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

Fillinger, M.J.

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.