On the Composition of Two-Prover Commitments, and Applications to Multi-Round Relativistic Commitments

(Extended Abstract)

Serge Fehr and Max Fillinger

Centrum Wiskunde & Informatica (CWI), Amsterdam, The Netherlands
{serge.fehr,M.J.Fillinger}@cwi.nl

Abstract. We consider the related notions of two-prover and of relativistic commitment schemes. In recent work, Lunghi *et al.* proposed a relativistic commitment scheme with a *multi-round sustain phase* that enables to keep the binding property alive as long as the sustain phase is running. They prove security of their scheme against classical attacks; however, the proven bound on the error parameter is very weak: it blows up *doubly exponentially* in the number of rounds.
In this work, we give a new analysis of the multi-round scheme of Lunghi *et al.*, and we show a *linear* growth of the error parameter instead (also considering classical attacks only). Our analysis is intuitively much simpler than the analysis provided by Lunghi *et al.* It is based on a new *composition theorem* for two-prover commitment schemes. The proof of our composition theorem is based on a better understanding of the binding property of two-prover commitments that we provide in the form of new definitions and relations among them. These new insights are certainly of independent interest and are likely to be useful in other contexts as well.
Finally, our work gives rise to several interesting open problems, for instance extending our results to the quantum setting, where the dishonest provers are allowed to perform measurements on an entangled quantum state in order to try to break the binding property.

Introduction

Two-Prover Commitment Schemes We consider the notion of *2-prover commitment schemes*, as originally introduced by Ben-Or, Goldwasser, Kilian and Wigderson in their seminal paper [BGKW88]. In a 2-prover commitment scheme, the prover (i.e., the entity that is responsible for preparing and opening the commitment) consists of two agents, P and Q, and it is assumed that these two agents cannot communicate with each other. With this approach, the classical and quantum impossibility results for unconditionally secure commitment schemes can be circumvented.

A simple 2-prover bit commitment schemes is the scheme proposed by Crépeau *et al.* [CSST11], which works as follows. The verifier V chooses a uniformly random $a \in \{0,1\}^n$ and sends it to P, who replies with $x := y \oplus a \cdot b$, where b is the bit to commit to, and $y \in \{0,1\}^n$ is a uniformly random string known (only) to P and Q. Furthermore, "$\oplus$" is bit-wise XOR, and "$\cdot$" is scalar multiplication (of the scalar b with the vector a). In order to open the commitment (to b), Q sends y to V, and V checks if $x \oplus y = a \cdot b$. It is clear that this scheme is hiding: the commitment $x = y \oplus a \cdot b$ is uniformly random and independent of a no matter what b is. On the other hand, the binding property follows from the observation that in order to open the commitment to $b = 0$, Q needs to announce $y = x$, and in order to open to $b = 1$, he needs to announce $y = x \oplus a$. Thus, in order to open to *both*, he must know x *and* $x \oplus a$, and thus a, which is a contradiction to the no-communication assumption, because a was sent to P only.

In the quantum setting, where the dishonest provers are allowed to share an entangled quantum state and can produce x and y by means of performing measurements on their respective parts of the state, the above reasoning for the binding property does not work anymore. Nevertheless, as shown in [CSST11], the binding property still holds (though with a weaker parameter).

Relativistic Commitment Schemes Roughly speaking, the idea of *relativistic commitment schemes*, as introduced by Kent [Ken99,Ken05], is to take a 2-prover commitment schemes as above, and enforce the no-communication assumption by means of relativistic effects: we place P and Q geographically very far apart, and execute the scheme quickly enough, so that by the finiteness of the speed of light, there is not enough time for them to communicate. The obvious downside of such a relativistic commitment scheme is that the binding property stays alive only for a very short time, i.e., the opening has to take place almost immediately after the committing, before the provers have the chance to exchange information.

Motivated by this limitation, Lunghi *et al.* [LKB+14] proposed what they call a *multi-round* scheme, where after the actual commit phase there is a *sustain phase*, during which the provers and the verifier keep exchanging messages, and as long as this sustain phase is running, the commitment stays binding (and hiding), until the commitment is finally opened. Their proposed scheme works as follows. The actual commit protocol is the commit protocol from the Crépeau *et al.* scheme: V sends a uniformly random string $a_0 \in \{0,1\}^n$ to P, who returns $x_0 := y_0 \oplus a_0 \cdot b$. Then, to sustain the commitment, before P has the chance to tell a_0 to Q, V sends a new uniformly random string $a_1 \in \{0,1\}^n$ to Q who replies with $x_1 := y_1 \oplus a_1 \cdot y_0$, where $y_1 \in \{0,1\}^n$ is another random string shared between P and Q, and the multiplication $a_1 \cdot y_0$ is in a suitable finite field. Then, to further sustain the commitment, V sends a new uniformly random string $a_2 \in \{0,1\}^n$ to P who replies with $x_2 := y_2 \oplus a_2 \cdot y_1$, etc. Finally, after the last sustain round where $x_m := y_m \oplus a_m \cdot y_{m-1}$ has been sent to V, in order to finally open the commitment, y_m is sent to V (by the other prover). See Figure 1. In order to verify the opening, V computes $y_{m-1}, y_{m-2}, \ldots, y_0$ inductively in the obvious way, and checks if $x_0 \oplus y_0 = a_0 \cdot b$.

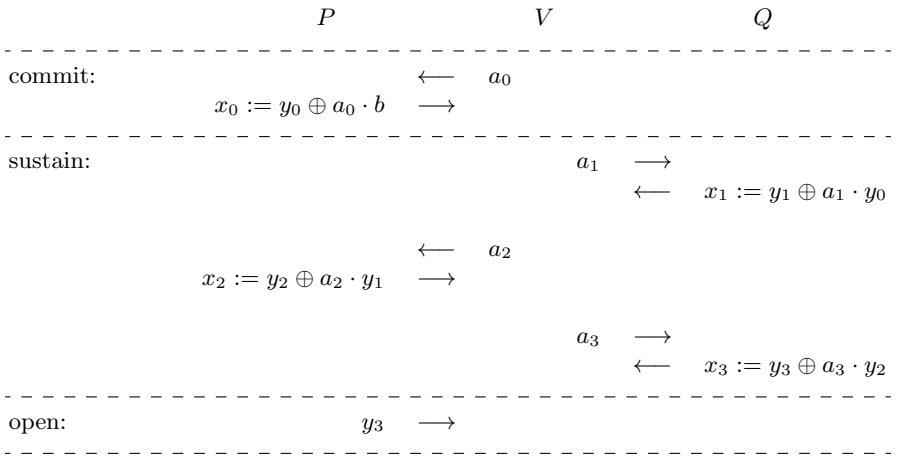

Fig. 1. The Lunghi *et al.* multi-round scheme (for $m = 3$).

What is important is that in round i (say for odd i), when preparing x_i, the prover Q must not know a_{i-1}, but he is allowed to know $a_1, \ldots, a_{i-2}$. Thus, execution must be timed in such a way that between subsequent rounds there is not enough time for the provers to communicate, but they may communicate over multiple rounds.

As for the security of this scheme, it is obvious that the hiding property stays satisfied up to the open phase: every single message V receives is one-time-pad encrypted. As for the binding property, Lunghi *et al.* prove that the scheme with a m-round sustain phase is ε_m-binding against classical attacks, where ε_m satisfies $\varepsilon_0 = 2^{-n}$ (this is just the standard Crépeau *et al.* scheme) and

$$\varepsilon_m \leq \frac{1}{2^{n+1}} + \sqrt{\varepsilon_{m-1}}$$

for $m \geq 1$. Thus, even when reading this recursive formula very liberally by ignoring the $2^{-(n+1)}$ term, we obtain that

$$\varepsilon_m \lesssim \sqrt[2^m]{\varepsilon_0} = 2^{-\frac{n}{2^m}} ,$$

i.e., the error parameter blows up *doubly exponentially.*[1] In other words, in order to have a non-trivial ε_m we need that n, the size of the strings that are communicated is *exponential* in m. This means that Lunghi *et al.* can only afford a very small number of rounds. For instance, in their practical implementation, fixing an error parameter of approximately $10^{-5} \approx 2^{-16}$, they can manage $m = 5$ and $n = 512$; beyond that, i.e. for larger m and thus larger n, the local computation takes too long. This allows them to keep a commitment alive for 2 ms.

[1] Lunghi *et al.* also provide a more complicated recursive formula for ε_m that is slightly better, but the resulting blow-up is still doubly exponential.

2. Our Results

Our main goal is to improve the bound on the binding parameter of the above multi-round scheme. Indeed, our results show that the binding parameter blows up only *linearly* in m, rather than doubly exponentially. Explicitly, our results show that (for classical attacks)

$$\varepsilon_m \leq 2(m+1) \cdot 2^{-\frac{n-1}{2}} \approx m \cdot 2^{-\frac{n}{2}} .$$

Using the same n and error parameter as in the implementation of Lunghi *et al.*, we can now afford approximately $m = 2^{240}$ rounds. Scaling up the 2 ms from the Lunghi *et al.* experiment for 5 rounds gives us a time that is larger (by far) than the age of the universe.

We use the following strategy to obtain our improved bound on ε_m. We observe that the first sustain round can be understood as committing on the opening information y_0 of the actual commitment, using an extended version of the Crépeau *et al.* scheme that commits to a *string* rather than to a bit. Similarly, the second sustain round can be understood as committing on the opening information y_1 of that commitment from the first sustain round, etc. Thus, thinking of the $m = 1$ version of the scheme, what we have to prove is that if we have two commitment schemes $\mathcal{S}$ and $\mathcal{S}'$, and we modify the opening phase of $\mathcal{S}$ in that we first commit to the opening information (using $\mathcal{S}'$) and then open that commitment, then the resulting commitment scheme is still binding; note that, intuitively, this is what one would indeed expect. Given such a general composition theorem, we can then apply it inductively and conclude security (i.e. the binding property) of the Lunghi *et al.* multi-round scheme.

Thus, our main result is such a composition theorem, which shows that if $\mathcal{S}$ and $\mathcal{S}'$ are respectively ε- and δ-binding (against classical attacks), then the composed scheme is $(\varepsilon + \delta)$-binding (against classical attacks), under some mild assumptions on $\mathcal{S}$ and $\mathcal{S}'$. Hence, the binding parameters simply add up; this is what gives us the linear growth. The proof of our composition theorem crucially relies on a new definition of the binding property of 2-prover commitment schemes, which seems to be handier to work with, but is actually equivalent to the $p_0 + p_1 \leq 1 + \varepsilon$ definition as for instance used by Lunghi *et al.*

One subtle issue is that the extended version of the Crépeau *et al.* scheme to strings, as it is used in the sustain phase, is not a fully secure string commitment scheme. The reason is that for *any* y that may be announced in the opening phase, there exists a string s such that $x \oplus y = a \cdot s$; as such, the provers can commit to some fixed string, and then can still decide to either open the commitment to that string (by running the opening phase honestly), or to open it to a random string that is out of their control (by announcing a random y). We deal with this by also introducing a *weak* version of the binding property, which captures this limited freedom for the provers, and we show that it is satisfied by the (extended version of the) Crépeau *et al.* scheme and that our composition theorem also holds for this weak version.[2] Finally, we observe that the composed weakly-binding string commitment scheme is a (strongly) binding *bit* commitment scheme in the natural way (i.e., when restricting the domain to a bit).

As such, we feel that our techniques and insights not only give rise to a drastically improved analysis of the Lunghi *et al.* multi-round scheme, but they significantly improve our understanding of the security of 2-prover commitment schemes, and as such are likely to find further applications.

3. Open Problems

Our work gives rise to a list of interesting open problems. For instance, our composition theorem only applies to pairs $\mathcal{S}, \mathcal{S}'$ of commitment schemes of a certain restricted form, e.g., only one prover should be involved in the commit phase (as it is the case in the Crépeau *et al.* scheme). Our proof crucially relies on this, but there seems to be no fundamental reason for such a restriction. Thus, we wonder if it is possible to generalize our composition theorem to a larger class of pairs of schemes, or, ultimately, to *all* pairs of schemes (that "fit together"). In another direction, some of our observations and results generalize immediately to the quantum setting, where the two dishonest provers are allowed to compute their messages by performing measurements on an entangled quantum state, but in particular our main result, the composition theorem, does not generalize. Also here, there seems to be no fundamental reason, and thus, generalizing our composition theorem to the quantum setting is an interesting open problem. Finally, in order to obtain security of the Lunghi *et al.* multi-round scheme against quantum attacks, beyond a quantum version of the composition theorem, one also needs to prove security against quantum attacks of the (extended version of the) original Crépeau *et al.* scheme as a (weakly binding) *string* commitment scheme; by [CSST11] we merely know its "quantum security" as a *bit* commitment scheme.

[2] Actually, we only prove it for this weak version, but the proof for the strong version goes along the same lines.

References

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

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

[Ken99] Adrian Kent. Unconditionally secure bit commitment. *Physical Review Letters*, 83(7):1447–1450, 1999.

[Ken05] Adrian Kent. Secure classical bit commitment using fixed capacity communication channels. *Journal of Cryptology*, 18(4):313–335, 2005. Available from: `http://arxiv.org/abs/quant-ph/9906103`.

[LKB+14] Tomaso Lunghi, Jędrzej Kaniewski, Felix Bussières, Raphael Houlmann, Marco Tomamichel, Stephanie Wehner, and Hugo Zbinden. Practical relativistic bit commitment. *ArXiv e-prints*, 2014. `http://arxiv.org/abs/1411.4917`.

On the Composition of Two-Prover Commitments, and Applications to Multi-Round Relativistic Commitments

(Full Version)

Serge Fehr and Max Fillinger

Centrum Wiskunde & Informatica (CWI), Amsterdam, The Netherlands
{serge.fehr,M.J.Fillinger}@cwi.nl

Abstract. We consider the related notions of two-prover and of relativistic commitment schemes. In recent work, Lunghi *et al.* proposed a relativistic commitment scheme with a *multi-round sustain phase* that enables to keep the binding property alive as long as the sustain phase is running. They prove security of their scheme against classical attacks; however, the proven bound on the error parameter is very weak: it blows up *doubly exponentially* in the number of rounds.
In this work, we give a new analysis of the multi-round scheme of Lunghi *et al.*, and we show a *linear* growth of the error parameter instead (also considering classical attacks only). Our analysis is intuitively much simpler than the analysis provided by Lunghi *et al.* It is based on a new *composition theorem* for two-prover commitment schemes. The proof of our composition theorem is based on a better understanding of the binding property of two-prover commitments that we provide in the form of new definitions and relations among them. These new insights are certainly of independent interest and are likely to be useful in other contexts as well.
Finally, our work gives rise to several interesting open problems, for instance extending our results to the quantum setting, where the dishonest provers are allowed to perform measurements on an entangled quantum state in order to try to break the binding property.

1 Introduction

Two-Prover Commitment Schemes. We consider the notion of *2-prover commitment schemes*, as originally introduced by Ben-Or, Goldwasser, Kilian and Wigderson in their seminal paper [BGKW88]. In a 2-prover commitment scheme, the prover (i.e., the entity that is responsible for preparing and opening the commitment) consists of two agents, P and Q, and it is assumed that these two agents cannot communicate with each other. With this approach, the classical and quantum impossibility results for unconditionally secure commitment schemes can be circumvented.

A simple 2-prover bit commitment schemes is the scheme proposed by Crépeau *et al.* [CSST11], which works as follows. The verifier V chooses a uniformly random $a \in \{0,1\}^n$ and sends it to P, who replies with $x := y \oplus a \cdot b$, where b is the bit to commit to, and $y \in \{0,1\}^n$ is a uniformly random string known (only) to P and Q. Furthermore, "$\oplus$" is bit-wise XOR, and "$\cdot$" is scalar multiplication (of the scalar b with the vector a). In order to open the commitment (to b), Q sends y to V, and V checks if $x \oplus y = a \cdot b$. It is clear that this scheme is hiding: the commitment $x = y \oplus a \cdot b$ is uniformly random and independent of a no matter what b is. On the other hand, the binding property follows from the observation that in order to open the commitment to $b = 0$, Q needs to announce $y = x$, and in order to open to $b = 1$, he needs to announce $y = x \oplus a$. Thus, in order to open to *both*, he must know x *and* $x \oplus a$, and thus a, which is a contradiction to the no-communication assumption, because a was sent to P only.

In the quantum setting, where the dishonest provers are allowed to share an entangled quantum state and can produce x and y by means of performing measurements on their respective parts of the state, the above reasoning for the binding property does not work anymore. Nevertheless, as shown in [CSST11], the binding property still holds (though with a weaker parameter).

Relativistic Commitment Schemes. Roughly speaking, the idea of *relativistic commitment schemes*, as introduced by Kent [Ken99,Ken05], is to take a 2-prover commitment schemes as above, and enforce the no-communication assumption by means of relativistic effects: we place P and Q geographically very far apart, and execute the scheme quickly enough, so that by the finiteness of the speed of light, there is not enough time for them to communicate. The obvious downside of such a relativistic commitment scheme is that the binding property stays alive only for a very short time, i.e., the opening has to

take place almost immediately after the committing, before the provers have the chance to exchange information.

Motivated by this limitation, Lunghi *et al.* [LKB+14] proposed what they call a *multi-round* scheme, where after the actual commit phase there is a *sustain phase*, during which the provers and the verifier keep exchanging messages, and as long as this sustain phase is running, the commitment stays binding (and hiding), until the commitment is finally opened. Their proposed scheme works as follows. The actual commit protocol is the commit protocol from the Crépeau *et al.* scheme: V sends a uniformly random string $a_0 \in \{0,1\}^n$ to P, who returns $x_0 := y_0 \oplus a_0 \cdot b$. Then, to sustain the commitment, before P has the chance to tell a_0 to Q, V sends a new uniformly random string $a_1 \in \{0,1\}^n$ to Q who replies with $x_1 := y_1 \oplus a_1 \cdot y_0$, where $y_1 \in \{0,1\}^n$ is another random string shared between P and Q, and the multiplication $a_1 \cdot y_0$ is in a suitable finite field. Then, to further sustain the commitment, V sends a new uniformly random string $a_2 \in \{0,1\}^n$ to P who replies with $x_2 := y_2 \oplus a_2 \cdot y_1$, etc. Finally, after the last sustain round where $x_m := y_m \oplus a_m \cdot y_{m-1}$ has been sent to V, in order to finally open the commitment, y_m is sent to V (by the other prover). See Figure 1. In order to verify the opening, V computes $y_{m-1}, y_{m-2}, \ldots, y_0$ inductively in the obvious way, and checks if $x_0 \oplus y_0 = a_0 \cdot b$.

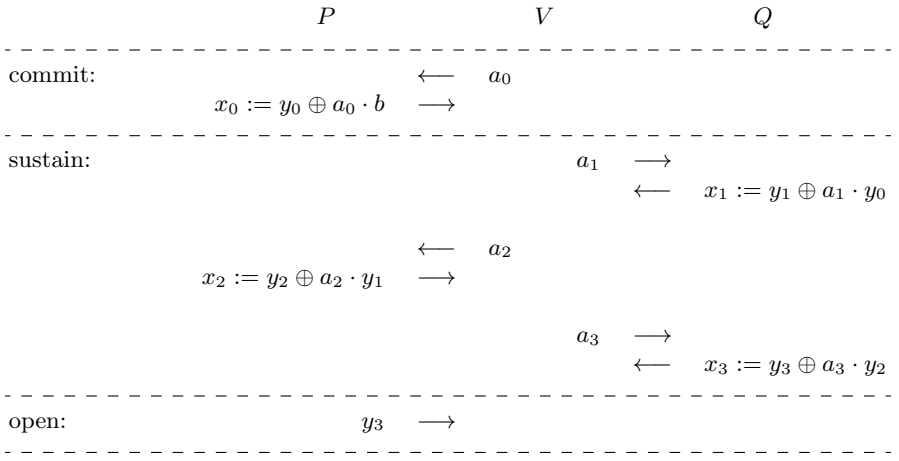

Fig. 1. The Lunghi *et al.* multi-round scheme (for $m = 3$).

What is important is that in round i (say for odd i), when preparing x_i, the prover Q must not know a_{i-1}, but he is allowed to know $a_1, \ldots, a_{i-2}$. Thus, execution must be timed in such a way that between subsequent rounds there is not enough time for the provers to communicate, but they may communicate over multiple rounds.

As for the security of this scheme, it is obvious that the hiding property stays satisfied up to the open phase: every single message V receives is one-time-pad encrypted. As for the binding property, Lunghi *et al.* prove that the scheme with a m-round sustain phase is ε_m-binding against classical attacks, where ε_m satisfies $\varepsilon_0 = 2^{-n}$ (this is just the standard Crépeau *et al.* scheme) and

$$\varepsilon_m \leq \frac{1}{2^{n+1}} + \sqrt{\varepsilon_{m-1}}$$

for $m \geq 1$. Thus, even when reading this recursive formula very liberally by ignoring the $2^{-(n+1)}$ term, we obtain that

$$\varepsilon_m \lesssim \sqrt[2^m]{\varepsilon_0} = 2^{-\frac{n}{2^m}},$$

i.e., the error parameter blows up *doubly exponentially.*[1] In other words, in order to have a non-trivial ε_m we need that n, the size of the strings that are communicated is *exponential* in m. This means that Lunghi *et al.* can only afford a very small number of rounds. For instance, in their practical implementation, fixing an error parameter of approximately $10^{-5} \approx 2^{-16}$, they can manage $m = 5$ and $n = 512$; beyond that, i.e. for larger m and thus larger n, the local computation takes too long. This allows them to keep a commitment alive for 2 ms.

[1] Lunghi *et al.* also provide a more complicated recursive formula for ε_m that is slightly better, but the resulting blow-up is still doubly exponential.

Our Results. Our main goal is to improve the bound on the binding parameter of the above multi-round scheme. Indeed, our results show that the binding parameter blows up only *linearly* in m, rather than doubly exponentially. Explicitly, our results show that (for classical attacks)

$$\varepsilon_m \leq 2(m+1) \cdot 2^{-\frac{n-1}{2}} \approx m \cdot 2^{-\frac{n}{2}} .$$

Using the same n and error parameter as in the implementation of Lunghi *et al.*, we can now afford approximately $m = 2^{240}$ rounds. Scaling up the 2 ms from the Lunghi *et al.* experiment for 5 rounds gives us a time that is larger (by far) than the age of the universe.

We use the following strategy to obtain our improved bound on ε_m. We observe that the first sustain round can be understood as committing on the opening information y_0 of the actual commitment, using an extended version of the Crépeau *et al.* scheme that commits to a *string* rather than to a bit. Similarly, the second sustain round can be understood as committing on the opening information y_1 of that commitment from the first sustain round, etc. Thus, thinking of the $m = 1$ version of the scheme, what we have to prove is that if we have two commitment schemes $\mathcal{S}$ and $\mathcal{S}'$, and we modify the opening phase of $\mathcal{S}$ in that we first commit to the opening information (using $\mathcal{S}'$) and then open that commitment, then the resulting commitment scheme is still binding; note that, intuitively, this is what one would indeed expect. Given such a general composition theorem, we can then apply it inductively and conclude security (i.e. the binding property) of the Lunghi *et al.* multi-round scheme.

Thus, our main result is such a composition theorem, which shows that if $\mathcal{S}$ and $\mathcal{S}'$ are respectively ε- and δ-binding (against classical attacks), then the composed scheme is $(\varepsilon + \delta)$-binding (against classical attacks), under some mild assumptions on $\mathcal{S}$ and $\mathcal{S}'$. Hence, the binding parameters simply add up; this is what gives us the linear growth. The proof of our composition theorem crucially relies on a new definition of the binding property of 2-prover commitment schemes, which seems to be handier to work with, but is actually equivalent to the $p_0 + p_1 \leq 1 + \varepsilon$ definition as for instance used by Lunghi *et al.*

One subtle issue is that the extended version of the Crépeau *et al.* scheme to strings, as it is used in the sustain phase, is not a fully secure string commitment scheme. The reason is that for *any* y that may be announced in the opening phase, there exists a string s such that $x \oplus y = a \cdot s$; as such, the provers can commit to some fixed string, and then can still decide to either open the commitment to that string (by running the opening phase honestly), or to open it to a random string that is out of their control (by announcing a random y). We deal with this by also introducing a *weak* version of the binding property, which captures this limited freedom for the provers, and we show that it is satisfied by the (extended version of the) Crépeau *et al.* scheme and that our composition theorem also holds for this weak version.[2] Finally, we observe that the composed weakly-binding string commitment scheme is a (strongly) binding *bit* commitment scheme in the natural way (i.e., when restricting the domain to a bit).

As such, we feel that our techniques and insights not only give rise to a drastically improved analysis of the Lunghi *et al.* multi-round scheme, but they significantly improve our understanding of the security of 2-prover commitment schemes, and as such are likely to find further applications.

Open Problems. Our work gives rise to a list of interesting open problems. For instance, our composition theorem only applies to pairs $\mathcal{S}, \mathcal{S}'$ of commitment schemes of a certain restricted form, e.g., only one prover should be involved in the commit phase (as it is the case in the Crépeau *et al.* scheme). Our proof crucially relies on this, but there seems to be no fundamental reason for such a restriction. Thus, we wonder if it is possible to generalize our composition theorem to a larger class of pairs of schemes, or, ultimately, to *all* pairs of schemes (that "fit together"). In another direction, some of our observations and results generalize immediately to the quantum setting, where the two dishonest provers are allowed to compute their messages by performing measurements on an entangled quantum state, but in particular our main result, the composition theorem, does not generalize. Also here, there seems to be no fundamental reason, and thus, generalizing our composition theorem to the quantum setting is an interesting open problem. Finally, in order to obtain security of the Lunghi *et al.* multi-round scheme against quantum attacks, beyond a quantum version of the composition theorem, one also needs to prove security against quantum attacks of the (extended version of the) original Crépeau *et al.* scheme as a (weakly binding) *string* commitment scheme; by [CSST11] we merely know its "quantum security" as a *bit* commitment scheme.

[2] Actually, we only prove it for this weak version, but the proof for the strong version goes along the same lines.

2 Preliminaries

2.1 Basic Notation

Probability Distributions. For the purpose of this work, a *(probability) distribution* is a function $p : \mathcal{X} \to [0,1]$, $x \mapsto p(x)$, where $\mathcal{X}$ is a finite non-empty set, with the property that $\sum_{x\in\mathcal{X}} p(x) = 1$. For specific choices $x_\circ \in \mathcal{X}$, we tend to write $p(x = x_\circ)$ instead of $p(x_\circ)$. For any subset $\Lambda \subset \mathcal{X}$, the probability $p(\Lambda)$ is naturally defined as $p(\Lambda) = \sum_{x\in\Lambda} p(x)$, and it holds that

$$p(\Lambda) + p(\Gamma) = p(\Lambda \cup \Gamma) + p(\Lambda \cap \Gamma) \leq 1 + p(\Lambda \cap \Gamma)$$

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

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

for all $\Lambda_1, \ldots, \Lambda_k \subset \mathcal{X}$. Similarly, for a distribution $p : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ on two (and similarly more) variables, probabilities like $p(x{=}y)$, $p(x{=}f(y))$, $p(x{\neq}y)$ etc. are naturally understood as

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

etc., and the *marginals* $p(x)$ and $p(y)$ are given by $p(x) = \sum_y p(x,y)$ and $p(y) = \sum_x p(x,y)$, respectively. Vice versa, given two distributions $p(x)$ and $p(y)$, we say that a distribution $p(x,y)$ on two variables is a *consistent joint distribution* if the two marginals of $p(x,y)$ coincide with $p(x)$ and $p(y)$, respectively. We will make use of the following property on the existence of a consistent joint distributions that maximizes the probability that $x = y$; the proof is given in the appendix.

Lemma 2.1. *Let $p(x)$ and $p(y)$ be two distribution on a common set $\mathcal{X}$. Then there exists a consistent joint distribution $p(x,y)$ such that $p(x = y = x_\circ) = \min\{p(x = x_\circ), p(y = x_\circ)\}$ for all $x_\circ \in \mathcal{X}$.*

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

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

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

For the purpose of this work, the algorithms prot_P and prot_Q for P and Q will be *deterministic* (i.e., have no local randomness), and all randomness will be provided by a "resource" res_{PQ}, which produces (two copies of) a uniformly random bit string r (of suitable length), and prot_P and prot_Q take r as (additional) input. Some, but not all, of our results extend to the quantum setting, where res_{PQ} produces a *bipartite entangled quantum state*, and prot_P and prot_Q produce their (classical) messages and respective (classical) outputs by performing *quantum operations*.

We will often consider protocols for which P and Q do not communicate (or their communication is limited). This simply means that all the messages that P prepares for Q are "empty" (i.e., some special symbol $\emptyset$).

We can *compose* two (interactive) algorithms prot_P and prot'_P in the obvious way, by applying prot'_P to the output of prot_P. The resulting interactive algorithm is denoted as $\mathsf{prot}_P \circ \mathsf{prot}'_P$. Composing the respective algorithms of two protocols $\mathsf{prot}_{PQV} = (\mathsf{prot}_P, \mathsf{prot}_Q, \mathsf{prot}_V)$ and $\mathsf{prot}'_{PQV} = (\mathsf{prot}'_P, \mathsf{prot}'_Q, \mathsf{prot}'_V)$ results in the composed protocol $\mathsf{prot}_{PQV} \circ \mathsf{prot}'_{PQV}$.

2.2 2-Prover Commitment Schemes

Definition 2.2. *A* 2-prover (string) commitment scheme $\mathcal{S}$ *consists of a* resource res_{PQ}*, and of two interactive protocols* $\mathsf{com}_{PQV} = (\mathsf{com}_P, \mathsf{com}_Q, \mathsf{com}_V)$ *and* $\mathsf{open}_{PQV} = (\mathsf{open}_P, \mathsf{open}_Q, \mathsf{open}_V)$ *between the two provers* P *and* Q *and the verifier* V*, with the following semantics. The resource* res_{PQ} *outputs a bit string* r*, chosen according to some distribution, to* P *and* Q*:*

$$(r\|r) \leftarrow \mathsf{res}_{PQ}\,.$$

The commit protocol com_{PQV} *takes as input the bit string* r *and a bit string* $s \in \{0,1\}^n$ *for* P *and* Q *(and no input for* V*), and outputs a* commitment *com to* V*, and some state information to* P *and* Q*:*

$$(com\|state_P\|state_Q) \leftarrow (\mathsf{com}_P(s,r)\|\mathsf{com}_Q(s,r)\|\mathsf{com}_V)\,.$$

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

$$(\emptyset\|\emptyset\|s) \leftarrow (\mathsf{open}_P(state_P)\|\mathsf{open}_Q(state_Q)\|\mathsf{open}_V(com))$$

with $s \in \{0,1\}^n \cup \{\perp\}$*. The set* $\{0,1\}^n$ *is called the* domain *of* $\mathcal{S}$*; if* $n = 1$ *then we refer to* $\mathcal{S}$ *as a* bit *commitment scheme instead, and we tend to use* b *rather than* s *to denote the committed bit.*

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

Definition 2.3. *A 2-prover commitment scheme is* η-sound *if in an honest execution* V*'s output* s *of* open_{PQV} *equals* P *and* Q*'s input* s *to* com_{PQV} *except with probability* η*. A* 0*-sound scheme is also called* perfectly sound*.*
A 2-prover commitment scheme is δ-hiding *if for any commit strategy* $\overline{\mathsf{com}}_V$*, the distribution of the commitment com, produced as* $(com\|state_P\|state_Q) \leftarrow (\mathsf{com}_P(s,r)\|\mathsf{com}_Q(s,r)\|\overline{\mathsf{com}}_V)$ *with* $(r\|r) \leftarrow \mathsf{res}_{PQ}$*, is* δ*-almost independent of* P *and* Q*'s input* s*, in the sense that the distributions* $p(s, com)$ *and* $p(s)\cdot p(com)$ *are* δ*-close in terms of statistical distance. A* 0*-hiding scheme is also called* perfectly hiding*.*

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

What exactly "allowed" means depends on the model and needs to be specified. Typically, in the 2-prover setting, we only allow strategies $\overline{\mathsf{com}}_{PQ}$ and $\overline{\mathsf{open}}_{PQ}$ with *no communication* between the two provers, but we may also be more liberal and allow some *well-controlled* communication, as we will see later. Furthermore, we may restrict to *classical* attacks, or we can consider *quantum* attacks. In the former case, $\overline{\mathsf{res}}_{PQ}$ produces shared randomness (as res_{PQ} does), and $\overline{\mathsf{com}}_P$ and $\overline{\mathsf{com}}_Q$ are classical interactive algorithms that compute the outgoing messages (and outputs) as deterministic functions[3] of their respective inputs and the incoming messages, and the same for $\overline{\mathsf{open}}_P$ and $\overline{\mathsf{open}}_Q$. In the case of quantum attacks, $\overline{\mathsf{res}}_{PQ}$ generates a bipartite quantum state and $\overline{\mathsf{com}}_P, \overline{\mathsf{com}}_Q, \overline{\mathsf{open}}_P$ and $\overline{\mathsf{open}}_Q$ produce their messages by performing measurements (that depend on the input and the previous incoming messages) on this quantum state. Our main result holds for classical attacks only, and so the unfamiliar reader can safely ignore the possibility of quantum attacks, but some of our results also hold for quantum attacks.

A commonly accepted definition for the binding property of a 2-prover *bit* commitment scheme, as it is for instance used in [LKB+14] (up to the factor 2 in the error parameter), is as follows. We assume it has been specified which attacks are *allowed* (e.g. those where P and Q do not communicate at all).

Definition 2.4. *A 2-prover bit commitment scheme is* ε-binding (in the sense of $p_0 + p_1 \leq 1 + 2\varepsilon$) *if for every allowed resource* $\overline{\mathsf{res}}_{PQ}$ *and commit strategy* $\overline{\mathsf{com}}_{PQ}$*, and for every pair of allowed opening strategies* $\overline{\mathsf{open}}^0_{PQ}$ *and* $\overline{\mathsf{open}}^1_{PQ}$*, fixing respective distributions* $p_0(b)$ *and* $p_1(b)$*, it holds that*

$$p_0(b=0) + p_1(b=1) \leq 1 + 2\varepsilon\,.$$

[3] It is without loss of generality to restrict to such *deterministic* protocols, because all the necessary randomness can be taken from the resource.

Example 2.5. Our main working example is the bit commitment scheme by Crépeau *et al*, where res_{PQ} produces a uniformly random $r \in \{0,1\}^n$ as shared randomness, com_{PQV} instructs V to sample and send to P a uniformly random $a \in \{0,1\}^n$ and P returns $x := r \oplus a \cdot b$ to V, where b is the bit to commit to, and open_{PQV} instructs Q to send $y := r$ to V and V outputs the (smaller) bit b that satisfies $x \oplus y = a \cdot b$, and $b := \perp$ in case no such bit exists. It is easy to see that the scheme is 2^{-n}-sound and perfectly hiding (soundness fails in case $a = 0$).

For *classical* provers that do not communicate at all, the scheme is $\frac{1}{2} \cdot 2^{-n}$ binding in the sense of $p_0 + p_1 \leq 1 + 2^{-n}$, i.e. according to Definition 2.4, and for *quantum* provers that do not communicate at all, the scheme is $2^{-n/2}$-binding in the sense of $p_0 + p_1 \leq 1 + 2 \cdot 2^{-n/2}$.

We also want to consider an extended version of the scheme, where the bit b is replaced by a string $s \in \{0,1\}^n$ in the obvious way (where the multiplication $a \cdot s$ is then understood in a suitable finite field), and we want to appreciate this extension as a *string* commitment scheme; we will refer to this scheme as $\mathcal{CSST}^n$. However, it is a priori not clear what definition is suitable for the binding property. Furthermore, in this particular scheme, the dishonest provers can always honestly commit to a string s, and can then decide to honestly open the commitment to s, or open to a *random* string by announcing a randomly chosen y — any y satisfies $x \oplus y = a \cdot s$ for *some* s (unless $a = 0$, which almost never happens).

3 On the Binding Property of 2-Prover Commitment Schemes

We introduce a new definition for the binding property of 2-prover commitment schemes. In the case of *bit* commitment schemes, it is equivalent to Definition 2.4, as we will show. However, we feel that our definition is closer to the intuition of what is expected from a commitment scheme, and as such it is easier to work with. Indeed, the proof of our composition result is heavily based on our new definition. We also introduce the notion of a *weak* binding property, which captures the binding property that is satisfied by the string commitment scheme $\mathcal{CSST}^n$.

Throughout this section, when quantifying over attacks against (the binding property of) a scheme $\mathcal{S}$, it is always understood that there is a notion of *allowed* attacks for that scheme (e.g., all attacks for which P and Q do not communicate), and that the quantification is over all such allowed attacks. Also, even though our focus is on classical attacks, much of Sections 3.1 and 3.2 hold in case of quantum attacks too, and we make it explicit when we (have to) restrict to classical attacks.

3.1 Defining The Binding Property

Intuitively, we say that a scheme is binding if after the commit phase there exists a string $\hat{s}$ so that no matter what the provers do in the opening phase, the verifier will output either $s = \hat{s}$ or $s = \perp$ (except with small probability). Formally, the definition is not in terms of $\hat{s}$, but in terms of its *distribution.*

Definition 3.1 (Binding property). *A 2-prover commitment scheme $\mathcal{S}$ is* ε-binding *if for all resources $\overline{\mathsf{res}}_{PQ}$ and commit strategies $\overline{\mathsf{com}}_{PQ}$ there exists a distribution $p(\hat{s})$ such that for every opening strategy $\overline{\mathsf{open}}_{PQ}$ (which then fixes the distribution $p(s)$ of V's output s) there is a consistent joint distribution $p(\hat{s}, s)$ such that $p(s \neq \hat{s} \wedge s \neq \perp) \leq \varepsilon$. In short:*

$$\forall\, \overline{\mathsf{res}}_{PQ}, \overline{\mathsf{com}}_{PQ}\ \exists\, p(\hat{s})\ \forall\, \overline{\mathsf{open}}_{PQ}\ \exists\, p(\hat{s}, s) : p(s \neq \hat{s} \wedge s \neq \perp) \leq \varepsilon\,. \tag{2}$$

The string commitment scheme $\mathcal{CSST}^n$ does *not* satisfy this definition (the bit commitment version does, as we will show): after the commit phase, the provers can still decide to open the commitment to a *fixed* string, chosen before the commit phase, or to a *random* string that is out of their control. We capture this by the following relaxed version of the binding property. In this relaxed version, we allow V's output s to be different to $\hat{s}$ and $\perp$, but in this case the provers should have little control over s: for any fixed *target string* $s_\circ$, it should be unlikely that $s = s_\circ$. Formally, this is captured as follows; we will show in Section 3.3 that $\mathcal{CSST}^n$ is weakly binding in this sense.

Definition 3.2 (Weak binding property). *A 2-prover commitment scheme $\mathcal{S}$ is* weakly ε-binding *if for all resources $\overline{\mathsf{res}}_{PQ}$ and commit strategies $\overline{\mathsf{com}}_{PQ}$ there exists a distribution $p(\hat{s})$ such that for every opening strategy $\overline{\mathsf{open}}_{PQ}$ (which then fixes the distribution $p(s)$ of V's output s) there is a consistent joint distribution $p(\hat{s}, s)$ so that for all $s_\circ \in \{0,1\}^n$ it holds that $p(s \neq \hat{s} \wedge s = s_\circ) \leq \varepsilon$. In short:*

$$\forall\, \overline{\mathsf{res}}_{PQ}, \overline{\mathsf{com}}_{PQ}\ \exists\, p(\hat{s})\ \forall\, \overline{\mathsf{open}}_{PQ}\ \exists\, p(\hat{s}, s)\ \forall\, s_\circ : p(s \neq \hat{s}\ \wedge\ s = s_\circ) \leq \varepsilon\,. \tag{3}$$

Remark 3.3. For weakly binding string commitment schemes (with large enough n), we may actually assume that V never outputs $\perp$, since he may output a randomly chosen string s instead.

Remark 3.4. For the (weak or ordinary) binding property, when considering classical attacks, it is sufficient to consider *deterministic* attacks, where the resource $\overline{\mathsf{res}}_{PQ}$ is trivial: $(\emptyset\|\emptyset) \leftarrow \overline{\mathsf{res}}_{PQ}$ (and $\overline{\mathsf{com}}_{PQ}$ and $\overline{\mathsf{open}}_{PQ}$ are deterministic). To see this, note that for a classical attack with a non-trivial resource $\overline{\mathsf{res}}_{PQ}$, every possible output r of $\overline{\mathsf{res}}_{PQ}$ induces a deterministic attack (which runs $\overline{\mathsf{com}}_{PQ}$ and $\overline{\mathsf{open}}_{PQ}$ on that particular choice of r). The binding property for deterministic attacks then implies the existence of distributions $p_r(\hat{s})$ and $p_r(\hat{s}, s)$ as required (in particular, $p_r(\hat{s})$ does not depend on $\overline{\mathsf{open}}_{PQ}$). It is then straightforward to verify that the distributions $p(\hat{s}) = \sum_r p(r) \cdot p_r(\hat{s})$ and $p(\hat{s}, s) = \sum_r p(r) \cdot p_r(\hat{s}, s)$ are as required for the original (randomized) attack.

Remark 3.5. Clearly, the ordinary binding property (i.e., binding in the sense of Definition 3.1) implies the weak binding property. Furthermore, in the case of *bit* commitment schemes it obviously holds that $p(b \neq \hat{b} \wedge b \neq \perp) = p(b \neq \hat{b} \wedge b = 0) + p(b \neq \hat{b} \wedge b = 1)$, and thus the weak binding property implies the ordinary one, up to a factor-2 loss. Furthermore, every weakly binding *string* commitment scheme gives rise to a ordinary-binding *bit* commitment scheme in a natural way, as shown by the following proposition.

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

Proof. Fix some $\overline{\mathsf{res}}_{PQ}$ and $\overline{\mathsf{com}}_{PQ}$ for $\mathcal{S}'$ and note that these can also be used to attack $\mathcal{S}$. Thus, there exists a distribution $p(\hat{s})$ as in Definition 3.2. We define a distribution $p(\hat{b}, \hat{s})$ by letting $\hat{b} = 0$ if $\hat{s} = s_0$ and $\hat{b} = 1$ otherwise. This defines $p(\hat{b})$ by taking the corresponding marginal. Now fix an opening strategy $\overline{\mathsf{open}}_{PQ}$ for $\mathcal{S}'$, which again is also a strategy against $\mathcal{S}$. Thus, it gives rise to a distribution $p(\hat{s}, s)$ such that $p(\hat{s} \neq s = s_\circ) \leq \varepsilon$ for any $s_\circ$ (and in particular $s_\circ = s_0$ or s_1). We define the distribution $p(\hat{b}, b, \hat{s}, s) = p(\hat{s}, s)p(b|s)p(\hat{b}|\hat{s})$ which gives us the desired distribution $p(\hat{b}, b)$. Indeed:

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

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

3.2 Relation To The Standard Definition

For bit commitment schemes, our binding property is equivalent to the $(p_0 + p_1)$-definition.

Theorem 3.7. *A bit-commitment scheme is ε-binding in the sense of $p_0 + p_1 \leq 1 + 2\varepsilon$ if and only if it is ε-binding (in the sense of Definition 3.1).*

Proof. First, consider a scheme that is ε-binding according to Definition 2.4. Fix a resource $\overline{\mathsf{res}}_{PQ}$, a commitment strategy $\overline{\mathsf{com}}_{PQ}$ and opening strategies $\overline{\mathsf{open}}^0_{PQ}$ and $\overline{\mathsf{open}}^1_{PQ}$ so that $p_0 = p(b_0 = 0)$ and $p_1 = p(b_1 = 1)$ are maximized, where $b_i \in \{0,1\} \cup \{\perp\}$ is V's output when the dishonest provers use opening strategy $\overline{\mathsf{open}}^i_{PQ}$. Let $p_0 + p_1 = 1 + 2\varepsilon'$. Since the scheme is ε-binding, we have $\varepsilon' \leq \varepsilon$. We define the distribution $p(\hat{b})$ as $p(\hat{b} = 0) := p_0 - \varepsilon'$ and $p(\hat{b} = 1) := p_1 - \varepsilon'$. To see that this is indeed a probability distribution, note that $p_0, p_1 \geq 2\varepsilon'$ (otherwise, we would have $p_0 > 1$ or $p_1 > 1$) and that $p(\hat{b} = 0) + p(\hat{b} = 1) = p_0 + p_1 - 2\varepsilon' = 1$. Now we consider an arbitrary opening strategy $\overline{\mathsf{open}}_{PQ}$ which fixes a distribution $p(b)$. By definition of p_0 and p_1, we have $p(b = i) \leq p_i$ and thus $p(b = i) \leq p(\hat{b} = i) + \varepsilon' \leq p(\hat{b} = i) + \varepsilon$. By Lemma 2.1, we conclude that there exists a consistent joint distribution $p(\hat{b}, b)$ with the property that $p(\hat{b} = b = i) = \min\{p(b = i), p(\hat{b} = i)\} \geq p(b = i) - \varepsilon$, and thus that $p(\hat{b} \neq b = i) = p(b = i) - p(\hat{b} = b = i) \leq \varepsilon$ for $i \in \{0,1\}$. This proves one direction of our claim.

For the other direction, consider a scheme that is ε-binding. Fix $\overline{\mathsf{res}}_{PQ}$ and $\overline{\mathsf{com}}_{PQ}$ and let $p(\hat{b})$ be a distribution such that for every opening strategy $\overline{\mathsf{open}}_{PQ}$, there is a joint distribution $p(\hat{b}, b)$ with

$p(\hat{b} \neq b \neq \perp) \leq \varepsilon$. Now consider two opening strategies $\overline{\mathsf{open}}^0_{PQ}$ and $\overline{\mathsf{open}}^1_{PQ}$ which give distributions $p(b_0)$ and $p(b_1)$. We need to bound $p(b_0 = 0) + p(b_1 = 1)$. There is a joint distribution $p(\hat{b}, b_0)$ such that $p(\hat{b} \neq b_0 \neq \perp) \leq \varepsilon$ and likewise for b_1. Thus,

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

which proves the other direction. □

3.3 Security of $\mathcal{CSST}^n$

In this section, we show that $\mathcal{CSST}^n$ is a weakly binding string commitment scheme against *classical* attacks.[4] To this end, we introduce yet another version of the binding property (which is meaningful only for classical attacks) and show that $\mathcal{CSST}^n$ satisfies this property. Then we show that this version of the binding property implies the weak binding property (up to some loss in the parameter).

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

Definition 3.8 (Simultaneous opening). *A two-prover commitment scheme $\mathcal{S}$ is* ε-binding in the sense of simultaneous opening *(against classical attacks) if for all $\overline{\mathsf{res}}_{PQ}$ and $\overline{\mathsf{com}}_{PQ}$, all pairs of opening strategies $\overline{\mathsf{open}}_{PQ}$ and $\overline{\mathsf{open}}'_{PQ}$, and all pairs $s_\circ, s'_\circ$ of distinct strings, we have $p(s = s_\circ \wedge s' = s'_\circ) \leq \varepsilon$.*

Proposition 3.9. *$\mathcal{CSST}^n$ is 2^{-n}-binding in the sense of simultaneous opening against classical attacks.*

Proof. Fix a resource $\overline{\mathsf{res}}_{PQ}$, a commit strategy $\overline{\mathsf{com}}_P$ and two opening strategies $\overline{\mathsf{open}}_Q$ and $\overline{\mathsf{open}}'_Q$.[5] This then fixes the distribution $p(a, x, y, s, y', s')$. Note that y and y' are produced by $\overline{\mathsf{open}}_Q$ and $\overline{\mathsf{open}}'_Q$ by means of acting on the shared randomness alone. As such, the pair y, y' is independent of a. Furthermore, s and s' satisfy $x \oplus y = a \cdot s$ and $x \oplus y' = a \cdot s'$. Thus, for any pair $s_\circ, s'_\circ$ of distinct strings, it holds that

$$\begin{aligned} p(s = s_\circ \wedge s' = s'_\circ) &\leq p(x \oplus y = a \cdot s_\circ \ \wedge \ x \oplus y' = a \cdot s'_\circ) \\ &= p(a \cdot s_\circ \ominus y = x = a \cdot s'_\circ \ominus y') \\ &\leq p(a = (y' \ominus y) \cdot (s_\circ \ominus s'_\circ)^{-1}) \\ &= \frac{1}{2^n} \end{aligned}$$

which proves the claim. □

Theorem 3.10. *Every scheme $\mathcal{S}$ that is ε-binding in the sense of simultaneous opening (against classical attacks) is weakly ε'-binding (against classical attacks) with $\varepsilon' = \sqrt{2\varepsilon}$.*

Proof. Fix $\overline{\mathsf{res}}_{PQ}$ and $\overline{\mathsf{com}}_{PQ}$ against $\mathcal{S}$. Enumerate all strings in the domain $\{0,1\}^n$ of $\mathcal{S}$ as $s_1, \ldots, s_{2^n}$, and for every $i \in \{1, \ldots, 2^n\}$ let $\overline{\mathsf{open}}^i_{PQ}$ be an opening strategy that maximizes $p_i := p(s = s_i)$, where s is the output of the verifier when P and Q use this strategy. We assume without loss of generality that the p_i's are in descending order. We define $p(\hat{s})$ as follows. Let $N \geq 2$ be an integer which we will fix later. By Definition 3.8 and inequality (1), it holds that

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

[4] It is understood that the allowed attacks against $\mathcal{CSST}^n$ are those where the provers do not communicate.

[5] Note that Q is inactive during the commit, and P during the opening phase.

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

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

for some $\tilde{\varepsilon} \leq \varepsilon$. We now set $p(\hat{s})$ to be $p(\hat{s} = s_i) := p_i - (N-1)\tilde{\varepsilon}/2 \geq p_i - (N-1)\varepsilon/2 \geq 0$ for all $i \leq N'$. Now consider an opening strategy $\overline{\mathsf{open}}_{PQ}$ and let $p(s)$ be the resulting output distribution. By definition of the p_i, it follows that $p(s = s_i) \leq p_i$ for all $i \leq 2^n$, and $p_i \leq p(\hat{s} = s_i) + (N-1)\varepsilon/2$ for all $i \leq N'$. By Lemma 2.1, we can conclude that there exists a consistent joint distribution $p(\hat{s}, s)$ with $p(\hat{s} = s = s_i) = \min\{p(s = s_i), p(\hat{s} = s_i)\} \geq p(s = s_i) - (N-1)\varepsilon/2$ for all $i \leq N'$, and thus $p(\hat{s} \neq s = s_i) = p(s = s_i) - p(\hat{s} = s = s_i) \leq (N-1)\varepsilon/2$ for all $i \leq N'$ Furthermore, when $N' < i \leq N$, we have $p(\hat{s} \neq s = s_i) = p(s = s_i) \leq p_i < (N-1)\varepsilon/2$ by definition of N'. Since the p_i are sorted in descending order, it follows that for all $i > N$

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

and thus, we have shown for all $s_\circ \in \{0,1\}^n$ that

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

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

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

for any $s_\circ \in \{0,1\}^n$, as claimed. □

Combining Proposition 3.9 and Theorem 3.10, we obtain the following corollary.

Corollary 3.11. *$\mathcal{CSST}^n$ is weakly ε-binding against classical attacks with $\varepsilon = 2^{-\frac{n-1}{2}}$.*

4 Composing Commitment Schemes

4.1 The Composition Operation

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

Definition 4.1. *Let $\mathcal{S} = (\mathsf{res}_{PQ}, \mathsf{com}_{PQV}, \mathsf{open}_{PQV})$ and $\mathcal{S}' = (\mathsf{res}'_{PQ}, \mathsf{com}'_{PQV}, \mathsf{open}'_{PQV})$ be two 2-prover string commitment schemes. We call the pair $(\mathcal{S}, \mathcal{S}')$* eligible *if the following three properties hold, or they hold with the roles of P and Q exchanged.*

1. *The respective opening and commit phases are of the form*

   $$\mathsf{com}_{PQV} = (\mathsf{com}_P, id_Q, \mathsf{com}_V)\,,\ \mathsf{open}_{PQV} = (\emptyset_P, \mathsf{open}_Q, \mathsf{open}_V) \text{ and } \mathsf{com}'_{PQV} = (id_P, \mathsf{com}'_Q, \mathsf{com}'_V),$$

   where id is the trivial protocol that outputs its inputs (and no communication takes place), and $\emptyset$ is the trivial protocol that has no output (and does not communicate). In other words, prover Q is inactive in the commit and P is inactive in the opening phase of $\mathcal{S}$, and P is inactive in the commit phase of $\mathcal{S}'$ (but both provers may be active in the opening phase).

2. *The opening phase of $\mathcal{S}$ is of the following simple form: Q sends a bit string $y \in \{0,1\}^m$ to V, and V computes s deterministically as $s = \mathrm{Extr}(y, a, x)$, where a is V's randomness for com_V and x collects the messages that P sent to V during the commit phase.*[6]
3. *The domain of $\mathcal{S}'$ contains (or equals) $\{0,1\}^m$.*

An example of an eligible pair of 2-prover commitments is the pair $(\mathcal{CSST}^n, \mathcal{XCSST}^n)$, where $\mathcal{XCSST}^n$ coincides with the Cépeau *et al.* scheme $\mathcal{CSST}^n$ except that the roles of P and Q are exchanged. The reader can safely think of this concrete example.

Remark 4.2. Given that for an eligible pair $(\mathcal{S}, \mathcal{S}')$, com_Q and open_P are fixed to id_Q and $\emptyset_P$, respectively, it is good enough to specify the commit and open phases of $\mathcal{S}$ by means of $\mathsf{com}_{PV} = (\mathsf{com}_P, \mathsf{com}_V)$ and $\mathsf{open}_{QV} = (\mathsf{open}_Q, \mathsf{open}_V)$, respectively. Vice versa, whenever we specify the commit and open phases of a 2-prover commitment scheme by means of $\mathsf{com}_{PV} = (\mathsf{com}_P, \mathsf{com}_V)$ and $\mathsf{open}_{QV} = (\mathsf{open}_Q, \mathsf{open}_V)$, we take it as understood that com_Q and open_P are fixed to id_Q and $\emptyset_P$; similarly for $\mathcal{S}'$.

Furthermore, for an eligible pair $(\mathcal{S}, \mathcal{S}')$, we take it as understood that when considering attacks against the binding property of $\mathcal{S}$, the *allowed* commit and open strategies also satisfy $\overline{\mathsf{com}}_Q = id_Q$ and $\overline{\mathsf{open}}_P = \emptyset_P$ (and thus are specified by $\overline{\mathsf{com}}_P$ and $\overline{\mathsf{open}}_Q$), and thus in particular no communication is allowed between P and Q. Similarly, for $\mathcal{S}'$, we take it as understood that any *allowed* commit strategy satisfies $\overline{\mathsf{com}}_P = id_P$, and thus in particular there is no communication before the start of the opening phase.

We now define the composition operation (see also Figure 2).

Definition 4.3. *Let $\mathcal{S} = (\mathsf{res}_{PQ}, \mathsf{com}_{PV}, \mathsf{open}_{QV})$ and $\mathcal{S}' = (\mathsf{res}'_{PQ}, \mathsf{com}'_{QV}, \mathsf{open}'_{PQV})$ form an eligible pair of 2-prover commitment schemes. Then, their composition $\mathcal{S} \star \mathcal{S}'$ is defined as the 2-prover commitment scheme $(\mathsf{res}_{PQ}, \mathsf{com}_{PV}, \mathsf{open}''_{PQV})$ with $\mathsf{open}''_{PQV} := \mathsf{open}_V \circ \mathsf{open}'_{PQV} \circ \mathsf{com}'_{QV} \circ \mathsf{open}_Q$. This means that committing is done by means of committing using $\mathcal{S}$, and to open the commitment, Q uses open_Q to locally compute the opening information y and he commits to y with respect to the scheme $\mathcal{S}'$, and then this commitment is opened (to y), and V computes and outputs $s = \mathrm{Extr}(a, x, y)$.*

Furthermore, when considering attacks against the binding property of $\mathcal{S} \star \mathcal{S}'$, we declare that the allowed attacks are those of the form $(\overline{\mathsf{res}}_{PQ}, \overline{\mathsf{com}}_P, \overline{\mathsf{open}}'_{PQ} \circ \mathsf{ptoq}_{PQ} \circ \overline{\mathsf{com}}'_Q)$, where $\overline{\mathsf{res}}_{PQ}$ is an allowed resource and $\overline{\mathsf{com}}_P$ an allowed commit strategy for $\mathcal{S}$, $\overline{\mathsf{com}}'_Q$ and $\overline{\mathsf{open}}'_{PQ}$ are allowed commit and opening strategies for $\mathcal{S}'$, and ptoq_{PQ} is a one-way communication protocol that allows P to send a message to Q (see also Figure 3).[7]

Remark 4.4. It is immediate that $\mathcal{S} \star \mathcal{S}'$ is a commitment scheme in the sense of Definition 2.2, and that it is sound and hiding if $\mathcal{S}$ and $\mathcal{S}'$ are (in both cases, the error parameters add up). Also, it is intuitively clear that $\mathcal{S} \star \mathcal{S}'$ *should be* binding if $\mathcal{S}$ and $\mathcal{S}'$ are: committing to the opening information y and then opening the commitment allows the provers to *delay* the announcement of y (which is the whole point of the exercise), but it does not allow them to *change* y, by the binding property of $\mathcal{S}'$; thus, $\mathcal{S} \star \mathcal{S}'$ should be (almost) as binding as $\mathcal{S}$. This intuition is confirmed by our composition theorem below.

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

Remark 4.5. We observe that if $(\mathcal{S}, \mathcal{S}')$ is an eligible pair, then so is $(\mathcal{XS}, \mathcal{S} \star \mathcal{S}')$, where $\mathcal{XS}$ coincides with $\mathcal{S}$ except that the roles of P and Q are exchanged. In particular, if $(\mathcal{S}, \mathcal{XS})$ is an eligible pair, then so is $(\mathcal{XS}, \mathcal{S} \star \mathcal{XS})$. As such, we can then compose $\mathcal{XS}$ with $\mathcal{S} \star \mathcal{XS}$, and obtain yet another eligible pair $(\mathcal{S}, \mathcal{XS} \star \mathcal{S} \star \mathcal{XS})$, etc. Applying this to the schemes $\mathcal{S} = \mathcal{CSST}^n$, we obtain the multi-round scheme from Lunghi *et al.* [LKB+14]. As such, our composition theorem below implies security of their scheme — with a *linear* blow-up of the error term (instead of doubly exponential).

[6] This restriction on $\mathcal{S}$ is actually without loss of generality: we may always assume that the commitment *com* consists of V's randomness and the incoming messages, and we may always assume that in the opening phase the provers just announce the shared randomness.

[7] This one-way communication models that in the relativistic setting, sufficient time has passed at this point for P to inform Q about what happened during the execution of $\overline{\mathsf{com}}_P$.

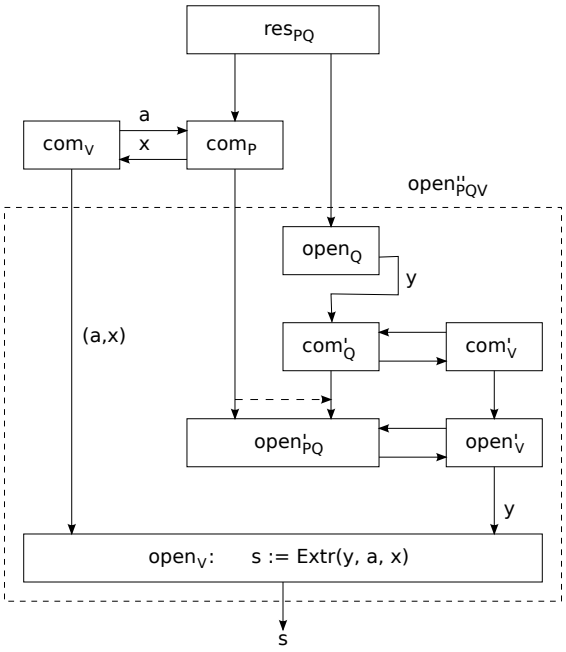

Fig. 2. The composition of $\mathcal{S}$ and $\mathcal{S}'$ (assuming single-round commit phases and that V's message equals its local randomness). The dotted arrow indicates communication allowed to the dishonest provers.

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

Definition 4.6. *Let $(\mathcal{S}, \mathcal{S}')$ be an eligible pair, which in particular means that V's action in the opening phase of $\mathcal{S}$ is determined by a function* Extr. *We define $k(\mathcal{S}) := \max_{a,x,s} |\{y \,|\, \mathrm{Extr}(y, a, x) = s\}|$.*

I.e., $k(\mathcal{S})$ counts the number of y's that are consistent with a given string s (in the worst case). Note that $k(\mathcal{CSST}^n) = 1$: for every $a, x, s \in \{0,1\}^n$ there is at most one $y \in \{0,1\}^n$ such that $x \oplus y = a \cdot s$.

4.2 The Composition Theorem

In the following composition theorem, we take it as understood that the assumed respective binding properties of $\mathcal{S}$ and $\mathcal{S}'$ hold with respect to a well-defined respective classes of allowed attacks. Furthermore, these allowed attacks need to be *classical* attacks; thus, our composition theorem only works for classical dishonest provers — extending it to quantum provers is left as an open problem.

Theorem 4.7. *Let $(\mathcal{S}, \mathcal{S}')$ be an eligible pair of 2-prover commitment schemes, and assume that $\mathcal{S}$ and $\mathcal{S}'$ are respectively weakly ε- and δ-binding against classical attacks. Then, their composition $\mathcal{S}'' = \mathcal{S} \star \mathcal{S}'$ is a weakly $(\varepsilon + k(\mathcal{S}) \cdot \delta)$-binding 2-prover commitment scheme against classical attacks.*

Proof. For simplicity, we assume that both schemes are such that V never outputs $\perp$ in the opening phase (see also Remark 3.3); the proof goes along the same lines in case of possible $\perp$-outputs.

We first consider the case $k(\mathcal{S}) = 1$. We fix an arbitrary attack $(\overline{\mathsf{res}}_{PQ}, \overline{\mathsf{com}}_P, \overline{\mathsf{open}}''_{PQ})$ against $\mathcal{S}''$, where $\overline{\mathsf{open}}''_{PQ}$ is of the form $\overline{\mathsf{open}}''_{PQ} = \overline{\mathsf{open}}'_{PQ} \circ \mathsf{ptoq}_{PQ} \circ \overline{\mathsf{com}}'_Q$, and we fix an arbitrary target string $s_\circ$. Without loss of generality, we may assume the attack to be *deterministic* (i.e., $\overline{\mathsf{res}}_{PQ}$ is trivial); as such, x is a function $x(a)$ of a. Such an attack fixes the distribution $p(a, y)$, and thus the distribution $p(s)$ for V's output $s = \mathrm{Extr}(y, a, x(a))$.

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

To show existence of such a joint distribution, we "decompose and reassemble" the attack strategy $(\overline{\mathsf{res}}_{PQ}, \overline{\mathsf{com}}_P, \overline{\mathsf{open}}'_{PQ} \circ \mathsf{ptoq}_{PQ} \circ \overline{\mathsf{com}}'_Q)$ for $\mathcal{S}''$ into an attack strategy $(\overline{\mathsf{res}}_{PQ}, \overline{\mathsf{com}}'_Q, \overline{\mathsf{newopen}}'_{PQ}(a))$ for $\mathcal{S}'$ with $\overline{\mathsf{newopen}}'_{PQ}(a) := \overline{\mathsf{open}}'_{PQ} \circ \mathsf{ptoq}_{PQ} \circ \overline{\mathsf{com}}_P(a)$, and where $\overline{\mathsf{com}}_P(a)$ locally runs $(\overline{\mathsf{com}}_P || \mathsf{com}_V)$ using the specific choice a for V's randomness (see also Figure 3).[8] Thus, we consider a fixed commit strategy and one opening strategy $\overline{\mathsf{newopen}}'_{PQ}(a)$ for every possible choice of a. Note that the resulting distribution of y is $p(y|a)$. Furthermore, we set $y_\circ$ to be the unique string such that $\mathrm{Extr}(y_\circ, a, x(a)) = s_\circ$; recall, we assume for the moment that $k(\mathcal{S}) = 1$.[9] It follows from the weak binding property of $\mathcal{S}'$ that there exists a distribution $p(\hat{y})$, only depending on $\overline{\mathsf{com}}'_Q$ so that for every choice of a there exists a consistent joint distribution $p(\hat{y}, y|a)$ so that $p(\hat{y} \neq y \wedge y = y_\circ|a) \leq \delta$. Note that here, consistency in particular means that $p(\hat{y}|a) = p(\hat{y})$. This joint conditional distribution $p(\hat{y}, y|a)$, together with the deterministic choice of $y_\circ$ when given a, and together with the distribution $p(a)$ of a, then naturally defines the distribution $p(a, \hat{y}, y, y_\circ)$, which is consistent with $p(a, y)$ considered above, and satisfies $p(\hat{y} \neq y \wedge y = y_\circ) \leq \delta$.

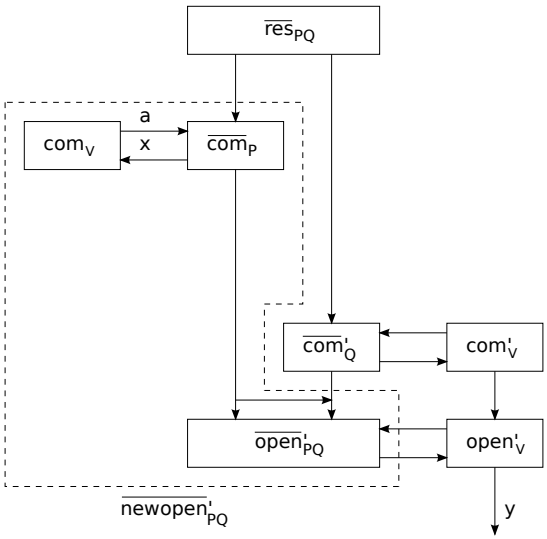

Fig. 3. Constructing the opening strategy $\overline{\mathsf{newopen}}'_{PQ}$ against $\mathcal{S}'$.

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

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

i.e., is consistent with the distribution $p(a, \hat{y}, y, y_\circ)$ above. By Definition 3.2, we know there exists a joint distribution $p(\hat{s}, \tilde{s})$, consistent with $p(\hat{s})$ fixed above and with $p(\tilde{s})$ determined by $\tilde{s} := \mathrm{Extr}(\hat{y}, a, x(a))$, and such that $p(\hat{s} \neq \tilde{s} \wedge \tilde{s} = s_\circ) \leq \varepsilon$. We can now "glue together" $p(\hat{s}, \tilde{s})$ and $p(a, \hat{y}, y, y_\circ, \tilde{s})$, i.e., find a joint distribution that is consistent with both, by setting

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

With respect to this distribution, it holds that

$$\begin{aligned}
p(\hat{s} \neq s \wedge s = s_\circ) \;&\leq\; p\big((\hat{s} \neq s \wedge s = s_\circ) \wedge (s = \tilde{s} \vee s \neq s_\circ)\big) + p\big(\neg(s = \tilde{s} \vee s \neq s_\circ)\big) \\
&=\; p(\hat{s} \neq s \wedge s = s_\circ \wedge s = \tilde{s}) + p(s \neq \tilde{s} \wedge s = s_\circ) \\
&=\; p(\hat{s} \neq \tilde{s} \wedge \tilde{s} = s_\circ \wedge s = \tilde{s}) + p\big(\mathrm{Extr}(y, a, x(a)) \neq \mathrm{Extr}(\hat{y}, a, x(a)) \;\wedge\; \mathrm{Extr}(y, a, x(a)) = s_\circ\big) \\
&\leq\; p(\hat{s} \neq \tilde{s} \wedge \tilde{s} = s_\circ) + p(y \neq \hat{y} \wedge y = y_\circ) \\
&\leq\; \varepsilon + \delta\,.
\end{aligned}$$

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

[9] Furthermore, we can safely ignore the case where there is no such $y_\circ$, because then s can anyway not hit $s_\circ$.

Thus, $p(a, y, \hat{s})$ gives rise to the required joint distribution $p(\hat{s}, s)$ (by setting $s = \mathrm{Extr}(y, a, x(a))$).

For the case where $k(\mathcal{S}) > 1$, we can reason similarly. The only difference is that we choose $y_\circ$ uniformly at random among those that satisfy $\mathrm{Extr}(y_\circ, a, x(a)) = s_\circ$. This then has the effect that in the bound on $p(\hat{s} \neq s \wedge s = s_\circ)$, the probability $p(y \neq \hat{y} \wedge y = y_\circ)$ is replaced by $k(\mathcal{S}) \cdot p(y \neq \hat{y} \wedge y = y_\circ)$, which then results in the claimed bound. □

References

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

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

[Ken99] Adrian Kent. Unconditionally secure bit commitment. *Physical Review Letters*, 83(7):1447–1450, 1999.

[Ken05] Adrian Kent. Secure classical bit commitment using fixed capacity communication channels. *Journal of Cryptology*, 18(4):313–335, 2005. Available from: `http://arxiv.org/abs/quant-ph/9906103`.

[LKB+14] Tomaso Lunghi, Jędrzej Kaniewski, Felix Bussières, Raphael Houlmann, Marco Tomamichel, Stephanie Wehner, and Hugo Zbinden. Practical relativistic bit commitment. *ArXiv e-prints*, 2014. `http://arxiv.org/abs/1411.4917`.

A Proof of Lemma 2.1

We start by introducing an event Δ in the respective probability spaces given by the distributions $p(x)$ and $p(y)$ by means of declaring that

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

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

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

for all $x_\circ \in \mathcal{X}$, and $p(x = x_\circ \wedge y = y_\circ|\Delta) := 0$ for all $x_\circ \neq y_\circ \in \mathcal{X}$ (this defines the "diagonal" of $p(x, y)$); and the latter as

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

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