Secure Bit Commitment From Relativistic Constraints

Jędrzej Kaniewski, Marco Tomamichel, Esther Hänggi and Stephanie Wehner

Abstract

We investigate two-party cryptographic protocols that are secure under assumptions motivated by physics, namely special relativity and quantum mechanics. In particular, we discuss the security of bit commitment in so-called split models, i.e. models in which at least one of the parties is not allowed to communicate during certain phases of the protocol. We find the minimal splits that are necessary to evade the Mayers-Lo-Chau no-go argument and present protocols that achieve security in these split models. Furthermore, we introduce the notion of local versus global command, a subtle issue that arises when the split committer is required to delegate non-communicating agents to open the commitment. We argue that classical protocols are insecure under global command in the split model we consider. On the other hand, we provide a rigorous security proof in the global command model for Kent’s quantum protocol [1]. The proof employs two fundamental principles of modern physics, the no-signalling property of relativity and the uncertainty principle of quantum mechanics.

Index Terms

bit commitment, special relativity, quantum theory.

I. INTRODUCTION

The goal of two-party cryptography is to enable two parties, Alice and Bob, to solve a task in cooperation even if they do not trust each other. An example of such a task is the cryptographic primitive known as bit commitment. A bit commitment protocol traditionally consists of two phases: In the commit phase, Bob commits a bit to Alice, who receives some form of confirmation that a commitment has been made. In the open phase, Bob reveals the bit to Alice. Security means that Bob should not be able to reveal anything but the committed bit, but nevertheless Alice cannot gain any information about the bit before the open phase. While many two-party cryptographic primitives have been defined, oblivious transfer and bit commitment are undoubtedly among the most important ones because they form essential building blocks for more complex problems [2].

Ideally, we would like to have protocols for such primitives that guarantee security without relying on any subjective (e.g. that a safe is difficult to open) or computational (e.g. that factoring a product of two large primes is difficult) assumptions. Unfortunately, however, it turned out that this is impossible, even if we allow quantum communication between Alice and Bob [3], [4], [5], [6]. Much work has thus been invested into determining what kind of assumptions allow us to obtain security. Of particular interest to this work are thereby assumptions of a physical nature, leading to information-theoretic security. Classical examples of such assumptions are, for example, access to some very special forms of shared randomness supplied in advance [7], access to a noisy communication channel [8], [9] or a limited amount of memory [10]. Similarly, it has been shown that security is possible if the attacker's quantum memory is bounded [11], [12], [13] or more generally noisy [14], [15], [16].

Another assumption is that of non-communication. More precisely, one imagines that each party is split up into multiple agents who cannot communicate with each other for at least some parts of the protocol. Intuitively, the use of non-communicating agents can evade the standard no-go argument because while all agents in total have enough information to cheat, no single agent can cheat on his own.

On one hand, such non-communicating models have received considerable attention in classical cryptography, where such agents are often referred to as servers [17] or provers [18]. For example, Ben-Or et al. [19] considered a simple protocol for bit commitment that is secure against classical attacks as long as the committer (Bob) is split up into two agents, Bob and Brian, who are not allowed to communicate throughout the protocol. This protocol can also be modified to give security against quantum adversaries [18]. Similarly, many classical protocols for other tasks have been proposed under the assumption of non-communication, such as distributed oblivious transfer [20], i.e. symmetric private information retrieval [21], [22], [17].
or simple private information retrieval [23]. In all such protocols it was assumed that the agents of one party can never communicate during any point in the protocol, or thereafter.

On the other hand, physicists have considered so-called relativistic assumptions for cryptography [24], [25], [26], [27], [1]. In essence, this takes the form of non-communicating models where the fact that a party’s agents cannot communicate is justified by their physical separation and the finite speed of light. The key difference to classical non-communicating models is that in relativistic models the separation is generally only imposed during certain periods of the protocol, whereas classical models generally assume a separation, i.e. non-communication, for all times. For example, relativistic protocols may only demand a split into several non-communicating agents after the commit phase of a bit commitment protocol is over [27], [1]. Another assumption based on relativity is the notion of guaranteed message delivery times (see Appendix C) or the assumption of an accelerated observer

Here, we will consider the security of bit commitment protocols under the assumption that one (or both) parties Alice and Bob, are forced to be split into non-communicating agents. Motivated by the relativistic protocols of [27], [1], we thereby do not demand that the parties are split into non-communicating agents for all time, but merely during certain phases of the protocol. A bit commitment protocol can be naturally divided into: the commit phase, the wait phase, the open phase, and the verification phase (see Section II-E). We thereby introduce the explicit notion of the wait and verification phases, which are usually only implicitly defined, in order to precisely divide the overall interaction between Alice and Bob into time frames. Our first contribution is

- A classification of non-communicating models into subclasses which are characterised by the phases in which Alice or Bob is split into non-communicating agents. We find that we can reduce our considerations to two minimal models, namely the one in which Alice is split during the commit and wait phases (α-split) and the one in which Bob is split during the wait and open phases (β-split). Either of these two models allows to evade the no-go theorem because the operations required for cheating are forbidden by the split.

Fig. 1: If Bob is required to perform two separate openings it becomes important whether the command which bit he is supposed to unveil is transmitted to just one or both agents.

It turns out that in certain split models a new, subtle issue needs to be addressed. If a cheating Bob is split into two agents, Bob and Brian, during the open phase of the commitment, who decides which bit should be opened? In standard bit commitment protocols this question does not arise, as there is only one cheating party. Bob will simply announce to Alice that he wishes to unveil a particular bit, and try to provide a matching proof. However, in a model of several distinct agents, Bob and Brian could conceivably base the decision about which bit to unveil on some external input. For example, depending on the latest stockmarket news they both decide to open a 0 or a 1, even though they themselves cannot communicate. Intuitively, we would like a bit commitment scheme to be secure in the latter setting, analogous to the case of a single party which can of course also base its decision on external events. To capture this subtlety, we introduce an external verifier, Victor, who dictates which bit should be unveiled. We thereby speak of local command if Victor only issues a command to one of the two agents, Bob. We speak of global command if Victor issues a matching command to both Bob and Brian. Note that Victor should be thought of as an external verifier invoked solely to quantify Bob’s cheating power and that he plays absolutely no role when both Alice and Bob are honest. The local and global command models will be defined in purely mathematical terms and the only reason to introduce Victor is to give these mathematical definitions some intuitive meaning. Note that a related concept has recently been introduced independently in [29] under the name of the oracle input model. In a model without separated agents, the local and global command models are equivalent but we will see that they differ in a relativistic setting. More precisely, our second contribution is to

- Introduce the distinction between local and global command in the models based on the β-split. We show that there is a simple classical protocol that is secure under the local command. However, we proceed to show that there exists no classical protocol that is secure under global command in the class of β-split models.

The latter naturally leads to the question, whether there is a quantum protocol that is secure even when Victor issues a global command. A quantum protocol that is likely to be secure under global command was given in [27]. Another quantum β-split

\[\text{The authors consider two inertial participants sharing a noiseless quantum channel in the presence of a uniformly accelerated eavesdropper. They show that any information the eavesdropper manages to acquire is inherently noisy which allows the two honest participants to communicate securely. It is well-known in cryptography that most cryptographic primitives can be implemented securely as long as an external source of guaranteed noise is present.}\]
protocol was proposed by Kent [1], which has the very appealing feature that it can be implemented by the honest parties using only single qubit measurements in BB84 [30] bases, without the use of any quantum memory. Yet, no explicit security bounds were provided in [1]. Our final contribution is to

- Provide a formal security proof and security bounds for the protocol proposed in [1] in the global command model.

We want to stress that a sketch of a security proof was given in [1] already; however, we were unable to derive explicit security bounds from the arguments provided there. We thus devised an alternative proof, which allows us to find these parameters explicitly.

Our proof requires two ingredients: First, we make use of the fact that the two agents cannot communicate. Second, we employ an uncertainty relation in terms of min- and max-entropies [31]. This relation was previously used to prove the security of quantum key distribution [32], and our result illustrates its power to prove security of other cryptographic primitives.

Outline: The paper is structured as follows. Section II contains some basic definitions and technical tools essential for the proof. We also remind the reader what a bit commitment protocol is and what conditions it should satisfy. In Section III we introduce the concept of split models and, by examining the standard no-go argument, we find the minimal split requirements that might give us security and for these we state generalised security requirements. We also show how certain splits arise from special relativity if we require certain parts of the protocol to take place at space-like separated points. Section IV presents simple protocols that achieve security in the minimal split models. Section V is entirely dedicated to the bit commitment protocol proposed by Kent [1]: first we describe the protocol and then we analyse its security to obtain explicit security bounds.

II. Preliminaries

A. Hamming distance

Let \( [n] = \{1, 2, \ldots, n\} \) and let \( x \) be an \( n \)-bit string, \( x \in \{0, 1\}^n \), and denote the \( k \)-th bit of \( x \) by \( x_k \). Define the Hamming distance between two strings \( x, y \in \{0, 1\}^n \) to be the number of positions at which they differ

\[
d_H(x, y) := |\{k \in [n] : x_k \oplus y_k = 1\}|.
\]

B. Probability distributions

Let \( X \) be a random variable taking values in \( \mathcal{X} \) and distributed according to \( P_X \). The Rényi entropy of order \( \alpha \in \mathbb{R}_+ \setminus \{0, 1, \infty\} \) is defined as [33]

\[
H_\alpha(X) := \frac{1}{1-\alpha} \log \left( \sum_{x \in \mathcal{X}} P_X(x)^\alpha \right).
\]

The special cases \( \alpha \in \{0, 1, \infty\} \) are defined as limits \( H_\alpha(X) = \lim_{\beta \rightarrow \alpha} H_\beta(X) \). Note that \( H_0(X) = \log|\{x \in \mathcal{X} : P_X(x) > 0\}| \) and that the Rényi entropies exhibit monotonicity

\[
H_\alpha(X) \geq H_\beta(X) \iff \alpha \leq \beta.
\]

For \( |\mathcal{X}| = 2 \) and \( \alpha = 1 \) we obtain the binary entropy

\[
h(q) := -q \log q - (1-q) \log(1-q).
\]

Let \( P_{XY|UV} \) be a joint conditional probability distribution. \( P_{XY|UV} \) satisfies no-signalling if for all \( u \in \mathcal{U}, x \in \mathcal{X} \) the value of the sum

\[
\sum_{y \in \mathcal{Y}} P_{XY|UV}(X = x, Y = y|U = u, V = v)
\]

does not depend on a particular choice of \( v \in \mathcal{V} \).

C. Quantum notation

Let \( \rho \) be a quantum state on a Hilbert space \( \mathcal{H} \), i.e. a positive semi-definite operator with \( \text{tr} \rho = 1 \) acting on \( \mathcal{H} \). Let \( \mathcal{S}(\mathcal{H}) \) be the set of all states on \( \mathcal{H} \). We say that \( \rho_{XA} \) is a classical-quantum (cq) state if it can be written in the form

\[
\rho_{XA} = \sum_{x \in \mathcal{X}} P_X(x) |x\rangle \langle x| \otimes \rho_x,
\]

where \( P_X \) is a probability distribution and \( \rho_x \in \mathcal{S}(\mathcal{H}_A) \). Then, we define the probability of guessing \( X \) given access to the quantum system \( A \) as

\[
p_{\text{guess}}(X|A) := \max_{\{M_x\}} \sum_{x \in \mathcal{X}} P_X(x) \text{tr}(M_x \rho_x),
\]
where the maximisation is taken over all positive operator-valued measurements (POVMs) on $\mathcal{H}_A$. The min-entropy of $X$ is defined as $H_{\min}(X) := H_\infty(X)$. The min-entropy of $X$ conditioned on $A$ is defined as

$$H_{\min}(X|A) := - \log p_{\text{guess}}(X|A).$$

We say that $\rho_{XY}$ is a classical-classical (cc) state if it can be written in the form

$$\rho_{XY} = \sum_{x \in X, y \in Y} P_{XY}(x, y) |x\rangle \langle x| \otimes |y\rangle \langle y|.$$

The max-entropy of $X$ is defined as $H_{\max}(X) := H_\frac{1}{2}(X)$. The max-entropy of $X$ conditioned on $Y$ is defined as

$$H_{\max}(X|Y) := \log \sum_{y \in Y} \text{Pr}[Y = y] \cdot 2^{H_{\max}(X|Y = y)}.$$

D. Uncertainty relation

Let $\rho_{ABC}$ be any tri-partite state and let $\{M_z\}_{z \in Z}$ and $\{N_x\}_{x \in X}$ be two POVMs on the $A$ subsystem whose measurement results are represented by classical random variables $Z$ and $X$. The following cq-states arise from performing the measurements mentioned above:

$$\rho_{ZB} := \sum_{z \in Z} |z\rangle \langle z| \otimes \text{tr}_{AC}(M_z \rho_{ABC})$$

and

$$\rho_{XC} := \sum_{x \in X} |x\rangle \langle x| \otimes \text{tr}_{AB}(N_x \rho_{ABC}).$$

Theorem II.1. [31] For any tri-partite state $\rho_{ABC}$ the following uncertainty relation holds

$$H_{\max}(Z|B) + H_{\min}(X|C) \geq \log \frac{1}{c},$$

where the entropies are evaluated for $\rho_{ZB}$ and $\rho_{XC}$, respectively, and $c := \max_{z,x} \|\sqrt{M_z} \sqrt{N_x}\|_\infty^2$.

E. Bit commitment

Bit commitment is a primitive that allows Bob to commit a bit $b$ to Alice in a way that is both binding (Bob cannot later convince Alice that he actually committed to $1 - b$) and hiding (Alice cannot figure out what $b$ is before Bob decides to unveil it). In this section we discuss how to describe a bit commitment protocol and how to formalise the desired security requirements.

Any action taken by Alice or Bob can be described by a completely positive, trace-preserving (CPTP) map and the entire protocol can be defined by specifying these maps. In this paper we will denote maps performed by Alice and Bob by $\Lambda$ and $\Phi$, respectively. The subscript $X \rightarrow Y$ means that the map acts on (reads and/or modifies) the existing register $X$ and creates a new register $Y$. Moreover, identity is assumed on any subsystems not explicitly mentioned within the map: $\Lambda_{X \rightarrow Y}(\rho_{XYZ})$ stands for $(\Lambda_{X \rightarrow Y} \otimes \text{id}_Z)(\rho_{XYZ})$.

The usual description of a bit commitment protocol divides it into two phases: commit and open. However, as our scenarios rely on timing and communication constraints, it is useful to be more explicit about the structure of the protocol. We divide the protocol into four phases: commit, wait, open and verify. The commit and open phases are the essence of the protocol: they are the only phases during which Alice and Bob interact. The wait phase acts merely as a separator (this is when the commitment is valid), while in the verify phase Alice uses the information collected in the previous phases to verify the commitment and decide whether to accept or reject it.

Let $\rho_{ABC}$ be the state that Alice and Bob share at the end of the commit phase if they are both honest. The subsystems $A$ and $B$ are controlled by Alice and Bob, respectively, while subsystem $C$ is a classical register in Bob’s possession indicating which bit Bob has (honestly) committed to. Let $\Phi_{BC \rightarrow P}^{\text{open}}$ be the quantum operation that Bob applies in the open phase and it should be thought of as extracting a proof of his commitment from the subsystems $B$ and $C$ and storing it in the (possibly quantum) subsystem $P$.

$$\rho_{ABPC} = \Phi_{BC \rightarrow P}^{\text{open}}(\rho_{ABC}).$$

In the last step of the open phase Bob passes the subsystems $P$ and $C$ to Alice. Note that as $C$ is a classical register Alice is automatically assumed to read it and, hence, she finds out what Bob claims to have committed to. Let $\Lambda_{APC \rightarrow F}$ be the

---

3To simplify the notation we will omit all the subsystems on which the projector equals identity. Hence, in our shorthand notation $M_z \rho_{ABC}$ stands for $(M_z \otimes I_{BC}) \rho_{ABC}$.

4Note that we do not consider the most general class of protocols as we assume that the open phase involves one-way communication from Bob to Alice only.

5Any private or shared randomness is included in the description of the state, hence, given a protocol we can extract a unique $\rho_{ABC}$.

8The honest opening map will simply read the value of the classical register $C$, hence, its state will not be affected.
quantum operation that Alice applies in the verify phase, which creates a classical binary register (flag), $F$, indicating whether the commitment is accepted or rejected

$$\rho_{ABPCF} = \Lambda^\text{verify}_{APC \rightarrow F}(\rho_{ABPC}),$$

and let us denote a (classical) basis of the subsystem $F$ by $\{|\text{accept}\}, |\text{reject}\}.$. Describing the honest protocol suffices to define correctness.

**Definition II.1.** A bit commitment protocol is perfectly correct if $\rho_{ABC}$ satisfies

$$\langle \text{accept} | \text{tr}_{AUPC} \Lambda^\text{verify}_{APC \rightarrow F}(\Phi^\text{open}_{BC \rightarrow P}(\rho_{ABC})) | \text{accept} \rangle = 1.$$  

If one of the parties is dishonest and does not follow the protocol then the state shared between Alice and Bob is no longer well-defined. We will use $\sigma$ to denote such a dishonest state\footnote{We make no assumptions on what the dishonest party stores in their part of the state. In particular it might contain some ancillary systems to be used later.} to distinguish them from the honest states denoted by $\rho$. Security guarantee for honest Bob states that Alice finds it difficult to guess the value of his commitment before the open phase. If Alice is dishonest and does not follow the protocol then the state shared at the end of the commit phase, $\sigma_{ABC}$, does not necessarily equal $\rho_{ABC}$. However, it is important to note that the classical register $C$ is still well-defined since Bob is honest. Let $K_A$ be the set of all tri-partite states that Alice might enforce at the end of the commit phase. Informally, a bit commitment is $\delta$-hiding if for any cheating strategy the probability that Alice guesses the committed bit correctly before the open phase is upperbounded by $\frac{1}{2} + \delta$.

**Definition II.2.** A bit commitment protocol is $\delta$-hiding if all $\sigma_{ABC} \in K_A$ satisfy

$$p^\text{guess}(C|A) \leq \frac{1}{2} + \delta.$$  

Similarly, if Bob is dishonest then different states may be reached at the end of the commit phase and let $K_B$ be the set of all states that he might enforce at the end of the commit phase. Note that the classical register $C$ is no longer well-defined so we will simply talk about bi-partite states $\sigma_{AB} \in K_B$. In order to cheat successfully Bob must be able to produce valid proofs for both values of $C$, which implies that there are two distinct dishonest opening maps: Bob applies $\Phi^\text{cheat,0}_{B \rightarrow PC}$ if he chooses to open 0 and $\Phi^\text{cheat,1}_{B \rightarrow PC}$ if he chooses to open 1. The cheating map $\Phi^\text{cheat,b}_{B \rightarrow PC}$ extracts the proof of having committed to $b$ from the subsystem $B$, stores it in the subsystem $P$ and stores $b$ in the newly-created register $C$

$$\sigma^b_{ABP} \otimes |b\rangle\langle b|_C = \Phi^\text{cheat,b}_{B \rightarrow PC}(\sigma_{AB}).$$

In the last step Bob gives $P$ and $C$ to Alice, who verifies the commitment using the honest map. Let $p_b$ be the probability that Alice accepts Bob’s unveiling of $b$

$$p_b = \langle \text{accept} | \text{tr}_{ABPC} \Lambda^\text{verify}_{APC \rightarrow F}(\Phi^\text{cheat,b}_{B \rightarrow PC}(\sigma_{AB})) | \text{accept} \rangle. \tag{2}$$

The security conditions on $p_0$ and $p_1$ depend on whether we are in the classical or quantum framework. Classically, we require that at the end of the commit phase at least one of $\{p_0, p_1\}$ is small. However, this requirement turns out to be too strong in the quantum world as explained in [34] and a weaker security condition is proposed in the same paper.

**Definition II.3.** A bit commitment protocol is $\varepsilon$-weakly binding if for all $\sigma_{AB} \in K_B$ and for all cheating maps $\{\Phi^\text{cheat,b}_{B \rightarrow PC}\}_{b \in \{0,1\}}$ we have $p_0 + p_1 \leq 1 + \varepsilon$.

Unfortunately, this definition does not give us composability (see Appendix B-A for a counter-example). On the other hand the usual composable definition used for quantum protocols introduced in [12] turns out to be too stringent for the scenarios considered in this paper (see Appendix B-B for details). Hence, throughout the paper we will stick to the weaker, non-composable definition.

### III. Relativistic models

Before considering relativistic models let us briefly examine the original no-go argument (for the full version please refer to [3], [4]) to see how it might be circumvented by imposing certain communication constraints.
A. The original no-go argument and the split models

First note that we can restrict ourselves to protocols in which the state shared between Alice and Bob is pure at all times. Let \( |\phi_{AB}^0\rangle \) be the state at the end of the commit phase if Bob has decided to commit to \( b \). We require that Alice should not be able to distinguish the two cases just by looking at her subsystem which implies that \( \rho_A^0 = \rho_A^1 \), where \( \rho_A^b = \text{tr}_B |\phi_{AB}^b\rangle \langle \phi_{AB}^b| \). By Uhlmann’s theorem [35] there exists a unitary \( U_B \) acting on the subsystem \( B \) alone such that \( U_B |\phi_{AB}^0\rangle = |\phi_{AB}^1\rangle \). Hence, if the states corresponding to both commitments are the same on Alice’s side then Bob can cheat perfectly. This argument can be extended to the case in which \( \rho_A^0 \) and \( \rho_A^1 \) are close in trace distance (which means that they are difficult to distinguish) and then one can show that Bob can still cheat with high probability (for the exact trade-off based on this idea refer to [36]; for the optimal bounds on quantum bit commitment see [37]).

What is a split model? Informally, a split model is a model in which at least one party is required to delegate multiple agents to perform certain parts of the protocol in a non-communicating fashion. In this paper we only consider models in which we require a party to delegate at most 2 agents. The basic rule of two-party cryptography is that there are no third parties: the world is split between Alice and Bob only, anything that does not belong to Alice is fully controlled by Bob. Now suppose that the split model requires that there are two agents of Bob (Bob and Brian). It is still true that Bob and Brian together control everything that does not belong to Alice. However, the class of operations they can perform in a non-communicating fashion is now restricted, which might give us security. It is clear that the only way to achieve security is to split Alice during the period for which security for Bob should hold or vice versa. Therefore, we arrive at two relevant splits.

- \( \alpha \)-split : Alice is split during the commit and wait phases.
- \( \beta \)-split : Bob is split during the wait and open phases.

\[
\begin{array}{|c|c|c|c|}
\hline
 & 1. commit & 2. wait & 3. open & 4. verify \\
\hline
\alpha & Alice is split & & & \\
\hline
\beta & Bob is split & & & \\
\hline
\end{array}
\]

Fig. 2: The two relevant types of separations: \( \alpha \) and \( \beta \).

The standard no-go does not apply to the \( \alpha \)-split model because while \( \rho_A^0 \) might be globally fully distinguishable from \( \rho_A^1 \) they might locally look the same for both Alice and Amy (her agent). The \( \beta \)-split evades the no-go because the global unitary \( U_B \) might be impossible to perform by Bob and Brian without communication. Note that whenever we say that a party is split during two (or more) consecutive phases of the protocol we mean one long split throughout the whole period rather than a sequence of short ones (the agents are not allowed to get together in between).

We treat the splits as a resource. Hence, we are interested in the minimal splits that give security and we will show that \( \alpha \) and \( \beta \) are such minimal splits. What about models that impose strictly more restrictions than those? On one hand any protocol secure in the minimal split will remain secure in the more split model, we only need to ensure it is still feasible. E.g. the protocol from [1] was originally proposed in the model in which both Alice and Bob are split during the wait and open phases, while our analysis applies to the \( \beta \)-split model (strictly less split). Therefore, our proof automatically extends to the original setting. On the other hand, imposing more split might allow for new, simpler protocols. E.g. for the case of Bob being split at all times there exists a number of protocols [19], [18], [24], [25].

The number of possible split models is rather large and examining all of them case-by-case is unlikely to give any valuable insight. Hence, in this paper we only focus on the minimal splits: \( \alpha \) and \( \beta \). It is clear that a split imposed on Alice will only affect her cheating power (not Bob’s) and it is only the security guarantee for honest Bob that needs to be generalised. In the \( \alpha \)-split Bob commits to a bit by talking to Alice and Amy (subsystems \( A \) and \( A' \), respectively) and a natural generalisation of the hiding condition is to require that neither of them acquires significant knowledge about the value of \( C \). In analogy to the non-split case let \( K_{AA'} \) be the set of states that dishonest Alice and Amy can enforce at the end of the commit phase. Then the split counterpart of Definition II.2 can be written as follows.

**Definition III.1.** An \( \alpha \)-split bit commitment protocol is \( \delta \)-hiding if all \( \sigma_{AA'BC} \in K_{AA'} \) satisfy

\[
p_{\text{guess}}(C|X) \leq \frac{1}{2} + \delta \quad \text{for} \quad X = \{A, A'\}.
\]

Similarly, in the \( \beta \)-split let \( K_{BB'} \) be the set of states that dishonest Bob and Brian can enforce at the end of the commit phase. In the introduction we mentioned the concept of an external verifier Victor who challenges Bob to open a particular bit and this is how we quantify Bob’s cheating power. In the case of Bob and Brian performing two openings separately we need to specify whether Victor only tells Bob what to unveil or both Bob and Brian receive the message. We call these two

\[[10]\]We assume that Alice and Bob start in a pure state and then all the actions can be performed coherently.
scenarios the local and global command models, respectively. The first variant corresponds to the situation in which Bob makes the decision while Brian intends to behave consistently. If \( b \) is the bit that Bob intends to unveil then the cheating maps in the local command model take the form
\[
\Phi_{BB'\rightarrow PP'CC'}^{\text{cheat, local}, b} = \Phi_{BB'\rightarrow PP'C}^{\text{cheat, b}} \otimes \Phi_{B'\rightarrow P'C'}^{\text{cheat, b}},
\]
i.e. Bob’s actions depend on \( b \) but Brian’s behaviour is independent of it.

The natural motivation for the second scenario is a situation in which the agents are not allowed to communicate with each other but they might receive information from the outside world, hence, they both know \( b \). The cheating maps in the global command model take the form
\[
\Phi_{BB'\rightarrow PP'CC'}^{\text{cheat, global}, b} = \Phi_{BB'\rightarrow PP'C}^{\text{cheat, b}} \otimes \Phi_{B'\rightarrow P'C'}^{\text{cheat, b}},
\]
i.e. both opening maps depend on the value of \( b \). Using the definition of \( p_b \), the probability of successfully opening \( b \), introduced in (2) we can state the security condition in the \( \beta \)-split model.

**Definition III.2.** A \( \beta \)-split bit commitment protocol is \( \varepsilon \)-weakly binding in the local (global) command model if for all \( \sigma_{ABB'} \in K_{BB'} \) and all the cheating maps allowed in the local (global) command model we have
\[
p_0 + p_1 \leq 1 + \varepsilon.
\]

The two variations of the \( \beta \)-split model turn out to be rather different from the security point of view: there exist simple classical protocols secure in the local command model, while no classical protocol can be secure in the global command model (for details please refer to Section IV-B). Hence, to satisfy this stronger security requirement one needs to resort to quantum protocols and we investigate one of them in Section V.

![Diagram](image-url)  
*Fig. 3: Light gray regions represent the light cones of \( Q \) and \( R \), while dark gray corresponds to the common past or future. \( P \) is the latest point of the common past, while \( T \) is the earliest point of the common future.*

**B. Relativistic motivation**

Special relativity states that information cannot travel faster than the speed of light. Hence, if we are guaranteed that sites \( X \) and \( Y \) are at some well-defined distance we can calculate the minimum time it takes for a message to travel from \( X \) to \( Y \) (or vice versa). This motivates guaranteed message delivery time models, in which transmitting messages between certain parties takes a finite amount of time. To the best of our knowledge, these were the first models in which relativistic bit commitment was proposed [24], [25] (please refer to Appendix C for a brief summary of what is known about these models). Special relativity can also motivate certain split models as explained below.

We consider the model proposed by Kent [27], [1]. Take the speed of light to be \( 1 \), let \( (x, t) \) be the coordinates for Minkowski space and define the following three points: \( P = (0, 0) \), \( Q = (-1, 1) \), \( R = (1, 1) \). It is clear that \( P \) is the latest point that belongs to the common past of \( Q \) and \( R \) (Fig. 3). Hence, no signal emitted after \( t = 0 \) (regardless of where it was emitted from) can reach both \( Q \) and \( R \). Kent’s bit commitment protocols take advantage of this scenario by assuming that each party has an agent at \( P \), \( Q \) and \( R \) and they are allowed to send information at the speed of light. The commit phase happens at \( P \) while the open phase happens at \( Q \) and \( R \). The resulting communication constraints are illustrated in Fig. 4. It is clear that the communication constraints following from this configuration in space-time are strictly stronger than those of the \( \beta \)-split. This serves as a proof of principle that at least certain split models can be physically realised by requiring different parts of the protocol to take place at different, space-like separated points.
IV. BIT COMMITMENT PROTOCOLS FOR THE MINIMAL SPLITS

In Section III-A we argued that either $\alpha$ or $\beta$-split needs to be imposed for security to be possible. In this section we give explicit examples of protocols which are secure in each of the two cases.

A. Protocols based on $\alpha$-split

The $\alpha$-split allows for a simple bit commitment protocol based on secret sharing. Such protocols will have the feature that once the commit phase is over, the combined systems of Alice and Amy determine the committed bit and the commitment only lasts as long as the separation is maintained. This is similar to the distributed oblivious transfer scenarios [20] in which security disappears as soon as the agents are allowed to communicate.

**Protocol 1:** Bit commitment from secret sharing

1) (commit) Bob commits to $b \in \{0, 1\}$ by generating a random bit $r$ and sending $b \oplus r$ to Alice and $r$ to Amy.
2) (open) Alice and Amy calculate $b = (b \oplus r) \oplus r$.

Security against classical adversaries follows directly from the properties of secret-sharing. It is also secure against quantum adversaries (see Appendix D-B for details). As there exists a classical protocol that is perfectly secure (even against quantum adversaries) in this scenario quantum mechanics gives us no advantage for the purpose of bit commitment.

B. Protocols based on $\beta$-split

In contrast to the $\alpha$-split case commitments based on the $\beta$-split can be made permanent—Bob and Brian can always refuse to participate in the open phase and Alice will learn nothing about their commitment. As discussed in Section III-A we need to distinguish between the local and global command models.
1) Security in the local command model: It turns out that in the $\beta$-split model under the local command there exists a simple classical protocol that achieves security.

**Protocol 2: Bit commitment in the local command model**

1) (commit) Bob chooses a bit $b$ and shares it with Brian.
2) (open) Bob and Brian independently send to Alice a bit they claim to have committed to (denote these bits by $x$ and $y$, respectively).
3) (verify) Alice accepts the commitment of $b$ if $b = x = y$, else she rejects.

It is easy to convince ourselves that the protocol is secure (according to the weakly binding definition). The problem that Bob and Brian face is to correlate the bits they are trying to unveil. In order to do that they either have to agree on the bit in advance (which corresponds to an honest commitment) or they would have to violate no-signalling. For a more detailed security analysis we refer to Appendix D-C (see also the independent discussion of this and related points in [29]).

2) Security in the global command model: We have seen that in the local command model there exists a very simple bit commitment protocol that achieves security. Unfortunately, as soon as we switch to the global command the protocol becomes insecure — Bob and Brian can cheat perfectly. Let us consider what is and what is not possible in the $\beta$-split model under the global command.

a) Classically: Classically, it is not possible to achieve security in the $\beta$-split model under the global command and the informal argument goes as follows. As the protocol needs to be correct Bob and Brian must be able to honestly commit to either bit, i.e. they must be able to agree on unveiling strategies\(^{11}\) that will make Alice accept either bit even without any further communication between Bob and Brian. Since the protocol is hiding the interaction during the commit phase cannot give away any information about the committed bit and, therefore, both strategies remain valid until the beginning of the open phase. Hence, whichever bit Bob and Brian are told to unveil they can always succeed.

b) Quantum mechanically: The informal argument presented above does not apply in the quantum world due to the no-cloning principle. The opening strategy may rely on some quantum system that is available to Bob right before the split — but cannot be shared with Brian without loss. The first protocols in the $\beta$-split model were proposed by Kent [27], [1] and Section V focuses on one of them.

V. BIT COMMITMENT BY TRANSMITTING MEASUREMENT OUTCOMES

We introduce a variant of the bit commitment protocol by Kent [1] and then present a security proof that leads to explicit security bounds.

A. The protocol

The original protocol presented in [1] uses BB84 states. However, for the purpose of the proof we analyse its purified analogue (which is equivalent from the security point of view). Denote the computational basis by $B_0 = \{ |0\rangle, |1\rangle \}$ and the Hadamard basis by $B_1 = \{ |+\rangle, |−\rangle \}$.

Note also that the original scenario described by Kent makes strictly more assumptions (because it requires both parties to be split rather than just one). However, we will see that whether Alice is split or not does not affect the security. Hence, the security proof for the $\beta$-split model presented here automatically applies to the setup originally proposed by Kent.

**Protocol 3: Bit commitment by transmitting measurement outcomes**

1) Alice creates $2n$ EPR pairs and sends one half of each pair to Bob.
2) (commit) Bob commits to a bit $b$ by measuring every qubit he receives in $B_0$. Denote the outcomes by $T$ (a classical bit string of length $2n$).
3) (end of commit) Bob splits up into two agents: Bob and Brian. Each of them holds a copy of $T$. **No more communication is allowed between Bob and Brian until the end of the protocol.**
4) (open) Bob opens the commitment by sending $b$ and $T$ to Alice. Brian does the same.
5) Alice picks a random subset $Z \subset [2n]$ of size $n$ and let $X := [2n] \setminus Z$. She measures the qubits from $Z$ in the computational basis and the qubits from $X$ in the Hadamard basis. Denote her measurement outcomes by $S$ (a classical bit string of length $2n$).

\(^{11}\)Bob and Brian agree on unveiling strategies during the commit phase, which they are allowed in the $\beta$-split model. This argument might not apply in the case of stronger splits (e.g. Bob and Brian split at all times).
6) (verify) Alice performs three checks:
- Alice checks whether the values of \( b \) submitted by Bob and Brian are the same.
- Alice checks whether the strings submitted by Bob and Brian are the same.
- Alice checks whether the strings submitted are consistent with \( S \) (consistency check on qubits she measured in \( B_b \) only).

If all three checks pass then the opening is accepted.

As mentioned in Section II-E a secure bit commitment protocol should satisfy three conditions. If Bob is honest he will choose a bit \( b \), perform the correct measurement to obtain the (classical) string \( T \). After the split Bob and Brian will both possess identical copies of \( b \) and \( T \), which they send to Alice during the open phase. Hence, the first two checks clearly go through. The third check goes through because honest Alice prepared perfect EPR pairs, measured them to obtain string \( S \) and so strings \( S \) and \( T \) must be perfectly correlated on the qubits measured in the same basis. Hence, the protocol is perfectly correct. Security for honest Bob is also easy to see. Alice does not receive any information before the open phase, hence, she cannot learn anything about Bob’s commitment by no-signalling and the protocol is \( \delta \)-hiding for \( \delta = 0 \). Therefore, we only analyse security for honest Alice, i.e. show the following result:

**Theorem V.1.** Protocol 3 in the \( \beta \)-split model under the global command is \( \varepsilon \)-weakly binding, where

\[
\varepsilon = \inf_{\delta \in (0, \frac{1}{2})} 2^{1-n(1-h(\delta))} + 2 \exp \left( -\frac{1}{2} n\delta^2 \right),
\]

where \( h(\cdot) \) is the binary entropy function as defined in Section II-B.

Note that not only does \( \varepsilon \) vanish in the limit \( n \to \infty \), but also the rate of decay is exponential in \( n \) (\( n \) is the number of rounds played, hence, the resources necessary to execute the protocol grow linearly in \( n \)). The fact that \( \varepsilon \) decays exponentially would be a great advantage if the protocol were to be implemented experimentally and shows that the protocol might be of practical interest.

**B. Security for honest Alice**

1) **Notation:** Let us denote the state of the system at the end of the commit phase by \( \sigma_{ABB'} \), where subsystems \( A \), \( B \) and \( B' \) belong to Alice, Bob and Brian, respectively. Alice is honest so we know the exact state of her subsystem— it contains \( 2n \) qubits, which have already been partitioned into sets \( \mathcal{Z} \) and \( \mathcal{X} \). This justifies a natural partition of the subsystem \( A \) into subsystems \( A_{\mathcal{Z}} \) and \( A_{\mathcal{X}} \), each containing exactly \( n \) qubits. Let quantum operation \( \Lambda^b_G \) for \( G \in \{ A_{\mathcal{Z}}, A_{\mathcal{X}} \} \), \( b \in \{0, 1\} \) correspond to measuring all qubits from the subsystem \( G \) in the basis \( B_b \). The relevant projectors can be formally defined as

\[
P^{b,s}_G := |H^{\otimes n}|^b \langle s | G | [H^{\otimes n}]^b \rangle,
\]

where \( s \in \{0, 1\}^n \). Denote the environment by \( E \) and the subsystem used to store the measurement outcomes by \( F \). Then \( \Lambda^b_G \) is defined as

\[
\rho_{FE} := \Lambda^b_G (\rho_{GE}) = \sum_s |s \rangle \langle s |_F \otimes \text{Tr}_G (P^{b,s}_G \rho_{GE}).
\]

The three relevant measurements are \( \Lambda^0_{A_{\mathcal{Z}}}, \Lambda^1_{A_{\mathcal{Z}}}, \Lambda^0_{A_{\mathcal{X}}} \) — the first two are actually performed in the honest protocol, while the third one is a virtual measurement, required for the proof only. Bob and Brian are expected to extract a string from their respective quantum systems. Let us simplify the notation introduced in Section II-E and denote Bob’s map intending to open \( b \) and producing string \( T \) as the output by \( \Phi^b_{B'} \). Similarly, for Brian denote the map intending to open \( b' \) by \( \Phi^b_{B'} \), and the output string by \( T' \). Observe that \( \Phi^b_B \) (\( \Phi^b_B \)) is restricted to operate on the subsystem \( B \) (\( B' \)) only. The string \( T \) corresponds to measuring all \( 2n \) qubits. Once Alice has chosen the partition into \( \mathcal{Z} \) and \( \mathcal{X} \) we can naturally split it into two substrings \( T = \{ T_{\mathcal{Z}}, T_{\mathcal{X}} \} \), which correspond to the outcomes obtained from the qubits from sets \( \mathcal{Z} \) and \( \mathcal{X} \), respectively. Splitting \( T \) into two substrings is useful because when Alice has to decide whether to accept or reject the commitment she will only look at one of the substrings (the one measured in the same basis). Clearly, analogous partition applies to \( T' = \{ T'_{\mathcal{Z}}, T'_{\mathcal{X}} \} \).

2) **No-signalling constraints:** Let us think of Alice as talking to Bob and Brian separately and making a separate decision (whether to accept or not) for each of them. We can see that this gives rise to a joint probability distribution with two inputs and two outputs: the inputs are the bits that Bob and Brian were asked by Victor to unveil (\( b \) and \( b' \), respectively), while the outputs are Alice’s binary (\{accept, reject\}) outcomes (one on each side). We have already defined the maps that Bob and Brian will apply so now we just need to specify what the tests on Alice’s side are. As described in the protocol Alice will

\footnote{We are in the global command model so both Bob and Brian know what they are trying to unveil.}
As Bob and Brian act on disjoint quantum systems and the tests performed by Alice are classical the probability distribution of outcomes must satisfy no-signalling. Note that we replaced certain fields \((a_{11} \text{ and } a_{14})\) by the probability of successfully opening 0 and 1 \((p_0 \text{ and } p_1)\), respectively. This follows from the definition of \(p_d\) in the global command model:

\[
p_d := \Pr\{\text{accept, accept} | b = d, b' = d\}.
\]

Also, we have replaced \(a_{14} \text{ by } \alpha\) because it turns out to be the quantity we will bound in the second part of the proof. The following lemma uses the no-signalling principle to find an upper bound on the sum of \(p_0 \text{ and } p_1\).

**Lemma V.1.** No-signalling between Bob and Brian implies that \(p_0 + p_1 \leq 1 + \alpha\).

**Proof:** Consider the following no-signalling constraints:

\[
\alpha + a_{24} = a_{34} + p_1 \text{ and } a_{21} + a_{22} = a_{23} + a_{24}.
\]

Moreover, we know that each quarter adds up to 1 so \(p_0 + a_{12} + a_{21} + a_{22} = 1\). Combining the two conditions gives

\[
p_0 + p_1 = 1 - a_{12} - a_{21} - a_{22} + a + a_{24} - a_{34} = 1 - a_{12} - a_{23} + \alpha - a_{34} \leq 1 + \alpha.
\]

Hence, it is enough to show that as the number of rounds \(n\) increases \(\alpha\) can be made arbitrarily small, which is the focus of the next section.

3) **Impossibility of guessing both strings:** The probability \(\alpha\) corresponds to Bob trying to unveil \(b = 0\), Brian trying to unveil \(b' = 1\) and both openings being accepted. Let \(\rho_{SZSTZT_2T_1T_X'}\) be the state after all three parties have performed their measurements (note that this state is purely classical)

\[
\rho_{SZSTZT_2T_1T_X'} := (A^0_{AZ} \otimes A^1_{AX} \otimes \Phi_B^0 \otimes \Phi_{B'}^0)\rho_{AZAXBB'}.
\]

As \(\alpha\) is the probability that \(\rho_{SZSTZT_2T_1T_X'}\) passes the relevant tests it can be written as

\[
\alpha = \tr(\Pi_B^0 \Pi_B^1 \rho_{SZSTZT_2T_1T_X'}). \tag{5}
\]

As operators acting on disjoint subsystems commute we can change the order slightly

\[
\alpha = \tr(\Pi_B^0 \Pi_B^1 \rho_{SZSTZT_2T_1T_X'}) = \tr(\Pi_B^0 \Pi_B^1 (A^0_{AZ} \otimes \Phi_B^0 \otimes \Phi_{B'}^0)\rho_{AZAXBB'}) = \tr(\Pi_B^0 (\Phi_{B'}^1) \tr(\Pi_B^0 (\Phi_B^0)\rho_{AZAXBB'})) = \tr(\Pi_B^1 (\Phi_{B'}^1) \tr(\Pi_B^0 (\Phi_B^0)\rho_{AZAXBB'})).
\]

Define

\[
p := \tr(\Pi_B^0 (\Phi_B^0)\rho_{AZAXBB'}),
\]

\[
p_{\text{pass}} := \frac{1}{p} \tr_{SZT_2} \left[\Pi_B^0 (\Phi_B^0)\rho_{AZAXBB'}\right].
\]

\(\text{The variables that do not appear in our argument have been replaced with placeholders.}\)
It is easy to see that $p$ is the probability that Bob passes his test and $\rho_{A_X T_X B'}^{\text{pass}}$ is the normalised state conditioned on passing. Hence, $\alpha$ can be written as

$$\alpha = \text{tr} \left( \Pi_B^1 (A_{A_X}^1 \otimes \Phi_{B'}^1) \rho_{A_X T_X B'}^{\text{pass}} \right) \cdot p.$$  

This way of writing $\alpha$ allows us to apply Theorem II.1 to the tri-partite state $\rho_{A_X T_X B'}^{\text{pass}}$.

**Lemma V.2.** For any strategy adopted by dishonest Bob

$$\alpha \leq \inf_{\delta \in (0,\frac{1}{2})} 2^{1-n(1-h(\delta))} + 2 \exp \left( -\frac{1}{2} n\delta^2 \right).$$

Proof: The trace on the right hand side of (6) corresponds to the probability that Brian guesses $S_X$ correctly by applying his opening map on his subsystem conditioned on Alice accepting Bob’s opening. The guessing probability using a fixed map $\Phi_{B'}^1$ is upperbounded by the optimal guessing probability [38] which can be written in terms of the min-entropy. Hence,

$$\frac{\alpha}{p} = \text{tr} \left( \Pi_B^1 (A_{A_X}^1 \otimes \Phi_{B'}^1) \rho_{A_X T_X B'}^{\text{pass}} \right) \leq 2^{-\min(S_X|B')} ,$$

where the min-entropy is evaluated on the state $\rho_{S_X T_X B'} := \text{tr}_{T_X} A_{A_X}^1 (\rho_{A_X T_X B'}^{\text{pass}})$. To use the uncertainty relation (1) we also need to consider $\rho_{\hat{S}_X T_X} := \text{tr}_{B'} A_{A_X}^0 (\rho_{A_X T_X B'}^{\text{pass}})$, which would be obtained if Alice decided to make the third (virtual) measurement in a complementary basis. Combining (1) with (8) gives

$$\frac{\alpha}{p} \leq 2^{H_{\max}(\hat{S}_X|T_X)-n},$$

where $H_{\max}(\hat{S}_X|T_X)$ is evaluated on $\rho_{\hat{S}_X T_X}$. Note that now we just need to bound the classical conditional max-entropy between two classical random variables (the state $\rho_{\hat{S}_X T_X}$ is purely classical). It turns out that it is enough to show that the Hamming distance between $\hat{S}_X$ and $T_X$ is small with high probability. To get such a bound we need to examine the (fully classical) state $\rho_{S_X \hat{S}_X T_X} := \text{tr}_{B'} \left[ (A_{A_X}^0 \otimes A_{A_X}^0 \otimes \Phi_B^0 \otimes \text{id}_{B'}) \rho_{ABB} \right]$. The fact that $Z$ and $X$ are random subsets of $[2n]$ allows us to derive the following inequality from the Hoeffding bound [39] (details in Appendix A).

$$\Pr \left[ \text{d}_H(\hat{S}_X, T_X) \geq \delta n \land \text{d}_H(S_Z, T_Z) = 0 \right] \leq \exp \left( -\frac{1}{2} n\delta^2 \right) =: \varepsilon .$$

We can also write it as conditional probability

$$\Pr \left[ \text{d}_H(\hat{S}_X, T_X) \geq \delta n | \text{d}_H(S_Z, T_Z) = 0 \right] \leq \frac{\varepsilon}{p},$$

because $\text{d}_H(S_Z, T_Z) = 0$ is equivalent to Bob passing the test (and happens with probability $p$ as defined in (V-B3)). Let $0 < \delta < \frac{1}{2}$ and define a binary event, $\Gamma$, such that

$$\Gamma := \begin{cases} 0 & \text{if } \text{d}_H(\hat{S}_X, T_X) < \delta n , \\ 1 & \text{if } \text{d}_H(\hat{S}_X, T_X) \geq \delta n . \end{cases}$$

If $\Gamma = 0$ then for any particular value of $T_X = t_X$ the Rényi entropy\footnote{All entropies are evaluated on $\rho_{\hat{S}_X T_X}^{\Gamma}$ except for $H_{\min}(S_X|B')$ which is evaluated on $\rho_{S_X B'}$.} of order 0 can be bounded by

$$H_0(\hat{S}_X|T_X = t_X, \Gamma = 0) \leq \log \left( \sum_{i=0}^{\lfloor n\delta \rfloor} \binom{n}{i} \right) \leq nh(\delta),$$

where the last inequality comes from a well-known bound (see e.g. Lemma 16.19 in [40]). The monotonicity of classical Rényi entropies implies that

$$H_{\max}(\hat{S}_X|T_X = t_X, \Gamma = 0) \leq H_0(\hat{S}_X|T_X = t_X, \Gamma = 0).$$

If $\Gamma = 1$ then we have no bound better than the maximal value $H_{\max}(\hat{S}_X|T_X = t_X, \Gamma = 1) \leq n$. It can be shown (see e.g. Section 4.3.2 in [41]) that the conditional max-entropy for classical states reduces to

$$H_{\max}(Z|Y) := \log \sum_{y \in Y} \Pr[Y = y] \cdot 2^{H_{\max}(Z|Y = y)} .$$

As neither of our bounds depends on the particular value of $T_X = t_X$, they will not be affected by averaging over all strings $t_X$. Hence, we only need to average over $\Gamma$

$$2^{H_{\max}(\hat{S}_X|T_X, \Gamma)} = \Pr[\Gamma = 0] \cdot 2^{H_{\max}(\hat{S}_X|T_X, \Gamma = 0)} + \Pr[\Gamma = 1] \cdot 2^{H_{\max}(\hat{S}_X|T_X, \Gamma = 1)} \leq \left( 1 - \frac{\varepsilon}{p} \right) 2^{n h(\delta)} + 2^n \frac{\varepsilon}{p} + 2^n \frac{\varepsilon}{p} .$$

(12)
One bit of information cannot decrease the entropy by more than 1 bit (see e.g. Proposition 5.10 in [41]), hence
\[ H_{\text{max}}(\hat{S}_X|T_X) \leq H_{\text{max}}(\hat{S}_X|T_X, \Gamma) + 1. \] (13)

Hence, from (9), (12) and (13) we get
\[ \alpha \leq 2p \left[ 2^{-n(1-h(\delta))} + \frac{\varepsilon}{p} \right] \leq 2^{1-n(1-h(\delta))} + 2 \exp \left( -\frac{1}{2} n\delta^2 \right), \]
which directly implies our claim. 
Finally, Theorem V.1 follows directly from Lemmas V.1 and V.2.

VI. CONCLUSIONS AND OPEN QUESTIONS

Our interest in bit commitment protocols based on the relativistic constraint was sparked by recent papers by Kent [27], [1]. While the author gave an intuition for the security of the protocol based on BB84 states, no explicit security bounds were given. Once we had proven the security of the protocol and calculated such bounds, we became interested in other split models: which of them can give us security and in which of them are quantum protocols more powerful than classical ones? We have investigated the minimal split assumptions that might allow for secure bit commitment and we have shown that they are indeed sufficient. We have found that in the \( \beta \)-split under the global command quantum protocols are more powerful than classical ones.

We have proven security of bit commitment with respect to the weakly binding definition, which is non-composable. We also know that the usual stronger definition (which would imply composability) is not achievable. We cannot hope for universal composability but maybe it is possible to prove some weaker form of composability. For example, is it possible to combine \( n \) bit commitment protocols [1] to obtain a secure string commitment scheme? If it is not secure one might investigate if there are some extra constraints (e.g. that the commit phases are executed sequentially or that the unveilings happen simultaneously at space-like separated points) that would guarantee composability.

One might also wonder whether these models allow us to construct other cryptographic primitives. Probably the most natural one to look at would be oblivious transfer [42], [43]. Unfortunately, the primitive of oblivious transfer requires the security to last forever. This would only be possible if certain parties remained split forever, which cannot be motivated by relativistic assumptions. Moreover, if certain parties were to remain split forever then oblivious transfer can be implemented even classically [20]. It is possible, however, that some weaker form of oblivious transfer (in which the security does not last forever) can be proven secure in relativistic models.

APPENDIX A

Hoeffding bound

In Lemma 7 we need to bound the probability that sampling a small, random substring gives rise to the statistics which is very different from the true statistics of the entire string. The Hoeffding bound is exactly the tool we need. Suppose that we have a string of length \( 2n \) which contains \( n_{\text{err}} \) errors and let \( \Lambda = \frac{n_{\text{err}}}{2n} \) denote the error fraction in the whole string. Let us take a random sample of the string of size \( k \) and denote the error fraction in the sample by \( \lambda \). Then, the Hoeffding bound [39] states that
\[ \Pr \left[ \Lambda \geq \lambda + \frac{\delta}{2} \right] \leq \exp \left( -\frac{1}{2} k\delta^2 \right). \]

Adding an extra event cannot increase the probability
\[ \Pr \left[ \Lambda \geq \lambda + \frac{\delta}{2} \land \lambda = 0 \right] \leq \exp \left( -\frac{1}{2} k\delta^2 \right). \]

The expression inside the square bracket can be rewritten, giving us
\[ \Pr \left[ n_{\text{err}} \geq \delta n \land \lambda = 0 \right] \leq \exp \left( -\frac{1}{2} k\delta^2 \right). \]

This is exactly the bound we use in (10).

APPENDIX B

Composability issues

For the sake of completeness we state some observations concerning composability. On one hand we show that the weak bindingness definition is not composable (by giving an explicit counter-example). On the other hand we argue that the usual stronger definition [12] cannot be satisfied in the split setting.
A. Counter-example to the composability of the weakly binding definition

In Section II-E we explained what it means that a bit commitment protocol is weakly binding and we also said that the definition does not guarantee composability, e.g. executing the protocol $n$ times does not necessarily give a secure string commitment (string commitment is an extension of bit commitment in which we are allowed to commit to a bitstring of length $n$ rather than just a single bit). Let us explain what the source of the problem is. Consider a bit commitment protocol which is binding in the sense that with probability $\frac{1}{2}$ Bob can unveil either bit successfully and with probability $\frac{1}{2}$ he will fail regardless of his intentions. Clearly, we would not call this protocol secure. However, as $p_0 = p_1 = \frac{1}{2}$ it satisfies the $\varepsilon$-weakly binding definition for $\varepsilon = 0$. To expose the problem even further consider the task of string commitment. Analogous to the bit commitment case suppose that at the end of the commit phase Alice and Bob share a state $\rho_{AB}$. Let $q_s(\rho_{AB})$ be the probability that Bob successfully unveils string $s$. Then it is natural to say that a string commitment protocol is $\delta$-weakly binding if for all states $\rho_{AB}$ it satisfies

$$\sum_s q_s(\rho_{AB}) \leq 1 + \delta.$$ 

Now consider a string commitment protocol such that Alice with probability $\frac{1}{2}$ accepts anything while with probability $\frac{1}{2}$ rejects everything. It is clear that this is not a secure string commitment box as $\sum_s q_s(\rho_{AB}) = \frac{1}{2}2^n = 2^{n-1}$. However, if we look at each bit separately we will find that $p_0 = p_1 = \frac{1}{2}$ and so each bit commitment is weakly binding. This shows that combining $n$ weakly binding bit commitments does not imply that the resulting string commitment is secure.

B. Impossibility of satisfying the stronger definition

**Definition B.1.** [12] A bit commitment protocol is $\epsilon$-binding if the fact that Alice is honest ensures that for any state at the beginning of the open phase, $\rho_{AB}$, there exists an extension of the form

$$\rho_{ABD} = P_D(0)|0\rangle_D\langle 0| \otimes \rho_{AB}^0 + P_D(1)|1\rangle_D\langle 1| \otimes \rho_{AB}^1,$$

where $D$ is a classical register and $P_D$ is a probability distribution, for which the conditioned states satisfy $p_{1-b}(\rho_{AB}^b) \leq \epsilon$ for $b \in \{0, 1\}$.

While this definition has proven useful in the bounded and noisy storage models [11], [14] we argue that it is generally inapplicable outside of these scenarios. The security in these models results from the fact that Alice and Bob cannot purify the protocol, as there is a subsystem, referred to as the environment, $E$, which they do not have access to. In other words $\rho_{AB}$ is not pure because we trace out the environment $E$, e.g. a pure state $|\phi\rangle_{ABE}$ leads to $\rho_{AB} = tr_E|\phi\rangle\langle \phi|_{ABE}$. The following argument shows that if the model does not prevent the parties from purifying the protocol then Definition B.1 can only be satisfied for $\epsilon \geq \frac{1}{2}$. Suppose that Bob commits to an equal superposition of 0 and 1 (as explained above). If Alice and Bob start in a pure state and execute a purified version of the protocol (i.e. implement all operations as unitaries, generate coherent randomness and keep all the measurements quantum) then the state at the beginning of the open phase is pure. One possible opening strategy for Bob is to measure the control qubit, which collapses the state. The collapsed state is exactly as if Bob had generated a random bit $b$ at the very beginning of the protocol and honestly committed to it. Such a strategy gives us a lower bound on how well Bob can open each bit, namely $p_b(\rho_{AB}) \geq \frac{1}{2}$ for $b \in \{0, 1\}$. As the overall state is pure at the beginning of the open phase, any classical register $D$ must necessarily be independent, which means that $\rho_{AB}^0 = \rho_{AB}^1 = \rho_{AB}$. Then $p_1(\rho_{AB}^0) = p_1(\rho_{AB}) \geq \frac{1}{2}$ so Definition B.1 can only hold for $\epsilon \geq \frac{1}{2}$. This argument shows that Definition B.1 cannot be satisfied by protocols that do not assume the presence of some external system inaccessible to either party.

APPENDIX C

GUARANTEED MESSAGE DELIVERY TIME MODELS

Suppose that Bob, based on Earth, exchanges messages with Alice, who is on the Moon. Special relativity states that no message can travel faster than the speed of light, hence the minimum delivery time equals about $1.26s$. This scenario motivates the study of models in which there are two separated sites and while intra-site communication can be instantaneous, any inter-site message takes at least $\Delta t$ to be delivered. We also assume that the inter-site (classical or quantum) channels are perfectly secure (neither party can read or alter anything that is on the wire).

A. One agent per site

![Diagram of one agent per site](image)

**Fig. 7:** The simplest guaranteed message delivery time model: one agent per site.
The simplest model (illustrated in Figure 7) assumes that each party controls one site. Clearly, if Bob sends a bit $b$ to Alice he is committed to it. The commitment is perfect because at time $t = 0$ Bob is fully committed (he cannot alter his commitment at any later time), while at the same time until time $t = \Delta t$ Alice is fully ignorant about the commitment. The drawback of such a scheme is the fact that the commitment only lasts for $\Delta t$ and then automatically opens. Such schemes have been studied before [44] but in a slightly different context. The conclusion is that for certain applications (e.g. constructing a strong coin flip, signing contracts) such timed commitments are good enough, while for others (e.g. Yao’s construction of OT using quantum communication [45], [43]) they are not. To illustrate the limitations of this model let us consider if it is possible to construct a commitment that lasts for longer than $\Delta t$. Classically, this is not possible and the intuitive argument is simple. In the absence of noise classical protocols are fully deterministic and no probabilities can arise. For each of the bits Bob either can ($p_b = 1$) or cannot ($p_b = 0$) unveil it. Hence, the distinction between being and not being committed is sharp (either $p_0 + p_1 = 2$ or $p_0 + p_1 = 1$). Bob being committed implies that the information beyond his control determines the bit. As Alice will have received all the messages in transit after time at most $\Delta t$ she will be able to learn the committed bit. Therefore, no commitment can be made longer than $\Delta t$. In the quantum world the situation is more complicated due to two things. First of all, quantum mechanics is a probabilistic theory so there is no sharp distinction between being and not being committed — Bob can be partially committed. The second complication is the no-cloning theorem. Suppose that at some point Bob becomes, to some extent, committed, which means that the information on Alice’s side combined with the messages on the wire give away some information about his commitment. Now, assume that Alice waits until the messages arrive (at most $\Delta t$) and does some measurements to learn something about Bob’s commitment. Clearly, the standard hiding-binding trade-off applies. However, the honest protocol might require Alice to return some states to Bob before the messages arrive and so by keeping them she takes a risk of being caught cheating. It is an open question if this time-constrained scenario gives us some advantage over the standard scenario for constructing cheat-sensitive bit commitments. It is clear, however, that no secure (hiding) bit commitment can last longer than $\Delta t$. Hence, for this specific purpose quantum and classical protocols are equally powerful.

B. Two agents per site

A slightly more complicated model (illustrated in Figure 8) assumes that each party has a trusted agent at each site (Bob trusts his agent Brian and Alice trusts her agent Amy). Protocols implementing bit commitment in such a scenario, in which the commitment can be sustained indefinitely as long as messages are exchanged at each site have been presented in [24], [25]. After the exchange stops the commitment remains valid for $\Delta t$ and then expires. These protocols have been shown to be secure against classical attacks and are conjectured to be secure against any quantum attack.

APPENDIX D
CLASSICAL PROTOCOLS AGAINST QUANTUM ADVERSARIES

Some of the protocols we present are purely classical but in order to determine whether they are secure against quantum adversaries we need to translate them into the quantum formalism. This section describes briefly how this can be achieved and analyses the security of these protocols in the quantum setting. While the actual security proofs may appear trivial, we have decided to include them for completeness.

A. Classical protocol in the quantum formalism

Sending a classical bit $b \in \{0, 1\}$ is equivalent to encoding it in the computational basis and sending the resulting state $|b\rangle$ to the other party. Receiving a classical bit corresponds to receiving a qubit and immediately measuring it in the computational basis.
B. Bit commitment from secret sharing

Here we analyse Protocol 1 from Section IV-A. If Alice and Amy are honest they will measure the qubits they receive immediately in the computational basis. Once the measurement outcomes are known Bob’s commitment is well-defined and he will not be able to cheat. If Bob is honest r will be a truly random bit. Then what Alice and Amy receive can be described by the following density matrix

$$\rho_{AA'}^0 = \frac{1}{2} [0\langle 0 |_A \otimes |d\rangle_A \langle d |_{A'} + |1\rangle_A \otimes |1 - d\rangle_{A'}].$$

It is easy to convince ourselves that while $\rho_{AA'}^0$ and $\rho_{AA'}^1$ are perfectly distinguishable the reduced states are fully mixed, $\rho_A^0 = \rho_A^1 = \rho_{A'}^1 = \frac{1}{2}$. Hence, Alice and Amy remain perfectly ignorant about Bob’s commitment as long as they are separated.

C. Bit commitment in the local command

Here we analyse Protocol 2 from Section IV-B1. Clearly, the protocol is perfectly hiding because Alice does not receive any messages until the beginning of the open phase. To show that it is also weakly binding we need to employ no-signalling between Bob and Brian.

**Lemma D.1.** Protocol 2 is weakly binding with $\varepsilon = 0$.

**Proof:** Suppose that Bob and Brian want to cheat. At the beginning of the open phase each of them picks an opening strategy from sets $R$ and $S$, respectively. Note that this has to be done independently because they are not allowed to communicate. Bob receives the command so his distribution will in general depend on the command and if the command is $b$ denote the probability of picking $r \in R$ by $p_R^b(r)$. For the second player the distribution has to be fixed and the probability of picking $s \in S$ equals $p_S(s)$, regardless of what the value of $b$ is. Recall from Section II-E that $p_b$ is the probability that Alice accepts the commitment if the command is $b$. Hence, we can write

$$p_b = \sum_{r \in R} \sum_{s \in S} p_R^b(r)p_S(s)p(x = b, y = b|r, s) \leq \sum_{r \in R} \sum_{s \in S} p_R^b(r)p_S(s)p(y = b|r, s).$$

By no-signalling we know that $p(y = b|r, s)$ does not depend on $r$ so we can write $p(y = b|s)$ instead. Then we get

$$p_0 + p_1 \leq \sum_{r \in R} \sum_{s \in S} \left[ p_R^b(r)p_S(s)p(y = 0|s) + p_R^b(r)p_S(s)p(y = 1|s) \right] = \sum_{s \in S} p_S(s) [p(y = 0|s) + p(y = 1|s)] = 1.$$

One might also wonder whether the protocol satisfies the stronger binding requirement (Definition B.1). However, a similar argument to the one sketched out in Section V-B shows that the stronger definition cannot hold.

**ACKNOWLEDGEMENTS:**

We thank Roger Colbeck, Fabian Furrer, Tomasz Paterek and Severin Winkler for helpful discussions and Adrian Kent for valuable comments on an earlier version of this manuscript.

**REFERENCES**

[1] A. Kent, “Unconditionally Secure Bit Commitment by Transmitting Measurement Outcomes,” *Phys. Rev. Lett.*, vol. 109, no. 13, p. 130501, Sep. 2012. [Online]. Available: [http://link.aps.org/doi/10.1103/PhysRevLett.109.130501](http://link.aps.org/doi/10.1103/PhysRevLett.109.130501)

[2] J. Kilian, “Founding cryptography on oblivious transfer,” in *Proc. ACM STOC*. New York, New York, USA: ACM Press, 1988, pp. 20–31. [Online]. Available: [http://portal.acm.org/citation.cfm?doid=62212.62215](http://portal.acm.org/citation.cfm?doid=62212.62215)

[3] D. Mayers, “Unconditionally Secure Quantum Bit Commitment is Impossible,” *Phys. Rev. Lett.*, vol. 78, no. 17, pp. 3414–3417, Apr. 1997. [Online]. Available: [http://link.aps.org/doi/10.1103/PhysRevLett.78.3414](http://link.aps.org/doi/10.1103/PhysRevLett.78.3414)

[4] H.-K. Lo and H. F. Chau, “Is Quantum Bit Commitment Really Possible?” *Phys. Rev. Lett.*, vol. 78, no. 17, pp. 3410–3413, Apr. 1997. [Online]. Available: [http://link.aps.org/doi/10.1103/PhysRevLett.78.3410](http://link.aps.org/doi/10.1103/PhysRevLett.78.3410)

[5] G. D’Ariano, D. Kretschmann, D. Schlingemann, and R. F. Werner, “Reexamination of Quantum Bit Commitment: The Possible and the Impossible,” *Phys. Rev. A.*, vol. 76, no. 3, pp. 1–27, Sep. 2007. [Online]. Available: [http://link.aps.org/doi/10.1103/PhysRevA.76.032328](http://link.aps.org/doi/10.1103/PhysRevA.76.032328)

[6] S. Winkler, M. Tomamichel, S. Hengl, and R. Renner, “Impossibility of Growing Quantum Bit Commitments,” *Phys. Rev. Lett.*, vol. 107, no. 9, Aug. 2011. [Online]. Available: [http://prl.aps.org/abstract/PRL/v107/i9/e090502](http://prl.aps.org/abstract/PRL/v107/i9/e090502)

[7] R. L. Rivest, “Unconditionally Secure Commitment and Oblivious Transfer Schemes Using Private Channels and a Trusted Initializer,” Nov. 1999.

[8] C. Crèpeau, “Efficient Cryptographic Protocols Based on Noisy Channels,” in *Proc. EUROCRYPT*, 1997, pp. 306–317.

[9] A. Winter, A. C. A. Nascimento, and H. Imai, “Commitment Capacity of Discrete Memoryless Channels,” in *Proc. Cryptography and Coding, Apr. 2003*, pp. 35–51. [Online]. Available: [http://arxiv.org/abs/cs/0304014](http://arxiv.org/abs/cs/0304014)

[10] U. M. Maurer, “A Provably-Secure Strongly-Randomized Cipher,” in *Proc. EUROCRYPT*, 1990, pp. 361–373.

[11] I. B. Damgård, S. Fehr, L. Salvail, and C. Schaffner, “Cryptography in the Bounded-Quantum-Storage Model,” *SIAM J. Comput.*, vol. 37, no. 6, p. 1865, Mar. 2008. [Online]. Available: [http://link.aip.org/link/SMJCAT/v37/i6/p1865/s1&Agg=doi](http://link.aip.org/link/SMJCAT/v37/i6/p1865/s1&Agg=doi)
Jędrzej Kaniewski was born on November 11, 1987 in Warsaw, Poland. He studied at Cambridge University and graduated in 2011 with a B.A. (Natural Sciences) and MMath degrees. He is currently a Ph.D. student at the Centre for Quantum Technologies, Singapore.

Marco Tomamichel was born on March 13, 1981, in St. Gallen (Switzerland). He studied Electrical Engineering at ETH Zurich (Switzerland), where he graduated in 2007 with a M.Sc. in Electrical Engineering and Information Technology degree. He then graduated with a Ph.D. at the Institute of Theoretical Physics at ETH Zurich (Switzerland). He is currently working as a post-doctoral research fellow at the Centre for Quantum Technologies at the National University of Singapore. His research interests include classical and quantum information theory in the non-asymptotic regime, and its applications to cryptography.
**Esther Hänggi** studied at EPF Lausanne (Switzerland) and graduated in 2005 with a Master’s degree in Physics. In 2010 she obtained a doctoral degree in Computer Science from ETH Zurich (Switzerland), before moving to the Centre for Quantum Technologies at the National University of Singapore to work as a post-doctoral research fellow. Her research interests are in the area of quantum information and (quantum) cryptography.

**Stephanie Wehner** is a physicist and computer scientist at the Centre for Quantum Technologies, National University of Singapore, born in Wuerzburg, Germany. She studied at the University of Amsterdam and obtained her Ph.D. at CWI, before moving to Caltech as a post-doctoral researcher under John Preskill. Since 2010 she is an assistant professor in the Department of Computer Science at the National University of Singapore and a Principal Investigator at the Centre for Quantum Technologies.