Chapter 6

Bit-commitment with Non-signaling Adversaries

6.1 Introduction

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

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

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

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

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

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

6.2 Bipartite Systems and Two-Prover Commitments

6.2.1 One-Round Bipartite Systems

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

6.2.2 Two-Round Systems

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

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

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

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

and the following two non-signaling constraints

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

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

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

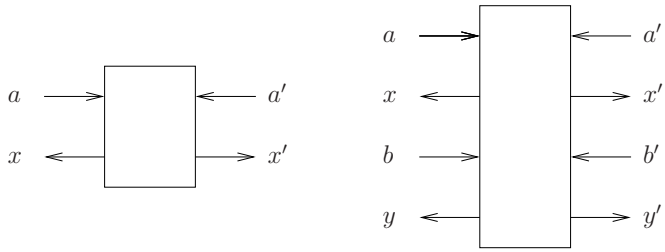

Figure 6.1: A one-round (left) and two-round (right) bipartite system.

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

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

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

6.2.3 Capturing the Non-Signaling Property

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

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

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

This follows from Lemma 6.5 below.

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

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

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

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

and thus $q(x|a,b,c,d) = q(x|a)$. □

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

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

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

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

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

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

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

6.2.4 Two-Prover Commitments

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

6.3 Impossibility of Two-Prover Commitments

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

6.3.1 Simple Schemes

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

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

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

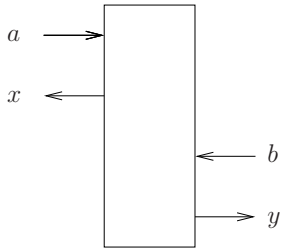

Figure 6.2: The adversaries' strategy $p(x, y|a, b)$ in case of a *simple* commitment scheme.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

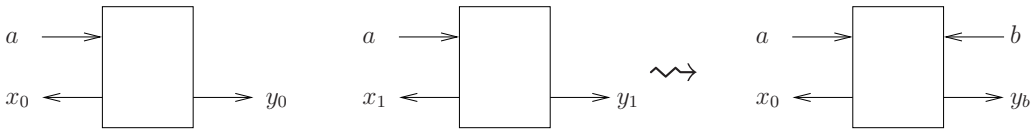

Figure 6.3: Defining the strategy q by gluing together $p(x_0, y_0|a)$ and $p(x_1, y_1|a)$.

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

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

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

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

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

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

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

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

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

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

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

6.3.2 Arbitrary Schemes

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

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

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

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

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

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

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

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

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

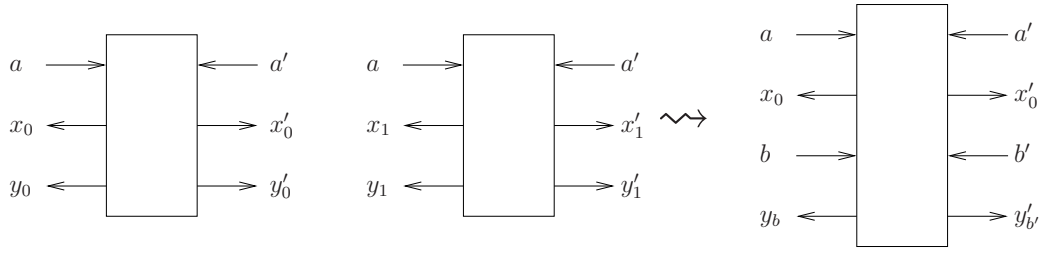

Figure 6.4: Defining the strategy q from $p(x_0, x_0', y_0, y_0'|a, a')$ and $p(x_1, x_1', y_1, y_1'|a, a')$ glued together.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

We now prove the two lemmas:

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

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

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

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

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

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

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

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

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

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

Thus, we set

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

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

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

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

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

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

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

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

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

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

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

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

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

6.3.3 Multi-Round Schemes

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

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

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

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

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

6.4 Possibility of Three-Prover Commitments

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

It follows that

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

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

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

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

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

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

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

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

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