

Unprovable Security of Perfect NIZK and Non-interactive Non-malleable Commitments

Rafael Pass*
Cornell University
rafael@cs.cornell.edu

December 17, 2012

Abstract

We present barriers to provable security of two fundamental (and well-studied) cryptographic primitives *perfect non-interactive zero knowledge (NIZK)*, and *non-malleable commitments*:

- Black-box reductions cannot be used to demonstrate *adaptive* soundness (i.e., that soundness holds even if the statement to be proven is chosen as a function of the common reference string) of any statistical (and thus also perfect) NIZK for \mathcal{NP} based on any “standard” intractability assumptions.
- Black-box reductions cannot be used to demonstrate non-malleability of non-interactive, or even 2-message, commitment schemes based on any “standard” intractability assumptions.

We emphasize that the above separations apply even if the construction of the considered primitives makes a *non-black-box* use of the underlying assumption.

As an independent contribution, we suggest a taxonomy of game-based intractability assumption based on 1) the *security threshold*, 2) the number of *communication rounds* in the security game, 3) the *computational complexity* of the game challenger, 4) the *communication complexity* of the challenger, and 5) the *computational complexity of the security reduction*.

*Original version from September 2011 contained only results regarding NIZK. The results on non-malleable commitments were added in February 2012. The current version is essentially identical to the February 2012 version. Pass is supported in part by a Alfred P. Sloan Fellowship, Microsoft New Faculty Fellowship, NSF Award CNS-1217821, NSF CAREER Award CCF-0746990, NSF Award CCF-1214844, AFOSR YIP Award FA9550-10-1-0093, and DARPA and AFRL under contract FA8750-11-2-0211. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the US Government.

1 Introduction

Modern Cryptography relies on the principle that cryptographic schemes are proven secure based on mathematically precise assumptions; these can be *general*—such as the existence of one-way functions—or *specific*—such as the hardness of factoring products of large primes. The security proof is a *reduction* that transforms any attacker A of the scheme into a machine that breaks the underlying assumption (e.g., inverts an alleged one-way function). This study has been extremely successful, and during the past three decades many cryptographic tasks have been put under rigorous treatment and numerous constructions realizing these tasks have been proposed under a number of well-studied complexity-theoretic hardness assumptions.

We here consider two fundamental cryptographic primitives—*perfect non-interactive zero-knowledge with adaptive statements* and *non-interactive non-malleable commitments*—for which security proofs based on well-studied intractability assumptions have remained elusive.

Perfect NIZK with Adaptive Inputs A non-interactive zero-knowledge (NIZK) protocol [BFM88] is protocol between two parties, a Prover, and a Verifier, through which the Prover can non-interactively (i.e., by sending a single message π) convince the Verifier of the validity of a statement x , only if x is true (this is called the *soundness* property), while at the same time revealing nothing beyond the fact that x is true (this is called the *zero-knowledge* property). To make such constructs possible both parties are additionally assumed to have access to a “Common Reference String” (CRS) that has been ideally sampled according to some distribution. The original definition of [BFM88] only considered *non-adaptive* notions of soundness and zero-knowledge: Roughly speaking, the (non-adaptive) soundness condition requires that for every false statement $x \notin L$, with high probability over the choice of the CRS, any proof π output by a malicious prover will be rejected by the verifier. The (non-adaptive) zero-knowledge property, on the other hand, requires that for every true statement $x \in L$, the joint distribution consisting of the reference string, and an honestly generated proof, can be reconstructed by a simulator. In both of these properties, the statement x is required to be *fixed* before the reference string is known. Feige, Lapidot and Shamir [FLS90] introduced stronger *adaptive* notions of both soundness and zero-knowledge; roughly speaking, here soundness and zero-knowledge should hold even if the statement x is adversarially chosen *as a function of* the reference string.

As with traditional zero-knowledge protocols, NIZKs come in several flavors: *computational NIZK*, *statistical NIZK*, and *perfect NIZK*. In the computational notion, the simulator’s output is only required to be computationally indistinguishable from an honestly generated view, whereas in the statistical (resp. perfect) variants, it is required to be statistically close (resp. identical) to an honestly generated view. Computational NIZK with adaptive zero-knowledge and soundness were constructed early on based on standard cryptographic intractability assumptions [FLS90, BY96], but constructions of statistical and perfect NIZK were elusive.

Only recently, a breakthrough result by Groth, Ostrovsky and Sahai (GOS) [GOS06] provided a construction of a perfect NIZK for \mathcal{NP} based on the hardness of a number theoretic assumption over bilinear groups. Their protocol satisfies the adaptive notion of zero-knowledge; however, it only satisfies the non-adaptive notion of soundness (that is, soundness is no longer guaranteed to hold if the attacker chooses a statement $x \notin L$ as a function of the common reference string). We here focus on whether there exists a perfect NIZK for \mathcal{NP} with both adaptive soundness and zero-knowledge.

A step towards answering this question appears in the work of Abe and Fehr [AF07], which presented a perfect NIZK for \mathcal{NP} with both adaptive soundness and zero-knowledge, using an

“knowledge-extraction” assumption (similar to the “knowledge-of-exponent” assumption of [Dam91]), as opposed to a computational-intractability assumption. Abe and Fehr also demonstrate that certain (arguably natural) types of proof techniques—which they refer to as “direct” black-box reductions—cannot be used to prove adaptive soundness of perfect NIZKs for \mathcal{NP} . Their notion of a “direct” proof, however, is quite restrictive (very roughly speaking, it requires the security reduction to “directly embed” some hard instance into the CRS in a “structure preserving way”).¹

Non-interactive Non-malleable Commitments 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. 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). In order to address the above concerns, Dolev, Dwork and Naor 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} .

More precisely, we consider a *man-in-the-middle* (MIM) attacker that participates in two concurrent executions of a commitment scheme Π ; in the “left” execution it interacts with an honest committer; in the “right” execution it interacts with an honest receiver. Additionally, we assume that the players have n -bit identities (where n is polynomially related to the security parameter), and that the commitment protocol depends only on the identity of the committer; we sometimes refer to this as the identity of the interaction. Intuitively, Π being non-malleable means that if the identity of the right interaction is different than the identity of the left interaction (i.e., A does not use the same identity as the left committer), the value A commits to on the right does not depend on the value it receives a commitment to on the left; this is formalized by requiring that for any two values v_1, v_2 , the value A commits to after receiving left commitments to v_1 or v_2 are indistinguishable.

The first non-malleable commitment protocol was constructed by Dolev, Dwork and Naor [DDN00] in 1991. The security of their protocol relies on the minimal assumption of one-way functions and requires $\Omega(\log n)$ rounds of interaction, where $k \in N$ is the length of party identities. The round-complexity of non-malleable commitments has since been extensively studied (see e.g., [Bar02, PR05b, PR05a, LPV08, LP09, PW10, Wee10]), leading up to constant round protocols based on one-way functions [LP11, Goy11].

The question of whether non-interactive, or even 2-round, non-malleable commitments exist, however, is wide open. (We note that in the Common Reference String model, constructions of non-interactive non-malleable commitments are known [CIO98]; we here focus on constructions in

¹Among other things, the structure preserving property requires that if the “hard instance” being directly embedded in the CRS is true, the CRS is valid, and if the hard instance is false, then the CRS is “invalid”. This property can never hold when considering NIZK in the Uniform Reference String model (as every CRS is valid), and as such their result holds vacuously when considering NIZK in the Uniform Reference String model.

the plain model, without any set-up.) Some initial progress towards this question can be found in [PPV08] where a construction of non-interactive non-malleable commitments based on a new hardness assumption is given; this assumption, however, has a strong non-malleability flavor; as such, it provides little insight into the question of whether non-malleability can be obtained from a “pure” hardness assumptions (such as e.g., the hardness of factoring).

1.1 Our results

The main result of this paper is showing that Turing (i.e., black-box) reductions cannot be used to base the security of the above-mentioned primitives, on a general class of intractability assumption.

More precisely, following Naor [Nao03] (see also [DOP05, HH09, RV10, Pas11, GW11]), we model an *intractability assumption* as an arbitrary game between a (potentially) unbounded challenger C , and an attacker A . A is said to break the assumption C with respect to the threshold t if it can make C output 1 with probability non-negligibly higher than the threshold t . All traditional cryptographic hardness assumptions (e.g., the hardness of factoring, the hardness of the discrete logarithm problem, the decisional Diffie-Hellman problem etc.) can be modeled as 2-round challengers C with the threshold t being either 0 (in case of the factoring or discrete logarithm problems) or $1/2$ (in case of the decisional Diffie-Hellman problem). In all these examples C is polynomial-time; Naor [Nao03] and Gentry and Wichs [GW11] refer to such assumptions as “falsifiable”. For generality, we here (following [Pas11]) refer to these as “efficient-challenger” assumptions. More generally, we refer to an assumption where the challenger can be implemented in time (resp. size) $T(\cdot)$ as a “ $T(\cdot)$ -time (resp. size) challenger assumption”. Note that more “esoteric” assumptions such as the “one-more discrete logarithm assumption” [BNPS03, BP02], or “adaptive one-way functions” [PPV08], are not efficient-challenger assumptions, but they are exponential-time challenger assumptions.

Our first result rules out basing statistical (and thus also perfect) NIZK with adaptive soundness on efficient-challenger (a.k.a falsifiable) assumptions.

Theorem 1 (Main Theorem 1—Informally stated). *Assume the existence of (non-uniformly hard) one-way functions. Then there exists an \mathcal{NP} -language L such that the following holds. Let Π be a statistical non-interactive adaptively zero-knowledge argument for L . Assume there exists a polynomial-time Turing reduction R such that R^A breaks the efficient-challenger assumption C w.r.t. the threshold t for every A that breaks adaptive soundness of Π . Then C can be broken in polynomial-time with respect to the threshold t .*

We next show that if we additionally assume the existence of sub-exponential one-way functions, and consider the constructions of NIZK for proving *any* polynomial-length (in the security parameter) statement in \mathcal{NP} based on a particular *exponential-time* challenger assumption (C, t) , then the assumption can already be broken in polynomial time.

Moving on to non-interactive non-malleable commitments, we show that if non-malleability of a non-interactive, or two message, commitment scheme Π can be based on a efficient-challenger (resp. $T(\cdot)$ -size) challenger assumption (C, t) using a polynomial-time (resp. $T(\cdot)$ -sized) security reduction, then C can be broken in polynomial-time (resp. by a $\text{poly}(T(\cdot))$ -sized circuit).

Theorem 2 (Main Theorem 2—Informally stated). *Let Π be a two-message commitment scheme. Assume there exists a polynomial-time (resp. $T(\cdot)$ -size) Turing reduction R such that R^A breaks the efficient-challenger (resp. $T(\cdot)$ -size) assumption C w.r.t. the threshold t for every A that breaks non-malleability of Π . Then C can be broken in polynomial-time (resp. by a $\text{poly}(T(\cdot))$ -sized circuit) with respect to the threshold t .*

We emphasize that for all the above-mentioned results, the *construction* of the protocols Π need not make use of the underlying assumption in a black-box way; the only restriction we impose is that the security *reduction* is a Turing (i.e., black-box) reduction.

Let us also remark that although we see only superficial similarities between the primitives of non-interactive statistical NIZK and non-interactive non-malleable commitments (e.g., they both refer to non-interactive primitives), the techniques used to prove the above impossibility results have significant overlap.

Uniform v.s. Non-uniform Security Reductions In this work we focus on ruling out *uniform* security reductions; that is, the security reduction is a Turing machine that gets no advice about the attacker. Nevertheless, a very recent work by Chung, Lin, Mahmoodi and Pass [CLMP13] provides techniques for extending certain types of separation results for the uniform setting also to the *non-uniform* setting (where we consider reductions that may receive a polynomial-length advice about the attacker). These techniques readily apply to our results, which thus also extend to rule out non-uniform security reductions.

A Taxonomy of Intractability Assumption As an independent contribution, we slightly generalize the notion of an intractability assumption from [Pas11] (see also [Nao03, DOP05, HH09, RV10, GW11]) and provide an, in our eyes, natural taxonomy of intractability assumptions based on 1) the *security threshold*, 2) the number of *communication rounds* in the security game, 3) the *computational complexity* of the game challenger, 4) the *communication complexity* of the challenger, and 5) the *computational complexity of the security reduction*. Our results, combined with [Pas11, GW11], demonstrate several natural primitives that may be (trivially) based on assumption of a certain type (e.g., the soundness condition of a perfect NIZK can trivially be viewed as a bounded-round assumption), but cannot be based on a different type of assumption (e.g., an assumption where the challenger is efficient). Our results focus on understanding limitations in terms of items 1, 2, 3 and 5; we leave open an exploration of item 4, i.e., the communication complexity of the challenger. More generally, we are optimistic that cryptographic tasks may be classified in this taxonomy, based on whether they can be achieved—even using a *non-black-box construction*—based on a class of assumptions in this taxonomy, but not on another (much like the celebrated taxonomy of Impagliazzo [Imp95] in the context of *black-box constructions*.)

A Note on Random Oracles Let us point out that in the Random Oracle model [BR93], both of the above-mentioned primitives are easy to construct. Perfect NIZK were constructed in [BR93] (by relying on the “Fiat-Shamir heuristics” [FS87]) and non-interactive non-malleable commitments in [Pas03a]. Indeed, many practical protocols rely on the assumption that a “good” hashfunction behaves like a non-interactive non-malleable commitment, and on non-interactive zero-knowledge arguments constructed by applying the “Fiat-Shamir heuristics” [FS87] to a three-message perfect zero-knowledge protocol. Our results show that such commonly used sub-protocols cannot be proven secure based on standard hardness assumptions. Note that these results are incomparable to those of e.g., [CGH04, GK03] on the “uninstantiability of random oracles”: the results of [CGH04, GK03] are stronger in the sense that any instantiation of their scheme with a concrete function can actually be *broken*, whereas we just show that the instantiated scheme cannot be *proven secure* using a Turing reduction based on standard assumptions. On the other hand, the separations of [CGH04, GK03] consider “artificial protocols”, whereas the protocols we consider are natural (and commonly used in practice).

1.2 Related Separation Results

There is a large literature on separation results between cryptographic primitives/assumptions. We distinguish between two types of results.

Separations for fully black-box constructions The seminal work of Impagliazzo and Rudich [IR88] provides a framework for proving black-box separations between cryptographic primitives. We highlight that this framework considers so-called “fully-black-box constructions” (see [RTV04] for a taxonomy of various black-box separations); that is, the framework considers both black-box *constructions* (i.e., the higher-level primitive only uses the underlying primitive as a black-box), and black-box *reductions*.

Separations for black-box reductions In recent years, new types of black-box separations have emerged. These types of separation apply even to non-black-box constructions, but still only rule out black-box proofs of security: Pass [Pas06] and Pass, Tseng and Venkatasubramanian [PTV11] (relying on the works of Brassard [Bra83] and Akavia et al [AGGM06], demonstrating limitations of “NP-hard Cryptography”²) demonstrate that under certain (new) complexity theoretic assumptions, various cryptographic task cannot be based on *one-way functions* using a black-box security reduction, even if the protocol uses the one-way function in a non-black-box way. Very recently, two independent works demonstrate similar types of separation bounds, but this time ruling our security reductions to a *general* set of intractability assumptions: Pass [Pas11] demonstrates impossibility of using black-box reductions to prove the security of several primitives (e.g., Schnorr’s identification scheme, commitment scheme secure under weak notions of selective opening, Chaum Blind signatures, etc) based on any “bounded-round” intractability assumption (where the challenger uses an a-priori bounded number of rounds, but is otherwise unbounded). Gentry and Wichs [GW11] demonstrate (assuming the existence of strong pseudorandom generators) impossibility of using black-box security reductions to prove soundness of “succinct non-interactive arguments” based on any “falsifiable” assumption (where the challenger is computationally bounded). Both of the above-mentioned work fall into the “meta-reduction” paradigm of Boneh and Venkatesan [BV98], which was earlier used to prove separations for restricted types of reductions (see e.g., [BMV08, HRS09, FS10]). Our separation results are in the vein of these two works, and follows some of their techniques.

1.3 Proof Overview: Ruling out Perfect NIZK with Adaptive Inputs

Assume there exists a perfect NIZK (P, V) for a hard-on-the average language L ; for simplicity, in this proof overview we focus on the case when the reference string is uniformly random (i.e., we consider only NIZK in the so-called Uniform Reference String (URS) Model). Assume further that there exists a Turing reduction R such that R^A breaks the assumption C (with respect to some thresholds t) whenever A breaks adaptive soundness of (P, V) . Following the “meta-reduction” paradigm by Boneh and Venkatesan [BV98] (which is used in both [Pas11] and [GW11], and also [AF07]), we want to use R to directly break C .

More precisely (just as in [Pas11, GW11]) we exhibit a particular attacker A to the adaptive soundness of (P, V) and next show how to “emulate” this attacker for R without disturbing R ’s interaction with C . Whereas in [Pas11] the emulation was statistically close (and thus the separation

²See also the results of Feigenbaum and Fortnow [FF93] and the result of Bogdanov and Trevisan [BT03] that demonstrate limitations of NP-hard cryptography for *restricted* types of reductions.

could be applied also to unbounded challengers), in [GW11] the emulation was only *computationally indistinguishable*, but this still suffices for convincing C as long as C is computationally efficient. We here follow the approach of [GW11].

Let us turn to describing our attacker A , and next explain how to emulate it. Given a CRS ρ , A first attempts to recover the random coins r used by the simulator S when outputting the CRS ρ ; since the simulation is perfect, such a string r exists (but finding r might require super-polynomial time). (Recall that since we are dealing with adaptive zero-knowledge, the zero-knowledge simulator needs to output a reference string ρ before knowing what statement it needs to simulate a proof of.) Next, A samples a false instance $x \notin L$ which is indistinguishable from a true instance (since L is hard-on-the average, this can be done efficiently). Finally, it runs the simulator S on the random coins r to generate ρ , and next feeds it the instance x , and lets π denote the proof output by S (again this final step is efficient).

Let us argue that the proof π of x is accepted by $V(\rho)$. Towards this, consider a hybrid attacker A' that performs exactly the same steps as A , but instead samples a *true* instance $x \in L$. It follows from the ZK property (combined with the completeness property) that V accepts the proofs output by A' . Now, intuitively, it should follow from the hard-on-the-average property of L that V also accepts the proofs output by A . But there is a catch: recall that A is *not efficient*. However, since it is only the first step of A that is inefficient, we can fix the random string r non-uniformly and still use the remaining steps of A and the efficient verifier V to contradict the hard-on-average property of L , as long as we assume that L is hard-on-average for non-uniform polynomial-time. Note that we here rely on the fact that A is allowed to choose the statement x *after* having seen the reference string ρ (i.e., we rely on A breaking *adaptive* soundness)—this is what allows us to non-uniformly choose r as a function of ρ , *before* sampling $x \in L$.

Now given this breaker A , let us see an attacker \tilde{A} that efficiently simulates it (in a computationally indistinguishable way). $\tilde{A}(\rho)$ simply picks a random true statement x together with a witness w , and next runs the honest prover strategy $P(\rho, x, w)$ to produce a proof π (this strategy is similar to the one used in [GW11]). It follows by the ZK property that the output of C when communicating with \tilde{A} and A' are indistinguishable, and we can then apply a similar argument as above (but more complicated) to argue that the output of C when communicating with A' and A are indistinguishable, and thus $R^{\tilde{A}}$ breaks C with roughly the same probability as R^A does.

Dealing with exponential-time challenger assumptions In case the running-time of the challenger C is super-polynomial in the security parameter k , the above approach seemingly fails: the fact that \tilde{A} generates computationally indistinguishable messages does not suffice to argue that C still accepts in the interaction with $R^{\tilde{A}}$. However, if we assume that the language L is hard-on-the-average for non-uniform subexponential time, then the above approach still works, as long as C is subexponential time; in fact, it rules out also subexponential-time reductions. To deal with also exponential-time challenger assumptions, we proceed as follows. If the *same* assumption C can be used to prove *any* statement in \mathcal{NP} of length polynomial in the security parameter, then if the language L is hard-on-the-average for non-uniform sub-exponential time, it suffices to pick statements x that are sufficiently long (but still of polynomial length) to ensure that \tilde{A} generates messages that are indistinguishable from those sent by \tilde{A} , even by C .

1.4 Proof Overview: Ruling out Non-interactive Non-malleable Commitments

Assume there exists a non-interactive commitment scheme Π ; for simplicity of exposition we here focus only on non-interactive, as opposed to two-message, commitments. Assume further that there

exists a Turing reduction R such that R^A breaks the assumption C (with respect to some thresholds t) whenever A breaks non-malleability of Π . Recall that an attacker A that breaks non-malleability of Π participates in two interactions—one on the “left” acting as a receiver, and one on the “right” acting as a committer. To be successful A needs to choose a different identity for the left and right interactions, and must commit to a value \tilde{v} which is related to the value v it receives a commitment to on the left. Consider a strong attacker A that chooses identity 0 on the left, and 1 on the right, and upon receiving a commitment c recovers (using brute force) the unique value v that c is a commitment to (if the value is not unique v is set to \perp), and next honestly commits to v on the right. Clearly A breaks non-malleability of Π , and thus R^A also breaks C w.r.t. t .

Let us now see how to efficiently emulate A . We simply consider a “trivial” adversary \tilde{A} that picks identity 0 on the left and 1 on the right (just as A), but instead of trying to commit to v on the right, it simply commits to 0 on the right. Now, intuitively, if the reduction R and the challenger C are polynomial-time, then it should follow by the hiding property of Π that $R^{\tilde{A}}$ still breaks C (w.r.t. t). Note, however, that R may be asking its oracle to break non-malleability of multiple commitments, and since A is not efficiently computable, we need to be a bit careful when doing the hybrid argument. Nevertheless, using a careful ordering of the hybrid (and as in the lower bound for statistical NIZK) relying on the *non-uniform* hiding property of Π we can show that $R^{\tilde{A}}$ still breaks C (w.r.t. t).

Note that the above proof idea applies to a very weak notion of “one-sided” non-malleability, where the attacker always uses identity 0 on the left and 1 on the right; Liskov et al [LLM⁺01] call commitments satisfying this weak notion of non-malleability, *mutually independent*. Interestingly, [LLM⁺01] shows a construction of a mutually independent commitment based on the existence of subexponentially hard one-way permutations. The idea is simple: Let Com_0 be a commitment scheme that is hard for subexponential time, and let Com_1 be a commitment scheme that can be fully broken in subexponential time. If a MIM upon receiving a commitment of v using Com_0 is able to output a commitment to a related value \tilde{v} using Com_1 , then we can violate the hiding of Com_0 by breaking Com_1 using brute-force. This security reduction, however, is super-polynomial (subexponential) time. A natural question is thus whether subexponential time/size reductions may be helpful for constructing “full-fledged” (as opposed to one-sided) non-interactive commitments.³ We proceed to rule out such reductions (or rather to show that if there exists such a reduction, then the reduction itself must already break the assumption).

Consider a $T(k)$ -sized reduction R , where $T(k)$ is super-polynomial, for basing non-malleability on an efficient challenger assumption C^4 , and consider the algorithms A and \tilde{A} described above. Note that if R has super-polynomial size, we have no guarantees that $R^{\tilde{A}}$ breaks C even if R^A does; since hiding of Π is only required to hold for polynomial-sized algorithms, $R^{\tilde{A}}$ ’s success probability may be very different from R^A success probability. But in this case, intuitively, R itself must be able to break the hiding of commitments using identity 1 (recall that A and \tilde{A} use identity 1 on the right).

So, if $R^{\tilde{A}}$ does not already convince C , we can use R (in conjunction with C) to obtain a circuit D that distinguishes, say commitments to 0^k and 1^k using identity 1.⁵ We may then use D to construct a man-in-the-middle attacker A' that chooses identity 1 on the *left* and 0 on the right (as

³Indeed, [PW10] rely on intuitions similar to those from mutually independent commitments to construct a “full-fledged” non-malleable commitment, but this construction requires multiple communication rounds.

⁴The assumption that C is an efficient challenger is only made here to simplify exposition; our actual proof also works when C is $T(k)$ -sized.

⁵As in the previous proof, to obtain a machine that breaks the hiding of the commitment, we need to rely a polynomial-length non-uniform advice to deal with the above-mentioned inefficiency issue in the hybrid argument; this is why we work with circuits here.

opposed to 0 on the left and 1 on the right, as A and \tilde{A} did) to break non-malleability of Π , and finally use R combined with A' to directly break C . So, summarizing, either $R^{\tilde{A}}$ works, or else, we use R in order to construct an MIM A' that breaks non-malleability, and then use $R^{A'}$ to convince C —in essence, we use R “on itself” to convince C .

1.5 Overview of the Paper

We provide some preliminaries and standard definitions in Section 2. We provide definitions of intractability assumptions and black-box reductions in Section 3; this section also contains our taxonomy of intractability assumptions. We formally state and prove our results about NIZK in Section 4. We formally state and prove our results about non-malleable commitments in Section 5.

2 Preliminaries

2.1 Notation

Integer, Strings and Vectors. We denote by N the set of natural numbers: $0, 1, 2, \dots$. Unless otherwise specified, a natural number is presented in its binary expansion (with no *leading* 0s) whenever given as an input to an algorithm. If $n \in N$, we denote by 1^n the unary expansion of n (i.e., the concatenation of n 1’s). Given a string x , we let $x|_i$ denote the i ’th bit of x . We denote by \vec{x} a finite sequence of elements x_1, x_2, \dots, x_n , and we let $|\vec{x}|$ denote the number of elements in the sequence.

Probabilistic notation. We employ the following probabilistic notation. We focus on probability distributions $X : S \rightarrow R^+$ over finite sets S .

Probabilistic assignments. If D is a probability distribution and p a predicate, then “ $x \stackrel{R}{\leftarrow} D$ ” denotes the elementary procedure consisting of choosing an element x at random according to D and returning x .

Probabilistic experiments. Let p be a predicate and D_1, D_2, \dots probability distributions, then the notation $\Pr \left[x_1 \stackrel{R}{\leftarrow} D_1; x_2 \stackrel{R}{\leftarrow} D_2; \dots : p(x_1, x_2, \dots) \right]$ denotes the probability that $p(x_1, x_2, \dots)$ will be true after the ordered execution of the probabilistic assignments $x_1 \stackrel{R}{\leftarrow} D_1; x_2 \stackrel{R}{\leftarrow} D_2; \dots$

New probability distributions. If D_1, D_2, \dots are probability distributions, the notation $\{x \stackrel{R}{\leftarrow} D_1; y \stackrel{R}{\leftarrow} D_2; \dots : (x, y, \dots)\}$ denotes the new probability distribution over $\{(x, y, \dots)\}$ generated by the ordered execution of the probabilistic assignments $x \stackrel{R}{\leftarrow} D_1, y \stackrel{R}{\leftarrow} D_2, \dots$.

Probability ensembles. A *probability ensemble* is a vector of random variables $X = \{X_n\}_{n \in N}$. We will consider ensembles of the form $X = \{X_n\}_{n \in N}$ where X_n ranges over strings of length $p(k)$, for some fixed, positive polynomial p .

In order to simplify notation, we sometimes abuse of notation and employ the following “short-cut”: Given a probability distribution X , we let x denote the random variable obtained by selecting $x \leftarrow X$ and outputting x .

Algorithms. We employ the following notation for algorithms.

Probabilistic algorithms. By a probabilistic algorithm we mean a Turing machine that receives an auxiliary random tape as input. If M is a probabilistic algorithm, then for any input x , the notation “ $M_r(x)$ ” denotes the output of the M on input x when receiving r as random tape. We let the notation “ $M(x)$ ” denote the probability distribution over the outputs of M on input x where each bit of the random tape r is selected at random and independently, and then outputting $M_r(x)$.

Oracle algorithms. An oracle algorithm is a machine that gets oracle access to another machine. Given a probabilistic oracle algorithm M and a probabilistic algorithm A , we let $M^A(x)$ denote the probability distribution over the outputs of the oracle algorithm M on input x , when given oracle access to A . We emphasize that if the algorithm A is probabilistic, M does not get to control the randomness of A , and each time A is invoked fresh randomness is used by A .⁶

Negligible functions. The term “negligible” is used for denoting functions that are asymptotically smaller than the inverse of any fixed polynomial. More precisely, a function $\nu(\cdot)$ from non-negative integers to reals is called *negligible* if for every constant $c > 0$ and all sufficiently large n , it holds that $\nu(k) < n^{-c}$.

2.2 Basic Definitions

In this section we provide some standard definitions and preliminaries.

2.3 Indistinguishability

The following definition of (computational) indistinguishability originates in the seminal paper of Goldwasser and Micali [GM84].

Let X be a countable set of strings. A *probability ensemble indexed by X* is a sequence of random variables indexed by X . Namely, any element of $A = \{A_x\}_{x \in X}$ is a random variable indexed by X .

Definition 1 (Indistinguishability). *Let X be a countable set. Two ensembles $\{A_{n,x}\}_{n \in N, x \in X}$ and $\{B_{n,x}\}_{n \in N, x \in X}$ are said to be computationally indistinguishable, if for every probabilistic machine D (the “distinguisher”) whose running time is polynomial in its first input, there exists a negligible function $\nu(\cdot)$ so that for every $n \in N, x \in X$:*

$$|\Pr [D(n, x, A_{n,x}) = 1] - \Pr [D(n, x, B_{n,x}) = 1]| < \nu(k)$$

In the above expression, D is simply given a sample from $A_{x,y}$ and $B_{x,y}$, respectively. $\{A_{n,x}\}_{n \in N, x \in X}$ and $\{B_{n,x}\}_{n \in N, x \in X}$ are said to be statistically indistinguishable over X if the above condition holds for all (possibly unbounded) machines D .

⁶The fact that we do not allow black-box reductions to control the randomness of the oracle they communicate with may seem restrictive. As we point out, however, our result apply even if we consider reductions that work only for *deterministic* attackers; see Remark 2.

2.4 Witness Relations

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

Definition 2 (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 $L = \{x : \exists w \text{ s.t. } (x, w) \in R_L\}$.*

We say that w is a witness for the membership $x \in L$ if $(x, w) \in R_L$. We will also let $R_L(x)$ denote the set of witnesses for the membership $x \in L$, i.e., $R_L(x) = \{w : (x, w) \in R_L\}$.

3 Intractability Assumptions and Black-box Reductions

Our definition of an intractability assumption closely follows [Pas11]. Following Naor [Nao03] (see also [DOP05, HH09, RV10]), we model an intractability assumption as an interaction (or game) between a probabilistic machine C —called the challenger—and an attacker A . Both parties get as input 1^k where k is the security parameter. Any such challenger C , together with a threshold function $t(\cdot)$ intuitively corresponds to the assumption:

For every polynomial-time adversary A , there exists a negligible function μ such that for all $k \in \mathbb{N}$, the probability that C outputs 1 after interacting with A is bounded by $t(k) + \mu(k)$.

We say that A breaks C w.r.t t with probability p on common input 1^k if $\Pr [\langle A, C \rangle(1^k) = 1] \geq t(k) + p$.

If the challenger C is polynomial-time in the length of the messages it receives, we say that the assumption is *efficient challenger*; such assumptions are referred to as *falsifiable* assumptions by Naor [Nao03] and Gentry and Wichs [GW11]. More generally, we refer to an assumption as having a $T(\cdot, \cdot)$ -time (resp. size) challenger if C can be implemented in time (resp. size) $T(k, \ell)$ on input the security parameter 1^k , and when receiving messages of length ℓ . (C, t) is an efficient challenger assumption iff C is a $T(\cdot, \cdot)$ -assumption where $T(k, \ell)$ is polynomial in both k and ℓ . For simplicity, we here consider either $\text{poly}(k, \ell)$ -time (or size) challengers, or $T(k, \ell) = T(k)$ -time (or size) challengers, where the running-time of the challenger is bounded only as a function of the security parameter.

We can easily model all “traditional” cryptographic assumptions as efficient challengers C and a threshold t . For instance, the assumption that a particular function f is (strongly) one-way corresponds to the threshold $t(k) = 0$ and the 2-round challenger C that on input 1^k pick a random input x of length k , sends $f(x)$ to the attacker, and finally outputs 1 iff the attacker returns an inverse to $f(x)$. Decisional assumptions (such as, e.g., the decisional Diffie-Hellman problem, or the assumption that a particular function g is a pseudorandom generator) can also easily be modelled as 2-round challengers but now we have the threshold $t(k) = 1/2$. More esoteric assumptions such as the “one-more discrete logarithm assumption” [BNPS03, BP02], or “adaptive one-way functions” [PPV08], are not efficient-challenger assumptions; however, they can be modeled as *exponential-time* challenger assumptions.

We may also consider other restricted types of intractability assumptions. For instance, [Pas11] considers challengers C that are computationally unbounded, but for which there exists a polynomial upper bound in the terms of the security parameter k on the number of communications rounds by C ; we refer to these assumptions as *bounded round* intractability assumptions. Another interesting class of assumptions is obtained by further restricting the communication complexity of

C ; for instance, we may require that there is a polynomial bound (again in terms of the security parameter k) on the communication complexity of C ; we refer to these assumptions as *bounded-communication intractability assumption*.

In this work we focus on establishing impossibility results for reductions from efficient-challenger (just as the results of [GW11]), and more generally time/size $T(\cdot)$ challenger assumptions where $T(\cdot)$ is a super-polynomial function. As mentioned, in [Pas11] impossibility results for reductions from bounded-round assumptions are presented. Since both non-malleability of a protocol, and adaptive soundness of a NIZK, is a bounded round assumption, we cannot hope to strengthen our result to rule out reductions also from bounded round assumptions. We leave open an exploration of bounded-communication intractability assumptions.

Finally, note that we can capture super-polynomial hardness of an assumption by allowing for super-polynomial-time reductions to the assumption.

A Taxonomy of Intractability Assumption The above way of modeling assumptions, provides an, in our eyes, natural taxonomy of intractability assumptions based on 1) the *security threshold* t , 2) the number of *communication rounds* used by C , 3) the *computational complexity* of C , 4) the *communication complexity* of C , and 5) the *computational complexity of the security reduction*. We are optimistic that cryptographic tasks may be classified in this taxonomy, based on whether they can be achieved—even using a *non-black-box construction*—based on class of assumptions in this taxonomy, but not on another (much like the celebrated taxonomy of Impagliazzo [Imp95] in the context of *black-box constructions*.)

Indeed, as mentioned above, the results of [Pas11, GW11] already yield some results in this direction, separating unbounded-round and bounded-round assumptions [Pas11] and unbounded-challenger and efficient-challenger assumptions [GW11]. The results in this paper further elucidate the landscape; among other things, separating unbounded challenger and exponential-time challenger assumptions, and exponential-time and efficient-challenger assumptions.

An interesting question for future work is obtaining separations for non-black-box constructions for more “structured” types of assumptions (such as the existence of one-way functions, one-way permutations). The results of [Pas06, PTV11] provide a first step in this direction, exhibiting separations from one-way functions for some natural cryptographic primitives, but rely on new complexity-theoretic assumptions.

Black-box Reductions We consider probabilistic polynomial time Turing reductions—i.e., *black-box reductions*. A black-box reduction refers to a probabilistic polynomial-time oracle algorithm. Roughly speaking, a black-box reduction for basing the security of a primitive P on the hardness of an assumption C , is a probabilistic polynomial-time oracle machine R such that whenever the oracle O “breaks” P with respect to the security parameter k , then R^O “breaks” C with respect to a polynomially related security parameter k' such that k' can be efficiently computed given k . We restrict to the case when $k' = k$. This is without loss of generality: we can always redefine the assumption C so that it on input k acts as if its input actually was k' (since k' can be efficiently computed given k). To formalize this notion, we thus restrict to oracle machines R that on input 1^k always query their oracle on inputs $(1^k, \cdot)$.

Definition 3. *We say that R is a valid black-box reduction if R is an oracle machine such that $R(1^k)$ only queries its oracle with inputs of the form $(1^k, y)$, where $y \in \{0, 1\}^*$.*

The reason that we (and as it standard in the literature) restrict R to only query its oracle on a single “input length” k , is that standard cryptographic definitions require ruling out the existence of

attackers that break some primitive even for *infinitely many* input lengths; as these inputs lengths can be very sparse, a *black-box* reduction must be successful even if it has access to an attacker that only succeeds on a single input length.⁷

4 Security of Perfect Adaptive NIZK

We recall the traditional definition of non-interactive proofs in the Common Reference String (CRS) model. For generality (and since we are proving a lower bound) we allow the CRS ρ be generated by an arbitrary polynomial-time distribution (as opposed to requiring it to be uniformly random). In the adaptively-sound notion of a non-interactive proof/argument, we require that soundness holds even if the attacker may adaptively pick a statement after having seen the CRS. We here focus only on proofs for languages in \mathcal{NP} where the *prover is efficient* when given an \mathcal{NP} -witness.

Definition 4 (Non-Interactive Proofs/Arguments). *A triple of algorithms, (\mathcal{D}, P, V) , is called a non-interactive proof system (with non-adaptive soundness) for a language L if the algorithm \mathcal{D} is probabilistic polynomial-time, the algorithm V is a deterministic polynomial-time, and P is probabilistic polynomial-time, such that the following two conditions hold:*

- *Completeness: There exists a negligible function μ such for every $x \in L$, every $w \in R_L(x)$ and every $k \in \mathbb{N}$,*

$$\Pr \left[\rho \leftarrow \mathcal{D}(1^k, 1^{|x|}); \pi \leftarrow P(1^k, x, w, \rho) : V(1^k, x, \rho, \pi) = 1 \right] \geq 1 - \mu(k)$$

- *Soundness: For every algorithm B and every polynomial q , there exists a negligible function μ such that for every $k \in \mathbb{N}$ and every $x \notin L$ such that $|x| \leq q(k)$*

$$\Pr \left[\rho \leftarrow \mathcal{D}(1^k, 1^{|x|}); \pi' \leftarrow B(1^k, x, \rho) : V(1^k, x, \rho, \pi') = 1 \right] \leq \mu(k)$$

If additionally the following condition holds, then we call (\mathcal{D}, P, V) an adaptively-sound non-interactive proof system:

- *Adaptive Soundness: For every algorithm B and every polynomial q , there exists a negligible function μ such that for every $k \in \mathbb{N}, n \in [q(k)]$*

$$\Pr \left[\rho \leftarrow \mathcal{D}(1^k, 1^n); (x, \pi') \leftarrow B(1^k, 1^n, \rho) : V(1^k, x, \rho, \pi') = 1 \wedge |x| = n \wedge x \notin L \right] \leq \mu(k)$$

Finally, if the soundness (resp adaptive soundness) condition only holds w.r.t polynomial-time adversaries B , we call (\mathcal{D}, P, V) a non-interactive argument (resp. an adaptively-sound non-interactive argument).

Let us turn to defining zero-knowledge. Also here there is a non-adaptive and an adaptive version. In the *non-adaptive* definition of zero-knowledge from [BFM88], there is a single simulator, which, after seeing the statement to be proven, generates both the CRS and the proof at the same time. In the *adaptive* definition from [FLS90], there are two simulators—the first of which must output a string before seeing any theorems. The stronger adaptive definition guarantees zero-knowledge even when the statement to be proved is chosen as a function of the CRS. We here focus only on adaptive zero-knowledge.

⁷For instance, consider an attacker that succeeds only on input lengths $2^c, 2^{2^c}, \dots$ (and outputs \perp on all other inputs); a black-box reduction that only accesses its oracle on a polynomially related security parameter can only access a single “non- \perp ” input length

Definition 5 (Non-Interactive Zero-Knowledge). *Let (\mathcal{D}, P, V) be a non-interactive proof system for the language L . We say that (\mathcal{D}, P, V) is (adaptively) zero-knowledge if there exists two probabilistic polynomial-time simulators S_1, S_2 such that for every polynomial q , every non-uniform polynomial-time statement-choosing algorithm $c(\cdot)$ that on input $(1^k, 1^n, \rho)$ outputs a n -bit statement $x \in L$, and every function $w(\cdot)$ such that $w(x) \in R_L(x)$, the following two ensembles are computationally indistinguishable*

$$\left\{ \rho \leftarrow \mathcal{D}(1^k, 1^n); x \leftarrow c(1^k, 1^n, \rho); w \leftarrow w(x); \pi \leftarrow P(1^k, x, w, \rho) : (\rho, x, \pi) \right\}_{k \in N, n \in [q(k)]}$$

$$\left\{ (\rho, \text{aux}) \leftarrow S_1(1^k, 1^n); x \leftarrow c(1^k, 1^n, \rho); \pi' \leftarrow S_2(1^k, x, \text{aux}) : (\rho, x, \pi') \right\}_{k \in N, n \in [q(k)]}$$

We furthermore say that (\mathcal{D}, P, V) is *perfect* (resp. *statistical*) zero-knowledge if the above ensembles are identically distributed (resp. statistically close).

We use the (common) acronym “NIZK” to denote a non-interactive zero-knowledge proof or argument. Feige, Lapidot and Shamir and Bellare and Yung [FLS90, BY96] (building on [BFM88]) show that the existence of enhanced trapdoor permutations implies that all of \mathcal{NP} has a adaptively-sound NIZK, but the zero-knowledge property is only computational. As mentioned, Groth, Ostrovsky and Sahai [GOS06] show (under some number theoretic assumptions) that all of \mathcal{NP} has a *perfect* NIZK with non-adaptive soundness. More recently, Abe and Fehr [AF07] present a perfect NIZK for \mathcal{NP} also with adaptive soundness but based the soundness property on a “knowledge” assumption (rather than an intractability assumption).

We aim to prove limitations of basing even weak notions of adaptive soundness for perfect or statistical NIZK for \mathcal{NP} on intractability assumptions. Let us first explicitly define what it means to break adaptive soundness of a NIZK.

Definition 6 (Breaking Adaptive Soundness). *We say that A breaks adaptive soundness of (\mathcal{D}, P, V) w.r.t the language L on input lengths $q(\cdot)$ with probability $\mu(\cdot)$ if for every $k \in N$,*

$$\Pr \left[\rho \leftarrow \mathcal{D}(1^k, 1^{q(k)}); (x, \pi') \leftarrow A(1^k, \rho) : V(1^k, x, \rho, \pi') \wedge |x| = q(k) \wedge x \notin L = 1 \right] \geq \mu(k)$$

Let us turn to defining what it means to base adaptive soundness on an intractability assumption C .

Definition 7 (Basing Adaptive Soundness on the Hardness of C). *We say that R is a black-box reduction for basing adaptive soundness of (\mathcal{D}, P, V) w.r.t. L and input lengths q , on the hardness of C w.r.t threshold $t(\cdot)$ if R is a valid black-box reduction and there exists a polynomial $p(\cdot, \cdot)$ such that for every probabilistic machine A that breaks adaptive soundness of (\mathcal{D}, P, V) w.r.t L and inputs lengths $q(\cdot)$ with probability $\mu(\cdot)$, for every $k \in N$, R^A breaks C w.r.t t with probability $p(\mu(k), 1/k)$ on input 1^k .*

Note that we here require that R^O breaks the assumption C on the security parameter k by querying O on the *same* security parameter k . As previously mentioned, a seemingly more general definition would allow R^O to break C on a polynomially-related security parameter k' (which can be efficiently computed given k), but this extra generality does not buy us anything as we can always re-define C to on input k act as its input was k' .

We now have the following theorem:

Theorem 3. *Assume the existence of non-uniformly hard one-way functions. Then there exists an \mathcal{NP} -language L such that the following holds. Let (\mathcal{D}, P, V) be a statistical non-interactive adaptively zero-knowledge argument for L , let $q(k)$ be polynomially related to k , and let (C, t)*

be any efficient-challenger assumption. If there exists a black-box reduction R for basing adaptive soundness of (\mathcal{D}, P, V) w.r.t L and input lengths q on the hardness of C w.r.t threshold t , then there exists a probabilistic polynomial-time machine B and a polynomial $p'(\cdot)$ such that for infinitely many $k \in \mathbb{N}$, B breaks C w.r.t t with probability $\frac{1}{p'(k)}$ on input 1^k . If furthermore assuming the existence of one-way functions secure against non-uniform subexponential-time algorithms, the above holds even if C is subexponential-time computable.

Let us remark that as shown in [Ost91, OW93, PS05], any (even computational) NIZK for a hard-on-the-average language (for non-uniform polynomial-time algorithms), implies the existence of non-uniformly hard one-way functions. So the assumption of one-way functions could actually be relaxed to assume that there exists a hard-on-the average language in \mathcal{NP} . Let us also remark that under the assumption of one-way functions secure against non-uniform subexponential-time algorithms, Theorem 3 directly extends also to a super-polynomial-time (SPS) [Pas03b] relaxation of the notion of a statistical NIZK, where the simulator may run in subexponential time. (Let us also briefly point out a very recent work by Chung, Lui, Mohammad and Pass [CLMP12] that presents barriers to *two-message* SPS zero-knowledge arguments.)

Note that in Theorem 3, we rule out statistical NIZK where adaptive soundness only needs to hold w.r.t. statements of a *particular* (polynomial) length $n = q(k)$.

Our next theorem rules out even *exponential-time* challenger assumptions C if the same assumption C can be used to prove adaptive soundness for *any polynomial length statement* (indeed, as far as we know, in all known NIZK constructions, the underlying intractability assumption depends only on the security parameter for the NIZK but not on the length of the statement to be proven).

Theorem 4. *Assume the existence of one-way functions secure against non-uniform subexponential-time algorithms. Then there exists an \mathcal{NP} -language L such that the following holds. Let (\mathcal{D}, P, V) be a statistical non-interactive adaptively zero-knowledge argument for L , and let (C, t) be any exponential-time challenger assumption. If for every polynomial q , there exists a black-box reduction R for basing adaptive soundness of (\mathcal{D}, P, V) w.r.t L and the input length q on the hardness of C w.r.t threshold t , then there exists a probabilistic polynomial-time machine B and a polynomial $p'(\cdot)$ such that for infinitely many $k \in \mathbb{N}$, B breaks C w.r.t t with probability $\frac{1}{p'(k)}$ on input 1^k .*

Note that Theorem 4 is weaker than Theorem 3 in that we require that the same assumption C can be used to prove any polynomial-length statement, whereas in Theorem 3 we rule out NIZK where the underlying hardness assumption may depend also on the length of the statement proved. This additional restriction is necessary: the assumption that a particular NIZK is adaptively sound for statements of length $q(k) = k$ can clearly be stated as an exponential-time challenger assumption.

4.1 Proof of Theorem 3

Proof. We start by considering a simplified case when the zero-knowledge property is *perfect* and the distribution sampled by \mathcal{D} is uniform over $\{0, 1\}^{\text{poly}(k)}$ —i.e., we consider perfect NIZK in the so-called “Uniform Reference String” (URS) model.

Ruling out Perfect NIZK in the URS Model Let $g : \{0, 1\}^* \rightarrow \{0, 1\}^*$ be a length-doubling PRG. Consider the language $L = \{g(s) | s \in \{0, 1\}^*\}$. Assume there exists a perfect NIZK (\mathcal{D}, P, V) for L in the URS model, where the reference string is of length $\ell(k)$ given the security parameter k , and assume there exists a black-box reduction R for basing adaptive soundness of (\mathcal{D}, P, V) w.r.t L and input lengths $q(\cdot)$ on the hardness of C w.r.t threshold t . In particular, this means that

for every A that breaks the adaptive soundness of (\mathcal{D}, P, V) w.r.t L and input lengths $q(\cdot)$ *with overwhelming probability*, there exists a polynomial $p(\cdot)$ such that for infinitely many $k \in \mathbb{N}$, R^A breaks C w.r.t t on common input 1^k with probability $\frac{1}{p(k)}$; i.e.,

$$\Pr \left[\langle R^A, C \rangle(1^k) = 1 \right] \geq t(k) + \frac{1}{p(k)}. \quad (1)$$

To be more concrete, R may feed $A(1^k)$ a reference string ρ , and will get in return a statement $x \in \{0, 1\}^{q(k)}$ —that with high probability is *false*—and a proof π of x —that with high probability is accepting; R may continue this process all throughout its interaction with C . Note that R is required to work even if A is probabilistic, and on each query made by R , A uses fresh random coins. (As we point out in Remark 2, at the cost of a minor complication, the proof can be adapted to work also if only considering reductions that work as long as the attacker is deterministic. However, for simplicity of exposition, we here consider also randomized attackers.)

Our goal is to present a polynomial-time algorithm that directly breaks C without the help of A . Towards this goal, we will first define a particular *randomized* attacker A , and next present an efficient “simulator” \tilde{A} for A , and show that $R^{\tilde{A}}$ still breaks C (w.r.t t).

Let us start by defining the attacker A . To simplify notation, let us assume that $q(k) = 2k$; it is easy to see that the same proof works as long as $q(k)$ is polynomially related to k . On input 1^k and a reference string ρ , A proceeds as follows:

- A first checks that $|\rho| = \ell(k)$; if not, it simply sends back \perp .
- Otherwise, it *uniformly* picks a random tape r such that $S_1(1^k, 1^n)$ outputs ρ , **aux** on input the random tape r . Since, by our assumption, the simulation is *perfect*, every string $\rho \in \{0, 1\}^{\ell(k)}$ is output by $S_1(1^k, 1^n)$ with positive (and the same) probability, so A will succeed in this task. *Note, however, that this step is not necessarily efficient.*
- Next, A uniformly picks a string $x \in \{0, 1\}^n$. Note that, except with probability 2^{-k} , $x \notin L$ (there are 2^{2k} strings, and at most 2^k can be in the range of the PRG g).
- Finally, A runs the simulator $S_2(1^k, 1^n, x, \mathbf{aux})$ to produce the proof π , and outputs (x, π) .

As noted above, with high probability, the statement x picked by A is false. But it remains to argue that the proof π of x output by A is accepting (for the reference string ρ). For now, let us simply assume that the proof is accepting with high probability, and let instead show how to emulate A in polynomial time (thus proving that C can be broken in polynomial time). We will then return to showing that A indeed is a “good” attacker, producing accepting proofs of false statements.

Consider the “simulator”, \tilde{A} that on input 1^k and a reference string ρ , proceeds as follows:

- Just as A , \tilde{A} first check that $|\rho| = \ell(k)$; if not it simply sends back \perp .
- Next, \tilde{A} uniformly picks a string $s \in \{0, 1\}^k$, and lets $x = g(s)$. Note that by definition $x \in L$.
- Finally, \tilde{A} runs the honest prover algorithm $P(1^k, \rho, x, w)$ to produce the proof π , and outputs (x, π) .

The following claim shows that \tilde{A} is a good simulator for A .

Claim 1. For every efficient C and R , there exists a negligible function μ such that for every $k \in N$,

$$\left| \Pr \left[\langle R^{\tilde{A}}, C \rangle(1^k) = 1 \right] - \Pr \left[\langle R^A, C \rangle(1^k) = 1 \right] \right| \leq \mu(k).$$

Proof. As a first attempt to proving the claim, consider a hybrid attacker A' that performs exactly the same steps as A , but samples a *true* statement $x \in L$ in exactly the same way as \tilde{A} (but otherwise runs the simulator, just as A). Note that the only difference between A' and \tilde{A} is that A' provides "simulated" proofs (of true statements), whereas \tilde{A} gives honestly generated proofs. Indeed, it follows from the perfect zero-knowledge property that A' perfectly emulates \tilde{A} . Furthermore, intuitively, it should follow from the fact that true and false statement are indistinguishable (by the pseudorandomness property of g) that A' correctly emulates A . But there is a problem: although, both C and R are efficient, A and A' are not, so *efficiently* contradicting the pseudorandomness property becomes problematic.

To circumvent this problem, we define a carefully ordered sequence of hybrid experiments, and rely on the fact that it is only the first step of A (and A') that is inefficient. With this careful ordering, the inefficient part of A can be dealt with using *non-uniformity* (and thus we finally contradict the pseudorandomness property of g w.r.t. non-uniform polynomial-time algorithms).

More precisely, assume for contradiction that the claim is false. That is, there exists a polynomial p' such that for infinitely many $k \in N$

$$\left| \Pr \left[\langle R^{\tilde{A}}, C \rangle(1^k) = 1 \right] - \Pr \left[\langle R^A, C \rangle(1^k) = 1 \right] \right| \geq \frac{1}{p'(k)}.$$

Let $m(k)$ be an upper-bound on the number of oracle queries by R on input 1^k , and fix a canonical k for which the above happens. Consider a sequence of hybrid experiments $H_0, \dots, H_{m(k)}$, where H_i is defined as the output of $C(1^k)$ after communicating with $R(1^k)$ where the first i oracle queries of R are answered by A , and the remaining ones are answered by the efficient \tilde{A} . Note that $H_0 = \langle R^{\tilde{A}}, C \rangle(1^k)$ and $H_{m(k)} = \langle R^A, C \rangle(1^k)$. It follows that there exists some j such that

$$\left| \Pr [H_{j+1} = 1] - \Pr [H_j = 1] \right| \geq \frac{1}{m(k)p'(k)}.$$

Define another hybrid H'_j which is identically defined to H_j , but where the statement x in the $j + 1$ oracle query is selected as a true statement (just as in H_{j+1}) but we still run the simulation (just as in H_j). It follows directly by the perfect zero-knowledge property of (\mathcal{D}, P, V) that the output of H'_j is identically distributed to the output of H_{j+1} . To reach a contradiction, let us finally argue that the output of H_j is indistinguishable to that of H'_j . Note that up until the point when R receives its $(j + 1)$ st proof back from the oracle, the two experiments proceed identically the same. Thus, if they are distinguishable, there exists some prefix τ of the execution of H_j ⁸, up until and including the $j + 1$ query of R , such that conditioned on this prefix τ , H_j and H'_j are distinguishable. We may now simply extend τ to also include the string aux picked by A in the $j + 1$ query, and conclude that there exists some extension τ' of τ such that even conditioned on τ' , H_j and H'_j are distinguishable. But now, note that given the prefix τ' , the continuations of H_j and H'_j (conditioned on τ') can be efficiently generated. And since the only difference between them is the choice of the statement x , if they can be distinguished, we violate the pseudorandomness property of g . Note that we here require that g is pseudorandom against non-uniform polynomial time (as we need the non-uniform advice τ'). This concludes the proof of Claim 1. \square

⁸Technically, the prefix includes the random tape of C and R and all the answers to the first j queries by R

So conclude the proof of the theorem, it only remains to show that A is a good attacker. Note that by the completeness property of (\mathcal{D}, P, V) it holds that, except with negligible probability, \tilde{A} provides accepting proofs. It now follows as a corollary of Claim 1 that except with negligible probability, A also provides accepting proofs: simply let R be the reduction that picks an honestly generated reference string ρ , and upon receiving back the pair (x, π) , outputs 1 iff $V(1^k, \rho, x, \pi)$ outputs 1, and let C be the algorithm that simply outputs whatever R outputs.

Dealing with Perfect NIZK in the CRS Model Let us turn to considering perfect NIZK (\mathcal{D}, P, V) in the “general” CRS model—that is, the reference string need no longer be uniform. Recall that the only place in the proof above where we used the fact that the reference string is uniform is when arguing that there always exists some string \mathbf{aux} such (ρ, \mathbf{aux}) is output by $S_1(1^k, 1^n)$ on some random tape. This is no longer true in the CRS model. We deal with it as follows. Whenever A receives a query $(1^k, \rho)$, it checks if ρ is in the range of $\mathcal{D}(1^k, 1^n)$: if it is, it proceed just as before; otherwise—we call such strings “bad”— A proceeds just as the efficient \tilde{A} ; that is it picks a true statement x , and honestly computes a proof of it. The proof of Claim 1 remains essentially unchanged; the only difference is that if in the hybrid H_j , conditioned on τ , the reference string τ is “bad”, we trivially have that H_j and H_{j+1} are identical (since A proceeds just as \tilde{A}). It only remains to argue that A is a good attacker: it follows just as before that except with negligible probability, its proofs are always accepting. But, it is no longer clear that it provides proofs of false statements. However, recall that (w.h.p) A only picks true statements when the reference string is “bad”, and to be a good attacker, A is only required to violate soundness in the event that the reference string is honestly picked (and thus not “bad”).

Dealing with Statistical NIZK in the CRS Model Note that the perfect zero-knowledge property is used in two places in the proof. First, we use it to guarantee that $S_1(1^k, 1^n)$ “hits” every reference string in the range of $\mathcal{D}(1^k, 1^n)$. Next, we use it to guarantee that the simulation is statistically close, even conditioned on the reference string ρ , for every ρ in the range of $\mathcal{D}(1^k, 1^n)$. Neither of those properties hold if the zero-knowledge property only is statistical. So, here is how we modify A to deal with them. By the statistical zero-knowledge property of (\mathcal{D}, P, V) , it follows using an averaging argument that there exists some negligible function $\mu'(\cdot)$ such that with probability $1 - \mu'(k)$ over the choice of $\rho \xleftarrow{R} \mathcal{D}(1^k, 1^n)$, it holds that there exists some string \mathbf{aux} such that the simulation is (at least) $(1 - \mu'(k))$ -statistically close for statements x generated as $g(s)$, also conditioned on ρ . Whenever A receives a query $(1^k, \rho)$, it now checks whether it is the case that *conditioned on* ρ being the reference string, S_1, S_2 provides a $1 - \mu'(k)$ -close simulation for a randomly chosen statement $x = g(s)$, and witness $w = s$. If so, A uniformly samples a random tape r such that $S_1(1^k, 1^n)$ outputs ρ, \mathbf{aux} on input the random tape r , and proceed exactly as before. Otherwise, A proceeds just as the efficient \tilde{A} . Claim 1 follows just as before. We just need to argue that A still is a good attacker. As before, by Claim 1, have that the proofs output by A are accepting with overwhelming probability, so we just need to argue that with high probability it provides proofs of false statements. Recall that in the case of Perfect NIZK in the CRS model, we only modified the behavior of A when receiving an invalid reference string ρ as input. Here, its behavior is modified also when receiving valid reference strings. But we only modify its behavior with probability $\mu'(k)$ over the choice of $\rho \xleftarrow{R} \mathcal{D}(1^k, 1^n)$ and thus with overwhelming probability over $\rho \xleftarrow{R} \mathcal{D}(1^k, 1^n)$, A still proves false statements with overwhelming probability.

Ruling out Subexponential-time Challenger Assumptions. If the challenger C is not efficient, then in the above hybrid argument, when switching the statement $x = g(s)$ from being

pseudorandom to being truly random, we can no longer directly argue that the probability of C outputting 1 does not change by much. However, if we could ensure that the pseudorandomness property held against sub exponential-time, then the same proof would go through as long as C is subexponential-time computable. Thus, if we assume the existence of one-way functions secure against subexponential time, we prove the theorem also when C is subexponential-time computable. \square

Remark 1. On Adaptive Culpable Soundness *Groth, Ostrovsky and Sahai [GOS06] also define a weaker definition of adaptive soundness which they call adaptive culpable soundness. Roughly speaking, 1) we restrict to languages in $\mathcal{NP} \cap \text{co}\mathcal{NP}$, and 2) require a successful attacker to not only prove a false statement, but also provide a an \mathcal{NP} proof of the fact that the statement is true.*

Let us remark that simply restricting to languages in $\mathcal{NP} \cap \text{co}\mathcal{NP}$ does not suffice for overcoming our lower-bound: assuming the existence of one-way permutations, our impossibility result rules out statistical NIZK with adaptive soundness also for $\mathcal{NP} \cap \text{co}\mathcal{NP}$: by the Blum-Micali-Goldreich-Levin construction of a pseudo random generator (see [Gol01]), the existence of one-way permutations implies the existence of a hard-on-the-average language in $\mathcal{NP} \cap \text{co}\mathcal{NP}$, which suffices to concluding the theorem.

It is also worthwhile to note why our proof strategy breaks down in case we required the attacker to also provide a witness for false statements (as in the definition of adaptive culpable soundness): a key component in our proof is that the attacker chooses statements that are false but look indistinguishable from true statements (and this is exactly the same method as the one used in [GW11]). In case the attacker also needs to provide an \mathcal{NP} witness for the statement being false, we can no longer simulate such an attacker by using a true statement.

Remark 2. Deterministic v.s. Randomized Attackers. *In the above proof, we consider a randomized oracle A , and thus only rule out reductions that work for randomized attackers. Following [Pas11], we can easily extend the proof to also rule out reductions that only work for deterministic attackers. First, if we consider a deterministic attacker, we may assume w.l.o.g. that R never asks the same query $(1^k, \rho)$ twice (since we can internally emulate responses to all repeated queries). Next, we define a deterministic attacker A^f that on input $(1^k, \rho)$ selects the random tape of A as $f(1^k, \rho)$ and next executes $A(1^k, \rho)$, as defined above, using the pre-selected random tape. Let $RO : \{0, 1\}^* \rightarrow \{0, 1\}^\infty$ be a uniformly distributed random oracle. Note that A^{RO} acts exactly as the attacker A defined in our proof as long as we never ask the attacker the same query twice. It follows that R 's view when talking to A^{RO} and A are identical. We can thus replace A with A^{RO} in Claim 1. Finally, by an averaging argument, with overwhelming probability over the choice of a random oracle $f \leftarrow RO$, A^f breaks adaptive soundness of (\mathcal{D}, P, V) with overwhelming probability, and for each such "good" choice of f we have that $R^{A^f}(1^k)$ breaks C with advantage $1/p(k)$, where $p(\cdot)$ is a polynomial; thus $R^{A^{RO}}(1^k)$ also breaks C with advantage negligibly close to $\frac{1}{p(k)}$.*

Remark 3. On Non-uniform Reductions *Our current lower bound only considers uniform security reductions R . However, by using the recent techniques of [CLMP13] (and relying on the fact that we can rule out reductions that only need to work for deterministic attackers), it readily extends also to rule out non-uniform reductions. We refer the reader to [CLMP13] for further details.*

4.2 Proof of Theorem 4

The proof of theorem 4 is essentially identical to the proof of theorem 3. The only obstacle we need to deal with is to ensure that true and false statements are indistinguishable for time 2^k . This is easily achieved if 1) assuming the existence of a pseudorandom generator secure against time 2^{n^ϵ} where n is its output length, and 2) letting A prove statements of length $q(k) = k^{\frac{1}{\epsilon}}$. Note that we here rely on the fact that the same assumption C (with the same security parameter k) can be used to prove adaptive soundness of *any* polynomial length statement.

5 Security of Non-interactive Non-malleable Commitments

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. In the rest of the paper, a commitment scheme always refers to a statistically-binding commitment.

Following [PR03, 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—id as common input.

Let us turn to defining non-malleable commitments. We use a definition essentially due to [PR05b, PR05a], which follows the earlier definition from [DDN00]. Let $\langle C, R \rangle$ be a tag-based commitment scheme, and let $k \in N$ be a security parameter. Consider a man-in-the-middle adversary A that, on inputs n and z (where z is received as an auxiliary input), participates in one “left” and one “right” interaction. In the left interaction, the man-in-the-middle adversary A interacts with C , receiving a commitment to the value v using an identity id of length $\ell(k)$ of its choice. In the right interaction A interacts with R attempting to commit a value \tilde{v} , again an identity $\tilde{\text{id}}$ of length $\ell(k)$ of its choice. If the right commitment is invalid, or undefined, its value is set to \perp ; furthermore, if the adversary uses the same identity on the left as on the right, the right interaction is considered invalid. Let $\text{MIM}^\Pi(A, \ell, v, z)$ denote a random variable that describes the value \tilde{v} in the above experiment.⁹

Definition 8. *A commitment scheme Π is said to be non-malleable (with respect to itself) for identities of length $\ell(\cdot)$ if for every \mathcal{PPT} man-in-the-middle adversary A the following ensembles are computationally indistinguishable.*

$$\begin{aligned} & \left\{ \text{MIM}^\Pi(A, \ell, v, z) \right\}_{k \in N, v \in \{0,1\}^n, z \in \{0,1\}^*} \\ & \left\{ \text{MIM}^\Pi(A, \ell, v, z) \right\}_{k \in N, v \in \{0,1\}^n, z \in \{0,1\}^*} \end{aligned}$$

We say that Π is non-malleable (with respect to itself) if it is non-malleable for identities of length $\ell(k)$.

Our lower bound will apply to non-malleability with respect to identities of length 1, and even if only considering the messages 0^k and 1^k . We now explicitly define what it means to break non-malleability *in this particular way*.

⁹We note that more recent definitions of non-malleability [LPV08, LP09] additionally output the view of A in the above experiment. Since we are proving a lower-bound we simply state the weaker definition.

Definition 9 (Breaking Non-malleability). Let Π be a tag-based commitment scheme. We say that A breaks non-malleability of Π with probability $\mu(\cdot)$ if for every $n \in N$,

$$|\Pr [\text{MIM}^\Pi(A, \ell, 0^k, \perp) = 0^k] - \Pr [\text{MIM}^\Pi(A, \ell, 1^k, \perp) = 0^k]| > \mu(k)$$

where $\ell(k) = 1$. We furthermore say that A breaks one-sided non-malleability of Π with probability $\mu(\cdot)$ if the above holds and A always picks identity 0 for the left interaction and identity 1 for the right one.

We call a commitment *weakly non-malleable* if no attacks of the above kind exist.

Definition 10 (Basing Weak Non-malleability on the Hardness of C). We say that R is a $T(\cdot)$ -black-box reduction for basing weak non-malleability of Π (resp. weak one-sided non-malleability) on the hardness of C w.r.t threshold $t(\cdot)$ if R is a time $T(\cdot)$ probabilistic oracle machine and there exists a polynomial $p(\cdot, \cdot)$ such that for every probabilistic machine A that breaks non-malleability (resp. one-sided non-malleability) of Π with probability $\mu(\cdot)$, for every $k \in N$, R^A breaks C w.r.t t with probability $p(\mu(k), 1/k)$ on input 1^k .

We are now ready to state our first lower bound for non-malleable commitments.

Theorem 5. Let Π be a two-round tag-based commitment scheme (i.e., the commit-phase consists of a single message from the receiver, followed by a single message from the committer), and let (C, t) be any efficient challenger assumption. If there exists a probabilistic polynomial-time black-box reduction R for basing weak one-sided non-malleability of Π on the hardness of C w.r.t threshold t , then there exists a probabilistic polynomial-time machine B and a polynomial $p'(\cdot)$ such that for infinitely many $k \in N$, B breaks C w.r.t t with probability $\frac{1}{p'(k)}$ on input 1^k .

Proof. Consider an unbounded attacker A that chooses identity 0 in the left interaction and 1 in the right and upon receiving a commitment to the value b^k commits to b^k , and otherwise (e.g., if the commitment is invalid) simply commits to 0^k . (Note that implementing A requires super-polynomial time by the hiding property of Π .) Now consider \tilde{A} that picks the same identities, but simply commits to 0^k in the right interaction (no matter what messages it receives on the left). We claim that $R^{\tilde{A}}$ breaks C w.r.t t with inverse polynomial probability for infinitely many k . Assume not; that is, there exists a negligible function μ such that for every $k \in N$, $R^{\tilde{A}}$ breaks C w.r.t t with probability at most $\mu(k)$. Consider a fixed k , and hybrids $H_1, \dots, H_{m(k)}$ where H_i denotes the success probability of R when the first i queries are answered by A and the remaining ones answered by \tilde{A} . By the triangle inequality there exists some i such $|H_i - H_{i+1}|$ is inverse polynomial (where the polynomial only depends on R). Since H_i and H_{i+1} proceed identically up until query i , this means there exists some prefix ρ of the executions up until query i such that 1) conditioned on ρ the gap in probability is as high, and 2) the executions of H_i and H_{i+1} are efficiently computable conditioned on this prefix (recall that \tilde{A} is efficient, so once we fix all the answers of A in all the first $i + 1$ queries, the rest of the experiment is efficient). Since the only difference between the hybrids is the commitment received by R in round $i + 1$, we contradict the *non-uniform* hiding of Π . Note that we here rely on the fact that Π only has two rounds (or else hiding may no longer hold since R could be rewinding its oracle), and that C is polynomial-time computable. \square

We remark that the above theorem is tight; we cannot hope to rule out also super-polynomial-time reductions for one-sided non-malleability: Liskov et al [LLM⁺01] present a construction of such commitments assuming the existence of one-way permutations with subexponential security.

Let us now turn to ruling out also super polynomial-time reductions for “two-sided” non-malleability (i.e., where the attacker may choose which identity to use).

Theorem 6. *Let Π be a two-round tag-based commitment scheme (i.e., the commit-phase consists of a single message from the receiver, followed by a single message from the committer), and let (C, t) be any $T(\cdot)$ -size computable assumption. If there exists a $T(\cdot)$ -size randomized black-box reduction R for basing weak non-malleability of Π on the hardness of C w.r.t threshold t , then there exists a $\text{poly}(T(\cdot))$ -sized attacker B and a polynomial $p'(\cdot)$ such that for infinitely many $k \in N$, B breaks C w.r.t t with probability $\frac{1}{p'(k)}$ on input 1^k .*

Proof. Consider the attacker A, \tilde{A} from the proof of Theorem 5. We consider two cases: Either $R^{\tilde{A}}$ breaks C w.r.t t with inverse polynomial probability for infinitely many k (as in the proof of Theorem 5) in which case we are done. Or, there exists a negligible function μ such that for every $k \in N$, $R^{\tilde{A}}$ breaks C w.r.t t with probability at most $\mu(k)$. As shown in the proof of Theorem 5, in this case R together with C and using a polynomial-size advice string may distinguish commitment to 0^k and 1^k with inverse polynomial probability for infinitely many k ; let D denote this “commitment distinguisher” (outputting a bit $b \in \{0, 1\}$). Now consider a third attacker A' that chooses identity 1 on the left and 0 on the right (instead of using 0 on the left and 1 on the right as A and \tilde{A}) and next applies D to the commitment it receives on the left in order to decide what value to commit to on the right. More formally, $A'(1^k)$ proceeds as follows:

- A' internally emulates $D(1^n)$, outputs the first output of D (the first message of the commitment¹⁰);
- Upon receiving a commitment c on the left, and a first message r on the right, A' simply forwards c to D ;
- Let b denote the final bit output by D ; output on the right a commitment to b^n (i.e., 0^n if $b = 0$ and 1^n if $b = 1$) using r as a first message.

It directly follows from the fact that D is a good distinguisher that A' breaks non-malleability of Π , and thus there exists a polynomial $p'(\cdot)$ such that for infinitely many $k \in N$, $R^{A'}$ breaks C w.r.t t with probability $\frac{1}{p'(k)}$ on input 1^k . The theorem follows by noting that both R and D are computable in size $\text{poly}(T(\cdot))$. \square

A Remark on Deterministic Attackers and Non-uniform Reductions Just as the proof of Theorem 3, the proofs of the above theorem readily extends to rule out reductions that only work with deterministic attackers, and, relying on the techniques from [CLMP13], also non-uniform reductions.

6 Acknowledgements

I am extremely grateful to Kai-min Chung and Mohammad Mahmoody for many helpful comments and definitional discussions.

References

- [AF07] Masayuki Abe and Serge Fehr. Perfect nizk with adaptive soundness. In *TCC*, pages 118–136, 2007.

¹⁰Recall that D is a commitment distinguisher for two-round commitment, and thus must first specify the first message of the commitment.

- [AGGM06] Adi Akavia, Oded Goldreich, Shafi Goldwasser, and Dana Moshkovitz. On basing one-way functions on NP-hardness. In *STOC '06*, pages 701–710, 2006.
- [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.
- [BFM88] Manuel Blum, Paul Feldman, and Silvio Micali. Non-interactive zero-knowledge and its applications (extended abstract). In *STOC*, pages 103–112, 1988.
- [BMV08] Emmanuel Bresson, Jean Monnerat, and Damien Vergnaud. Separation results on the "one-more" computational problems. In *CT-RSA*, pages 71–87, 2008.
- [BNPS03] Mihir Bellare, Chanathip Namprempre, David Pointcheval, and Michael Semanko. The one-more-rsa-inversion problems and the security of chaum's blind signature scheme. *J. Cryptology*, 16(3):185–215, 2003.
- [BP02] Mihir Bellare and Adriana Palacio. Gq and schnorr identification schemes: Proofs of security against impersonation under active and concurrent attacks. In *CRYPTO*, pages 162–177, 2002.
- [BR93] Mihir Bellare and Phillip Rogaway. Random oracles are practical: A paradigm for designing efficient protocols. In *ACM Conference on Computer and Communications Security*, pages 62–73, 1993.
- [Bra83] Gilles Brassard. Relativized cryptography. *IEEE Transactions on Information Theory*, 29(6):877–893, 1983.
- [BT03] Andrej Bogdanov and Luca Trevisan. On worst-case to average-case reductions for np problems. In *FOCS*, pages 308–317, 2003.
- [BV98] Dan Boneh and Ramarathnam Venkatesan. Breaking rsa may not be equivalent to factoring. In *EUROCRYPT*, pages 59–71, 1998.
- [BY96] Mihir Bellare and Moti Yung. Certifying permutations: Noninteractive zero-knowledge based on any trapdoor permutation. *J. Cryptology*, 9(3):149–166, 1996.
- [CGH04] Ran Canetti, Oded Goldreich, and Shai Halevi. The random oracle methodology, revisited. *J. ACM*, 51(4):557–594, 2004.
- [CIO98] Giovanni Di Crescenzo, Yuval Ishai, and Rafail Ostrovsky. Non-interactive and non-malleable commitment. In *STOC*, pages 141–150, 1998.
- [CLMP12] Kai-min Chung, Edward Lui, Mohammad Mahmoody, and Rafael Pass. Unprovable security of two-message zero-knowledge. Manuscript, 2012.
- [CLMP13] Kai-min Chung, Huijia Lin, Mohammad Mahmoody, and Rafael Pass. On the power of non-uniform proof of security. In *ITCS'13*, 2013.
- [Dam91] Ivan Damgård. Towards practical public key systems secure against chosen ciphertext attacks. In *CRYPTO*, pages 445–456, 1991.

- [DDN00] Danny Dolev, Cynthia Dwork, and Moni Naor. Nonmalleable cryptography. *SIAM Journal on Computing*, 30(2):391–437, 2000.
- [DOP05] Yevgeniy Dodis, Roberto Oliveira, and Krzysztof Pietrzak. On the generic insecurity of the full domain hash. In *CRYPTO*, pages 449–466, 2005.
- [FF93] Joan Feigenbaum and Lance Fortnow. Random-self-reducibility of complete sets. *SIAM Journal on Computing*, 22(5):994–1005, 1993.
- [FLS90] Uriel Feige, Dror Lapidot, and Adi Shamir. Multiple non-interactive zero knowledge proofs based on a single random string. In *FOCS '90*, pages 308–317, 1990.
- [FS87] Amos Fiat and Adi Shamir. How to prove yourself: practical solutions to identification and signature problems. In *Proceedings on Advances in cryptology—CRYPTO '86*, pages 186–194, London, UK, 1987. Springer-Verlag.
- [FS10] Marc Fischlin and Dominique Schröder. On the impossibility of three-move blind signature schemes. In *EUROCRYPT*, pages 197–215, 2010.
- [GK03] Shafi Goldwasser and Yael Tauman Kalai. On the (in)security of the fiat-shamir paradigm. In *FOCS '03*, pages 102–111, 2003.
- [GM84] Shafi Goldwasser and Silvio Micali. Probabilistic encryption. *J. Comput. Syst. Sci.*, 28(2):270–299, 1984.
- [Gol01] Oded Goldreich. *Foundations of Cryptography — Basic Tools*. Cambridge University Press, 2001.
- [GOS06] Jens Groth, Rafail Ostrovsky, and Amit Sahai. Perfect non-interactive zero knowledge for np. In *EUROCRYPT*, pages 339–358, 2006.
- [Goy11] Vipul Goyal. Constant round non-malleable protocols using one way functions. In *STOC*, pages 695–704, 2011.
- [GW11] Craig Gentry and Daniel Wichs. Separating succinct non-interactive arguments from all falsifiable assumptions. In *STOC*, pages 99–108, 2011.
- [HH09] Iftach Haitner and Thomas Holenstein. On the (im)possibility of key dependent encryption. In *TCC*, pages 202–219, 2009.
- [HRS09] Iftach Haitner, Alon Rosen, and Ronen Shaltiel. On the (im)possibility of arthur-merlin witness hiding protocols. In *TCC '09*, pages 220–237, 2009.
- [Imp95] Russell Impagliazzo. A personal view of average-case complexity. In *Structure in Complexity Theory '95*, pages 134–147, 1995.
- [IR88] Russell Impagliazzo and Steven Rudich. Limits on the provable consequences of one-way permutations. In *CRYPTO '88*, pages 8–26, 1988.
- [LLM⁺01] Moses Liskov, Anna Lysyanskaya, Silvio Micali, Leonid Reyzin, and Adam Smith. Mutually independent commitments. In *ASIACRYPT*, pages 385–401, 2001.
- [LP09] Huijia Lin and Rafael Pass. Non-malleability amplification. In *STOC '09*, pages 189–198, 2009.

- [LP11] Huijia Lin and Rafael Pass. Constant-round non-malleable commitments from any one-way function. In *STOC*, pages 705–714, 2011.
- [LPV08] Huijia Lin, Rafael Pass, and Muthuramakrishnan Venkatasubramanian. Concurrent non-malleable commitments from any one-way function. In *TCC '08*, pages 571–588, 2008.
- [Nao03] Moni Naor. On cryptographic assumptions and challenges. In *CRYPTO*, pages 96–109, 2003.
- [Ost91] Rafail Ostrovsky. One-way functions, hard on average problems, and statistical zero-knowledge proofs. In *Structure in Complexity Theory Conference*, pages 133–138, 1991.
- [OW93] Rafail Ostrovsky and Avi Wigderson. One-way functions are essential for non-trivial zero-knowledge. In *Theory and Computing Systems, 1993*, pages 3–17, 1993.
- [Pas03a] Rafael Pass. On deniability in the common reference string and random oracle model. In *CRYPTO*, pages 316–337, 2003.
- [Pas03b] Rafael Pass. Simulation in quasi-polynomial time, and its application to protocol composition. In *EUROCRYPT*, pages 160–176, 2003.
- [Pas06] Rafael Pass. Parallel repetition of zero-knowledge proofs and the possibility of basing cryptography on np-hardness. In *IEEE Conference on Computational Complexity*, pages 96–110, 2006.
- [Pas11] Rafael Pass. Limits of provable security from standard assumptions. In *STOC*, pages 109–118, 2011.
- [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.
- [PR03] Rafael Pass and Alon Rosen. Bounded-concurrent secure two-party computation in a constant number of rounds. In *FOCS '03*, pages 404–, 2003.
- [PR05a] Rafael Pass and Alon Rosen. Concurrent non-malleable commitments. In *FOCS '05*, 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.
- [PS05] Rafael Pass and Abhi Shelat. Unconditional characterizations of non-interactive zero-knowledge. In Victor Shoup, editor, *CRYPTO*, volume 3621 of *Lecture Notes in Computer Science*, pages 118–134. Springer, 2005.
- [PTV11] Rafael Pass, Wei-Lung Dustin Tseng, and Muthuramakrishnan Venkatasubramanian. Towards non-black-box lower bounds in cryptography. In *TCC*, pages 579–596, 2011.
- [PW10] Rafael Pass and Hoeteck Wee. Constant-round non-malleable commitment from strong one-way functions. In *Eurocrypt '10*, 2010.

- [RTV04] Omer Reingold, Luca Trevisan, and Salil P. Vadhan. Notions of reducibility between cryptographic primitives. In *TCC*, pages 1–20, 2004.
- [RV10] Guy N. Rothblum and Salil P. Vadhan. Are peeps inherent in efficient arguments? *Computational Complexity*, 19(2):265–304, 2010.
- [Wee10] Hoeteck Wee. Black-box, round-efficient secure computation via non-malleability amplification. To appear in FOCS 2010, 2010.