# Non-Malleability Amplification

Huijia Lin[*] Rafael Pass[†]

### Abstract

We show a technique for amplifying commitment schemes that are non-malleable with respect to identities of length $t$, into ones that are non-malleable with respect to identities of length $\Omega(2^t)$, while only incurring a constant overhead in round-complexity. As a result we obtain a construction of $O(1)^{log^* n}$-round (i.e., "essentially" constant-round) non-malleable commitments from any one-way function, and using a black-box proof of security.

[*] Cornell University, E-Mail: `huijia@cs.cornell.edu`.

[†] Cornell University, E-Mail: `rafael@cs.cornell.edu`.

# 1 Introduction

Commitment schemes are one of the most fundamental cryptographic building blocks. Often described as the "digital" analogue of sealed envelopes, commitment schemes enable a *sender* to commit itself to a value while keeping it secret from the *receiver*. This property is called *hiding*. Furthermore, the commitment is *binding*, and thus in a later stage when the commitment is opened, it is guaranteed that the "opening" can yield only a single value determined in the committing stage. Their applications range from coin flipping [Blu83] to the secure computation of any efficiently computable function [GMW91, GMW87]. In light of their importance, commitment schemes have received a considerable amount of attention. This has resulted in a fairly comprehensive understanding of the hardness assumptions under which they can be realized [HILL99, Nao91, DPP94].

For many applications, however, the most basic security guarantees of commitments are not sufficient. For instance, the basic definition of commitments does not rule out an attack where an adversary, upon seeing a commitment to a specific value $v$, is able to commit to a related value (say, $v - 1$), even though it does not know the actual value of $v$. This kind of attack might have devastating consequences if the underlying application relies on the *independence* of committed values (e.g., consider a case in which the commitment scheme is used for securely implementing a contract bidding mechanism). The state of affairs is even worsened by the fact that many of the known commitment schemes are actually susceptible to this kind of attack. In order to address the above concerns, Dolev, Dwork and Naor (DDN) introduced the concept of *non-malleable commitments* [DDN00]. Loosely speaking, a commitment scheme is said to be non-malleable if it is infeasible for an adversary to "maul" a commitment to a value $v$ into a commitment to a related value $\tilde{v}$.

The first non-malleable commitment protocol was constructed by Dolev, Dwork and Naor [DDN00] in the early 90's. The security of their protocol relies on the minimal assumption of one-way functions and requires $O(\log n)$ rounds of interaction, where $n \in N$ is the length of party identities (or alternatively, a security parameter). More recent results improved the round complexity by making stronger assumptions: Barak [Bar02] presents a constant-round protocol for non-malleable commitments whose security relies on the existence of trapdoor permutations and hash functions that are collision-resistant against circuits of sub-exponential size. Pass and Rosen [PR05b] subsequently showed that collision resistant hash functions secure against polynomially sized circuits are sufficient to obtain a constant-round protocol.

But, despite two decades of research, there have been no improvements over the original DDN construction, when only assuming the existence of one-way functions, leaving open the following question.

> *Does there exist a sub-logarithmic non-malleable commitment scheme, assuming only one-way functions?*

Additionally, whereas the original DDN-construction only relies on "elementary" techniques and has a black-box proof of security, the more recent works of [Bar02, PR05b] rely on *non-black box techniques* [Bar01] and inherit the "heavy" machinery (e.g., the PCP theorem) associated with them. Yet more recent work obtains round-efficient constructions using black-box proofs of security by instead relying on non-standard hardness assumptions: Pandey, Pass and Vaikuntanathan [PPV08] provide a construction of constant-round non-malleable commitments based on a new hardness assumption with a strong non-malleability flavor[1]. Thus, given the current state-of-the-art, it is unknown whether the round complexity of the DDN construction can be improved under

[1] More precisely, they assume the existence of, so called, adaptive one-way permutations—namely permutations which remain one-way even when the adversary has access to an inversion oracle.

*any* "standard" hardness assumptions, if restricting our attention to constructions with black-box proofs of security.

## 1.1 Our Results

In this work we provide an affirmative answer to the above questions. We show the existence of an $O(1)^{\log^* n}$-round—i.e., "almost" constant-round—non-malleable commitment from only one-way functions and using a black-box proof of security.

**Theorem 1.** *Assume the existence of one-way functions. Then there exists a $O(1)^{\log^* n}$-round non-malleable commitment with a black-box proof of security.*

Subsequent to the original publications of our work [LP09], Hoeteck Wee [Wee10] (following our high-level approach, but using a different construction; see Section 1.4 for more details) improved the round complexity to $O(\log^* n)$. The round-complexity of our protocol can also be improved to $O(\log^* n)$ (by running some of our sub-protocols in parallel). We thus have the following theorem (first established by Wee).

**Theorem 2.** *Assume the existence of one-way functions. Then there exists a $O(\log^* n)$-round non-malleable commitment with a black-box proof of security.*

## 1.2 A New Technique: Non-malleability Amplification

To establish our main result we develop a new approach for obtaining number of possible identities players can take (e.g., the DDN construction requires $O(\log n)$ rounds if the identity length is $n$ bits). Thus, if we consider only a small set of possible identities—say 8—then the DDN construction results in a constant-round protocol. The problem, however, is that once the number of identities grows the protocol also becomes more complicated (and requires more rounds).

We here consider the question of whether a $t$-non malleable commitment—i.e., a commitment scheme that is non-malleable with respect to identities of length $t$—can be amplified into a $t'$-non malleable commitment, where $t' > t$; we call this task *non-malleability amplification.* Although *hardness* amplification has been extensively studied in both cryptography and complexity theory, we are not aware of any prior works that consider the question of amplifying non-malleability.

Our main result shows a general technique for amplifying "robust" $t$-non malleable commitments into "robust" $2^{t-1}$-non malleable commitments, while only incurring a constant additive overhead in round-complexity.[2] We describe the notion of a robust non-malleable commitment shortly; formal definitions can be found in Section 2.8 and 2.10.

**Theorem 3. (Non-malleability Amplification)** *Let $\langle C, R\rangle$ be a $k(n)$-round robust $t(n)$-non malleable commitment scheme with computational complexity $p(n)$. Then, there exists a robust $(k(n)+8)$-round $2^{t(n)-1}$-non malleable commitment scheme $\langle C', R'\rangle$ with computational complexity $2^{t(n)}p(n) + k(n)\text{poly}(n)$.*

By repeatedly applying the non-malleability amplification theorem, we get the following generalized version.

**Corollary 1. (Generalized Non-malleability Amplification)** *Let $\langle C, R\rangle$ be a $k(n)$-round robust $t(n)$-non malleable commitment scheme, s.t. $3 \le t(n) \le n$. Then, there exists a $(O(l(n)) + k(n))$-round robust non-malleable commitment scheme $\langle \hat{C}, \hat{R}\rangle$, where $l(n) = O(\log^* n - \log^* t(n))$.*

[2] In the conference version of this work [LP09], the amplification procedure incurred a constant *multiplicative* overhead; this is why the final construction required $O(1)^{\log^* n}$ round instead of $O(\log^* n)$.

Finally, by applying our amplification procedure $O(\log^* n)$ times to any constant-round robust 3-non-malleable commitment based on one-way functions, we get a $O(\log^* n)$-round non-malleable commitment from any one-way function.[3] Theorem 2 is concluded by observing that e.g., the non-malleable commitment protocols of [DDN00, LPV08] are robust. Given that we only need a constant-round protocol that is non-malleable with respect a constant number of identities, it actually suffices to start off with an exponential-round protocol (i.e., a protocol where the number of rounds is exponential in the length of the identity); as we show in Appendix A, such a protocol is significantly easier to construct.

**Our Amplification Technique.** In analogy with recent amplification techniques by Reingold [Rei05] and Dinur [Din07] (see also the survey by Goldreich [Gol05]), our amplification consists of two separate operations; each one performs "amplification" in a different "axis." For instance, Reingold [Rei05] considers as axes the connectivity, and the degree, of a graph: the first operation improves the connectivity (by using a direct product), but degrades the degree. The second operation improves (i.e., lowers) the degree of the graph, but degrades the connectivity. Combined together, however, these operations improve the connectivity without loosing in degree.[4] In our case the two axes are:

1. the ratio between the number of identities and the number of rounds, and
2. the "quality" of non-malleability.

The first step shows how to (essentially) use a direct product to amplify non-malleability, while keeping the round complexity constant, thus improving the ratio. This transformation, however, degrades the "quality" of non-malleability. The second transformation shows how to restore the original quality by only incurring a constant overhead in round complexity. Combined they yield the desired amplification. More precisely, we proceeds as follows.

- **Step 1: Improving identity/round ratio.** Dolev, Dwork and Naor [DDN00] presented a method for collapsing rounds in their protocol; the idea is to appropriately run $n$ parallel repetitions of a $O(\log n)$-non malleable protocol. As recently demonstrated by Lin, Pass and Venkitasubramaniam [LPV08], this approach—sometimes referred to as the "$\log n$ trick"—can, in fact, be applied to any *concurrently non-malleable* commitment (using the definition of [LPV08]); roughly speaking, a commitment scheme is concurrently non-malleable if non-malleability holds even if the adversary is participating in an unbounded number of concurrent executions. A first idea would be to simply iteratively apply this technique. The problem is that the "$\log n$-trick" requires the initial commitment scheme—the one that only needs to work for "small" identities—to be concurrently secure, but the resulting commitment scheme—the one that works for larger identities—will no longer be concurrently secure; in fact, it is easily seen that it can be broken under a concurrent attack.

- **Step 2: Restoring the "quality": from stand-alone to concurrent security.** To be able to iteratively apply the "$\log n$-trick" we present a method for compiling "robust" stand-alone secure non-malleable commitment schemes into "robust" concurrently non-malleable commitment schemes. Our main technical contribution lies in this step. The compilation technique only requires the existence of one-way functions (which anyway is implied by the existence of commitment schemes) and only increases the round complexity by a constant. A

[3] Note that if only start with a 2-non-malleable commitment our amplification theorem only gives us a commitment that is non-malleable with respect to $2^{2-1} = 2$ bit identities; thus, no amplification.

[4] Dinur [Dam00] instead considers as axes the ratio of unsatisfied clauses in a SAT formula, and the alphabet size.

first idea enabling this transformation is the fact that any (stand-alone) secure non-malleable commitment (satisfying the definition of [LPV08]) already satisfies a "weak" notion of *concurrent* non-malleability. Using this insight, we next show how to turn a weak concurrent non-malleable commitment into a "traditional" concurrent non-malleable commitment. The key conceptual insight enabling this step is to consider a notion of non-malleability with respect to *arbitrary* $k$-round protocols—that is, instead of only considering non-malleability in a setting where both the left and the right interactions consist of executions of the same protocol, we consider a scenario where the left interaction might be any arbitrary $k$-round protocol. In fact, in our construction, it suffices to consider non-malleability with respect to arbitrary 4-round protocols; we call such protocols *robust* commitment scheme. Relying on the above two notions, we finally show a transformation from any *robust* stand-alone non-malleable commitment scheme, into one that is both *robust* and concurrently non-malleable.

The complete amplification procedure is obtained by combining the above two steps, and the fact that step 1 also preserves robustness of the non-malleable commitment scheme. More precisely, the amplification theorem follows from the following two lemmas.

**Lemma 1 (Improving identity/round).** *Let $\langle C, R\rangle$ be a $k(n)$-round robust concurrent $t(n)$-non malleable commitment scheme with computational complexity $p(n)$. Then, there exists a $k(n)$-round robust $2^{t(n)-1}$-non malleable commitment scheme $\langle \tilde{C}, \tilde{R}\rangle$ with computational complexity $2^{t(n)-1}p(n) + \mathrm{poly}(n)$.*

**Lemma 2 (Improving concurrent security).** *Let $\langle \tilde{C}, \tilde{R}\rangle$ be a $k(n)$-round robust $t(n)$-non malleable commitment scheme with computational complexity $p(n)$. Then, there exists a $(k(n) + 8)$-round robust concurrent $t(n)$-non malleable commitment scheme $\langle \hat{C}, \hat{R}\rangle$ with computational complexity $2p(n) + k(n)\mathrm{poly}(n)$.*

We mention that Lemma 2 is interesting in its own right: it shows that it is sufficient to study stand-alone non-malleable commitments to conclude the existence of concurrent non-malleable commitments. Furthermore, as we discuss further in Section 1.4, we believe that the notion of *robust non-malleability* (i.e., non-malleability with respect to arbitrary $k$-round protocols) is a fundamental concept that might have other applications.

## 1.3 Applications to Round-efficient Secure Computation

Goldreich, Micali and Wigderson's [GMW91] original work on secure multi-party computation showed a $O(n)$-round multi-party computation protocol based on the existence of enhanced trapdoor permutations (TDPs), where $n$ is the number of players in the execution. Subsequent works improved the round-complexity by making stronger assumptions. Katz, Ostrovsky, and Smith [KOS03] obtained a $O(\log n)$-round protocol assuming TDPs and dense-crypto systems. By additionally assuming the existence of hash-function collision-resistant against circuits of subexponential size (and non-black-box technique), they also obtained a $O(1)$-round protocol. The latter results was subsequently improved to Pass [Pas04], showing the existence of a $O(1)$-rounds protocol assuming only TDPs and (standard) collision resistant hash functions (but still using non-black box techniques). But so far, no (asymptotic) improvements to the round-complexity of multi-party computation have been established assuming only TDPs.

Lin, Pass and Venkitasubramaniam [LPV09] recently show that the round complexity of non-malleable commitments is intimately connected with the round complexity of protocols for secure multi-party computation. More precisely, [LPV09] shows that the existence of $k(n)$-round robust non-malleable commitments and the existence of TDPs implies the existence of $O(k(n))$–round secure multi-party computation. (In fact, as shown in [LPV09], a similar result also applies to *universally composable* [Can01] secure computation in a number of set-up models or relaxed models of security.) By combining their result with our new construction of non-malleable commitments, we get that $O(\log^* n)$ rounds suffice to construct a secure multi-party computation protocol, based on only TDPs.

**Theorem 4.** *Assume the existence of enhanced trapdoor permutations. Then there exists a $O(\log^* n)$-round protocol for secure multi-party computation.*

## 1.4 Subsequent Work and Perspective

Several recent works extend our technique.

**Towards Constant-round Non-malleable Commitment from One-way function.** Pass and Wee [PW10] recently provide a construction of a *constant-round* non-malleable commitment scheme assuming sub-exponential one-way functions. Their protocol is obtained by first constructing a constant-round protocol that is $\log\log\log n$-non malleable, and next applying our amplification technique to obtain a full-fledged non-malleable commitment. The question of obtaining constant-round commitments based on standard one-way functions is still open; by our work, it suffices to construct a constant-round protocol that is non-malleable for $\log\log\log\ldots\log n$ length identities.

**Black-box Non-malleability Amplification.** The elegant recent work by Wee [Wee10] presents a simplified non-malleability amplification technique. As mentioned, [Wee10] was first to achieve a $O(\log^* n)$-round (as opposed to a $O(1)^{log^* n}$-round) protocol. Instead of amplifying robust non-malleable commitments, Wee's amplification procedure operates directly on robust *concurrent* non-malleable commitments. Doing this simplifies the analysis of the amplification procedure. (On the downside, we now need to bootstrap the amplification with a concurrently secure protocol; known such protocols requires a quite complex analysis (see [LPV08]) even if we only need a protocol that is $O(1)$-non malleable. In contrast, as we show in Appendix A, constructing a robust *stand-alone* secure $O(1)$-non-malleable protocols is significantly simpler.) Even more interestingly, Wee shows that his amplification procedure can be made *black box*—i.e., the new commitment scheme (that works for longer identities) only invokes the original one as a black box. Unfortunately, the notion of non-malleability achieved by this black-box construction is somewhat weaker than the traditional definitions [DDN00, PR05b, LPV08]; however, as shown by Wee, this weaker notion of non-malleability is actually sufficient for the purpose of secure multiparty computation. As a consequence, Wee concludes a black-box version of Theorem 4. The problem of obtaining black-box non-malleability amplification, while preserving the traditional notion of non-malleability, seems like an interesting open question (that could lead to practical constructions of non-malleable commitments).

**Applications of Robust Non-malleable Commitment.** We believe that the notion of *robust non-malleability* (i.e., non-malleability with respect to arbitrary $k$-round protocols) is the key conceptual notion enabling our amplification procedure. A-posteriori, it seems quite natural to consider a notion of non-malleability with respect to classes of protocols; it also seems natural that such a notion is helpful when using non-malleable commitments as sub-protocols. We have recently exploited this notion in other works where non-malleable commitments are used as sub-protocols [LPV09, LPTV10].

### 1.5 Outline

In Section 2 we provide some preliminaries and definitions of the new notions of non-malleability we introduce. In Section 3 we prove Lemma 1, which, as mentioned, essentially follows using [DDN00, LPV08]. Section 4, which contains the proof of Lemma 2, is the main technical part of the paper. In Section 5, we show how to apply our non-malleability amplification theorem to establish the generalized version of the non-malleability amplification theorem and conclude Theorem 2. Finally, for completeness, in Appendix A we provide a simple construction of a robust, but exponential-round, non-malleable commitment.

## 2 Preliminaries and Definitions

### 2.1 Notations

Let $N$ denote the set of all positive integers. For any integer $n \in N$, let $[n]$ denote the set $\{0, 2, \ldots, n-1\}$, and let $\{0,1\}^n$ denote the set of $n$-bit strings. We denote by $\mathcal{PPT}$ probabilistic polynomial time Turing machines. We assume familiarity with interactive Turing machines, denoted ITM. Given a pair of ITMs, $A$ and $B$, we denote by $\langle A(x), B(y)\rangle(z)$ the random variable representing the (local) output of $B$, on common input $z$ and private input $y$, when interacting with $A$ with private input $x$, when the random tape of each machine is uniformly and independently chosen.

### 2.2 Witness Relations

We recall the definition of a witness relation for a $\mathcal{NP}$ language [Gol01].

**Definition 1** (Witness relation). *A* witness relation *for a language $L \in \mathcal{NP}$ is a binary relation $R_L$ that is polynomially bounded, polynomial time recognizable and characterizes $L$ by $L = \{x : \exists y\, s.t.\, (x,y) \in R_L\}$*

We say that $y$ is a witness for the membership $x \in L$ if $(x,y) \in R_L$. We will also let $R_L(x)$ denote the set of witnesses for the membership $x \in L$, i.e., $R_L(x) = \{y : (x,y) \in L\}$. In the following, we assume a fixed witness relation $R_L$ for each language $L \in \mathcal{NP}$.

### 2.3 Indistinguishability

**Definition 2** (Computational Indistinguishability). *Let $Y$ be a countable set. Two ensembles $\{A_{n,y}\}_{n\in N, y\in Y}$ and $\{B_{n,y}\}_{n\in N, y\in Y}$ are said to be* computationally indistinguishable *(denoted by $\{A_{n,y}\}_{n\in N, y\in Y} \approx \{B_{n,y}\}_{n\in N, y\in Y}$), if for every $\mathcal{PPT}$ "distinguishing" machine $D$, there exists a negligible function $\nu(\cdot)$ so that for every $n \in N, y \in Y$:*

$$\left|\Pr\left[a \leftarrow A_{n,y} \;:\; D(1^n, y, a) = 1\right] - \Pr\left[b \leftarrow B_{n,y} \;:\; D(1^n, y, b) = 1\right]\right| < \nu(n)$$

### 2.4 Interactive Proofs

We use the standard definitions of interactive proofs (and interactive Turing machines) [GMR89] and arguments (a.k.a computationally-sound proofs) [BCC88]. Given a pair of interactive Turing machines, $P$ and $V$, we denote by $\langle P(w), V\rangle(x)$ the random variable representing the (local) output of $V$, on common input $x$, when interacting with machine $P$ with private input $w$, when the random input to each machine is uniformly and independently chosen.

**Definition 3** (Interactive Proof System)**.** *A pair of interactive machines $\langle P, V\rangle$ is called an* interactive proof system *for a language $L$ if there is a negligible function $\nu(\cdot)$ such that the following two conditions hold :*

- Completeness: *For every $x \in L$, and every $w \in R_L(x)$,* $\Pr[\langle P(w), V\rangle(x) = 1] = 1$
- Soundness: *For every $x \notin L$, and every interactive machine $B$,* $\Pr[\langle B, V\rangle(x) = 1]$ $\leq \nu(|x|)$

*In case that the soundness condition is required to hold only with respect to a computationally bounded prover, the pair $\langle P, V\rangle$ is called an interactive* argument *system.*

## 2.5 Witness Indistinguishable Proofs

The notion of *witness indistinguishability* ($\mathcal{WI}$) was introduced by Feige and Shamir in [FS90]. Roughly speaking, an interactive proof is said to be $\mathcal{WI}$ if the verifier's output is "computationally independent" of the witness used by the prover for proving the statement. In this context, we focus on languages $L \in \mathcal{NP}$ with a corresponding witness relation $R_L$. Namely, we consider interactions in which, on common input $x$, the prover is given a witness in $R_L(x)$. By saying that the output is computationally independent of the witness, we mean that for any two possible $\mathcal{NP}$-witnesses that could be used by the prover to prove the statement $x \in L$, the corresponding outputs are computationally indistinguishable.

**Definition 4** (Witness-indistinguishability)**.** *Let $\langle P, V\rangle$ be an interactive proof system for a language $L \in \mathcal{NP}$. We say that $\langle P, V\rangle$ is* witness-indistinguishable *for $R_L$, if for every $\mathcal{PPT}$ ITM $V^*$ and for every two sequences $\{w^1_{n,x}\}_{n\in N, x\in L\cap\{0,1\}^n}$ and $\{w^2_{n,x}\}_{n\in N, x\in L\cap\{0,1\}^n}$, such that $w^1_{n,x}, w^2_{n,x} \in R_L(x)$ for every $x$, the following probability ensembles are computationally indistinguishable.*

- $\{\langle P(w^1_{n,x}), V^*(z)\rangle(x)\}_{n\in N, x\in L\cap\{0,1\}^n, z\in\{0,1\}^*}$
- $\{\langle P(w^2_{n,x}), V^*(z)\rangle(x)\}_{n\in N, x\in L\cap\{0,1\}^n, z\in\{0,1\}^*}$

## 2.6 Special-sound $\mathcal{WI}$ proofs

A 4-round public-coin interactive proof for the language $L \in \mathcal{NP}$ with witness relation $R_L$ is special-sound with respect to $R_L$, if for any two transcripts $(\delta, \alpha, \beta, \gamma)$ and $(\delta', \alpha', \beta', \gamma')$ such that the initial two messages, $\delta, \delta'$ and $\alpha, \alpha'$, are the same but the challenges $\beta, \beta'$ are different, there is a deterministic procedure to extract the witness from the two transcripts and runs in polynomial time. Special-sound $\mathcal{WI}$ proofs for languages in $\mathcal{NP}$ can be based on the existence of 2-round commitment schemes, which in turn can be based on one-way functions [GMW91, FS90, HILL99, Nao91].

## 2.7 Commitment Schemes

Commitment schemes are used to enable a party, known as the *sender*, to commit itself to a value while keeping it secret from the *receiver* (this property is called hiding). Furthermore, the commitment is binding, and thus in a later stage when the commitment is opened, it is guaranteed that the "opening" can yield only a single value determined in the committing phase. In this work, we consider commitment schemes that are statistically-binding, namely while the hiding property only holds against computationally bounded (non-uniform) adversaries, the binding property is required to hold against unbounded adversaries. We refer the reader to [Gol01] for a formal definition.

Two-round (i.e., a single message from the receiver followed by a single message from the committer) commitment schemes are known to exist based on the minimal assumption of one-way functions [Nao91, HILL99].

**Initial-binding.** In this work we consider a specific type of statistically binding commitment schemes, where the first message sent by the committer already determines the value committed to. We call this property *initial binding.* In the sequel of the paper, a commitment scheme always refers to a statistically-binding commitment with initial binding.

**Tag-based Commitment Scheme.** Following [PR05a, DDN00], we consider *tag-based commitment schemes* where, in addition to the security parameter, the committer and the receiver also receive a "tag"—a.k.a. the identity—$\mathsf{id}$ as common input.

## 2.8 Concurrent Non-Malleable Commitments

We recall the definition of concurrent non-malleability from [LPV08]. For convenience, we use a slightly different presentation (based on indistinguishability instead of simulation); equivalence follows using a standard argument (c.f. [GM84, PR05a]). Let $\langle C, R\rangle$ be a tag-based commitment scheme, and let $n \in N$ be a security parameter. Consider a man-in-the-middle adversary $A$ (as shown in figure 1) that, on inputs $n$ and $z$ (where $z$ is received as an auxiliary input), participates in $m$ left and right interactions simultaneously. In the left interactions the man-in-the-middle adversary $A$ interacts with $C$, receiving commitments to values $v_1, \ldots, v_m$, using identities $\mathsf{id}_1, \ldots, \mathsf{id}_m$ of its choice. In the right interactions $A$ interacts with $R$ attempting to commit to a sequence of related values $\tilde{v}_1, \ldots, \tilde{v}_m$, again using identities $\tilde{\mathsf{id}}_1, \ldots, \tilde{\mathsf{id}}_m$ of its choice. If any of the right commitments are invalid, or undefined, its value is set to $\perp$. For any $i$ such that $\tilde{\mathsf{id}}_i = \mathsf{id}_j$ for some $j$, set $\tilde{v}_i = \perp$—i.e., any commitment where the adversary uses the same identity as one of the left interactions is considered invalid. Let $\mathsf{mim}^A_{\langle C,R\rangle}(v_1, \ldots, v_m, z)$ denote a random variable that describes the values $\tilde{v}_1, \ldots, \tilde{v}_m$ and the view of $A$, in the above experiment.

**Definition 5.** *A commitment scheme $\langle C, R\rangle$ is said to be* concurrent non-malleable (with respect to itself) *if for every polynomial $p(\cdot)$, and every $\mathcal{PPT}$ man-in-the-middle adversary $A$ that participates in at most $m = p(n)$ concurrent executions, the following ensembles are computationally indistinguishable.*

$$\left\{\mathsf{mim}^A_{\langle C,R\rangle}(v_1, \ldots, v_m, z)\right\}_{n\in N, v_1,\ldots,v_m\in\{0,1\}^n, v'_1,\ldots,v'_m\in\{0,1\}^n, z\in\{0,1\}^*}$$
$$\left\{\mathsf{mim}^A_{\langle C,R\rangle}(v'_1, \ldots, v'_m, z)\right\}_{n\in N, v_1,\ldots,v_m\in\{0,1\}^n, v'_1,\ldots,v'_m\in\{0,1\}^n, z\in\{0,1\}^*}$$

Further, a commitment scheme is said to be concurrent $t(n)$-non malleable, if it is secure against all $\mathcal{PPT}$ man-in-the-middle adversaries that only select identities of length $t(n)$. Below for convenience, when referring to concurrent non-malleable commitments, we mean concurrent $n$-non-malleable commitments.

We also consider relaxed notions of concurrent non-malleability: one-one, one-many, and many-one secure non-malleable commitments (See Figure 2 below.) In a one-one (a.k.a., a stand-alone secure) non-malleable commitment, we consider only adversaries $A$ that participate in one left and one right interaction; in one-many, $A$ participates in one left and many right, and in many-one, $A$ participates in many left and one right.

As shown in [LPV08], any protocol that is one-many non-malleable is also concurrent non-malleable.

$$\begin{tikzcd}
C & & R \\
Com(v_1) \arrow[ddr] & & Com(\tilde{v}_1) \\
\cdots & & \cdots \\
Com(v_i) \arrow[r] & A \arrow[uur] \arrow[r] \arrow[ddr] & Com(\tilde{v}_i) \\
\cdots & & \cdots \\
Com(v_m) \arrow[uur] & & Com(\tilde{v}_m)
\end{tikzcd}$$

Figure 1: A concurrent man-in-the-middle adversary.

$$\begin{tikzcd}
Com(v) \arrow[r] & A \arrow[r] & Com(\tilde{v})
\end{tikzcd}$$

(i) one-one

$$\begin{tikzcd}
 & & Com(\tilde{v}_1) \\
 & & \cdots \\
Com(v) \arrow[r] & A \arrow[uur] \arrow[r] \arrow[ddr] & Com(\tilde{v}_i) \\
 & & \cdots \\
 & & Com(\tilde{v}_m)
\end{tikzcd}$$

(ii) one-many

$$\begin{tikzcd}
Com(v_1) \arrow[ddr] & & \\
\cdots & & \\
Com(v_i) \arrow[r] & A \arrow[r] & Com(\tilde{v}) \\
\cdots & & \\
Com(v_m) \arrow[uur] & &
\end{tikzcd}$$

(iii) many-one

Figure 2: Restricted man-in-the-middle adversaries.

**Proposition 1** ([LPV08]). *Let $\langle C, R \rangle$ be a one-many concurrent non-malleable commitment. Then, $\langle C, R \rangle$ is also a concurrent non-malleable commitment.*

We remark that our notion of non-malleability (from [LPV08]) is stronger than that in [DDN00, PR05b, PR05a] in that we consider not only the values committed to by the adversary, but also the view of the adversary. This property is used when establishing proposition 1, and will be used extensively in the remainder of the paper.

### 2.9 Weak Concurrent Non-Malleability

The standard concurrent non-malleability requires that no matter what values an adversary receives commitments to, the view of the adversary, combined with the values it commits to (on the right) are indistinguishable. This means that for any property $P$, the probability that $P$ holds on the view $\mathcal{V}$ and committed values $\tilde{v}_1, \ldots, \tilde{v}_m$ changes by at most a negligible amount. We here consider a relaxed notion of non-malleability, which focuses on a restricted class of properties $P$; furthermore we only require that the probability that $P$ holds does not "jump" from negligible to non-negligible. More precisely, we will be interested only in properties $\mathsf{exist}_Q$, which hold if the adversary manages to generate a view $\mathcal{V}$ and commit to some value $v$, such that $Q(\mathcal{V}, v)$ holds; formally, given a predicate $Q$, define $\mathsf{exists}_Q(\mathcal{V}, \tilde{v}_1, \ldots, \tilde{v}_m) = 1$ iff $\exists 1 \leq i \leq m$, s.t. $Q(\mathcal{V}, \tilde{v}_i) = 1$.

**Definition 6.** *A commitment scheme $\langle C, R \rangle$ is said to be* weak one-many non-malleable (with respect to itself) *if for every polynomial $p(\cdot)$, every $\mathcal{PPT}$ man-in-the-middle adversary $A$ that participates in one left interaction and at most $m = p(n)$ right interactions, and every binary predicate $Q$, there exists a negligible function $\varepsilon(\cdot)$, such that, for every $n \in N, v_1, v_2 \in \{0,1\}^n, z \in \{0,1\}^*$ it holds that:*

$$\Pr\left[\mathsf{exists}_Q(\mathsf{mim}^A_{\langle C,R \rangle}(v_1, z)) = 1\right] \leq m \times \Pr\left[\mathsf{exists}_Q(\mathsf{mim}^A_{\langle C,R \rangle}(v_2, z)) = 1\right] + \varepsilon(n)$$

Whereas one-one non-malleability does not imply one-many non-malleability, it does imply weak one-many non-malleability.

**Proposition 2.** *Let $\langle C, R\rangle$ be a stand-alone non-malleable commitment. Then, $\langle C, R\rangle$ is also a weak one-many non-malleable commitment.*

*Proof.* Assume, for contradiction, that $\langle C, R\rangle$ is not weak one-many non-malleable; that is, there exists an adversary $A$, a predicate $Q$, and a polynomial $p(\cdot)$ such that for infinitely many $n \in N$, there exists $v_1, v_2 \in \{0,1\}^n$, and $z \in \{0,1\}^*$ such that,

$$\Pr\left[\mathsf{exists}_Q(\mathsf{mim}^A_{\langle C,R\rangle}(v_1, z)) = 1\right] \geq m \times \Pr\left[\mathsf{exists}_Q(\mathsf{mim}^A_{\langle C,R\rangle}(v_2, z)) = 1\right] + \frac{1}{p(n)}$$

We construct a machine $B$ that violates the one-one non-malleability property of $\langle C, R\rangle$. $B(z)$ acts as a man-in-the-middle adversary, participating in one left and one right interaction of $\langle C, R\rangle$. Internally, it simulates a one-many man-in-the-middle execution of $\langle C, R\rangle$ with $A(z)$: on the left, it simply forwards messages from the external committer to $A$, while on the right, it picks one of the right interactions at random and forwards it to the outside receiver; the rest of the right interactions are honestly emulated. Since $B$ perfectly simulates the one-many man-in-the-middle execution for $A$, the probability that $\mathsf{exists}_Q$ holds in simulation by $B$ is

- $p_1 = \Pr\left[\mathsf{exists}_Q(\mathsf{mim}^A_{\langle C,R\rangle}(v_1, z_n)) = 1\right]$, when the external committer commits to $v_1$, and
- $p_2 = \Pr\left[\mathsf{exists}_Q(\mathsf{mim}^A_{\langle C,R\rangle}(v_2, z_n)) = 1\right]$, when the external committer commits to $v_2$.

Since the value $B$ commits to is simply the value $A$ commits to in a randomly picked right interaction, the probability that $Q$ holds on the value $B$ commits to is

- at least $\frac{p_1}{m}$ when receiving a commitment to $v_1$, and
- at most $p_2$, on receiving a commitment to $v_2$.

Thus, $Q$ distinguishes the value $B$ commits to, on receiving commitments to $v_1$ or $v_2$, with probability $\frac{p_1}{m} - p_2 \geq \frac{1}{mp(n)}$, which contradicts the stand-alone non-malleability of $\langle C, R\rangle$. □

### 2.10 Robustness: Non-Malleability w.r.t. $k$-round Protocols

The concept of non-malleability is traditionally only considered in a setting where a man-in-the middle adversary is participating in two (or more) executions of the *same* protocol. We here consider a notion of non-malleability with respect to arbitrary $k$-round protocols.

Consider a one-many man-in-the-middle adversary $A$ (as shown in figure 3) that participates in one left interaction—communicating with a machine $B$—and many right interactions—acting as a committer using the commitment scheme $\langle C, R\rangle$. As in the standard definition of non-malleability, $A$ can adaptively choose the identities in the right interactions. We denote by $\mathsf{mim}^{B,A}_{\langle C,R\rangle}(y, z)$ the random variable consisting of the view of $A(z)$ in a man-in-the-middle execution when communicating with $B(y)$ on the left and honest receivers on the right, combined with the values $A(z)$ commits to on the right. Intuitively, we say that $\langle C, R\rangle$ is one-many non-malleable w.r.t $B$ if $\mathsf{mim}^{B,A}_{\langle C,R\rangle}(y_1, z)$ and $\mathsf{mim}^{B,A}_{\langle C,R\rangle}(y_2, z)$ are indistinguishable, whenever interactions with $B(y_1)$ and $B(y_2)$ cannot be distinguished.

**Definition 7.** *Let $\langle C, R\rangle$ be a commitment scheme, and $B$ a $\mathcal{PPT}$ ITM. We say the commitment scheme $\langle C, R\rangle$ is* one-many non-malleable w.r.t. *$B$, if for every two sequences $\{y^1_n\}_{n\in N}$ and $\{y^2_n\}_{n\in N}$, $y^1_n, y^2_n \in \{0,1\}^n$, such that, for all $\mathcal{PPT}$ ITM $\tilde{A}$, it holds that*

$$\left\{\langle B(y^1_n), \tilde{A}(z)\rangle(1^n)\right\}_{n\in N, z\in\{0,1\}^*} \approx \left\{\langle B(y^2_n), \tilde{A}(z)\rangle(1^n)\right\}_{n\in N, z\in\{0,1\}^*}$$

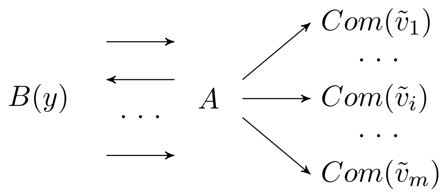


Figure 3: A concurrent man-in-the-middle adversary with respect to protocol $B$ on input $y$.

*then it also holds that, for every $\mathcal{PPT}$ one-many man-in-the-middle adversary $A$,*

$$\left\{\mathsf{mim}^{B,A}_{\langle C,R\rangle}(y^1_n, z)\right\}_{n\in N, z\in\{0,1\}^*} \approx \left\{\mathsf{mim}^{B,A}_{\langle C,R\rangle}(y^2_n, z)\right\}_{n\in N, z\in\{0,1\}^*}$$

We say that $\langle C, R\rangle$ is one-many non-malleable w.r.t $k$-round protocols if $\langle C, R\rangle$ is one-many non-malleable w.r.t any machine $B$ that interacts with the man-in-the-middle adversary in $k$ rounds. Such commitment schemes are easy to construct: any commitment scheme that is "extractable" and has more than $k$ "rewinding slots" is directly one-many non-malleable w.r.t. $k$-round protocols. (A formal proof of this is obtained identically to the proof of Claim 5; See Section 4.2.2). In this work, we focus on non-malleability w.r.t 4-round protocols; we call such protocols *robust*.

As mentioned earlier, to bootstrap the amplification technique, we need a constant-round robust $O(1)$-non-malleable commitment scheme. By observing that the $O(\log n)$-round non-malleable commitment protocol of [LPV08] (which is based on [DDN00]) contains $O(\log n)$ "rewinding slots", we directly get:

**Theorem 5.** *Assume the existence of one-way functions. Then there exists a constant-round robust $O(1)$-non-malleable commitment scheme.*

But, to obtain a constant-round $O(1)$-non-malleable commitment schemes, it suffices to device an protocol whose round complexity grows exponentially with the length of the identities. For self-containment, we provide a simple construction of such a protocol in Appendix A (thus providing an alternative, simpler, proof of Theorem 5).

## 3 Proof of Lemma 1—The DDN "log n trick"

Given a robust $t(n)$-concurrently non-malleable commitment scheme $\langle C, R\rangle$, we construct a robust $2^{t(n)-1}$-non malleable commitment scheme $\langle \tilde{C}, \tilde{R}\rangle$, following the construction by Dolev, Dwork and Naor [DDN00].

**Description of the Protocol $\langle \tilde{C}, \tilde{R}\rangle(\mathsf{id})$ [DDN00, LPV08]:** Let $l$ be the length of the identity $\mathsf{id}$, $l = |\mathsf{id}| = 2^{t(n)-1}$. To commit to value $v \in \{0,1\}^n$, choose $l$ random shares $r_0, \dots, r_{l-1} \in \{0,1\}^n$, such that $v = r_0 \oplus \dots \oplus r_{l-1}$. Then for each $0 \le i \le l-1$, commit to $r_i$ (in parallel) using $\langle C, R\rangle$ with identity $(i, \mathsf{id}_i)$, where $\mathsf{id}_i$ is the $i$th bit of $\mathsf{id}$.

The protocol $\langle \tilde{C}, \tilde{R}\rangle$ consists of $2^{t(n)-1}$ parallel executions of $\langle C, R\rangle$ with $t(n)$-bit identities. Thus, the round complexity of $\langle \tilde{C}, \tilde{R}\rangle$ is the same as $\langle C, R\rangle$, and the computational complexity increases by a factor of $2^{t(n)-1}$. Hiding and initial-binding follows using standard techniques. Moreover, as is shown in [LPV08], $\langle \tilde{C}, \tilde{R}\rangle$ is stand-alone non-malleable.

**Proposition 3** ([LPV08])**.** *If $\langle C, R\rangle$ is concurrently non-malleable, then $\langle \tilde{C}, \tilde{R}\rangle$ is stand-alone non-malleable.*

Furthermore, it is easy to see that $\langle \tilde{C}, \tilde{R} \rangle$ is also robust.

**Proposition 4.** *If $\langle C, R \rangle$ is one-many non-malleable w.r.t. arbitrary $k$-round protocols, then $\langle \tilde{C}, \tilde{R} \rangle$ is also one-many non-malleable w.r.t. arbitrary $k$-round protocols.*

*Proof.* Consider some adversary $A$, machine $B$, distinguisher $D$, inputs $x_1$, $x_2$, and $z$, such that $D$ distinguishes the view of $A$ and the values it commits to using $\langle \tilde{C}, \tilde{R} \rangle$, after interacting with $B$ on inputs $x_1$ or $x_2$, with probability $\epsilon$. Note that $A$ can also be viewed as a one-many adversary for $\langle C, R \rangle$ (with respect to $B$). Now, consider the distinguisher $D'$ that on input the view and values $\{\tilde{v}_i^j\}$ committed to by $A$ using $\langle C, R \rangle$—where $\tilde{v}_i^j$ denotes the $j$th "share" of the $i$th value committed to by $A$ using $\langle \tilde{C}, \tilde{R} \rangle$, lets $\tilde{v}_i = \oplus_{j \in [l]} \tilde{v}_i^j$ and outputs $D(\mathcal{V}, \tilde{v}_1, \ldots, \tilde{v}_m)$. It follows that $D'$ distinguishes the view of $A$ and the values it commits to using $\langle C, R \rangle$, after interacting with $B$ on inputs $x_1$ or $x_2$ with probability $\epsilon$, which therefore must be negligible (in $n$) by the robustness of $\langle C, R \rangle$. □

We note, however, that $\langle \tilde{C}, \tilde{R} \rangle$ is not concurrently non-malleable. Consider the following man-in-the-middle adversary $A$ that participates in one left and two right interactions: on the left, $A$ receives a commitment of $\langle \tilde{C}, \tilde{R} \rangle$ using identity $\mathsf{id}_L$, where $\mathsf{id}_L$ is $l$ bits long and $l$ is even; on the right, $A$ chooses identities $\mathsf{id}_R^1$ and $\mathsf{id}_R^2$, such $\mathsf{id}_L \neq \mathsf{id}_R^1, \mathsf{id}_L \neq \mathsf{id}_R^2$, but the first $l/2$ bits of $\mathsf{id}_R^1$ are the same as those in $\mathsf{id}_L$, and the last $l/2$ bits of $\mathsf{id}_R^2$ are the same as those in $\mathsf{id}_L$. To complete the two right commitments, $A$ simply forwards the first $l/2$ parallel executions of $\langle C, R \rangle$ in the left interaction to the first right commitment (as the first $l/2$ parallel executions), and the remaining $l/2$ parallel executions in the left to the second right interaction (as the last $l/2$ parallel executions); for the remaining parallel executions on the right, $A$ commits to 0 using $\langle C, R \rangle$. By construction, the XOR of the two values committed to in the the right executions equals to the value committed to on the left, and thus $\langle \tilde{C}, \tilde{R} \rangle$ cannot be concurrently non-malleable.

In the next section, we show how to restore concurrent non malleability.

# 4 Proof of Lemma 2—The Strengthening Technique

The $\log n$ trick in the previous section allows us to increase the length of identities used in a non-malleable commitment protocol while keeping the round complexity constant. The transformation, however, weakens the quality of non-malleability; namely, we need to start off with a concurrently non-malleable protocol, but after the transformation we only get back a stand-alone non-malleable protocol. We provide a technique for restoring concurrent non malleability without sacrificing too much in round complexity.

More specifically, the strengthening technique to be introduced shows how to transform any robust stand-alone non-malleable commitment into a robust concurrent non-malleable commitment, while only increasing the round complexity by a constant. Below we first present a strengthening technique that increases the round complexity by a constant *multiplicative* factor, and then show how to improve the round complexity further by running different subprotocols in parallel.

## 4.1 The Protocol $\langle \hat{C}, \hat{R} \rangle$

For simplicity of exposition, our description below relies on the existence of one-way functions with efficiently recognizable range, but the protocol can be easily modified to work with any arbitrary one-way function (by simply providing a witness hiding proof that an element is in the range of the one-way function; see Remark 4.2 for more details). Given an $k(n)$-round robust stand-alone non-malleable commitment protocol $\langle \tilde{C}, \tilde{R} \rangle$, the construction of $\langle \hat{C}, \hat{R} \rangle$ proceeds as follows. To commit to a value $v$, the Committer and the Receiver, on common input the identity $\mathsf{id} \in \{0,1\}^l$, and $1^n$, where $n$ is the security parameter, proceed in six stages:

**Stage 1** the Receiver picks a random string $r \in \{0,1\}^n$, and sends its image $s = f(r)$, through a one-way function $f$ with an efficiently recognizable range, to the Committer. The Committer checks that $s$ is in the range of $f$ and aborts otherwise. Furthermore, the receiver also sends in parallel the first message of a commitment of $\mathsf{com}$, where $\mathsf{com}$ is any two-round statistically-binding string commitment.

**Stage 2** the Committer sends the second message of a commitment of $\mathsf{com}$ to $v$.

**Stage 3** the Committer commits to $v$ (again) using $\langle \tilde{C}, \tilde{R} \rangle$ and identity $\mathsf{id}$.

**Stage 4** the Committer commits to $0^n$ using $\langle \tilde{C}, \tilde{R} \rangle$ and identity $\mathsf{id}$.

**Stage 5** the Committer performs a 4-round $\mathcal{WISSP}$ proof $\langle P, V \rangle$ of the statement that it has committed to $0^n$ in Stage 4, *or* it knows a pre-image of $s$.

**Stage 6** the Committer proves that it has committed to value $v$ in both Stage 2 and 3, or a pre-image of $s$ in Stage 4. This is proved using $k(n) + 1$ sequential invocations of the 4-round $\mathcal{WISSP}$ protocol $\langle P, V \rangle$.

Let $\Delta$ be a transcript of the protocol $\langle \tilde{C}, \tilde{R} \rangle$. We define the value committed to in the transcript to be the value committed to in Stage 2 (using $\mathsf{com}$) of $\Delta$. That is, in the reveal stage, the committer needs to reveal the committed value and the randomness used in Stage 2 of the protocol as the decommitment. A formal description of the protocol appears in Figure 4.
It is easy to see that $\langle \hat{C}, \hat{R} \rangle$ has $6k(n)+10 \leq 16k(n)$ rounds. Furthermore, it follows using standard techniques that $\langle \hat{C}, \hat{R} \rangle$ is a commitment scheme.

**Proposition 5.** *$\langle \hat{C}, \hat{R} \rangle$ is a statistically binding commitment scheme, with initial binding.*

*of Proposition 5.* The binding property follows directly from the statistically binding property of $\mathsf{com}$ used in Stage 2. The protocol is initial binding by construction, since the first message from the committer is (the second message of) the $\mathsf{com}$ commitment. For the hiding property, we show that any adversary $R^*$ that violates the hiding property of $\langle \hat{C}, \hat{R} \rangle$ can be used to violate the hiding property of $\langle \tilde{C}, \tilde{R} \rangle$. More precisely, given any adversary $R^*$ (without loss of generality, deterministic) that distinguishes commitments made using $\langle \hat{C}, \hat{R} \rangle$, we construct a machine $R'$ that distinguishes commitments made using $\langle \tilde{C}, \tilde{R} \rangle$. Let $s$ be the first message sent by $R^*$. $R'$ on auxiliary input $r$, such that, $s = f(r)$, proceeds as follows. It internally incorporates $R^*$ and simulates a $\langle \hat{C}, \hat{R} \rangle$ commitment to $R^*$ by committing to 0 in Stage 2 and forwarding the external commitment made using $\langle \tilde{C}, \tilde{R} \rangle$ to $R^*$ in Stage 3. For the rest of the commitment, it uses the "fake witness" $r$. More specifically, it commits to $r$ instead of $0^n$ in Stage 4, and then gives $\mathcal{WI}$ proofs that it has committed to a pre-image of $s$ in Stage 6; in Stage 5, it proves its knowledge of a pre-image of $s$. Finally, it outputs whatever $R^*$ outputs. From the hiding property of $\mathsf{com}$, $\langle \tilde{C}, \tilde{R} \rangle$ and the $\mathcal{WI}$ property of Stage 5 and 6, it follows that $R'$ distinguishes the commitments made using $\langle \tilde{C}, \tilde{R} \rangle$, if $R^*$ distinguishes the commitments made using $\langle \hat{C}, \hat{R} \rangle$. □

## 4.2 Proof of Concurrent Non-malleability

To prove that $\langle \hat{C}, \hat{R} \rangle$ is a robust concurrent non-malleable commitment we need to show the following two claims.

**Protocol $\langle \hat{C}, \hat{R} \rangle$**

**Common Input:** A security parameter $1^n$ and an identity $\mathsf{id} \in \{0,1\}^l$.

**Auxiliary Input for Committer:** A string $v \in \{0,1\}^n$.

**Commit Stage**

**Stage 1**

R uniformly chooses $r \in \{0,1\}^n$

R → C: $s = f(r)$ and the first message $m$ of a commitment of $\mathsf{com}$.

C aborts if $s$ is not in the range of $f$.

**Stage 2**

C → R: the second message $c$ of a commitment of $\mathsf{com}$ to $v$, in reply to $m$.

**Stage 3**

C → R: a commitment to $v$ using the protocol $\langle \tilde{C}, \tilde{R} \rangle$. Let $\mathcal{T}_1$ be the transcript generated.

**Stage 4**

C → R: a commitment to $0^n$ using the protocol $\langle \tilde{C}, \tilde{R} \rangle$. Let $\mathcal{T}_2$ be the transcript generated.

**Stage 5**

C → R: a $\mathcal{WISSP}$ proof of $\langle P, V \rangle$ of the statement that:

- *either* $\mathcal{T}_2$ is a valid a commitment to $0^n$,
- or there exists a value $r$, s.t $s = f(r)$.

**Stage 6**

C → R: $k(n) + 1$ $\mathcal{WISSP}$ proofs of $\langle P, V \rangle$ of the statement:

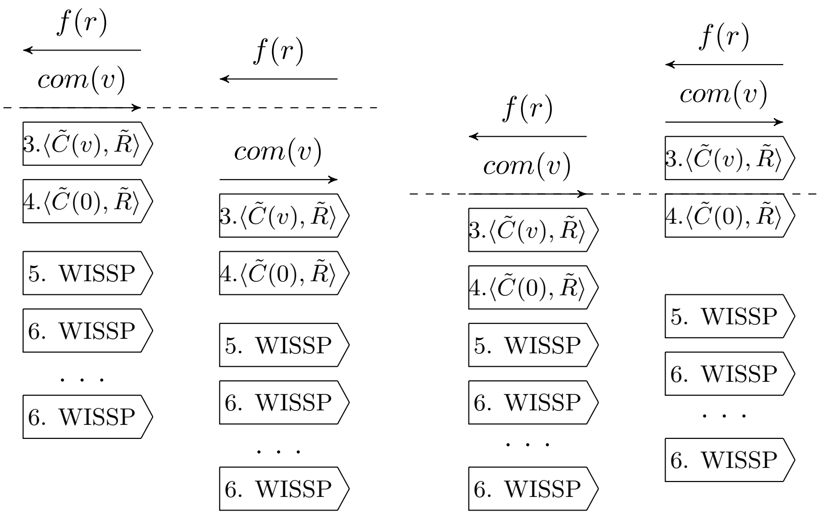
- *either* there exists a value $v$, s.t $(m, c)$ is a valid commitment to $v$ using $\mathsf{com}$ and $\mathcal{T}_1$ is a valid commitment to $v$ using $\langle \tilde{C}, \tilde{R} \rangle$,
- *or* there exists a value $r$ s.t $s = f(r)$ and $\mathcal{T}_2$ is a valid commitment to $r$ using $\langle \tilde{C}, \tilde{R} \rangle$.

**Reveal Stage**

C → R: $(v, \sigma)$

$R$ outputs 1 if the honest committer of $\mathsf{com}$, on receiving private input $v$, random tape $\sigma$, and message $m$ in the first round, outputs $c$; it outputs 0 otherwise.

Figure 4: A Concurrent Non-Malleable Commitment Scheme $\langle \hat{C}, \hat{R} \rangle$



(i) The simplified case (ii) The general case

Figure 5: The simplified case and the general case.

**Proposition 6.** *$\langle \hat{C}, \hat{R} \rangle$ is one-many non-malleable w.r.t 4-round protocols.*

**Proposition 7.** *$\langle \hat{C}, \hat{R} \rangle$ is one-many non-malleable.*

The proof of Proposition 6 follows using standard techniques (by reducing the robustness property of $\langle \hat{C}, \hat{R} \rangle$ to that of $\langle \tilde{C}, \tilde{R} \rangle$) and we thus postpone it to Section 4.2.1. We here turn to the proof of Proposition 7. Below, for simplicity of notation, unless otherwise specified, we view the Stage 2 commitment as a non-interactive commitment; however, the reader should keep in mind that it is actually a *two-round* commitment where the first round sent in Stage 1.

*of Proposition 7.* We want to prove that for any man-in-the-middle adversary $A$ that participates in one left execution and many right executions, the following ensembles are indistinguishable:

$$\left\{ \mathsf{mim}^{A}_{\langle \hat{C}, \hat{R} \rangle}(v_1, z) \right\}_{n \in N, v_1, v_2 \in \{0,1\}^n, z \in \{0,1\}^*}$$

$$\left\{ \mathsf{mim}^{A}_{\langle \hat{C}, \hat{R} \rangle}(v_2, z) \right\}_{n \in N, v_1, v_2 \in \{0,1\}^n, z \in \{0,1\}^*}$$

Assume, for contradiction, that there exists a $\mathcal{PPT}$ $A$, a distinguisher $D$ and a polynomial $p(n)$, such that for infinitely many $n \in N$, there exist $v_1, v_2, z$ such that $D$ distinguishes $\mathsf{mim}^{A}_{\langle \hat{C}, \hat{R} \rangle}(v_1, z)$ and $\mathsf{mim}^{A}_{\langle \hat{C}, \hat{R} \rangle}(v_2, z)$ with probability $\frac{1}{p(n)}$. Towards reaching a contradiction, we first consider a simplified case, in which $A$ is restricted to always send its first message (i.e, the Stage 2 commitment) in the right interactions after receiving the first message (again, the Stage 2 commitment) in the left interaction, as shown in Figure 5 (i). Intuitively, it suffices to consider this simplified case as the values $A$ commits to (in Stage 2 on the right) before receiving the first message on the left are trivially independent of the value committed to on the left.

**Analysis of the simplified case:** We proceed to analyze the case when $A$ always sends its first message in the right interaction after receiving the first message of the left interaction. We define a sequence of hybrid experiments $H_0, \ldots, H_5$. In each of these experiments, we show that the view of $A$, combined with the values $A$ commits to in *Stage 3* on the right are, indistinguishable. (Note that we here consider the values that $A$ commits to in Stage 3, and not the values committed to in Stage 2 as in the definition of $\mathsf{mim}^A_{\langle \hat{C},\hat{R}\rangle}$.)

Furthermore, in each of these experiments, we also show that, except with negligible probability, $A$ is never able to commit to a "fake" witness $r$—that is, a pre-image of $s$ (the Stage 1 message from a right receiver)—in Stage 4 of any successful right interactions that has a different identity from the left interaction. Below, we let $\mathsf{hyb}_i(v,z)$ denote the random variable describing the view of $A(z)$, combined with the values it commits to in *Stage 3* of the right interactions in hybrid $H_i$ (as usual, the committed value is replaced with $\perp$ if the right interaction fails or if $A$ has copied the identity of the left interaction).

**Hybrid** $H_0$ Hybrid $H_0$ consists of an "honest" emulation of $\mathsf{mim}^A_{\langle \hat{C},\hat{R}\rangle}(v,z)$. Note that it follows directly from the one-wayness of $f$ and the proof of knowledge property of Stage 5, that, except with negligible probability, $A$ never commits to a "fake" witness $r$ in any successful right execution (or else we can use $A$ to invert $f$). Furthermore, note that $\mathsf{hyb}_0(v,z)$ is identically defined to $\mathsf{mim}^A_{\langle \hat{C},\hat{R}\rangle}(v,z)$ except that in $\mathsf{mim}^A_{\langle \hat{C},\hat{R}\rangle}(v,z)$ we consider the values committed to by $A$ in Stage 2, whereas in $\mathsf{hyb}_0(v,z)$ we consider the values committed to in Stage 3. However, by the soundness of the Stage 6 proof system, it holds that, in every successful right interaction, either the values committed to in Stage 2 and 3 are the same, or $A$ has committed to a "fake" witness in Stage 4. Since the latter only happens with negligible probability, we have that the combined and view and values committed by $A$ in $\mathsf{hyb}_0(v,z)$ and $\mathsf{mim}^A_{\langle \hat{C},\hat{R}\rangle}(v,z)$ are statistically close.

**Claim 1.** *For every $\mathcal{PPT}$ adversary $A$, it holds that:*

$$\left\{\mathsf{mim}^A_{\langle \hat{C},\hat{R}\rangle}(v,z)\right\}_{n\in N, v\in\{0,1\}^n, z\in\{0,1\}^*} \approx \left\{\mathsf{hyb}_0(v,z)\right\}_{n\in N, v\in\{0,1\}^n, z\in\{0,1\}^*}$$

**Hybrid** $H_1$ Hybrid $H_1$ proceeds identically to $H_0$ except that the left execution is emulated by finding a fake witness $r$, using a brute-force search, and next using this $r$ when emulating Stage 5; everything else remains the same. It then follows from the one-many non-malleability with respect to 4-round protocols of Stage 3 and the witness indistinguishability property of Stage 5 of the left interaction, that the view of $A$ and the values it commits to in Stage 3 on the right are indistinguishable to those in $H_0$. To formalize this, recall that the notion of non-malleability with respect to 4-round protocols only guarantees indistinguishability if the left 4-round interaction is efficiently computable. This holds here since, as we only consider the simplified case, the Stage 2 commitments on the right are always sent after the Stage 2 commitment on the left, and hence, so are the Stage 3 commitments on the right. Thus the fake witness $r$ is determined, and can be fixed non-uniformly, before any Stage 3 commitment on the right begins; then, given $r$, the rest of the left interaction, and in particular the 4-round $\mathcal{WI}$ proof, can be generated efficiently. Therefore, robustenss can be applied here.

**Claim 2.** *For every $\mathcal{PPT}$ adversary $A$, it holds that:*

$$\{\mathsf{hyb}_0(v,z)\}_{n\in N, v\in\{0,1\}^n, z\in\{0,1\}^*} \approx \{\mathsf{hyb}_1(v,z)\}_{n\in N, v\in\{0,1\}^n, z\in\{0,1\}^*}$$

The proof of Claim 2 (which simply formalizes the above argument) can be found in Section 4.2.2. Additionally, by the same property of the protocol in Stage 4, it follows using exactly the same proof as for Claim 2 that, except with negligible probability, $A$ never commits to a "fake" witness in any successful right interactions (since this was the case in $H_0$.)

**Claim 3.** *There exists some negligible function $\mu$, such that for every $\mathcal{PPT}$ adversary $A$, and every inputs $v,\ z \in \{0,1\}^*$, the probability that $A$ commits to a "fake" witness, in a successful right interaction that uses a different identity from the left interaction, in $\mathsf{hyb}_1(v,z)$, is smaller than $\mu(n)$.*

**Hybrid** $H_2$ Hybrid $H_2$ proceeds identically to $H_1$ except that the left execution is emulated by committing to the fake witness $r$ in Stage 4. Since (by the scheduling constraint on $A$ in the simplified case) Stage 4 of the right interactions start after Stage 2 of the left interaction, it follows from the one-many *weak* non-malleability of Stage 4 that $A$ can only commit to a fake witness with negligible probability. A formal proof of Claim 4 below appears in Section 4.2.2.

**Claim 4.** *There exists some negligible function $\mu$, such that for every $\mathcal{PPT}$ adversary $A$, and every input $v,\ z \in \{0,1\}^*$, the probability that $A$ commits to a "fake" witness, in a successful right interaction that uses a different identity from the left interaction, in $\mathsf{hyb}_2(v,z)$, is smaller than $\mu(n)$.*

It only remains to show that the combined view and values committed to by $A$ in Stage 3 are indistinguishable to those in $H_1$. On a high-level, this follows from the fact that the commitment in Stage 4 only has $k(n)$ rounds, but Stage 6 consists of $k(n)+1$ special-sound proofs. Thus, for each execution on the right, there exists some special-sound proof that does not contain any of the Stage 4 messages from the left execution; thus we can use rewindings to extract the witness used in this proof *without rewinding* Stage 4 on the left. Since we already showed that $A$ can only commit to a fake witness with negligible probability, it follows from the special soundness of the proof that, except with negligible probability, the value extracted out are the values committed to by $A$ in Stage 3. It now follows from the hiding property of Stage 4 on the left (and the fact that Stage 4 of the left interaction is never rewound) that the the view combined with the values extracted out (which in turn are statistically close to the values committed to by $A$ in Stage 3) are indistinguishable to those in hybrid $H_1$. The proof of Claim 5 below can be found in Section 4.2.2.

**Claim 5.** *For every $\mathcal{PPT}$ adversary $A$, it holds that:*

$$\{\mathsf{hyb}_1 v, z\}_{n\in N, v\in\{0,1\}^n, z\in\{0,1\}^*} \approx \{\mathsf{hyb}_2 v, z\}_{n\in N, v\in\{0,1\}^n, z\in\{0,1\}^*}$$

**Hybrids** $H_3^1$ **to** $H_3^{k(n)+1}$ In hybrids $H_3^1$ to $H_3^{k(n)+1}$, we change the witness used in the $k(n)+1$ $\mathcal{WISSP}$ proofs in Stage 6 of the left interaction. More specifically, the experiment $H_3^i$ proceeds identically to $H_2$, except that in the first $i$ Stage 6 proofs on the left, we use the fact that we have committed to a fake witness $r$ in Stage 4. As the only difference between two consecutive hybrids $H_3^i$ and $H_3^{i+1}$ is the witness used in a single $\mathcal{WISSP}$ proof (i.e., the $i+1$'th proof) on the left, it follows using the same argument as in hybrid $H_1$ (*i.e.*, from the one-many non-malleability of Stage 3 and 4 w.r.t to 4-round protocols and the witness indistinguishability of the 4-round proofs in Stage 6) that the combined view and values committed to by $A$ in Stage 3 are indistinguishable, and that $A$ almost never commits to a fake witness in any successful right interaction that uses a different identity from the left interaction. We conclude by a hybrid argument that this holds also in hybrid $H_3^{k(n)+1}$.

**Hybrid** $H_4$ Hybrid $H_4$ proceeds identically to $H_3^{k(n)+1}$ except that Stage 3 of the left execution is emulated by committing to 0. It follows using the same argument as in hybrid $H_2$ that the combined view and values committed to by $A$ in Stage 3 are indistinguishable (from those in

$H_3^{k(n)+1}$), and that $A$ only commits to a fake witness in a successful right interaction with a different identity from the left interaction with negligible probability.

**Hybrid** $H_5$ Hybrid $H_5$ proceeds identically to $H_4$ except that the Stage 2 commitment of the left execution is emulated by committing to 0. It follows using the same argument as in hybrid $H_1$ (relying on the non-malleability w.r.t. 4-round protocols property of Stage 3 on the right) and the the hiding property of the Stage 2 commitment on the left that the combined view and values committed to by $A$ in Stage 3 are indistinguishable from those in $H_4$. (Note that "non-malleability w.r.t. 1-round protocols" (of Stage 3 on the right) suffices here since Stage 2 only contains a single message (i.e., the second message of the commitment using $\mathsf{com}$); furthermore, note that the hiding property of this (two-round) commitments holds no matter what first message the receiver has sent in Stage 1).

It follows by a hybrid argument that,

$$\left\{\mathsf{mim}^{A}_{\langle\hat{C},\hat{R}\rangle}(v,z)\right\}_{n\in N,v\in\{0,1\}^n,z\in\{0,1\}^*} \approx \left\{\mathsf{hyb}_5(v,z)\right\}_{n\in N,v\in\{0,1\}^n,z\in\{0,1\}^*}$$

Since the above holds for every value $v$, we have

$$\begin{aligned}\left\{\mathsf{mim}^{A}_{\langle\hat{C},\hat{R}\rangle}(v_1,z)\right\}_{n\in N,v_1,v_2\in\{0,1\}^n,z\in\{0,1\}^*} &\approx \left\{\mathsf{hyb}_5(v_1,z)\right\}_{n\in N,v_1,v_2\in\{0,1\}^n,z\in\{0,1\}^*}\\ \left\{\mathsf{mim}^{A}_{\langle\hat{C},\hat{R}\rangle}(v_2,z)\right\}_{n\in N,v_1,v_2\in\{0,1\}^n,z\in\{0,1\}^*} &\approx \left\{\mathsf{hyb}_5(v_2,z)\right\}_{n\in N,v_1,v_2\in\{0,1\}^n,z\in\{0,1\}^*}\end{aligned}$$

Finally, since by definition of $\mathsf{hyb}_5$, it holds that for every $v_1$, $v_2$ and $z$, $\mathsf{hyb}_5(v_1,z) = \mathsf{hyb}_5(v_2,z)$, the following ensembles are indistinguishable.

$$\left\{\mathsf{mim}^{A}_{\langle\hat{C},\hat{R}\rangle}(v_1,z)\right\}_{n\in N,v_1,v_2\in\{0,1\}^n,z\in\{0,1\}^*}$$

$$\left\{\mathsf{mim}^{A}_{\langle\hat{C},\hat{R}\rangle}(v_1,z)\right\}_{n\in N,v_1,v_2\in\{0,1\}^n,z\in\{0,1\}^*},$$

which concludes the proof for the simplified case.

**Analysis of the General Case:** Next we consider the general case where the adversary might send some Stage 2 commitments on the right earlier than the Stage 2 commitment of the left interaction, as shown in Figure 5 (ii). Let $\Gamma(A,v,z)$ denote the distribution of the joint views $\tau$ of $A$ and the right receivers, such that after $\tau$, the left committer immediately sends the Stage 2 commitment in the left interaction. By our hypothesis, it follows using an averaging argument that for at least a fraction $\frac{1}{2p(n)}$ of the joint views $\tau$ in $\Gamma(A,v,z)$,

- $D$ distinguishes $\left\{\mathsf{mim}^{A}_{\langle\hat{C},\hat{R}\rangle}(v_1,z)|\tau\right\}$ and $\left\{\mathsf{mim}^{A}_{\langle\hat{C},\hat{R}\rangle}(v_2,z)|\tau\right\}$ with probability at least $\frac{1}{2p(n)}$, where $\left\{\mathsf{mim}^{A}_{\langle\hat{C},\hat{R}\rangle}(v,z)|\tau\right\}$ denotes the output of $\mathsf{mim}^{A}_{\langle\hat{C},\hat{R}\rangle}(v,z)$ conditioned on the event that the execution is consistent with $\tau$.

Furthermore, by the one-wayness of $f$, and the proof of knowledge property of Stage 5, it follows that, except with negligible probability, $A$ never commits to a "fake" witness in any successful right executions in an honest man-in-the-middle execution. Therefore, for at least a fraction $\frac{1}{3p(n)}$ of the joint views $\tau$ in $\Gamma(A,v,z)$, the following two conditions hold:

- $D$ distinguishes $\left\{\mathsf{mim}^{A}_{\langle\hat{C},\hat{R}\rangle}(v_1,z)|\tau\right\}$ and $\left\{\mathsf{mim}^{A}_{\langle\hat{C},\hat{R}\rangle}(v_2,z)|\tau\right\}$ with probability at least $\frac{1}{2p(n)}$;

- $A$ only commits to a fake witness in any successful right interactions in the experiment $\left\{\mathsf{mim}^A_{\langle\hat{C},\hat{R}\rangle}(v_1,z)|\tau\right\}$ or $\left\{\mathsf{mim}^A_{\langle\hat{C},\hat{R}\rangle}(v_2,z)|\tau\right\}$ with negligible probability.

We call such a joint view $\tau$ "good". Consider any "good" view $\tau$. Without loss of generality (by a renumbering of right interactions) we assume that only the first $\ell$ right interactions have their Stage 2 commitments sent inside $\tau$; in the remaining $m-\ell$ right interactions, $A$ sends the Stage 2 commitment after $\tau$. By the statistically binding property of $\mathsf{com}$, except with negligible probability, the values $A$ commits to in Stage 2 in the first $\ell$ right interactions are decided by $\tau$; let $\tilde{v}_1,\ldots,\tilde{v}_\ell$ be the committed values. Now consider a man-in-the-middle execution in which $A$ and the right receivers are fed with their view in $\tau$. We show that there exists a distinguisher $D'$ that distinguishes the view and the values $A$ commits to, after $\tau$, with probability at least $\frac{1}{2p(n)}$. $D'$, on input $\rho,\tilde{v}_{\ell+1},\ldots,\tilde{v}_m$, and auxiliary input $\tau,\tilde{v}_1,\ldots,\tilde{v}_\ell$, simply outputs $D(\tau\|\rho,\tilde{v}_1,\ldots,\tilde{v}_m)$. As $D$ distinguishes $\left\{\mathsf{mim}^A_{\langle\hat{C},\hat{R}\rangle}(v_1,z)|\tau\right\}$ and $\left\{\mathsf{mim}^A_{\langle\hat{C},\hat{R}\rangle}(v_2,z)|\tau\right\}$ with probability at least $\frac{1}{2p(n)}$, $D'$ also distinguishes the view and values committed to by $A$, after $\tau$, with probability at least $\frac{1}{2p(n)}$. Finally, recall that the proof of the simplified case only relies on the following two facts about $A$:

- $A$ always sends the Stage 2 commitments on the right after receiving the Stage 2 commitment from the left committer, and
- $A$ only never commits to a "fake" witness in any successful right interactions with negligible probability.

For every "good" $\tau$, it holds that, in a man-in-the-middle execution consistent with $\tau$,

- In the last $m-\ell$ right interactions, $A$ sends the Stage 2 commitments after receiving the Stage 2 commitment from the left committer, and
- $A$ only commits to a "fake" witness in any successful right interactions with negligible probability.

Therefore, it follows using exactly the same proof as that for the simplified case that the view and the value $A$ commits to in the last $m-\ell$ right interactions, i.e., after $\tau$, are indistinguishable, which gives a contradiction.

□

**Remark.** *The protocol $\langle\hat{C},\hat{R}\rangle$ described above uses a one-way function with efficiently recognizable range in its first stage. It can be easily modified to work with any arbitrary one-way function, by adding one more stage after Stage 1. In this new stage, the receiver provides a witness hiding proof that the message it sends in Stage 1 is in the range of the one-way function; the committer then verifies the proof and aborts if it is not convincing. It follows by the same proofs as above that the modified protocol is a robust concurrent non-malleable commitment scheme.*

### 4.2.1 Proof of Proposition 6

*Proof.* Assume that there exists some adversary $A$, machine $B$, distinguisher $D$ and a polynomial $p$, such that for infinitely many $n$ there exist $x_1,x_2,z$ such that $D$ distinguishes the combined view and values committed to by $A(z)$ when interacting with $B$ on input either $x_1$ or $x_2$, with probability $\frac{1}{p(n)}$. We show that there exist an adversary $A'$ and a distinguisher $D'$ such that $D'$ distinguishes only the view of $A'(z)$ in interactions with $B$ on inputs $x_1$ and $x_2$.

Fix one such $n$, $x_1$, $x_2$ and $z$. Given $A$, we construct an adversary $\tilde{A}$ attacking the non-malleability of $\langle \tilde{C}, \tilde{R} \rangle$ with respect to $B$. More specifically, $\tilde{A}(z)$ externally interacts with $B$ on input either $x_1$ or $x_2$; internally, it incorporates $A(z)$ and forwards all the messages from $B$ to $A$; furthermore, it forwards externally all the $\langle \tilde{C}, \tilde{R} \rangle$ commitments in Stage 3 from the right interactions of $A$; finally it outputs whatever $A$ outputs. We show that there exists a distinguisher $\tilde{D}$ that distinguishes the combined view and committed values of $\tilde{A}$, with probability $\frac{1}{p(n)}$. $\tilde{D}$, on input the view $\mathcal{V}$ and values $\tilde{v}_1, \ldots, \tilde{v}_m$ committed to by $\tilde{A}$, reconstructs the view and values committed to by $A$ in the emulation by $\tilde{A}$. This is done by reconstructing the view $\mathcal{V}_A$ of $A$ from $\mathcal{V}$, and setting $\tilde{v}'_k = \tilde{v}_k$ if interaction $k$ succeeds and $\perp$ otherwise. It then invokes the distinguisher $D$ on the reconstructed view $\mathcal{V}_A$ and values $\tilde{v}'_k$, and outputs the output of $D$. Notice that the view of $A$ in simulation by $\tilde{A}$ is identically distributed to that in a real execution. By the one-wayness of $f$ and the proof of knowledge property of Stage 5, we know that, except from negligible probability, none of the Stage 4 commitment in any successful right interaction would be a commitment to a "fake" witness (i.e., a pre-image of the first message $s$ sent by the receiver). Therefore, it follows from the soundness of the Stage 6 argument that, in simulation by $\tilde{A}$, except with negligible probability, the value $A$ commits to in Stage 3 of any successful right interaction is the same as that in Stage 2 (and hence, by definition, the value $A$ commits to in the corresponding right interaction). Thus, except with negligible probability, $\mathcal{V}_A$ combined with the values $\tilde{v}'_k$ are identically distributed to the view and values committed to by $A$ in a real execution. Therefore $\tilde{D}$ distinguishes the view and values committed to by $\tilde{A}$ with probability at least $\frac{1}{2p(n)}$.

It then follows from the robustness of $\langle \tilde{C}, \tilde{R} \rangle$ that there exists an adversary $A'$, a distinguisher $D'$, and a polynomial $q$, such that for infinitely many $n$, $D'$ distinguishes, just the view of $A'$ in interactions with $B$ on input either $x_1$ or $x_2$ with probability $\frac{1}{q(n)}$.

□

### 4.2.2 Proof of Claims

In the proofs below, we say that an adversary is "well-behaved" if it always sends its first message in the right interactions after receiving the first message in the left interaction (i.e., the adversary respect the scheduling constraint for the Simplified Case).

*of Claim 2.* Assume, for contradiction, that there exists a well-behaved adversary $A$, a distinguisher $D$ and a polynomial $p$, such that, for infinitely many $n$, $v$, and $z$, $D$ distinguishes $\mathsf{hyb}_0(v, z)$ and $\mathsf{hyb}_1(v, z)$ with probability $\frac{1}{p(n)}$. We show how this violates robustness of $\langle \tilde{C}, \tilde{R} \rangle$.

Towards this goal, first note that the two experiments $H_0$ and $H_1$ proceed identically before Stage 2 commitment of the left interaction is sent. Therefore, there must exist a partial joint view $\tau$ of all parties that defines the execution before Stage 2 of the left interaction, such that $D$ distinguishes $\{\mathsf{hyb}_0(v, z) \| \tau\}$ and $\{\mathsf{hyb}_1(v, z) \| \tau\}$ with probability at least $\frac{1}{p(n)}$, where $\{\mathsf{hyb}_i(v, z) \| \tau\}$ denotes the outcome of $\mathsf{hyb}_i(v, z)$ conditioned on that the event that the execution is consistent with joint view $\tau$; let $s$ be denote the first left message in $\tau$ and let $r$ denote a pre-image of $s$ through $f$. Consider the machine $B(b^n)$, which upon receiving a statment $x$, and two witnesses for $x$, $w_0, w_1$, provides a $\mathcal{WISSP}$ proof of the statement $x$ using witness $w_b$. Since $\mathcal{WISSP}$ is a 4-round protocol (where the first message is sent from the verifier), $B$ thus also defines a 4-round protocol (as $x, w_0, w_1$ can be sent with the first message from the $\mathcal{WISSP}$ verifier). It follows directly from the $\mathcal{WI}$ property of the $\mathcal{WISSP}$ proof that no (non-uniform) $\mathcal{PPT}$ adversary can distinguish interactions with $B(0^n)$ and $B(1^n)$.

We now construct an adversary $\tilde{A}$ such that the view and values that $\tilde{A}$ commitment to after interacting with $B(0^n)$ and $B(1^n)$ can be distinguished by a distinguisher $\tilde{D}$ (that appropriately incorporates $D$). $\tilde{A}$, upon receiving $v$, $z$, $\tau$ and $r$ as auxiliary input, internally emulates a man-in-the-middle execution with $A$ from $\tau$ as follows. It honestly emulates the left committer and right receivers for $A$ with the following two exceptions:

- To emulate the Stage 5 proof of the left interaction, it externally sends $B$ the statement $x$ and the witnesses $w_0, w_1$—$w_0$, is a valid decommitment to the Stage 4 commitment protocol (since $\tilde{A}$ honestly emulated Stage 4, it knows this decommitment information), $w_1$ instead is the "fake witness" $r$. It next forwards the $\mathcal{WISSP}$ proof from $B$ to $A$.
- In Stage 3 of the right interactions, it externally forwards messages from $A$ to an honest receiver of $\langle \tilde{C}, \tilde{R} \rangle$. (Note that since $A$ is well-behaved the Stage 3 commitment has not yet stated in $\tau$ so this can be done.)

The distinguisher $\tilde{D}$, on input the view $\mathcal{V}$ and the values $v_1, \ldots, v_m$ committed to by $\tilde{A}$, reconstructs the view and values committed to by $A$ in Stage 3 in emulation by $\tilde{A}$, by extracting the view $\mathcal{V}_A$ of $A$ from $\mathcal{V}$, and setting the values $\tilde{v}'_k$ committed to by $A$ to $v_k$ if the right interaction $k$ is accepting and has a different identity from the left interaction in $\mathcal{V}_A$ and $\perp$ otherwise. $\tilde{D}$ then invokes the distinguisher $D$ on the reconstructed view $\mathcal{V}_A$ and committed values $\tilde{v}'_k$, and outputs the output of $D$. Since $\tilde{A}$, in interaction with $B(b^n)$, perfectly emulates the view of $A$ in the hybrid experiment $H_b$, the extracted view $\mathcal{V}_A$ and committed values $\tilde{v}'_k$ are identically distributed to $\{\mathsf{hyb}_b(v,z) \| \tau\}$. It follows that $\tilde{D}$ distinguishes the view and the values committed to by $\tilde{A}$ using $\langle \tilde{C}, \tilde{R} \rangle$ with probability $\frac{1}{p(n)}$, which contradicts with the robustness of $\langle \tilde{C}, \tilde{R} \rangle$. □

*of Claim 4.* Below, we say that the man-in-the-middle adversary $\mathsf{cheats}$, if it manages to commit to a "fake" witness in any successful right interaction that uses a different identity from the left interaction. Now, assume for contradiction that there exists a well-behaved adversary $A$, and a polynomial $p$, such that, for infinitely many $n$, $v$ and $z$, the probability that $A$ $\mathsf{cheats}$ in $H_2$ is $1/p(n)$. We show that there exists a $\mathcal{PPT}$ adversary $B$ that violates the weak one-many non-malleability of $\langle \tilde{C}, \tilde{R} \rangle$.

Take any such $n$, $v$ and $z$. Note that (since $A$ is well-behaved) the hybrids $H_1$ and $H_2$ proceed identically before the Stage 2 commitment of the left interaction is sent; furthermore, by Claim 3, except from negligible probability, $A$ never $\mathsf{cheats}$ in $H_1$. It follows using an averaging argument (just as in analysis of the General Case in Section 4.2), there exists a partial joint view $\rho$ of all parties that defines the execution before Stage 2 of the left interaction, such that, conditioned on that the execution being consistent with $\rho$, the probability that $A$ $\mathsf{cheats}$ in $H_2$ is at least $1/2p(n)$, whereas the probability in $H_1$ is negligible; let $r$ be a pre-image of the first left-message in $\rho$.

Then, the adversary $B$, upon receiving $\rho$ and $r$ as auxiliary inputs, internally emulates a man-in-the-middle execution with $A$ from $\rho$ as follows. It emulates the left committer and the right receivers honestly (from the joint view $\rho$), except that it externally forwards messages from $A$ in Stage 4 of the left and right interactions to an honest committer (on the left), and honest receivers (on the right) of the commitment $\langle \tilde{C}, \tilde{R} \rangle$. By construction, $B$ perfectly emulates the view of $A$ in $H_1$ whenever it it receives a commitment to $0^n$ from the external committer, and perfectly emulates the view of $A$ in $H_2$ whenever it it receives a commitment to the "fake" witness $r$. By our assumption, it follows that the probability that $A$ cheats in the emulation by $B$ changes from being negligible (when receiving a commitment to $0^n$ to (at least) $1/p(n)$ (when receiving a commitment to the fake witness). Now, consider the local property $Q$ defined as follows: $Q$, on input a view $\mathcal{V}$ of $B$ and one of its committed values $\tilde{v}$, outputs 1, if and only if, (1) $\tilde{v}$ is a "fake" witness of some right interaction $i$ in the emulated man-in-the-middle execution in $\mathcal{V}$, and (2) the interaction is successful and uses a different identity from the left interaction. The probability that $\mathsf{exists}_Q$ holds on the view and values committed to by $B$ is identical to the probability that $A$ $\mathsf{cheats}$ in the emulation by $B$. But, as noted above, this probability changes from being negligible to $\frac{1}{p(n)}$ when changing the external left commitment from $0^n$ to $r$. This contradicts with the weak one-many non-malleability of $\langle \tilde{C}, \tilde{R} \rangle$. $\square$

*of Claim 5.* Assume, for contradiction, that there exists a well-behaved adversary $A$, a distinguisher $D$, and a polynomial function $p(\cdot)$, such that, for infinitely many $n \in N$, $v \in \{0,1\}^n$ and $z \in \{0,1\}^*$, the distinguisher $D$ distinguishes $\mathsf{hyb}_1(v,z)$ and $\mathsf{hyb}_2(v,z)$ with probability $\frac{1}{p(n)}$. We construct an expected $\mathcal{PPT}$ adversary $B$ that breaks the hiding property of $\langle \tilde{C}, \tilde{R} \rangle$. Below we are only concerned with the last three messages of a $\mathcal{WISSP}$ proof and, thus, for simplicity of notation, we view them as 3-round protocols $(\alpha, \beta, \gamma)$.

Consider any $n$, $v$ and $z$ for which the above happens. Notice that $H_1$ and $H_2$ proceed identically before the Stage 2 commitment of the left interaction is sent; furthermore, by Claim 3 and 4, it holds that, except from negligible probability, $A$ never $\mathsf{cheats}$ in $H_1$ and $H_2$. It again follows using an averaging argument (just as in the analysis of the General Case in Section 4.2) that, there exists a partial joint view $\rho$ of all parties that defines the execution before Stage 2 of the left interaction, such that, $D$ distinguishes $\{\mathsf{hyb}_1(v,z) \| \tau\}$ and $\{\mathsf{hyb}_2(v,z) \| \tau\}$ with probability $1/2p(n)$, and conditioned on the execution being consistent with $\rho$, the probability that $A$ $\mathsf{cheats}$ is negligible; let $r$ be a pre-image of the first left-message in $\rho$.

On a high-level, $B$, on receiving $\tau$ and $r$ as auxiliary input, externally acts as a receiver of a $\langle \tilde{C}, \tilde{R} \rangle$ commitment, and tries to distinguish commitments to 0 and $r$. Internally, $B$ incorporates $A$ and proceeds in the following three phases.

1. In the first phase, $B$ emulates a man-in-the-middle execution with $A$ from $\rho$, by emulating the left committer and the right receivers honestly, from $\rho$, except that it forwards messages from $A$ in Stage 4 of the left interaction externally. We call this phase the *"Main Execution Phase"* and denote $\Delta$ the transcript of messages of this phase.

2. In the second phase, $B$ attempts to extract the values $A$ commits to on the right, via rewinding the special sound proofs in Stage 6 of the protocol. More precisely, for each right interaction, $B$ finds the first $\mathcal{WISSP}$ proof $(\alpha, \beta, \gamma)$ in $\Delta$, such that, during its the execution, no messages belonging to Stage 4 of the left interaction are exchanged. (Such a $\mathcal{WISSP}$ proof must exist since there are more $\mathcal{WISSP}$ proofs than the number of rounds of Stage 4.) $B$ then rewinds the proof by sending new random challenges $\beta'$ until a second transcript $(\alpha, \beta', \gamma')$ is obtained. In each rewinding, $B$ emulates the left and right interactions for $A$ exactly the same as in the Main Execution Phase, except that it cancels every rewinding in which $A$ expects a new message in Stage 4 of the left interaction. Finally, it computes the witness of the proof if $\beta'$ differs from $\beta$; it aborts and outputs $\mathsf{fail}$ if $\beta' = \beta$ or the extracted witness is not a valid decommitment to Stage 2. Additionally, $B$ cuts off its execution after $2^n$ steps; it halts and output $\mathsf{fail}$ in this case. This phase is called the *"Extraction Phase"*.

3. Finally, in the *"Output Phase"*, $B$ outputs the view of $A$ and all the values extracted, with the following exception: a value $v_i$ is replaced by $\perp$, if in the right interaction, $A$ fails, or uses the same identity as the left interaction.

A formal description of the procedure of $B$ appears in figure 6. We have the following two claims.

**Subclaim 1.** *$B$ runs in expected polynomial time.*

**Subclaim 2.** *The output of $B$ satisfies the following:*

**Description of $B$**
**Input:** $B$ receives auxiliary input $\tau$, $z$ and $r$
**Procedure:** $B$ interacts externally as a receiver using $\langle \tilde{C}, \tilde{R} \rangle$. Internally it incorporates $A(z)$ and proceeds in the following three phases.

**Main Execution Phase** Emulates a one-many man-in-the-middle execution of $\langle \hat{C}, \hat{R} \rangle$.

- Feed the view in $\tau$ to $A$ and all right receivers.
- Forward the external $\langle \tilde{C}, \tilde{R} \rangle$ commitment in as Stage 4 of the left interaction.
- Emulate the rest of all interactions and complete the execution with $A$ as in hybrid $H_1$.

Let $\Delta$ be the transcript of messages obtained.

**Extraction Phase** For $k \in [m]$, if interaction $k$ is convincing, and its identity is different from the left interaction, do:

- In $\Delta$, find the first $\mathcal{WISSP}$ proof $(\alpha, \beta, \gamma)$ in Stage 6, such that no left Stage 4 messages are exchanged during its execution.
- Repeat until a second proof transcript $(\alpha, \beta', \gamma')$ is obtained:

  Emulate all interactions as in the Main-Execution Phase, except that for the left interaction, if $A$ expects a message in Stage 4 of the left interaction, cancel the execution, rewinds to the beginning of $\beta$ and continue.

  Note that, since (by the scheduling constraint in the simplified case) interactions $k$ has its Stage 6 occurring after $\tau$, none of the rewinding can make $A$ request a message appearing in $\tau$ again.
- If $\beta \neq \beta'$, extract witness $w$ from $(\alpha, \beta, \gamma)$ and $(\alpha, \beta', \gamma')$. Otherwise halt and output fail.
- If $w$ contains is a valid decommitment $(v, r'')$ for the Stage 3 commitment of $\langle \hat{C}, \hat{R} \rangle$ in interaction $k$, then set $\hat{v}_k = v$. Otherwise halt and output fail.

**Output Phase** For every interaction $k$ that is not convincing, or whose identity is the same as the left interaction, set $\hat{v}_k = \perp$. Output $(\hat{v}_1, \ldots, \hat{v}_m)$ and the view of $A$ from the Main Execution Phase.

Finally, if $B$ runs for more than $2^n$ steps, halt and output fail.

Figure 6: The construction of $B$.

- *If $B$ receives a commitment to $0$, the output is statistically close to the outcome of* $\{\mathsf{hyb}_1(v,z)\|\tau\}$.
- *If $B$ receives a commitment to $r$, the output is statistically close to the outcome of* $\{\mathsf{hyb}_2(v,z)\|\tau\}$.

Then, by the assumption that, $D$ distinguishes $\{\mathsf{hyb}_1(v,z)\|\tau\}$ and $\{\mathsf{hyb}_2(v,z)\|\tau\}$ with probability at least $1/2p(n)$, $D$ also distinguishes the outputs of $B$ with probability at least $1/3p(n)$. Since this holds for infinitely many $n$, $B$ violates the hiding property of $\langle \tilde{C}, \tilde{R}\rangle$.

*of Subclaim 1.* We show that $B$ runs in expected polynomial time. This follows using essentially the same proof as in [LPV08]; for completeness, we provide the analysis below.

Note that the time spent by $B$ in the Main Execution Phase is poly$(n)$, since $A$ is a strict polynomial time machine. We show below that the expected time spent by $B$ in the Extraction Phase is poly$(n)$. To bound the expected running time, we assume for simplicity that $B$ does not check the $\mathsf{fail}$ conditions and may run for more than $2^n$ steps (since this only increases the running time).

Recall that in the Extraction Phase, $B$ rewinds $A$ from the first $\mathcal{SSP}$ proof that is not interleaved with the Stage 4 of the left interaction. Let $T_k(i)$ be the random variable that describes the time spent rewinding the $i^{\text{th}}$ $\mathcal{SSP}$ proof in Stage 6 of interaction $k$. We show that $E[T_k(i)] \leq \text{poly}(n)$ and then, by linearity of expectation, conclude that the expected time spent by $B$ in the Extraction phase is

$$\sum_k \sum_i E[T_k(i)] \leq \sum_k \sum_i \text{poly}(n) \leq \text{poly}(n),$$

We proceed to bound $E[T_k(i)]$. Let $(\alpha, \beta, \gamma)$ be the $i^{\text{th}}$ $\mathcal{SSP}$ proof in Stage 6 of the $k$th right interaction; and $\Gamma$ the set of all (partial) transcript $\rho$, such that, after $\rho$, the challenge $\beta$ of the $i$'th slot of the $k$'th interaction is immediately sent. Given one $\rho \in \Gamma$, let $\Pr[\rho]$ denote the probability that $\rho$ occurs as a prefix of the execution emulated in the Main Execution phase. Furthermore, let $p_\rho$ denote the probability that $\rho$ is rewound in $B$, i.e. $p_\rho$ is the probability that, conditioned on the prefix $\rho$ occurring, the right interaction $k$ is convincing, and $(\alpha, \beta, \gamma)$ is the first $\mathcal{SSP}$ proof in the interaction such that no left Stage 4 message is exchanged during its execution. Recall that $B$ rewinds until it finds another transcript of the proof $(\alpha, \beta, \gamma)$; it cancels every rewinding for which $A$ fails, or requests a left Stage 4 message. We claim that the probability of cancelling a rewinding from $\rho$ is at most $1 - p_\rho$, since in every rewinding that is cancelled, either $A$ fails or a left Stage 4 message is about to be exchanged, and conditioned on $\rho$, the probability of a view occurring in a rewinding from $\rho$ is same as occurring in the Main Execution phase (as $B$ emulates the left committer and right receivers in a rewinding the same as it does in the Main Execution phase). Thus, the expected number of rewindings is at most $\frac{1}{p_\rho}$. Therefore, the expected number of rewindings from $\rho$ is at most $p_\rho \cdot \frac{1}{p_\rho} = 1$ and each rewinding takes at most poly$(n)$ steps, i.e.

$$E[T_k(i)|\rho] \leq \text{poly}(n)$$

Thus,

$$\begin{aligned} E[T_k(i)] &= \sum_{\rho \in \Gamma} E[T_k(i)|\rho] \times \Pr[\rho] \\ &\leq \text{ poly}(n) \times \sum_{\rho\in\Gamma} \Pr[\rho] \leq \text{ poly}(n) \end{aligned}$$

□

*of Subclaim 2.* We now argue that the output distribution of $B$ is correct. This follows from the following two claims.

1. Upon receiving a commitment to 0, the output of $B$—conditioned on not outputting $\mathsf{fail}$—is statistically close to the outcome of $\{\mathsf{hyb}_1(v,z)\|\tau\}$. Furthermore, upon receiving a commitment to $r$, the output of $B$—conditioned on not outputting $\mathsf{fail}$—is statistically close to the outcome of $\{\mathsf{hyb}_2(v,z)\|\tau\}$.

2. $B$ outputs $\mathsf{fail}$ with only negligible probability.

For the first of these claims, note that $B$ perfectly emulates the execution of $H_1$ from $\tau$ (upon receiving a commitment to 0) or $H_2$ (upon receiving a commitment to $r$) for $A$. Furthermore, for every convincing right interaction $k$ that has a different identity than the left interaction, $B$ finds the first $\mathcal{SSP}$ proof that does not interleave with Stage 4 of the left interaction, (as previously mentioned, such a proof must exist, since there are more $\mathcal{SSP}$ proofs than messages in Stage 4) and will rewind that interaction, and eventually output $\mathsf{fail}$ or a valid decommitment of Stage 3. Conditioned on the event that $B$ does not output $\mathsf{fail}$, by the statistical-binding property of $\langle\tilde{C},\tilde{R}\rangle$, it follows that, except with negligible probability, the witnesses extracted by $B$ are the values committed to by $A$ in Stage 3.

For the second claim, recall that $B$ outputs $\mathsf{fail}$ only in the following cases:

**$B$ runs for more than $2^n$ steps:** We know that the expected running time of $B$ is $\mathit{poly}(n)$. Using Markov inequality, we conclude that the probability that $B$ runs more than $2^n$ steps is at most $\frac{\mathit{poly}(n)}{2^n}$.

**The same proof transcript is obtained:** This case occurs if $B$ picks some challenge $\beta$ in the Extraction Phase that appeared as a challenge in the Main Execution Phase. As $B$ runs for at most $2^n$ steps, it picks at most $2^n$ challenges. Furthermore, the length of each challenge is $2n$. By applying the union bound, we obtain that the probability that a $\beta$ is picked twice is at most $\frac{2^n}{2^{2n}}$. Since there are at most polynomially many challenges picked in the Main Execution Phase, using the union bound again, we conclude that the probability of obtaining the same transcript is negligible.

**The witness extracted is not a valid decommitment:** Suppose, for some successful interaction $k$ that has a different identity from the left interaction, the witness extracted is not the decommitment information. Then, by the special-soundness, it follows that it must be a "fake" witness $r$ for the interaction, and that $A$ must have committed to $r$ in Stage 4 of the interaction. However, this contradicts with the fact that conditioned on $\tau$, $A$ can only commit to a fake witness with negligible probability in such a right interaction.

□

□

### 4.3 Improving the Round Complexity

**Editorial Remark:** The improvements in this section were added after Hoeteck Wee announced a (different) construction of a $O(\log^* n)$-round non-malleable commitment scheme from any one-way function [Wee10]. Our original amplification procedure in [LP09] did not use these improvements and resulted in a $O(1)^{\log^* n}$ round non-malleable commitment.

The strengthening technique presented in Section 4 transforms a $k(n)$-round robust stand-alone non-malleable commitment $\langle \tilde{C}, \tilde{R} \rangle$ into a robust concurrent non-malleable commitment $\langle \hat{C}, \hat{R} \rangle$, with $6k(n) + 10$ rounds. Iterating this procedure would yield a $O(1)^{log^* n}$-round full-fledged non-malleable commitment from one-way function. But, by simply running the last four stages of $\langle \hat{C}, \hat{R} \rangle$ in "parallel", we can get a construction which only increases the round complexity by an additive constant. Let $\langle \tilde{C}, \tilde{R} \rangle$ be a robust stand-alone non-malleable commitment scheme with initial binding; without loss of generality, we assume that the protocol has an even number of rounds, with $k/2$ messages from each of the committer and the receiver, and the first message is from the receiver. Next, replace the last four stages in the construction of $\langle \hat{C}, \hat{R} \rangle$ with the following two stages (i.e., we keep the first two stages untouched).

**Stage 3** the Committer and the Receiver exchange the first two messages of a commitment to $v$ and a commitment to $0^n$, from the Committer to the Receiver, using $\langle \tilde{C}, \tilde{R} \rangle$ and $\mathsf{id}$, *in parallel.* Let $(m_1, m_2)$ and $(m'_1, m'_2)$ be the transcript generated. By the initial binding property of $\langle \tilde{C}, \tilde{R} \rangle$, $m_2$ and $m'_2$ determines the values committed to in the two commitments.

**Stage 4** the Committer and the Receiver perform the following in parallel:

- exchange the remaining $k(n) - 2$ messages of the commitment to $v$ using $\langle \tilde{C}, \tilde{R} \rangle$ and $\mathsf{id}$.
- exchange the remaining $k(n) - 2$ messages of the commitment to $0^n$ using $\langle \tilde{C}, \tilde{R} \rangle$ and $\mathsf{id}$.
- the Committer provides a 4-round $\mathcal{WISSP}$ proof $\langle P, V \rangle$ of the statement that the committed value decided by $(m'_1, m'_2)$ is $0^n$, *or* it knows a pre-image of $s$.
- the Committer proves that it has committed to value $v$ in both Stage 2 and $(m_1, m_2)$, or a pre-image of $s$ in $(m'_1, m'_2)$. This is proved using $l(n) = \frac{k(n)}{2} + 1$ invocations of a 4-round $\mathcal{WISSP}$ protocol $\langle P, V \rangle$. More specifically, let $(\delta, \alpha, \beta, \gamma)$ denote a transcript of $\langle P, V \rangle$; the Committer and the Receiver first exchange the first two messages $(\delta, \alpha)$ of the $l(n)$ arguments of $\langle P, V \rangle$, *in parallel*; next, they exchange the last two messages, $(\beta, \gamma)$, of the $l(n)$ arguments, *in sequence.*

It is easy to see that the protocol $\langle \hat{C}, \hat{R} \rangle$ has at most $k(n) + 8$ rounds. Furthermore, it follows exactly as in Section 4.2 that $\langle \hat{C}, \hat{R} \rangle$ is a robust concurrent non-malleable commitment (note that the proof in Section 4.2 never relied on the fact that the sub-protocols in $\langle \hat{C}, \hat{R} \rangle$ after Stage 2 were sequentially run) .

# 5 Concluding Theorem 1

**Corollary 1 (Restated).** *Let $\langle C, R \rangle$ be a $k(n)$-round robust stand-alone $t(n)$-non malleable commitment scheme, s.t. $3 \le t(n) \le n$. Then, there exists a $k(n) + O(l(n))$-round robust non-malleable commitment scheme $\langle \hat{C}, \hat{R} \rangle$, where $l(n) = O(\log^* n - \log^* t(n))$.*

*Proof.* Define $\langle C^j, R^j \rangle$ to be the commitment scheme obtained after applying Theorem 3 $j$ times on $\langle C, R \rangle$. Let $\mathsf{round}^j(n)$ and $\mathsf{complex}^j(n)$ be the round and computational complexity of $\langle C^j, R^j \rangle$, and $\mathsf{id}^j(n)$ be the maximal length of the identities that $\langle C^j, R^j \rangle$ can accommodate. By assumption, we have that $\mathsf{round}^0(n) = k(n)$, $\mathsf{complex}^0(n) = poly(n)$ and $\mathsf{id}^0(n) = t(n)$. Furthermore, by Theorem 3, it holds that, for any $j$

1. $\mathsf{round}^j(n) \le \mathsf{round}^{j-1}(n) + 8$.
2. $\mathsf{id}^j(n) = 2^{\mathsf{id}^{j-1}(n) - 1}$.

3. $\mathsf{complex}^j(n) = 2\mathsf{id}^j(n)\mathsf{complex}^{j-1}(n) + \mathsf{round}^{j-1}(n)p(n)$

where $p(n) = \mathrm{poly}(n)$. Let $l(n) = max(\{j : \mathsf{id}^j(n) \leq n\})$. By the second condition above, we have that $\mathsf{id}^{l(n)}(n) \geq \log n + 1$ and $\mathsf{id}^{l(n)-1}(n) \leq \log n + 1$. Furthermore, since $\mathsf{id}^{j+2}(n) > 2^{\mathsf{id}^j(n)}$, we have that $l(n) = O(\log^* n - \log^* t(n))$. It follows that,

- $\log n + 1 \leq \mathsf{id}^{l(n)} \leq n$.
- $\mathsf{round}^{l(n)}(n) \leq \mathsf{round}^0(n) + 8l(n) \leq k(n) + 8l(n)$

Next we show that $\mathsf{complex}^{l(n)}(n)$ is also polynomially bounded. As $\mathsf{round}^j \leq \mathsf{round}^{j+1}$, we have that for every $j \leq l(n)$, $\mathsf{round}^{j+1}(n)$ is bounded by a polynomial $q$. It follows that

$$\mathsf{complex}^j(n) = 2\mathsf{id}^j\mathsf{complex}^{j-1}(n) + p(n)q(n)$$

Expand the equation for $\mathsf{complex}^{l(n)}(n)$.

$$\mathsf{complex}^{l(n)}(n) \quad = \quad 2^{l(n)}\left(\prod_{j=1}^{l(n)} \mathsf{id}^j(n)\right)\mathsf{complex}^0(n) \; + \; \sum_{l'=2}^{l(n)} 2^{l(n)-l'+1}\left(\left(\prod_{j=l'}^{l(n)} \mathsf{id}^j(n)\right)p(n)q(n)\right)$$

Since the second term of the right hand side is polynomially related to the the first term, it suffices to bound the first term.

$$\begin{aligned}
&2^{l(n)}\left(\prod_{j=1}^{l(n)} \mathsf{id}^j(n)\right)\mathsf{complex}^0(n) = 2^{l(n)}poly(n)\prod_{j=1}^{l(n)} \mathsf{id}^j(n)\\
\leq& poly(n)\mathsf{id}^{l(n)}\left(\mathsf{id}^{l(n)-1}(n)\right)^{l(n)}\\
\leq& poly(n)n(\log n + 1)^{O(\log^* n - \log^* t(n))}\\
\leq& poly(n)n(\log n + 1)^{\log\log n} = poly(n)
\end{aligned}$$

Therefore $\langle C^{l(n)}, R^{l(n)}\rangle$ is a robust $\mathsf{id}^{l(n)}$-non malleable commitment scheme, and $\mathsf{id}^{l(n)} \geq \log n + 1$. To obtain a robust $n$-non malleable commitment scheme $\langle \hat{C}, \hat{R}\rangle$, we simply view $\langle C^{l(n)}, R^{l(n)}\rangle$ as a $(\log n + 1)$-non malleable commitment, and apply the amplification theorem once. □

By setting $t(n)$ to 3 in Corollary 1, we get:

**Lemma 3.** *Let $\langle C, R\rangle$ be a constant-round robust 3-non malleable commitment scheme with computational complexity $poly(n)$. Then, there exists a $O(\log^* n)$-round robust non-malleable commitment scheme $\langle C', R'\rangle$.*

## 6 Acknowledgements

We are very grateful to Hoeteck Wee for many helpful discussions. As mentioned, we only realized that we could parallelize the execution of our sub-protocols, yielding an amplification procedure with a constant *additive*, instead of *multiplicative* overhead, after Hoeteck told us about obtaining such a result (using a different protocol). We are also grateful to Oded Goldreich for a very helpful conversation, and, in particular, for encouraging us to include the simple construction of a constant-round $O(1)$-non-malleable protocol, and to highlight the importance of the notion of robust non-malleability. Finally, we are grateful to Cynthia Dwork, Danny Dolev, Moni Naor, Alon Rosen and Muthuramakrishan Venkitasubramaniam for discussions about non malleability over the years.

# References

[Bar01] Boaz Barak. How to go beyond the black-box simulation barrier. In *FOCS '01*, volume 0, pages 106–115, 2001.

[Bar02] Boaz Barak. Constant-round coin-tossing with a man in the middle or realizing the shared random string model. In *FOCS '02: Proceedings of the 43rd Symposium on Foundations of Computer Science*, pages 345–355, Washington, DC, USA, 2002. IEEE Computer Society.

[BCC88] Gilles Brassard, David Chaum, and Claude Crépeau. Minimum disclosure proofs of knowledge. *J. Comput. Syst. Sci.*, 37(2):156–189, 1988.

[Blu83] Manuel Blum. Coin flipping by telephone a protocol for solving impossible problems. *SIGACT News*, 15(1):23–27, 1983.

[Can01] Ran Canetti. Universally composable security: A new paradigm for cryptographic protocols. In *FOCS '01: Proceedings of the 42nd IEEE symposium on Foundations of Computer Science*, page 136, Washington, DC, USA, 2001. IEEE Computer Society.

[Dam00] Ivan Damgård. Efficient concurrent zero-knowledge in the auxiliary string model. In *EUROCRYPT '00*, pages 418–430, 2000.

[DDN00] Danny Dolev, Cynthia Dwork, and Moni Naor. Nonmalleable cryptography. *SIAM Journal on Computing*, 30(2):391–437, 2000.

[Din07] Irit Dinur. The pcp theorem by gap amplification. *J. ACM*, 54(3):12, 2007.

[DPP94] Ivan Damgård, Torben P. Pedersen, and Birgit Pfitzmann. On the existence of statistically hiding bit commitment schemes and fail-stop signatures. In *CRYPTO '93: Proceedings of the 13th Annual International Cryptology Conference on Advances in Cryptology*, pages 250–265, London, UK, 1994. Springer-Verlag.

[FS90] Uriel Feige and Adi Shamir. Witness indistinguishable and witness hiding protocols. In *STOC '90*, pages 416–426, 1990.

[GM84] Shafi Goldwasser and Silvio Micali. Probabilistic encryption. *J. Comput. Syst. Sci.*, 28(2):270–299, 1984.

[GMR89] Shafi Goldwasser, Silvio Micali, and Charles Rackoff. The knowledge complexity of interactive proof systems. *SIAM Journal on Computing*, 18(1):186–208, 1989.

[GMW87] Oded Goldreich, Silvio Micali, and Avi Wigderson. How to play any mental game. In *STOC '87: Proceedings of the nineteenth annual ACM symposium on Theory of computing*, pages 218–229, New York, NY, USA, 1987. ACM.

[GMW91] Oded Goldreich, Silvio Micali, and Avi Wigderson. Proofs that yield nothing but their validity or all languages in NP have zero-knowledge proof systems. *J. ACM*, 38(3):690–728, 1991.

[Gol01] Oded Goldreich. *Foundations of Cryptography — Basic Tools*. Cambridge University Press, 2001.

[Gol05] Oded Goldreich. Bravely, moderately: A common theme in four recent results. *Electronic Colloquium on Computational Complexity (ECCC)*, (098), 2005.

[HILL99] Johan Håstad, Russell Impagliazzo, Leonid Levin, and Michael Luby. A pseudorandom generator from any one-way function. *SIAM Journal on Computing*, 28:12–24, 1999.

[KOS03] Jonathan Katz, Rafail Ostrovsky, and Adam Smith. Round efficiency of multi-party computation with a dishonest majority. In *EUROCRYPT '03*, pages 578–595, 2003.

[LP09] Huijia Lin and Rafael Pass. Non-malleability amplification. In *STOC '09*, pages 189–198, 2009.

[LPTV10] Huijia Lin, Rafael Pass, Wei-Lung Dustin Tseng, and Muthuramakrishnan Venkitasubramaniam. Concurrent non-malleable zero knowledge proofs. In *CRYPTO*, pages 429–446, 2010.

[LPV08] Huijia Lin, Rafael Pass, and Muthuramakrishnan Venkitasubramaniam. Concurrent non-malleable commitments from any one-way function. In *TCC '08*, pages 571–588, 2008.

[LPV09] Huijia Lin, Rafael Pass, and Muthuramakrishnan Venkitasubramaniam. A unified framework for concurrent security: universal composability from stand-alone non-malleability. In *STOC '09*, pages 179–188, 2009.

[Nao91] Moni Naor. Bit commitment using pseudorandomness. *Journal of Cryptology*, 4:151–158, 1991.

[Pas04] Rafael Pass. Bounded-concurrent secure multi-party computation with a dishonest majority. In *STOC '04: Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*, pages 232–241, New York, NY, USA, 2004. ACM.

[PPV08] Omkant Pandey, Rafael Pass, and Vinod Vaikuntanathan. Adaptive one-way functions and applications. In *CRYPTO 2008: Proceedings of the 28th Annual conference on Cryptology*, pages 57–74, Berlin, Heidelberg, 2008. Springer-Verlag.

[PR05a] Rafael Pass and Alon Rosen. Concurrent non-malleable commitments. In *FOCS*, pages 563–572, 2005.

[PR05b] Rafael Pass and Alon Rosen. New and improved constructions of non-malleable cryptographic protocols. In *STOC '05*, pages 533–542, 2005.

[PW10] Rafael Pass and Hoeteck Wee. Constant-round non-malleable commitment from strong one-way functions. In *Eurocrypt '10*, 2010.

[Rei05] Omer Reingold. Undirected st-connectivity in log-space. In *STOC '05: Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, pages 376–385, New York, NY, USA, 2005. ACM.

[Wee10] Hoeteck Wee. Black-box, round-efficient secure computation via non-malleability amplification. In *FOCS*, 2010.

**Protocol $\langle \overline{C}, \overline{R} \rangle$**

**Common Input:** An identity $\mathsf{id} \in \{0,1\}^t$ and $1^n$.

**Auxiliary Input for Committer:** A string $v \in \{0,1\}^n$.

**Stage 1:**

C uniformly chooses $v_1 \in \{0,1\}^n$, $\sigma_1, \sigma_2 \in \{0,1\}^{poly(n)}$; and sets $v_2 = v \oplus v_1$

C → R: $c_1 = \mathsf{com}(v_1, \sigma_1)$ and $c_2 = \mathsf{com}(v_2, \sigma_2)$.

**Stage 2:**

C → R: $\mathsf{id} + 6$ $\mathcal{WISSP}$ proofs of the statement that there exist values $v_1$ and $\sigma_1$ s.t. $c_1 = \mathsf{com}(v_1, \sigma_1)$.

**Stage 3:**

C → R: $2^t - \mathsf{id} + 6$ $\mathcal{WISSP}$ proofs of the statement that there exist values $v_2$ and $\sigma_2$ s.t. $c_2 = \mathsf{com}(v_2, \sigma_2)$.

Figure 7: A $O(2^t)$-Round Robust $t$-Non-Malleable Commitment Scheme $\langle \hat{C}, \hat{R} \rangle$

# A A Simple Exponential-round Robust Non-Malleable Commitment

In this section we present a $O(2^t)$-round $t$-non-malleable robust commitment. The commitment scheme $\langle \overline{C}, \overline{R} \rangle$ proceeds as follows: to commit to a value $v$, the Committer and the Receiver, on common input a $t$-bit identity $\mathsf{id} \in \{0,1\}^t$, and $1^n$, where $n$ is the security parameter, proceed in 3 stages:

**Stage 1** the Committer chooses 2 random shares $v_1, v_2 \in \{0,1\}^n$, such that $v = v_1 \oplus v_2$, and sends $c_1 = \mathsf{com}(v_1)$ and $c_2 = \mathsf{com}(v_2)$, where $\mathsf{com}$ is any two-round statistically-binding string commitment.

**Stage 2** the Committer proves that $c_1$ is a valid commitment to value $v_1$. This is proved using $\mathsf{id} + 5$ sequential invocations of a 4-round $\mathcal{WISSP}$ protocol $\langle P, V \rangle$.

**Stage 3** the Committer proves that $c_2$ is a valid commitment to value $v_2$. This is proved using $2^t - \mathsf{id} + 5$ sequential invocations of the 4-round $\mathcal{WISSP}$ protocol $\langle P, V \rangle$.

A formal description of the protocol can be found in Figure 7.

Below we show that the protocol is robust and t-non-malleable. The robustness essentially follows from the fact that both random shares $v_1$ and $v_2$ are extractable (from the $\mathcal{WISSP}$ proofs in Stage 2 and 3 respectively) and have at least 5 "rewinding slots" (a rewinding slot consists of the challenge and reply pair of a $\mathcal{WISSP}$ proof). Then, since the committed value $v$ is completely decided by $v_1$ and $v_2$, we conclude that the protocol is non-malleable w.r.t. any 4-round protocols. A formal proof proceeds just as in Claim 5.

We proceed to show that the protocol is $t$-non malleable. Note that the commitment scheme would still be hiding if we were to reveal one of the shares committed to (since each of the shares is individually random). To prove non-malleability, we show that both the shares commited in the right interaction can be extracted out (by rewinding), while only rewinding one of the Stage 2 and

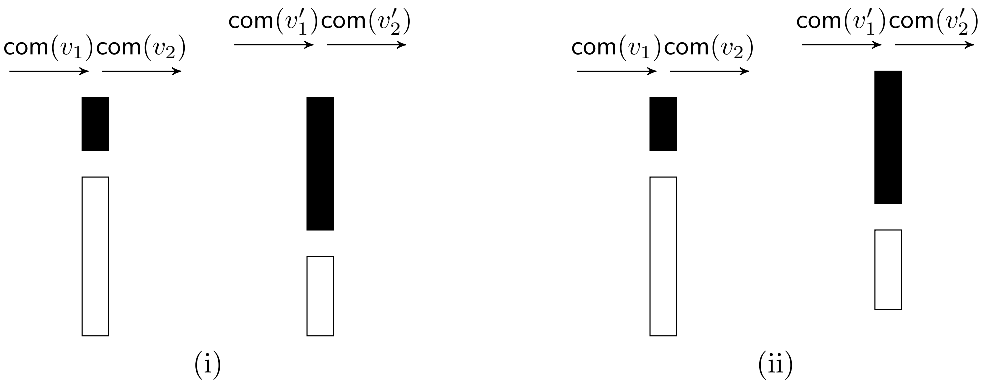


Figure 8: Two schedulings of a man-in-the-middle execution of $\langle \overline{C}, \overline{R} \rangle$.

3 of the left interaction. But since each stage proves knowledge only of one of the shares on the left, such a rewinding can reveal only one of the share and thus the left interaction still remains hiding and non-malleability follows. To show this, we consider two case: either the man-in-the-middle attacker $A$ "aligns" the left and right interactions (that is, it sends the challenge of the $i$'th left-proof, in between the challenge and reply pair of the $i$'th right-proof), or it does not (see Figure 8 (i) and (ii) respectively). In case $A$ aligns (and the left and right interactions have different identities), there must be one stage in the left interaction that contains more proofs than its counterpart in the right interaction; let it be Stage $a$ (for instance, $a = 3$ in Figure 8 (i)). This means we can extract (using rewinding) both the shares committed to on the right while only rewinding Stage $a$ on the left. In case the left and right interactions are not aligned, there must be one right-proof that does not "contain" any left challenge messages (e.g. the first right-proof in Figure 8 (ii)); hence we can rewind this right-proof to extract one of the shares, without rewinding any left-proofs. The other share can then be easily extracted while only rewinding one of the stages on the left.

We proceed to a formal proof of non-malleability. On a high-level, the proof follows from the same paradigm as [LPV08], (which, in turn, follows [DDN00]). We first consider another commitment scheme $\langle C', R' \rangle$ that is a close variant of the original protocol $\langle \overline{C}, \overline{R} \rangle$. $\langle C', R' \rangle$ proceeds identically to $\langle \overline{C}, \overline{R} \rangle$, except that the receiver is allowed to request an arbitrary number of proofs for each of the random shares; to request a proof for $v_b$, it simple sends $b$ to the committer. Furthermore, at any point of the protocol execution, the receiver can request the decommitment information for one of commitments, $c_b$, by sending $(\mathsf{open}, b)$. It is easy to see that $\langle C', R' \rangle$ is hiding as each share is individually random. (A formal proof proceed by simulation using a standard argument.)

We next construct an extractor $E$ that after interacting with $C'$, outputs a simulated view of $A$ (that is identically distributed to the view of $A$ after interacting with $\bar{C}$), and furthermore extracts out the value $A$ commits to on the right (via rewinding the $\mathcal{WISSP}$ proofs in Stage 2 and 3). $E$ externally interacts with committer $C'$, and internally proceeds in the following three phases.

- In the Main Execution Phase, it emulates a man-in-the-middle execution with $A$, where messages in the left interaction are emulated by forwarding appropriate messages from the external commitment of $\langle C', R' \rangle$. Let $\Delta$ denote the simulated view of $A$.
- In the Rewinding Phase, it attempts to extract the value $A$ commits to in $\Delta$, if the right interaction is successful and has a different identity than the left interaction. To do so, it finds one "rewinding slot" in each of Stage 2 and 3 of the right interaction, such that, they are both "aligned" with proofs belonging to the same stage of the left interaction (see figure 8 (i)), or one of them is not "aligned" with any proof on the left (see figure 8 (ii)). (As argued

before, there must exist such two slots, if the left and right identities are different.) Let Stage $j$ be the stage on the left that is aligned with the two rewinding slots, and Stage $k$ be the one that is not. $E$, then, rewinds each of the two slots by sending new random challenges to $A$, until a second transcript of the slot is obtained, and hence the corresponding witness can be computed[5]. In each rewinding, it emulates the man-in-the-middle execution with $A$ as follows.

  - If $A$ expects a new Stage 1 message, it starts a new session with $C'$, and forwards the Stage 1 message of the new session to $A$.
  - If $A$ expects a message in Stage $j$ of the left interaction, it requests $C'$ to reveal the random share corresponding to Stage $j$, and emulates $\bar{C}$ in Stage $j$ honestly using $v$.
  - If $A$ expects a new complete proof in Stage $k$ of the left interaction, it simply requests for a new proof from $C'$ and forwards the proof to $A$. However, if $A$ sends (only) a challenge message and expects a new reply in Stage $k$, it cancels the rewinding.

- Finally, in the Output Phase, it outputs the view $\Delta$ and $v = w_1 \oplus w_2$, where $w_1$ and $w_2$ are the two witnesses extracted; in the case where the right interaction fails or uses the same identity as the left interaction, it sets $v = \perp$.

It is easy to see that the simulated view $\Delta$ of $A$ by $E$ is perfect and that, whenever the right interaction in $\Delta$ is successful and uses a different identity from the left interaction, $E$ would (almost always) extract out two witnesses $w_1$ and $w_2$; by the special soundness of $\langle P, V \rangle$, these are the shares committed to and thus the value committed to is extracted. Hence the output of $E$ is statistically close to the combined view and committed value by $A$ in a real execution with $\bar{C}$. Furthermore, it follows using the same argument as in Claim 5 that the expected running time of $E$ is bounded by a polynomial. Therefore, it follows from the hiding property of $\langle C', R' \rangle$ that the combined view and value committed to by $A$ in the real execution is indistinguishable, no matter which value $A$ receives commitment to in the left interaction. It follows that $\langle \overline{C}, \overline{R} \rangle$ is non-malleable.

[5] In the rare case, where the same transcript in $\Delta$ is obtained again, $E$ simply aborts the extraction phase and set $v = \mathsf{fail}$