Two-Prover Bit-Commitments:

Classical, Quantum and Non-Signaling

Max Fillinger

Two-Prover Bit-Commitments:
Classical, Quantum and Non-Signaling

Proefschrift
ter verkrijging van
de graad van Doctor aan de Universiteit Leiden
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 19 maart 2019
klokke 10:00 uur

door

Maximilian Johannes Fillinger
geboren te Wuppertal, Duitsland,
in 1988

Promotor:

Prof. dr. Serge Fehr (CWI, Amsterdam & Universiteit Leiden)

Samenstelling van de promotiecommissie:

Dr. Stacey Jeffery (CWI, Amsterdam)

Prof. dr. Adrian Kent (University of Cambridge)

Prof. dr. Bart de Smit (Universiteit Leiden)

Prof. dr. Aad van der Vaart (Universiteit Leiden)

Prof. dr. Stefan Wolf (Università della Svizzera italiana, Lugano)

This work was supported by the *NWO Free Competition* grant 617.001.203
and carried out at CWI, Amsterdam.

Universiteit Leiden

CWI

Contents

Chapter 1

Introduction

1.1 Background

1.1.1 Cryptography with Mutually Distrusting Parties

Cryptology began as the art of designing and breaking ciphers that allow two parties to communicate in such a way that an outsider who observes the messages does not learn the actual content. During the second half of the 20th century, cryptology has changed significantly: it became more rigorous and scientific, and also expanded its scope.

Replacing mechanical with digital ciphers made it possible to create much stronger encryption schemes. Also, entirely new concepts of encryption were introduced: While previous encryption schemes all required both parties to have the same secret key, *public-key cryptography* makes it possible to generate a key pair consisting of a *public key* and a *private key* so that messages can be encrypted using the public key but decrypted *only* with the private key. Thus, the receiver can safely publish the public key and receive encrypted messages from anyone.

In traditional cryptographic applications, we want to protect two communicating parties from outside attackers. However, modern cryptography also considers situations where *mutually distrusting* parties want to cooperate in a secure way, meaning that no party needs to disclose more information than strictly necessary and that each party is protected if the other one turns out to be dishonest. A classic example is the *Millionaires' Problem* where two millionaires want to know who is richer, without disclosing any additional information about their wealth to the other one. Situations like this are considered in *multi-party computation* where n players each hold one input x_i to a function f and want to compute $f(x_1, \ldots, x_n)$ so that the other parties do not learn anything about x_i except for what they can deduce from the function value.

Another example of such a situation is a so-called *coin-flipping* protocol, i.e., a protocol between two parties, called Alice and Bob, that generates a uniformly random bit b which is output to both of them. A coin-flipping protocol needs to ensure that neither party can bias the "coin": if one party is dishonest and deviates from the protocol, the output distribution for the honest party must still be a uniformly random bit.

If we can ensure that both parties send their messages simultaneously, this problem is easy to solve: each of them sends a uniformly random bit to the other and then outputs the XOR of the bit they sent and the one they received. The output will then be a uniformly random bit, as long as at least one party is honest. However, if we cannot ensure that the messages are sent simultaneously, this protocol is not secure. Suppose that Bob receives Alice's message before sending his own. If he is dishonest, he can then choose his bit depending on the one he received from Alice and enforce any output distribution he likes for the protocol.

In both of the applications that we discussed so far, all participants have the same security concerns. However, there are also applications where that is not the case, such as zero-knowledge proofs: Suppose that one party (called the *prover*) knows a satisfying assignment for a Boolean formula. In a *zero-knowledge proof system* for the satisfiability problem, the prover wants to convince another party (the *verifier*) that a formula is satisfiable, but he does not want to reveal any further information (such as the satisfying assignment). The verifier on the other hand wants to be certain that a dishonest prover cannot deceive him about the satisfiability of the formula, but is not concerned with the secrecy of the satisfying assignment.

Traditional cryptographic primitives, like encryption or message authentication, can still be useful in the context of mutually distrusting parties – for example, multi-party computation protocols generally require that each participant has a confidential and authenticated channel to each other participant. However, on their own, these tools are not sufficient for building protocols for coin-flipping, multi-party computation, or zero-knowledge proofs. New cryptographic primitives were required to solve these problems. In the next section, we introduce such a cryptographic primitive.

1.1.2 Commitment Schemes

A *commitment scheme* is a cryptographic primitive which was first formally defined in 1988 by Gilles Brassard, David Chaum and Claude Crépeau in order to build a zero-knowledge protocol for proving that a Boolean formula has a satisfying assignment [BCC88]. Since determining the decidability of a Boolean formula is an NP-complete problem, this also shows that there are zero-knowledge protocols for all problems in the complexity class NP. However, the general idea of bit commitments has already been used in the early 80s in work on coin-flipping protocols and similar applications [SRA81, Blu82,

EGL83].

A commitment scheme allows a *prover* to select an element s of a publicly known set D so that he can later reveal it to another party, called the *verifier*, in such a way that the verifier can be certain that the revealed value is the same as the originally selected one. If D is the set $\{0, 1\}$ we speak of a *bit-commitment* scheme. If we want to emphasize that a commitment scheme has a larger domain, we call it a *string-commitment* scheme (even when the elements of D are not actually strings). If the domain has size (at least) 2^n, we call the commitment scheme an n-bit string commitment scheme.

More formally, a commitment scheme consists of two interactive protocols between the prover and the verifier, called the *commit phase* and *opening phase*. The commit phase takes as input an element $s \in D$ from the prover and no input from the verifier. It outputs some state information to the two parties. The verifier's state is commonly called the *commitment* to s. The opening phase takes the state information of the prover and verifier as input, and outputs an element $s' \in D$ or the failure symbol $\perp$ to the verifier; we say that the prover opened the commitment to s', or if the output is $\perp$, that he failed to open the commitment. The opening phase is often non-interactive in the sense that the prover sends some *opening information* to the verifier who then determines the output by local computation.

A basic requirement of a commitment scheme is that if both parties follow the protocols, the input s to the commit phase and the output s' of the opening phase are identical. This property is called *completeness*. Informally, the security requirements for a commitment scheme are as follows:

- The *hiding* property: by the execution of the commit phase, the verifier does not learn the prover's input s. In particular, this holds even if the verifier is dishonest and deviates from the commit protocol in arbitrary ways.
- The *binding* property: after the commit phase has been executed, there is *at most one* element $s' \in D$ that the prover can open to in the opening phase. This holds even if the prover is dishonest and deviates from the protocols in arbitrary ways.

To see how bit-commitment schemes can be used for coin-flipping, consider the following protocol, which does not require that the parties send their messages simultaneously: First, Alice samples a random bit and commits to it. Then, Bob samples a random bit and sends it to Alice. Finally, Alice opens the commitment and both parties output the XOR of the two random bits. In this protocol, the hiding property of the commitment scheme ensures that Bob does not know Alice's bit before sending his bit. The binding property ensures that Alice can not choose her bit after learning Bob's: after the commit phase, there is at most one value she can open to.

There remains one subtle issue: how should the case where Alice fails to open the commitment be handled? After Alice receives the bit from Bob, she

knows what the outcome of the coin flipping protocol will be, but Bob does not. A dishonest Alice could thus decide whether or not to open the commitment depending on the outcome. If one outcome of the coin-flipping protocol is considered favorable for Alice, and the other unfavorable, it makes sense to stipulate that the protocol outputs the unfavorable outcome in that case. But then, the protocol only implements *weak* coin-flipping, meaning that dishonest parties *can* bias the output distribution, but only towards the outcome that is unfavorable for them.

1.1.3 Capabilities of Adversaries

Rigorous security claims of cryptographic schemes always require some model that specifies the capabilities and limitations of the honest parties and of the adversaries. Most commonly, the honest parties and adversaries are modeled as Turing machines that are limited to *efficient* computations. Efficiency here is understood asymptotically: algorithms and protocols are parametrized by a *security parameter* n, and the running time for the honest parties must be polynomial in n, while the adversaries must not be able to break the scheme in time polynomial in n. Security proofs in this model typically use *computational hardness assumptions*, i.e., assumptions that certain computational problems, such as factoring large integers, can not be solved efficiently. We then speak of computational security. No such assumptions have been proven.

The security of cryptographic primitives that are relied on in practice is typically based on problems that have been studied for a long time, without an efficient solution being discovered. This is considered empirical evidence that the problem in fact does not have an efficient solution.

The security of many popular cryptosystems (e.g., RSA and the Diffie-Hellman Key Exchange) is based on problems that are believed to have no efficient solution in the Turning machine model, but do have an efficient solution on a quantum computer [Sho97]. Since quantum computers might become practical in the near future, there is a lot of interest in *quantum-safe* (also called *post-quantum*) cryptosystems which are hard to break even with a quantum computer, but can be executed on classical computers. Formally, one would then model the honest parties as Turing machines and the adversary as a family of quantum circuits with a polynomial number of gates.

In the *information-theoretic* model, we remove the efficiency requirement from the adversary – we say that the adversary is *computationally unbounded*. For example, an encryption scheme would only be considered secure in the information-theoretic model if the adversary cannot recover any amount of information about the message, even given unlimited time.[1] In other words, the plaintext is (almost) statistically independent of the information that the adversary has. If a scheme is proven to be secure in the information-theoretic

[1] One can also relax this requirement and allow the adversary to obtain a very small amount of information.

model, it is secure against adversaries with unlimited computing power. We also call such schemes *unconditionally* secure.

Most schemes that are used in practice are not unconditionally secure. For example, if we consider a symmetric cipher with an n-bit key, a computationally unbounded adversary could decrypt a given ciphertext under every possible key. He then knows that one of the 2^n outputs must be the original plaintext. If the size of the plaintext space is greater than 2^n, this gives the adversary a significant amount of information. Furthermore, if the plaintext is known to be, e.g., English text, the adversary is likely able to rule out all but one of the candidate keys. Unconditionally secure encryption is only possible if the key has at least as much entropy as the message, as Claude Shannon proved in 1949 [Sha49].

Note that up to now, we discussed the *standard model* where the participants in the scheme or protocol can only communicate classical information via a completely unsecured channel. More results can be achieved if the honest parties have access to additional resources. An example is Quantum Key Distribution (QKD), introduced by Charles Bennett and Gilles Brassard in 1984 [BB84]. It allows two parties to securely establish a shared random key using an authenticated classical channel and a completely insecure channel for quantum information. The generated key can then be used for a classical information-theoretically secure encryption scheme like the One-time Pad. QKD offers information-theoretic security beyond the Shannon bound, but it requires that the honest parties are able to produce, transmit, and measure quantum states, e.g., in the form of polarized photons.[2]

As another example, if two parties can communicate via a noisy channel, they can also transmit messages securely, as Aaron Wyner proved in [Wyn75]. The noisy channel here is modeled as a channel that flips every bit that is sent with a known probability ε and leaves it unchanged otherwise. In particular, if an adversary taps the channel, the bits he receives are flipped with the same probability, but independently of the bits received by the intended recipient. Cryptographic primitives that are useful for cryptography with mutually distrusting parties, like oblivious transfer, can be implemented as well using a noisy channel [CK88].

One can also impose non-computational restrictions on the adversary, such as limited classical memory [CM97, CCM98], or limited or noisy quantum memory [DFSS05, WCSL10]. Storing quantum information is a difficult problem, and thus schemes where an adversary needs to store large amounts of quantum information while the honest parties can measure the quantum states as they arrive are of interest. A different kind of restriction is to split the prover into two separate parties and restrict the communication between them. We discuss this in more detail in Section 1.2.2.

[2]Note that QKD does not require the honest parties to have a quantum computer or quantum memory. A channel for transmitting quantum information suffices.

1.2 Two-Prover Commitment Schemes

1.2.1 (In)security of Commitment Schemes

The existence of commitment schemes in the computational model follows from very week assumptions: if a pseudo-random generator[3] exists, then there exists a commitment scheme that is both hiding and binding [Nao91]. The existence of pseudo-random generators has not been proven, although there are many candidates in both theory and practice. But if they do not exist, then there are no secure cryptographic schemes in the computational model [IL89].[4] Or conversely, if computational cryptography is at all possible, then commitment schemes exist.

Let us now consider commitment schemes in the information-theoretic model, typically referred to as *unconditionally secure* commitment schemes. It is well known that in the standard communication model, bit-commitment schemes can not be both unconditionally hiding and unconditionally binding. Consider a bit-commitment scheme that is unconditionally binding. It is easy to see that a computationally unbounded dishonest verifier can break the hiding property as follows.

First, both parties execute the commit phase, which outputs state information $state_P$ and $state_V$ to the prover and verifier, respectively. If the opening phase is executed with inputs $state_P$ and $state_V$, the output is the bit b that the prover committed to. The binding property requires that the prover can open to at most one bit, so if $state_P$ is replaced with a different input, the output will be either $\perp$ or b (except possibly with some small probability). A computationally unbounded verifier can simulate the opening phase for every possible value of $state_P$, and thus determine the bit b that the prover committed to.

Since unconditionally secure bit-commitment is impossible in the standard communication model, we have to move to a different one. There was some hope that unconditionally secure bit-commitment schemes could be achieved using quantum communication, but eventually, an impossibility result was proved also in that setting [May97, LC97]. If the dishonest players have *bounded memory* [CCM98], then unconditionally secure bit-commitment is possible.[5] The same holds in the *bounded quantum storage* model where adversaries have unlimited classical memory, but only a limited amount of quantum memory [DFSS05].

[3] A function that maps a short string of random bits to a longer string so that the longer string cannot be distinguished from a truly random string in polynomial time.

[4] The cited paper argues that computational cryptography cannot exist if there are no one-way functions; it is possible to implement a pseudo-random generator with one-way functions.

[5] The topic of the cited paper is not bit-commitment, but a different cryptographic primitive called *oblivious transfer*. However, bit-commitment schemes can be implemented using oblivious transfer.

1.2.2 Adding a Second Prover

Another way to circumvent the impossibility result is to *split up* the prover into two entities that are assumed to be unable to communicate with each other. This so-called *two-prover setting* was introduced by Michael Ben-Or, Shafrira Goldwasser, Joe Kilian and Avi Wigderson [BGKW88]. The provers *can* communicate before the start of the commit phase to generate shared randomness, and, if they are dishonest, agree on a cheating strategy, but from the start of the commit phase until the end of the opening phase, they cannot communicate.

As an example for a scheme in this model, we consider a scheme that was introduced by Jean-Raymond Simard in [Sim07] and further explored by Claude Crépeau, Louis Salvail, Simard and Alain Tapp in [CSST11]. This scheme will also play an important role in the remainder of this thesis. We call this scheme $\mathcal{CHSH}^q$ where q is a prime power. It works as follows: Let b be the bit that the provers want to commit to and r a uniformly random element of the finite field $\mathbb{F}_q$ that the provers agree on as shared randomness before the execution of the commit phase. In the commit phase, the verifier V sends a uniformly random element a of the finite field $\mathbb{F}_q$ to the first prover P, who sends back $x = a \cdot b + r$. In the opening phase, the second prover Q sends the bit b and $y = r$ to V. Then, V outputs b if $x - y = a \cdot b$, and the failure symbol $\perp$ otherwise.[6]

Let us verify that this scheme satisfies the properties that we want a commitment scheme to have. The *hiding* property requires that a (possibly dishonest) verifier can learn nothing about the committed bit before the opening phase. The scheme is *perfectly hiding* because P's message x is always a uniformly random field element, independent of the value of b.

Completeness requires that if all parties are honest, the verifier opens to the bit that the provers committed to. It is easy to see that this requirement is satisfied.

The *binding* property requires that even dishonest provers can open to at most one value. Let a and x be the messages exchanged between V and P in the commit phase. Since we consider dishonest provers, x does not have to be computed as specified in the protocol – in fact, the dishonest provers do not need to have any specific bit b in mind while executing the commit phase. The following argument works for any value of x and makes no assumptions on how it is computed. If Q wants to open to $b = 0$, he needs to send $b = 0$ and $y = x$ to V; if he wants to open to $b = 1$, he needs to send $b = 1$ and $y = x - a$. Thus, if Q can open to *both* bits, it follows that he knows a. But a was sent only to P, and by assumption, P and Q cannot communicate. Therefore, Q can only open to both bits if he correctly guesses a. This happens only with

[6] This version of CHSH^q differs slightly from the version we use later on, where Q does not send the bit b. In that case, V has to check whether the equation $x - y = a \cdot b$ holds for $b = 0$ or $b = 1$.

probability q^{-1}.

We emphasize that $\mathcal{CHSH}^q$ as described above and as analyzed in previous work is a bit-commitment scheme, but it can be naturally extended to a string-commitment scheme by letting b be an arbitrary element of $\mathbb{F}_q$. This extension has been used as part of a larger protocol in [LKB+15], but prior to our work in [FF16], it has not been analyzed as a stand-alone string-commitment scheme.

Analyzing it as a string commitment scheme turns out to be somewhat subtle: for instance, it is not clear a priori what the right formal definition of the binding property is for a *string-commitment* scheme in the two-prover setting. This thesis will answer those kinds of questions.

1.2.3 Capabilities of the Provers

As discussed above, the security of $\mathcal{CHSH}^q$ relies on the assumption that the provers cannot communicate. However, it turns out that what this precisely means is more subtle than the previous section makes it appear. As in [BGKW88], we implicitly assumed in the argument above that the only type of information that the provers can share before the commit phase is classical information. However, as pointed out in [Sim07, CSST11], the argument falls apart when we consider provers that share an entangled quantum state. The reason for that is *non-locality*, one of the counterintuitive properties of quantum mechanics, which is studied by means of Bell inequalities and non-local games [EPR35, Bel64, CHSH69].

Formally, the point where the argument from the previous section fails in the quantum case is the part where we conclude that a prover who can *choose* to output either x or $x - a$ must also know a. This does not follow in the quantum case. It is generally not possible to measure (i.e., extract information from) a quantum state without irreversibly *changing* the state. If Q could produce $y = x$ using one measurement and $y = x - a$ using another measurement, then he could open to any bit he likes, but it does not follow that he could produce *both* x and $x - a$ at the same time. Hence, it does not follow that he knows a.

[CSST11] shows that $\mathcal{CHSH}^{2^n}$ is secure in the quantum case. On the other hand, the same paper also shows that a slight variation of this scheme is secure only against classical adversaries: an error-tolerant version where V only checks that 85% of the bits in x and y or x and $y + a$ are equal is secure against classical adversaries, but completely insecure against quantum adversaries. This is a consequence of the fact that players with an entangled quantum state can win the non-local game known as the CHSH game with probability ≈ 0.85 (see Section 5.2.4). This connection with the CHSH game is the reason why we refer to the commitment scheme as $\mathcal{CHSH}^q$.

Thus, the seemingly sole assumption that the provers cannot communicate during the execution of the protocol is actually underspecified. To truly base security *only* on the no-communication assumption, one needs to consider

general non-signaling adversaries, i.e., adversaries that are equipped with a hypothetical resource that allows them to correlate their behavior in arbitrary ways as long as no communication is implied. Such hypothetical resources are known as non-local boxes.

1.3 Relativistic Cryptography

1.3.1 Enforcing the No-Communication Assumption

There are two-prover commitment schemes that are secure if the two provers cannot communicate and can correlate their behavior only through shared randomness or entangled quantum states. This leaves open the question of how one might actually prevent the provers from communicating. In *relativistic commitment schemes*, we exploit the fact that information does not travel faster than light, and thus, messages from one prover to the other arrive only with some delay.

In [BC96], Gilles Brassard and Claude Crépeau briefly discuss the idea, communicated by Louis Salvail, of applying special relativity to two-prover bit-commitment schemes that rely on the no-communication assumption, as described in Section 1.2.2. If the provers are n light-seconds apart, the laws of physics ensure that the commitment is binding, *but only for a limited time*: The commitment will "live" for n seconds, starting when the first message from the verifier arrives at a prover. If the provers open within this time-span, the verifier can be assured that the provers can open to at most one value, since the no-communication assumption is guaranteed by the fact that information can not be transmitted faster than the speed of light. If the commitment is not opened within this time-span, it is possible that the provers have communicated with each other. Thus, they might be able to open to multiple values.

Adrian Kent introduced the concept of *relativistic* commitment schemes that can remain binding indefinitely as long as the provers can only communicate with some delay [Ken99, Ken05]. This is achieved by introducing an additional *sustain phase* between the commit and opening phase. During this phase, additional communication between the verifier and the provers takes place that is meant to ensure that the commitment remains binding. The hiding property should still apply during this phase.

1.3.2 Previous and Related Work

The first relativistic commitment scheme was introduced by Kent in [Ken99]. He argues that the scheme is secure against classical adversaries, and he reasons that dishonest provers with quantum capabilities can not break the commitment scheme on its own, but might gain an advantage if it is part of a larger protocol.

A major issue with Kent's original scheme is that the length of the messages that need to be communicated in each round of the sustain phase grows exponentially in the number of rounds. Furthermore, the security arguments are rather informal and not in terms of rigorous definitions. As such, it cannot be considered a mathematical security proof. However, his work demonstrated that it is possible, or at least plausible, to base the security of a commitment scheme on the fact that information does not travel faster than light and thus laid the foundation for subsequent work in this area.

In [Ken05], he introduced an improved scheme where the same number of bits is communicated in every round. Additionally, the security proofs are more formal, using the sum-binding definition (see Definition 2.14). However, the results are mostly of an asymptotic nature and clearly not practical, although some concrete parameter choices are also discussed.

Kent also considered commitment schemes that involve quantum communication. Concretely, he presented a scheme where the players transmit quantum states instead of classical bits [Ken11], and a scheme where the verifier sends a quantum state to the provers, and the provers return classical bits [Ken12]. These schemes do not require a sustain phase. Furthermore, the former one requires only one prover and one verifier. The latter requires a prover that is split into *three* agents.

A security proof for the latter scheme was later published in a joint work of Sarah Croke and Kent [CK12]. See [KTHW13] for an alternative proof. This scheme was implemented in 2013 by Tomaso Lunghi, Jędrzej Kaniewski, Felix Bussières, Raphael Houlmann, Marco Tomamichel, Adrian Kent, Nicolas Gisin, Stephanie Wehner and Hugo Zbinden [LKB$^+$13, Kan15].

In later work [Ken13], Kent isolated the "game" whose hardness underlies the security of the quantum bit-commitment schemes. In the *summoning* problem, Alice gives a quantum state to Bob. The description of the state is known to her, but not to Bob. She will later ask Bob to "summon" it to some point. For simplicity, one may assume that she fixes two points P_0 and P_1 that are known to Bob, and selects one of them uniformly at random. After receiving Alice's summoning request, Bob has a short amount of time to produce a quantum state at this point that Alice can not distinguish from her original state.

Kent shows that this task is (in general) impossible by combining the no-cloning theorem [Par70, WZ82], which states that it is in general not possible to create a perfect copy of quantum states, with special relativity: Bob cannot send copies of the state to both locations due to the no-cloning theorem. But if the two points are far enough apart, he cannot position the state in such a way that he can always "summon" it to the point Alice chooses within the time constraint. Further work on the summoning problem can be found in [AK16, Ken18].

Deterministic quantum bit commitment schemes that do not rely on secret randomness have been proposed by Emily Adlam and Kent [AK15a]. The se-

curity proofs are based on relativity and monogamy of entanglement [CKW00].

For many quantum-cryptographic tasks, such as Quantum Key Distribution, device-independent protocols have been discovered [MY98]. In such protocols, the participants do not even need to trust the devices that carry out the preparation and measurement of quantum states. Quantum bit-commitment schemes with this property have been discovered as well [AK15b].

But there also has been progress for classical schemes: Lunghi, Kaniewski, Bussières, Houlman, Tomamichel, Wehner and Zbinden proposed a new multi-round commitment scheme where the honest parties communicate classically [LKB$^+$15]. They provided a rigorous, non-asymptotic, security analysis with respect to the sum-binding definition. However, their analysis only guarantees an error term that worsens double-exponentially in the number of rounds of communication. Furthermore, their security proof only applies to classical provers, i.e., provers with no quantum capabilities.

The fact that information does not travel faster than light has also been applied in related areas of cryptography: Roger Colbeck and Adrian Kent introduced variable-bias coin-tossing schemes, where the probability distribution of the outcome is secretly determined by one of the parties [CK06, Col06]. On the other hand, Colbeck showed that unconditionally secure two-party computation is not possible (for most functions) even with the combined power of quantum information and relativity [Col06, Col07].

A relativistic quantum key distribution scheme has also been proposed [RKKM14]. While QKD schemes like the one introduced by Charles Bennet and Gilles Brassard [BB84] can be proven secure only on the basis of quantum mechanics, implementations can often be broken due to imperfections in the physical apparatus (see e.g. [LWW$^+$10]). While not being device-independent, the relativistic scheme has a higher tolerance for the type of imperfections that occur in practice, and is more efficient than device-independent protocols.

A different conjectured application of relativity is *position-based quantum cryptography*, also known as *quantum tagging*. Here, a prover wants to demonstrate that he is at a specific location. This claim is checked by a set of verifiers which are positioned at different points in space. The verifiers use the response time of the prover to determine his position. However, just sending a nonce to the prover and requiring him to send it back is insufficient: a group of adversaries could pretend to be a single prover at the correct position, even though none of them are actually there. Therefore, techniques to prevent this attack using quantum information have been studied.

The first position-based cryptography scheme, patented in 2006 [KMSB06], was invented by Kent, William Munro, Timothy Spiller, and Raymond Beausoleil. Robert Malaney was the first to publish such a scheme in the scientific literature in 2010 [Mal10a, Mal10b]. None of these schemes were proven secure, and, in fact, were later broken: in 2011, Kent, Munro and Spiller published a proof that all schemes proposed so far were insecure if the dishonest provers have shared entanglement [KMS11]. Harry Buhrman, Nishanth Chan-

dran, Serge Fehr, Ran Gelles, Vipul Goyal, Rafail Ostrovsky and Christian Schaffner proved a general impossibility result: position-based cryptography schemes cannot be secure if dishonest provers can have an unlimited amount of pre-shared entanglement [BCF+11]. Further research in the field of position-based cryptography aimed at finding a scheme where dishonest provers need large amounts of pre-shared entanglement compared to the number of qubits that the honest parties need to exchange.

1.4 Contributions of this Thesis

This thesis is based on the following publications and follow-up work:

- Serge Fehr and Max Fillinger. *Multi-Prover Commitments Against Non-Signaling Attacks.* In *Advances in Cryptology - CRYPTO 2015, part II*, pages 403-421.
  Also presented at *QCRYPT 2015* .
- Serge Fehr and Max Fillinger. *On the Composition of Two-Prover Commitments, and Applications to Multi-round Relativistic Commitments.* In *Advances in Cryptology - EUROCRYPT 2016, part II*, pages 477-496. An earlier version was presented at *QCRYPT 2015*. An extended version is available at
  `https://arxiv.org/abs/1507.00240`.

The author published two other papers during his PhD studies. These are not featured in this thesis because they are about very different subjects and would not allow for a coherent presentation. A full list of publications can be found on page 109.

1.4.1 New Definitions for the Binding Property

Finding good security definitions is a crucial part of cryptography. A good definition has to capture the intuitive notion of the desired security property in a precise mathematical way. It should be as strong as possible, but it also should be possible to prove that this definition can be met, or to reduce it to standard hardness assumptions in the case of computational security.

In the case of two-prover commitment schemes, the main topic of this thesis, the information-theoretic hiding property is straightforward to define: the commitment that the verifier receives should be distributed independently of the bit or string that the provers commit to. One can also relax this definition somewhat by allowing the distributions to have some small statistical distance.

However, defining the information-theoretic binding property turns out to be tricky: a somewhat accepted definition is the sum-binding property (see Definition 2.14). This definition requires that, for any strategy that the

dishonest provers may use, it holds that $p_0 + p_1 \leq 1 + 2\varepsilon$, where p_b is the probability that the provers successfully open to b. However, this definition suffers from several limitations. An immediately obvious limitation is that this definition only applies to bit-commitment schemes, and there is some ambiguity in how to extend it to string-commitment schemes. It also does not fully capture the intuitive requirement of the binding property: Suppose that a scheme rejects all opening attempts of dishonest provers with probability 1/2 and allows dishonest provers to open to an arbitrary bit with probability 1/2. This scheme then has a "perfect" parameter of $\varepsilon = 0$ (see Remark 3.13), but the intuitive requirement that the provers should only be able to open a commitment to at most one value is violated with probability 1/2. Finally, it turns out that the sum-binding definition is inconvenient to work with, e.g., it does not seem to compose well.

Contribution 1.1. *We propose several new definitions for the binding property of bit and string commitment schemes and analyze their relations with each other and with the sum-binding definition.*

These definitions overcome many of the shortcomings of the sum-binding definition. They are applicable to bit and string commitment schemes, they are closer to the intuitive definition that a commitment has only one bit or string "inside" and they are more convenient to work with. Indeed, one of the main results of this thesis crucially relies on the use of these new definitions.

Depending on the version of our definition, we end up with weaker or stronger notions of the binding property. When we restrict the domain to one bit, the weaker definition is equivalent to the sum-binding definition, while the stronger definition is *strictly* stronger.

Our definitions also include relaxed versions, called *fairly-binding*, which allow dishonest provers to open to a value other than the one they committed to, but if they do, the resulting string will be random and out of their control. This relaxation will play an important role later on.

Naturally, a new definition is only useful if there are schemes that actually satisfy it:

Contribution 1.2. *We show that for every binding property that we define, there exists a variant of $\mathcal{CHSH}^q$ that satisfies it. In particular, this is the first time that the security of this scheme is analyzed as a* string-commitment *scheme.*

Recall that $\mathcal{CHSH}^q$ is easily understood as a string-commitment scheme when the bit b is replaced with an arbitrary field element. We show in Section 3.2.6 that this scheme satisfies the fairly-binding property. This in turn allows us to analyze the scheme from [LKB+15] using our composition theorem (see Section 1.4.2 and Chapter 4). Furthermore, if we change $\mathcal{CHSH}^q$ so that the second prover sends the string s that he wants to open to along

with the opening information, it satisfies our stronger definition of the binding property.

1.4.2 Composition Theorem for Multi-Round Schemes

Contribution 1.3. *We prove a composition theorem for two-prover commitment schemes: if a pair of two-prover commitment schemes $\mathcal{S}$ and $\mathcal{S}'$ satisfies some mild requirements, they can be composed into a new secure commitment scheme with delayed opening.*

The composition works as follows: In the first round, the first prover commits to some string s using $\mathcal{S}$. In the second round, instead of sending the opening information, the second prover *commits* to the opening information using $\mathcal{S}'$. Then, the opening phase of $\mathcal{S}'$ is executed, the verifier learns the opening information and uses it to open the commitment produced by $\mathcal{S}$. Note that the opening phase in $\mathcal{S}'$ can itself be multi-round, so this composition operation can be applied iteratively.

Intuitively, one would expect such a composition to work. Committing to the opening information before revealing it should not affect the security of the commitment, as long as $\mathcal{S}'$ is secure. However, proving that this is indeed the case turns out to be nontrivial. In particular, there seems to be no straightforward way to prove this result using the sum-binding definition.

As we mentioned in Section 1.3.1, a two-prover commitment scheme that is binding for non-communicating provers is binding for a limited time in the relativistic setting, i.e., when the provers can only communicate with some delay. By means of the above composition, it is possible to *delay* the opening of the original commitment. Thus, the multi-round schemes that are generated by the composition operation are binding in the relativistic setting, if the rounds are timed correctly.

Using our new definitions of the binding property, we formally prove this composition theorem. The failure probabilities of the component schemes add up. That is, if the binding property in $\mathcal{S}$ fails with probability at most ε and in $\mathcal{S}'$ with probability at most ε', the composed scheme fails with probability at most $\varepsilon + \varepsilon'$.

Given that, the bit-commitment scheme presented in [LKB$^+$15] can be viewed as a composition of multiple instances of $\mathcal{CHSH}^q$. Contribution 1.3 gives us a means to analyze that scheme. Thus, we obtain the following result:

Contribution 1.4. *The binding error of the Lunghiet al. relativistic commitment scheme grows linearly in the number of rounds, instead of double-exponentially, as previously proven in [LKB$^+$15]. Furthermore, security holds with respect to a stronger definition of the binding property instead of the commonly-used sum-binding definition.*

To put this difference in real-world terms: The authors of [LKB$^+$15] implemented their scheme with provers in Bern and Geneva (distance: 129.2

km). Based on their analysis, they concluded that this commitment would stay binding for at least 2 ms. Based on our analysis, this time can be scaled up to 10^{56} years, or, speaking more practically, until the devices run out of memory. Alternatively, one can also decrease the distance: Verbanis *et al.* executed the Lunghi *et al.* scheme for 24 hours across a distance of 7 km, based on our security analysis [VMH+16].

Finally, we also show that our analysis of the scheme is essentially tight, i.e., the binding error probability can not be better than linear in m.

1.4.3 Partial Progress towards Quantum Safety

So far, no multi-round relativistic commitment scheme has been proven to be "post-quantum" or quantum-safe in the sense that honest parties are protected against adversaries with quantum capabilities, while not having such capabilities themselves. Our notion of quantum-safety is different from the one used in computational cryptography: In the computational setting, quantum-safe cryptography is concerned with adversaries that use a quantum computer. In the information-theoretic setting, quantum computers are irrelevant because the set of computable functions is the same for classical and quantum computers. Quantum computers are believed to provide speedups for computing some functions, e.g., factoring integers, but this makes no difference when we consider computationally unbounded adversaries. However, when we impose restrictions on the communication between adversaries, a different aspect of quantum information becomes relevant: using quantum entanglement, the adversaries can correlate their behavior in ways that are not possible with shared randomness only.

The basic intuition of the composition theorem still applies. If we have two commitment schemes that are binding for provers with shared entanglement, then the composed scheme should be binding as well: the provers commit to the opening information of the first scheme, so they can delay revealing it without being able to change it. In Chapter 5, we show some partial progress towards proving that the Lunghi *et al.* scheme is quantum safe.

The first hurdle for proving quantum-safety is that some of our new definitions do not make sense for entangled adversaries, since they assume that the provers can only use shared randomness.

Contribution 1.5. *We provide quantum analogues for our new definitions of the binding property and prove a composition theorem for these definitions with respect to adversaries that have a shared entangled quantum state.*

We define a quantum analogue for our stronger binding property and show a composition theorem based on this definition. We also show that $\mathcal{CHSH}^q$ satisfies the weaker definition of the binding property against provers with quantum capabilities.

The proof of the composition theorem for the quantum case follows an approach that is slightly different from the composition theorem for the classical case: In the classical case, we extend a multi-round scheme by prefixing it with a one-round scheme. Both schemes are assumed to be binding according to the same definition, and the composed scheme is binding according to that definition as well.

In the quantum case, we start with a multi-round scheme that is binding according to the weaker definition and a one-round scheme that is binding according to the stronger one. The one-round scheme is *appended* to the multi-round scheme. The composed scheme is binding according to the *weaker* definition.

There remains one missing piece for actually proving that there is a multi-round scheme which is binding in the quantum setting: we do not know if $\mathcal{CHSH}^q$ (or any other one-round scheme) also satisfies the stronger binding property and thus, the question whether the composed scheme is binding for provers with quantum capabilities is left open.

1.4.4 Impossibility of Two-Prover Commitments with Security against Non-Signaling Attacks

In Chapter 6, we leave the topic of relativistic commitment schemes and discuss whether there are two-prover commitment schemes whose security depends *only* on the assumption that the provers can not communicate. In the classical and quantum case, we make assumptions about the physical laws that the provers can use to correlate their behavior. To remove these assumptions, we need to consider non-signaling provers. That is, we allow any input-output behavior of the provers as long as it does not imply transmission of information.

An example of a non-signaling system is the *NL-box* (non-local box), also known as the *PR-box* which was introduced by Sandu Popescu and Daniel Rohrlich in [PR94]. Let $p(x, y|a, b)$ be a conditional distribution where x, y, a and b are bits. Suppose that the marginals $p(x|a, b)$ and $p(y|a, b)$ are both uniformly random, but for any values of a and b, $p(x \oplus y = a \cdot b|a, b) = 1$. Now consider two provers with joint access to a "black box" that samples this distribution. The first prover supplies the input a and receives the output x. The second prover supplies b and receives the output y. We assume that the box immediately returns the output once the input is entered. It is impossible for them to use this box to communicate with each other, since each prover only sees a uniformly random bit, no matter what input the other prover has entered. However, both provers know that the outputs they receive are always correlated so that $x \oplus y = a \cdot b$.

Implementing such a box appears to be physically impossible without the two "halves" of it exchanging information. Using classical shared randomness, such a box could at most achieve $p(x \oplus y = a \cdot b|a, b) = 0.75$. Quantum entanglement increases this probability to ≈ 0.85 (see Section 5.2.4). But if

we are not willing to make any assumptions about the laws of physics that constrain them, we need to assume that the provers could use such a box. Note that this box renders the $\mathcal{CHSH}^q$ scheme non-binding: If the verifier's first message is $a = a_1 \dots a_n$, the first prover puts each bit in an independent copy of the box and receives $x = x_1 \dots x_n$ as output. The second prover then can open to $b \in \{0, 1\}$ by inputting b to every copy. The output $y = y_1 \dots y_n$ then satisfies $x_i \oplus y_i = a_i \cdot b$, so the provers can open to any bit they want.

This shows that $\mathcal{CHSH}^q$ is insecure against general non-signaling provers. This is related to the fact that the CHSH game, like all XOR games[7], has non-signaling value of 1, meaning that there is a non-signaling strategy for this game that always wins. However, there exist two-player non-local games that have a non-signaling value strictly lower than 1. For example, the Fortnow-Feige-Lovász game [For98, FL92] has a non-signaling, quantum and classical value of $2/3$ (see Appendix A in [Hol09] for a proof). Thus, one might hope that there is a commitment scheme that is secure against general non-signaling provers. However, we show that this is not the case.

Contribution 1.6. *We show that a two-prover commitment scheme that is hiding can not be binding for general non-signaling provers.*

If the scheme is perfectly hiding, then non-signaling dishonest provers can perfectly emulate the behavior of honest provers. As an example, consider a simple bit-commitment scheme where, in the commit phase, the verifier sends a message a to the first prover who replies with a message x, and in the opening phase, the second prover sends some opening information y to the verifier. The verifier then computes his output as a function of a, x and y.

Let $p_b(x, y|a)$ be the distribution that describes the input-output behavior of the honest provers when committing and opening to $b \in \{0, 1\}$. Then, if the scheme is perfectly hiding, $p(x, y|a, b) := p_b(x, y|a)$ is a bi-partite non-signaling distribution, where a is the input for P and b the input for Q. Thus, it is possible for non-signalling provers to sample this distribution and exactly replicate the input-output behavior of the honest provers. But unlike the honest provers, they can choose the bit that they want to open to *after* the commit phase, since b is an input for the second prover who is inactive in the commit phase.

Furthermore, we show that if the scheme is close to perfectly hiding, there is a bi-partite non-signaling distribution that is *statistically close* to the input-output behavior of the honest provers. Thus, the dishonest provers can emulate the honest provers *almost* perfectly in that case.

We prove similar results for more general schemes where both provers are active in the commit and opening phase. Here, the proof is somewhat more involved, because when we adapt the approach used in simple schemes to this problem, the outcome is *not* a non-signaling distribution. We also investigate

[7] A XOR game is a two-player non-local game where the players output bits and the outcome only depends on the exclusive-or of the players' outputs.

the case where the commit phase can consist of multiple rounds of communication. Here, we again show an impossibility result, but only for perfectly hiding schemes.

We also present a positive result:

Contribution 1.7. *We show the existence of a three-prover commitment scheme that is perfectly hiding and at the same time binding for non-signalling provers.*

A scheme that achieves this property works as follows: Take the $\mathcal{CHSH}^q$ bit-commitment scheme and add a third prover that mimics the behavior of the second prover in the opening phase. In the opening phase, the verifier computes the output as usual from the messages of the first two provers, and also checks if the second and third prover sent the same message. If that is not the case, he outputs $\perp$.

This construction is reminiscent of a result by Masanes, Acin and Gisin [MAG06] which implies that for every two-player game $\mathcal{G}$ where the second player has two possible inputs, there is a three-player game $\mathcal{G}'$ whose non-signaling value is the same as the classical value of $\mathcal{G}$. That is, non-signaling provers have the same chance of winning $\mathcal{G}'$ as classical provers have of winning $\mathcal{G}$. The scheme $\mathcal{G}'$ is constructed by having the first two players play $\mathcal{G}$ and requiring that the third player produces the same output as the second player.

Chapter 2

Preliminaries

2.1 Probabilities

2.1.1 Basic Notation

A (finite) *probability distribution* is a function $p : \mathcal{X} \to [0,1]$, $x \mapsto p(x)$, where $\mathcal{X}$ is a finite non-empty set such that $\sum_{x\in\mathcal{X}} p(x) = 1$. For $x_\circ \in \mathcal{X}$, we write $p(x = x_\circ)$ instead of $p(x_\circ)$. For any subset $\Lambda \subset \mathcal{X}$, called an *event*, the probability $p(\Lambda)$ (or $p(x \in \Lambda)$) is naturally defined as $p(\Lambda) = \sum_{x\in\Lambda} p(x)$, and it holds that

$$p(\Lambda) + p(\Gamma) = p(\Lambda \cup \Gamma) + p(\Lambda \cap \Gamma) \leq 1 + p(\Lambda \cap \Gamma) \tag{2.1}$$

for all $\Lambda, \Gamma \subset \mathcal{X}$, and, more generally, that

$$\sum_{i=1}^{k} p(\Lambda_i) \leq p(\Lambda_1 \cup \ldots \cup \Lambda_k) + \sum_{i<j} p(\Lambda_i \cap \Lambda_j) \leq 1 + \sum_{i<j} p(\Lambda_i \cap \Lambda_j) \tag{2.2}$$

for all $\Lambda_1, \ldots, \Lambda_k \subset \mathcal{X}$. For a distribution $p : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ on two variables and a relation R on $\mathcal{X} \times \mathcal{Y}$ (e.g., $x = y$, $x = f(y)$, $x \neq y$) the probability $p(R(x,y))$ is defined by

$$p(R(x,y)) = p\big(\{(x,y) \in \mathcal{X} \times \mathcal{Y} \,|\, R(x,y)\}\big) = \sum_{\substack{x\in\mathcal{X}, y\in\mathcal{Y} \\ \text{s.t. } R(x,y)}} p(x,y)\ .$$

The *marginals* $p(x)$ and $p(y)$ are given by $p(x) = \sum_y p(x,y)$ and $p(y) = \sum_x p(x,y)$, respectively. Vice versa, given two distributions $p(x)$ and $p(y)$, we say that a distribution $p(x,y)$ on two variables is a *consistent joint distribution* if the two marginals of $p(x,y)$ coincide with $p(x)$ and $p(y)$, respectively.

Definition 2.1. *Let $p(x)$ and $p(y)$ be two distributions over a common set $\mathcal{X}$. The* statistical distance *of the two distributions is defined as*

$$d(p(x), p(y)) = \frac{1}{2} \sum_{x_\circ \in \mathcal{X}} |p(x = x_\circ) - p(y = x_\circ)|$$

2.1.2 Some Basic Technical Results

We will make use of the following lemma regarding the existence of a consistent joint distribution that maximizes the probability that $x = y$.

Lemma 2.2. *Let $p(x)$ and $p(y)$ be two distributions on a common set $\mathcal{X}$. Then there exists a consistent joint distribution $p(x, y)$ such that $p(x = y = x_\circ) = \min\{p(x{=}x_\circ), p(y{=}x_\circ)\}$ for all choices of $x_\circ \in \mathcal{X}$. Additionally, $p(x, y)$ satisfies $p(x, y|x \neq y) = p(x|x \neq y) \cdot p(y|x \neq y)$.*

Proof. We first extend the respective probability spaces given by the distributions $p(x)$ and $p(y)$ by introducing an event Δ and declaring that

$$p(x{=}x_\circ \wedge \Delta) = \min\{p(x = x_\circ), p(y = x_\circ)\} = p(y{=}x_\circ \wedge \Delta)$$

for every $x_\circ \in \mathcal{X}$. Note that $p(\Delta)$ is well defined (by summing over all $x_\circ$). As we will see below, Δ will become the event $x = y$. In order to find a consistent joint distribution $p(x, y)$, it suffices to find a consistent joint distribution $p(x, y|\Delta)$ for $p(x|\Delta)$ and $p(y|\Delta)$, and a consistent joint distribution $p(x, y|\neg\Delta)$ for $p(x|\neg\Delta)$ and $p(y|\neg\Delta)$. The former, we choose as

$$p(x = x_\circ \wedge y = x_\circ|\Delta) := \min\{p(x = x_\circ), p(y = x_\circ)\}/p(\Delta)$$

for all $x_\circ \in \mathcal{X}$, and $p(x = x_\circ \wedge y = y_\circ|\Delta) := 0$ for all $x_\circ \neq y_\circ \in \mathcal{X}$, and the latter we choose as

$$p(x = x_\circ \wedge y = y_\circ|\neg\Delta) := p(x = x_\circ|\neg\Delta) \cdot p(y = y_\circ|\neg\Delta)$$

for all $x_\circ, y_\circ \in \mathcal{X}$. It is straightforward to verify that these are indeed *consistent* joint distributions, as required, so that $p(x, y) = p(x, y|\Delta) \cdot p(\Delta) + p(x, y|\neg\Delta) \cdot p(\neg\Delta)$ is also consistent. Furthermore, note that $p(x{=}y|\Delta) = 1$ and $p(x{=}y|\neg\Delta) = 0$; the latter holds because we have $p(x{=}x_\circ \wedge \Delta) = p(x = x_\circ)$ or $p(y{=}x_\circ \wedge \Delta) = p(y = x_\circ)$ for each $x_\circ \in \mathcal{X}$, and thus $p(x{=}x_\circ \wedge \neg\Delta) = 0$ or $p(y = x_\circ \wedge \neg\Delta) = 0$. As such, Δ is the event $x = y$, and therefore $p(x = y = x_\circ) = p(x = x_\circ \wedge \Delta) = \min\{p(x = x_\circ), p(y = x_\circ)\}$ for every $x_\circ \in \mathcal{X}$ as required. Finally, the claim regarding $p(x, y|x \neq y)$ holds by construction. □

The following property of the statistical distance is well known (see e.g. [RK05]) and can easily be proved using Lemma 2.2.

Corollary 2.3. *Let $p(x)$ and $p(y)$ be two distributions over the same set $\mathcal{X}$ with statistical distance $d\big(p(x),p(y)\big)=\varepsilon$. Then, there exists a consistent joint distribution $p(x,y)$ over $\mathcal{X}\times\mathcal{X}$ such that $p(x\neq y)=\varepsilon$.*

Proof. We apply Lemma 2.2 to obtain a consistent joint distribution $p(x,y)$ such that for all $x_\circ\in\mathcal{X}$, $p(x=y=x_\circ)=\min\{p(x=x_\circ),p(y=x_\circ)\}$. We then have

$$\begin{aligned}
p(x\neq y) &= 1-\sum_{x_\circ\in\mathcal{X}} p(x=y=x_\circ)\\
&= 1-\sum_{x_\circ\in\mathcal{X}} \min\{p(x=x_\circ),p(y=x_\circ)\}\\
&= \frac{1}{2}\sum_{x_\circ\in\mathcal{X}} p(x=x_\circ)-\min\{p(x=x_\circ),p(y=x_\circ)\}\\
&\quad+\frac{1}{2}\sum_{x_\circ\in\mathcal{X}} p(y=x_\circ)-\min\{p(x=x_\circ),p(y=x_\circ)\}\\
&= \frac{1}{2}\sum_{x_\circ\in\mathcal{X}} |p(x=x_\circ)-p(y=x_\circ)|\\
&= \varepsilon
\end{aligned}$$

as claimed. □

The following is an immediate consequence.

Lemma 2.4. *Let $p(x_0,y_0)$ and $p(x_1,y_1)$ be distributions with $d\big(p(x_0),p(x_1)\big)=\varepsilon$. Then, there exists a consistent joint distribution $p(x_0,x_1,y_0,y_1)$ such that $p'(x_0\neq x_1)=\varepsilon$ and, as a consequence, $d\big(p(x_0,y_1),p(x_1,y_1)\big)\leq\varepsilon$.*

Proof. We first apply Corollary 2.3 to $p(x_0)$ and $p(x_1)$ to obtain a consistent joint distribution $p(x_0,x_1)$, and then we set

$$p(x_0,x_1,y_0,y_1)=p(x_0,x_1)\cdot p(y_0|x_0)\cdot p(y_1|x_1)\,.$$

It is easy to see that this distribution is consistent with $p(x_0,y_0)$ and $p(x_1,y_1)$. For the last claim, we note that

$$\begin{aligned}
p(x_0,y_1) &= p(x_0=x_1)\cdot p(x_0,y_1|x_0=x_1)+p(x_0\neq x_1)\cdot p(x_0,y_1|x_0\neq x_1)\\
&= p(x_0=x_1)\cdot p(x_1,y_1|x_0=x_1)+p(x_0\neq x_1)\cdot p(x_0,y_1|x_0\neq x_1)
\end{aligned}$$

and

$$p(x_1,y_1)=p(x_0=x_1)\cdot p(x_1,y_1|x_0=x_1)+p(x_0\neq x_1)\cdot p'(x_1,y_1|x_0\neq x_1)$$

and the claim follows because $p(x_0\neq x_1)=\varepsilon$. □

When applying the above lemma, we say that we "glue together" the distributions $p(x_0, y_0)$ and $p(x_1, y_1)$ along x_0 and x_1.

Remark 2.5. *In the special case where $p(x_0)$ and $p(x_1)$ are distributed identically, we obviously have $p(x_0, y_1) = p(x_1, y_1)$.*

Remark 2.6. *It is easy to see from the proof of Lemma 2.4 that the following natural property holds. If $p(x_0, x_1, y_0, y_1, y_0', y_1')$ is obtained by gluing together $p(x_0, y_0, y_0')$ and $p(x_1, y_1, y_1')$ along x_0 and x_1, then the marginal $p(x_0, x_1, y_0, y_1)$ coincides with the distribution obtained by gluing together the marginals $p(x_0, y_0)$ and $p(x_1, y_1)$ along x_0 and x_1.*

Let $p(x, y, z)$ be a distribution over $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ and let $\Lambda \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. We then write $x \to y \to z$ to express that $p(x)$ and $p(z)$ are independent when conditioned on y, i.e., $p(x, z|y) = p(x|y)p(z|y)$. Similarly, we write $x \to \Lambda \to y$ to express that $p(x, y|\Lambda) = p(x|\Lambda)p(y|\Lambda)$, etc. We show the following property for conditionally independent variables.

Lemma 2.7. *If $x \to y \to z$ and $x \to x \neq y \to y$, then $x \to x \neq y \to z$.*

Proof. We assume that $x \to y \to z$ and $x \to x \neq y \to y$. We first observe that

$$\begin{aligned} p(x, x \neq y, z|y) &= p(x, x \neq y|y)\, p(z|x, y, x \neq y) \\ &= p(x, x \neq y|y)\, p(z|x, y) \\ &= p(x, x \neq y|y)\, p(z|y)\,, \end{aligned}$$

which means that $(x, x \neq y) \to y \to z$, and, by summing over x, implies $x \neq y \to y \to z$. It follows that

$$p(z|x, y, x \neq y) = p(z|y) = p(z|y, x \neq y)\,,$$

which actually means that $x \to (y, x \neq y) \to z$. Therefore,

$$\begin{aligned} p(x, z|x \neq y) = \sum_y p(x, y, z|x \neq y) &= \sum_y p(x, y|x \neq y)\, p(z|x, y, x \neq y) \\ &= p(x|x \neq y) \sum_y p(y|x \neq y)\, p(z|y, x \neq y) \\ &= p(x|x \neq y) \sum_y p(y, z|x \neq y) \\ &= p(x|x \neq y)\, p(z|x \neq y)\,, \end{aligned}$$

which was to be proven. □

2.2 Two-Prover Commitment Schemes

2.2.1 Protocols

In this work, we will consider 3-party interactive *protocols*, where the parties are named P, Q and V (the two "provers" and the "verifier"). Such a protocol prot_{PQV} consists of a triple $(\mathsf{prot}_P, \mathsf{prot}_Q, \mathsf{prot}_V)$ of L-round *interactive algorithms* for some $L \in \mathbb{N}$. Each interactive algorithm takes an input, and for every round $\ell \leq L$ computes the messages to be sent to the other algorithms/parties in that round as deterministic functions of its input, the messages received in the previous rounds, and the local randomness. In the same way, the algorithms produce their respective outputs after the last round. We write

$$(out_P \| out_Q \| out_V) \leftarrow \big(\mathsf{prot}_P(in_P) \| \mathsf{prot}_Q(in_Q) \| \mathsf{prot}_V(in_V)\big)$$

to denote the execution of the protocol prot_{PQV} on the respective inputs in_P, in_Q and in_V, and that the respective outputs out_P, out_Q and out_V are produced. Clearly, for any protocol prot_{PQV} and any input in_P, in_Q, in_V, the probability distribution $p(out_P, out_Q, out_V)$ of the output is naturally well defined.

If we want to make the local randomness explicit, we write $\mathsf{prot}_P[\xi_P](in_P)$ etc., and understand that ξ_P is correctly sampled. Furthermore, we write $\mathsf{prot}_P[\xi_{PQ}](in_P)$ and $\mathsf{prot}_Q[\xi_{PQ}](in_Q)$ to express that prot_P and prot_Q use *the same* randomness, in which case we speak of *joint randomness*.

We can *compose* two interactive algorithms prot_P and prot'_P in the obvious way, by applying prot'_P to the output of prot_P. The resulting interactive algorithm is denoted as $\mathsf{prot}'_P \circ \mathsf{prot}_P$. Composing the respective algorithms of two protocols $\mathsf{prot}_{PQV} = (\mathsf{prot}_P, \mathsf{prot}_Q, \mathsf{prot}_V)$ and $\mathsf{prot}'_{PQV} = (\mathsf{prot}'_P, \mathsf{prot}'_Q, \mathsf{prot}'_V)$ results in the composed protocol $\mathsf{prot}'_{PQV} \circ \mathsf{prot}_{PQV}$. If prot_P is a *non-interactive* algorithm, then $\mathsf{prot}'_{PQV} \circ \mathsf{prot}_P$ is naturally understood as the protocol $\mathsf{prot}'_{PQV} \circ \mathsf{prot}_P = (\mathsf{prot}'_P \circ \mathsf{prot}_P, \mathsf{prot}'_Q, \mathsf{prot}'_V)$, and similarly $\mathsf{prot}'_{PQV} \circ \mathsf{prot}_{QV}$ in case prot_{QV} is a protocol among Q and V only.

2.2.2 Defining Commitment Schemes

We model a two-prover commitment scheme as two protocols, one for the commit phase, one for the opening phase. In the commit phase, the input for the provers is an element s of the scheme's domain D which is the value they want to commit to. The verifier has no input. All three participants may output some state information. The protocol for the opening phase then takes this state information as input, and the verifier outputs an element of $D \cup \{\perp\}$. An output of $s \in D$ indicates that the provers successfully opened their commitment to s while an output of $\perp$ indicates that they failed to open their commitment.

Dishonest parties may execute arbitrary protocols. Dishonest provers do not have an input in the commit phase: intuitively speaking, they do not know which value they eventually want to open to at this time.

Definition 2.8. *A* 2-prover (string) commitment scheme $\mathcal{S}$ *with domain* D *consists of a pair of interactive protocols* $\mathsf{com}_{PQV} = (\mathsf{com}_P, \mathsf{com}_Q, \mathsf{com}_V)$ *and* $\mathsf{open}_{PQV} = (\mathsf{open}_P, \mathsf{open}_Q, \mathsf{open}_V)$ *between the provers* P *and* Q *and the verifier* V*, with the following syntactics. The* commit protocol com_{PQV} *uses joint randomness* ξ_{PQ} *for* P *and* Q *and takes an element* $s \in D$ *as input for* P *and* Q *(and independent randomness and no input for* V*), and it outputs a* commitment *com to* V *and some state information to* P *and* Q*:*

$$(state_P \| state_Q \| c) \leftarrow \left(\mathsf{com}_P[\xi_{PQ}](s) \| \mathsf{com}_Q[\xi_{PQ}](s) \| \mathsf{com}_V(\emptyset)\right).$$

The opening protocol open_{PQV} *uses joint randomness* η_{PQ} *and outputs a string or a rejection symbol to* V*, and nothing to* P *and* Q*:*

$$(\emptyset \| \emptyset \| s) \leftarrow \left(\mathsf{open}_P[\eta_{PQ}](state_P) \| \mathsf{open}_Q[\eta_{PQ}](state_Q) \| \mathsf{open}_V(c)\right)$$

with $s \in \{0,1\}^n \cup \{\perp\}$*. If* $D = \{0,1\}$*, we refer to* $\mathcal{S}$ *as a* bit*-commitment scheme instead, and we tend to use* b *rather than* s *to denote the committed bit.*

Remark 2.9. *Note that we still speak of string-commitment schemes even if the domain* D *does not consist of bit-strings.*

Remark 2.10. *By convention, we assume throughout the paper that the commitment* c *output by* V *equals the* communication *that takes place between* V *and the provers during the commit phase. This is without loss of generality since, in general,* c *is computed as a (possibly randomized) function of the communication, which* V *just as well can apply in the opening phase.*

Remark 2.11. *Note that we specify that* P *and* Q *use* fresh *joint randomness* η_{PQ} *in the opening phase, and, if necessary, the randomness* ξ_{PQ} *from the commit phase can be "handed over" to the opening phase via* $state_P$ *and* $state_Q$*; this will be convenient later on.*

Whenever we refer to such a 2-prover commitment scheme, we take it as understood that the scheme is complete and hiding, as defined below, for "small" values of γ and δ. Since our focus will be on the binding property, we typically do not make the parameters γ and δ explicit.

Definition 2.12. *A 2-prover commitment scheme is* γ-complete *if in an honest execution* V*'s output* s *of* open_{PQV} *equals* P *and* Q*'s input* s *to* com_{PQV} *except with probability* η*, for any choice of* P *and* Q*'s input* $s \in D$*.*

The standard definition for the hiding property is as follows:

Definition 2.13. *A 2-prover commitment scheme is* δ-hiding *if for any commit strategy* $\overline{\mathsf{com}}_V$ *and any two strings* s_0 *and* s_1*, the distribution of the commitments* c_0*,* c_1*, produced as*

$$(state_P \| state_Q \| c_b) \leftarrow (\mathsf{com}_P[\xi_{PQ}](s_b) \| \mathsf{com}_Q[\xi_{PQ}](s_b) \| \overline{\mathsf{com}}_V(\emptyset)), b = 0, 1$$

have statistical distance at most δ*. A* 0*-hiding scheme is also called* perfectly hiding.

Defining the binding property is more subtle. First, note that an attack against the binding property consists of a possible commit strategy $\overline{\mathsf{com}}_{PQ} = (\overline{\mathsf{com}}_P, \overline{\mathsf{com}}_Q)$ and a possible opening strategy $\overline{\mathsf{open}}_{PQ} = (\overline{\mathsf{open}}_P, \overline{\mathsf{open}}_Q)$ for P and Q. Any such attack fixes $p(s)$, the distribution of $s \in \{0,1\}^n \cup \{\perp\}$ that is output by V after the opening phase, in the obvious way.

A somewhat accepted definition for the binding property of a 2-prover *bit* commitment scheme, as it is for instance used in [CSST11, LKB$^+$15, FF15] (up to the factor 2 in the error parameter), is the *probability sum-binding* property defined below. Here, we assume it has been specified which attacks are *possible*, e.g., those where P and Q do not communicate during the course of the scheme.

Definition 2.14. *A 2-prover bit-commitment scheme is* ε-sum-binding *if for every possible commit strategy* $\overline{\mathsf{com}}_{PQ}$*, and for every pair of possible opening strategies* $\overline{\mathsf{open}}^0_{PQ}$ *and* $\overline{\mathsf{open}}^1_{PQ}$*, which fix distributions* $p(b_0)$ *and* $p(b_1)$ *for* V*'s respective outputs, it holds that*

$$p(b_0 = 0) + p(b_1 = 1) \leq 1 + 2\varepsilon\,.$$

In the literature (see e.g. [CSST11] or [LKB$^+$15]), the two probabilities $p(b_0 = 0)$ and $p(b_1 = 1)$ above are usually referred to as p_0 and p_1, respectively.

In the information theoretic setting, a commitment scheme can not be both hiding and binding with good (i.e., low) parameters. Thus, we have to assume some restriction on the provers, e.g., that they are unable to communicate during the execution of the scheme. However, we might also be more liberal and allow some limited communication during the protocol, as in the Lunghi *et al.* multi-round scheme.

If we rely on relativity to enforce the communication restrictions, we need to make sure that the provers are at the appropriate distance from each other. We do not address this problem in this thesis and assume that the provers are at fixed known positions. However, depending on the commitment scheme, it can be possible to also split the verifier into two separate agents that each only need to communicate with one prover. These are then placed next to the provers they communicate with, and brought together at the end of the opening phase to compute the result. For the Lunghi *et al.* scheme, this is possible – in fact, it is how the scheme is presented in [LKB$^+$15]. For simplicity, we describe protocols with only one verifier.

Our proof of the commitment scheme relies on different, but stronger or equivalent, notions of a binding commitment. We explain these, and the relations among them in Chapter 3.

2.2.3 The $\mathcal{CHSH}^q$ Scheme

Our main example is a generalization of the bit-commitment scheme by Crépeau *et al.* [CSST11]. Let q be a prime power and let $\mathbb{F}_q$ be the finite field with q elements. The bit-commitment scheme $\mathcal{CHSH}^q$ works as follows: The commit phase com_{PQV} instructs V to sample and send to P a uniformly random $a \in \mathbb{F}_q$, and it instructs P to return $x := r + a \cdot b$ to V, where $r \in \mathbb{F}_q$ is the provers' joint randomness and b is the bit to commit to. The opening phase open_{PQV} instructs Q to send $y := -r$ to V, and V outputs the (smaller) bit b that satisfies $x + y = a \cdot b$, and $b := \perp$ in case no such bit exists. Note that the provers in this scheme use the same randomness in the commit and opening phase; thus, formally, Q needs to output the shared randomness as $state_Q$. The opening phase uses no fresh randomness.

It is easy to see that this scheme is q^{-1}-complete and perfectly hiding (completeness fails in case $a = 0$). For *classical* provers that do not communicate at all, the scheme is $\left(q^{-1}/2\right)$-sum-binding. As for *quantum* provers, Crépeau *et al.* showed that the scheme $\mathcal{CHSH}^{2^n}$ is $2^{-n/2}$-binding; this was recently minorly improved to $2^{-(n+1)/2}$ by Sikora, Chailloux and Kerenidis [SCK14].

We also want to consider an extended version of the scheme where, instead of a bit, the provers commit to an arbitrary field element $s \in \mathbb{F}_q$, thus making $\mathbb{F}_q$ the domain of the scheme. In the opening phase, the verifier's output is picked as above, except that it is selected from the set $\mathbb{F}_q$ instead of $\{0, 1\}$. If $a \neq 0$, the output is $s = a^{-1}(x + y)$. In general, we will thus view $\mathcal{CHSH}^q$ as a string-commitment scheme, and explicitly mention when we restrict its domain to $\{0, 1\}$.

However, it is a priori not clear what a suitable definition for the binding property is, especially because for this particular scheme, the dishonest provers can always honestly commit to a string s, and can then decide to correctly open the commitment to s by announcing $y := r$, or open to a *random* string by announcing a randomly chosen y — any y satisfies $x + y = a \cdot s$ for *some* s (unless $a = 0$, which almost never happens).[1]

Due to its close relation to the CHSH game [CHSH69], in particular to the arbitrary-finite-field version considered in [BS15], we will refer to this *string* commitment scheme as $\mathcal{CHSH}^q$.

[1] This could easily be prevented by requiring Q to announce s (rather than letting V compute it), but we want the information announced during the opening phase to fit into the domain of the commitment scheme.

Chapter 3

The Hiding and Binding Properties

3.1 Introduction

Formal security definitions form a crucial part of modern cryptography, where the aim is to mathematically prove the security of cryptographic schemes. Such definitions should capture and refine the informal intuition about the desired security requirements. For example, the informal goal of an encryption scheme is to keep a message secret from an adversary that does not know the appropriate key. The informal requirement guides the development of precise definitions (such as the modern game-based ones), but in formalizing it, one needs to also fill in many details that are left vague in the informal intuition.

Typically, we want security definitions to be as strong as possible while still being satisfiable, in order to offer security guarantees that are as strong as possible. Ideally, they should also be easy to work with. It is also desirable for them to be composable: Informally, this means that the stand-alone security of a scheme implies that security is still satisfied when the scheme is used as a building block in a larger system, and the security of the scheme propagates as expected to the larger system. The security proof for the larger system would not need to concern itself with the internal details of the components if the components satisfy composable security definitions.

In Chapter 1, we have discussed the informal security properties that a bit-commitment scheme should have. They should be hiding, meaning that a dishonest verifier cannot learn the committed value before the opening phase, and binding, meaning that after the commit phase, there is at most one value that can be revealed.

The (information-theoretic) hiding property is straightforward to define formally: even if the verifier is dishonest and arbitrarily deviates from the pro-

tocol, we require that the messages that he sees in the commit phase are statistically independent of the committed value. This definition can be relaxed to allow a limited amount of information by requiring the distributions to have small statistical distance from each other. Defining the information-theoretic binding property for two-prover schemes is more involved. The naive approach of requiring that the value to which the commitment can be opened should be uniquely determined by the verifier's view after the commit phase obviously leads to a contradiction with the information-theoretic binding property (see Section 1.2.1). Thus special care is necessary here.

In this chapter, we study several different and new definitions of the binding property which vary in certain technical aspects, and we analyze how they relate to each other. Our definitions vary in how we formalize the bit or string that the provers supposedly are committed to. One of our definitions, when restricted to bits, turns out to be equivalent to the sum-binding definition, while another one is strictly stronger. Our definitions also vary in how strict we are in not allowing the adversary to open to anything else than the committed value. Schemes that satisfy the less strict definition can quite easily be transformed into schemes that satisfy the stricter one by simply restricting their domain.

Naively, one might think that it suffices to consider the strongest achievable notion. However, some of our weaker definitions play a crucial role in our analysis of multi-round schemes (see Chapter 4).

We also prove that all of the definitions are satisfied by variants of the $\mathcal{CHSH}$ commitment scheme. This in particular is the first time the $\mathcal{CHSH}$ commitment scheme is proven secure as a *string* commitment scheme.

3.2 Defining The Binding Property

3.2.1 Possible Strategies

Like the sum-binding property, the binding properties we discuss in this chapter can only hold with respect to some restricted class of strategies. We call these strategies the *possible strategies*. In this chapter, we assume that all possible strategies $(\overline{\mathsf{com}}_{PQ}, \overline{\mathsf{open}}_{PQ})$ are *classical* interactive algorithms with access to joint randomness. We consider strategies that use quantum entanglement in Chapter 5. Our main result holds only in the classical case.

We assume that the set of possible strategies is the *convex hull* of a set of deterministic strategies. That is, if $(\overline{\mathsf{com}}_{PQ}, \overline{\mathsf{open}}_{PQ})$ is a possible randomized strategy, then the deterministic strategies that result from replacing the randomness with fixed values are possible as well. Conversely, if a strategy $(\overline{\mathsf{com}}_{PQ}, \overline{\mathsf{open}}_{PQ})$ instructs the provers to execute a possible strategy selected according to some probability distribution, then $(\overline{\mathsf{com}}_{PQ}, \overline{\mathsf{open}}_{PQ})$ is itself a possible strategy.

We make this assumption because we think of the set of possible strategies not as some arbitrary set, but as the strategies that are permitted by some constraints on the communication between the provers. If a set of strategies is permitted by those constraints, then executing a random strategy from this set should be possible as well. If a randomized strategy is permitted by those constraints, then deterministic strategies that result from replacing the randomness with fixed values should not violate the constraints either.

In the remainder of this chapter, we usually leave the set of possible strategies implicit and take it as understood that when we quantify over strategies, we refer only to possible strategies.

3.2.2 The (Strong) Binding Property

Intuitively, we say that a scheme is binding if after the commit phase there exists a string $\hat{s}$ so that no matter what the provers do in the opening phase, the verifier will output either $s = \hat{s}$ or $s = \perp$ (except with small probability). We consider two definitions of the binding property which interpret this intuitive requirement in two different ways. In the first definition, which we introduce in this section, $\hat{s}$ is a function of the provers' combined view immediately after the commit phase. In the second one, which we introduce in Section 3.2.3, $\hat{s}$ is specified by its distribution only. Both of these definitions admit a composition theorem.

Definition 3.1 (Binding property)**.** *A 2-prover commitment scheme $\mathcal{S}$ is* ε-binding *if for every commit strategy $\overline{\mathsf{com}}_{PQ}[\bar{\xi}_{PQ}]$ there exists a function $\hat{s}(\bar{\xi}_{PQ}, c)$ of the joint randomness $\bar{\xi}_{PQ}$ and the commitment[1] c such that for every opening strategy $\overline{\mathsf{open}}_{PQ}$ it holds that $p(s \neq \hat{s}(\bar{\xi}_{PQ}, c) \wedge s \neq \perp) \leq \varepsilon$. In short:*

$$\forall\, \overline{\mathsf{com}}_{PQ}\ \exists\, \hat{s}(\bar{\xi}_{PQ}, c)\ \forall\, \overline{\mathsf{open}}_{PQ} : p(s \neq \hat{s} \wedge s \neq \perp) \leq \varepsilon\,. \tag{3.1}$$

The string-commitment scheme $\mathcal{CHSH}^q$ does *not* satisfy this definition (the bit-commitment version does, as we will show): after the commit phase, the provers can still decide to open the commitment to a *fixed* string, chosen before the commit phase, or to a *random* string that is out of their control. We capture this property of $\mathcal{CHSH}^q$ by the following relaxed version of the binding property: we allow V's output s to be different from $\hat{s}$ and $\perp$, but in this case the provers should have little control over s; for any fixed *target string* $s_\circ$, it should be unlikely that $s = s_\circ$. Formally, this is captured as follows; we will show in Section 3.2.6 that $\mathcal{CHSH}^q$ is fairly-binding in this sense.

Definition 3.2 (Fairly binding property)**.** *A 2-prover commitment scheme $\mathcal{S}$ is* ε-fairly-binding *if for every commit strategy $\overline{\mathsf{com}}_{PQ}[\bar{\xi}_{PQ}]$ there exists a*

[1] Recall that by convention (Remark 2.10), c equals the communication between V and the provers during the commit phase.

function $\hat{s}(\bar{\xi}_{PQ}, c)$ *such that for every opening strategy* $\overline{\mathsf{open}}_{PQ}[\bar{\eta}_{PQ}]$ *and all* $s_\circ \in D$ *it holds that* $p(s \neq \hat{s}(\bar{\xi}_{PQ}, c) \wedge s = s_\circ) \leq \varepsilon$. *In short:*

$$\forall\, \overline{\mathsf{com}}_{PQ}\ \exists\, \hat{s}(\bar{\xi}_{PQ}, c)\ \forall\, \overline{\mathsf{open}}_{PQ}\ \forall\, s_\circ(\bar{\xi}_{PQ}, \bar{\eta}_{PQ}) : p(s \neq \hat{s} \wedge s = s_\circ) \leq \varepsilon\,. \quad (3.2)$$

If we want to show that a scheme is ε-(fairly-)binding (with respect to all possible strategies), it suffices to show that it is binding with respect to all possible *deterministic* strategies, as the following lemma shows.

Lemma 3.3. *Let* $\mathcal{S}$ *be a commitment scheme that is* ε*-(fairly-)binding with respect to all possible deterministic strategies. Then it also is* ε*-(fairly-)binding with respect to* all *possible strategies.*

Proof. We prove the lemma for ε-binding schemes. It is easy to see how the proof can be adapted for fairly-binding schemes. Let $\overline{\mathsf{com}}_{PQ}[\bar{\xi}_{PQ}]$ be a possible commit strategy. By our assumptions from Section 3.2.1, it follows that the strategy $\overline{\mathsf{com}}^{r_c}_{PQ}$ where we set the joint randomness $\bar{\xi}_{PQ}$ to the value r_c is also possible for every r_c. By the assumed binding property, for every r_c, there exists a function $\hat{s}_{r_c}$ such that for every deterministic opening strategy $\overline{\mathsf{open}}_{PQ}$, we have $p(s \neq \hat{s}_{r_c} \wedge s \neq \bot) \leq \varepsilon$.

We define $\hat{s}(\bar{\xi}_{PQ}, c) = \hat{s}_{\bar{\xi}_{PQ}}(c)$. If the provers use $\overline{\mathsf{com}}_{PQ}[\bar{\xi}]$ and any possible deterministic opening strategy, we have

$$p(s \neq \hat{s} \wedge s \neq \bot) = \sum_{r_c} p(\bar{\xi}_{PQ} = r_c) p(s \neq \hat{s}_{r_c}(c) \wedge s \neq \bot | \bar{\xi}_{PQ} = r_c) \leq \varepsilon\,.$$

It is straightforward to extend the above inequality to randomized opening strategies: the above inequality holds when we set the randomness to any particular value, and thus it also holds for the randomized strategy. □

The next lemma shows that in Definition 3.2, instead of quantifying over strings $s_\circ$, we may also quantify over functions of the provers' randomness.

Lemma 3.4. *Let* $\mathcal{S}$ *be an* ε*-fairly-binding scheme and* $\overline{\mathsf{com}}_{PQ}[\bar{\xi}_{PQ}]$ *a commit strategy. There is a function* $\hat{s}(\bar{\xi}_{PQ}, c)$ *such that for every opening strategy* $\overline{\mathsf{open}}_{PQ}[\bar{\eta}_{PQ}]$ *and every* function $s_\circ(\bar{\xi}_{PQ}, \bar{\eta}_{PQ})$ *with values in* D*, it holds that* $p(s \neq \hat{s}(\bar{\xi}_{PQ}, c) \wedge s = s_\circ(\bar{\xi}_{PQ}, \bar{\eta}_{PQ})) \leq \varepsilon$.

Proof. Let $\overline{\mathsf{com}}^{r_c}_{PQ}$ and $\overline{\mathsf{open}}^{r_o}_{PQ}$ be the deterministic strategies that results from fixing the randomness in $\overline{\mathsf{com}}_{PQ}[\bar{\xi}_{PQ}]$ to r_c and the randomness in $\overline{\mathsf{open}}_{PQ}[\bar{\eta}_{PQ}]$ to r_o. Fix an arbitrary function $s_\circ(\bar{\xi}_{PQ}, \bar{\eta}_{PQ})$. By the binding property, for every r, there is a function $\hat{s}_{r_c}(c)$ such that $p(s \neq \hat{s}_r(c) \wedge s = s_\circ(r_c, r_o) | \bar{\xi}_{PQ} = r_c, \bar{\eta}_{PQ} = r_o) \leq \varepsilon$. Setting $\hat{s}(\bar{\xi}_{PQ}, c) = \hat{s}_{\bar{\xi}_{PQ}}(c)$, we have

$$\begin{aligned}
&p(s \neq \hat{s}(\bar{\xi}_{PQ}, c) \wedge s = s_\circ(\bar{\xi}_{PQ}, \bar{\eta}_{PQ})) \\
&= \sum_{r_c, r_o} p(\bar{\xi}_{PQ} = r_c) p(\bar{\eta}_{PQ} = r_o) p(s \neq \hat{s}(r_c, c) \wedge s = s_\circ(r_c, r_o) | \bar{\xi}_{PQ} = r_c, \bar{\eta}_{PQ} = r_o) \\
&\leq \varepsilon
\end{aligned}$$

which proves our claim. □

Remark 3.5. *Clearly, the binding property implies the fairly binding property. Furthermore, in the case of* bit *commitment schemes it obviously holds that* $p(b \neq \hat{b} \wedge b \neq \perp) = p(b \neq \hat{b} \wedge b = 0) + p(b \neq \hat{b} \wedge b = 1)$*, and thus the fairly-binding property implies the binding property with a factor-2 loss in the parameter. Furthermore, every fairly-binding* string *commitment scheme gives rise to a binding* bit*-commitment scheme in a natural way, as shown by the following proposition.*

Proposition 3.6. *Let* $\mathcal{S}$ *be an* ε*-fairly-binding string-commitment scheme with domain* D*. Fix any two distinct strings* $s_0, s_1 \in D$ *and consider the bit-commitment scheme* $\mathcal{S}'$ *defined as follows. To commit to* $b \in \{0,1\}$*, the provers commit to* s_b *using* $\mathcal{S}$*, and in the opening phase* V *checks if* $s = s_b$ *for some bit* $b \in \{0,1\}$ *and outputs this bit if it exists and else outputs* $b = \perp$*. Then,* $\mathcal{S}'$ *is a* 2ε*-binding bit-commitment scheme.*

Proof. Fix some commit strategy $\overline{\mathsf{com}}_{PQ}$ for $\mathcal{S}'$ and note that it can also be used to attack $\mathcal{S}$. Thus, there exists a function $\hat{s}(\bar{\xi}_{PQ}, c)$ as in Definition 3.2. We define

$$\hat{b}(\bar{\xi}_{PQ}, c) = \begin{cases} 0 & \text{if } \hat{s}(\bar{\xi}_{PQ}, c) = s_0 \\ 1 & \text{otherwise} \end{cases}$$

Now fix an opening strategy $\overline{\mathsf{open}}_{PQ}$ for $\mathcal{S}'$, which again is also a strategy against $\mathcal{S}$. Thus, we have $p(\hat{s} \neq s = s_\circ) \leq \varepsilon$ for any $s_\circ$ (and in particular $s_\circ = s_0$ or s_1). This gives us

$$\begin{aligned} p(\hat{b} \neq b \neq \perp) &= p(\hat{b} = 1 \wedge b = 0) + p(\hat{b} = 0 \wedge b = 1) \\ &= p(\hat{s} \neq s_0 \wedge s = s_0) + p(\hat{s} = s_0 \wedge s = s_1) \\ &\leq p(\hat{s} \neq s_0 \wedge s = s_0) + p(\hat{s} \neq s_1 \wedge s = s_1)\,, \\ &\leq 2\varepsilon \end{aligned}$$

and thus $\mathcal{S}'$ is a 2ε-binding bit-commitment scheme. □

Remark 3.7. *The proof of Proposition 3.6 generalizes in a straightforward way: given an* ε*-fairly-binding commitment scheme* $\mathcal{S}$ *with domain* D*, and a subset* $D' \subseteq D$*, we define a commitment scheme* $\mathcal{S}_{D'}$ *with domain* D' *as follows: In the commit phase, the players use* $\mathcal{S}$ *to produce a commitment to* $s \in D'$*. In the opening phase, the players run the opening phase of* $\mathcal{S}$*. If the result is in* D'*,* V *outputs it, and otherwise outputs* $\perp$*. Then,* $\mathcal{S}_{D'}$ *is* $|D'|\varepsilon$*-binding.*

When $D' \not\subseteq D$*, but* $|D'| < |D|$*, we can define a similar scheme by fixing an injection from* D' *to* D*. In particular, any* ε*-fairly-binding* n*-bit string-commitment scheme can be turned into a* $2^k\varepsilon$*-binding* k*-bit string-commitment scheme for any* $k < n$*.*

3.2.3 The Weak Binding Property

Here, we introduce yet another definition for the binding property. It is similar in spirit to Definition 3.1, but weaker. One advantage of this weaker notion is that it is also meaningful when considering quantum attacks, whereas Definition 3.1 is not. Note, however, that in the quantum setting, it does *not* suffice to only consider deterministic attacks. Therefore, results that depend on this property do not automatically carry over to the quantum setting. That includes Theorem 4.13, the composition theorem. In Section 3.2.4, we will see that for *bit*-commitment schemes, this weaker notion of the binding property is equivalent to the sum-binding definition, i.e., Definition 2.14.

Definition 3.8 (Weak binding property)**.** *A 2-prover commitment scheme $\mathcal{S}$ is* ε-weak-binding *if for all commit strategies $\overline{\mathsf{com}}_{PQ}$ there exists a distribution $p(\hat{s})$ such that for every opening strategy $\overline{\mathsf{open}}_{PQ}$ (which then fixes the distribution $p(s)$ of V's output s) there is a consistent joint distribution $p(\hat{s}, s)$ such that $p(s \neq \hat{s} \wedge s \neq \perp) \leq \varepsilon$. In short:*

$$\forall\, \overline{\mathsf{com}}_{PQ}\ \exists\, p(\hat{s})\ \forall\, \overline{\mathsf{open}}_{PQ}\ \exists\, p(\hat{s}, s) : p(s \neq \hat{s} \wedge s \neq \perp) \leq \varepsilon\,. \tag{3.3}$$

We also consider a related, i.e., "fairly", version of this binding property, similar to Definition 3.2.

Definition 3.9 (Fairly weak binding property)**.** *We say that a 2-prover commitment scheme $\mathcal{S}$ is* ε-fairly-weak-binding *if for all commit strategies $\overline{\mathsf{com}}_{PQ}$ there exists a distribution $p(\hat{s})$ such that for every opening strategy $\overline{\mathsf{open}}_{PQ}$ (which then fixes the distribution $p(s)$ of V's output s) there is a consistent joint distribution $p(\hat{s}, s)$ so that for all $s_\circ \in \{0,1\}^n$, $p(s \neq \hat{s} \wedge s = s_\circ) \leq \varepsilon$. In short:*

$$\forall\, \overline{\mathsf{com}}_{PQ}\ \exists\, p(\hat{s})\ \forall\, \overline{\mathsf{open}}_{PQ}\ \exists\, p(\hat{s}, s)\ \forall\, s_\circ : p(s \neq \hat{s}\ \wedge\ s = s_\circ) \leq \varepsilon\,. \tag{3.4}$$

Remark 3.10. *Lemma 3.3 and Remark 3.5 also hold for the weak binding properties. Furthermore, it is easy to see that the binding and fairly-binding properties imply their weak counterparts.*

Proposition 3.11. *Let $\mathcal{S}$ be a string-commitment scheme and define $\mathcal{S}'$ as in Proposition 3.6. If $\mathcal{S}$ is ε-fairly-weak-binding, then $\mathcal{S}'$ is a 2ε-weak-binding* bit*-commitment scheme.*

Proof. The proof of Proposition 3.6 can be easily adapted: Let $p(\hat{s})$ be as required by Definition 3.9. We define $p(\hat{b})$ by taking the marginal of $p(\hat{s}, \hat{b})$ where $\hat{b} = 0$ if $\hat{s} = s_0$, and $\hat{b} = 1$ otherwise. An opening strategy $\overline{\mathsf{open}}_{PQ}$ for $\mathcal{S}'$ can also be viewed as a strategy for $\mathcal{S}$. As such, there is a joint distribution $p(\hat{s}, s)$ as required by Definition 3.8 which we can extend to $p(\hat{s}, s, b)$ by setting $b = 0$ if $s = s_0$, $b = 1$ if $s = s_1$ and $b = \perp$ otherwise. We define $p(\hat{b}, b) := \sum_{\hat{s},s} p(\hat{s}, \hat{b}) \cdot p(s, b|\hat{s})$. As in the proof of Proposition 3.6, one can easily check that $p(\hat{b} \neq b \neq \perp) \leq 2\varepsilon$ holds. □

3.2.4 Relations Between The Definitions

Here, we show that in case of *bit*-commitment schemes, the weak binding property as introduced in Definition 3.8 above is actually *equivalent* to the sum-binding-definition. Even though our focus in this chapter is on classical attacks, the proof immediately carries over to quantum attacks as well.

Theorem 3.12. *A 2-prover bit-commitment scheme is ε-sum-binding if and only if it is ε-weak-binding.*

Proof. First, consider a scheme that is ε-binding according to Definition 2.14. Fix an arbitrary commit strategy $\overline{\mathsf{com}}_{PQ}$. Let $\overline{\mathsf{open}}^0_{PQ}$ and $\overline{\mathsf{open}}^1_{PQ}$ be opening strategies so that $p_0 = p(b_0 = 0)$ and $p_1 = p(b_1 = 1)$ are both maximized, where $b_i \in \{0, 1, \perp\}$ is V's output when the dishonest provers use the commit strategy $\overline{\mathsf{com}}_{PQ}$ and opening strategy $\overline{\mathsf{open}}^i_{PQ}$. Let ε' be such that $p_0 + p_1 = 1 + 2\varepsilon'$. Since the scheme is ε-binding, we have $\varepsilon' \leq \varepsilon$. We define the distribution $p(\hat{b})$ as $p(\hat{b} = 0) := p_0 - \varepsilon'$ and $p(\hat{b} = 1) := p_1 - \varepsilon'$. To see that this is indeed a probability distribution, note that $p_0, p_1 \geq 2\varepsilon'$ (otherwise, we would have $p_0 > 1$ or $p_1 > 1$) and that $p(\hat{b} = 0) + p(\hat{b} = 1) = p_0 + p_1 - 2\varepsilon' = 1$. Now we consider an arbitrary opening strategy $\overline{\mathsf{open}}_{PQ}$ which fixes a distribution $p(b)$. By definition of p_0 and p_1, we have $p(b = i) \leq p_i$ and thus $p(b = i) \leq p(\hat{b} = i) + \varepsilon' \leq p(\hat{b} = i) + \varepsilon$. By Lemma 2.2, there exists a consistent joint distribution $p(\hat{b}, b)$ such that $p(\hat{b} = b = i) = \min\{p(b = i), p(\hat{b} = i)\}$. We wish to bound $p(\hat{b} \neq b \wedge b \neq \perp) = p(\hat{b} = 0 \wedge b = 1) + p(\hat{b} = 1 \wedge b = 0)$. For $i \in \{0, 1\}$, it holds that

$$\begin{aligned} p(\hat{b} = 1 - i \wedge b = i) &= p(b = i) - p(\hat{b} = b = i) \\ &= p(b = i) - \min\{p(\hat{b} = i), p(b = i)\} \\ &= \max\{0, p(b = i) - p(\hat{b} = i)\} \\ &\leq \varepsilon \end{aligned}$$

and furthermore, there is at most *one* $i \in \{0, 1\}$ such that $p(b = i) > p(\hat{b} = i)$, for if $p(b = i) > p(\hat{b} = i)$ for both $i = 0$ and $i = 1$, then $p(b = 0) + p(b = 1) > p(\hat{b} = 0) + p(\hat{b} = 1) = 1$ which is a contradiction. Thus, we have $p(\hat{b} \neq b \wedge b \neq \perp) \leq \varepsilon$. This proves one direction of our claim.

For the other direction, consider a scheme that is ε-weak-binding. Fix $\overline{\mathsf{com}}_{PQ}$ and let $p(\hat{b})$ be a distribution such that for every opening strategy $\overline{\mathsf{open}}_{PQ}$, there is a joint distribution $p(\hat{b}, b)$ with $p(\hat{b} \neq b \neq \perp) \leq \varepsilon$. Now consider two opening strategies $\overline{\mathsf{open}}^0_{PQ}$ and $\overline{\mathsf{open}}^1_{PQ}$ which give distributions $p(b_0)$ and $p(b_1)$. We need to bound $p(b_0 = 0) + p(b_1 = 1)$. There is a joint

distribution $p(\hat{b}, b_0)$ such that $p(\hat{b} \neq b_0 \neq \perp) \leq \varepsilon$ and likewise for b_1. Thus,

$$\begin{aligned} & p(b_0 = 0) + p(b_1 = 1) \\ = \; & p(\hat{b} = 0, b_0 = 0) + p(\hat{b} = 1, b_0 = 0) + p(\hat{b} = 0, b_1 = 1) + p(\hat{b} = 1, b_1 = 1) \\ \leq \; & p(\hat{b} = 0) + p(\hat{b} = 1) + p(\hat{b} \neq b_0 \neq \perp) + p(\hat{b} \neq b_1 \neq \perp) \\ \leq \; & 1 + 2\varepsilon \end{aligned}$$

which proves the other direction. □

Remark 3.13. *By Remark 3.10, it follows that Definition 3.1 also implies the sum-binding-definition. In fact, Definition 3.1 is* strictly *stronger (and hence, also strictly stronger than the weak-binding definition). Consider the following (artificial and very non-complete) scheme: In the commit phase, V chooses a uniformly random bit and sends it to the provers, and then accepts anything or rejects anything during the opening phase, depending on that bit. Then, $p_0 + p_1 = 1$, yet a commitment can be opened to $1 - \hat{b}$ (no matter how $\hat{b}$ is defined) with probability $\frac{1}{2}$.*

Since a non-complete separation example may not be fully satisfying, we note that it can be converted into a complete (but even more artificial) scheme. Fix a "good" (i.e., complete, hiding and binding with low parameters) scheme and call our example scheme above the "bad" scheme. We define a combined *scheme as follows: At the start, the first prover can request either the "good" or "bad" scheme to be used. The honest prover is instructed to choose the former, guaranteeing completeness. The dishonest prover may choose the latter, so the combined scheme inherits the binding properties of the "bad" scheme: it is binding according to the sum-binding-definition, but not according to Definition 3.1.*

3.2.5 Simultaneous Opening

The binding definitions from the previous sections are useful for proving our composition theorem, but it is not clear how to prove in a straightforward way that a commitment scheme satisfies those definitions. In this section, we propose another definition which is easier to check and which implies the binding properties from the previous sections (with some loss in the parameter). We then use this result in Section 3.2.6 to prove that $\mathcal{CHSH}^q$ is fairly-binding.

This binding property is based on the intuition that it should not be possible to open a commitment to two different values *simultaneously* (except with small probability). For this, we observe that when considering a commit strategy $\overline{\mathsf{com}}_{PQ}$, as well as *two* opening strategies $\overline{\mathsf{open}}_{PQ}$ and $\overline{\mathsf{open}}'_{PQ}$, we can run both opening strategies *simultaneously* on the produced commitment with two independent copies of open_V, by applying $\overline{\mathsf{open}}_{PQ}$ and $\overline{\mathsf{open}}'_{PQ}$ to two copies of the respective internal states of P and Q). This gives rise to a

joint distribution $p(s, s')$ of the respective outputs s and s' of the two copies of open_V.

Definition 3.14 and Theorem 3.18 were first considered in [Sca16].

Definition 3.14. *A 2-prover commitment scheme $\mathcal{S}$ is* ε-binding in the sense of simultaneous opening *if for all $\overline{\mathsf{com}}_{PQ}$ and all pairs of opening strategies $\overline{\mathsf{open}}_{PQ}$ and $\overline{\mathsf{open}}'_{PQ}$, it holds that $p(s \neq s' \wedge s \neq \bot \wedge s' \neq \bot) \leq \varepsilon$.*

Definition 3.15. *A 2-prover commitment scheme $\mathcal{S}$ is* ε-fairly-binding in the sense of simultaneous opening *if for all $\overline{\mathsf{com}}_{PQ}$, all pairs of opening strategies $\overline{\mathsf{open}}_{PQ}$ and $\overline{\mathsf{open}}'_{PQ}$, and all pairs $s_\circ, s'_\circ$ of distinct strings, it holds that $p(s = s_\circ \wedge s' = s'_\circ) \leq \varepsilon$.*

Remark 3.16. *Also for this notion of binding, it is sufficient to consider* deterministic *strategies, as can easily be seen.*

Remark 3.17. *It follows directly from Eq.* (2.1) *that every* bit*-commitment scheme that is ε-fairly-binding in the sense of simultaneous opening (against classical attacks) is $\varepsilon/2$-sum-binding (and thus also according to Definition 3.8). The converse is not true though: the schemes from Remark 3.13 again serve as counterexamples.*

Theorem 3.18. *Let $\mathcal{S} = (\mathsf{com}_{PQV}, \mathsf{open}_{PQV})$ be a 2-prover commitment scheme. If it is ε-binding in the sense of simultaneous opening, and open_V is deterministic, then $\mathcal{S}$ is $2\sqrt{\varepsilon}$-binding.*

Proof. By Lemma 3.3, it suffices to consider deterministic strategies for the provers. We fix some deterministic commit strategy $\overline{\mathsf{com}}_{PQ}$ and an enumeration $\{\overline{\mathsf{open}}^i_{PQ}\}_{i=1}^N$ of all deterministic opening strategies. Since we assume that open_V is deterministic, for any fixed deterministic opening strategy for the provers, the verifier's output s is a *function* of the commitment c. Thus, for each opening strategy $\overline{\mathsf{open}}^i_{PQ}$ there is a function f_i such that the verifier's output is $s = f_i(c)$. We will now define the function $\hat{s}(c)$ that satisfies the properties required by Definition 3.1. We will now define the function $\hat{s}(c)$ that satisfies the properties required by Definition 3.1. Our definition depends on a parameter $\alpha > 0$ which we fix later. To define $\hat{s}$, we partition the set C of all possible commitments into disjoint sets $C = R \cup \bigcup_i C_i$ that satisfy the following three properties for every i:

- $f_i(c) \neq \bot$ for all $c \in C_i$,
- $p(c \in C_i) \geq \alpha$ or $C_i = \emptyset$,
- and $p(c \in R \wedge f_i(c) \neq \bot) < \alpha$

The second property implies that there are at most α^{-1} non-empty sets C_i. It is easy to see that such a partitioning exists: Start with $R = C$ and while there exists some i with $p(c \in R \wedge f_i(c) \neq \bot) \geq \alpha$, let $C_i = \{c \in R \mid f_i(c) \neq \bot\}$

and remove the elements of C_i from R. For any $c \in C$, we now define $\hat{s}(c)$ as follows. We set $\hat{s}(c) = f_i(c)$ if $c \in C_i$ and $\hat{s}(c) = 0$ for $c \in R$.

Now fix some opening strategy $\overline{\mathsf{open}}^i_{PQ}$ and write s_i for the verifier's output. It follows that

$$\begin{aligned}
& p(s_i \neq \hat{s}(c) \wedge s_i \neq \perp) \\
= \; & p(f_i(c) \neq \hat{s}(c) \wedge f_i(c) \neq \perp) \\
\leq \; & p(c \in R \wedge f_i(c) \neq \perp) + \sum_j p(f_i(c) \neq \hat{s}(c) \wedge f_i(c) \neq \perp \wedge c \in C_j) \\
< \; & \alpha + \sum_{j : C_j \neq \emptyset} P(f_i(c) \neq f_j(c) \wedge f_i(c) \neq \perp \wedge f_j(c) \neq \perp) \\
\leq \; & \alpha + \alpha^{-1} \cdot \varepsilon
\end{aligned}$$

where the final inequality holds because $p(f_i(c) \neq f_j(c) \wedge f_i(c) \neq \perp \wedge f_j(c) \neq \perp) \leq \varepsilon$ by the assumed binding property. It is easy to see that the upper bound $\alpha + \alpha^{-1} \cdot \varepsilon$ is minimized by setting $\alpha = \sqrt{\varepsilon}$. We conclude that $p(s_i \neq \hat{s}(c) \wedge s_i \neq \perp) < 2\sqrt{\varepsilon}$. □

Theorem 3.19. *Let $\mathcal{S} = (\mathsf{com}_{PQV}, \mathsf{open}_{PQV})$ be a 2-prover commitment scheme. If $\mathcal{S}$ is ε-fairly-binding in the sense of simultaneous opening and open_V is deterministic, then $\mathcal{S}$ is $2\sqrt{\varepsilon}$-fairly-binding.*

Proof. It again suffices to consider deterministic strategies for the provers. As in the previous proof, we fix a deterministic commit strategy $\overline{\mathsf{com}}_{PQ}$ and an enumeration $\{\overline{\mathsf{open}}^i_{PQ}\}_{i=1}^N$ of the deterministic opening strategies. The verifier's output when the provers use $\overline{\mathsf{open}}^i_{PQ}$ is a function $f_i(c)$ of the commitment. We now define the function $\hat{s}(c)$ that satisfies the properties required by Definition 3.2. Our definition again depends on a parameter $\alpha > 0$. We partition the set C of all possible commitments into disjoint sets $R \cup \bigcup_{s,i} C_{s,i} = C$ that satisfy the following three properties for every i and every s:

$$C_{s,i} \subseteq f_i^{-1}(\{s\}), \;\; p(c \in C_{s,i}) \geq \alpha \text{ or } C_{s,i} = \emptyset, \text{ and } p(c \in R \wedge f_i(c) = s) < \alpha\,.$$

The second property implies that there are at most α^{-1} non-empty sets $C_{s,i}$. Similar to the previous proof, it is easy to see that such a partitioning exists. For any $c \in C$, we now define $\hat{s}(c)$ as follows. We set $\hat{s}(c) = s$ for $c \in C_{s,i}$ and $\hat{s}(c) = 0$ for $c \in R$.

Now fix some opening strategy $\overline{\mathsf{open}}^i_{PQ}$ and a string $s_\circ$, and write s_i for the verifier's output. Using $C_{\neq s_\circ}$ as a shorthand for $\bigcup_{s \neq s_\circ} \bigcup_j C_{s,j}$, we note

that if $\hat{s}(c) \neq s_\circ$ then $c \in R \cup C_{\neq s_\circ}$. Thus, it follows that

$$\begin{aligned}
& p(s_i \neq \hat{s}(c) \wedge s_i = s_\circ) \\
= \; & p(\hat{s}(c) \neq s_\circ \wedge s_i = s_\circ) \\
\leq \; & p\big(c \in (R \cup C_{\neq s_\circ}) \wedge f_i(c) = s_\circ\big) \\
= \; & p(c \in R \wedge f_i(c) = s_\circ) + \sum_{s \neq s_\circ, j} p(c \in C_{s,j} \wedge f_i(c) = s_\circ) \\
\leq \; & p(c \in R \wedge f_i(c) = s_\circ) + \sum_{\substack{s \neq s_\circ, j \\ \text{s.t. } C_{s,j} \neq \emptyset}} p(f_j(c) = s \wedge f_i(c) = s_\circ) \\
< \; & \alpha + \alpha^{-1} \cdot \varepsilon
\end{aligned}$$

where the final inequality holds because $p(f_j(c) = s \wedge f_i(c) = s_\circ) \leq \varepsilon$ by the assumed binding property. Again, we minimize the upper bound by setting $\alpha = \sqrt{\varepsilon}$ which completes the proof. □

For the fairly-weak-binding property, we can get better parameters. Also note that we do not require open_V to be deterministic here.

Theorem 3.20. *Every 2-prover commitment scheme $\mathcal{S}$ that is ε-fairly-binding in the sense of simultaneous opening is $\sqrt{2\varepsilon}$-fairly-weak-binding.*

Proof. Fix a commit strategy $\overline{\mathsf{com}}_{PQ}$ against $\mathcal{S}$. Enumerate all strings in the domain D of $\mathcal{S}$ as $s^1_\circ, \ldots, s^d_\circ$, and for every i, let $\overline{\mathsf{open}}^i_{PQ}$ be an opening strategy maximizing $p_i := p(s_i = s^i_\circ)$, where s_i is the output of the verifier when P and Q use this strategy. We assume without loss of generality that the p_i are in descending order. We define $p(\hat{s})$ as follows. Let $N \geq 2$ be an integer which we will fix later. By Definition 3.14 and Inequality (2.2), it holds that

$$\sum_{i=1}^{N} p_i \leq 1 + \binom{N}{2} \cdot \varepsilon = 1 + \frac{N(N-1)}{2} \cdot \varepsilon$$

where we let $p_i = 0$ for $i > d$ in case $N > d$. We would like to define $p(\hat{s})$ as $p(\hat{s} = s^i_\circ) := p_i - (N-1)\varepsilon/2$ for all $i \leq N, d$; however, this is not always possible because $p_i - (N-1)\varepsilon/2$ may be negative. To deal with this, let N' be the largest integer such that $N' \leq N$ and $p_1, \ldots, p_{N'} \geq (N-1)\varepsilon/2$. (We take $N = 0$ if $p_1 < (N-1)\varepsilon/2$.) It follows that

$$\sum_{i=1}^{N'} p_i \leq 1 + \frac{N'(N'-1)}{2} \cdot \varepsilon \leq 1 + \frac{N'(N-1)}{2} \cdot \varepsilon$$

and thus

$$\sum_{i=1}^{N'} p_i = 1 + \frac{N'(N-1)}{2} \cdot \tilde{\varepsilon}$$

for some $\tilde{\varepsilon} \leq \varepsilon$. We now set $p(\hat{s})$ to be $p(\hat{s} = s_i) := p_i - (N-1)\tilde{\varepsilon}/2 \geq p_i - (N-1)\varepsilon/2 \geq 0$ for all $i \leq N'$. Now consider an opening strategy $\overline{\mathsf{open}}_{PQ}$ and let $p(s)$ be the resulting output distribution. By definition of the p_i, it follows that $p(s = s^i_\circ) \leq p_i$ for all $i \leq d$, and $p_i \leq p(\hat{s} = s^i_\circ) + (N-1)\varepsilon/2$ for all $i \leq N'$. By Lemma 2.2, we can conclude that there exists a consistent joint distribution $p(\hat{s}, s)$ with $p(\hat{s} = s = s^i_\circ) = \min\{p(s = s^i_\circ), p(\hat{s} = s^i_\circ)\} \geq p(s = s_i) - (N-1)\varepsilon/2$ for all $i \leq N'$, and thus $p(\hat{s} \neq s = s^i_\circ) = p(s = s^i_\circ) - p(\hat{s} = s = s^i_\circ) \leq (N-1)\varepsilon/2$ for all $i \leq N'$. Furthermore, when $N' < i \leq N$, we have $p(\hat{s} \neq s = s^i_\circ) = p(s = s^i_\circ) \leq p_i < (N-1)\varepsilon/2$ by definition of N'. Since the p_i are sorted in descending order, it follows that for all $i > N$

$$p(\hat{s} \neq s = s^i_\circ) = p(s = s^i_\circ) \leq p_i \leq p_N \leq \frac{1}{N}\sum_{i=1}^{N} p_i \leq \frac{1}{N} + \frac{N-1}{2} \cdot \varepsilon$$

and thus, we have shown for all $s_\circ \in D$ that

$$p(\hat{s} \neq s = s_\circ) \leq \frac{1}{N} + \frac{N-1}{2} \cdot \varepsilon.$$

We now select N so that this value is minimized: it is easy to verify that the function $f : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$, $x \mapsto 1/x + (x-1)\varepsilon/2$ has its global minimum in $\sqrt{2/\varepsilon}$; thus, we pick $N := \lceil \sqrt{2/\varepsilon} \rceil$, which gives us

$$p(\hat{s} \neq s = s_\circ) \leq \frac{1}{N} + \frac{N-1}{2} \cdot \varepsilon \leq \frac{1}{\sqrt{2/\varepsilon}} + \frac{\sqrt{2/\varepsilon}}{2} \cdot \varepsilon = \sqrt{2\varepsilon}$$

for any $s_\circ \in D$, as claimed. □

3.2.6 Security of $\mathcal{CHSH}^q$

Using the results from the previous section, we now show that $\mathcal{CHSH}^q$ is a fairly-binding string-commitment scheme. It is understood that the possible attacks against $\mathcal{CHSH}^q$ are those where the provers do not communicate.

Proposition 3.21. *The string-commitment scheme $\mathcal{CHSH}^q$ is q^{-1}-fairly-binding in the sense of simultaneous opening.*

Proof. By Remark 3.16, it suffices to consider deterministic attack strategies. Fix a deterministic strategy $\overline{\mathsf{com}}_{PQ}$ and two deterministic opening strategies $\overline{\mathsf{open}}_{PQ}$ and $\overline{\mathsf{open}}'_{PQ}$. The strategy $\overline{\mathsf{com}}_{PQ}$ specifies P's output x as a function $f(a)$ of the verifier's message a. The opening strategies are described by constants y and y'. By definition of $\mathcal{CHSH}^q$, $s = s_\circ$ implies $f(a) + y = a \cdot s_\circ$ and likewise, $s' = s'_\circ$ implies $f(a) + y' = a \cdot s'_\circ$. The condition $s = s_\circ \wedge s' = s'_\circ$ can hold only if $a = (y - y')/(s_\circ - s'_\circ)$. It follows that $p(s = s_\circ \wedge s' = s'_\circ) \leq p\big(a = (y - y')/(s_\circ - s'_\circ)\big) \leq q^{-1}$, which proves our claim. □

From Proposition 3.21, Theorem 3.19 and Theorem 3.20, we conclude that the following corollaries hold.

Corollary 3.22. *$\mathcal{CHSH}^q$ is $2\sqrt{q^{-1}}$-fairly-binding.*

Corollary 3.23. *$\mathcal{CHSH}^q$ is $\sqrt{2q^{-1}}$-fairly-weak-binding.*

Remark 3.24. *It is not too hard to see that Corollary 3.23 above implies an upper bound on the classical value ω of the game CHSH_{2^n} considered in [BS15] of $\omega(\mathsf{CHSH}_{2^n}) \leq 2^{-\frac{n-1}{2}} + 2^{-n}$. As such, Theorem 1.3 in [BS15] implies that the above ε is asymptotically optimal for odd n, i.e., the square root loss to the binding property of the* bit*-commitment version is unavoidable (for odd n).*

As for security against quantum attacks, we point out that [BS15, RAM16] provide an upper bound on the quantum value $\omega^*(\mathsf{CHSH}_q)$ of general finite-field CHSH; however, this does not directly imply security against quantum attacks of $\mathcal{CHSH}^q$ as a (fairly-weak-binding) string-commitment scheme.

Furthermore, we show that a variant of $\mathcal{CHSH}^q$ is $2\sqrt{q^{-1}}$-binding. However, this variant requires the opening information to be twice as large as the domain of $\mathcal{CHSH}^q$, so it is not possible to compose multiple instances of this variant using our composition theorem (see Definition 4.1).

Corollary 3.25. *Let $\mathcal{CHSH}^q_+$ be the scheme defined as follows: The commit phase is the same as in $\mathcal{CHSH}^q$. In the opening phase, Q sends the opening information and the string s that the provers committed to. Then, V opens the commitment as in $\mathcal{CHSH}^q$ and checks if the result equals the string s he received from Q. If yes, he outputs s and if not, $\perp$. This scheme is $2\sqrt{q^{-1}}$-binding and $\sqrt{2q^{-1}}$-weak-binding.*

Proof. Let $\overline{\mathsf{open}}_{PQ}$ be a dishonest strategy for the opening phase of $\mathcal{CHSH}^q_+$. Let $s_\circ$ be the string that Q sends along with the opening information. Since V does not send any messages to Q in $\mathcal{CHSH}^q_+$, the string $s_\circ$ is computed as a function of the provers' randomness. From the strategy $\overline{\mathsf{open}}_{PQ}$, a strategy $\overline{\mathsf{open}}'_{PQ}$ for $\mathcal{CHSH}^q$ can be extracted by simply leaving out $s_\circ$. By Lemma 3.4 and Corollary 3.22, we conclude that if the provers use $\overline{\mathsf{open}}'_{PQ}$ in $\mathcal{CHSH}^q$, we have $p(s \neq \hat{s} \wedge s = s_\circ) \leq 2\sqrt{q^{-1}}$. It follows that when they use $\overline{\mathsf{open}}_{PQ}$ in $\mathcal{CHSH}^q_+$, we have $p(s \neq \hat{s} \wedge s \neq \perp) \leq 2\sqrt{q^{-1}}$. The result for the weak-binding property follows similarly, using Remark 3.10 and Corollary 3.23. □

While it may seem like a similar idea could be used to transform *any* fairly-binding scheme into a binding scheme at the cost of increasing the size of the opening information, the proof above relies on the assumption that the second prover can not choose the message $s_\circ$ depending on any information sent by the verifier. Otherwise, Lemma 3.4 does not apply.

As a counter-example, consider another variant of $\mathcal{CHSH}^q$ similar to $\mathcal{CHSH}^q_+$ where in the opening phase, P sends the string s and Q sends the opening information. The following strategy breaks the binding property of this variant:

In the commit phase, P sends $x = 1$. The provers can open to 0 by sending 0 and 1 respectively in the opening phase. Since this strategy always opens to 0, $p(\hat{s} = 0)$ needs to be large. On the other hand, if P sends $s_\circ = a^{-1}$ (if it exists) and Q sends 0, it holds that $p(s = s_\circ)$ is large and $p(s_\circ \neq \hat{s})$ is small.

Chapter 4

The Composition Theorem

4.1 Composition of Commitment Schemes

4.1.1 Introduction

In this chapter, we present one of the central results of this thesis: the composition theorems. We define a *composition operation* in Definition 4.3 which constructs a *composed scheme* out of two *component schemes*. This is achieved as follows: The provers use the first scheme to commit, but instead of opening the commitment, they instead commit to the opening information using the second scheme. They then open this second commitment, which allows the verifier to obtain the opening information of the first scheme and thus open the first commitment.

The composition operation cannot be applied to any two schemes – we define a notion of *eligible pairs* (Definition 4.1) that it can be applied to. Some of the requirements in that definition are necessary for the composition operation to make sense. For example, the composition operation requires that the first scheme is structured so that in the opening phase, a prover sends some *opening information* to the verifier who then computes the result. Other requirements are necessary for our proofs of the composition theorems to work.

The composition theorems show that if the first scheme is ε-binding (according to some binding definition) and the second one is δ-binding (according to the same definition), then the composed scheme is $(\epsilon + \delta)$-binding.

While this is what one would expect from this composition operation, it is non-trivial to formally prove. It is unclear how one would prove this result using the sum-binding definition directly, rather than our newly-introduced definitions.

Together with our analysis of the CHSH scheme in Section 3.2.6, the composition theorem implies that the binding parameter of the Lunghi *et al.* scheme decays linearly in the number m of rounds, rather than double-exponentially,

as [LKB+15] suggests.

4.1.2 Eligible Pairs and the Composition Operation

We consider two 2-prover commitment schemes $\mathcal{S}$ and $\mathcal{S}'$ of a restricted form, and we compose them to a new 2-prover commitment scheme $\mathcal{S}'' = \mathcal{S} \star \mathcal{S}'$ in a well-defined way; our composition theorem then shows that $\mathcal{S}''$ is secure (against classical attacks) if $\mathcal{S}$ and $\mathcal{S}'$ are. We start by specifying the restriction to $\mathcal{S}$ and $\mathcal{S}'$ that we impose.

Definition 4.1. *Let $\mathcal{S}$ and $\mathcal{S}'$ be two 2-prover string-commitment schemes with domains D and D', respectively. We call the pair $(\mathcal{S}, \mathcal{S}')$* eligible *if the following two properties hold, or they hold with the roles of P and Q exchanged.*

1. *The commit phase of $\mathcal{S}$ is a protocol com_{PQV} that involves communication between P and V only, and the opening phase of $\mathcal{S}$ is a protocol $\mathsf{open}_{QV} = (\mathsf{open}_Q, \mathsf{open}_V)$ between Q and V only. In other words, com_Q and open_P are both trivial and do nothing.*[1] *Similarly, the commit phase of $\mathcal{S}'$ is a protocol com'_{PQV} that involves communication between Q and V only (but both provers may be active in the opening phase).*

2. *The opening phase open_{QV} of $\mathcal{S}$ is of the following simple form: Q sends some $y \in D'$ to V, and V computes s deterministically as $s = \mathsf{Extr}(y, c)$, where c is the commitment.*[2] *We call this message y the* opening information*.*

Furthermore, we specify that the possible attacks on $\mathcal{S}$ are so that P and Q do not communicate during the course of the entire scheme, and the possible attacks on $\mathcal{S}'$ are so that P and Q do not communicate during the course of the commit phase but there may be limited communication during the opening phase.

An example of an eligible pair of 2-prover commitments is $(\mathcal{CHSH}^q, \mathcal{XCHSH}^q)$, where $\mathcal{XCHSH}^q$ coincides with scheme $\mathcal{CHSH}^q$ except that the roles of P and Q are exchanged.

Remark 4.2. *For an eligible pair $(\mathcal{S}, \mathcal{S}')$, it will be convenient to understand open_Q and open_V as* non-interactive *algorithms, where open_Q produces y as its* output*, and open_V takes y as additional* input *(rather than viewing the pair as a protocol with a single one-way communication round).*

We now define the composition operation. Informally, committing is done by means of committing using $\mathcal{S}$, and to open the commitment, Q uses open_Q

[1] Except that com_Q may output state information to the opening protocol open_Q, e.g., in order to pass on the commit phase randomness.

[2] Our composition theorem also works for a randomized Extr, but for simplicity, we restrict to the deterministic case.

to locally compute the opening information y and he commits to y with respect to the scheme $\mathcal{S}'$, and then this commitment is opened (to y), and V computes and outputs $s = \mathsf{Extr}(y, c)$. Formally, this is captured as follows (see also Fig. 4.1).

Definition 4.3. *Let $\mathcal{S} = (\mathsf{com}_{PV}, \mathsf{open}_{QV})$ and $\mathcal{S}' = (\mathsf{com}'_{QV}, \mathsf{open}'_{PQV})$ be an eligible pair of 2-prover commitment schemes. Then, their composition $\mathcal{S} \star \mathcal{S}'$ is defined as the 2-prover commitment scheme consisting of $\mathsf{com}_{PV} = (\mathsf{com}_P[\xi_{PQ}], \mathsf{com}_V)$ and*

$$\mathsf{open}''_{PQV} = (\mathsf{open}'_P,\ \mathsf{open}'_Q \circ \mathsf{com}'_Q \circ \mathsf{open}_Q,\ \mathsf{open}_V \circ \mathsf{open}'_V \circ \mathsf{com}'_V)\,,$$

where we make it explicit that com_P and open_Q use joint randomness, and so do com'_Q and open'_P.

When considering attacks against the binding property of the composed scheme $\mathcal{S} \star \mathcal{S}'$, we declare that the possible deterministic attacks[3] are those of the form $(\overline{\mathsf{com}}_P, \overline{\mathsf{open}}'_{PQ} \circ \mathsf{ptoq}_{PQ} \circ \overline{\mathsf{com}}'_Q)$, where $\overline{\mathsf{com}}_P$ is an possible deterministic commit strategy for $\mathcal{S}$, $\overline{\mathsf{com}}'_Q$ and $\overline{\mathsf{open}}'_{PQ}$ are possible deterministic commit and opening strategies for $\mathcal{S}'$, and ptoq_{PQ} is the one-way communication protocol that communicates P's input to Q (see also Fig. 4.2).[4]

Remark 4.4. *We point out that the composition $\mathcal{S} \star \mathcal{S}'$ can be naturally defined for a* larger *class of pairs of schemes (e.g. where* both *provers are active in the commit phase of both schemes), and the above intuition still holds. However, our proof only works for this restricted class of pairs of schemes. Extending the composition result in that direction is an open problem.*

Remark 4.5. *We observe that if $(\mathcal{S}, \mathcal{XS})$ is an eligible pair, where $\mathcal{XS}$ coincides with $\mathcal{S}$ except that the roles of P and Q are exchanged, then so is $(\mathcal{XS}, \mathcal{S} \star \mathcal{XS})$. As such, we can then compose $\mathcal{XS}$ with $\mathcal{S} \star \mathcal{XS}$, and obtain yet another eligible pair $(\mathcal{S}, \mathcal{XS} \star \mathcal{S} \star \mathcal{XS})$, etc. We write $\mathcal{S}_m$ for the m-fold composition of $\mathcal{S}$ with itself, i.e., $\mathcal{S}_{\Updownarrow} = \mathcal{S} \star \mathcal{XS} \star \mathcal{S} \star \ldots$ for m terms. Applying this to the schemes $\mathcal{S} = \mathcal{CHSH}^q$, we obtain the multi-round scheme from Lunghi et al. [LKB+15]. As such, our composition theorem below implies security of their scheme — with a* linear *blow-up of the error term (instead of double exponential).*

We point out that formally we obtain security of the Lunghi et al. scheme as a 2-prover commitment scheme *under an* abstract restriction *on the provers' communication: in every round, the active prover cannot access the message that the other prover received in the previous round. As such, when the rounds of the protocol are executed fast enough so that it is ensured that there is no*

[3] The possible *randomized* attacks are then naturally given as those that pick one of the deterministic attacks according to some distribution.

[4] This one-way communication models that in the relativistic setting, sufficient time has passed at this point for P to inform Q about what happened during com_P.

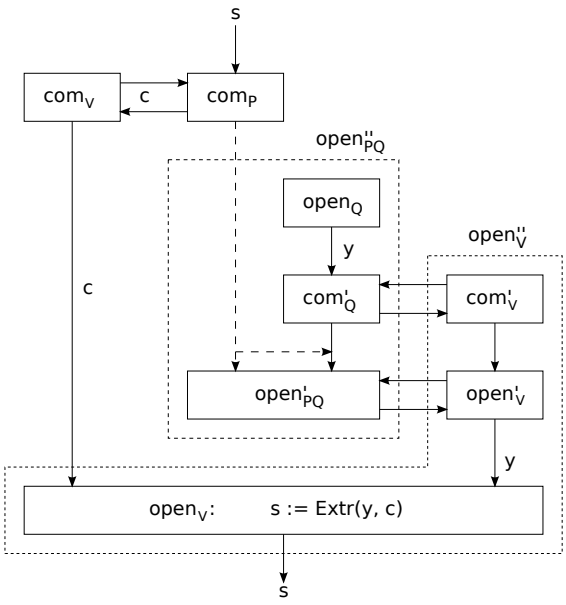

Figure 4.1: The composition of $\mathcal{S}$ and $\mathcal{S}'$ (assuming single-round commit phases). The dotted arrows indicate communication possible to the dishonest provers.

time for the provers to communicate between subsequent rounds, then security as a relativistic commitment scheme *follows immediately.*

Remark 4.6. *In Section 2.2.2, we noted that the verifier in a relativistic commitment scheme also needs to be convinced that the provers keep the appropriate distance from each other. In some schemes, this can be achieved by splitting up the verifier into two entities and placing them at the same locations as the provers. If $(\mathcal{S}, \mathcal{XS})$ is an eligible pair, then $\mathcal{S} \star \mathcal{XS}$, and more generally $\mathcal{S}_m$, allow for the verifier to be split up in this way. This holds because in each round before the last, the verifier initiates a new commitment.*

Remark 4.7. *It is immediate that $\mathcal{S} \star \mathcal{S}'$ is a commitment scheme in the sense of Definition 2.8, and that it is complete if $\mathcal{S}$ and $\mathcal{S}'$ are, with the error parameters adding up. It is intuitively clear that $\mathcal{S} \star \mathcal{S}'$* should be *binding if $\mathcal{S}$ and $\mathcal{S}'$ are: committing to the opening information y and then opening the commitment allows the provers to* delay *the announcement of y (which is the whole point of the exercise), but it does not allow them to* change *y, by the binding property of $\mathcal{S}'$; thus, $\mathcal{S} \star \mathcal{S}'$ should be (almost) as binding as $\mathcal{S}$. This intuition is confirmed by our composition theorem below.*

4.1.3 Hiding Property for Composed Schemes

The hiding property is obviously inherited from $\mathcal{S}$, i.e., $\mathcal{S} \star \mathcal{S}'$ is δ-binding in the sense of Definition 2.13 if and only if $\mathcal{S}$ is δ-hiding. However, this definition is not suitable for schemes with a multi-round opening phase: a scheme that reveals the committed string s in the first round of the opening phase would still satisfy Definition 2.13, but clearly, doing so defeats the entire purpose of a multi-round commitment scheme. Recall that, using the terminology used in context of relativistic commitments, the rounds of the opening phase up to before the last are referred to as the *sustain phase*, and only the last round is considered the opening phase proper. We show in this section that $\mathcal{S} \star \mathcal{S}'$ is hiding up to before the last round, with the error parameters adding up.

Definition 4.8. *Let $\mathcal{S} = (\mathsf{com}_{PQV}, \mathsf{open}_{PQV})$ be a commitment scheme. We write v for the verifier's view immediately before the last round of communication in open_{PQV}. We say that a scheme is* ε-hiding until the last round *if for any (possibly dishonest) verifier V and any two inputs s_0 and s_1 to the honest provers, we have $d(p(v|s_0), p(v|s_1)) \leq \varepsilon$.*

Theorem 4.9. *Let $\mathcal{S}$ be a ε-hiding commitment scheme and $\mathcal{S}'$ a scheme that is δ-hiding until the last round. If $(\mathcal{S}, \mathcal{S}')$ is eligible, then the composed scheme $\mathcal{S}'' = \mathcal{S} \star \mathcal{S}'$ is $(\varepsilon + \delta)$-hiding until the last round.*

Proof. Fix a strategy against the hiding-until-the-last-round property of $\mathcal{S}''$. We consider the distribution $p(v, y, v'|s)$ where s is the string that the provers commit to, v the verifier's view after $\overline{\mathsf{com}}_{PQV}$ has been executed, y the opening

information to which Q commits using the scheme $\mathcal{S}'$, and v' the verifier's view immediately before the last round of communication. We need to show that $d(p(v'|s_0), p(v'|s_1)) \leq \varepsilon + \delta$ for any s_0 and s_1.

First, note that $p(v'|v, y, s_b) = p(v'|v, y)$ since v' is produced by P, Q and V acting on y and v only. From any strategy against $\mathcal{S}''$, we can obtain a strategy against $\mathcal{S}'$ by fixing v. Thus, by the hiding property of $\mathcal{S}'$, for any y_0 and y_1, we have $d(p(v'|v, y = y_0), p(v'|v, y = y_1)) \leq \delta$ and it follows by the convexity of the statistical distance in both arguments that

$$p(v'|v, s_0) = \sum_y p(y|v, s_0)p(v'|v, y) \approx_\delta \sum_y p(y|v, s_1)p(v'|v, y) = p(v'|v, s_1)$$

where we use $\approx_\delta$ to indicate that the two distributions have statistical distance at most δ. Since we have $d(p(v|s_0), p(v|s_1)) \leq \varepsilon$ by the hiding property of $\mathcal{S}$, it follows that

$$\begin{aligned} p(v'|s_0) = p(v, v'|s_0) = p(v|s_0)p(v'|v, s_0) &\approx_\delta p(v|s_0)p(v'|v, s_1) \\ &\approx_\varepsilon p(v|s_1)p(v'|v, s_1) \\ &= p(v, v'|s_1) \\ &= p(v'|s_1) \end{aligned}$$

where the first and last equality hold because v' contains v since v' is the view of V at a later point in time. □

4.2 The Composition Theorems

4.2.1 A Composition Theorem for (Strongly) Binding Schemes

Before stating and proving the composition theorem, we need to single out one more relevant parameter.

Definition 4.10. *Let $(\mathcal{S}, \mathcal{S}')$ be an eligible pair, which in particular means that V's action in the opening phase of $\mathcal{S}$ is determined by a function* Extr. *We define $k(\mathcal{S}) := \max_{c,s} |\{y \,|\, \mathsf{Extr}(y, c) = s\}|$.*

I.e., $k(\mathcal{S})$ counts the number of y's that are consistent with a given string s in the worst case. Note that $k(\mathcal{CHSH}^q) = 1$: for every $a, x, s \in \mathbb{F}_q$ there is at most one $y \in \mathbb{F}_q$ such that $x + y = a \cdot s$.

In the following composition theorems, we take it as understood that the assumed respective binding properties of $\mathcal{S}$ and $\mathcal{S}'$ hold with respect to a well-defined respective classes of possible attacks. We start with the composition theorem for the fairly-binding property, which is easier to prove than the one for the fairly-weak-binding property.

Theorem 4.11. *Let $(\mathcal{S}, \mathcal{S}')$ be an eligible pair of 2-prover commitment schemes, and assume that $\mathcal{S}$ and $\mathcal{S}'$ are respectively ε-fairly-binding and δ-fairly-binding. Then, their composition $\mathcal{S}'' = \mathcal{S} \star \mathcal{S}'$ is $(\varepsilon + k(\mathcal{S}) \cdot \delta)$-fairly-binding.*

Proof. We first consider the case $k(\mathcal{S}) = 1$. We fix an attack $(\overline{\mathsf{com}}_P, \overline{\mathsf{open}}''_{PQ})$ against $\mathcal{S}''$. Without loss of generality, the attack is deterministic, so $\overline{\mathsf{open}}''_{PQ}$ is of the form $\overline{\mathsf{open}}''_{PQ} = \overline{\mathsf{open}}'_{PQ} \circ \mathsf{ptoq}_{PQ} \circ \overline{\mathsf{com}}'_Q$.

Note that $\overline{\mathsf{com}}_P$ is also a commit strategy for $\mathcal{S}$. As such, by the fairly-binding property of $\mathcal{S}$, there exists a function $\hat{s}(c)$, only depending on $\overline{\mathsf{com}}_P$, so that the property specified in Definition 3.2 is satisfied for every opening strategy $\overline{\mathsf{open}}_Q$ for $\mathcal{S}$. We will show that it is also satisfied for the (arbitrary) opening strategy $\overline{\mathsf{open}}''_{PQ}$ for $\mathcal{S}''$, except for a small increase in ε: we will show that $p(\hat{s}(c) \neq s \wedge s = s_\circ) \leq \varepsilon + \delta$ for every fixed target string $s_\circ$. This then proves the claim.

To show this property on $\hat{s}(c)$, we "decompose and reassemble" the attack strategy $(\overline{\mathsf{com}}_P, \overline{\mathsf{open}}'_{PQ} \circ \mathsf{ptoq}_{PQ} \circ \overline{\mathsf{com}}'_Q)$ for $\mathcal{S}''$ into an attack strategy $(\overline{\mathsf{com}}'_Q, \overline{\mathsf{newopen}}'_{PQ})$ for $\mathcal{S}'$ with $\overline{\mathsf{newopen}}'_{PQ}$ formally defined as

$$\overline{\mathsf{newopen}}'_{PQ}[c](\overline{state}'_Q) := \overline{\mathsf{open}}'_{PQ}\big(\overline{state}_P(c) \| (\overline{state}_P(c), \overline{state}'_Q)\big)$$

where

$$(\overline{state}_P(c) \| c) \leftarrow \big(\overline{\mathsf{com}}_P || \mathsf{com}_V\big)\,.$$

Informally, this means that ahead of time, P and Q *simulate* an execution of $(\overline{\mathsf{com}}_P(\emptyset) || \mathsf{com}_V(\emptyset))$ and take the resulting communication/commitment[5] c as shared randomness, and then $\overline{\mathsf{newopen}}'_{PQ}$ computes $\overline{state}_P$ from c as does $\overline{\mathsf{com}}_P$, and runs $\overline{\mathsf{open}}'_{PQ}$ (see Fig. 4.2).[6] It follows from the fairly-binding property that there is a function $\hat{y}(c')$ of the commitment c' so that $p(\hat{y}(c') \neq y \wedge y = y_\circ(c)) \leq \delta$ for every function $y_\circ(c)$.

The existence of $\hat{y}$ now gives rise to an opening strategy $\overline{\mathsf{open}}_Q$ for $\mathcal{S}$; namely, simulate the commit phase of $\mathcal{S}'$ to obtain the commitment c', and output $\hat{y}(c')$. By Definition 3.2, for $\tilde{s} := \mathsf{Extr}(\hat{y}(c'), c)$ and every $s_\circ$, $p(\hat{s}(c) \neq \tilde{s} \wedge \tilde{s} = s_\circ) \leq \varepsilon$.

We are now ready to put things together. Fix an arbitrary target string $s_\circ$. For any c we let $y_\circ(c)$ be the unique string such that $\mathsf{Extr}(y_\circ(c), c) = s_\circ$ (and some default string if no such string exists); recall, we assume for the moment that $k(\mathcal{S}) = 1$. Omitting the arguments in $\hat{s}(c), \hat{y}(c')$ and $y_\circ(c)$, it follows that

$$\begin{aligned}
p(\hat{s} \neq s \wedge s = s_\circ) &\leq p(\hat{s} \neq s \wedge s = s_\circ \wedge s = \tilde{s}) + p(s = s_\circ \wedge s \neq \tilde{s}) \\
&\leq p(\hat{s} \neq \tilde{s} \wedge \tilde{s} = s_\circ) + p\big(\mathsf{Extr}(y, c) \neq \mathsf{Extr}(\hat{y}, c) \wedge \mathsf{Extr}(y, c) = s_\circ\big) \\
&\leq p(\hat{s} \neq \tilde{s} \wedge \tilde{s} = s_\circ) + p(y \neq \hat{y} \wedge y = y_\circ) \\
&\leq \varepsilon + \delta.
\end{aligned}$$

[5] Recall that by convention (Remark 2.10), the commitment c equals the communication between V and, here, P.

[6] We are using here that Q is inactive during $\overline{\mathsf{com}}_{PQ}$ and P during $\overline{\mathsf{com}}'_{PQ}$, and thus the two "commute".

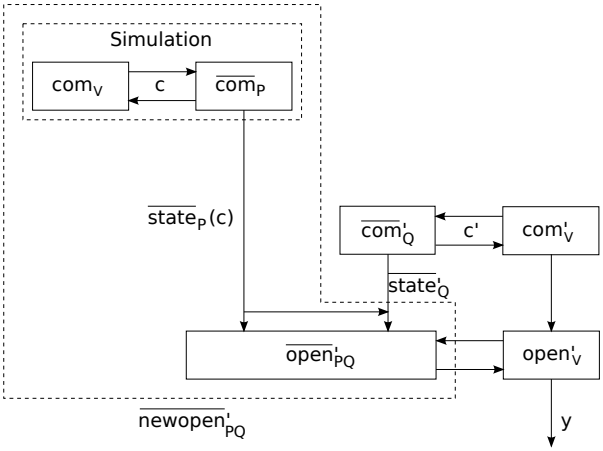

Figure 4.2: Constructing the opening strategy $\overline{\mathsf{newopen}}'_{PQ}$ against $\mathcal{S}'$.

Thus, $\hat{s}$ is as required.

For the general case where $k(\mathcal{S}) > 1$, we can reason similarly, except that we then list the $k \leq k(\mathcal{S})$ possibilities $y_\circ^1(c), \ldots, y_\circ^k(c)$ for $y_\circ(c)$, and conclude that $p(s \neq \tilde{s} \wedge s = s_\circ) \leq \sum_i p\big(y \neq \hat{y} \wedge y = y_\circ^i\big) \leq k(\mathcal{S}) \cdot \delta$, which then results in the claimed bound. □

Remark 4.12. *Putting things together, we can now conclude the security (i.e., the binding property) of the Lunghi et al. multi-round commitment scheme. Corollary 3.22 ensures the fairly-binding property of $\mathcal{CHSH}^q$, i.e., the Crépeau et al. scheme as a string-commitment scheme, with parameter $2\sqrt{q^{-1}}$. The composition theorem (Theorem 4.11) then guarantees the fairly-binding property of the m-fold composition as a string-commitment scheme, with parameter $(m+1) \cdot 2\sqrt{q^{-1}}$. Finally, Proposition 3.6 implies that the m-fold composition of $\mathcal{CHSH}^q$ with itself is a ε_m-binding bit-commitment scheme with error parameter $\varepsilon_m = (m+1) \cdot 4\sqrt{q^{-1}}$ as claimed in the introduction, or, more generally, and by taking Remark 3.7 into account, a $(m+1) \cdot 2^{k+1}\sqrt{q^{-1}}$-binding k-bit-string-commitment scheme.*

4.2.2 Composition Theorem for Weakly Binding Schemes

We now show the composition theorem for the weak version of the binding property. Since this notion makes sense also against quantum attacks, we emphasize the restriction to classical attacks — extending the theorem to quantum attacks is an open problem. See Chapter 5 for some partial progress in this direction.

Theorem 4.13. *Let $(\mathcal{S}, \mathcal{S}')$ be an eligible pair of 2-prover commitment schemes, and assume that $\mathcal{S}$ and $\mathcal{S}'$ are respectively ε-fairly-weak-binding and δ-fairly-weak-binding against classical attacks. Then, their composition $\mathcal{S}'' = \mathcal{S} \star \mathcal{S}'$ is a $(\varepsilon + k(\mathcal{S}) \cdot \delta)$-fairly-weak-binding 2-prover commitment scheme against classical attacks.*

Proof. We first consider the case $k(\mathcal{S}) = 1$. We fix an arbitrary deterministic attack $(\overline{\mathsf{com}}_P, \overline{\mathsf{open}}''_{PQ})$ against $\mathcal{S}''$, where $\overline{\mathsf{open}}''_{PQ}$ is of the form $\overline{\mathsf{open}}''_{PQ} = \overline{\mathsf{open}}'_{PQ} \circ \mathsf{ptoq}_{PQ} \circ \overline{\mathsf{com}}'_Q$. Let a be V's randomness in com_V. Then, c is a function $c(a)$ of a, and the distribution $p(a, y)$ is well defined. Since $\overline{\mathsf{com}}_P$ is also an attack strategy against $\mathcal{S}$, there exists a distribution $p(\hat{s})$ (only depending on $\overline{\mathsf{com}}_P$) such that Definition 3.9 is satisfied for every opening strategy $\overline{\mathsf{open}}_Q$ for $\mathcal{S}$.

Similar to the proof of Theorem 4.11, we decompose and reassemble the attack strategy $(\overline{\mathsf{com}}_P, \overline{\mathsf{open}}'_{PQ} \circ \mathsf{ptoq}_{PQ} \circ \overline{\mathsf{com}}'_Q)$ for $\mathcal{S}''$ into an attack strategy $(\overline{\mathsf{com}}'_Q, \overline{\mathsf{newopen}}'_{PQ})$ for $\mathcal{S}'$. Concretely, for every fixed choice of a, we obtain a deterministic opening strategy $\overline{\mathsf{newopen}}'_{PQ,a}$ given by

$$\overline{\mathsf{newopen}}'_{PQ,a}(\overline{state}'_Q) := \overline{\mathsf{open}}'_{PQ}\big(\overline{state}_P(c(a)) \| (\overline{state}_P(c(a)), \overline{state}'_Q)\big)\,,$$

and the distribution of the verifier's output y when the provers use $\overline{\mathsf{newopen}}'_{PQ,a}$ is $p(y|a)$. It follows from the fairly-weak-binding property of $\mathcal{S}'$ that there exists a distribution $p(\hat{y})$, only depending on $\overline{\mathsf{com}}'_Q$, so that for every choice of a there exists a consistent joint distribution $p(\hat{y}, y|a)$ so that $p(\hat{y} \neq y \wedge y = y_\circ|a) \leq \delta$ for every fixed target string $y_\circ$. Note that here, consistency in particular means that $p(\hat{y}|a) = p(\hat{y})$. This joint conditional distribution $p(\hat{y}, y|a)$ together with the distribution $p(a)$ of a then naturally defines the distribution $p(a, \hat{y}, y)$, which is consistent with $p(a, y)$ considered above.

The existence of $p(\hat{y})$ now gives rise to an opening strategy $\overline{\mathsf{open}}_Q$ for $\mathcal{S}$; namely, sample $\hat{y}$ according to $p(\hat{y})$ and output $\hat{y}$. Note that the joint distribution of a and $\hat{y}$ in this "experiment" is given by

$$p(a) \cdot p(\hat{y}) = p(a) \cdot p(\hat{y}|a) = p(a, \hat{y})\,,$$

i.e., is consistent with the distribution $p(a, \hat{y}, y)$ above. By Definition 3.9, we know there exists a joint distribution $p(\hat{s}, \tilde{s})$, consistent with $p(\hat{s})$ fixed above and with $p(\tilde{s})$ determined by $\tilde{s} := \mathsf{Extr}(\hat{y}, c(a))$, and such that $p(\hat{s} \neq \tilde{s} \wedge \tilde{s} = s_\circ) \leq \varepsilon$ for every $s_\circ$. We can now "glue together" $p(\hat{s}, \tilde{s})$ and $p(c, \hat{y}, y, \tilde{s})$, i.e., find a joint distribution that is consistent with both, by setting

$$p(a, \hat{y}, y, \tilde{s}, \hat{s}) := p(a, \hat{y}, y, \tilde{s}) \cdot p(\hat{s}|\tilde{s})\,.$$

We now fix an arbitrary target string $s_\circ$. Furthermore, for any a we let $y_\circ(a)$ be the unique string such that $\mathsf{Extr}(y_\circ(a), c(a)) = s_\circ$ (and to some default string if no such string exists); recall, we assume for the moment that $k(\mathcal{S}) = 1$. With

respect to the above joint distribution, it then holds that

$$\begin{aligned} p(\hat{s} \neq s \wedge s = s_\circ) &= p(\hat{s} \neq s \wedge s = s_\circ \wedge s = \tilde{s}) + p(s = s_\circ \wedge s = s_\circ \wedge s \neq \tilde{s}) \\ &\leq p(\hat{s} \neq \tilde{s} \wedge s = s_\circ \wedge s = \tilde{s}) + p(s \neq \tilde{s} \wedge s = s_\circ) \\ &\leq p(\hat{s} \neq \tilde{s} \wedge \tilde{s} = s_\circ) \\ &\quad + p\big(\mathsf{Extr}(y, c(a)) \neq \mathsf{Extr}(\hat{y}, c(a)) \wedge \mathsf{Extr}(y, c(a)) = s_\circ\big) \\ &\leq p(\hat{s} \neq \tilde{s} \wedge \tilde{s} = s_\circ) + p(y \neq \hat{y} \wedge y = y_\circ(a)) \\ &\leq p(\hat{s} \neq \tilde{s} \wedge \tilde{s} = s_\circ) + \textstyle\sum_a p(a) \cdot p(y \neq \hat{y} \wedge y = y_\circ(a) | a) \\ &\leq \varepsilon + \delta \,. \end{aligned}$$

Thus, the distribution $p(\hat{s}, s)$ is as required.

For the case where $k(\mathcal{S}) > 1$, we can reason similarly, except that we then list the $k \leq k(\mathcal{S})$ possibilities $y_\circ^1(a), \ldots, y_\circ^a(a)$ for $y_\circ(a)$, and conclude that $p(s \neq \tilde{s} \wedge s = s_\circ) \leq \sum_i p\big(y \neq \hat{y} \wedge y = y_\circ^i(a)\big) \leq k(\mathcal{S}) \cdot \delta$, which then results in the claimed bound. □

Remark 4.14. *Analogously to Remark 4.12, we can conclude from Corollary 3.23 and Theorem 4.13 that $\mathcal{CHSH}^q$ is $(m{+}1)\cdot\sqrt{2q^{-1}}$-fairly-weak-binding. It follows from Proposition 3.11 that $\mathcal{CHSH}^q$ is a $(m+1) \cdot 2^{3/2}\sqrt{q^{-1}}$-weak-binding bit-commitment scheme. More generally, we can conclude that for any $k < n$, it is a $(m+1) \cdot 2^{k+1/2}\sqrt{q^{-1}}$-weak-binding k-bit string-commitment scheme. Below, we show how to avoid the factor 2 introduced by invoking Proposition 3.11.*

4.3 Variations

In this section, we show two variants of the composition theorems. The first one says that if we compose a weak-binding with a fairly-weak-binding scheme, we obtain a weak-binding scheme. This allows us to slightly improve the parameter in Remark 4.14. The proof crucially relies on the fact that, in the weak definition, there is some freedom in "gluing together" the distributions $p(s)$ and $p(\hat{s})$. The second variant says that composing two binding (or weak-binding) schemes yields a binding (or weak-binding, respectively) scheme.

We start by proving the following two properties for fairly-weak-binding commitment schemes. The first property shows that one may assume the joint distribution $p(\hat{s}, s)$ to be such that s and $\hat{s}$ are independent conditioned on $s \neq \hat{s}$.

Lemma 4.15. *Let $\mathcal{S}$ be a ε-fairly-weak-binding commitment scheme. Then, for any $\overline{\mathsf{com}}_{PQ}$ and $\overline{\mathsf{open}}_{PQ}$ there exists a joint distribution $p(\hat{s}, s)$ as required by Definition 3.9, but with the additional property that*

$$p(\hat{s}, s | s \neq \hat{s}) = p(\hat{s} | s \neq \hat{s}) \cdot p(s | s \neq \hat{s}) \,.$$

Proof. Since the scheme is ε-fairly-weak-binding, it follows that there exists a consistent joint distribution $p(\hat{s}, s)$ such that $p(s \neq \hat{s} \wedge s = s_\circ) \leq \varepsilon$ for every $s_\circ$. Because of this, we have

$$\begin{aligned} p(s = s_\circ) &= p(s{=}s_\circ \wedge \hat{s}{=}s_\circ) + p(s{=}s_\circ \wedge \hat{s}{\neq}s_\circ) \\ &= p(s{=}s_\circ \wedge \hat{s}{=}s_\circ) + p(s{\neq}\hat{s} \wedge s{=}s_\circ) \\ &\leq p(\hat{s} = s_\circ) + \varepsilon. \end{aligned}$$

We apply Lemma 2.2 to the marginal distributions $p(\hat{s})$ and $p(s)$. The resulting joint distribution $\tilde{p}(\hat{s}, s)$ satisfies $\tilde{p}(\hat{s} = s_\circ \wedge s = s_\circ | s = \hat{s}) = \min\{p(s = s_\circ), p(\hat{s} = s_\circ)\}$ and $\tilde{p}(\hat{s}, s | s \neq \hat{s}) = \tilde{p}(\hat{s} | s \neq \hat{s}) \cdot \tilde{p}(s | s \neq \hat{s})$. It remains to show that $\tilde{p}(s \neq \hat{s} \wedge s = s_\circ) \leq \varepsilon$ for all $s_\circ$. Indeed, we have

$$\begin{aligned} \tilde{p}(s \neq \hat{s} \wedge s = s_\circ) &= \tilde{p}(s = s_\circ) - \tilde{p}(s = \hat{s} \wedge s = s_\circ) \\ &= \tilde{p}(s = s_\circ) - \tilde{p}(\hat{s} = s_\circ \wedge s = s_\circ) \\ &= p(s = s_\circ) - \min\{p(\hat{s} = s_\circ), p(s = s_\circ)\} \\ &\leq p(s = s_\circ) - (p(s = s_\circ) - \varepsilon) \\ &= \varepsilon \end{aligned}$$

as claimed. □

The second property shows that the quantification over all *fixed* $s_\circ$ in Definition 3.9 of the fairly-weak-binding property can be relaxed to $s_\circ$ that may depend on $\hat{s}$, but only on $\hat{s}$. Note that we can obviously not allow $s_\circ$ to depend (arbitrarily) on s, since then one could choose $s_\circ = s$.

Proposition 4.16. *Let $\mathcal{S}$ be a ε-fairly-weak-binding commitment scheme. Then*

$$\forall\, \overline{\mathsf{com}}_{PQ}\ \exists\, p(\hat{s})\ \forall\, \overline{\mathsf{open}}_{PQ}\ \exists\, p(\hat{s}, s)\ \forall\, p(s_\circ | \hat{s}) : p(s \neq \hat{s}\ \wedge\ s = s_\circ) \leq \varepsilon\,,$$

where it is understood that $p(\hat{s}, s, s_\circ) := p(\hat{s}, s) \cdot p(s_\circ | \hat{s})$. Thus, the joint distribution $p(\hat{s}, s)$ is such that $p(s \neq \hat{s}\ \wedge\ s = s_\circ) \leq \varepsilon$ holds in particular for any function $s_\circ = f(\hat{s})$ of $\hat{s}$.

Proof. For given $\overline{\mathsf{com}}_{PQ}$ and $\overline{\mathsf{open}}_{PQ}$, let $p(\hat{s}, s)$ be as guaranteed by the fairly-weak-binding property. By Lemma 4.15, we may assume without loss of generality that $p(\hat{s}, s | s \neq \hat{s}) = p(\hat{s} | s \neq \hat{s})\, p(s | s \neq \hat{s})$. Then, by Lemma 2.7, we also

have that $p(s, s_\circ | s \neq \hat{s}) = p(s|s \neq \hat{s})\, p(s_\circ | s \neq \hat{s})$. It follows that

$$\begin{aligned}
p(s \neq \hat{s} \wedge s = s_\circ) &= p(s \neq \hat{s}) \cdot p(s = s_\circ | s \neq \hat{s}) \\
&= p(s \neq \hat{s}) \sum_{s^*_\circ} p(s = s^*_\circ \wedge s_\circ = s^*_\circ | s \neq \hat{s}) \\
&= p(s \neq \hat{s}) \sum_{s^*_\circ} p(s = s^*_\circ | s \neq \hat{s}) \cdot p(s_\circ = s^*_\circ | s \neq \hat{s}) \\
&= \sum_{s^*_\circ} p(s \neq \hat{s} \wedge s = s^*_\circ) \cdot p(s_\circ = s^*_\circ | s \neq \hat{s}) \\
&\leq \varepsilon \cdot \sum_{s^*_\circ} p(s_\circ = s^*_\circ | s \neq \hat{s}) \\
&= \varepsilon
\end{aligned}$$

where the inequality follows from the fact that $p(s \neq \hat{s} \wedge s = s^*_\circ) \leq \varepsilon$ for every *fixed* $s^*_\circ$. □

For the rest of the section, we take it as understood that we only consider classical attacks.

Theorem 4.17. *Let $(\mathcal{S}, \mathcal{S}')$ be an eligible pair of 2-prover commitment schemes, where $\mathcal{S}$ is ε-weak-binding and $\mathcal{S}'$ is δ-fairly-weak-binding, and let D be the domain of $\mathcal{S}$. Then, the composition $S \star \mathcal{S}'$ is a $(\varepsilon + (|D|-1) \cdot k(\mathcal{S}) \cdot \delta)$-weak-binding commitment scheme.*
In particular, if $\mathcal{S}$ is a bit *commitment scheme then $S \star \mathcal{S}'$ is a $(\varepsilon + k(\mathcal{S}) \cdot \delta)$-weak-binding.*

Proof. We follow the proof of Theorem 4.13, up to when it comes to choosing $y_\circ$. Let us first consider the case $m = 1$, i.e., $\mathcal{S}$ is a *bit*-commitment scheme. In that case, and assuming for the moment that $k(\mathcal{S}) = 1$, we let $y_\circ$ be the unique string that satisfies $\mathsf{Extr}(y_\circ, c) = s_\circ$, but where now $s_\circ := 1 - \tilde{s}$. We emphasize that for a fixed c, this choice of $y_\circ$ is *not* fixed anymore (in contrast to the choice in the proof of Theorem 4.13); namely, it is a function of $\tilde{s} = \mathsf{Extr}(\hat{y}, c)$, which in turn is a function of $\hat{y}$. Therefore, by Proposition 4.16, it still holds that $p(y \neq \hat{y} \wedge y = y_\circ | a) \leq \delta$, and we can conclude that

$$\begin{aligned}
p(\hat{s} \neq s \wedge s \neq \bot) &\leq p(\hat{s} \neq s \wedge s \neq \bot \wedge s = \tilde{s}) + p(s \neq \tilde{s} \wedge s \neq \bot) \\
&= p(\hat{s} \neq \tilde{s} \wedge s \neq \bot \wedge s = \tilde{s}) + p(s \neq \tilde{s} \wedge s = 1 - \tilde{s}) \\
&\leq p(\hat{s} \neq \tilde{s} \wedge \tilde{s} \neq \bot) + p(y \neq \hat{y} \wedge y = y_\circ) \\
&\leq p(\hat{s} \neq \tilde{s} \wedge \tilde{s} \neq \bot) + \textstyle\sum_a p(a)\, p(y \neq \hat{y} \wedge y = y_\circ | a) \\
&\leq \varepsilon + \textstyle\sum_a p(a)\, \delta \\
&= \varepsilon + \delta\,.
\end{aligned}$$

In the case that $k(\mathcal{S}) > 1$, we instead randomly select one of the at most $k(\mathcal{S})$ strings $y_\circ$ that satisfy $\mathsf{Extr}(y_\circ, c) = s_\circ = 1 - \tilde{s}$. Then, conditioned on a, $y_\circ$ is still independent of y given $\hat{y}$, so that Proposition 4.16 still applies, and we can argue as above, except that we get a factor $k(\mathcal{S})$ blow-up from $p(s \neq \tilde{s} \wedge s = 1 - \tilde{s}) \leq k(\mathcal{S}) \cdot p(y \neq \hat{y} \wedge y = y_\circ)$.

Finally, for the case $m > 1$, we first pick a random $s_\circ \in D \setminus \{\tilde{s}\}$, and then choose $y_\circ$ such that $\mathsf{Extr}(y_\circ, c) = s_\circ$, uniquely or at random, depending of $k(\mathcal{S})$. Conditioned on a, $y_\circ$ is still independent of y given $\hat{y}$, and therefore Proposition 4.16 still applies, but now we get an additional factor $(|D| - 1)$ blow-up from $p(s \neq \tilde{s} \wedge s \neq \perp) \leq (|D| - 1)\, p(s \neq \tilde{s} \wedge s = s_\circ)$. □

Remark 4.18. *Theorem 4.17 allows us to slightly improve the bound we obtain in Remark 4.14 on the Lunghi et al. multi-round commitment scheme. By Theorem 4.13, we can compose m instances of $\mathcal{CHSH}^n$ to obtain a $m \cdot 2^{-(n-1)/2}$-fairly-weak-binding string-commitment scheme. Then, we can compose the Crépeau et al.* bit *commitment scheme (i.e., the bit-commitment version of $\mathcal{CHSH}^n$), which is $2^{-(n-1)}$-weak-binding, with this fairly-weak-binding string-commitment scheme; by Theorem 4.17, this composition, which is the Lunghi et al. multi-round bit-commitment scheme, is $\left(m \cdot 2^{-(n-1)/2} + 2^{-(n-1)}\right)$-weak-binding.*

Finally, for completeness, we point out that the composition theorem also applies to two ordinary binding or weak-binding commitment schemes.

Theorem 4.19. *Let $(\mathcal{S}, \mathcal{S}')$ be an eligible pair of 2-prover commitment schemes, where $\mathcal{S}$ is ε-binding and $\mathcal{S}'$ is δ-binding. Then, the composition $S \star \mathcal{S}'$ is $(\varepsilon + \delta)$-binding. The same holds for the weak-binding property.*

Proof. The proof is almost the same as in Theorem 4.11 or Theorem 4.13, respectively, except that now there are no $s_\circ$ and $y_\circ$, and in the end we can simply conclude that

$$\begin{aligned} p(s \neq \hat{s} \wedge s \neq \perp) &\leq p(s \neq \hat{s} \wedge s \neq \perp \wedge s = \tilde{s}) + p(s \neq \tilde{s} \wedge s \neq \perp) \\ &\leq p(\tilde{s} \neq \hat{s} \wedge \tilde{s} \neq \perp) + p(y \neq \hat{y} \wedge y \neq \perp) \\ &\leq \varepsilon + \delta\,, \end{aligned}$$

where the second inequality holds since $y = \perp$ implies that $s = \mathsf{Extr}(y, c) = \perp$. □

4.4 Tightness

We now show that our composition result is nearly tight for $\mathcal{CHSH}^q$. Let $\mathcal{CHSH}^q_m$ be the m-fold composition of $\mathcal{CHSH}^q$ with itself, as defined in Remark 4.5. We show that if $q = p^{2k}$ for some prime p, this composed scheme can be ε-weak-binding as a bit-commitment scheme only if $\varepsilon \gtrsim \frac{1}{4} m \sqrt{q^{-1}}$. A

slightly weaker result was proved in [BC16], which shows that $\varepsilon \gtrsim \frac{1}{6}m2^{-n/2}$ for $q = 2^n$ with n even.[7] Furthermore, we show that, as a string-commitment scheme, $\mathcal{CHSH}^n_m$ can be ε-fairly-weak-binding only if $\varepsilon \gtrsim \frac{1}{2}m\sqrt{q^{-1}}$ (for $q = p^{2k}$).

Lemma 4.20. *Consider functions $X_q, Y_q : \mathbb{F}_q \times R_q \to \mathbb{F}_q$ where R_q is some finite set. Let*

$$\lambda_q = \max_{X_q, Y_q} \; p(X_q(a,r) + Y_q(s,r) = a \cdot s) \tag{4.1}$$

where a, s and r are selected uniformly at random. It holds that:

1. *There are X_q and Y_q such that $p(X_q(a,r) + Y_q(s,r) = a \cdot s) = \lambda_q$ for* all *$a, s \in \mathbb{F}_q$.*
2. *If $q = p^{2k}$ for some prime p, we have $\lambda_q = \Omega\big(\sqrt{q^{-1}}\big)$. Otherwise, we have $\lambda_q = \Omega\big(q^{-2/3}\big)$.*

Proof. Fix X'_q and Y'_q that achieve the maximum in Equation (4.1). We show that there also are functions X_q and Y_q such that for *any* a and s, $p(X_q(a,r) + Y_q(s,r) = a \cdot s) = \lambda_q$: Without loss of generality, X'_q and Y'_q depend only on a and s, not on r. Intuitively, X_q and Y_q do the following: they randomize their inputs a and s by adding uniformly random elements $r_a, r_s \in \mathbb{F}_q$, then apply X'_q and Y'_q, and finally remove the random terms again from the output. Formally, we let

$$\begin{aligned} X_q(a,(r_a,r_s)) &= X'_n(a + r_a) - ar_s - r_a r_s \\ Y_q(a,(r_a,r_s)) &= Y'_n(s + r_s) - r_a s \end{aligned}$$

For r_a and r_s uniformly random, we have $p(X'_q(a+r_a)+Y'_q(s+r_s) = as+ar_s+r_ar_s+sr_a) = \lambda_q$. Thus, it is easy to see that $p(X_q(a,(r_a,r_s))+Y_q(s,(r_a,r_s)) = as) = \lambda_q$.

The functions X_q and Y_q in Equation (4.1) describe classical strategies for the CHSH_q game and λ_q is the maximal winning probability that classical players can achieve in this game. As shown in [BS15], it holds that $\lambda_q = \Omega\big(\sqrt{q^{-1}}\big)$ for $q = p^{2k}$, and $\lambda_q = \Omega\big(q^{-2/3}\big)$ otherwise. □

The following lemma can be seen as a generalization of Theorem 3.12 to string-commitment schemes. Intuitively, it bounds the winning probability of the provers in the following game: First, they have to produce a commitment. Then, they receive a uniformly random string $s_\circ$ and, in order to win, they have to open the commitment to $s_\circ$. The winning probability in this game is at most $\varepsilon + 2^{-n}$, when the scheme is an ε-fairly-weak-binding n-bit string-commitment scheme.

[7] The paper states $\varepsilon \gtrsim \frac{1}{3}m2^{-n/2}$, but their binding definition is $p_0 + p_1 \leq 1 + \varepsilon$; to convert their bound to our definition (equivalent to $p_0 + p_1 \leq 1 + 2\varepsilon$), it must be multiplied by 1/2.

Lemma 4.21. *Let $\mathcal{S}$ be an ε-fairly-weak-binding commitment scheme with domain D. Fix a possible commit strategy $\overline{\mathsf{com}}_{PQ}$ for $\mathcal{S}$ and, for each $s_\circ \in D$, a possible opening strategy $\overline{\mathsf{open}}_{PQ}(s_\circ)$. Let $p(s|s_\circ)$ be the output distribution of $\mathcal{S}$ if the provers use $\overline{\mathsf{com}}_{PQ}$ and $\overline{\mathsf{open}}_{PQ}(s_\circ)$. Let $p(s_\circ)$ be the uniform distribution over D. Then, $p(s = s_\circ) = \sum_{s_\circ \in D} p(s_\circ)p(s = s_\circ|s_\circ) \leq \varepsilon + |D|^{-1}$.*

Proof. Let $p(\hat{s})$ be a distribution that satisfies Equation (3.4) for the commit strategy $\overline{\mathsf{com}}_{PQ}$. Now consider any consistent joint distribution $p(s, \hat{s}|s_\circ)$. Here, consistency also means that $p(\hat{s}|s_\circ) = p(\hat{s})$. Thus, for a uniformly random $s_\circ$, $p(\hat{s} = s_\circ) = |D|^{-1}$. By the ε-fairly-weak-binding property of $\mathcal{S}$, we have

$$\varepsilon \geq p(s \neq \hat{s} \wedge s = s_\circ) \geq p(s = s_\circ) - p(\hat{s} = s_\circ) = p(s = s_\circ) - |D|^{-1}$$

and thus our claim follows. □

With the help of the lemma above, is easy to see that λ_q limits the binding parameter of the one-round scheme $\mathcal{CHSH}^q$: If P sends $X_n(a, r)$ and Q sends $Y_n(s_\circ, r)$ for uniformly random r, then we have $p(s = s_\circ|a \neq 0) = \lambda_q$, and thus $p(s = s_\circ) \geq \lambda_q - q^{-1}$ for every $s_\circ$. Thus, by Lemma 4.21, $\mathcal{CHSH}^q$ can be ε-fairly-weak-binding only if $\varepsilon \geq \lambda_q - 2q^{-1}$. We now show that this bound scales approximately linearly with the number of rounds.

Theorem 4.22. *Let λ_q as in Lemma 4.20. For odd m, the $\mathcal{CHSH}^q_m$ commitment scheme can be ε-fairly-weak-binding as a string-commitment scheme only if*

$$\varepsilon \geq \frac{(m+1)\lambda_q}{2} - \frac{\left(m^2-1\right)\lambda_q^2}{8} - (m+1)q^{-1}.$$

If $m = o\left(\lambda_q^{-1}\right)$, it holds that $\varepsilon \geq \Omega(m\lambda_q)$. If, furthermore, $q = p^{2k}$, we have $\varepsilon \geq \Omega\left(m\sqrt{q^{-1}}\right)$; otherwise, $\varepsilon \geq \Omega\left(mq^{-2/3}\right)$.

Proof. Let $X_q(a, r)$ and $Y_q(b, r)$ be functions as in Lemma 4.20. We define a commit strategy $\overline{\mathsf{com}}_{PQ}$ and an opening strategy $\overline{\mathsf{open}}_{PQ}(s_\circ)$ for every $s_\circ$ which aims to open to $s_\circ$.

We assume that the provers have m uniformly random $r_i \in \mathbb{F}_q$ and $(m+1)/2$ uniformly random inputs r'_i, i odd, for X_q and Y_q as shared randomness. We write $c_i = (a_i, x_i)$ for the communication between the verifier and the active prover in round i, where the x_i are specified below. The dishonest provers exchange their communications as fast as possible, so in round $i + 2$, the active prover knows $c_1, \ldots, c_i$. Let $y_0 = s_\circ$ and for $i > 0$, let y_i such that $\mathsf{Extr}(y_i, c_i) = y_{i-1}$. Such a y_i exists and is unique if $a_i \neq 0$. We only specify our strategy for the case where the verifier's messages a_i are all non-zero and assume that the provers fail to open to $s_\circ$ otherwise. One can compute y_i from $c_1, \ldots, c_i$, so in round $i + 2$, the active prover can compute y_i.

If in any round i, the commitment is $(a_i, r_i + a_i \cdot y_{i-1})$, the provers can open to $s_\circ$ simply by following the honest strategy for $\mathcal{CHSH}^q_m$ from that round on.

The strategy described below is such that the provers have $(m+1)/2$ chances to bring about this situation with probability λ_q.

- Round 1 (commit): P produces a "fake commitment" $x_1 = X_q(a_1, r'_1)$.
- Round i, i even: Q computes $y'_{i-1} = Y_q(y_{i-2}, r'_{i-1})$, hoping that $x_{i-1} + y'_{i-1} = a_{i-1} \cdot y_{i-2}$, i.e., $y'_{i-1} = y_{i-1}$. He honestly commits to y'_{i-1} by computing $x_i = a_i \cdot y'_{i-1} + r_i$.
- Round $i+1$, i even: P checks if $y_{i-1} = y'_{i-1}$. If yes, both provers proceed honestly from this round on, i.e., they follow the honest strategy for $\mathcal{CHSH}^q_m$ in all subsequent rounds.[8] If not, P again produces a "fake commitment" $x_{i+1} = X_q(a_{i+1}, r'_{i+1})$.
- Round $m+1$: Q sends $y'_m = Y_q(y_{m-1}, r'_m)$ to V.

By definition, it holds that $y'_{i-1} = y_{i-1}$ if and only if $X_q(a_{i-1}, r'_{i-1}) + Y_q(y_{i-2}, r'_{i-1}) = a_{i-1} \cdot y_{i-2}$, which happens with probability λ_q. In this case, we have $c_i = (a_i, r_i + a_i \cdot y_{i-1})$, so the provers can indeed open to $s_\circ$ by proceeding honestly (ignoring completeness errors for now).

By definition of X_q, Y_q, and λ_q, if the provers use the strategy $\overline{\mathsf{open}}_{PQ}(s_\circ)$, then for

$$\lambda = 1-(1-\lambda_q)^{(m+1)/2} \geq \frac{(m+1)\lambda_q}{2} - \binom{(m+1)/2}{2}\lambda_q^2 = \frac{(m+1)\lambda_q}{2} - \frac{\big(m^2-1\big)\lambda_q^2}{8}$$

we have $p(s = s_\circ | a_1, \ldots, a_m \neq 0) = \lambda$. Thus, $p(s = s_\circ) \geq \lambda - mq^{-1}$ for all $s_\circ$. Applying Lemma 4.21, we conclude that the scheme can be ε-fairly-weak-binding only if

$$\varepsilon \geq \lambda - (m+1)q^{-1} \geq \frac{(m+1)\lambda_q}{2} - \frac{\big(m^2-1\big)\lambda_q^2}{8} - (m+1)q^{-1}$$

which is in $\Omega(m\lambda_q)$ if $m = o\big(\lambda_q^{-1}\big)$. Finally, we have $\Omega(m\lambda_q) = \Omega\big(m\sqrt{q^{-1}}\big)$ if $q = p^{2k}$ and $\Omega(m\lambda_q) = \Omega\big(mq^{-2/3}\big)$ otherwise, by claim 2 of Lemma 4.20. □

From the analysis in the above proof, we can also derive a version of the theorem for the bit-commitment scheme described in Proposition 3.11.

Corollary 4.23. *For even m, the commitment scheme $\mathcal{CHSH}^q_m$ can be ε-binding as a bit-commitment scheme only if*

$$\varepsilon \geq \frac{m\lambda_q}{4} - \frac{(m^2-2m)\lambda_q^2}{16} - (m+1)q^{-1}.$$

If $m = o\big(\lambda_q^{-1}\big)$, it holds that $\varepsilon \geq \Omega(m\lambda_q)$. If $q = p^{2k}$, we have $\varepsilon \geq \Omega\big(m\sqrt{q^{-1}}\big)$ and if it is odd, $\varepsilon \geq \Omega\big(mq^{-2/3}\big)$.

[8] Q can compute y_{i-1} in round $i+2$ and thus he too knows whether the provers should proceed honestly or not.

Proof. Let $\overline{\mathsf{com}}_P = \mathsf{com}_P(0)$, i.e., P produces an honest commitment to 0. Let $\overline{\mathsf{open}}_{PQ}(0) = \mathsf{open}_{PQ}$, i.e., the honest opening strategy. Since the provers play honestly, they are successful with probability at least $1 - (m+1)q^{-1}$.

For $\overline{\mathsf{open}}_{PQ}(1)$, let $s_\circ$ such that $\mathsf{Extr}(s_\circ, c_1) = 1$. The provers then use the strategy in the proof of Theorem 4.22 to produce a fake commitment c_1 and open it to $s_\circ$. Then, we have

$$p(b = 1 | a_1, \ldots, a_m \neq 0) \geq \frac{m\lambda_q}{2} - \frac{(m^2 - 2m)\lambda_q^2}{8} - q^{-1}$$

and thus,

$$p(b = 1) \geq \frac{m\lambda_q}{2} - \frac{(m^2 - 2m)\lambda_q^2}{8} - (m+1)q^{-1}.$$

It follows that

$$p(b = 0) + p(b = 1) \geq 1 + \frac{m\lambda_q}{2} - \frac{(m^2 - 2m)\lambda_q^2}{8} - (m+1)q^{-1}$$

and, by Theorem 3.12, the scheme can be ε-weak-binding only if

$$\varepsilon \geq \frac{m\lambda_q}{4} - \frac{(m^2 - 2m)\lambda_q^2}{16} - (m+1)q^{-1}.$$

□

Chapter 5

Towards Quantum Safety

5.1 Introduction

In Chapter 4, we proved composition theorems for two-prover commitment schemes. Those theorems crucially rely on the assumption that dishonest provers can *only* use classical shared randomness, and not entangled quantum states: The definition of our (strong) binding property does not apply if the provers use quantum entanglement instead of classical randomness. While the weak binding property is well-defined for adversaries with quantum capabilities, our proof of the composition theorem for this binding property, i.e., Theorem 4.13, still requires that we can assume without loss of generality that the adversaries' strategy is deterministic. This is not true if we consider adversaries with quantum capabilities.

In this chapter, we take some steps towards arguing that the Lunghi *et al.* scheme is binding for provers with quantum capabilities.

In Section 5.4, we show that $\mathcal{CHSH}^q$ satisfies the fairly-weak-binding definition as a string-commitment scheme even when the adversaries can share entangled quantum states. Our intuitive argument in Chapter 4 thus suggests that $\mathcal{CHSH}^q_m$ also satisfies the binding property with parameter linear in m for such adversaries. However, since our composition theorem only applies to classical provers, this intuition remains without a rigorous proof.

Approaching the problem from another direction, we introduce an analogue of the strong binding definition for the quantum case, and prove a composition theorem using this definition which applies to quantum provers. However, we currently do not know if $\mathcal{CHSH}^q$ (or any other scheme) satisfies this stronger definition. Thus, the question whether there exists a multi-round scheme binding for quantum adversaries remains open.

5.2 Quantum Information Theory

We start with a very brief introduction to quantum information theory where we fix our notation. We refer readers who are not familiar with the subject to introductory textbooks such as [NC00] or [Wil13]. Quantum information theory is based on quantum mechanics, but takes a somewhat different point of view. While quantum mechanics is about the evolution of quantum systems over time, quantum information theory views them as static carriers of information which only change when acted upon by an experimenter.

We begin with defining the required mathematical concepts and then introduce quantum states, measurements, and entanglement. Many properties that are taken for granted in classical information do not apply to quantum information: in particular, it is generally not possible to extract information from a quantum state without changing it, or to make a perfect copy of a quantum state. However, we also show how to express classical information and computation in the formalism of quantum information theory. Thus, quantum information can be viewed as an *extension* of classical information.

5.2.1 Definitions

In this section, we recall the mathematical concepts that are used to describe quantum information. We let $\mathcal{H}$ be a *finite-dimensional complex Hilbert space*, i.e., a complex vector space with an inner product $\langle\cdot|\cdot\rangle$ that is conjugate-symmetric and linear in the second argument. We write vectors in $\mathcal{H}$ using the *bra-ket notation* introduced by Paul Dirac [Dir39]: A vector in $\mathcal{H}$ is written as a *ket-vector* $|\phi\rangle$. Every ket-vector $|\phi\rangle$ has a corresponding *bra-vector* $\langle\phi|$ in the dual space $\mathcal{H}^*$:

$$\langle\phi| : \mathcal{H} \to \mathbb{C}, \quad |\psi\rangle \mapsto \langle\phi|\psi\rangle .$$

Thus, a bra- and a ket-vector "fit together" notationally to form the inner product: $\langle\phi|\,|\psi\rangle = \langle\phi|\psi\rangle$.

Definition 5.1. *A vector $|\phi\rangle \in \mathcal{H}$ is called a* state vector *if $\| \, |\phi\rangle \, \| := \sqrt{\langle\phi|\phi\rangle} = 1$.*

Definition 5.2. *Let $\{|i\rangle\}_{i\in I}$ be a basis of $\mathcal{H}$. It is called an* orthonormal basis *if all vectors have norm 1 and are mutually orthogonal. That is, for all $i, j \in I$:*

$$\langle i|j\rangle \;=\; \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

We write $\mathcal{L}(\mathcal{H})$ to denote the vector space of linear operators on $\mathcal{H}$. For any pair of vectors $|\phi\rangle, |\psi\rangle \in \mathcal{H}$, the *outer product* is defined as the linear operator $|\phi\rangle\langle\psi| : \mathcal{H} \to \mathcal{H}, |\delta\rangle \mapsto |\phi\rangle\langle\psi|\delta\rangle = \langle\psi|\delta\rangle\,|\phi\rangle$. The space $\mathcal{L}(\mathcal{H})$ is spanned by the set of outer products. In fact, if $\{|i\rangle\}_{i\in I}$ is a basis of $\mathcal{H}$, then $\{|i\rangle\langle j|\}_{i,j\in I}$ is a basis of $\mathcal{L}(\mathcal{H})$. We recall the following two linear operators acting on $\mathcal{L}(\mathcal{H})$:

Definition 5.3. *The* conjugate transpose *is the linear operator* $\mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ *that maps* $|\phi\rangle\langle\psi|$ *to* $|\phi\rangle\langle\psi|^\dagger := |\psi\rangle\langle\phi|$ *for all* $|\phi\rangle, |\psi\rangle \in \mathcal{H}$.

Definition 5.4. *The* trace *is the linear operator* $\mathrm{Tr} : \mathcal{L}(\mathcal{H}) \to \mathbb{C}$ *defined by* $\mathrm{Tr}(|\phi\rangle\langle\psi|) = \langle\psi|\phi\rangle$ *for all* $|\phi\rangle, |\psi\rangle \in \mathcal{H}$.

For any $A \in \mathcal{L}(\mathcal{H})$ and any orthonormal basis $\{|i\rangle\}_{i\in I}$, $\sum_{i\in I} \langle i| A |i\rangle = \sum_{i\in I} \mathrm{Tr}(A |i\rangle\langle i|) = \mathrm{Tr}(A)$, using the equality $\sum_{i\in I} |i\rangle\langle i| = \mathbb{I}$. Thus, the above definition of the trace coincides with the more usual one where $\mathrm{Tr}(A)$ is defined as the sum of the diagonal elements of a matrix representation of A.

Definition 5.5. *A linear operator* $\rho : \mathcal{H} \to \mathcal{H}$ *is called a* density matrix *if* ρ *is positive semi-definite (i.e.,* $\langle\phi| \rho |\phi\rangle \geq 0$ *for all* $|\phi\rangle \in \mathcal{H}$*) and* $\mathrm{Tr}(\rho) = 1$.

Definition 5.6. *An operator* $U \in \mathcal{L}(\mathcal{H})$ *is called* unitary *if* $U^\dagger U = \mathbb{I}$.

An equivalent characterization of unitaries is that they map an orthonormal basis to an orthonormal basis, i.e., U is unitary if and only if there are orthonormal bases $\{|i\rangle\}_{i\in I}$ and $\{|\phi_i\rangle\}_{i\in I}$ such that $U = \sum_{i\in I} |\phi_i\rangle\langle i|$.

Definition 5.7. *Let* $P : \mathcal{H} \to \mathcal{H}$ *be a linear operator. We say that* P *is a projector if* $P^2 = P$ *and* $P^\dagger = P$. *We say that two projectors* P *and* Q *are* (mutually) orthogonal *if* $PQ = 0$.[1]

Definition 5.8. *Suppose that* $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. *The* partial trace *is defined as the linear operator*

$$\mathrm{Tr}_A : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}_B),\ \rho_A \otimes \rho_B \mapsto \mathrm{Tr}(\rho_A)\rho_B$$

The partial trace Tr_B *is defined symmetrically.*

5.2.2 Quantum States and Measurements

A (finite-dimensional) quantum system A is associated with a *state space* $\mathcal{H}_A$, which is a finite-dimensional, complex Hilbert space. The *state* of a quantum system is represented as a density matrix ρ over $\mathcal{H}_A$. We identify the state of a quantum system with the density matrix that describes it and also call ρ a quantum state.

A quantum state can be acted on in the following two ways: The first is to *apply a unitary* U, transforming the state ρ into $U\rho U^\dagger$. The second way is to *perform a measurement*, which is the only way to extract classical information from a quantum state. A (projective) measurement is described by a collection $\{P_i\}_{i\in I}$ of mutually orthogonal projectors such that $\sum_{i\in I} P_i = \mathbb{I}$,

[1] Usually, projectors are defined just by the property $P^2 = P$. If they also satisfy $P^\dagger = P$, they are usually called *orthogonal projectors*. However, since all projectors we consider are orthogonal projectors, we reserve the term *orthogonal* for mutually orthogonal pairs of projectors.

where I is some finite index set. The measurement produces outcome i with probability $p_i := \mathrm{Tr}(P_i \rho P_i^\dagger) = \mathrm{Tr}(P_i \rho)$. If the state is measured and outcome i is observed, the state *collapses* to $\frac{1}{p_i} P_i \rho P_i$. There are more general formalisms for measurements, but we can restrict to projective measurements without loss of generality (see *Naimark's Dilation Theorem*).

If we can write $\rho = |\phi\rangle\langle\phi|$, ρ is called a *pure* state and we can use $|\phi\rangle$ as a representation of the quantum state. In this representation, applying a unitary U maps $|\phi\rangle$ to $U|\phi\rangle$. When performing a measurement, we observe outcome i with probability $p_i = \| P_i |\phi\rangle \|^2$ and obtain post-measurement state $\frac{1}{p_i} P_i |\phi\rangle$. If a quantum system is in the pure state $|\phi_i\rangle$ with probability p_i, it is represented by the density matrix $\sum_i p_i |\phi_i\rangle\langle\phi_i|$.

For every orthonormal basis $B = \{|i\rangle\}_{i \in I}$ of $\mathcal{H}$, $\{|i\rangle\langle i|\}_{i \in I}$ is a projective measurement, called the *total projective measurement* with respect to B or simply the measurement in basis B.

5.2.3 Bi-partite Quantum States and Entanglement

A crucial concept for this chapter is *quantum entanglement* shared between two parties. Let A and B be two quantum systems with respective state spaces $\mathcal{H}_A$ and $\mathcal{H}_B$. We may consider the two systems together as a single joint quantum system AB. The state space of AB is $\mathcal{H}_{AB} := \mathcal{H}_A \otimes \mathcal{H}_B$. If A and B are prepared independently in states ρ_A and ρ_B, respectively, the state of the joint system is $\rho_A \otimes \rho_B$. If the state of the joint system is a pure tensor like this, we say that it is *in product state*. A product state can be seen as an analogue to a pair of independent random variables. A state is called *separable* if it can be written as $\rho_{AB} = \sum_i p_i \rho_{A,i} \otimes \rho_{B,i}$ where the $\rho_{A,i}$ and $\rho_{B,i}$ are density matrices, $p_i \geq 0$ and $\sum_i p_i = 1$. A separable state is analogous to a pair of possibly correlated random variables. Finally, states that are not separable are called *entangled* and do not have any classical analogue.

If we perform an action on system A, the joint system AB is affected as follows: If a unitary U_A is applied on A, it acts as $U_A \otimes \mathbb{I}_B$ on the whole system. A measurement $\{P_i\}_{i \in I}$ on A acts as $\{P_i \otimes \mathbb{I}_B\}_{i \in I}$ on the joint system. Symmetrically, applying a unitary U_B on system B acts as $\mathbb{I}_A \otimes U_B$ and a measurement $\{P_i\}_{i \in I}$ on B acts as $\{\mathbb{I}_A \otimes P_i\}_{i \in I}$ on AB.

A final operation that can be performed on a joint system is to *remove* a part of it. When we have a joint system AB in state ρ_{AB}, the state of the subsystem B on its own is described by the *partial trace* $\mathrm{Tr}_A(\rho_{AB})$.

The above generalizes to tripartite (and, more generally, n-partite) states. Entanglement does not depend on physical proximity. Two agents that are far apart – like the provers in a relativistic bit-commitment scheme – can each keep one part of an entangled quantum state and apply unitaries and measurements to their part. Sharing an entangled quantum state allows two parties to correlate their behavior without communicating in a way that is not possible classically. We give an example of this phenomenon in the next

section.

5.2.4 Example: A Strategy for CHSH

To illustrate the effects of quantum entanglement, we now describe a strategy for the CHSH game using an entangled quantum state, which has a better success probability than any strategy that relies solely on shared randomness. Recall that the CHSH game works as follows [CHSH69]: The players Alice and Bob each receive an input bit a and b, respectively, and they each have to output a bit x and y, respectively, without communicating. They win if $x+y = a \cdot b$. In other words, if $a = b = 1$, they have to output *different* bits, and in all other cases, they have to output *identical* bits. It is easy to see that the maximal success probability for strategies using shared classical randomness is 0.75. However, there is a strategy that uses quantum entanglement and achieves a success probability of $\cos(\pi/8)^2 \approx 0.85$.

Let A and B be quantum systems with state spaces $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^2$. We write the standard basis as $\{|0\rangle, |1\rangle\}$ and define the following vectors:

$$
\begin{aligned}
|+\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\
|-\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \\
|\phi_0\rangle &= \cos(\pi/8)\,|0\rangle + \sin(\pi/8)\,|1\rangle \\
|\phi_1\rangle &= \sin(\pi/8)\,|0\rangle - \cos(\pi/8)\,|1\rangle \\
|\psi_0\rangle &= \cos(\pi/8)\,|0\rangle - \sin(\pi/8)\,|1\rangle \\
|\psi_1\rangle &= \sin(\pi/8)\,|0\rangle + \cos(\pi/8)\,|1\rangle
\end{aligned}
$$

$\{|+\rangle, |-\rangle\}$, $\{|\phi_0\rangle, |\phi_1\rangle\}$, and $\{|\psi_0\rangle, |\psi_1\rangle\}$ are orthonormal bases of $\mathbb{C}^2$. The measurements in the standard basis and each of those bases are defined as follows.

$$
\begin{aligned}
M_0^A &= \{|0\rangle\langle 0|, |1\rangle\langle 1|\} \\
M_1^A &= \{|+\rangle\langle +|, |-\rangle\langle -|\} \\
M_0^B &= \{|\phi_0\rangle\langle \phi_0|, |\phi_1\rangle\langle \phi_1|\} \\
M_1^B &= \{|\psi_0\rangle\langle \psi_0|, |\psi_1\rangle\langle \psi_1|\}
\end{aligned}
$$

The system AB is prepared in state $|\Phi\rangle = \frac{1}{\sqrt{2}}(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle)$. Alice keeps system A and Bob keeps system B. The strategy now works as follows. The players apply the measurement M_a^A or M_b^B, respectively, on their part of the quantum state. Their outputs are their respective measurement outcomes.

First, let us consider the case $a = 0$, $b = 0$. Alice applies the measurement M_0^A, which results in a uniformly random outcome $x \in \{0, 1\}$ with post-measurement state $|x\rangle \otimes |x\rangle$. Bob applies M_0^B. In order to win the game, he needs to output the same bit as Alice, and thus the winning probability is

$$\|(\mathbb{I}_A \otimes |\phi_x\rangle \langle\phi_x|)(|x\rangle \otimes |x\rangle)\|^2 = |\langle\phi_x|x\rangle|^2 = \cos(\pi/8)^2 \approx 0.85$$

If $a = 0$, $b = 1$, the two players again need to output the same bit. Thus, the probability that Bob produces the correct output is $|\langle\psi_x|x\rangle|^2 = \cos(\pi/8)^2 \approx 0.85$. If $a = 1$, $b = 0$, Alice outputs a uniformly random $x \in \{0, 1\}$, and the post-measurement state is $|+\rangle \otimes |+\rangle$ if $x = 0$ and $|-\rangle \otimes |-\rangle$ if $x = 1$. They need to output the same bit again, and thus the success probability is

$$\begin{aligned} |\langle +|\phi_0\rangle|^2 = |\langle -|\phi_1\rangle|^2 &= \left(\frac{1}{\sqrt{2}}(\sin(\pi/8) + \cos(\pi/8))\right)^2 \\ &= (\cos(\pi/4)\cos(\pi/8) + \sin(\pi/4)\sin(\pi/8))^2 \\ &= \cos(\pi/8)^2 \approx 0.85 \end{aligned}$$

Finally, we consider the case $a = 1$, $b = 1$. Here, the provers win if they produce *different* outputs. The winning probability is

$$|\langle +|\psi_1\rangle|^2 = |\langle -|\psi_0\rangle|^2 = \left(\frac{1}{\sqrt{2}}(\sin(\pi/8) + \cos(\pi/8))\right)^2 = \cos(\pi/8)^2 \approx 0.85$$

and thus, for all possible inputs a and b, the players can win with probability $\cos(\pi/8)^2 \approx 0.85$, without communicating.

5.2.5 Representing Classical Information and Randomness

The formalism from the preceding sections can also be used to capture classical information. We can represent a single bit as a state of a *qubit* system, i.e., a quantum system with state space $\mathbb{C}^2$. We write $\{|0\rangle, |1\rangle\}$ for the standard basis of $\mathbb{C}^2$ and represent the bit b as the state vector $|b\rangle$. More generally, we can represent elements of a finite set S via a bijection mapping each element s to a standard basis vector $|s\rangle$ of $\mathbb{C}^{|S|}$. This representation can also be understood as an encoding, where decoding is done by measuring the quantum state in the standard basis. Note that because the quantum state itself is one of the basis vectors, performing this measurement has a deterministic outcome and does not alter the state. A distribution $p(s)$ over some set S is then naturally represented by the density matrix $\sum_{s\in S} p(s)\,|s\rangle\langle s|$. Shared randomness can be represented as the separable bipartite state $\sum_{s\in S} p(s)\,|s\rangle\langle s| \otimes |s\rangle\langle s|$.

It is possible to represent all classical computations as unitaries. Let $\mathcal{X}$ be a finite set and $\mathcal{Y}$ a finite group, where the group operation is denoted

by + and the neutral element by 0. Let $f : \mathcal{X} \to \mathcal{Y}$ be some function. Let $U_f : \mathbb{C}^{|X|} \otimes \mathbb{C}^{|Y|} \to \mathbb{C}^{|X|} \otimes \mathbb{C}^{|Y|}, |x\rangle \otimes |y\rangle \mapsto |x\rangle \otimes |y + f(x)\rangle$. It is easy to see that U_f is unitary because it maps an orthonormal basis to an orthonormal basis. We can represent the computation of $f(x)$ as follows in the quantum formalism: Given $|x\rangle$, we first append $|0\rangle \in \mathbb{C}^{|Y|}$ and then apply U_f. The second quantum system now holds the desired result.

5.3 Protocols

In this section, we adapt our formal definitions of interactive protocols from Section 2.2.1 to the quantum setting. We again consider protocols involving three parties, the provers P and Q and the verifier V. A protocol $\mathsf{prot}_{PQV} = (\mathsf{prot}_P, \mathsf{prot}_Q, \mathsf{prot}_V)$ consists of a triple of l-round interactive algorithms operating on a quantum state ρ_{PQV} over a Hilbert space $\mathcal{H}_{PQV} = \mathcal{H}_P \otimes \mathcal{H}_Q \otimes \mathcal{H}_V$. The subscripts P, Q and V indicate the player that controls the system. The players can perform the following actions on their respective parts of the quantum states:

- apply unitaries and perform measurements,
- prepare an additional quantum system in some initial state and add it to their system,
- discard a part of their quantum state,
- transmit part of their quantum state to another player, making them part of that player's state in the next round.

The outcome of this procedure is a quantum state ρ'_{PQV} over a Hilbert space $\mathcal{H}'_{PQV} = \mathcal{H}'_P \otimes \mathcal{H}'_Q \otimes \mathcal{H}'_V$. Note that the state spaces may change due to players exchanging parts of their states and preparing additional subsystems. We write

$$\rho'_{PQV} \leftarrow \mathsf{prot}_{PQV}(\rho_{PQV})$$

to denote an execution of the protocol on the input state ρ_{PQV}. As in the classical case, we can compose protocols by using the output state of one as the input for the other. We separate shared entanglement from the input: to denote shared entanglement between P and Q, we write $\mathsf{prot}_{PQV}[\sigma_{PQ}]$.

A commitment scheme with domain D consists of two interactive quantum algorithms: The first, $\mathsf{com}_{PQV}[\sigma_{PQ}] = (\mathsf{com}_P, \mathsf{com}_Q, \mathsf{com}_V)$ takes an input state $|s\rangle_P \otimes |s\rangle_Q \in \left(\mathbb{C}^{|D|}\right)^{\otimes 2}$ for P and Q, and V has no input. The output is some state ρ'_{PQV}. $\mathsf{open}_{PQV}[\sigma'_{PQ}]$ then takes the output of com_{PQV} as input and V produces a (quantum representation of a) single bit as output, indicating whether the commitment was opened correctly or not. P and Q produce no output.

Of particular interest is the security of a classical commitment scheme (i.e., one that requires only classical information processing on the part of the honest players) against dishonest provers with quantum capabilities. In this case, they can not send quantum information to the honest V. We may model this as V immediately measuring all quantum states he receives in the standard basis of their respective state space.

For classical commitment schemes, the definitions of soundness and the hiding property simply carry over. If the scheme requires the verifier to store quantum information, the hiding property needs to be adapted: we say that the scheme is δ-hiding if no measurement that V can perform on his quantum state allows him to distinguish between any pair of strings s_0 and s_1 with probability better than δ.[2]

5.4 Binding Properties

5.4.1 Definition

As we already mentioned, the weak-binding properties defined in Section 3.2.3 carry over to the quantum setting without change. We define the following strong binding property for the quantum case, essentially replacing the function $\hat{s}$ with a measurement. As usual, the binding property is defined with respect to some set of possible strategies for the dishonest players, e.g., strategies where they do not communicate.

Definition 5.9. *Let S be a commitment scheme. We say that S is* ε-binding *if for every dishonest opening strategy $\overline{\mathsf{com}}_P$ there exists a measurement $\mathsf{Eval} = \{M_{\hat{s}}\}_{\hat{s}}$ such that for every possible shared quantum states σ_{PQ} and every dishonest opening strategy $\overline{\mathsf{open}}_Q$ we have $p(s \neq \hat{s} \wedge s \neq \perp) \leq \varepsilon$ where s is the output of the verifier after the opening phase and $\hat{s}$ is the outcome of the measurement Eval applied to the state $\rho'_{PV} = \mathrm{Tr}_Q\left(\rho'_{PQV}\right)$ where ρ'_{PQV} is the state after the execution of $\overline{\mathsf{com}}_{PQV}[\sigma_{PQ}]$. We say that it is* fairly ε*-binding if $p(s \neq \hat{s} \wedge s = s_\circ) \leq \varepsilon$ holds for all $s_\circ$.*

In order to actually perform the measurement Eval, P needs to obtain (a copy of) V's quantum state. This is possible if only P communicates with V in the commit phase, V sends (a copy of) his initial state to P, and all the communication with V is classical, so P can simply store copies of all messages he sends to V. If only Q is active in the opening phase, the measurement can be performed without affecting the outcome of the protocol execution. These conditions are all satisfied in the $\mathcal{CHSH}^q$ scheme.

We do not currently know whether there exists a scheme that satisfies Definition 5.9. However, we were able to show that $\mathcal{CHSH}^q$ satisfies the weak-binding definition as a string-commitment scheme.

[2] To be more formal, one might use the *trace distance* here.

5.4.2 The Finite Field CHSH Game With Restricted Inputs

We now prove that $\mathcal{CHSH}^q$ is weak-binding with respect to quantum adversaries. Our proof is based on an analysis of the following CHSH-like game: Let $S \subseteq \mathbb{F}_q$ with $|S| = N$. In the game CHSH^q_S, the two players Alice and Bob receive uniformly random inputs $a \in \mathbb{F}_q$ and $b \in S$ and, without communicating, produce outputs x and y in $\mathbb{F}_q$. They win if $x + y = a \cdot b$.

Theorem 5.10. *The game CHSH^q_S with $|S| = N$ has quantum value*

$$\frac{1}{N} + \frac{N-1}{\sqrt{q}}.$$

We use the following lemma:

Lemma 5.11. *Let $\{\Pi_i\}_{i\in\mathcal{I}}$ be a collection of projectors on a Hilbert space $\mathcal{H}$ and let $|\phi\rangle$ be a unit vector in $\mathcal{H}$. For $i, j \in \mathcal{I}$, let $\varepsilon_{i,j} = \|\Pi_i\Pi_j |\phi\rangle\|$. Then, $\| \sum_{i\in\mathcal{I}} \Pi_i |\phi\rangle \| \leq 1 + \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{I}\setminus\{i\}} \varepsilon_{i,j}$.*

Proof. If $\| \sum_i \Pi_i |\phi\rangle \| \leq 1$, the conclusion obviously holds, so we assume that $\| \sum_i \Pi_i |\phi\rangle \| > 1$. We have

$$\begin{aligned}
\left\| \sum_i \Pi_i |\phi\rangle \right\|^2 &= \left| \sum_{i,j} \langle\phi| \Pi_i\Pi_j |\phi\rangle \right| \\
&\leq \sum_i \langle\phi| \Pi_i |\phi\rangle + \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{I}\setminus\{i\}} |\langle\phi| \Pi_i\Pi_j |\phi\rangle| \\
&\leq \left\| \sum_i \Pi_i |\phi\rangle \right\| + \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{I}\setminus\{i\}} \underbrace{\||\phi\rangle\|}_{=1} \cdot \|\Pi_i\Pi_j |\phi\rangle\| \text{ by Cauchy-Schwarz} \\
&= \left\| \sum_i \Pi_i |\phi\rangle \right\| + \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{I}\setminus\{i\}} \varepsilon_{i,j}
\end{aligned}$$

which implies that

$$\left\| \sum_i \Pi_i |\phi\rangle \right\| = 1 + \frac{\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{I}\setminus\{i\}} \varepsilon_{i,j}}{\|\sum_i \Pi_i |\phi\rangle\|} \leq 1 + \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{I}\setminus\{i\}} \varepsilon_{i,j}$$

as we claimed. □

Proof of Theorem 5.10. A quantum strategy for Alice and Bob is (without loss of generality) described by a bipartite quantum state $|\phi\rangle \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$, a projective measurement $\{P^a_x\}_{x\in\mathbb{F}_q}$ on $\mathcal{H}_A$ for every $a \in \mathbb{F}_q$, and a projective measurement $\{Q^b_y\}_{y\in\mathbb{F}_q}$ for every $b \in S$. The output distribution of Alice and

Bob on inputs a and b is $p(x,y|a,b) = \langle\phi| (P_x^a \otimes Q_y^b) |\phi\rangle$. Let win be the event that $x + y = a \cdot b$. The winning probability conditioned on Bob's input b is

$$p(\mathsf{win}|b) = \frac{1}{q} \sum_{a,x} \langle\phi| (P_x^a \otimes Q_{a\cdot b-x}^b) |\phi\rangle = \frac{1}{q} \sum_a \langle\phi| \Pi_{a,b} |\phi\rangle$$

where $\Pi_{a,b} = \sum_x P_x^a \otimes Q_{a\cdot b-x}^b$. Note that $\Pi_{a,b}$ is a projector since it is the sum of a set of mutually orthogonal projectors. In order to bound $p(\mathsf{win}|b)$, we consider an extended version of the game where Bob receives two distinct inputs b and b' and produces two outputs y and y'. Here, the players win if $x + y = a \cdot b$ *and* $x + y' = a \cdot b'$. We write win for the former event and win' for the latter. It is easy to see that if $b \neq b'$, then no matter what strategy the players use, $p(\mathsf{win}, \mathsf{win}'|b,b') = 1/q$: If both winning conditions are satisfied, then $y - y' = a \cdot (b - b')$, so $a = (y - y') \cdot (b - b')^{-1}$. Since Bob does not know a, this holds with probability $1/q$. One strategy for the extended game is that Bob first applies the measurement $\{Q_y^b\}_{y\in\mathbb{F}_q}$ and then the measurement $\{Q_{y'}^{b'}\}_{y'\in\mathbb{F}_q}$ on the post-measurement state. Thus,

$$\begin{aligned}
\frac{1}{q} = p(\mathsf{win}, \mathsf{win}'|b,b') &= \frac{1}{q} \sum_{a,x} \langle\phi| \left(P_x^a \otimes Q_{a\cdot b'-x}^{b'} Q_{a\cdot b-x}^b\right) |\phi\rangle \\
&= \frac{1}{q} \sum_a \langle\phi| \left(\sum_x P_x^a \otimes Q_{a\cdot b'-x}^{b'} Q_{a\cdot b-x}^b\right) |\phi\rangle \\
&= \frac{1}{q} \sum_a \left\| \sum_x \left(P_x^a \otimes Q_{a\cdot b'-x}^{b'} Q_{a\cdot b-x}^b\right) |\phi\rangle \right\|^2
\end{aligned}$$

where we use that the $\{P_x^a\}_x$ are mutually orthogonal projectors. Using that

$$\sum_x P_x^a \otimes Q_{a\cdot b'-x}^{b'} Q_{a\cdot b-x}^b = \sum_{x,x'} (P_x^a \otimes Q_{a\cdot b'-x}^{b'})(P_{x'}^a \otimes Q_{a\cdot b-x'}^b) = \Pi_{a,b}\Pi_{a,b'}$$

we conclude that $\sum_a \|\Pi_{a,b'}\Pi_{a,b} |\phi\rangle\|^2 = 1$. It follows that $\sum_a \|\Pi_{a,b'}\Pi_{a,b} |\phi\rangle\| \leq \sqrt{q}$.

It holds that

$$\begin{aligned}
p(\mathsf{win}) = \frac{1}{N}\sum_b p(\mathsf{win}|b) &= \frac{1}{qN}\sum_{a,b} \langle\phi|\,\Pi_{a,b}\,|\phi\rangle \\
&= \frac{1}{qN}\sum_a \langle\phi|\sum_b \Pi_{a,b}\,|\phi\rangle \\
&\leq \frac{1}{qN}\sum_a \left\|\sum_b \Pi_{a,b}\,|\phi\rangle\right\| \text{ by Cauchy-Schwarz} \\
&\leq \frac{1}{qN}\sum_a \left(1 + \sum_{b\neq b'} \|\Pi_{a,b}\Pi_{a,b'}\,|\phi\rangle\|\right) \text{ by Lemma 5.11} \\
&= \frac{1}{N} + \frac{1}{qN}\sum_{b\neq b'}\sum_a \|\Pi_{a,b}\Pi_{a,b'}\,|\phi\rangle\,\| \\
&\leq \frac{1}{N} + \frac{N(N-1)\sqrt{q}}{qN} \\
&\leq \frac{1}{N} + \frac{N-1}{\sqrt{q}}
\end{aligned}$$

□

5.4.3 Binding Property of the Commitment Scheme

Theorem 5.12. *The $\mathcal{CHSH}^q$ scheme is ε-fairly-weak-binding for $\varepsilon = 2/\sqrt[4]{q}$.*

Proof. Fix a commit strategy $\overline{\mathsf{com}}_{PQ}$ against the scheme. Enumerate the elements of $\mathbb{F}_q$ as $s_1, \ldots, s_q$, and for every $i \in \{1, \ldots, q\}$ let $\overline{\mathsf{open}}^i_{PQ}$ be an opening strategy maximizing $p_i := p(s = s_i)$, where s is the output of the verifier when P and Q use this strategy. We assume without loss of generality that the p_is are in descending order. We define $p(\hat{s})$ as follows. Let $N \geq 2$ be an integer which we will fix later. By Theorem 5.10, it holds that

$$\sum_{i=1}^{N} p_i \leq 1 + \frac{N(N-1)}{\sqrt{q}}$$

where we let $p_i = 0$ for $i > q$ in case $N > q$. To see that this inequality holds, consider the game CHSH^q_S with $S = \{s_1, \ldots, s_N\}$. We let Alice produce her output x using the strategy $\overline{\mathsf{com}}_{PQ}$ and we let Bob use the strategy $\overline{\mathsf{open}}^i_{PQ}$ on input s_i. Given that the input is s_i, this strategy succeeds with probability p_i. From our bound on the quantum value of this game, our bound on the sum of probabilities follows.

We would like to define $p(\hat{s})$ as $p(\hat{s} = s_i) := p_i - (N-1)/\sqrt{q}$ for all $i \leq N, q$; however, this is not always possible because $p_i - (N-1)/\sqrt{q}$ may be

negative. To deal with this, let N' be the largest integer such that $N' \leq N$ and $p_1, \ldots, p_{N'} \geq (N-1)/\sqrt{q}$. (We let $N' = 0$ if $p_1 < (N-1)/\sqrt{q}$.) It follows that

$$\sum_{i=1}^{N'} p_i \leq 1 + \frac{N'(N'-1)}{\sqrt{q}} \leq 1 + \frac{N'(N-1)}{\sqrt{q}} \quad \text{and thus} \quad \sum_{i=1}^{N'} p_i = 1 + N'(N-1)\cdot\varepsilon$$

for some $\varepsilon \leq 1/\sqrt{q}$. We now set $p(\hat{s})$ to be $p(\hat{s} = s_i) := p_i - (N-1)\varepsilon \geq 0$ for all $i \leq N'$. Now consider an opening strategy $\overline{\mathsf{open}}_{PQ}$ and let $p(s)$ be the resulting output distribution. By definition of the p_i, it follows that $p(s = s_i) \leq p_i$ for all $i \leq q$, and $p_i \leq p(\hat{s} = s_i) + (N-1)/\sqrt{q}$ for all $i \leq N'$. It follows that there is a joint distribution $p(\hat{s}, s)$ with $p(\hat{s} = s = s_i) = \min\{p(s = s_i), p(\hat{s} = s_i)\} \geq p(s = s_i) - (N-1)/\sqrt{q}$ for all $i \leq N'$, and thus $p(\hat{s} \neq s = s_i) = p(s = s_i) - p(\hat{s} = s = s_i) \leq (N-1)/\sqrt{q}$ for all $i \leq N'$. Furthermore, when $N' < i \leq N$, we have $p(\hat{s} \neq s = s_i) = p(s = s_i) \leq p_i < (N-1)/\sqrt{q}$ by definition of N'. Since the p_i are sorted in descending order, it follows that for all $i > N$

$$p(\hat{s} \neq s = s_i) = p(s = s_i) \leq p_i \leq p_N \leq \frac{1}{N}\sum_{i=1}^{N} p_i \leq \frac{1}{N} + \frac{N-1}{\sqrt{q}}$$

and thus, we have shown for all $s_\circ \in \mathbb{F}_q$ that

$$p(\hat{s} \neq s = s_\circ) \leq \frac{1}{N} + \frac{N-1}{\sqrt{q}}.$$

We now select N so that this value is minimized: it is easy to verify that the function $f : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$, $x \mapsto 1/x + (x-1)/\sqrt{q}$ has its global minimum in $\sqrt[4]{q}$; thus, we pick $N := \lceil \sqrt[4]{q} \rceil$, which gives us

$$p(\hat{s} \neq s = s_\circ) \leq \frac{1}{N} + \frac{N-1}{\sqrt{q}} \cdot \varepsilon \leq \frac{1}{\sqrt[4]{q}} + \frac{\sqrt[4]{q}}{\sqrt{q}} = \frac{2}{\sqrt[4]{q}}$$

for any $s_\circ \in \mathbb{F}_q$, as claimed. □

5.5 The Composition Theorem

5.5.1 The Composition Operation

Besides considering adversaries with quantum capabilities, this composition theorem also differs from the previous one in that it composes a weak-binding and binding scheme to produce a weak-binding scheme. The structure of the proof is somewhat different as well: In the proof of the classical composition theorem, we composed a "small" $\mathcal{S}$ (in the sense that only one prover communicates with the verifier in the commit phase and only the other in the

opening phase) with a "big" $\mathcal{S}'$. In the proof of our composition theorem for the quantum case, we instead compose a "big" $\mathcal{S}$ with a small $\mathcal{S}'$. This requires a slightly different definition of eligible pairs $(\mathcal{S}, \mathcal{S}')$ where some properties of $\mathcal{S}$ and $\mathcal{S}'$ are reversed:

Definition 5.13. *Let $\mathcal{S}$ and $\mathcal{S}'$ be two 2-prover string-commitment schemes with domains D and D', respectively. We call the pair $(\mathcal{S}, \mathcal{S}')$* eligible *if they are both classical (i.e., no quantum communication or computation is required on the part of the honest verifier and provers) and the following two properties hold, or they hold with the roles of P and Q exchanged.*

1. *The commit phase of $\mathcal{S}$ is a protocol com_{PQV} that involves communication between P and V only. The commit phase of $\mathcal{S}'$ is a protocol com'_{PQV} that involves communication between Q and V only and the opening phase is a protocol open'_{PQV} involving communication between P and V only.*

2. *The last round of open_{PQV} is of the following simple form: Q sends some $y \in D'$ to V who computes the output deterministically as $s = \mathsf{Extr}(y, \bar{c})$ where $\bar{c}$ consists of all communication in the previous rounds (including the commit phase). We call this message y the* final opening information*.*

Furthermore, we specify that the possible attacks on $\mathcal{S}$ are so that P and Q do not communicate during the course of the commit phase, but there may be some limited communication during the opening phase. The possible attacks on $\mathcal{S}'$ are so that P and Q do not communicate during the course of the entire execution of the scheme.

The composition operation $\mathcal{S} \star \mathcal{S}'$ is defined in the same way as in the classical setting: instead of sending the opening information y in the last round, the provers instead use $\mathcal{S}'$ to commit to y and then open the commitment. We define the possible attacks $(\overline{\mathsf{com}}_{PQ}, \overline{\mathsf{open}}''_{PQ})$ on $\mathcal{S} \star \mathcal{S}'$ to be those where $\overline{\mathsf{com}}_{PQ}$ is a commit strategy for $\mathcal{S}$ and $\overline{\mathsf{open}}''_{PQ}$ can be decomposed as $\overline{\mathsf{open}}'_{PQ} \circ \overline{\mathsf{com}}'_{PQ} \circ \mathsf{ptoq} \circ \overline{\mathsf{open}}^*_{PQ}$ where ptoq allows P to send a quantum state to Q, $\overline{\mathsf{open}}^*_{PQ}$ is an opening strategy for $\mathcal{S}$ excluding the last round. For any input state ρ_{PQ}, $\overline{\mathsf{com}}'_{PQ}(\rho_{PQ})$ is a commit strategy and $\overline{\mathsf{open}}'_{PQ}$ is an opening strategy for $\mathcal{S}'$.

5.5.2 The Composition Theorem

In the following composition theorem, we take it as understood that the assumed respective binding properties of $\mathcal{S}$ and $\mathcal{S}'$ hold with respect to a well-defined respective classes of possible attacks.

Theorem 5.14. *Let $(\mathcal{S}, \mathcal{S}')$ be an eligible pair of 2-prover commitment schemes, and assume that $\mathcal{S}$ is ε-fairly-weak-binding and that $\mathcal{S}'$ is δ-fairly-binding. Then, their composition $\mathcal{S}'' = \mathcal{S} \star \mathcal{S}'$ is a $(\varepsilon + k(\mathcal{S}) \cdot \delta)$-fairly-weak-binding 2-prover commitment scheme.*

Proof. We first consider the case where $k(\mathcal{S}) = 1$. We fix an arbitrary possible strategy $(\overline{\mathsf{com}}_P, \overline{\mathsf{open}}''_{PQ})$ against $\mathcal{S}''$, where $\overline{\mathsf{open}}''_{PQ}$ is of the form $\overline{\mathsf{open}}''_{PQ} = \overline{\mathsf{open}}'_{PQ} \circ \overline{\mathsf{com}}'_{PQ} \circ \mathsf{ptoq} \circ \overline{\mathsf{open}}^*_{PQ}$. We decompose and reassemble the strategy into strategies against $\mathcal{S}$ and $\mathcal{S}'$. The strategy for $\mathcal{S}'$ uses a shared quantum state ρ_{PQ} of appropriate size and works by executing $\overline{\mathsf{com}}'_{PQ}(\rho_{PQ})$ and then $\overline{\mathsf{open}}'_{PQ}$. By the δ-fairly-binding property of $\mathcal{S}'$, there exists a measurement $\mathsf{Eval} = \{M_{\hat{y}}\}_{\hat{y}}$ depending only on $\overline{\mathsf{com}}'_{PQ}$ which Q can apply after $\overline{\mathsf{com}}$ such that $p(y \neq \hat{y} \wedge y = y_\circ) \leq \delta$ for every $y_\circ$, where y is the output of $\overline{\mathsf{open}}_{PQV}$. This holds for every possible shared quantum state ρ_{PQ}, and in particular for the ones generated as follows: Let σ_{PQV} be the quantum state after an execution of $\overline{\mathsf{com}}_{PQV}$ and $\overline{\mathsf{open}}^*_{PQV}$. This execution involves sending classical messages $\bar{c}$ between the provers and the verifier. Thus, we may write $\sigma_{PQV} = \sum_{\bar{c}} p(\bar{c})\sigma^{\bar{c}}_{PQV}$. Writing $p(y, \hat{y}|\bar{c})$ for the distribution of the output and the measurement outcome when using the shared quantum state $\rho_{PQ} = \sigma^{\bar{c}}_{PQ}$, it holds that $p(y \neq \hat{y} \wedge y = y_\circ|\bar{c}) \leq \delta$.

Note that $\overline{\mathsf{com}}_P$ is a commit strategy for $\mathcal{S}$. As such, by the weak-binding property of $\mathcal{S}$, there exists a distribution $p(\hat{s})$, only depending on $\overline{\mathsf{com}}_P$, so that Definition 3.9 is satisfied for every opening strategy $\overline{\mathsf{open}}_Q$ for $\mathcal{S}$. In particular, this holds for the following opening strategy:

1. The provers execute $\overline{\mathsf{open}}^*_{PQ}$.
2. Q simulates an execution of $\overline{\mathsf{com}}'_Q$ and performs the measurement Eval with result $\hat{y}$.
3. Q sends $\hat{y}$ to V.

By the ε-fairly-weak binding property of $\mathcal{S}$, there is a consistent joint distribution $p(s, \hat{s})$ such that for every $s_\circ$, $p(\tilde{s} \neq \hat{s} \wedge \tilde{s} = s_\circ) \leq \varepsilon$ where $\tilde{s} = \mathsf{Extr}(\hat{y}, \bar{c})$. Let $y_\circ(\bar{c})$ be such that $\mathsf{Extr}(y_\circ, \bar{c}) = s_\circ$ (or some default value if no such $y_\circ$ exists). Since we assume that $k(\mathcal{S}) = 1$, there is only one possible value for $y_\circ(\bar{c})$. Recalling that $s = \mathsf{Extr}(y, \bar{c})$, it holds that

$$\begin{aligned}
p(\hat{s} \neq s \wedge s = s_\circ) &= p(\hat{s} \neq s \wedge s = s_\circ \wedge s = \tilde{s}) + p(\hat{s} \neq s \wedge s = s_\circ \wedge s \neq \tilde{s}) \\
&\leq p(\hat{s} \neq \tilde{s} \wedge \tilde{s} = s_\circ) + p(\mathsf{Extr}(y, \bar{c}) \neq \mathsf{Extr}(\hat{y}, \bar{c}) \wedge \mathsf{Extr}(y, \bar{c}) = s_\circ) \\
&\leq p(\hat{s} \neq \tilde{s} \wedge \tilde{s} = s_\circ) + \sum_{\bar{c}} p(\bar{c}) p(y \neq \hat{y} \wedge y = y_\circ(\bar{c})|\bar{c}) \\
&\leq \varepsilon + \delta
\end{aligned}$$

and in the case that $k(\mathcal{S}) > 1$, it is easy to see that we can upper-bound this probability with $\varepsilon + k(\mathcal{S})\delta$. □

Chapter 6

Bit-commitment with Non-signaling Adversaries

6.1 Introduction

In this chapter, we shift our focus to the no-communication assumption. Specifically, we consider if (1-round) two-prover bit-commitment schemes can be secure based on the *sole assumption* that the provers cannot communicate. In the previous chapters, we imposed some limit on the resources that the provers could use to correlate their behaviour. In Chapters 3 and 4, we assumed that the provers can only use classical shared randomness and in Chapter 5, we considered provers that can use an entangled quantum state to correlate their actions. But even for single-round schemes, the assumptions on the provers' resources can make a difference for the security of the scheme. The $\mathcal{CHSH}^q$ scheme is binding for classical provers, and also against quantum provers, albeit with a weaker parameter. However, in [CSST11], it is shown that there exists a bit-commitment scheme that is binding in the classical setting, but not in the quantum setting. The scheme in question is essentially an error-tolerant version of $\mathcal{CHSH}^{2^n}$: instead of requiring that $x + y = a \cdot b$, we require that when we parse $x + y$ and $a \cdot b$ as bit-strings, 85% of the positions are equal.

This condition means that for 85% of indices $i \in \{1, \ldots, n\}$, the CHSH winning condition $x_i \oplus y_i = a_i \cdot b$ has to be satisfied. In other words, the provers can open to an arbitrary bit if, in a series of CHSH games, they can win 85% of the time. Since there is a strategy that wins a CHSH game with probability $\approx 85\%$ using quantum entanglement (see Section 5.2.4), there is a high probability that dishonest provers in the quantum setting can open to any bit they want. Furthermore, as discussed in Section 1.4.4, $\mathcal{CHSH}^q$ is insecure against arbitrary non-signalling provers.

Thus, it is natural to ask if there is a commitment scheme whose binding

property require only the assumption that the provers can not transmit any information to one another, i.e., that it is truly based on the non-signaling assumption only.

We answer this question in the negative. If a commitment scheme is perfectly hiding, then it is possible for non-signaling adversaries to perfectly emulate the behavior of the honest provers, thus breaking the binding property. If it is close to perfectly hiding, they can act in a way that is hard to distinguish from the behavior of honest provers, and thus, cheating provers can succeed with near-certainty.

We also show a positive result: there exists a *three*-prover bit commitment scheme that is perfectly hiding and binding with a strong parameter. We prove this result by describing a three-prover commitment scheme with these properties: The first two provers execute the $\mathcal{CHSH}^q$ protocol with the verifier. The third prover has to send the same output as the second one to the verifier.

6.2 Bipartite Systems and Two-Prover Commitments

6.2.1 One-Round Bipartite Systems

Informally, a *bipartite system* consists of two subsystem, which we refer to as the left and the right subsystem. Upon input a to the left and input a' to the right subsystem, the left subsystem outputs x and the right subsystem outputs x' (see Fig. 6.1, left). Formally, the behavior of such a system is given by a conditional distribution $q(x, x'|a, a')$, with the interpretation that given input (a, a'), the system outputs a specific pair (x, x') with probability $q(x, x'|a, a')$. Note that we leave the sets $\mathcal{A}, \mathcal{A}', \mathcal{X}$ and $\mathcal{X}'$, from which a, a', x and x' are respectively sampled, implicit.

If we do not put any restriction upon the system, then *any* conditional distribution $q(x, x'|a, a')$ is eligible, i.e., describes a bipartite system. However, we are interested in systems where the two subsystems cannot communicate with each other. How exactly this requirement restricts $q(x, x'|a, a')$ depends on the available "resources". For instance, if the two subsystems are deterministic, i.e., compute x and x' as *deterministic* functions of a and a' respectively, then this restricts $q(x, x'|a, a')$ to be of the form $q(x, x'|a, a') = \delta(x|a) \cdot \delta(x'|a')$ for conditional Dirac distributions $\delta(x|a)$ and $\delta(x'|a')$. If in addition to allowing them to compute deterministic functions, we give the two subsystem *shared randomness*, then $q(x, x'|a, a')$ may be of the form

$$q(x, x'|a, a') = \sum_r p(r) \cdot \delta(x|a, r) \cdot \delta(x'|a', r)$$

for a distribution $p(r)$ and conditional Dirac distributions $\delta(x|a, r)$ and $\delta(x'|a', r)$. Such a system is called *classical* or *local*. Interestingly, this is not the end of

the story. By the laws of *quantum mechanics*, if the two subsystems share an entangled quantum state and obtain x and x' without communication as the result of local measurements that may depend on a and a', respectively, then this gives rise to conditional distributions $q(x, x'|a, a')$ of the form

$$q(x, x'|a, a') = \left\langle \psi \right| \left(E_x^a \otimes F_{x'}^{a'} \right) \left| \psi \right\rangle ,$$

where $|\psi\rangle$ is a quantum state and $\{E_x^a\}_x$ and $\{F_{x'}^{a'}\}_{x'}$ are so-called POVMs.[1] This is typically referred to as a *violation of Bell inequalities* [Bel64], and is nicely captured by the notion of *non-local games*. A famous example is the so-called CHSH-game [CHSH69], which is closely connected to the example two-prover commitment scheme from the introduction, and which shows that the variant considered in [CSST11] is insecure against quantum attacks.

The largest possible class of bipartite systems that is compatible with the requirement that the two subsystem do not communicate, but otherwise does not assume anything on the available resources and/or the underlying physical theory, are the so-called *non-signaling* systems, defined as follows.

Remark 6.1. *By convention, we write $p(x|a, b) = p(x|a)$ to express that $p(x|a, b)$ does not depend on b, i.e., that $p(x|a, b_1) = p(x|a, b_2)$ for all b_1 and b_2, and as such $p(x|a)$* is *well defined and equals $p(x|a, b)$.*

Definition 6.2. *A conditional distribution $q(x, x'|a, a')$ is called a* non-signaling (one-round) bipartite system *if it satisfies*

$$q(x|a, a') = q(x|a) \qquad \text{(NS)}$$

as well as with the roles of the primed and unprimed variables exchanged, i.e.,

$$q(x'|a, a') = q(x'|a') \qquad \text{(NS}'\text{)}$$

We emphasize that this is the *minimal* necessary condition for the requirement that the two subsystems do not communicate. Indeed, if e.g. $q(x|a, a'_1) \neq q(x|a, a'_2)$, i.e., if the input-output behavior of the left subsystem depends on the input to the right subsystem, then the system can be used to communicate by giving input a'_1 or a'_2 to the right subsystem, and observing the input-output behavior of the left subsystem. Thus, in such a system, communication does take place.

The non-signaling requirement for a bipartite system is — conceptually and formally — equivalent to requiring that the two subsystems can (in principle) be queried *in any order*. Conceptually, it holds because the left subsystem should be able to deliver its outputs *before* the right subsystem has received any input if and only if the output does not depend on the right subsystem's input (which means that no information is communicated from right to left),

[1] A POVM is essentially a measurement where only the measurement outcome is recorded and the post-measurement state is ignored.

and similarly the other way round. And, formally, we see that the non-signaling requirement from Definition 6.2 is equivalent to asking that $q(x, x'|a, a')$ can be written as

$$q(x, x'|a, a') = q(x|a) \cdot q(x'|x, a, a') \quad \text{and} \quad q(x, x'|a, a') = q(x'|a') \cdot q(x|x', a, a')$$

for some respective conditional distributions $q(x|a)$ and $q(x'|a')$. This characterization is a convenient way to "test" whether a given bipartite system is non-signaling.

Clearly, all classical systems are non-signaling. Also, any quantum system is non-signaling.[2] But there are non-signaling systems that are not quantum (and thus in particular not classical). The typical example is the *NL-box* (nonlocal box; also known as *PR-box*) [PR94], which, upon input bits a and a' outputs *random* output bits x and x' subject to

$$x \oplus x' = a \cdot a' \,.$$

This system is indeed non-signaling, as it can be queried in any order: submit a to the left subsystem to obtain a uniformly random x, and then submit a' to the right subsystem to obtain $x' := x \oplus a \cdot b$, and correspondingly the other way round.

6.2.2 Two-Round Systems

We now consider bipartite systems as discussed above, but where one can interact with the two subsystems multiple times. We restrict to two rounds: after having input a to the left subsystem and obtained x as output, one can now input b into the left subsystem and obtain output y, and similarly with the right subsystem (see Fig. 6.1, right). In such a two-round setting, the non-signaling condition needs to be paired with *causality*, which captures that the output of the first round does not depend on the input that will be given in the second round.

Definition 6.3. *A conditional distribution $q(x, x', y, y'|a, a', b, b')$ is called a* non-signaling two-round bipartite system *if it satisfies the following two causality constraints*

$$q(x, x'|a, a', b, b') = q(x, x'|a, a') \qquad \text{(C1)}$$

$$\textit{and} \quad q(x'|x, y, a, a', b, b') = q(x'|x, y, a, a', b) \qquad \text{(C2)}$$

and the following two non-signaling constraints

$$q(x, y|a, a', b, b') = q(x, y|a, b) \qquad \text{(NS1)}$$

$$\textit{and} \quad q(y|x, x', a, a', b, b') = q(y|x, x', a, a', b) \qquad \text{(NS2)}$$

[2] Indeed, the two parts of an entangled quantum state can be measured in any order, and the outcome of the first measurement does not depend on how the other part is going to be measured.

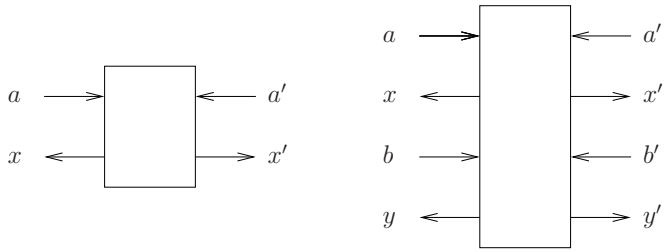

Figure 6.1: A one-round (left) and two-round (right) bipartite system.

as well as with the roles of the primed and unprimed variables exchanged.

(C1) captures causality of the overall system, i.e., when considering the left and the right system as one "big" multi-round system. (C2) captures that no matter what interaction there is with the left system, the right system still satisfies causality. Similarly, (NS1) captures that the left and the right system are non-signaling over both rounds, and (NS2) captures that no matter what interaction there was in the first round, the left and the right system remain non-signaling in the second round.

It is rather clear that these are *necessary* conditions; we argue that they are *sufficient* to capture a non-signaling two-round system in the following section.

6.2.3 Capturing the Non-Signaling Property

To see that Definition 6.3 is not only necessary but also sufficient to capture the non-signaling constraint, consider a two-round bipartite system that conforms to Definition 6.3. We show that the two subsystems can be queried *in any order* without altering the output distribution, as long as the order of rounds for each subsystem individually is respected. Thus, it is impossible to obtain information about the right side of the system by observing only the behaviour on the left side (and vice versa), which shows that Definition 6.3 is indeed sufficient. First, we point out the following.

Remark 6.4. *(C1) and (NS1) together imply that* $q(x|a,b)$ *and* $q(x|a,a')$ *are well-defined and satisfy*

$$q(x|a,b) = q(x|a) \quad \textit{(C3)} \qquad \textit{and} \qquad q(x|a,a') = q(x|a) \qquad \textit{(NS3)}\,.$$

This follows from Lemma 6.5 below.

Lemma 6.5. *Any conditional distribution* $q(x|a,b,c,d)$ *such that* $q(x|a,b,c,d) = q(x|a,b)$ *and* $q(x|a,b,c,d) = q(x|a,c)$*, must also satisfy* $q(x|a,b,c,d) = q(x|a)$*.*

Proof. Recall that, by convention, $q(x|a,b,c,d) = q(x|a,b)$ means $q(x|a,b,c,d) = q(x|a,b,c',d')$ for all x,a,b,c,c',d,d', and similarly for $q(x|a,b,c,d) = q(x|a,c)$.

As such, for arbitrary x, a, b, b', c, c', d, d' it holds that

$$q(x|a,b,c,d) = q(x|a,b,c',d') = q(x|a,b',c',d')$$

and thus $q(x|a,b,c,d) = q(x|a)$. □

If $q(x,x',y,y'|a,a',b,b')$ is a non-signaling two-round bipartite system, it can be written as

$$\begin{aligned} q(x,x',y,y'|a,a',b,b') &= q(x,y|a,b) \cdot q(x',y'|x,y,a,a',b,b') \\ &= q(x|a) \cdot q(y|x,a,b) \cdot q(x'|x,y,a,a',b) \cdot q(y'|x,y,a,a',b,b') \end{aligned}$$

where the first equality uses (NS1), and the second uses (C3) and (C2), and as

$$\begin{aligned} & q(x,x',y,y'|a,a',b,b') \\ = \; & q(x,x'|a,a') \cdot q(y,y'|x,x',a,a',b,b') \\ = \; & q(x|a) \cdot q(x'|x,a,a') \cdot q(y|x,x',a,a',b) \cdot q(y'|x,x',y,a,a',b,b') \end{aligned}$$

where the first equality uses (C1), and the second uses (NS3) and (NS2), and the second equality can also be replaced by

$$= q(x|a) \cdot q(x'|x,a,a') \cdot q(y'|x,x',a,a',b') \cdot q(y|x,x',y,a,a',b,b') \,.$$

And, similarly, with the roles of the primed and unprimed variables exchanged. This shows that the two subsystems can be queried in any order. For instance, one can first query the left subsystem to get x on input a, distributed according to $q(x|a)$, and then y on input b, distributed according to $q(y|x,a,b)$, and then then one can query the right subsystem twice to get x' and y', distributed according to $q(x'|x,y,a,a',b)$ and $q(y'|x,y,a,a',b,b')$, respectively.[3] Or, one can first query the left subsystem once to obtain x, then query the right subsystem to obtain x' etc. It is straightforward to verify that all six eligible orderings are possible.

6.2.4 Two-Prover Commitments

In this section, we consider commitment schemes with a one-round commit and opening phase not as interactive algorithms but as abstract bipartite systems. That is, we redefine them as follows:

Definition 6.6. *A* single-round two-prover bit-commitment scheme $\mathcal{S}$ *consists of a probability distribution* $p(a,a')$*, conditional distributions* $p_0(x,x',y,y'|a,a')$

[3] Note that in order to sample, say, x' according to $q(x'|x,y,a,a',b)$, it seems like that the right subsystem needs to know a, x etc., i.e., that communication is necessary, contradicting the non-signaling requirement. However, this reasoning merely shows that in general, such a non-signaling system is not classical.

and $p_1(x, x', y, y'|a, a')$*, and a function* $\mathsf{Extr}(c, y, y')$ *with range* $\{0, 1, \perp\}$*, where* c *is the commitment, i.e.,* $c = (a, a', x, x')$.[4]
We say that $\mathcal{S}$ *is classical/quantum/non-signaling if* $p_b(x, x', y, y'|a, a')$ *for* $b = 0, 1$ *are both classical/quantum/non-signaling when parsed as bipartite one-round systems* $p_b((x, y), (x', y')|a, a')$*. By default, any two-prover commitment scheme* $\mathcal{S}$ *is assumed to be non-signaling.*

Formulating it a bit more algorithmically, V samples messages a and a' for the two provers according to the distribution $p(a, a')$. In order to commit and open to a bit b, the honest provers input a and a' into a bipartite system described by the distribution $p_b(x, x', y, y'|a, a')$. This system then produces the respective messages x and x' that P and Q send in the commit phase, and the messages y and y' that they send in the opening phase. The verifier V computes the function $\mathsf{Extr}(c, y, y')$ to determine to which bit the provers opened, or if they failed to open to any bit.

In this formalism, the completeness property can be restated as follows:

Definition 6.7. *A commitment scheme* $\mathcal{S}$ *is* γ-complete *if* $p_b(\mathsf{Extr}(c, y, y') \neq b) \leq \gamma$ *for all* $b \in \{0, 1\}$*,*

Writing $p(x_b, x'_b, y_b, y'_b|a, a')$ for $p_b(x, x', y, y'|a, a')$, we can restate the definition of the hiding property as follows:

Definition 6.8. $\mathcal{S}$ *is* δ-hiding *if* $d\big(p(x_0, x'_0|a, a'), p(x_1, x'_1|a, a')\big) \leq \delta$ *for all* a, a'*. If* $\mathcal{S}$ *is* 0*-hiding, we also say it is* perfectly hiding.

The definition for the binding property that we use here is essentially the sum-binding definition (Definition 2.14), restated in the form of the following game between the (honest) verifier V and the adversarial provers P, Q.

1. The commit phase is executed: V samples a and a' according to $p(a, a')$, and sends a to P and a' to Q, upon which P and Q send x and x' back to V, respectively.
2. V sends a bit $b \in \{0, 1\}$ to P and Q.
3. P and Q try to open the commitment to b: they prepare y and y' and send them to V.
4. The provers win if $b = \mathsf{Extr}(c, y, y')$.

We emphasize that even though in the actual binding game above, *the same* bit b is given to the two provers, we require that the response of the provers is well determined by their strategy even in the case that they receive different bits. Of course, if the provers are allowed to communicate, they are

[4] Recall that we assume without loss of generality that the commitment is the entire communication during the commit phase.

able to detect when they receive different bits b and b' and could reply with, e.g., $y = y' = \perp$ in that case. However, if we restrict to non-signaling provers, we assume that it is *physically* impossible for them to communicate with each other and distinguish the case of $b = b'$ from $b \neq b'$.

A *non-signaling strategy* for dishonest provers is described by a non-signaling bipartite system $q(x, x', y, y'|a, a', b, b')$ as specified in Definition 6.3. Together with the distribution $p(a, a')$ and the bit b sent by the verifier, this system defines the distributions

$$q_b(x, x', y, y') = \sum_{a,a'} p(a, a')q(x, x', y, y'|a, a', b, b)$$

for $b \in \{0, 1\}$. Writing win_b for the event that $b = \mathsf{Extr}(c, y, y')$, the binding property requires a bound on the sum $q_0(\mathsf{win}_0) + q_1(\mathsf{win}_1)$.

Definition 6.9. *A two-prover commitment scheme $\mathcal{S}$ is* ε-binding *(against non-signaling attacks) if it holds for any non-signaling two-round bipartite system $q(x, x', y, y'|a, a', b, b')$ that $q_0(\mathsf{win}_0) + q_1(\mathsf{win}_1) \leq \varepsilon$.*

In other words, a scheme is ε-binding if in the above game the dishonest provers win with probability at most $1/2 + \varepsilon$ when $b \in \{0, 1\}$ is selected uniformly at random.

If a commitment scheme is binding (for a small ε) in the sense of Definition 6.9, then for any strategy q for P and Q, they can just as well *honestly* commit to a bit $\hat{b}$, where $\hat{b}$ is set to 0 with probability $p_0 = q_0(\mathsf{win}_0)$ and to 1 with probability $p_1 = 1 - p_0 \approx q_1(\mathsf{win}_1)$, and they will have essentially the same respective success probabilities in opening the commitment to $b = 0$ and to $b = 1$.

6.3 Impossibility of Two-Prover Commitments

In this section, we show impossibility of secure single-round two-prover commitments against arbitrary non-signaling attacks. We start with the analysis of a restricted class of schemes which are easier to understand and for which we obtained stronger results.

6.3.1 Simple Schemes

We first consider a special, yet natural, class of schemes. We call a two-prover commitment scheme $\mathcal{S}$ *simple* if it has the same communication pattern as the scheme described in the introduction. More formally, it is called simple if a', x' and y are "empty" (or fixed), i.e., if $\mathcal{S}$ is given by $p(a)$, $p_0(x, y'|a)$, $p(x, y'|a)$ and $\mathsf{Extr}(c, y')$ with $c = (a, x)$; to simplify notation, we then write y instead of y'. In other words, P is only involved in the commit phase, where, in order to commit to bit b, he outputs x upon input a, and Q is only involved in the

opening phase, where he outputs y. The non-signaling requirement for $\mathcal{S}$ then simplifies to $p_b(y|a) = p_b(y)$.

In case of such a simple two-prover commitment scheme $\mathcal{S}$, a non-signaling two-prover strategy reduces to a non-signaling *one-round* bipartite system as specified in Definition 6.2 (see Figure 6.2).

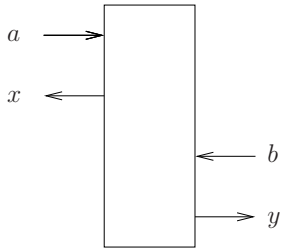

Figure 6.2: The adversaries' strategy $p(x, y|a, b)$ in case of a *simple* commitment scheme.

As a warm-up exercise, we first consider a simple two-prover commitment scheme that is *perfectly hiding* and *0-complete*. Recall that the perfect hiding property means that $p_0(x|a) = p_1(x|a)$ for any a. To show that such a scheme cannot be binding, we have to show that there exists a non-signaling one-round bipartite system $q(x, y|a, b)$ such that $q_0(\mathsf{win}_0) + q_1(\mathsf{win}_1)$ is significantly larger than 1. But this is actually trivial: we can simply set $q(x, y|a, b) := p_b(x, y|a)$. It then holds trivially that $q_b(x, y) = p_b(x, y)$, so the dishonest provers are as successful in opening the commitment as the honest provers in opening an honestly prepared commitment. Thus, the binding property is broken as badly as it can get. The only thing that needs to be verified is that $q(x, y|a, b)$ is actually non-signaling, i.e., that $q(x|a, b) = q(x|a)$ and $q(y|a, b) = q(y|b)$.

To see that the latter holds, note that $q(y|a, b) = p_b(y|a)$, and because $\mathcal{S}$ is non-signaling we have that $p_b(y|a) = p_b(y)$, i.e., does not depend on a. Thus, the same holds for $q(y|a, b)$ and we have $q(y|a, b) = q(y|b)$. The former condition follows from the (perfect) hiding property: $q(x|a, b) = p_b(x|a) = p_{b'}(x|a) = q(x|a, b')$ for arbitrary $b, b' \in \{0, 1\}$, and thus $q(x|a, b) = q(x|a)$.

Below, we show how to extend this result to non-perfectly-binding simple schemes. In this case, we cannot simply set $q(x, y|a, b) := p_b(x, y|a)$, because such a q would not be non-signaling anymore — it would merely be "almost non-signaling". Instead, we have to find a strategy $q(x, y|a, b)$ that is (perfectly) non-signaling and close to $p_b(x, y|a)$; we will find such a strategy with the help of Lemma 2.4. In Section 6.3.2, we will then consider general schemes where *both* provers interact with the verifier in *both* phases. In this general case, further complications arise.

Theorem 6.10. *Consider a simple two-prover commitment scheme $\mathcal{S}$ that is*

δ-hiding. Then, there exists a non-signaling strategy $q(x, y|a, b)$ such that

$$q_0(\mathsf{win}_0) = p_0(\mathsf{win}_0) \quad \textit{and} \quad q_1(\mathsf{win}_1) > p_1(\mathsf{win}_1) - \delta\,.$$

If $\mathcal{S}$ is 0-complete, it follows that

$$q_0(\mathsf{win}_0) + q_1(\mathsf{win}_1) > 1 + (1 - \delta)$$

and thus it cannot be ε-binding for $\varepsilon \leq (1 - \delta)/2$.

Proof. Recall that $\mathcal{S}$ is given by $p(a)$, $p_b(x, y|a)$ and $\mathsf{Extr}(c, y)$, and we write $p(x_b, y_b|a)$ instead of $p_b(x, y|a)$. Because $\mathcal{S}$ is δ-hiding, it holds that

$$d\big(p(x_0|a), p(x_1|a)\big) \leq \delta$$

for any fixed a. Using Lemma 2.4 for every a, we can glue together $p(x_0, y_0|a)$ and $p(x_1, y_1|a)$ along x_0 and x_1 to obtain a distribution $p(x_0, x_1, y_0, y_1|a)$ such that $p(x_0 \neq x_1|a) \leq \delta$, and in particular $d\big(p(x_0, y_1|a), p(x_1, y_1|a)\big) \leq \delta$.

We define a strategy q for the dishonest provers by setting $q(x, y|a, b) := p(x_0, y_b|a)$ (see Fig. 6.3). First, we show that q is non-signaling. Indeed, we have $q(x|a, b) = p(x_0|a)$ for any b, so $q(x|a, b) = q(x|a)$, and we have $q(y|a, b) = p(y_b|a) = p(y_b)$ for any a, and thus $q(y|a, b) = q(y|b)$.

As for the winning probability, for $b = 0$ we have $q(x, y|a, 0) = p(x_0, y_0|a)$ and as such $q_0(\mathsf{win}_b)$ equals $p_0(\mathsf{win}_b)$. For $b = 1$, we have

$$d\big(q(x, y|a, 1), p(x_1, y_1|a)\big) = d\big(p(x_0, y_1|a), p(x_1, y_1|a)\big) \leq \delta$$

and since the statistical distance does not increase under data processing, it follows that $q_1(\mathsf{win}_1)$ and $p_1(\mathsf{win}_1)$ are δ-close; this proves the claim. □

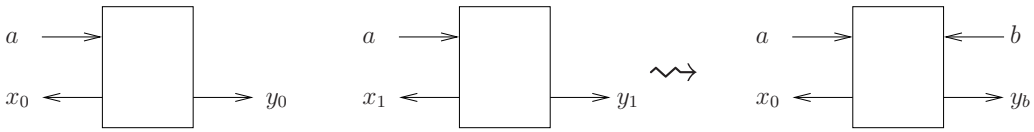

Figure 6.3: Defining the strategy q by gluing together $p(x_0, y_0|a)$ and $p(x_1, y_1|a)$.

The bound on the binding property in Theorem 6.10 is tight, as the following theorem shows.

Theorem 6.11. *For all $\delta \in \mathbb{Q}$ such that $0 < \delta \leq 1$ there exists a classical simple two-prover commitment scheme that is perfectly sound, δ-hiding and $(1 - \delta)/2$-binding against non-signaling adversaries.*

Proof. We construct a scheme where the first prover reveals the bit b right at the beginning with probability δ. For simplicity, we first assume that $\delta = 1/n$ for some integer $n \geq 1$ and then indicate how to extend the proof to arbitrary rational numbers.

The scheme works as follows. Let $[n] = \{0, \ldots, n-1\}$. The shared randomness of the provers is $r \in [n]$ selected uniformly at random. The verifier selects $a \in [n]$ uniformly at random and sends it to prover P. If $a = r$ then P reveals $x := b$ to the verifier. Otherwise, he sends back $x := \perp$. In the opening phase, Q sends r to the verifier. The verifier accepts if and only if P revealed b or the output y of Q satisfies $y \in [n]$ and $y \neq a$.

It is clear that this scheme is sound and δ-hiding. Now consider dishonest provers that follow some non-signaling strategy $q(x, y|a, b)$. This then defines $q_b(a, x, y) = p(a) \cdot q(x, y|a, b)$ with $p(a) = 1/n$, and it holds that $q_b(\mathsf{win}_b) = q_b(x{=}b) + q_b(x{=}\perp, y \neq a)$. Since $q(y|a, b) = q(y|b)$, we have

$$q_b(y{\neq}a) = \sum_{\substack{a,y \\ a \neq y}} q_b(a, y) = \sum_{\substack{a,y \\ a \neq y}} p(a)\, q_b(y) = \sum_y \frac{n-1}{n} q_b(y) = 1 - \delta\,.$$

Therefore, using that $q(x|a, b) = q(x|a)$ and hence $q_0(x) = q_1(x)$, we calculate

$$\begin{aligned} q_0(\mathsf{win}_0) + q_1(\mathsf{win}_1) &= q_0(x{=}0) + q_0(x{=}\perp, y \neq a) + q_1(x{=}1) + q_1(x{=}\perp, y \neq a) \\ &\leq q_0(x{=}0) + q_1(x{=}1) + q_0(x{=}\perp) + q_1(y{\neq}a) \\ &= 1 + (1 - \delta)\,. \end{aligned}$$

We now adapt this argument to $\delta = m/n$, where m and n are integers such that $0 < m \leq n$. For every $a \in [n]$, we define a subset S_a of $[n]$ as

$$S_a = \{a + i \bmod n \mid i \in \{0, \ldots, m-1\}\}.$$

We adapt our scheme by replacing the condition $r = a$ with $r \in S_a$. Clearly, the scheme is still sound. Since every S_a has exactly m elements, the scheme is δ-hiding: the probability that the first prover reveals b is $m/n = \delta$; otherwise, he does not give any information about b. The proof that the scheme is $(1-\delta)/2$-binding goes through as before if we can show that $q(y \notin S_a|a, b) = 1 - \delta$ for any non-signaling strategy q. Indeed, for every $y \in [n]$, there are exactly m values for a such that $y \in S_a$. Since $a \in [n]$ is selected randomly and $q(y|a, b)$ is independent of a, we have $q(y \notin S_a|a, b) = 1 - m/n = 1 - \delta$. □

6.3.2 Arbitrary Schemes

We now remove the restriction on the scheme to be simple. As before, we first consider the case of a perfectly hiding scheme.

Theorem 6.12. *Let $\mathcal{S}$ be a single-round two-prover commitment scheme. If $\mathcal{S}$ is perfectly hiding, then there exists a non-signaling two-prover strategy $q(x, x', y, y'|a, a', b, b')$ such that $q_b(\mathsf{win}_b) = p_b(\mathsf{win}_b)$ for $b \in \{0, 1\}$.*

Proof. $\mathcal{S}$ being perfectly hiding means that $d(p(x_0, x_0'|a, a'), p(x_1, x_1'|a, a')) = 0$ for all a and a'. Gluing together the distributions $p(x_0, x_0', y_0, y_0'|a, a')$ and $p(x_1, x_1', y_1, y_1'|a, a')$ along (x_0, x_0') and (x_1, x_1') for every (a, a'), we obtain a distribution $p(x_0, x_0', x_1, x_1', y_0, y_0', y_1, y_1'|a, a')$ with the correct marginals and $p((x_0, x_0') \neq (x_1, x_1')|a, a') = 0$. That is, we have $x_0 = x_1$ and $x_0' = x_1'$ with certainty. We now define a strategy for dishonest provers as (Figure 6.4)

$$q(x, x', y, y'|a, a', b, b') := p(x_0, x_0', y_b, y_{b'}'|a, a') .$$

Since $p(x_0, x_0', y_b, y_b'|a, a') = p(x_b, x_b', y_b, y_b'|a, a')$, it holds that $q_b(\mathsf{win}_b) = p_b(\mathsf{win}_b)$. It remains to show that this distribution satisfies the non-signaling and causality constraints (C1) up to (NS2) of Definition 6.3. This is done below.

- For (C1), note that summing up over y and y' yields $q(x, x'|a, a', b, b') = p(x_0, x_0'|a, a')$, which indeed does not depend on b and b'.
- For (NS1), note that $q(x, y|a, a', b, b') = p(x_0, y_b|a, a') = p(x_b, y_b|a, a') = p(x_b, y_b|a)$, where the last equality holds by the non-signaling property of $p(x_b, y_b|a, a')$.
- For (C2), first note that

  $$q(x, x', y|a, a', b, b') = p(x_0, x_0', y_b|a, a') \tag{6.1}$$

  which does not depend on b'. We then see that (C2) holds by dividing by $q(x, y|a, a', b, b') = p(x_0, y_b|a, a')$.
- For (NS2), divide Equation (6.1) by $q(x, x'|a, a', b, b') = p(x_0, x_0'|a, a')$

The properties (C1) to (NS2) with the roles of the primed and unprimed variables exchanged follows from symmetry. This concludes the proof. □

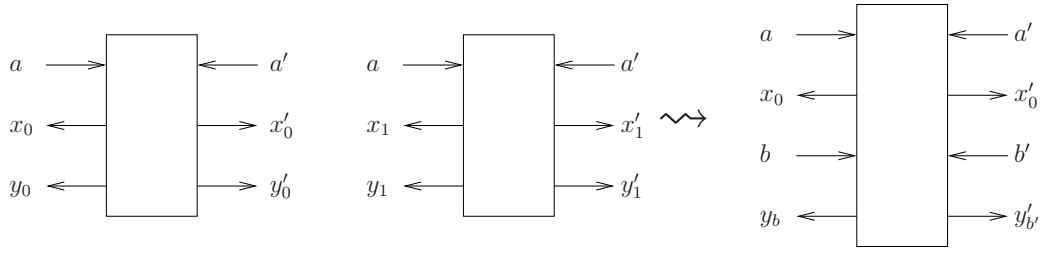

Figure 6.4: Defining the strategy q from $p(x_0, x_0', y_0, y_0'|a, a')$ and $p(x_1, x_1', y_1, y_1'|a, a')$ glued together.

The case of non-perfectly hiding schemes is more involved. At first glance, one might expect that by proceeding analogously to the proof of Theorem 6.12 – that is, gluing together $p(x_0, x_0', y_0, y_0'|a, a')$ and $p(x_1, x_1', y_1, y_1'|a, a')$ along (x_0, x_0') and (x_1, x_1') and defining q the same way – one can obtain a strategy q that succeeds with probability $1-\delta$ if the scheme is δ-hiding. Unfortunately, this approach fails because in order to show (NS1) we use that $p(x_0, y_1|a, a') = p(x_1, y_1|a, a')$ which in general does not hold for commitment schemes that are not perfectly hiding. As a consequence, our proof is more involved, and we have a constant-factor loss in the parameter.

Theorem 6.13. *Let $\mathcal{S}$ be a single-round two-prover commitment scheme and suppose that it is δ-hiding. Then there exists a non-signaling two-prover strategy $q(x, x', y, y'|a, a', b, b')$ such that*

$$q_0(\mathsf{win}_0) = p_0(\mathsf{win}_0) \quad \textit{and} \quad q_1(\mathsf{win}_1) \geq p_1(\mathsf{win}_1) - 5\delta\,.$$

Thus, if $\mathcal{S}$ is perfectly sound, it is at best $(1-5\delta)/2$-binding.

To prove this result, we use two technical lemmas. In the first one, we add the additional assumptions that $p(x_0|a, a') = p(x_1|a, a')$ and $p(x_0'|a, a') = p(x_1'|a, a')$. The second one shows that we can tweak an arbitrary scheme in such a way that these additional conditions hold. We give the proofs after Theorem 6.13.

Lemma 6.14. *Let $\mathcal{S}$ be a δ-hiding two-prover commitment scheme with the additional property that $p(x_0|a, a') = p(x_1|a, a')$ and $p(x_0'|a, a') = p(x_1'|a, a')$. Then, there exists a non-signaling $p'(x_1, x_1', y_1, y_1'|a, a')$ such that*

$$d\big(p'(x_1, x_1', y_1, y_1'|a, a'), p(x_1, x_1', y_1, y_1'|a, a')\big) \leq \delta$$

and $p'(x_1, x_1'|a, a') = p(x_0, x_0'|a, a')$.

As usual, the non-signaling requirement on $p'(x_1, x_1', y_1, y_1'|a, a')$ is to be understood as $p'(x_1, y_1|a, a') = p'(x_1, y_1|a)$ and $p'(x_1', y_1'|a, a') = p'(x_1', y_1'|a')$.

Lemma 6.15. *Let $\mathcal{S}$ be a δ-hiding two-prover commitment scheme. Then, there exists a non-signaling $\tilde{p}(x_1, x_1', y_1, y_1'|a, a')$ such that*

$$d\big(\tilde{p}(x_1, x_1', y_1, y_1'|a, a'), p(x_1, x_1', y_1, y_1'|a, a')\big) \leq 2\delta$$

which has the property that $\tilde{p}(x_1|a, a') = p(x_0|a, a')$ and $\tilde{p}(x_1'|a, a') = p(x_0'|a, a')$.

With these two lemmas, Theorem 6.13 is easy to prove.

Theorem 6.13. We start with a δ-hiding non-signaling bit-commitment scheme $\mathcal{S}$. We apply Lemma 6.15 and obtain $\tilde{p}(x_1, x_1', y_1, y_1'|a, a')$ that is 2δ-close to $p(x_1, x_1', y_1, y_1'|a, a')$ and satisfies $\tilde{p}(x_1|a, a') = p(x_0|a, a')$ and $\tilde{p}(x_1'|a, a') = p(x_0'|a, a')$. Furthermore, by triangle inequality

$$d\big(\tilde{p}(x_1, x_1'|a, a'), p(x_0, x_0'|a, a')\big) \leq 3\delta\,.$$

Thus, replacing $p(x_1, x_1', y_1, y_1|a, a')$ by $\tilde{p}(x_1, x_1', y_1, y_1'|a, a')$ gives us a 3δ-hiding two-prover commitment scheme that satisfies the extra assumption in Lemma 6.14. As a result, we obtain a distribution $p'(x_1, x_1', y_1, y_1'|a, a')$ that is 3δ-close to $\tilde{p}(x_1, x_1', y_1, y_1'|a, a')$, and thus 5δ-close to $p(x_1, x_1', y_1, y_1'|a, a')$, with the property that $p'(x_1, x_1'|a, a') = p(x_0, x_0'|a, a')$. Therefore, replacing $\tilde{p}(x_1, x_1', y_1, y_1'|a, a')$ by $p'(x_1, x_1', y_1, y_1'|a, a')$ gives us a *perfectly-hiding* two-prover commitment scheme, to which we can apply Theorem 6.12. As a consequence, there exists a non-signaling strategy $q(x, x', y, y'|a, a')$ with $q_0(\mathsf{win}_0) = p_0(\mathsf{win}_0)$ and $q_1(\mathsf{win}_1) \geq p_1(\mathsf{win}_1) - 5\delta$, as claimed. □

Remark 6.16. *If $\mathcal{S}$ already satisfies $p(x_0|a, a') = p(x_1|a, a')$ and $p(x_0'|a, a') = p(x_1'|a, a')$, we can apply Lemma 6.14 right away and thus get a strategy q with $q_0(\mathsf{win}_0) = p_0(\mathsf{win}_0)$ and $q_1(\mathsf{win}_1) = p_1(\mathsf{win}_1) - \delta$. Thus, with this additional condition, we still obtain a tight bound as in Theorem 6.10.*

We now prove the two lemmas:

Proof of Lemma 6.14. For arbitrary a and a', we glue together the distributions $p(x_0, x_0', y_0, y_0'|a, a')$ and $p(x_1, x_1', y_1, y_1'|a, a')$ to obtain a joint distribution $p(x_0, x_0', x_1, x_1', y_0, y_0', y_1, y_1'|a, a')$ such that

$$p\big((x_0, x_0') \neq (x_1, x_1')|a, a'\big) \leq \varepsilon\,,$$

and thus $d(p(x_0, x_0', y_1, y_1'|a, a'), p(x_1, x_1', y_1, y_1'|a, a')) \leq \varepsilon$. Let Λ be the event that both $x_0 = x_1$ and $x_0' = x_1'$. We define $p'(x_1, x_1', y_1, y_1'|a, a')$ as follows, where x_0 is associated with x_1 and x_0' with x_1':

$$\begin{aligned} p'(x_1, x_1', y_1, y_1'|a, a') &:= p(\Lambda, x_0, x_0'|a, a') \cdot p(y_1, y_1'|\Lambda, x_1, x_1', a, a') \\ &\quad + p(\overline{\Lambda}, x_0, x_0'|a, a') \cdot r(y_1|x_0, a, a') \cdot r(y_1'|x_0', a, a') \\ &= p(\Lambda, x_1, x_1', y_1, y_1'|a, a') \\ &\quad + p(\overline{\Lambda}, x_0, x_0'|a, a') \cdot r(y_1|x_0, a, a') \cdot r(y_1'|x_0', a, a') \end{aligned}$$

where $r(y_1|x_0, a, a')$ and $r(y_1'|x_0', a, a')$ are to be defined later, and the last equality holds by definition of Λ.[5]

The claim about the closeness to $p(x_1, x_1', y_1, y_1'|a, a')$ follows from the fact that $p(\overline{\Lambda}|a, a') \leq \varepsilon$. Furthermore, we have $p'(x_1, x_1'|a, a') = p(\Lambda, x_0, x_0'|a, a') + p(\overline{\Lambda}, x_0, x_0'|a, a') = p(x_0, x_0'|a, a')$ as claimed.

It remains to show that we can achieve p' to be non-signaling. For that, we simply define $r(y_1|x_0, a, a')$, and similarly $r(y_1'|x_0', a, a')$, in such a way

[5] Algorithmically, the distribution p' should be understood as follows. First, x_0, x_0', x_1 and x_1' are sampled according to the glued-together distribution p. Then, if the event Λ occurred (i.e. $x_0 = x_1$ and $x_0' = x_1'$), y_1 and y_1' are sampled according to the corresponding conditional distribution; otherwise, they are chosen *independently* according to distributions that depend only on x_0 and x_0', respectively.

that $p'(x_1, y_1|a, a') = p(x_1, y_1|a, a')$; this does the job since $p(x_1, y_1|a, a') = p(x_1, y_1|a)$, and as such $p'(x_1, y_1|a, a') = p'(x_1, y_1|a)$. Note that

$$p'(x_1, y_1|a, a') = p(\Lambda, x_1, y_1|a, a') + p(\overline{\Lambda}, x_0|a, a') \cdot r(y_1|x_0, a, a') \,. \tag{6.2}$$

Thus, we set

$$r(y_1|x_0, a, a') := \frac{p(x_1, y_1|a, a') - p(\Lambda, x_1, y_1|a, a')}{p(\overline{\Lambda}, x_0|a, a')} = \frac{p(\overline{\Lambda}, x_1, y_1|a, a')}{p(\overline{\Lambda}, x_0|a, a')}$$

It remains to show that $r(y_1|x_0, a, a')$ as defined is indeed a probability distribution, and that things work out also in case $p(\overline{\Lambda}, x_0|a, a') = 0$.

In the latter case, we have $p'(x_1, y_1|a, a') = p(\Lambda, x_1, y_1|a, a')$, independent of the choice of r; thus, it remains to show that $p(\Lambda, x_1, y_1|a, a') = p(x_1, y_1|a, a')$. For that, we observe that $p(\Lambda, x_1|a, a') = p(\Lambda, x_0|a, a') = p(x_0|a, a') = p(x_1|a, a')$, where the first equality is due to the definition of Λ and the last holds by our additional assumption on Com. It follows that

$$\sum_{y_1} p(\Lambda, x_1, y_1|a, a') = p(\Lambda, x_1|a, a') = p(x_1|a, a') = \sum_{y_1} p(x_1, y_1|a, a')$$

and since $p(\Lambda, x_1, y_1|a, a') \leq p(x_1, y_1|a, a')$, it holds that $p(\Lambda, x_1, y_1|a, a') = p(x_1, y_1|a, a')$ as required.

Finally, to show that $r(y_1|x_0, a, a')$ is a probability distribution, we observe that $r(y_1|x_0, a, a') \geq 0$, and, summing over y_1 and using that $p(x_0|a, a') = p(x_1|a, a')$, we see that

$$\begin{aligned}\sum_{y_1} r(y_1|x_0, a, a') = \frac{p(x_1|a, a') - p(\Lambda, x_1|a, a')}{p(\overline{\Lambda}, x_0|a, a')} &= \frac{p(x_0|a, a') - p(\Lambda, x_0|a, a')}{p(\overline{\Lambda}, x_0|a, a')} \\ &= \frac{p(\overline{\Lambda}, x_0|a, a')}{p(\overline{\Lambda}, x_0|a, a')} \\ &= 1 \,.\end{aligned}$$

In the same way, it is possible to choose $r(y_1'|x_0', a, a')$ so that $p'(x_1', y_1'|a, a') = p(x_1', y_1'|a, a') = p(x_1', y_1'|a')$, using the assumption that $p(x_0'|a, a') = p(x_1'|a, a')$. This concludes the proof. □

Proof of Lemma 6.15. We begin by adjusting the distribution of x_1. By the hiding property of Com, $p(x_0, x_0'|a, a')$ and $p(x_1, x_1'|a, a')$ are ε-close, and thus in particular $d(p(x_0|a, a'), p(x_1|a, a')) \leq \varepsilon$. Gluing together the distributions $p(x_0|a, a')$ and $p(x_1, x_1', y_1, y_1'|a, a')$ along x_0 and x_1, we obtain a distribution $p(x_0, x_1, x_1', y_1, y_1'|a, a')$ such that

$$p'(x_1, x_1', y_1, y_1'|a, a') := p(x_0, x_1', y_1, y_1'|a, a')$$

satisfies $d\big(p'(x_1, x_1', y_1, y_1'|a, a'), p(x_1, x_1', y_1, y_1'|a, a')\big) \leq \varepsilon$ and also $p'(x_1|a, a') = p(x_0|a, a')$.

We show that p' is non-signaling. Since $p'(x_1', y_1'|a, a') = p(x_1', y_1'|a, a')$ and p is non-signaling, it follows that $p'(x_1', y_1'|a, a') = p'(x_1', y_1'|a')$. Showing that $p'(x_1, y_1|a, a') = p'(x_1, y_1|a)$ is equivalent to showing that $p(x_0, y_1|a, a') = p(x_0, y_1|a)$. By the observation in Remark 2.6, the marginal $p(x_0, x_1, y_1|a, a')$ is obtained by gluing together $p(x_0|a, a')$ and $p(x_1, y_1|a, a')$ along x_0 and x_1. Since Com is non-signaling, we have $p(x_0|a, a') = p(x_0|a)$ and $p(x_1, y_1|a, a') = p(x_1, y_1|a)$. It follows that $p(x_0, x_1, y_1|a, a') = p(x_0, x_1, y_1|a)$, and therefore that $p(x_0, y_1|a, a') = p(x_0, y_1|a)$.

In order to obtain $\tilde{p}$ as claimed, we repeat the above process. Note that the modification from p to p' did not change the distribution of x_1', y_1', i.e., $p'(x_1', y_1'|a, a') = p(x_1', y_1'|a, a')$, and in particular $d\big(p(x_0'|a, a'), p'(x_1'|a, a')\big) = d\big(p(x_0'|a, a'), p(x_1'|a, a')\big) \leq \varepsilon$. Therefore, exactly as above, we can now adjust the distribution of x_1' in p' and obtain a non-signaling $\tilde{p}(x_1, x_1', y_1, y_1'|a, a')$ that is ε-close to $p'(x_1, x_1', y_1, y_1'|a, a')$ and thus 2ε-close to $p(x_1, x_1', y_1, y_1'|a, a')$, and which satisfies $\tilde{p}(x_1'|a, a') = p(x_0'|a, a')$ and $\tilde{p}(x_1|a, a') = p'(x_1|a, a') = p(x_0|a, a')$, as claimed. □

6.3.3 Multi-Round Schemes

We briefly discuss a limited extension of our impossibility results for single-round schemes to schemes where during the commit phase, there is multi-round interaction between the verifier V and the two provers P and Q. We still assume the opening phase to be one-round; this is without loss of generality in case of *classical* two-prover commitment schemes (where the honest provers are restricted to be classical). In this setting, we have the following impossibility result, which is restricted to perfectly-hiding schemes.

Theorem 6.17. *Let $\mathcal{S}$ be a multi-round two-prover commitment scheme. If $\mathcal{S}$ is perfectly hiding, then there exists a non-signaling two-prover strategy that completely breaks the binding property, in the sense of Theorem 6.12.*

A formal proof of this statement requires a definition of n-round non-signaling bipartite systems for arbitrary n. Such a definition can be based on the intuition that it must be possible to query the left and right subsystem in any order. With this definition, the proof is a straightforward extension of the proof of Theorem 6.12: the non-signaling strategy is obtained by gluing together $p(\mathbf{x}_0, \mathbf{x}_0'|\mathbf{a}, \mathbf{a}')$ and $p(\mathbf{x}_1, \mathbf{x}_1'|\mathbf{a}, \mathbf{a}')$ along $(\mathbf{x}_0, \mathbf{x}_0')$ and $(\mathbf{x}_1, \mathbf{x}_1')$, and setting $q(\mathbf{x}, \mathbf{x}', y, y'|\mathbf{a}, \mathbf{a}', b, b') := p(\mathbf{x}_0, \mathbf{x}_0', y_b, y_{b'}'|\mathbf{a}, \mathbf{a}')$, where we use bold-face notation for the vectors that collect the messages sent during the multi-round commit phase: $\mathbf{a}$ collects all the messages sent by the verifier to the prover P, etc.

As far as we see, the proof of the non-perfect case, i.e. Theorem 6.13, does not generalize immediately to the multi-round case. As such, proving

the impossibility of *non-perfectly-hiding multi-round* two-prover commitment schemes remains an open problem.

6.4 Possibility of Three-Prover Commitments

It turns out that we can overcome the impossibility results by adding a third prover. We will describe a scheme that is perfectly sound, perfectly hiding and 2^{-n}-binding with communication complexity $O(n)$. We now define what it means for three provers to be non-signaling; since our scheme is similar to a simple scheme, we can simplify this somewhat. We consider distributions $q(x, y, z|a, b, c)$ where a and x are input and output of the first prover P, b and y are input and output of the second prover Q and c and z are input and output of the third prover R.

Definition 6.18. *A conditional distribution $q(x, y, z|a, b, c)$ is called a* non-signaling (one-round) tripartite system *if it satisfies*

$$q(x|a,b,c) = q(x|a)\ , \quad q(y|a,b,c) = q(y|b)\ , \quad q(z|a,b,c) = q(z|c)\ ,$$
$$q(x,y|a,b,c) = q(x,y|a,b)\ , \quad q(x,z|a,b,c) = q(x,z|a,c)$$
$$\text{and} \quad q(y,z|a,b,c) = q(y,z|b,c)\,.$$

In other words, for any way of viewing q as a bipartite system by dividing in- and outputs consistently into two groups, we get a non-signaling bipartite system. Actually, by means of Lemma 6.5, it is not hard to see that the first three requirements follow by the (union of the) latter three.

We restrict to *simple* schemes, where during the commit phase, only P is active, sending x upon receiving a from the verifier, and during the opening phase, only Q and R are active, sending y and z to the verifier, respectively.

Definition 6.19. *A* simple three-prover commitment scheme $\mathcal{S}$ *consists of a probability distribution $p(a)$, two distributions $p_0(x, y, z|a)$ and $p_1(x, y, z|a)$, and a function $\mathsf{Extr}(c, y, z)$ with range $\{0, 1, \perp\}$ where $c = (a, x)$.*
It is called classical/quantum/non-signaling if $p_b(x, y, z|a)$ is, when understood as a tripartite system $p_b(x, y, z|a, \emptyset, \emptyset)$ with two "empty" inputs.

Soundness and the hiding-property are defined in the obvious way. As for the binding property, for a simple three-prover commitment scheme $\mathcal{S}$ and a non-signaling strategy $q(x, y, z|a, b, c)$, let $q_b(a, x, y, z) := p(a)q(x, y, z|a, b, b)$. Like before, we define win_b as the event $b = \mathsf{Extr}(c, y, z)$. We say that $\mathcal{S}$ is ε-binding if

$$q_0(\mathsf{win}_0) + q_1(\mathsf{win}_1) \leq 1 + \varepsilon.$$

Theorem 6.20. *For every prime power q, there exists a classical simple three-prover commitment scheme that is perfectly sound, perfectly hiding and $q^{-1}/2$-binding. The verifier communicates $\lceil \log q \rceil$ bits to the first prover and receives the same number of bits from each prover.*

The scheme that achieves this is essentially the $\mathcal{CHSH}^q$ scheme, except that we add a third prover that imitates the actions of the second. To be more precise: The provers P, Q and R have as shared randomness a uniformly random $r\mathbb{F}_q$. The verifier V chooses a uniformly random $a \in \mathbb{F}_q$ and sends it to P. As commitment, P returns $x := r + a \cdot b$. To open the commitment to b, Q and R send $y := r$ and $z := r$ to V. The output of $\mathsf{Extr}((a, x), y, z)$ is defined as follows: if $y = z$, it is the smallest b such that $x - y = a \cdot b$, and if $y \neq z$, or no such b exists, it is $\perp$.

Before beginning with the formal proof that this scheme has the properties stated in our theorem, we give some intuition. Let a and x be the input and output of the dishonest first prover, P. To succeed, the second prover Q has to produce output $x + a \cdot b$ where b is the second prover's input and the third prover R has to produce $x + a \cdot c$ where c is the third prover's input. Our theorem implies that a strategy which always produces these outputs must be signaling. Why is that the case?

In the game that defines the binding-property, we always have $b = c$, but the dishonest provers must obey the non-signaling constraint even in the "impossible" case that $b \neq c$. Let us consider the difference between Q's output and R's output in the case that $b \neq c$: we get $(x + a \cdot b) - (x + a \cdot c) = a \cdot b - a \cdot c = \pm a$. But in the non-signaling setting, the joint distribution of Q's and R's output may not depend on a. Thus, the strategy we suggested does not satisfy the non-signaling constraint. Let us now prove the theorem.

Theorem 6.20. It is easy to see that the scheme is q^{-1}-complete, like $\mathcal{CHSH}^q$. For every fixed a and b, $p_b(x|a)$ is uniform, so the scheme is perfectly hiding. Now consider a non-signaling strategy q for dishonest provers. The provers succeed if and only if $y = z = a \cdot b - x$. Define $q(a, x, y, z|b, c) = p(a) \cdot q(x, y, z|a, b, c)$. The non-signaling property implies that

$$q(y = a \cdot b - x|a, b, c = 0) = q(y = x \oplus a \cdot b|a, b, c = 1) \quad \text{and} \tag{6.3}$$

$$q(z = a \cdot c - x|a, b = 0, c) = q(z = x \oplus a \cdot c|a, b = 1, c)\,. \tag{6.4}$$

It follows that

$$\begin{aligned}
&q_0(\mathsf{win}_0) + q_1(\mathsf{win}_1) \\
&\quad = q(y = a \cdot b - x, z = a \cdot c - x|b = 0, c = 0) \\
&\qquad\quad + q(y = a \cdot b - x, z = a \cdot c - x|b = 1, c = 1) \\
&\quad \leq q(y = a \cdot b - x|b = 0, c = 0) + q(z = a \cdot c - x|b = 1, c = 1) \\
&\quad = q(y = a \cdot b - x|b = 0, c = 1) + q(z = a \cdot c - x|b = 0, c = 1) \\
&\qquad \text{by Equations (6.3) and (6.4)} \\
&\quad \leq 1 + q(y = a \cdot b - x, z = a \cdot c - x|b = 0, c = 1) \text{ by Equation (2.2)}
\end{aligned}$$

It now remains to upper-bound $q(y = a \cdot b - x, z = a \cdot c - x|b = 0, c = 1)$. Since

$p(a)$ is uniform and $q(y, z|a, b, c)$ is independent of a, we have

$$q(y = a \cdot b - x, z = a \cdot c - x|b = 0, c = 1) \leq q(y \oplus z = a|b = 0, c = 1) = \frac{1}{q}$$

and thus our scheme is $q^{-1}/2$-binding. □

This result is reminiscent of a result by Masanes, Acin and Gisin [MAG06] where they show that if a non-signaling distribution $p(x, y|a, b)$ with $b \in \{0, 1\}$ is 2-shareable, then it is also local, i.e., it can be sampled using classical shared randomness.[6] Being 2-shareable means that there is a non-signaling distribution $p(x, y_1, y_2|a, b_1, b_2)$ such that $p(x, y|a, b) = p(x, y_1|a, b) = p(x, y_2|a, b)$ for all $b \in \{0, 1\}$.

This relates to our commitment scheme as follows: Suppose that the dishonest provers' strategy $q(x, y, z|a, b, c)$ is such that $q(y = z|b = c) = 1$. It follows that $q(x, y|a, b = x) = q(x, z|a, c = x)$ for $x = 0, 1$. Thus, the distribution $q(x, y|a, b)$ is 2-shareable and hence local. Hence, the dishonest provers can not succeed with a better probability than classical provers.

Remark 6.21. *The three-prover scheme above has the drawback that* two *provers are involved in the opening phase; as such, there needs to be* agreement *on whether to open the commitment or not; if there is disagreement then this may be problematic in certain applications. However, P and Q are not allowed to communicate. One possible solution is to have V forward an* authenticated *"open" or "not open" message from P to Q and R. This allows for some communication from P to Q and R, but if the size of the authentication tag is small enough compared to the security parameter of the scheme, i.e., n, then security is still ensured.*

[6] More generally, they show that if $b \in \{0, \ldots, m-1\}$ and $p(x, y|a, b)$ is m-shareable, then $p(x, y|a, b)$ is local.

Bibliography

[AK15a] Emily Adlam and Adrian Kent. Deterministic relativistic quantum bit commitment. *International Journal of Quantum Information*, 13(05):1550029, aug 2015.

[AK15b] Emily Adlam and Adrian Kent. Device-independent relativistic quantum bit commitment. *Physical Review A*, 92:022315, Aug 2015.

[AK16] Emily Adlam and Adrian Kent. Quantum paradox of choice: More freedom makes summoning a quantum state harder. *Physical Review A*, 93:062327, Jun 2016.

[BB84] Charles Bennet and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. In *IEEE International Conference on Computers, Systems and Signal Processing*, volume 1, pages 175–179, 1984.

[BC96] Gilles Brassard and Claude Crépeau. 25 years of quantum cryptography. *SIGACT News*, 27(3):13–24, 1996.

[BC16] Rémi Bricout and André Chailloux. Recursive cheating strategies for the relativistic $\mathbb{F}_Q$ bit commitment protocol. *ArXiv e-prints*, arXiv:1608.03820 [quant-ph], 2016. `https://arxiv.org/abs/1608.03820`.

[BCC88] Gilles Brassard, David Chaum, and Claude Crépeau. Minimum disclosure proofs of knowledge. *Journal of Computer and System Sciences*, 37(2):156–189, October 1988.

[BCF+11] Harry Buhrman, Nishanth Chandran, Serge Fehr, Ran Gelles, Vipul Goyal, Rafail Ostrovsky, and Christian Schaffner. Position-Based Quantum Cryptography: Impossibility and Constructions. In Phillip Rogaway, editor, *Advances in Cryptology - CRYPTO 2011*, pages 429–446, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.

[Bel64] John Stewart Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1:195–200, 1964.

[BGKW88] Michael Ben-Or, Shafi Goldwasser, Joe Kilian, and Avi Wigderson. Multi-Prover Interactive Proofs: How to Remove Intractability Assumptions. In Janos Simon, editor, *STOC 1988*, pages 113–131. ACM, 1988.

[Blu82] Manuel Blum. Coin flipping by telephone. In Allen Gersho, editor, *Advances in Cryptology: A Report on CRYPTO 81*, pages 11–15. U. C. Santa Barbara, Dept. of Elec. and Computer Eng., 1982.

[BS15] Mohammad Bavarian and Peter W. Shor. Information Causality, Szemerédi-Trotter and Algebraic Variants of CHSH. In Tim Roughgarden, editor, *ITCS 2015*, pages 123–132. ACM, 2015.

[CCM98] Christian Cachin, Claude Crépeau, and Julien Marcil. Oblivious transfer with a memory-bounded receiver. In *FOCS 1998* [DBL98], pages 493–502.

[CHSH69] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. Proposed Experiment to Test Local Hidden-Variable Theories. *Physical Review Letters*, 23:880–884, 1969.

[CK88] Claude Crépeau and Joe Kilian. Achieving oblivious transfer using weakened security assumptions (extended abstract). In *FOCS 1988*, pages 42–52. IEEE Computer Society, 1988.

[CK06] Roger Colbeck and Adrian Kent. Variable-bias coin tossing. *Physical Review A*, 73:032320, Mar 2006.

[CK12] Sarah Croke and Adrian Kent. Security details for bit commitment by transmitting measurement outcomes. *Physical Review A*, 86:052309, Nov 2012.

[CKW00] Valerie Coffman, Joydip Kundu, and William K. Wootters. Distributed entanglement. *Physical Review A*, 61:052306, Apr 2000.

[CM97] Christian Cachin and Ueli M. Maurer. Unconditional security against memory-bounded adversaries. In Burton S. Kaliski Jr., editor, *Advances in Cryptology - CRYPTO '97*, volume 1294 of *Lecture Notes in Computer Science*, pages 292–306. Springer, 1997.

[Col06] Roger Colbeck. *Quantum And Relativistic Protocols For Secure Multi-Party Computation*. PhD thesis, University of Cambridge, 2006.

[Col07] Roger Colbeck. Impossibility of secure two-party classical computation. *Physical Review A*, 76:062308, Dec 2007.

[CSST11] Claude Crépeau, Louis Salvail, Jean-Raymond Simard, and Alain Tapp. Two Provers in Isolation. In Dong Hoon Lee and Xiaoyun Wang, editors, *Advances in Cryptology - ASIACRYPT 2011*, volume 7073 of *Lecture Notes in Computer Science*, pages 407–430. Springer, 2011.

[DBL98] *FOCS 1998*. IEEE Computer Society, 1998.

[DFSS05] Ivan Damgård, Serge Fehr, Louis Salvail, and Christian Schaffner. Cryptography in the bounded quantum-storage model. In *FOCS 2005*, pages 449–458. IEEE Computer Society, 2005.

[Dir39] Paul A. M. Dirac. A new notation for quantum mechanics. *Mathematical Proceedings of the Cambridge Philosophical Society*, 35(3):416–418, 1939.

[EGL83] Shimon Even, Oded Goldreich, and Abraham Lempel. A randomized protocol for signing contracts. In David Chaum, Ronald L. Rivest, and Alan T. Sherman, editors, *Advances in Cryptology: Proceedings of CRYPTO '82*, pages 205–210. Plenum Press, New York, 1983.

[EPR35] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47:777–780, May 1935.

[FF15] Serge Fehr and Max Fillinger. Multi-Prover Commitments Against Non-Signaling Attacks. In Rosario Genaro and Matthew Robshaw, editors, *Advances in Cryptology - CRYPTO 2015, part II*, volume 9216 of *Lecture Notes in Computer Science*, pages 403–421. Springer, 2015.

[FF16] Serge Fehr and Max Fillinger. On the Composition of Two-Prover Commitments, and Applications to Multi-round Relativistic Commitments. In Marc Fischlin and Jean-Sébastien Coron, editors, *Advances in Cryptology - EUROCRYPT 2016, part II*, volume 9665 of *Lecture Notes in Computer Science*, pages 477–496. Springer, 2016.

[FL92] Uriel Feige and László Lovász. Two-prover one-round proof systems: Their power and their problems (extended abstract). In S. Rao Kosaraju, Mike Fellows, Avi Wigderson, and John A. Ellis, editors, *Proceedings of the 24th Annual ACM Symposium on Theory of Computing*, pages 733–744. ACM, 1992.

[For98] Lance Fortnow. *Complexity-Theoretic Aspects of Interactive Proof Systems*. PhD thesis, Massachusetts Institute of Technology, 1998.

[Hol09] Thomas Holenstein. Parallel Repetition: Simplification and the No-Signaling Case. *Theory of Computing*, 5(8):141–172, 2009.

[IL89] Russell Impagliazzo and Michael Luby. One-way functions are essential for complexity based cryptography (extended abstract). In *FOCS 1989*, pages 230–235. IEEE Computer Society, 1989.

[Kan15] Jędrzej Kaniewski. *Relativistic quantum cryptography*. PhD thesis, University of Cambridge, 2015.

[Ken99] Adrian Kent. Unconditionally Secure Bit Commitment. *Physical Review Letters*, 83(7):1447–1450, 1999.

[Ken05] Adrian Kent. Secure Classical Bit Commitment Using Fixed Capacity Communication Channels. *Journal of Cryptology*, 18(4):313–335, 2005.

[Ken11] Adrian Kent. Unconditionally secure bit commitment with flying qudits. *New Journal of Physics*, 13(11):113015, 2011.

[Ken12] Adrian Kent. Unconditionally Secure Bit Commitment by Transmitting Measurement Outcomes. *Physical Review Letters*, 109:130501, Sep 2012.

[Ken13] Adrian Kent. A no-summoning theorem in relativistic quantum theory. *Quantum Information Processing*, 12(2):1023–1032, 2013.

[Ken18] Adrian Kent. Summoning, No-Signaling and Relativistic Bit Commitments. *ArXiv e-prints*, arXiv:1804.05246 [quant-ph], April 2018. `https://arxiv.org/abs/1804.05246`.

[KMS11] Adrian Kent, William Munro, and Timothy Spiller. Quantum tagging: Authenticating location via quantum information and relativistic signaling constraints. *Phys. Rev. A*, 84:012326, Jul 2011.

[KMSB06] Adrian Kent, William Munro, Timothy Spiller, and Raymond Beausoleil. Tagging systems, 2006. Patent No. US 7075438.

[KTHW13] Jędrzej Kaniewski, Marco Tomamichel, Esther Hänggi, and Stephanie Wehner. Secure bit commitment from relativistic constraints. *IEEE Transactions on Information Theory*, 59:4687–4699, 2013.

[LC97] Hoi-Kwong Lo and Hoi Fung Chau. Is quantum bit commitment really possible? *Physical Review Letters*, 78:3410–3413, 1997.

[LKB+13] Tommaso Lunghi, Jędrzej Kaniewski, Felix Bussières, Raphael Houlmann, Marco Tomamichel, Adrian Kent, Nicolas Gisin, Stephanie Wehner, and Hugo Zbinden. Experimental bit commitment based on quantum communication and special relativity. *Physical Review Letters*, 111:180504, 2013.

[LKB+15] Tommaso Lunghi, Jędrzej Kaniewski, Felix Bussières, Raphael Houlmann, Marco Tomamichel, Stephanie Wehner, and Hugo Zbinden. Practical Relativistic Bit Commitment. *Physical Review Letters*, 115, 2015.

[LWW+10] Lars Lydersen, Carlos Wiechers, Christoffer Wittmann, Dominique Elser, Johannes Skaar, and Vadim Makarov. Hacking commercial quantum cryptography systems by tailored bright illumination. *Nature Photonics*, 4:686–689, Aug 2010.

[MAG06] Ll. Masanes, A. Acin, and N. Gisin. General properties of nonsignaling theories. *Physical Review A*, 73:012112, Jan 2006.

[Mal10a] Robert Malaney. Location-dependent communications using quantum entanglement. *Phys. Rev. A*, 81:042319, Apr 2010.

[Mal10b] Robert Malaney. Quantum location verification in noisy channels. In *2010 IEEE Global Telecommunications Conference GLOBECOM 2010*, pages 1–6, 2010.

[May97] Dominic Mayers. Unconditionally Secure Quantum Bit Commitment is Impossible. *Physical Review Letters*, 18:3414–3417, 1997.

[MY98] Dominic Mayers and Andrew Chi-Chih Yao. Quantum cryptography with imperfect apparatus. In *FOCS 1998* [DBL98], pages 503–509.

[Nao91] Moni Naor. Bit commitment using pseudorandomness. *Journal of Cryptology*, 4(2):151–158, 1991.

[NC00] Michael Nielsen and Isaac Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[Par70] James L. Park. The concept of transition in quantum mechanics. *Foundations of Physics*, 1(1):23–33, Mar 1970.

[PR94] Sandu Popescu and Daniel Rohrlich. Quantum nonlocality as an axiom. *Foundations of Physics*, 24(3):379–385, 1994.

[RAM16] Ravishankar Ramanathan, Remigiusz Augusiak, and Gláucia Murta. Generalized xor games with d outcomes and the task of nonlocal computation. *Physical Review A*, 93:022333, Feb 2016.

[RK05] Renato Renner and Robert König. Universally Composable Privacy Amplification Against Quantum Adversaries. In Joe Kilian, editor, *Theory of Cryptography*, volume 3378 of *Lecture Notes in Computer Science*, pages 407–425. Springer, 2005.

[RKKM14] I. V. Radchenko, K. S. Kravtsov, S. P. Kulik, and S. N. Molotkov. Relativistic quantum cryptography. *Laser Physics Letters*, 11(6):065203, 2014.

[Sca16] Giada Scalpelli. On the Binding Property of Two-Prover Commitment Schemes: Definitions and Composability. Master's thesis, Università degli Studi Di Padova, 2016.

[SCK14] Jamie Sikora, André Chailloux, and Iordanis Kerenidis. Strong Connections Between Quantum Encodings, Non-Locality and Quantum Cryptography. *Physical Review A*, page 9, 2014.

[Sha49] Claude Shannon. Communication Theory of Secrecy Systems. *The Bell System Technical Journal*, 28(4):656–715, Oct 1949.

[Sho97] Peter Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, October 1997.

[Sim07] Jean-Raymond Simard. Classical and quantum strategies for bit commitment schemes in the two-prover model. Master's thesis, McGill University, Montréal, Québec, 2007.

[SRA81] Adi Shamir, Ronald L. Rivest, and Leonard M. Adleman. Mental poker. In David A. Klarner, editor, *The Mathematical Gardner*, pages 37–43. Springer US, Boston, MA, 1981.

[VMH+16] Ephanielle Verbanis, Anthony Martin, Raphaël Houlmann, Gianluca Boso, Félix Bussieres, and Hugo Zbinden. 24-Hour Relativistic Bit Commitment. *Physical Review Letters*, 117(14):140506, 2016. ID: unige:88082.

[WCSL10] Stephanie Wehner, Marcos Curty, Christian Schaffner, and Hoi-Kwong Lo. Implementation of two-party protocols in the noisy-storage model. *Physical Review A*, 81(5):052336, May 2010.

[Wil13] Mark Wilde. *Quantum Information Theory*. Cambridge University Press, New York, NY, USA, 2013.

[Wyn75] Aaron D. Wyner. The wire-tap channel. *The Bell System Technical Journal*, 54(8):1355–1387, Oct 1975.

[WZ82] W. K. Wootters and W. H. Zurek. A single quantum cannot be cloned. *Nature*, 299:802, October 1982.

Summary

This dissertation makes different contributions to the theory of multi-prover commitment schemes; in particular relativistic commitment schemes. A *commitment scheme* is an important cryptographic primitive crucial for cryptographic protocols that allow two or more parties that do not fully trust each other to cooperate in a secure way. More concretely, we consider multi-prover commitment schemes whose security relies on an assumed restriction on the communication between the provers, and not on computational hardness assumptions.

A commitment scheme is a means to solve the following problem: Alice has selected a message which she wants to keep secret at the moment, but which she may want to reveal to Bob at a later time. However, if she simply sends the message to Bob later, he has no guarantee that the message he received is the same one that Alice selected earlier. A commitment scheme ensures the secrecy of the message until Alice chooses to reveal it, but also prevents Alice from revealing a different message than the one she selected originally.

Formally, a commitment scheme consists of a pair of interactive protocols between (usually) two parties called the *prover* and the *verifier*. The first protocol is called the *commit phase*; it takes a message from the prover as input and no input from the verifier. The second protocol, called the *opening phase*, then outputs either a message or the failure symbol $\perp$ to the verifier. If the output is a message m, we say that the prover opened the commitment to m; if the output is $\perp$, we say that the prover failed to open the commitment. Often, the opening phase just consists in the prover sending some *opening information* to the verifier who then computes the output locally.

To be secure, a commitment scheme needs to have the following three properties: It should be *complete*, meaning that the input to the commit phase and the output of the opening phase are equal if both parties follow the protocols. It should be *hiding*, meaning that the verifier cannot learn the prover's input message before the opening phase is executed, even if the verifier is dishonest and deviates from the protocols. Finally, it should be *binding*, meaning that after the commit phase, there is at most one message that the prover can successfully open to, even if the prover is dishonest and deviates from the protocols. The set of possible messages that a commitment

scheme can take as input is called its *domain*. If the domain is the set $\{0, 1\}$, we speak of a *bit*-commitment scheme.

The standard notion of commitment schemes can offer security only if the dishonest prover or the dishonest verifier is *computationally bounded*, i.e. limited in the amount of computation he can perform. However, this is only true for commitment schemes with a single prover. Ben-Or, Goldwasser, Kilian and Wigderson showed in 1988 how to overcome this limitation by considering a variant of the notion of commitment schemes where the prover is split into two (or more) separate entities and it is assumed that they cannot communicate during the execution of the commitment scheme. Related to this approach is the notion of *relativistic commitment schemes*, introduced by Kent in 1999, where this non-communication assumption is temporarily enforced through spatial separation of the provers.

The first main contribution of this dissertation is a set of new definitions of the binding property for multi-prover commitment schemes. These new definitions have several advantages over the *sum-binding* definition which has been used so far: They are not restricted to *bit*-commitment schemes but are applicable to commitment schemes with arbitrary finite domains. When restricted to bits, some of our definitions are *strictly stronger* than the sum-binding definition. Finally, our definitions are closer to the intuitive notion of a commitment scheme being binding and are more convenient to work with. We introduce these new definitions and study how they relate to each other.

As a testing ground for our new definitions, we consider the bit-commitment scheme $\mathcal{CHSH}^q$, introduced by Crépeau, Salvail, Simard and Tapp, which can be extended in a natural way to a commitment scheme with domain $\mathbb{F}_q$ (where q is a prime power). We analyze it with respect to this larger domain for the first time and show that different variations of the scheme satisfy our different definitions of the binding property.

Our new definitions enable us to prove a rather general *composition theorem* for two-prover commitment schemes, which is the second main contribution of this dissertation. We compose two commitment schemes by having the provers commit to the opening information of the first scheme instead of sending it to the verifier, and then they open this second commitment, revealing the opening information of the first scheme so that the original commitment is opened. Under some mild assumptions about the two original schemes, we prove that the composed scheme is binding if the two original schemes are binding (with the cheating probabilities adding up).

The purpose of this composition is to *delay* the opening of the commitment. This is important in the context of relativistic commitment schemes where the no-communication assumption is only enforced temporarily and so the binding property holds only for a limited time.

Very concretely, our composition theorem allows us to give a tight analysis of the relativistic commitment scheme introduced by Lunghi, Kaniewski, Brussières, Houlman, Tomamichel and Wehner in 2015. Their original anal-

ysis showed an upper bound on the cheating probability of dishonest provers that was doubly-exponential in the number of communication rounds where the latter determines for how long the scheme remains binding. The scheme can be understood as an iterated composition of $\mathcal{CHSH}^q$ with itself and so we can analyze its security using our composition theorem, achieving a significant improvement over the original analysis by Lunghi *et al.*: it follows that the cheating probability for dishonest provers can be bounded by a term that is only *linear* in the number of rounds, rather than double-exponential. We also show the optimality of our bound up to a small constant factor.

To put this difference into more concrete terms: Lunghi *et al.* implemented their scheme with provers in Bern and Geneva (distance: 129.2 km). Their analysis guaranteed that the commitment would stay binding (with a reasonably low cheating probability) for about 2 ms. Based on our analysis, this time scales up to 10^{56} years, or, speaking more practically, until the devices run out of memory.

The third main contribution is an impossibility result about two-prover commitment schemes with *general non-signaling adversaries*. As Crépeau, Salvail, Simard and Tapp pointed out, the assumption that the provers cannot communicate needs further specification. Some two-prover commitment schemes are secure against classical non-communicating provers but insecure if they have quantum capabilities and can use entangled quantum states shared among them. Furthermore, if we want to truly base security on the sole assumption that the provers cannot communicate, we need to consider general non-signaling provers, i.e., provers whose behavior may be correlated in arbitrary ways as long as no communication between them is implied. The $\mathcal{CHSH}^q$ scheme is secure against provers with quantum entanglement, but insecure against general non-signaling provers. This raises the question whether any other commitment schemes are secure against such general non-signaling provers.

We show that for *two-prover* commitment schemes the answer is no: any commitment scheme that is complete and hiding is by necessity not binding against general non-signaling provers. On the other hand, we show a positive answer for *three-prover* commitment schemes: we prove that a simple extension of $\mathcal{CHSH}^q$ to three provers is complete, hiding against an arbitrary dishonest verifier and binding against general non-signaling provers.

Samenvatting

Deze dissertatie draagt bij aan de theorie van multi-prover commitment schemes, in het bijzonder relativistische commitment schemes. Een *commitment scheme* is een cruciale bouwsteen voor cryptografische protocollen welke meerdere partijen die elkaar niet vertrouwen in staat stellen om op een veilige manier samen te werken. Meer specifiek beschouwen wij multi-prover commitment schemes waarvan de veiligheid gebaseerd is op veronderstelde beperkingen op de communicatie tussen de provers, en niet op aannamen over de computationele complexiteit van wiskundige problemen.

Een commitment scheme is een middel om het volgende probleem op te lossen: Alice heeft een bericht geschreven die zij eerst geheim wil houden, maar die zij mogelijk later aan Bob wil onthullen. Echter, als zij het bericht eenvoudigweg op een latere tijd aan Bob zou sturen, weet Bob niet zeker of het bericht dat hij ontving gelijk is aan het bericht dat zij eerder heeft geschreven. Een commitment scheme zorgt ervoor dat het bericht geheim blijft totdat Alice het wil onthullen, maar ook dat Alice geen ander bericht aan Bob kan onthullen.

Een commitment scheme is formeel gedefinieerd als een tweetal interactieve protocollen tussen (meestal) twee partijen, genoemd de *prover* en de *verifier*. Het eerste protocol, genoemd de *commit* fase, neemt een bericht van de prover als input en geen input van de verifier. Het tweede protocol, genoemd de *opening* fase, geeft dan een bericht uit aan de verifier, of $\perp$ om aan te geven dat de prover er niet in geslaagd is om het commitment te openen. Als de output een bericht m is, zegt men dat de prover het commitment naar m geopend had. Vaak stuurt de prover in de opening fase alleen zogenoemde *opening information* aan de verifier die dan lokaal de output berekent.

Om veilig te zijn moet een commitment scheme de volgende eigenschappen hebben: het moet *compleet* zijn, d.w.z. dat de input van de commit fase gelijk is aan de output van de opening fase als beide partijen de protocollen volgen. Het moet *verbergend* zijn, d.w.z. dat de verifier de input van de prover niet voor de opening fase kan uitvinden, ook als de verifier oneerlijk is en van de protocollen afwijkt. Het moet *bindend* zijn, d.w.z. dat de prover na de commit fase maar naar één bericht kan openen, ook als de prover oneerlijk is en van de protocollen afwijkt. De verzameling van mogelijke inputs van een commitment

scheme wordt zijn *domein* genoemd. Als het domein de verzameling $\{0, 1\}$ is, spreken we van een *bit*-commitment scheme.

Een standaard commitment scheme kan alleen veilig zijn als tenminste één van de partijen beperkte rekenkracht heeft. Dit is echter alleen waar als er maar één prover is. Ben-Or, Goldwasser, Kilian en Wigderson bewezen in 1988 dat deze beperking omzeild kan worden door een variant van commitment schemes waar de prover opgesplitst wordt in twee (of meer) aparte entiteiten die per aanname gedurende de uitvoering van het commitment scheme niet kunnen communiceren. Gerelateerd aan dit idee is het idee van *relativistische commitment schemes*, dat in 1999 door Kent werd voorgesteld: door de provers ver van elkaar te plaatsen is het mogelijk om tijdelijk aan de aanname te voldoen dat de provers niet kunnen communiceren.

De eerste hoofdbijdrage van deze dissertatie zijn nieuwe definities voor het bindend-zijn van een multi-prover commitment scheme. Deze nieuwe definities hebben meerdere voordelen boven de tot nu gebruikte *sum-binding* definitie: ze zijn niet beperkt tot *bit*-commitment schemes, maar van toepassing voor commitment schemes met arbitraire eindige domeinen. Voor bit-commitment schemes zijn sommige van onze definities *strikt sterker* dan de sum-binding definitie. Verder blijven onze definities dichter bij de intuïtie en zijn makkelijker te gebruiken. Wij stellen deze definities voor en bestuderen de relaties tussen hen.

Om onze definities te testen gebruiken wij het $\mathcal{CHSH}^q$ bit-commitment scheme – geïntroduceerd door Crépeau, Salvail, Simard en Tapp – die op een natuurlijke manier uitgebreid kan worden naar een commitment scheme met domein $\mathbb{F}_q$ (waar q een macht van een priemgetal is). Wij analyseren het voor het eerst als een commitment scheme voor dit grotere domein en bewijzen dat verschillende varianten van het commitment scheme aan verschillende definities voldoen.

Onze nieuwe definities stellen ons in staat om een vrij algemene *compositie-stelling* voor two-prover commitment schemes te bewijzen, de tweede hoofdbijdrage van deze dissertatie. Wij voegen twee commitment schemes samen door de provers aan de *opening information* van het eerste scheme te laten committen met het tweede scheme, en dan dit tweede commitment te openen. Dit onthult de opening information en het eerste commitment wordt dus geopend. Wij bewijzen dat het samengestelde scheme bindend is als de twee originele schemes bindend zijn en aan enkele lichte verdere eisen voldoen. (De succeskans voor oneerlijke provers in het samengestelde commitment scheme is de som van hun succeskansen in de twee originele schemes.)

Het doel van deze compositie is het openen van het commitment te vertragen. Dit is belangrijk in de context van relativistische commitment schemes, waar de aanname dat de provers niet kunnen communiceren alleen tijdelijk geldig is.

Concreet maakt onze compositie-stelling een betere analyse mogelijk van het relativistische commitment scheme dat door Lunghi, Kaniewski, Brussiè-

res, Houlman, Tomamichel en Wehner in 2015 werd voorgesteld. Hun originele analyse bewees een bovengrens aan de succeskans van oneerlijke provers die dubbel exponentieel was in het aantal communicatie-ronden, dat bepaalt hoe lang het commitment scheme bindend zal blijven. Het scheme kan beschouwd worden als een iteratieve compositie van $\mathcal{CHSH}^q$ met zichzelf, en dus kunnen wij het met behulp van onze compositie-stelling analyseren. Op deze manier bereiken wij een enorme verbetering van het resultaat van Lunghi *et al.*: wij bewijzen dat de succeskans van oneerlijke provers lineair is in het aantal communicatie-ronden, en niet dubbel exponentieel. We bewijzen ook dat ons resultaat optimaal is, op een klein constante factor na.

Meer concreet: Lunghi *et al.* hebben hun commitment scheme met provers in Bern en Genève (afstand: 192.2 km) geïmplementeerd. Volgens hun analyse bleef het scheme bindend (met een redelijk lage succeskans voor oneerlijke provers) voor 2 ms. Onze analyse bewijst dat een duur van 10^{56} jaar met dezelfde grens aan de succeskans mogelijk is – of, meer praktisch, totdat de geheugens van de apparaten vol zijn.

De derde hoofdbijdrage is een bewijs dat er geen two-prover commitment schemes bestaan die veilig zijn tegen algemene *non-signaling provers*. Zoals bewezen door Crépeau, Salvail, Simard en Tapp moet de niet-communicatie premisse verder worden uitgewerkt. Sommige commitment schemes zijn veilig tegen klassieke oneerlijke provers, maar onveilig tegen provers die een verstrengelde kwantumtoestand delen. Wil men nog verder gaan en de veiligheid van een commitment scheme echt alleen bouwen op de aanname dat de provers niet kunnen communiceren, moet men algemene non-signaling provers beschouwen. Dit betekent dat het gedrag van de provers op welke manier dan ook gecorreleerd kan zijn, zolang het geen communicatie tussen hen impliceert. $\mathcal{CHSH}^q$ is veilig tegen provers met kwantumverstrengeling, maar niet tegen algemene non-signaling provers. Dit werpt de vraag op of een ander commitment scheme wel veilig tegen zulke provers is.

Wij bewijzen dat dit voor *two-prover* commitment schemes niet het geval is: een commitment scheme dat compleet en verbergend is kan niet bindend zijn voor algemene non-signaling provers. Anderzijds hebben wij ook een positief resultaat: wij bewijzen dat een eenvoudige uitbreiding van $\mathcal{CHSH}^q$ naar een *three-prover* commitment scheme compleet, verbergend voor een arbitraire oneerlijke verifier, en bindend voor algemene non-signaling provers is.

Acknowledgements

First of all, I want to thank my promotor Serge Fehr. It was a pleasure to work with him and taught me a lot about both research and communicating ideas effectively. During our discussions, he always insisted on not only producing results, but also providing an intuitive understanding of *why* the results are correct. I realized that this is not only a convenience for the reader – which already is important enough by itself – but can also lead to stronger results. I am also grateful for his patience especially when the final part of my dissertation turned out to take much longer than anticipated.

I want to thank Léo Ducas for collaborating on our sideproject on homomorphic encryption, and Ronald Cramer for many interesting discussions ranging from mathematics to politics.

I would like to thank Christian Schaffner whose cryptography course at the Universiteit van Amsterdam sparked my interest in the subject, and Marc Stevens for the opportunity to follow up on this interest with a Master thesis as an intern at the CWI Cryptology Group.

I would like to thank my fellow PhD students at the CWI, especially Gabriele Spini and Diego Mirandola, for many fun evenings at the pub, movie nights, table soccer matches, and so on.

Finally, I want to thank my family and friends!

Curriculum Vitae

Maximilian Fillinger was born in Wuppertal, Germany, on March 22, 1988 and grew up first in the nearby city Remscheid, and later in Wuppertal.

He began his studies at the Heinrich-Heine-Universität Düsseldorf in 2005 and obtained Bachelor degrees in philosophy and mathematics in 2008 and 2009, respectively.

He then enrolled in the Master of Logic programme at the Institute for Logic, Language and Computation of the Universiteit van Amsterdam. During his studies there, he became interested in cryptography. In 2013, he obtained his Master degree. His thesis, titled "Reconstructing the Cryptanalytic Attack behind the Flame Malware", was written during an internship at the cryptology group of the Centrum Wiskunde & Informatica (CWI) in Amsterdam, supervised by Christian Schaffner and Marc Stevens. Later in 2013, he obtained a PhD position at the cryptology group of the CWI under the supervision of Serge Fehr.

Since 2018, he works as a software developer at Fox-IT in Delft.

Publications

- Guillaume Bonnoron, Léo Ducas and Max Fillinger. *Large FHE Gates from Tensored Homomorphic Accumulator.* In *Progress in Cryptology - AFRICACRYPT 2018*, pages 217-251.
- Serge Fehr and Max Fillinger. *On the Composition of Two-Prover Commitments, and Applications to Multi-round Relativistic Commitments.* In *Advances in Cryptology - EUROCRYPT 2016, part II*, pages 477-496. An earlier version was presented at *QCRYPT 2015*. An extended version is available at `https://arxiv.org/abs/1507.00240`.
- Max Fillinger and Marc Stevens. *Reverse-engineering of the cryptanalytic attack used in the Flame super-malware.* In *Advances in Cryptology - ASIACRYPT 2015, part II*, pages 586-611.
- Serge Fehr and Max Fillinger. *Multi-Prover Commitments Against Non-Signaling Attacks.* In *Advances in Cryptology - CRYPTO 2015, part II*, pages 403-421. Also presented at *QCRYPT 2015* .