UNIVERSITY OF TARTU
Faculty of Science and Technology
Institute of Computer Science
Computer Science Curriculum

Mattias Lass

Post-Quantum Security Definitions for Relativistic Commitment Protocols

Master's Thesis (30 ECTS)

Supervisor: Dominique Unruh, PhD

Tartu 2021

Post-Quantum Security Definitions for Relativistic Commitment Protocols

Abstract:
Relativistic commitment protocols are a promising candidate for post-quantum information theoretically secure commitments. Unfortunately, the security analysis of these protocols has often been informal.
We propose a novel formal framework for proving the security of relativistic bit commitment protocols and use it to prove security for a well established protocol. In addition, we also study an alternative definition of binding for relativistic commitment protocols based on the collapse-binding definition introduced in (Unruh, Eurocrypt 2016).

Keywords:
Commitments, relativistic protocols, collapse-binding, post-quantum cryptography, information theoretical security

CERCS: P170 – Computer science, numerical analysis, systems, control

Postkvant-turvanõuete definitsioonid relatiivsetele kinnistusprotokollidele

Lühikokkuvõte: Relatiivsed kinnistusprotokollid lubavad saavutada informatsiooniteoreetiliselt postkvant-turvalisust. Kahjuks on nende turvalisust seni peamiselt analüüsitud mitteformaalsete meetoditega.

Me esitame uue formaalse raamistiku, mille abil on võimalik analüüsida relatiivsete bitikinnistusprotokollide turvalisust ning tõestame selles raamistikus ühe tunnustatud protokolli turvalisuse. Lisaks uurime me ka alternatiivset siduvuse definitsiooni, mis põhineb niinimetatud kollaps-siduvuse (Unruh, Eurocrypt 2016) definitsioonil.

Võtmesõnad:
Kinnistus, relatiivsed protokollid, kollaps-siduvus, postkvant-krüptograafia, informatsiooniteoreetiline turvalisus

CERCS:P170 – Arvutiteadus, arvutusmeetodid, süsteemid, juhtimine (automaatjuhtimisteooria

Contents

1 Introduction

Relativistic commitment protocols are a promising candidate for information theoretically secure commitments. This is contrary to the fact that information theoretically secure commitments have been proven impossible barring certain strong assumptions. It would therefore seem that relativism, which is achievable in practice, is one such strong assumption.

Unfortunately, the security analysis of these protocols has often been mostly informal. There have been few attempts to actually model relativistic commitment protocols, instead most have used non-relativistic security definitions. Such definitions however, disregard the complexities of relativism.

In this thesis we use the framework of spacetime circuits to formulate explicitly relativistic security definitions and prove security of a state of the art relativistic commitment protocol using this framework. To our knowledge, this is the first time security of a relativistic protocol has been proven to such a degree. We also define the notion of relativistic collapse-binding and study the usefulness of this notion by trying to prove collapse-binding for the beforementioned commitment protocol. We believe that the result we achieve indicates both that the collapse-binding definition is useful and that it is provable that the protocol we investigate is collapse-binding.

In the second section of this thesis we give an overview of the literature relevant to relativistic commitment protocols and spacetime circuits. In the third section, we define our model for relativistic commitment protocols and using this model define security definitions for relativistic bit commitment protocols. We then prove security for an existing relativistic bit commitment scheme and briefly discuss what the definitions informally imply. In the final section, we investigate an alternative, more promising approach to define security by defining relativistic collapse-binding. We show that an existing relativistic commitment scheme is at least partially secure for this definition and briefly consider how the collapse-binding relates to security definitions proposed in section three.

2 Background

2.1 Commitment protocols

Commitment protocols are two party protocols where the first party sends some message to the second party which commits them to some *value* without revealing the value to the second party. We call this message the *commitment* and the party that sends it the *sender*. The party which receives the commitment we call the *recipient*. At some later point the sender can choose to reveal the value by sending an *opening* to the recipient. The recipient can then uncover the value such that they can be certain that this value is indeed what the sender originally committed to. The value committed to is strictly classical, but the messages exchanged can be either quantum or classical. Based on the nature of the messages these protocols are classified either as *quantum* or *classical* commitment protocols. Another useful class of commitment protocols are *bit commitment protocols*. These are protocols where the value committed to is a single bit. Committing to multiple bits in parallel can be used to commit on some larger value, thus a bit commitment protocol can actually be used to commit on any classical value. [1]

For a commitment protocol to be *secure*, it then follows that the protocol must fulfil two properties *binding* and *hiding*. Informally, a protocol is hiding if until the sender sends the opening, the recipient is unable to learn the value committed to. A protocol is binding if the sender, having sent the commitment, can only open to a single value. In the computational setting, achieving security is quite possible. Hiding and binding are strongly opposed if not outright contradictory, therefore in the information theoretical setting, which this thesis focuses on, achieving security is a much more difficult problem.

It is generally agreed that a *perfectly secure* protocol does not exist, i.e. a protocol where both binding and hiding properties are satisfied completely. For some time however, it was believed that quantum commitment protocols could achieve information theoretical security close to perfect security, in the literature this is often called *unconditional security*. In practice this could mean, for example, that the binding property is satisfied completely, and that the hiding property is satisfied to such a degree that a dishonest recipient would only have an exponentially small advantage in guessing the value committed to. One such protocol was believed to be the scheme BCJL [BCJL93], which was claimed to be provably secure meaning that it had been proven to achieve information theoretical security. Unfortunately, according to a review on the impossibility of quantum bit commitment [BCMS97] in 1995 Mayers found a subtle flaw in BCJL which proved the protocol to be insecure. Mayers did not publicise the result until a similar result was discovered independently by Lo and Chau in 1996. Mayers' discovery led him to then generalize his result into an impossibility theorem [May97] in 1997, which we refer to as the Mayers-Lo-Chau impossibility theorem. According to this

[1] such a composition is not actually as straightforward as implied here, particularly when defining security for commitment protocols, but the basic principle still applies

theorem an information theoretically secure bit commitment protocol is impossible to achieve in general, at least without some third party mutually trusted by both the sender and the recipient.[2]

As for the applications for commitment protocols, there are multiple. One standalone application is that a commitment protocol allows the sender to commit to some prediction on some future event without simultaneously revealing their prediction. Somewhat more interestingly, commitment protocols are used as a primitive for many larger cryptographic protocols including, for example, zero-knowledge proofs and secure multiparty computation. Making some simplifications, this means that Mayers-Lo-Chau impossibility theorem also implies that protocols built from commitment protocols can not be information theoretically secure.

2.2 Relativistic commitment protocols

Relativistic commitment protocols are a special subset of commitment protocols which rely on the impossibility of faster-than-light communication. In such protocols the parties are split into multiple agents and the restriction on the speed at which information can propagate is used to limit what information the agents have access to, which can in turn be used to guarantee the security of a protocol. Importantly, relativistic protocols have been claimed to evade the Mayers-Lo-Chau impossibility theorem meaning that they might achieve information theoretical security. The intuition behind this is that, in some sense, the impossibility of faster-than-light communication is actually the third "party" mutually trusted both by the recipient and the sender. On the other hand, such protocols are often somewhat limited in their practical application. For example, usually in these protocols the commitment only stays binding within some short window of time. Such protocols might then not be usable as a primitives for building larger protocols.

The first relativistic commitment scheme RBC1[Ken99] was published by Kent in 1999 as a theoretical challenge to the Mayers-Lo-Chau impossibility theorem. RBC1 is a classical bit-commitment protocol. The protocol itself consist of some number of rounds. In the first round one of the sender's agents performs the commitment. Subsequent rounds take place at alternating agents within some time window such that the communication from the previous round would not have time to reach the current round. In these subsequent rounds the sender has a choice to either open the commitment or to extend the commitment at the cost of exponentially growing rate of communication. If the sender's agent would take too long to do this, meaning that they could have access to the communication from the previous round, then the commitment would stop being binding. We give here a simplified description of RBC1 where the sender decides to open

[2] A simple example of a protocol using a third trusted party would be as follows. The sender commits by announcing the value they wish to commit to some middle-man. To open the sender request the middle-man to reveal the committed value to the recipient.

immediately, this illustrates well the basic principle of relativistic commitment protocols.

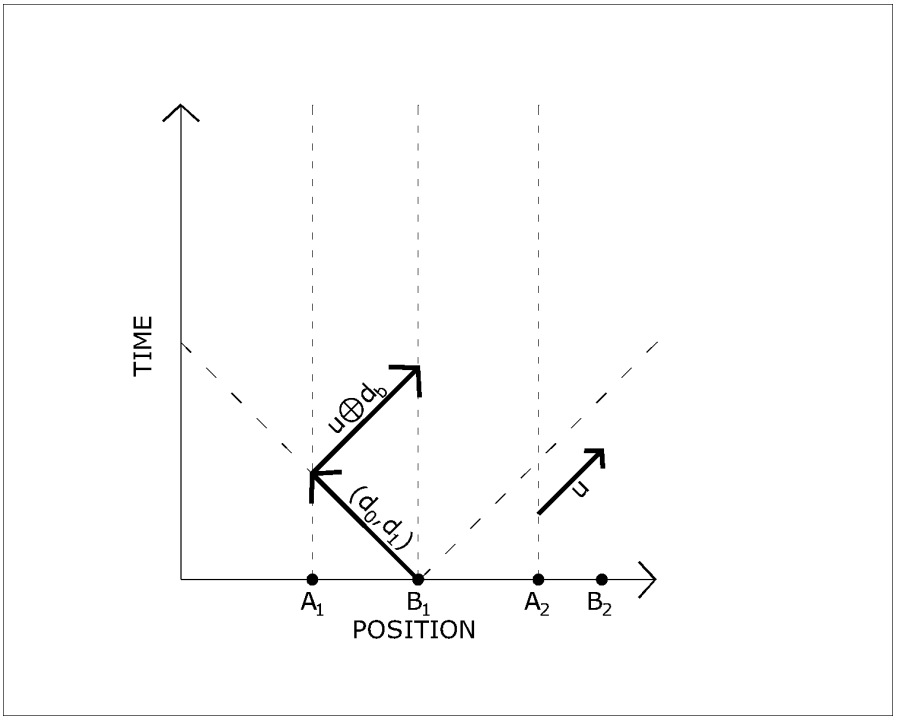


Figure 1. Message flow for "single-round RBC1" in the 1-D case

Single-round RBC1
commit:

1. For two distinct points in space both parties send an agent to these points resulting in two pairs of agents of different parties located at different points.
2. Let A_1 be the sender's agent at the first point and B_1 the recipient's agents at the same point, and let A_2 and B_2 be the respective agents at the second point.
3. Before the protocol begins the sender generates some secret n-bit string u which is given to both A_1 and A_2.
4. The protocol starts by B_1 revealing a challenge d, this consist of two different uniformly randomly chosen n-bit strings d_0 and d_1.
5. For the bit b the sender wishes to commit, A_1 then promptly replies with $c = d_b \oplus u$.

unveil:

6. To open A_2 reveals u within some time-window such that they are guaranteed to not know d.

7. The recipient then gathers u and c and d, allowing them to calculate $c \oplus u$ which should evaluate to d_b revealing the bit committed.

In practice the agents within a pair would not need to be at the exact same point. It is sufficient if they are close to each other and far away from the other pair. Note that in figure 1 this is actually the case.

To extend the protocol, instead of opening the sender's agent can commit to each of the bits of u using a new instance of the same protocol. This would then require n times more shared randomness between the sender's agents causing the exponential growth mentioned before.

Kent's arguments for the security of the protocol are essentially as follows. If A_2 decides not reveal u, the only information revealed by the sender is the uniformly random $d_b \oplus u$. Informally, this makes the protocol hiding. To break binding a malicious sender would have to output u such that $c \oplus u = d_b$ for some bit b. But as d_b is a uniformly randomly determined bit-string that the sender can not possibly know at the point at which they reveal u achieving this impossible. Thus the protocol should also be binding. Against classical adversaries these informal arguments are quite convincing, but as Kent himself also notes, in the quantum setting the case for the protocol being binding[3] is not so obvious. After providing some informal reasoning, Kent does however conclude his analysis by conjecturing that the protocol is secure against quantum adversaries.

Further notable progress in the field has also been mostly made by Kent. In 2005 he published a similar classical protocol RBC2 [Ken05] which had the improvement that it only required a fixed rate of communication. The analysis of the security of RBC2 was also more formal than the analysis for RBC1. This time Kent provided a clear definition for security and claimed to prove that both RBC1 and RBC2 are secure in the classical setting. Kent also claimed to prove that the protocol evades the Mayers-Lo-Chau impossibility theorem. However Kent did still not prove the security against general quantum attacks.

While we do not investigate the following schemes further, for the sake of completeness we note that in 2011 he published the first relativistic quantum bit commitment scheme [Ken11] and in 2012 a variant of this scheme implementable with current technology [Ken12].

[3] it is trivial to see that the protocol is hiding even in the quantum setting

At the time Kent introduced the concept of relativistic commitments, the idea to split the sender into multiple agents was not novel. In our knowledge, this was first done in 1988 in the BGKW commitment protocol [BOGK88]. In this protocol the sender's agents were simply required to not be able to communicate. However it was later shown that assumption of no communication is not sufficient to guarantee security [CSST11]. In essence, even if the agents can not communicate, they could still cooperate using some correlated events, for example their communication with the recipient. A stronger requirement of *isolation* was proposed, and while we do not investigate this claim further, we believe that secure relativistic protocols satisfy this assumption.

In 2014 Lunghi et al. [LKB$^+$15] proved the security of a relativistic version of BGKW against quantum adversaries. They claimed that to break binding the adversary would have to win the $CHSH_n$ game[4]. While their claim seems to hold, as we show in section 3.3, the reasoning behind it was somewhat informal. For one, it was not explicitly explained how the a **non-relativistic** $CHSH_n$ game models cheating in the **relativistic** BGKW protocol. The claim that relativistic BGKW is hiding is trivially provable.

Relativistic BGKW is in nature similar to the single round version of RBC1, therefore it only allows for the commitment to be binding for a very short time. Using this protocol as a basis, Lunghi et al. also proposed and performed a practical demonstration of a novel multi-round bit-commitment protocol, which allowed for the commitment time to be extended. The security for the extended protocol was proven only against classical adversaries.

Given that relativistic BGKW has a security proof against quantum adversaries we will be using it throughout the thesis as the state of the art protocol we study. The description of relativistic BGKW as presented by Lunghi et al [LKB$^+$15] is as follows.

Relativistic BGKW

commit:

1. For two distinct points in space both parties send an agent to these points resulting in two pairs of agents of different parties located at different points.
2. Let A_1 be the sender's agent at the first point and B_1 the recipient's agents at the same point, and let A_2 and B_2 be the respective agents at the second point.
3. Before the protocol begins the sender generates some secret n-bit string u which is given to both A_1 and A_2.
4. The protocol starts by B_1 revealing a challenge d a uniformly randomly chosen n-bit string.

[4] section 2.3 covers the $CHSH_n$ game

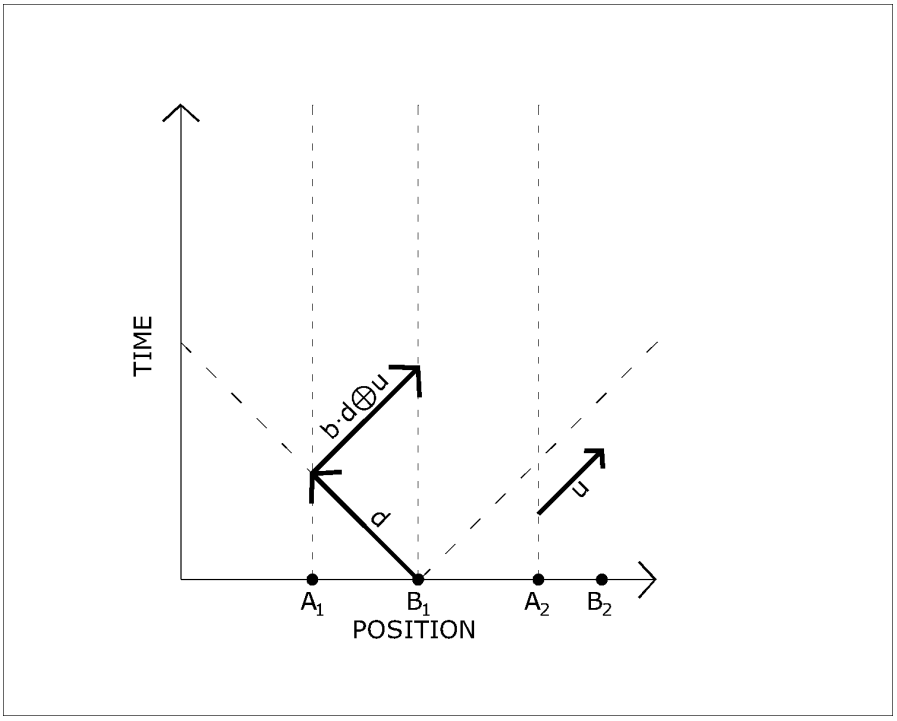


Figure 2. Message flow for "relativistic BGKW" in the 1-D case

5. For the bit b the sender wishes to commit, A_1 then promptly replies with $c = b \cdot d \oplus u$, where $0 \cdot d = 0...0$ and $1 \cdot d = d$.

unveil:

6. To open A_2 reveals u and b within some time-window such that they are guaranteed to not know d.

7. The recipient then gathers all of the information transmitted, allowing them to check whether $c = b \cdot d \oplus u$. If it does they accept that b was committed otherwise they reject the commitment.

There have also been some advances toward defining formal security for relativistic commitment protocols, i.e. the main goal of this thesis. Kaniewski et al.[KTHW13] proposed a formal model for proving security of relativistic commitment protocols. It differs from our method as it does not provide a clear way to translate the relativistic constraints into a non-relativistic setting. In 2017 Portmann et al. [PMM$^+$17] proposed the first generally composable security framework for relativistic protocols. Related to

this in 2019 Vilasini et al. [VPdR19] produced multiple impossibility results for general composability of relativistic bit commitment schemes.

2.3 $CHSH_n$ game

The $CHSH_n$ game is a generalization of the influential $CHSH$ game which for our purposes is important as it closely relates to the BGKW protocol. We provide here for context its definition as fomulated by Sikora et al. [SCK14]

Definition 1. *The $CHSH_n$ game is a game between Alice and Bob where:*

- *Alice and Bob are allowed to create and share an entangled state $|\psi\rangle$ before the game starts. Once the game starts, there is no further communication between Alice and Bob.*
- *Alice receives a random string $x \in \{0,1\}^n$ and Bob a random bit y.*
- *Alice outputs $a \in \{0,1\}^n$ and Bob outputs $b \in \{0,1\}^n$.*
- *Alice and Bob win if $a_i \oplus b_i = yx_i$ for all $i \in \{1, ..., n\}$*

2.4 Spacetime circuits

As was the case for relativistic BGKW, the security analysis for relativistic commitment schemes have mostly just simply stated that the relativistic constraints are enough to make the agents not be able to communicate. We see at least two problems with this approach. The claim itself is quite vague or informal, it is not always trivial to follow what information different agents might have access to. Secondly, as we noted before it has been shown that simply limiting communications might not be enough to guarantee security [CSST11]. Furthermore, not using a well established formal framework for security proofs is error-prone, especially in the case of post-quantum cryptography. Therefore in this thesis we demonstrate how spacetime circuits could be used to improve the confidence we have in the security of relativistic protocols.

The concept of *spacetime circuits* is a methodology for analyzing quantum circuits in spacetime introduced by Unruh [Unr14] for proving security of position verification schemes. Position verification schemes are schemes where a party attempts to prove that they are within some region of space. As an example of why using such a methodology is important, Unruh also demonstrated that one previously existing proof of a position verification scheme did not actually apply to cases where there was more than one positional dimension. As this aspect of the position verification protocols is quite similar to relativistic commitments, it is not improbable that proofs for security of relativistic protocols also share a similar issue.

We take an abstract approach to spacetime which is consistent with our current understanding of physics on Earth. For us spacetime is simply some set of coordinates and a partial order relation *precede*. The actual structure of the coordinates is not important for our analysis, but a reader can think of a coordinate as a three values which determine the location in space and a single value which determines the location in time. We call such a coordinate a *point*[5]. If information can flow from point p_1 to point p_2, as in practice defined by the speed of light, then p_1 *precedes* p_2. Precede is *reflexive*, i.e. every point precedes itself. Any subset of spacetime we call a *region*. The *causal past* of a point p or $C^-(p)$ is a region consisting of all of the points which precede p. The *causal future* of p or $C^+(p)$ is a region consisting of all of the points which p precedes. A region P is *past closed* if its equal to the union of causal pasts of all of the points in P. P is *future closed* if its equal to the union of causal futures of all of the points in P. Note that the complement of a past closed region is a future closed region and the complement of a future closed region is a past closed region.

A spacetime circuit then is a quantum circuit where every gate is at a particular point of spacetime. Wires from gate g_0 to g_1 can only exist if g_0 is in the causal past of g_1. As precede is a partial order, there can then also be no cycles between gates, which means that if we forget the points at which the gates are located, we can transform a spacetime circuit into a valid quantum circuit.

We also define some new terminology for spacetime circuits. A circuit might have some inputs or outputs which we refer to as *dangling wires*. For the spacetime circuit A with some dangling input wires indexed as $1, 2, ..., n$, the notation $A(a_1, a_2, ..., a_m)$, where $n \geq m$, denotes a spacetime circuit derived from A by connecting the wire i to a gate which outputs a_i. For the spacetime circuit B with some output dangling wires indexed as $1, 2, ..., n$, the notation $B \rightarrow a_1, a_2, ..., a_n$ indicates that the i-th wire outputs the state a_i.

We can compose two spacetime circuits A and B into the circuit $\langle A, B \rangle$. We call such a composition an interaction. This means that some dangling input/output wires of A are connected to some dangling output/input wires of B. Such connections have to result in a legal spacetime circuit. The notation itself is ambiguous to which wires are connected, and is to be used in conjunction with a list of connections. We provide here an example of this. Let A a be a spacetime circuit with dangling input wires 1 and 2 and B be a spacetime circuit with dangling output wires 1 and 2. Let $\langle A(a), B \rangle \rightarrow b$ denote an interaction where the only input wire of $A(a)$ is connected to the output wire 1 of B.[6] Note that it is then unambiguous which wire a is given to as input (input wire 1 of A) and as the interaction only has a single dangling output wire, what wire the value b is obtained from (output wire 2 of B).

[5] in their paper Unruh refers to this as an *event* as is the norm in physics, but to avoid confusion with *event* from probability theory we will be using different terminology

[6] we assume here that A and B are circuits where such a connection is legal

Finally, for some region of spacetime P and a spacetime circuit A, we also define the circuit A_P or the circuit A restricted to the region P. This is the circuit which consists only of the gates of A which are located in P. Note that restricting the circuit to a region can both add or remove dangling wires from it. We say that spacetime circuits A and B are identical when restricted to a region P if A_P is identical to B_P and any wires of A and B that were turned into a dangling output wire by the restriction to P lead to a gate at the same point.

3 Security definitions for relativistic commitments

In this section we define a model based on spacetime circuits for relativistic commitment protocols along with security definitions for relativistic bit-commitment protocols. We will be demonstrating that coming up with such definitions is non-trivial by investigating an earlier version of a binding definition we formulated. We will then show that our definitions are useful by proving the security of the relativistic BGKW protocol. Finally we will be comparing BGKW to some toy relativistic commitment protocols that also achieve security according to our definitions and will formulate a metric for determining usefulness of a relativistic commitment protocol.

3.1 Model for relativistic commitment protocols

Definition 2 (Commitment protocol)**.** *Let M (message space[7]) be a set of bit-strings. Let SND (sender) and RCP (recipient) be spacetime circuits, which both have an input wire which accepts as input a bit-string from M. Additionally let the circuit RCP have an output wire which outputs a bit.*

A commitment protocol is the then the tuple $(\mathsf{M}, \mathsf{SND}, \mathsf{RCP})$.

Definition 3 (Correctness)**.** *A commitment protocol $(\mathsf{M}, \mathsf{SND}, \mathsf{RCP})$ is correct for some interaction, if such an interaction exists, such that for all $m \in \mathsf{M}$, $\langle \mathsf{SND}(m), \mathsf{RCP}(m)) \rangle \to 1$.*

We clarify the intuition behind the interaction $\langle \mathsf{SND}(m), \mathsf{RCP}(m') \rangle \to b$. Here m is the message that the sender commits to. The bit m' indicates the bit the recipient uncovers, i.e. the bit they believe the recipient committed to. The bit b then indicates whether the recipient accepts the commitment, it should be 1 if they do and 0 otherwise. Note that the circuits SND and RCP are not limited to dangling wires defined in definition 2, those are just the minimal set of dangling wires they are required to have. It is actually expected that the circuits have numerous additional dangling wires, which are connected by the interaction from definition 3 in a way that models the communication between the sender and the recipient as specified by an actual protocol description.

Generally the models for commitment protocols define two phases usually named *commit* and *unveil*. In the relativistic setting such phases are clearly defined by restricting the circuits to some regions of spacetime. However as we will show these regions are different depending on whether we are discussing the hiding property or the binding property. Therefore, we can not make such a distinction on the level of the model.

Even though our model is restricted to a classical message space, it allows for quantum communication between the parties. While with the provided definition of

[7] while this is called the *message* space, this is actually the set of values that can be committed to not the set of messages that could be exchanged by the parties

correctness, the model excludes protocols where the honest execution of the protocol could result in the recipient rejecting the commitment with some negligible probability, it could be remedied by introducing an additional, more lenient definition for correctness.

3.2 Relativistic security definitions

Having defined a suitable model, we can now define relativistic security, the notions of binding and hiding. For now, we do this only for bit-commitment protocols, i.e. protocols with message space $\{0, 1\}$. We formulate two equivalent definitions for binding and a single definition for hiding.

Let P be some past-closed region of spacetime and $\varepsilon \in \mathbb{R}$ such that $0 \leq \varepsilon \leq 1$. Let $(\{0, 1\}, \mathsf{SND}, \mathsf{RCP})$ be some correct commitment protocol for some interaction $\langle \mathsf{SND}(m), \mathsf{RCP}(m)) \rangle \to 1$.

Definition 4 (ε-binding for P)**.** *Let A be a spacetime circuit that can have any number of dangling input wires leading to gates outside P which are given as input a uniformly randomly determined bit m. Let p denote the maximal probability that for some valid interaction $\langle A(m, ..., m), \mathsf{RCP}(m)) \rangle \to 1$. If no such valid interaction exists we say that $p = 0$. The protocol is ε-binding for P iff for all A, $p \leq \frac{1}{2} + \frac{\varepsilon}{2}$.*

Definition 5 (ε-binding for P)**.** *Let A_0 and A_1 be some spacetime circuits which are identical when restricted to region P. Let p_m denote the maximal probability that for some valid interaction $\langle A_m, \mathsf{RCP}(m)) \rangle \to 1$. If no such valid interaction exists we say that $p = 0$. The protocol is ε-binding for P iff for all (A_0, A_1), $p_0 + p_1 \leq 1 + \varepsilon$.*

Definition 6 (ε-hiding for P)**.** *Let B be a spacetime circuit which outputs a bit. For a uniformly random bit m, let p denote the maximal probability that for some valid interaction $\langle \mathsf{SND}_P(m), B) \rangle \to m$. If no such valid interaction exists we say that $p = 0$. The protocol is ε-hiding for P iff for all B, $p \leq \frac{1}{2} + \frac{\varepsilon}{2}$.*

If a protocol is 0-binding/hiding for P, we say that the protocol is *perfectly* binding/hiding for P.

The intuition behind the region parameter of the binding definitions or simply *the binding region* is that this is the the region of spacetime from which a sender can influence what bit is opened. Outside this region the sender is committed. Therefore the smaller this region is the stronger the binding is.

The *hiding region* indicates two things. Firstly, this is the region from which the sender can decide not to open by not executing some part of the protocol, thus keeping the bit committed a secret indefinitely. Secondly, even if the sender performs the protocol in its entirety, in this region the recipient will be unable to learn what the bit committed is. Contrary to the binding region, the larger the hiding region the stronger the hiding .

Lemma 1. *Definition 4 and definition 5 are equivalent.*

Proof. We start by showing that binding by definition 4 implies binding by definition 5 by constructing an adversary that breaks definition 5 from an adversary that breaks definition 4.

1. Suppose a protocol $(\mathsf{M}, \mathsf{SND}, \mathsf{RCP})$ is not ε-binding for P by definition 5, then there exists A_0 and A_1 for which $p_0 + p_1 > 1 + \varepsilon$.

2. Let A be a circuit which restricted to P is identical to A_0 and A_1 restricted to P. Let A restricted to complement of P simulate A_m where m is some bit given to A at all gates outside the region P.

3. If m is determined uniformly randomly then there exist some valid interaction $\langle A(m, ..., m), \mathsf{RCP}(m)\rangle) \to 1$ with probability $0.5p_0 + 0.5p_1$.

4. $0.5 * p_0 + 0.5 * p_1 = 0.5(p_0 + p_1) > 0.5(1 + \varepsilon) = \frac{1}{2} + \frac{\varepsilon}{2}$

5. Therefore the protocol is also not ε-binding for P by definition 4.

For the protocols to be equivalent it is now sufficient to prove the implication in the opposite direction. We do this using the same technique as before.

1. Suppose a protocol $(\mathsf{M}, \mathsf{SND}, \mathsf{RCP})$ is not ε-binding for P by definition 4, then there exists some A for which $p > \frac{1}{2} + \frac{\varepsilon}{2}$.

2. Let A_0 and A_1 be circuits which restricted to P are identical to A restricted to P. For a bit m let A_m restricted to the complement of P simulate $A(m, ..., m)$.

3. There must now exist some valid interactions $\langle A_m, \mathsf{RCP}(m))\rangle \to 1$ for which $p = 0.5p_0 + 0.5p_1$ or equivalently $2p = p_0 + p_1$.

4. $p_0 + p_1 = 2p > 2(\frac{1}{2} + \frac{\varepsilon}{2}) = 1 + \varepsilon$.

5. Therefore the protocol is also not ε-binding for P by definition 5.

□

3.2.1 Relativistic definitions are non-trivial

The equivalency of the binding definitions might seem trivial to a sceptical reader. Indeed similar definitions for non-relativistic binding are often considered equivalent without further investigation. Consider then an earlier candidate for definition 4 we studied, the only difference being the number of points that A is given m at.

Definition 7. *Let A be a spacetime circuit that can have a single dangling input wire leading to a gate outside P which is given as input a uniformly randomly determined bit m. Let p denote the maximal probability that for some valid interaction $\langle A(m), \mathsf{RCP}(m)) \rangle \to 1$. If no such valid interaction exists we say that $p = 0$. The protocol is ε-binding for P iff for all A, $p \leq \frac{1}{2} + \frac{\varepsilon}{2}$.*

This is actually a weaker notion of binding than definition 4 as illustrated by the following example.

Consider the toy commitment protocol "reveal message twice".

Reveal message twice

1. Let p and p' be two points of spacetime such that neither precedes the other.
2. The sender announces the bit m they wish to commit to at both of these points.
3. The recipient gathers these announced bits and accepts the commitment if they are equal and rejects the commitment otherwise.

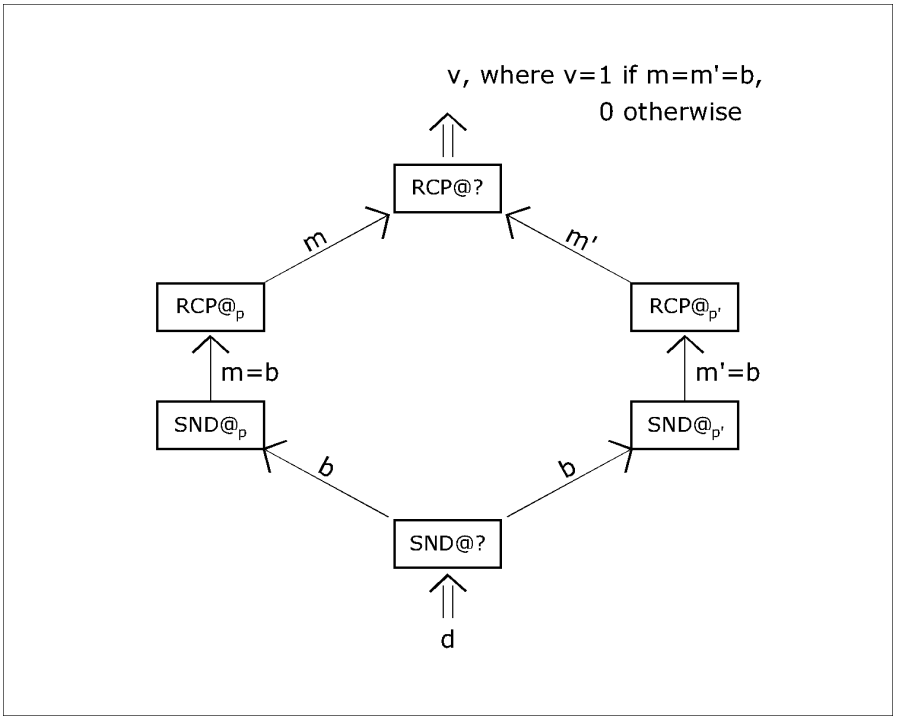


Figure 3. The interaction $\langle \mathsf{SND}(b), \mathsf{RCP}(b')) \rangle \to v$ for reveal message twice

Let SND be a spacetime circuit which at some point preceding both p and p' accepts as input a bit b from wire 1. At points p and p' SND outputs b from wires 1 and 2. Let RCP be a spacetime circuit which at points p and p' expects respectively a bit m and m' to be input from wires 2 and 3. At some point which both p and p' precede RCP expects as input a bit from wire 1 and outputs 1 if this bit is equal to both m and m' and 0 otherwise. As illustrated by figure 3 let $\langle \mathsf{SND}(b), \mathsf{RCP}(b') \rangle \to v$ denote an interaction where the output wire 1 of SND is connected to the input wire 2 of RCP and the output wire 2 of SND is connected to the input wire 3 of RCP. It is trivial to see that the commitment protocol $(\{0,1\}, \mathsf{SND}, \mathsf{RCP})$ is then correct for this interaction.

Ignoring the fact that this protocol is not really hiding in any useful sense, we can still look at whether the protocol is binding for the region $P := (C^-(p) \backslash \{p\}) \cup (C^-(p') \backslash \{p'\})$. By definition 4 it is clearly not binding, as an adversary can simply output whatever bit it is given at both p and p'. By definition 7 it is however binding, as the adversary can not know the input bit at both p and p'.

While this protocol is definitely a contrived example, it demonstrates some hidden complexities that we had to overcome when formulating our definitions. Additionally the fact that even slight differences in wording can have a meaningful definition on a relativistic definition is another indication that studying such protocols requires a well defined formal framework.

3.3 Relativistic BGKW is secure

To prove the security of relativistic BGKW we first have to define spacetime circuits which model relativistic BGKW and also fit our definitions. To keep the proof cleaner we use here a slightly modified version of relativistic BGKW where the sender skips the step where they directly reveal the bit they committed to (as specified by step 6. of the protocol). Practically this would mean that if the recipient happens to generate the challenge to be $0...0$ they will be unable to uncover what the bit committed actually was, but otherwise the usefulness of the protocol remains unchanged. As this version requires the sender to reveal less data, this means that any proof that this version of the relativistic BGKW is binding would also imply that the full version of relativistic BGKW is binding.

Let RCP be a circuit with input wires 1, 2, 3 and output wires 1 and 2. At point p'_d RCP outputs a uniformly randomly determined n-bit string d from wire 1. At points p'_c and p'_u respectively RCP expects an input n-bit strings c and u from wires 2 and 3. At some point which p'_d, p'_u and p'_c all precede RCP expects as input a bit m' from wire 1 and outputs from wire 2 a bit which is 1 if $c = m' \cdot d \oplus u$ and 0 otherwise. In addition, matching the protocol description RCP selects p'_d such that it does not precede p'_u.

Let SND be a circuit with input wires 1, 2 and output wires 1 and 2. SND generates a uniformly random n-bit string at some point preceding points p_c and p_u and accepts as input from wire 1 a bit m. At point p_d which p'_d precedes SND expects as input d from

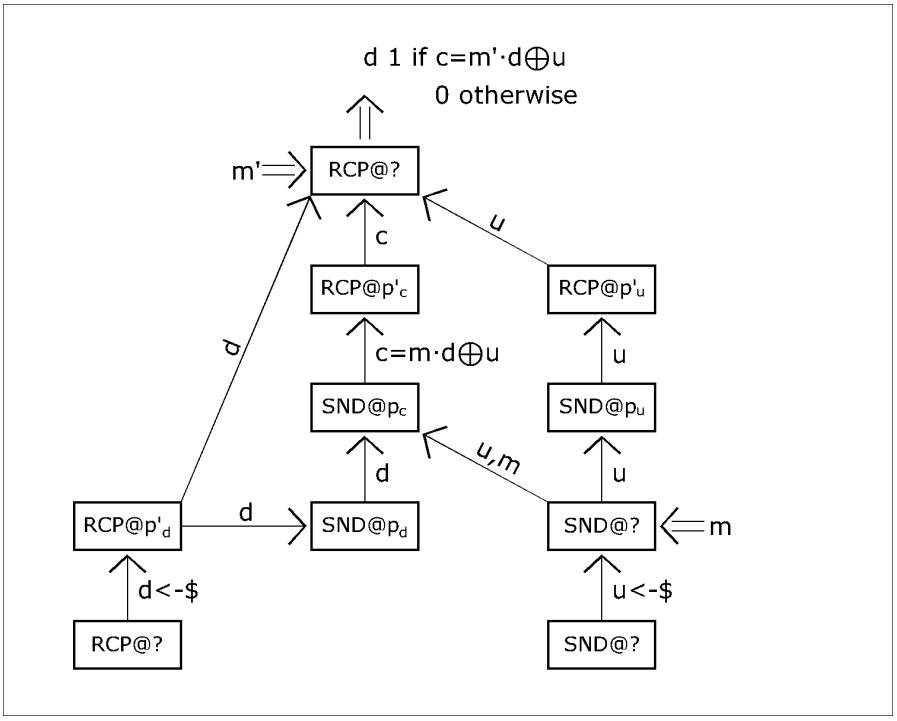


Figure 4. The interaction $\langle \mathsf{SND}(b), \mathsf{RCP}(b')) \rangle \to v$ for relativistic BGKW

wire 2. At point p_c preceding p'_c SND outputs $c := m \cdot d \oplus u$ from wire 1. At point p_u preceding p'_u SND outputs u.

Let the interaction $\langle \mathsf{SND}(b), \mathsf{RCP}(b')) \rangle \to v$ be defined as in the figure 4. It is trivial to see that the commitment protocol $(\{0,1\}, \mathsf{SND}, \mathsf{RCP})$ is then correct for this interaction.

Theorem 1. *Relativistic BGKW is perfectly hiding for* $P_h := (C^-(p_u) \setminus \{p_u\})$.

Proof. Note that as $p_u \notin P_h$, then the only output of SND_{P_h} is the commitment c. For any adversary that does not know u, c is indistinguishable from an uniformly random output. Given that u is not revealed at any point, the best an adversary could do is to guess the bit committed to with probability 0.5. Therefore the protocol must be perfectly hiding for P_h. □

Note that we have proven perfect hiding for a maximal region. The only point that could be added to P_h while keeping the resulting region past-closed is p_u. For such a region the protocol would clearly not be hiding. An alternative region which hiding could be proven for is $(C^-(p_c) \setminus \{p_c\})$, but this is trivial as this cuts out the part of SND which performs the commitment.

Theorem 2. *Relativistic BGKW is* $\frac{2}{\sqrt{2^n+1}}$*-binding for* $P_b = C^-(p'_c)$*, where* n *is the length of the challenge.*

Proof. Let $A(m, ..., m)$ denote some spacetime circuit constructed by giving some circuit A an uniformly randomly determined bit m as input in any points in region $\overline{P_b}$. Without loss of generality, let $A(m, ..., m)$ dangling wires such that in the interaction $\langle A(m, ..., m), \mathsf{RCP}(m)\rangle \to b$, $A(m, ..., m)$ accepts as input from $\mathsf{RCP}(m)$ the challenge d and outputs to $\mathsf{RCP}(m)$ a commitment c and an opening u.

We split A into four subcircuits using the regions $P_d := C^+(p'_d)$ and P_b. The important intuition here is that in the region $\overline{P_d}$ A does not have access to the challenge d and from the region $\overline{P_b}$ the adversary can not influence the commitment c.

- $A1_{pre}$ - A restricted to $P_b \cap \overline{P_d}$
- $A1_{post}$ - A restricted to $P_b \cap P_d$
- $A2_{pre}$ - A restricted to $\overline{P_b} \cap \overline{P_d}$
- $A2_{post}$ - A restricted to $\overline{P_b} \cap P_d$

We then have that the following wires are not possible according to the spacetime restrictions (these restrictions are essentially summarized by noting that information can not flow in the direction $2 \to 1$ nor in the direction $post \to pre$):

- $A1_{post} \not\to A1_{pre}, A2_{pre}$
- $A2_{pre} \not\to A1_{pre}, A1_{post}$
- $A2_{post} \not\to A1_{pre}, A1_{post}, A2_{pre}$

Either $A1_{pre}$ or $A1_{post}$ has to be the circuit which outputs c to RCP's gate at point p'_c as no points in $A2_{pre}$ or $A2_{post}$ precede p'_c. Now note that the point p'_c is located in the region $P_b \cap P_d$, the region which $A1_{post}$ is restricted to. Also note that all points in $P_b \cap \overline{P_d}$, the region where $A1_{pre}$ is restricted to, precede p'_c. Therefore for any adversary what would output c from $A1_{pre}$, there exists an adversary that forwards it to p'_c and has $A1_{post}$ output it. Without loss of generality let $A1_{post}$ output c.

Note that in the model we specified that RCP selects p'_d such that it does not precede p'_u. This means that any that p'_d precedes would also not precede p'_u which is another way of saying that no point in P_d precedes p'_u. As RCP expects to receive u at p'_u this means that either $A1_{pre}$ or $A2_{pre}$ has to output it. Now note that the point p'_u is not located in the region P_b as the opening is revealed after the commitment. Therefore p'_u must be in the region $\overline{P_b} \cap \overline{P_d}$ or in other words the region where $A2_{pre}$ is restricted to. A can only possibly output u from a point preceding p'_u, therefore once again, for any adversary that would output u from $A1_{pre}$, there exists an adversary that forwards it to p'_c and has $A2_{pre}$ output it. Then without loss of generality let $A2_{pre}$ output u.

We can now have confidence, that the quantum circuit on figure 5 models any successful adversary. As the sub-circuit $A2_{post}$ does not have any outputs, it gives an

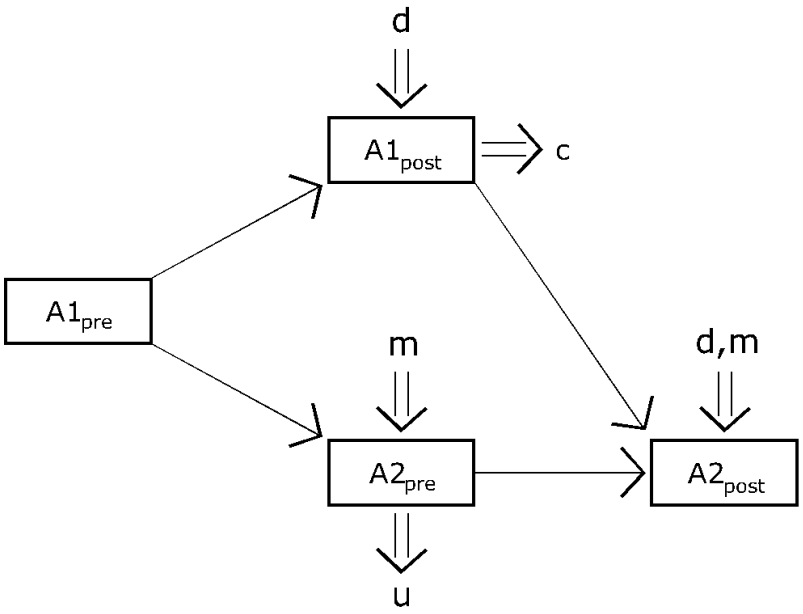


Figure 5. Splitting the circuit A into a quantum circuit

adversary no additional power, we can remove it without loss of generality, resulting in the circuit from figure 6.

Recall that the goal of an adversary is to output u and c such that $c = m \cdot d \oplus u$ or equivalently such that $c \oplus u = m \cdot d$. As pointed out by Lunghi et al [LKB$^+$15] this is actually the $CHSH_n$ game. As proved by Sikora et al. [SCK14] the probability of winning this game is bound by $\frac{1}{2} + \frac{1}{\sqrt{2^{n+1}}}$. □

Note that binding for relativistic BGKW can only be proven for regions which do not include a point that precedes p'_c. This then means that P_b is the minimal region for which relativistic BGKW is binding as this region includes no points which do not precede p'_c.

3.4 Interaction between the binding and hiding regions

Consider the trivial protocol "reveal message".

Reveal message

1. The sender reveals the bit they wish to commit.

At first glance this protocol might seem useless, but it turns out that this protocol is both perfectly hiding and binding according to our definitions. Namely if the point of

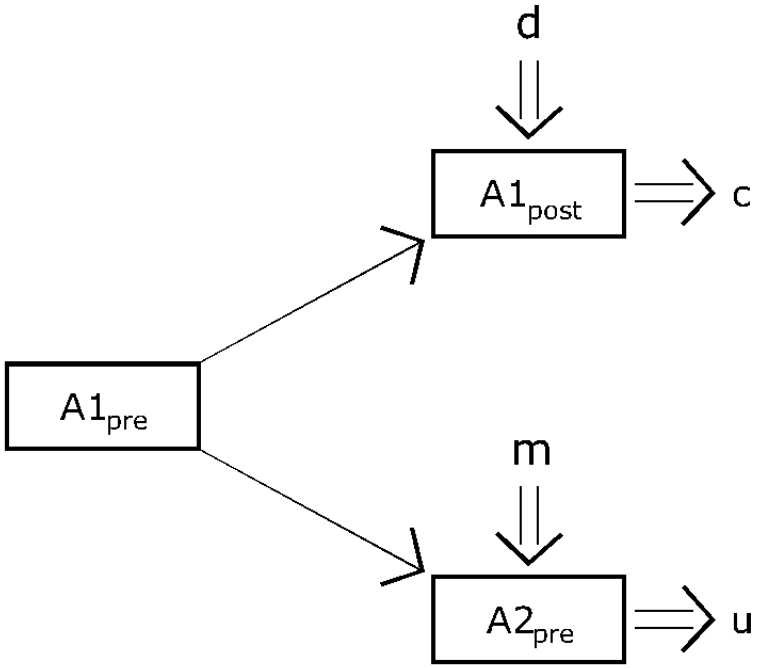


Figure 6. Reduction of A equivalent to the $CHSH_n$ game

spacetime at which the sender reveals their message is p, then the protocol is perfectly hiding for $P_h = C^-(p) \setminus \{p\}$ and perfectly binding for $P_b = C^-(p)$. Showing this is trivial, the sender restricted to P_h has no outputs, therefore they give an adversary no advantage. Similarly outside P_b the sender does not output anything, therefore an adversary given some bit m outside P_b has no outputs it could influence to open to m.

This is quite counterintuitive, how could this protocol in some sense be just as hiding and binding as BGKW [8]. More formally there seems to be no obvious difference between the binding regions of BGKW and "reveal message", and the same holds for their hiding regions. However we should consider the security as a whole, or in other words, if a protocol is binding for P_b and hiding for P_h what are the regions $P_b \setminus P_h$ and $P_h \setminus P_b$. For "reveal message" these evaluate respectively to $\{p\}$ and $\{\}$. For relativistic BGKW these evaluate to $C^-(p'_c) \setminus C^-(p_u)$ and $C^-(p_u) \setminus (C^-(p'_c) \cup \{p_u\})$ which seem much more promising regions.

To investigate this notion further consider a different toy protocol "simple relativistic commitment".

[8] although we did not prove BGKW to be perfectly binding, doing this for a larger region that includes p'_u is trivial

Simple relativistic commitment

1. Let p and p' be two points of spacetime such that neither precedes the other.
2. The sender determines uniformly randomly a bit u.
3. At point p the sender reveals $c = m \oplus u$, where m is the bit they wish to commit to.
4. At point p' the sender reveals u
5. The recipient gathers c and u and calculates $m = c \oplus u$ to unveil the commitment.

At first glance this protocol seems much more promising than "reveal message". It should be trivial to see that this protocol is perfectly hiding for both regions $C^-(p) \setminus \{p\}$ and $C^-(p') \setminus \{p'\}$. The (perfect) binding region for this protocol however turns out to be the trivial region $P_b = C^-(p) \cup C^-(p')$ which similarly to the binding region of "reveal message" contains the entire protocol from the sender's perspective. Adding any points to this region would result in a region that is not binding at all. For example if an adversary would be asked to commit to bit m which is given to them at point p', they could simply output 0 at p and m at p'. Thus this is a maximal binding region.

As the binding regions are "symmetrical", using either would give a similar end result, let $P_h = C^-(p') \setminus \{p'\}$. Then $P_b \setminus P_h = (C^-(p) \setminus C^-(p')) \cup \{p'\}$, which is quite similar to the respective region of BGKW. On the other hand $P_h \setminus P_b = \{\}$, the same result we had for "reveal message".

It is not clear whether the simple relativistic commitment is less useful than relativistic BGKW. Recall however Kent's method for extending RBC1, that is that instead of opening a party could commit to the opening to extend the commitment. This method does not work for "simple relativistic commitment", but seems to work for relativistic BGKW. We conjecture then that for a protocol to be useful it must be binding for region P_b and hiding for region P_h such that neither the region $P_b \setminus P_h$ nor the region $P_h \setminus P_b$ are empty.

4 Relativistic collapse-binding

In the previous section we formulated definitions for relativistic biding and hiding but multiple issues where left unsolved. Firstly, the definitions only apply to bit-commitment protocols, and it is not trivial to generalize the definitions for protocols with larger message spaces. Secondly, we have gained no insight into how these relativistic protocols could be used as building blocks in some larger protocols. As a step toward solving these issues, in this section we explore a different approach to binding, collapse-binding.

4.1 Background

Collapse-binding is a definition for computationally binding quantum commitment schemes introduced by Unruh [Unr16]. It has multiple properties that we are interested in achieving for a relativistic definition. It is parallelly composable, meaning that it is possible to construct protocols with larger message spaces from a collapse-binding bit-commitment protocol. Additionally Unruh also showed how to construct zero-knowledge quantum arguments of knowledge using a collapse-binding commitment protocol, indicating that the definition is useful when considering composition.

Collapse-binding is a definition for the non-interactive setting. Essentially, this means that commitment protocols are modeled as a pair of functions (*com, verify*). The function *com* applies to some message m and outputs a commitment c, i.e this is the function that the sender uses to produce a commitment. $Verify$ is the function that the recipient uses to verify an opening u, it takes as input m, c and u and outputs 1 if the tuple is a valid commitment and 0 otherwise. For context we provide Unruh's [Unr16] definition here:

Definition 8. *For algorithms A, B consider the following games:*

$$Game_1 : (S, M, U, c) \leftarrow A, ok \leftarrow V_c(M, U), m \leftarrow M_{ok}(M), b \leftarrow B(S, M, U)$$

$$Game_2 : (S, M, U, c) \leftarrow A, ok \leftarrow V_c(M, U), \qquad b \leftarrow B(S, M, U)$$

Here S, M, U are quantum registers. V_c is a measurement whether M, U contains a valid opening, formally V_c is defined through the projector $\sum_{\substack{m,u \\ verify(c,m,u)=1}} |m\rangle\langle m| \otimes |u\rangle\langle u|$.

M_{ok} a measurement of M in the computational basis if $ok = 1$, and does nothing if $ok = 0$ (i.e. it sets $m :=\perp$ and does not touch the register M).

A commitment scheme is collapse-binding iff for any quantum-polynomial-time algorithms A, B, the difference $|Pr[b = 1 : Game_1|| - |Pr[b = 1 : Game_2||$ is negligible.

Essentially the goal of an adversary is to be able to output some M which is a non-trivial superposition of some messages, such that the V_c measurement does not collapse

M into a trivial superposition by measuring it. If such a collapse would happen, then the measurement M_{ok} would have no effect, thus the games could not be differentiated by the adversary. In other words, these games test whether the adversary is capable of finding some commitment c which they can open to different messages.

4.2 Relativistic collapse-binding

Translating the collapse-binding definition into the relativistic setting comes with multiple challenges. Relativistic protocols are inherently interactive, therefore the model used for the non-relativistic collapse-binding definitions does not suit us. Fortunately, Unruh also discussed how his definition could be generalized for interactive commitments. We can therefore combine their method with our model for relativistic commitments. A second issue we have is that, the $verify$ function which the collapse-binding games use in the V_c projector, requires a non-quantum commitment and opening. Although Unruh also briefly discussed how to construct a definition for quantum protocols, for the sake of simplicity, as the protocols we have described in this thesis are all classical, our relativistic collapse-binding definition will be restricted to classical protocols. Finally, note that Unruh's definition is for the computational setting, our goal is to construct a definition for the information theoretical setting.

As as side note, Unruh did not investigate their interactive definition further. They neither proved a protocol secure according to this definition nor showed that a protocol fulfilling their definition would be useful as a building block for zero-knowledge arguments. Therefore our proof for BGKW using our relativistic collapse-binding definition also implies that Unruh's interactive definition is at least somewhat useful.

For constructing a relativistic collapse-binding definition we now need to overcome two issues, within the definition we need to model the relativistic constraints, and we also need to define the function $verify$ that a collapse-binding definition requires. To define $verify$ consider a commitment protocol $(\mathsf{M}, \mathsf{SND}, \mathsf{RCP})$ which is correct for some interaction $\langle \mathsf{SND}(m), \mathsf{RCP}(m')) \rangle \to b$. This interaction comes with some list of wires which connect SND and RCP. To use this interaction within a collapse-binding definition we require these wires to transmit only classical information. Secondly, we require these wires to be additionally designated either as *commitment* or *opening* wires such that gates connected to an opening wire must not precede any gates connected to a commitment wire and that opening wires are only in the direction SND to RCP. For example, in the BGKW protocol the commitment wires would be the wires which transmit the challenge and the commitment, and the opening wire would be the wire which transmits the opening. Collectively we denote the commitment wires with w_c and the opening wires w_u, the tuples which are transmitted via these wires we denote respectively with c and u. There must then exist some function $verify$ such that $verify(m', c, u) = b$, intuitively this would be the function which RCP uses to determine whether to accept or reject the commitment. We then say that $verify$ is a verification function

extracted from the commitment protocol $(\mathsf{M}, \mathsf{SND}, \mathsf{RCP})$ correct for some interaction $\langle \mathsf{SND}(m), \mathsf{RCP}(m')) \rangle \to b$ for commitment wires w_c and opening wires w_u.

Definition 9. *Let $verify$ be a verification function extracted from the commitment protocol $(\mathsf{M}, \mathsf{SND}, \mathsf{RCP})$ which is correct for some interaction $\langle \mathsf{SND}(m), \mathsf{RCP}(m')) \rangle \to b$ for commitment wires w_c and opening wires w_u. Then let P be the region defined as the union of causal pasts of all output gates of wires w_c[9]. For some quantum circuits A which connects to all of the dangling wires of RCP_P and B, and a past closed region P' consider then the following games:*

$$Game_1 : (S, M, U) \leftarrow \langle A, \mathsf{RCP}_P \rangle, ok \leftarrow V_c(M, U), m \leftarrow M_{ok}(M), b \leftarrow B(S, M, U)$$

$$Game_2 : (S, M, U) \leftarrow \langle A, \mathsf{RCP}_P \rangle, ok \leftarrow V_c(M, U), \qquad b \leftarrow B(S, M, U)$$

Here S, M, and U are output by A within the region P'.

The value c is obtained from the wires w_c. V_c is a measurement whether the state (M, U) would pass the verification function. This is formally defined by the projector $\sum_{\substack{m,u \\ verify(c,m,u)=1}} |mu\rangle \langle mu|$. It outputs $ok = 1$ if the measurement succeeds and $ok = 0$ otherwise.

M_{ok} performs a measurement of M in the computational basis when $ok = 1$ and does nothing when $ok = 0$.

A commitment protocol is ε-relativistic collapse-binding for P' iff for all adversaries (A, B) the difference $|Pr[b = 1 : Game_1] - Pr[b = 1 : Game_2]| \leq \varepsilon$.

4.3 Relativistic BGKW and relativistic collapse-binding

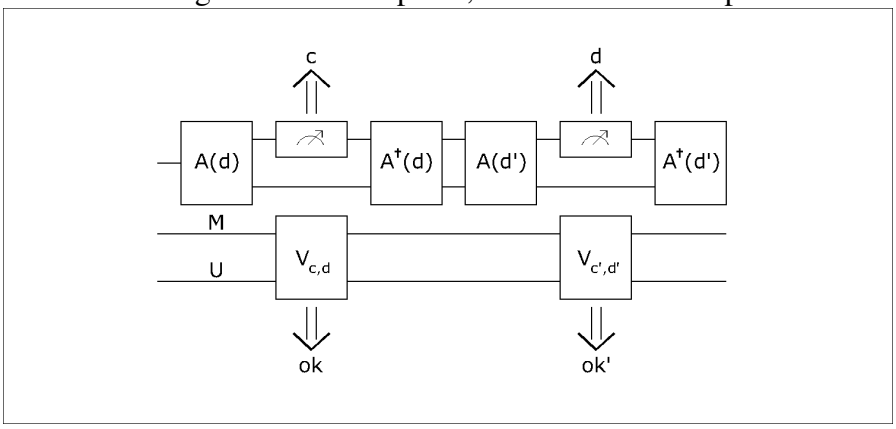
In this section we demonstrate that there is no adversary that can *completely break* the relativistic collapse-binding games for relativistic BGKW. By a complete break we mean that there exists no successful adversary for which the verification measurement always outputs $ok = 1$. We believe that this implies that it is provable that for some non-trivial region P', relativistic BGKW is ε-relativistic collapse-binding for P', where ε is negligible.

To prove this we must first discuss how to extract $verify$ for use within the collapse-binding games. Recall the spacetime circuits RCP and SND we defined in section three which model the relativistic BGKW protocol. These circuits are connected by three wires. Let the wires which transmit the challenge c and the commitment d be designated as the *commitment* wires w_c. Then let the wire which transmits the opening be designated as the *opening* wire w_u. Note that if p'_c is the point at which RCP expects to receive the commitment, then $P = C^-(p'_c)$ is the union of causal pasts of all output gates of wires w_c. This follows from the fact that the wire which transmits the challenge precedes the wire

[9] by output gates we mean here the gates to which the wires lead to

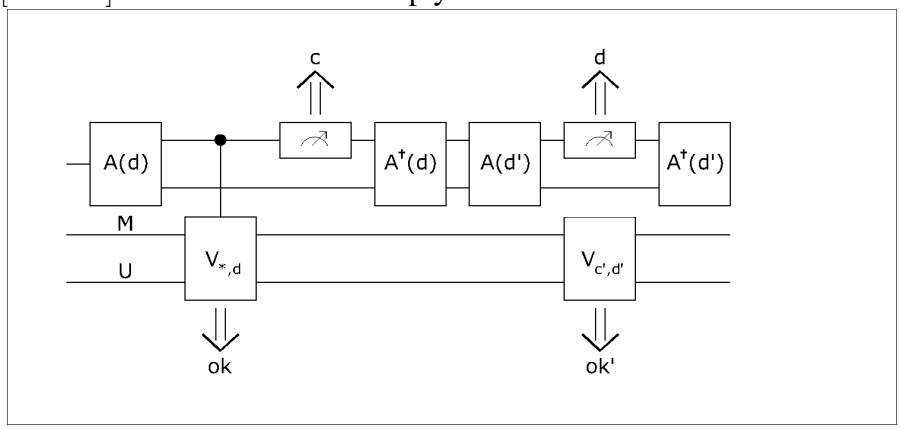
which transmits the commitment. Now let $verify(m, (c, d), u)$ output 1 if $c = m \cdot d \oplus u$ and 0 otherwise. The measurement $V_{c,d}$ in the collapse-binding games for relativistic BGKW is then defined by the projector $|0c\rangle \langle 0c| + |1(c \oplus d)\rangle \langle 1(c \oplus d)|$.

Before we can get to our main proof, we first define and prove a useful lemma.

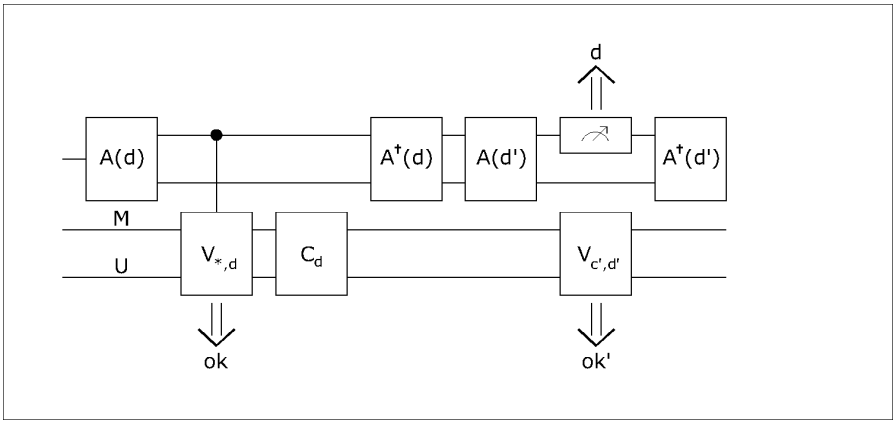


Lemma 2. *Consider the circuit above where A is some arbitrary quantum circuit that takes a bit-string as input and $V_{c,d}$ is a measurement defined by the projector $|0c\rangle \langle 0c| + |1(c \oplus d)\rangle \langle 1(c \oplus d)|$. For any initial state and where inputs d and d' are determined uniformly randomly such that $d \neq d'$, if in this circuit $V_{c,d}$ outputs $ok = 1$ with probability, then $V_{c'd'}$ also outputs $ok' = 1$ with probability 1.*

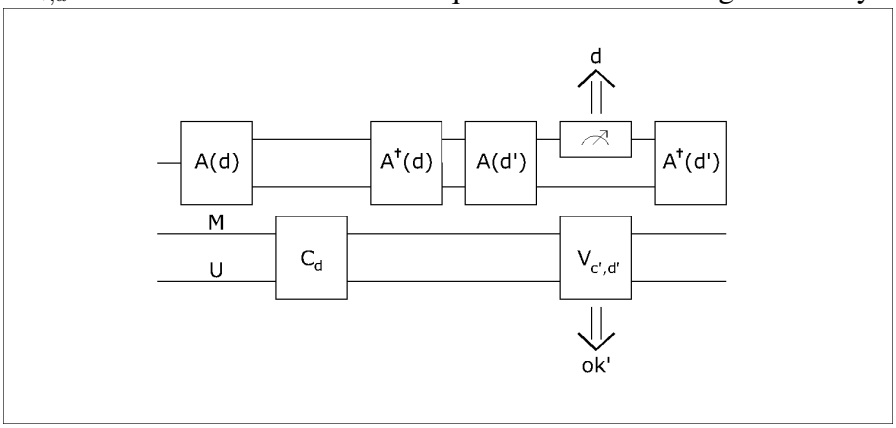
Proof. We prove this lemma using a series of transformations which do not change $Pr[ok' = 1]$. The final transformation will result in a circuit for which it is clear that $Pr[ok' = 1] = 1$. This will then imply that the same also holds for the circuit above.



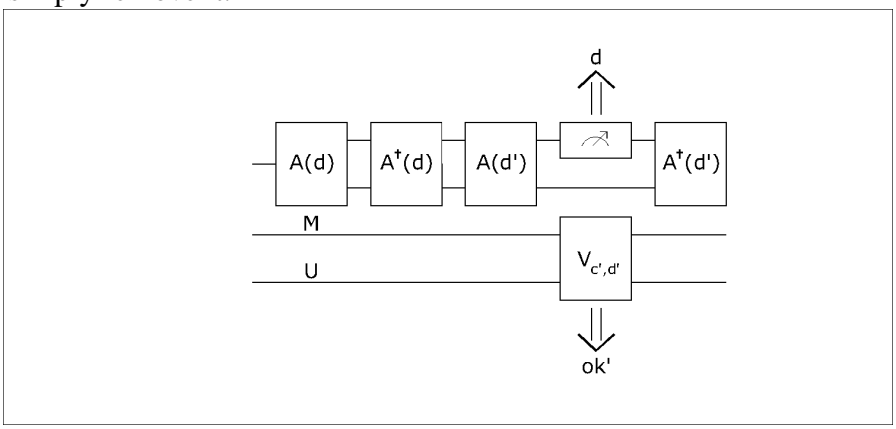
In the first transformation we replace the $V_{c,d}$ circuit with the $C-V_{*,d}$ circuit. $C-V_{*,d}$ is formally defined by the projector $\sum_c |c\rangle \langle c| \otimes V_{c,d}$. From $Pr[ok = 1] = 1$ we know that regardless of the result of the measurement c, $V_{c,d}$ did not change the state of the system. Therefore $C - V_{*,d}$ also has no effect on the state of the system which in turn allows us to perform the replacement without influencing $Pr[ok' = 1]$.



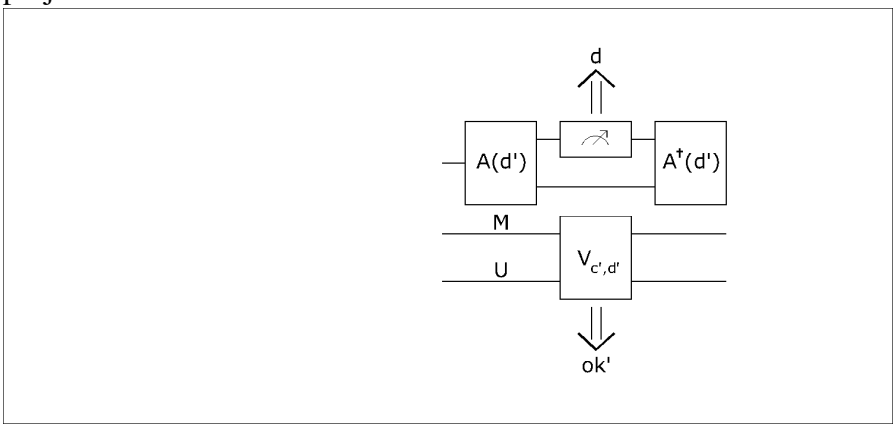
For the second transformation we replace the measurement c on wire C with the circuit C_d on wires MU. C_d is defined by a set of $2^{|C|}$ projectors, where the c-th projector has the form $|0c\rangle\langle 0c| + |1(c \oplus d)\rangle\langle 1(c \oplus d)|$. From the state of the system after the $C - V_{*,d}$ circuit it is clear that this is equivalent to measuring C directly.



For the third transformation we remove the $C - V_{*,d}$ circuit. We have already seen that $C - V_{*,d}$ has no effect on the state of the system. Now that we no longer need it we can simply remove it.



In the fourth transformation we switch the order of C_d and $V_{c',d'}$. For this it is sufficient to see that these circuits commute which can be trivially shown by multiplying the projectors which define these circuits.

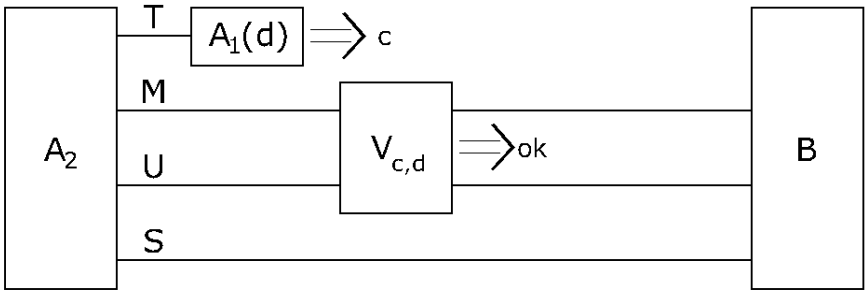


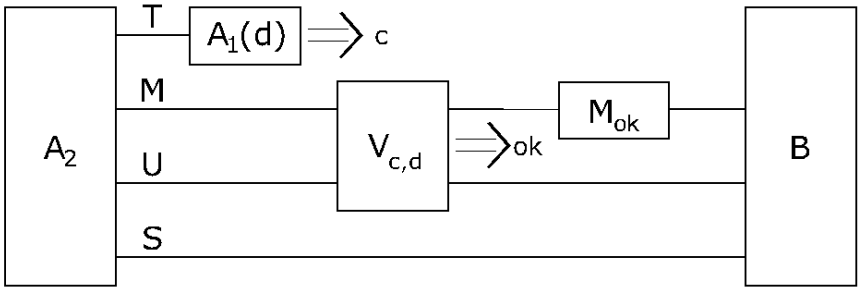
In the final transformation we remove $A_1(d)$ and its inverse $A_1^\dagger(d)$ which applied together are simply the identity circuit. Note that the final circuit we have obtained is the same as the circuit which produced $ok = 1$ with probability 1. Therefore this circuit must also produce $ok' = 1$ with probability 1. □

Having this lemma as an useful tool we can then get to the main theorem of this secion.

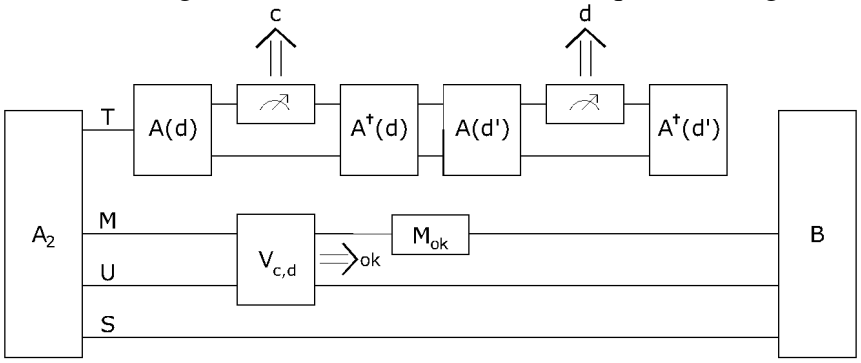
Theorem 3. *There exists no adversary (A, B) which can differentiate the collapse-binding games of relativistic BGKW for region $P' = \overline{C^+(p'_d)}$, such the $V_{c,d}$ measurement in the games always outputs $ok = 1$.*

Proof. We split A into two subcircuits, such that A_2 is A restricted to P' and A_1 is A restricted to $\overline{P'}$. Note that the challenge d is then given as input to the circuit A_1 and the registers S, M and U have to be output by A_2. As the point p'_c is within $\overline{P'}$, we can say without loss of generality that A_1 outputs c. The following quantum circuits then illustrate the adversary ((A_1, A_2), B) in games 1 and 2 of the collapse-binding games of relativistic BGKW for region $\overline{P'}$.

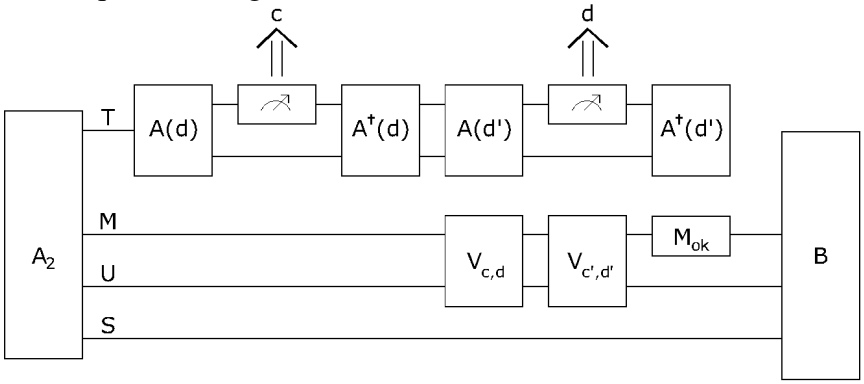




We now describe a series of games starting with game 2. Each of these are equivalent from the perspective of the adversary $((A_1, A_2), B)$. The circuit below depicts game 3 which is obtained from game 2. Note that here the $A(d)$ circuit and the measurement on wire C that directly follows it are the circuit A_1 with input d. As game is then just game 2 with some extra gates added to the end of it, it is equivalent to game 2.



On the next circuit we can see the game 4. Here we have added the measurement $V_{c',d'}$ which, according to lemma 2 and the assumption that the measurement $V_{c,d}$ always outputs $ok = 1$, always succeeds, i.e it does not affect the state of the system. Therefore game 3 is equivalent to game 4.



The state of the system after the measurement $V_{c,d}$ must be $\sum_{t,s} a_{t,s} |t\rangle |0c\rangle |s\rangle +$

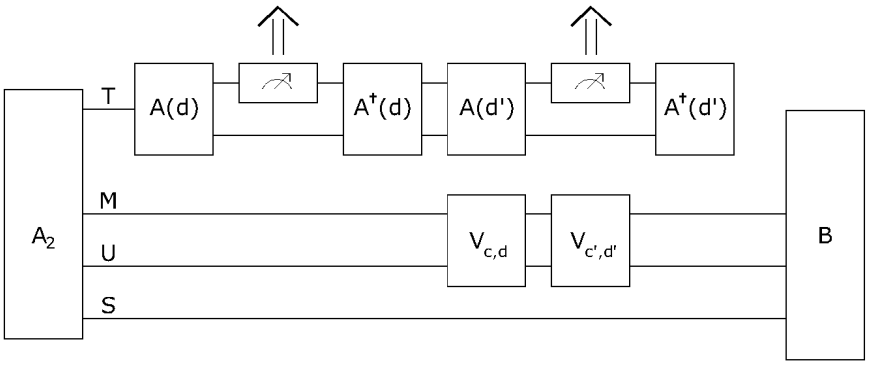
$b_{t,s} |t\rangle |1(c \oplus d)\rangle |s\rangle$.

From $Pr[ok' = 1] = 1$ we have that the state of the system before the measurement $V_{c',d'}$ must be $\sum_{t,s} a'_{t,s} |t\rangle |0c'\rangle |s\rangle + b'_{t,s} |t\rangle |1(c' \oplus d')\rangle |s\rangle$.

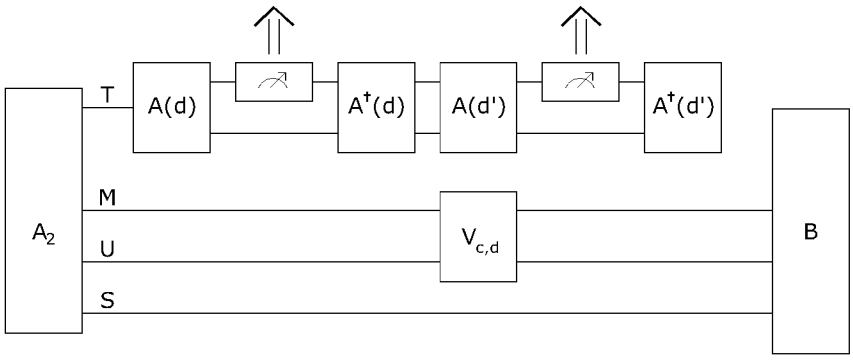
Therefore $\sum_{t,s} a_{t,s} |t\rangle |0c\rangle |s\rangle + b_{t,s} |t\rangle |1(c \oplus d)\rangle |s\rangle = \sum_{t,s} a'_{t,s} |t\rangle |0c'\rangle |s\rangle + b'_{t,s} |t\rangle |1(c' \oplus d')\rangle |s\rangle$. This means that either $c = c'$ or all of the $a_{t,s}$ equal 0. Likewise either $c \oplus d = c' \oplus d'$ or all of the $b_{t,s}$ equal 0.

Note that $c \oplus d = c' \oplus d'$ and $d \neq d'$ implies that $c \neq c'$. Therefore either all of the $a_{t,s}$ equal 0 or all of the $b_{t,s}$ equal 0. In other words the state on wire M is either $|0\rangle$ or $|1\rangle$.

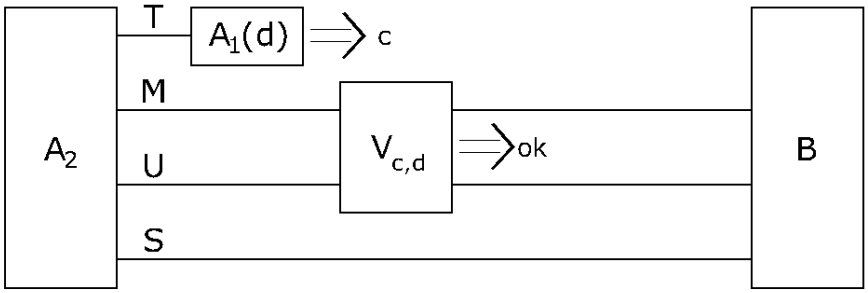
Therefore let game 5 be game 4 with the M_{ok} omitted. As M was already measured this has no effect.



We can then remove the circuit $V_{c',d'}$ for game 6. Recall that it did not influence the state of the system.



Finally we replace the top two wires with the circuit with A_1 again, which actually gives us game 1. Therefore we have shown that for the adversary $((A_1, A_2), B)$ are equivalent.



□

4.4 Usefulness of relativistic collapse-binding

Consider the toy protocols "reveal message" and "simple relativistic commitment" we investigated in section 3.4. These are trivially impossible to be proven as collapse-binding. To see this, consider that as the recipient in these protocols always accepts the commitment, any verification function extracted from these protocols would always have to output 1. Thus the verification measurement V_c would essentially be a no-operation. Then any state that an adversary would output as M would pass verification, thus the adversary could have the maximal advantage at differentiating the games. Therefore the collapse-binding definition is able to exclude some useless protocols, that were provable to be binding by our previous definitions.

On the other hand, it seems that if a protocol is relativistic collapse-binding for some region P and relativistic binding for some region P', these regions are not equal. This is definitely the case for BGKW. Furthermore we were able to give an intuitive meaning to the binding region P', for the region P we are unable to do this. Note however, that it seems still to be possible to derive the binding region from the collapse-binding definition. At least for BGKW this was exactly the region which was the causal past of all of the wires labeled as commitment wires. We conjecture that this is also true for other protocols.

5 Open issues

We note here a list of open issues that were introduced but left unsolved by this thesis.

- The usefulness of security definitions from section three could be studied further by trying to prove security for a multi-round relativistic commitment protocol.
- The security definitions from section three could be generalized for commitment protocols other than bit-commitment protocols.
- The relation between the hiding and binding regions could be thoroughly analyzed. Currently we only superficially looked into this issue, but there seems to be some hidden complexities in this as illustrated by section 3.4.
- The properties that the non-relativistic collapse-binding definition has (e.g. parallel composition) could be proven or disproven for the relativistic collapse-binding definition.
- While we proved that there exists no adversary that completely breaks collapse-binding for BGKW, it remains to be proven relativistic that it is actually collapse-binding.

References

[BCJL93] Gilles Brassard, Claude Crépeau, Richard Jozsa, and Denis Langlois. A quantum bit commitment scheme provably unbreakable by both parties. In *Proceedings of 1993 IEEE 34th Annual Foundations of Computer Science*, pages 362–371. IEEE, 1993.

[BCMS97] Gilles Brassard, Claude Crépeau, Dominic Mayers, and Louis Salvail. A brief review on the impossibility of quantum bit commitment. *arXiv preprint quant-ph/9712023*, 1997.

[BOGK88] M Ben-Or, S Goldwasser, and J Kilian. A. wigderson. multi-prover interactive proofs: How to remove intractability. In *Proc*, pages 113–131, 1988.

[CSST11] Claude Crépeau, Louis Salvail, Jean-Raymond Simard, and Alain Tapp. Two provers in isolation. In *International Conference on the Theory and Application of Cryptology and Information Security*, pages 407–430. Springer, 2011.

[Ken99] Adrian Kent. Unconditionally secure bit commitment. *Physical Review Letters*, 83(7):1447, 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(13):130501, 2012.

[KTHW13] Jędrzej Kaniewski, Marco Tomamichel, Esther Hänggi, and Stephanie Wehner. Secure bit commitment from relativistic constraints. *IEEE Transactions on Information Theory*, 59(7):4687–4699, 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(3):030502, 2015.

[May97] Dominic Mayers. Unconditionally secure quantum bit commitment is impossible. *Physical review letters*, 78(17):3414, 1997.

[PMM+17] Christopher Portmann, Christian Matt, Ueli Maurer, Renato Renner, and Björn Tackmann. Causal boxes: quantum information-processing systems closed under composition. *IEEE Transactions on Information Theory*, 63(5):3277–3305, 2017.

[SCK14] Jamie Sikora, André Chailloux, and Iordanis Kerenidis. Strong connections between quantum encodings, nonlocality, and quantum cryptography. *Physical Review A*, 89(2):022334, 2014.

[Unr14] Dominique Unruh. Quantum position verification in the random oracle model. In *Annual Cryptology Conference*, pages 1–18. Springer, 2014.

[Unr16] Dominique Unruh. Computationally binding quantum commitments. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 497–527. Springer, 2016.

[VPdR19] Venkatesh Vilasini, Christopher Portmann, and Lídia del Rio. Composable security in relativistic quantum cryptography. *New Journal of Physics*, 21(4):043057, 2019.

Appendix

I. Licence

Non-exclusive licence to reproduce thesis and make thesis public

I, **Mattias Lass**,

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,

   Post-Quantum Security Definitions for Relativistic Commitment Protocols, supervised by Dominique Unruh.

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Mattias Lass
11/11/2021