.lanl.gov/ps/quant-ph/9604040.

[BBBV97] BENNETT, C. H., E. BERNSTEIN, G. BRASSARD AND U. VAZIRANI, "Strengths and weaknesses of quantum computation", *SIAM J. Computing* **26**, 5 (1997), pp. 1510–1523.

[BGW88] BEN-OR, M., S. GOLDWASSER AND A. WIGDERSON, "Completeness theorems for fault-tolerant distributed computing", In *Proc. 20th ACM Symp. on Theory of Computing* (Chicago, 1988), ACM, pp. 1–10.

[BV97] BERNSTEIN, E. AND U. VAZIRANI, "Quantum complexity theory", *SIAM J. Computing* **26**, 5 (1997), pp. 1411–1473.

[BBBGM98] BIHAM, E., M. BOYER, G. BRASSARD, J. VAN DE GRAAF AND T. MOR, "Security of quantum key distribution against all collective attacks", Tech. Rep. 9801022, LANL Quant-ph archives, 1998, available at http://xxx.lanl.gov/ps/quant-ph/9801022.

[BL82] BLUM, M., "Coin flipping by telephone", In *Proc. IEEE Spring COMPCOM* (1982), IEEE, pp. 133–137.

[BO97] BODOR, A., "Is quantum bit commitment really impossible", Tech. Rep. 9711040, LANL Quant-ph Archives, 1997, available at http://xxx.lanl.gov/ps/quant-ph/9711040.

[BCC88] BRASSARD, G., D. CHAUM AND C. CRÉPEAU, "Minimum disclosure proofs of knowledge", *Journal of Computer and System Sciences* **37**, 2 (1988), pp. 156–189.

[BC89] BRASSARD, G. AND C. CRÉPEAU, "Sorting out zero-knowledge", In *Proc. EUROCRYPT 89* (1989), J.-J. Quisquater and J. Vandewalle, Eds., Springer, pp. 181–191, Lecture Notes in Computer Science No. 434.

[BCJL93] BRASSARD, G., C. CRÉPEAU, R. JOZSA AND D. LANGLOIS, "A quantum bit commitment scheme provably unbreakable by both parties", In *Proc.* 34*th IEEE Symp. on Foundations of Comp. Science* (1993), IEEE, pp. 362–371.

[BCS96] BRASSARD, G., C. CRÉPEAU AND M. SÁNTHA, "Oblivious transfers and intersecting codes", *IEEE Transactions on Information Theory* **42**, 6 (1996), pp. 1769–1780.

[BC97] BRASSARD, G. AND C. CRÉPEAU, "Oblivious transfers and privacy amplification", In *Proc. EUROCRYPT 97* (1997), Springer-Verlag, pp. 334–347, Lecture Notes in Computer Science No. 1233.

[BCMS97] BRASSARD, G., C. CRÉPEAU, D. MAYERS AND L. SALVAIL, "Quantum attacks on classical bit commitment schemes", Tech.

Rep. 9712023, LANL Quant-ph archives, 1997, available at http://xxx.lanl.gov/ps/quant-ph/9712023.

[CA96] CANETTI, R., *Studies in Secure Multiparty Computation and Applications*, PhD thesis, The Weizmann Institute of Science, 1996, version dated: "June 1995; Revised: March 1996".

[CFGN96] CANETTI, R., U. FEIGE, O. GOLDREICH AND M. NAOR, "Adaptively secure multi-party computation", Tech. Rep. 682, Laboratory of Computer Science, Massachusetts Institute of Technology, 1996, Report on line available from http://www.wisdom.weizmann.ac.il/people/homepages/oded/papers.html.

[CH81] CHAUM, D., "Untraceable electronic mail, return addresses, and digital pseudonyms", *Communications of the ACM* **24** (1981), pp. 84–88.

[CCD88] CHAUM, D., C. CREPEAU AND I. DAMGÅRD, "Multi-party unconditionally secure protocols", In *Proc.* 20*th ACM Symp. on Theory of Computing* (Chicago, 1988), ACM.

[CG97] CIRAC, J. I. AND N. GISIN, "Coherent eavesdropping strategies for the 4 state quantum cryptography protocol", Tech. Rep. 9702002, LANL Quant-ph Archives, 1997, available at http://xxx.lanl.gov/ps/quant-ph/9702002.

[CT91] COVER, T. M. AND J. A. THOMAS, *Elements of Information Theory*, Wiley Series in Telecommunications, New York: John Wiley & Sons, 1991.

[CK88] CRÉPEAU, C. AND J. KILIAN, "Weakening security assumptions and oblivious transfer", In *Proc. CRYPTO 88* (1988), S. Goldwasser, Ed., Springer-Verlag, pp. 2–7, Lecture Notes in Computer Science No. 403.

[CR90] CRÉPEAU, C., *Correct and Private Reductions Among Oblivious Transfers*, PhD thesis, Massachusetts Institute of Technology, 1990.

[CR94] CRÉPEAU, C., "Quantum oblivious transfer", *Journal of Modern Optics* **41**, 12 (1994), pp. 2445–2454.

[CGT95] CRÉPEAU, C., J. VAN DE GRAAF AND A. TAPP, "Committed oblivious transfer and private multi-party computation", In *Proc. CRYPTO 95* (1995), Springer-Verlag, pp. 110–123, Lecture Notes in Computer Science No. 963.

[CR97] CRÉPEAU, C., "Efficient cryptographic protocols based on noisy channels", In *Proc. EUROCRYPT 97* (1997), Springer-Verlag, pp. 306–317, Lecture Notes in Computer Science No. 1233.

[DE85] DEUTSCH, D., "Quantum theory, the Church-Turing principle and the universal quantum computer", *Proc. Royal Society London* **Series A**, 400 (1985), pp. 97–117.

[EPR35] EINSTEIN, A., B. PODOLSKI AND N. ROSEN, "Can the quantum-mechanical description of physical reality be considered complete?", *Physical Review* **47** (1935), pp. 777–780, Reprinted in [WZ83].

[FE82] FEYNMAN, R., "Simulating physics with computers", *Int. Journal of Theoretical Physics* **21**, 6/7 (1982), pp. 467–488.

[FR93] FRANKLIN, M., *Complexity and Security of Distributed Protocols*, PhD thesis, Computer Science Department, of Columbia University, New York, 1993.

[FS90] FEIGE, U. AND A. SHAMIR, "Witness indistinguishable and witness hiding protocols", In *Proc.* 22*nd ACM Symp. on Theory of Computing* (Baltimore, MY, 1990), B. Awerbuch, Ed., ACM Press, pp. 416–426.

[FC94] FUCHS, C. A. AND C. M. CAVES, "Ensemble-dependent bounds for accessible information in quantum mechanics", *Physical Review Letters* **73**, 23 (1994), pp. 3047–3050.

[FC95] FUCHS, C. A. AND C. M. CAVES, "Mathematical techniques for quantum communication theory", *Open Systems and Information Dynamics* **3**, 3 (1995), pp. 345–356.

[FU95] FUCHS, C. A., *Distinguishability and Accessible Information in Quantum Theory*, PhD thesis, University of New Mexico, 1995.

[FU96] FUCHS, C. A., "Information gain vs. state disturbance in quantum theory", *to appear in Fortschritte der Physik* (1996), available at http://xxx.lanl.gov/ps/quant-ph/9611010.

[FGGNP97] FUCHS, C. A., N. GISIN, R. B. GRIFFITHS, C. S. NIU AND A. PERES, "Optimal eavesdropping in quantum cryptography. I. Information bound and optimal strategy", *Physical Review A* **56**, 2 (1997), pp. 1163–1172.

[GO94] GOLDREICH, O. AND Y. OREN, "Definitions and properties of zero-knowledge proof systems", *Journal of Cryptology* **7**, 1 (1994), pp. 1–32.

[GMR85] GOLDWASSER, S., S. MICALI AND C. RACKOFF, "The knowledge complexity of interactive proof-systems", In *Proc. 17th ACM Symp. on Theory of Computing* (Providence, 1985), ACM, pp. 291–304.

[GMR89] GOLDWASSER, S., S. MICALI AND C. RACKOFF, "The knowledge complexity of interactive proof-systems", *SIAM. J. Computing* **18**, 1 (1989), pp. 186–208.

[GR95] VAN DE GRAAF, J., "The matchmaking problem", *CWI Quarterly* **8**, 2 (1995), pp. 129–146.

[HR70] HELLMAN, M. E. AND J. RAVIV, “Probability of error, equivocation, and the Chernoff bound”, *IEEE Transactions on Information Theory* **IT-16**, 4 (1970), pp. 368–372.

[HE76] HELSTROM, C. W., *Quantum Detection and Estimation Theory*, Mathematics in Science and Engineering, vol. 123, Academic Press, 1976.

[HU89] HUGHES, R. I. G., *The Structure and Interpretation of Quantum Mechanics*, Harvard University Press, 1989.

[IS95] ISHAM, C., *Lectures on Quantum Theory*, Imperial College Press, 1995.

[JO94] JOZSA, R., “Fidelity for mixed quantum states”, *Journal of Modern Optics* **41**, 12 (1994), pp. 2315–2323.

[KA67] KAILATH, T., “The divergence and Bhattacharyya distance measures in signal selection”, *IEEE Transactions on Communication Technology* **COM-15**, 1 (1967), pp. 52–60.

[KE97] KENT, A., “Unconditionally secure quantum bit commitment”, Tech. Rep. 9711023, LANL Quant-ph Archives, 1997, available at http://xxx.lanl.gov/ps/quant-ph/9711023.

[KI88] KILIAN, J., “Founding cryptography on oblivious transfer”, In *Proc. 20th ACM Symp. on Theory of Computing* (Chicago, 1988), ACM, pp. 20–31.

[KI90] KILIAN, J., *Uses of Randomness in Algorithms and Protocols*, MIT Press, 1990.

[LO97] LO, H.-K., “Insecurity of quantum secure computations”, Tech. Rep. 9611031, LANL Quant-ph Archives, 1997, available at http://xxx.lanl.gov/ps/quant-ph/9611031.

[LC97A] LO, H.-K. AND H. F. CHAU, "Is quantum bit commitment really possible?", *Physical Review Letters* **78** (1997), pp. 3410–3413.

[LC97B] LO, H.-K. AND H. F. CHAU, "Why quantum bit commitment and ideal quantum coin tossing are impossible", Tech. Rep. 9711065, LANL Quant-ph archives, 1997, available at http://xxx.lanl.gov/ps/quant-ph/9711065.

[LV93] LI, M. AND P. VITÁNYI, *An introduction to Kolmogorov Complexity and its Applications*, Springer, 1993.

[LI56] LINDLEY, D. V., "On a measure of the information provided by an experiment", *Ann. Math. Statist.* **27** (1956), pp. 986–1005.

[MA97] MAYERS, D., "Unconditionally secure quantum bit commitment is impossible", *Physical Review Letters* **78** (1997), pp. 3414–3417.

[ME96] MERMIN, N. D., "The Ithaca Interpretation of Quantum Mechanics", Tech. Rep. 9609013, LANL Quant-ph Archives, 1996, available at http://xxx.lanl.gov/ps/quant-ph/9609013.

[MR91] MICALI, S. AND P. ROGAWAY, "Secure Computation", preliminary version in CRYPTO 91.

[MO97] MOR, T., *Quantum Memory in Quantum Cryptography*, PhD thesis, Technion, Haifa, 1997.

[MS97] MAYERS, D. AND L. SALVAIL, *personal communication*, October 1997.

[NOVY92] NAOR, M., R. OSTROVSKY, R. VENKATESAN AND M. YUNG, "Perfect zero-knowledge arguments for np can be based on general complexity assumptions", In *Proc. CRYPTO 92* (1992), E. F. Brickell, Ed., Springer-Verlag, pp. 196–214, Lecture Notes in Computer Science No. 740.

[OR87] OREN, Y., "On the cunning power of cheating verifiers: Some observations about zero knowledge proofs (extended abstract)", In *28th Annual Symposium on Foundations of Computer Science* (Los Angeles, California, 1987), IEEE, pp. 462–471.

[OR96] ORŁOWSKI, A., "Measures of distance between quantum states", In *Proceedings of the Fourth Workshop on Physics and Computation — PhysComp '96* (Boston, 1996), New England Complex Systems Institute, pp. 239–242.

[PE93] PERES, A., *Quantum Theory: Concepts and Methods*, Kluwer Academic Publishers, 1993.

[PR97] PRESKILL, J., "Reliable quantum computer", Tech. Rep. 9705031, LANL Quant-ph Archives, 1997, available at http://xxx.lanl.gov/ps/quant-ph/9705031.

[RA81] RABIN, M., "How to exchange secrets by oblivious transfer", Tech. Rep. TR-81, Harvard Aiken Computation Laboratory, 1981.

[RB89] RABIN, T. AND M. BEN-OR, "Verifiable secret sharing and multiparty protocols with honest majority", In *STOC89* (1989), pp. 73–85.

[RS72] REED, M. AND B. SIMON, *Methods of Modern Mathematical Physics — Part I: Functional Analysis*, Academic Press, 1972.

[RO91] ROGAWAY, P., *The Round Complexity of Secure Protocols*, PhD thesis, Massachusetts Institute of Technology, 1991.

[SC91] SCHNORR, C. P., "Efficient signature generation by smart cards", *Journal of Cryptology* **4** (1991), pp. 161–174.

[SC35] SCHRÖDINGER, E., "Die gegenwärtige Situation in der Quantenmechanik", *Naturwissenschaften* **23** (1935), pp. 807–812, 823–828, 844–849, English translation in [WZ83].

[SRA81] SHAMIR, A., R. L. RIVEST AND L. M. ADLEMAN, "Mental poker", In *The Mathematical Gardner* (Belmont, California, 1981), D. Klarner, Ed., Wadsworth, pp. 37–43.

[SH48] SHANNON, C. E., "A mathematical theory of communication", *Bell Sys. Tech. J.* **27** (1948), pp. 379–423 and 623–656.

[SH49] SHANNON, C. E., "Communication theory of secrecy systems", *Bell Sys. Tech. J.* **28** (1949), pp. 657–715.

[SH96] SHOR, P. W., "Fault-tolerant quantum computation", In *37th Annual Symposium on Foundations of Computer Science* (Burlington, Vermont, 14–16 Oct. 1996), IEEE, pp. 56–65.

[ST97] STEANE, A., "Quantum computing", Tech. Rep. 9708022, LANL Quant-ph Archives, 1997, available at http://xxx.lanl.gov/ps/quant-ph/9708022.

[SU86] SUDBERY, *Quantum Mechanics and the Particles of Nature — an Outline for Mathematicians*, Cambridge University Press, 1986.

[SV97] SVOZIL, K., "The Church-Turing thesis as a guiding principle for physics", Tech. Rep. 9710055, LANL Quant-ph archives, 1997, available at http://xxx.lanl.gov/ps/quant-ph/9710055.

[TO71] TOUSSAINT, G. T., "Some functional lower bounds on the expected divergence for multihypothesis pattern recognition, communication, and radar systems", *IEEE Transactions on Systems, Man, and Cybernetics* **SMC-1** (1971), pp. 384–385.

[TO72] TOUSSAINT, G. T., "Comments on 'The divergence and Bhattacharyya distance measures in signal selection' ", *IEEE Transactions on Communication Technology* **COM-20** (1972), p. 485.

[UH76] UHLMANN, A., "The 'transition probability' in the state space of a *-algebra", *Reports on Mathematical Physics* **9** (1976), pp. 273–279.

[WO81] WOOTTERS, W., "Statistical distance and Hilbert space", *Physical Review D* **23** (1981), pp. 357–362.

[WZ82] WOOTTERS, W. AND W. ZUREK, "A single quantum cannot be cloned", *Nature* **299** (1982), pp. 802–803.

[WZ83] WHEELER, J. AND W. ZUREK, *Quantum Theory and Measurement*, Princeton University Press, 1983.

[YA82A] YAO, A. C.-C., "Protocols for secure computations", In *Proc. 23rd IEEE Symp. on Foundations of Comp. Science* (Chicago, 1982), IEEE, pp. 160–164.

[YA82B] YAO, A. C.-C., "Theory and application of trapdoor functions", In *Proc. 23rd IEEE Symp. on Foundations of Comp. Science* (Chicago, 1982), IEEE, pp. 80–91.

[YA95] YAO, A. C.-C., "Security of quantum protocols against coherent measurements", In *Proc. 27th ACM Symp. on Theory of Computing* (1995), ACM, pp. 67–82.