Private Coins versus Public Coins in Zero-Knowledge Proof Systems

Rafael Pass* and Muthuramakrishnan Venkitasubramaniam

Cornell University,
{rafael,vmuthu}@cs.cornell.edu

Abstract. Goldreich-Krawczyk (Siam J of Comp'96) showed that only languages in BPP have constant-round *public-coin* black-box zero-knowledge protocols. We extend their lower bound to "fully black-box" *private-coin* protocols based on one-way functions. More precisely, we show that only languages in $\mathsf{BPP}^{\mathsf{Sam}}$—where Sam is a "collision-finding" oracle in analogy with Simon (Eurocrypt'98) and Haitner et. al (FOCS'07)—can have constant-round fully black-box zero-knowledge proofs; the same holds for constant-round fully black-box zero-knowledge *arguments* with sublinear verifier communication complexity. We also establish near-linear lower bounds on the round complexity of fully black-box concurrent zero-knowledge proofs (or arguments with sublinear verifier communication) for languages outside $\mathsf{BPP}^{\mathsf{Sam}}$.
The technique used to establish these results is a transformation from private-coin protocols into Sam-relativized public-coin protocols; for the case of fully black-box protocols based on one-way functions, this transformation preserves zero knowledge, round complexity and communication complexity.

1 Introduction

Roughly speaking, interactive proofs, introduced by Goldwasser, Micali and Rackoff [9] and Babai and Moran [1]), are protocols that allow one party P—called the *Prover* (or Merlin)—to convince a computationally-bounded party V—called the *Verifier* (or Arthur)—of the validity of some statement $x \in L$. While, the notion of interactive proofs introduced by Goldwasser, Micali and Rackoff considers arbitrary probability polynomial time verifiers, the notion introduced by Babai and Moran, called *Arthur-Merlin games* considers verifiers that only send truly random messages; such proof systems are also called *public coin*. Soon after their introduction, a surprisingly result by Goldwasser and Sipser [11] showed that the two notions in fact are equivalent in their expressive power: Any private coin protocol $\langle P, V \rangle$ for a language L can be transformed into a public-coin $\langle \widehat{P}, \widehat{V} \rangle$ for L with the same round-complexity. Their result has played an important role in subsequent complexity-theoretic work. However,

* Supported in part by a Microsoft New Faculty Fellowship, NSF CAREER Award CCF-0746990, AFOSR Award FA9550-08-1-0197 and BSF Grant 2006317

from a cryptographic perspective, the transformation is somewhat unsatisfactory as it does not preserve the efficiency of the prover—and can thus not be applied to "computationally-sound" protocols (a.k.a. *arguments*)—or properties such as *zero-knowledge*—the principal notion introduced in [9]. By a result of Vadhan [26], any transformation that uses the original private-coin protocol $\langle P, V \rangle$ as a black-box, in fact, must require the prover to run in super-polynomial time.

In this work, we provide different and "robust" transformations from private-coin protocols to public-coin protocols. Our transformations preserve zero-knowledge, computational and communication complexity, but instead require the prover and the verifier to have oracle access to a certain "collision-finding" oracle [25, 13], denoted Sam. Our transformation is black-box and thus by Vadhan's results we are required to use a super-polynomial time oracle. Nevertheless, the Sam oracle is not "too" powerful; in particular, as shown by Haitner, Hoch, Reingold and Segev [13] it cannot be used to invert *one-way functions*. Therefore, if the security properties (namely, zero-knowledge and computational soundness) of the private-coin protocol are based on the hardness of inverting one-way functions (or even *trapdoor permutations*), we can use our transformation to extended lower bounds for public-coin protocols to private-coin protocols.

More precisely, Goldreich and Krawczyk [8] showed that only languages in BPP can have constant-round public-coin *black-box* zero-knowledge protocols. Recently, Pass, Tseng and Wikström [21] extended this results to include all (even super-constant round) black-box zero-knowledge protocols that remain secure under concurrent (or even parallel) composition (a.k.a *concurrent zero-knowledge* protocols). Combining our transformation with these results, we obtain new lower bounds for *fully black-box constructions* of general, potentially private-coin, black-box zero-knowledge protocols based on the existence of one-way permutations.

Theorem 1 (Lower Bounds for Fully Black-Box Zero Knowledge—Informally stated). *Let $\langle P, V \rangle$ be a fully black-box construction of a zero-knowledge proof (or argument) for the language L from one-way permutations. Then, $L \in \mathsf{BPP}^{\mathsf{Sam}}$ if any of the following hold:*

1. *$\langle P, V \rangle$ is an $O(1)$-round proof.*
2. *$\langle P, V \rangle$ is an $O(1)$-round argument with $o(n)$ verifier communication complexity.*
3. *$\langle P, V \rangle$ is an $o(\frac{n}{\log n})$-round concurrent zero-knowledge proof.*
4. *$\langle P, V \rangle$ is an $o(\frac{n}{\log n})$-round concurrent zero-knowledge argument with $o(n)$ verifier communication complexity.*

We remark that all the above type of protocols can be achieved for languages in NP, assuming the existence of *collision-resistant hash-functions* [7, 17, 20]. Assuming only one-way permutations, however, the best zero-knowledge proofs require a super-constant number of rounds [10], and $O(n/\log n)$-rounds for concurrent zero-knowledge [22]. As such, assuming $\mathsf{NP} \not\subseteq \mathsf{BPP}^{\mathsf{Sam}}$, Theorem 1 is tight.

In Section 3, we discuss the complexity of $\mathsf{BPP}^{\mathsf{Sam}}$. We observe that the class SZK, of languages having *statistical* zero-knowledge proofs, is contained in $\mathsf{BPP}^{\mathsf{Sam}}$. This should not be surprising as Ong and Vadhan provide unconditional constructions of constant-round black-box zero-knowledge proofs for languages in SZK [19]. By extending the result of [13] we also observe that $\mathsf{BPP}^{\mathsf{Sam}}$ does not "generically" decide all NP languages, and seems thus like an interesting and natural complexity class in its own right.

We finally mention that the techniques used in our transformation are interesting in their own right. First, it directly follows that that there is no fully black-box construction of a one-way function, that compresses its input by more than a constant factor, from one-way permutations. Next, as pointed out to us by Haitner, it would seem that by our techniques, the black-box lower bounds from [14] can be extended also to honest-but-curious protocols; see the proof of Lemma 2 for more details.

2 Preliminaries and Definitions

We assume familiarity with the basic notions of an Interactive Turing Machine (ITM for brevity) and a protocol (in essence a pair of ITMs). We denote by $\mathcal{PPT}$ the class of probabilistic polynomial time Turing machines and $n.u.\mathcal{PPT}$, the class of non-uniform $\mathcal{PPT}$ machines. We denote by $M^\bullet$ an oracle machine; we sometimes drop $\bullet$ when it is clear from the context. As usual, if $M^\bullet$ is an oracle machine, M^O denotes the joint execution of M with oracle access to O. Let $\mathcal{O}$ be a random variable over functions from $\{0,1\}^* \to \{0,1\}^*$. Then, $M^{\mathcal{O}}$ denotes the execution of M^O, where O is sampled according to $\mathcal{O}$. Let $\mathit{\Pi}_n$ denote the set of all permutations on $\{0,1\}^n$ and $\mathit{\Pi}$ denote the set of all permutations $\{0,1\}^* \to \{0,1\}^*$ (obtained by choosing a π_n from $\mathit{\Pi}_n$ for every n).

2.1 Fully Black-Box Constructions

A construction of a cryptographic primitive p from a primitive q is said to be *fully black-box* if both the implementation and the proof of correctness are black-box. (See [24] for more details on black-box constructions and reductions.) Here, we focus on fully black-box constructions from one-way permutations. For simplicity, we show our results only for one-way permutations, but analogous to [13], our results extend to trapdoor permutations as well. We proceed to define fully black-box constructions of arguments and zero-knowledge.

Definition 1 (Fully black-box interactive arguments) *Let $\langle P^\bullet, V^\bullet \rangle$ be an interactive argument for a language $L \subseteq \{0,1\}^*$. We say that $\langle P^\bullet, V^\bullet \rangle$ is a fully black-box construction from one-way permutations, if there exists a $\mathcal{PPT}$ machine $A^\bullet$, and a polynomial $q(\cdot)$ such that for every permutation $\pi = \{\pi_n\}_{n=1}^\infty$, malicious prover $P^{*\bullet}$, sequence $\{x_n\}_{n=1}^\infty$ where $x_n \in \bar{L} \cap \{0,1\}^n$ and polynomial $p(\cdot)$, if $\Pr[\langle P^\pi, V^\pi \rangle(x_n) = 1] \geq \frac{1}{p(n)}$ for infinitely many n, then*

$$Pr[A^{\pi, P^{*\pi}(x_n)}(1^n, y) = \pi_n^{-1}(y)] > \frac{1}{q((p(n))}$$

for infinitely many n, where the probability is taken uniformly over $y \in \{0,1\}^n$ *(and over all the internal coin tosses of A).*

Definition 2 (Fully black-box computational zero-knowledge) *Let* $\langle P^\bullet, V^\bullet \rangle$ *be an interactive proof (or argument) system for a language L. We say that* $\langle P^\bullet, V^\bullet \rangle$ *is a* fully black-box *construction of a computational zero-knowledge proof (or argument) from one-way permutations, if there exists an expected* $\mathcal{PPT}$ *simulator* $S^\bullet$*, a* $\mathcal{PPT}$ *machine* $A^\bullet$*, and a polynomial* $q(\cdot)$ *such that for every permutation* $\pi = \{\pi_n\}_{n=1}^\infty$*, a distinguisher D, malicious verifier* $V^{*\bullet}$*, sequence* $\{(x_n, z_n)\}_{n=1}^\infty$ *where* $x_n \in L \cap \{0,1\}^n$*,* $z_n \in \{0,1\}^*$ *and polynomial* $p(\cdot)$*, if for infinitely many n, D distinguishes* $\left\{S^{\pi, V_r^{*\pi}(x_n, z_n)}(x_n)\right\}$ *and* $\left\{\langle P^\pi, V_r^{*\pi}(z_n)\rangle(x_n)\right\}$ *with probability at least* $\frac{1}{p(n)}$ *where* $\langle P^\pi, V_r^{*\pi}(z)\rangle(x)$ *denotes the output of* $V^{*\pi}$ *in an interaction between* P^π *and* $V_r^{*\pi}(z)$ *on common input x, then*

$$Pr[A^{\pi, V_r^{*\pi}(x_n, z_n)}(1^n, y) = \pi_n^{-1}(y)] > \frac{1}{q(p((n))}$$

for infinitely many values of n, where the probability is taken uniformly over $y \in \{0,1\}^n$ *(and over all the internal coin tosses of A).*

Remark 1. Note that in Definition 2, the simulator S unconditionally runs in expected polynomial time. One can consider a weaker definition where A is required to invert π when the expected running time of S exceeds polynomial time. For simplicity (and due to the fact that all known black-box zero-knowledge proofs satisfy this property), we consider the stronger definition, but our results extend also to the weaker definition.

A fully black-box construction of a computational zero-knowledge arguments refers to a construction that is a fully black-box construction in the argument sense and the zero-knowledge sense.

3 The Collision Finding Class

Our transformation makes use of a "collision finding" oracle. Such an oracle was introduced by Simon [25]. In this work, we require a slightly stronger oracle that finds "collisions" in interactive protocols. Such an oracle—referred to as Sam—was recently introduced by Haitner, Hoch, Reingold and Segev [13]. The oracle comes with a permutation π and a parameter d; the depth parameter d denotes the number of rounds in the protocol on which it finds collisions. We denote the oracle by Sam_d^π. Below, we recall the Sam oracle from [13].

3.1 The Oracle Sam_d^π

Informally, Sam_d^π is an oracle, that takes as input a probabilistic interactive turing machine (ITM) $M^\bullet$ and a partial transcript *trans* of an interaction with M of d or fewer rounds, and

– If $trans$ was an output of a previous query, Sam samples a random tape τ for M^π among all random tapes that are consistent with $trans$, and generates M^π's next message m using τ and outputs $trans :: m$.
– Otherwise, outputs $\perp$.

Description of $\mathsf{Sam}^\pi_{d(n)}$: Let $\pi = \{\pi_n\}_{n=1}^\infty$ be a permutation and $M^\bullet$ be a probabilistic oracle ITM that runs a d-round protocol and has access to π. Let $trans_i = (a_1, b_1, \ldots, a_i, b_i)$ be a partial transcript of the messages exchange with M^π in an execution; Define $R_{trans_i}(M^\pi)$ to be the set of all random tapes τ for which $M^\pi_\tau(a_1, b_1, \ldots, b_{j-1}) = a_j$ for all $j < i$; we say that such a τ is *consistent* w.r.t $trans_i$. Without loss of generality, we assume that M^π sends the first message (i.e. outputs a message on initiation). An input query for $\mathsf{Sam}^\pi_{d(n)}$ is of the form $Q = (M^\pi, trans_i, r)$ where $trans_{i-1} = (a_1, b_1, \ldots, b_{i-1})$ and $r \in \{0,1\}^*$. It outputs $(\tau', trans_{i-1} :: a_i)$ such that $\tau' \in R_{trans_j}(M^\pi)$ and $M^\pi_{\tau'}(trans_i) = a_i$, with the following restrictions:

1. If $i > 1$, then $(a_1, b_1, \ldots, a_{i-1})$ was the result of a previous query of the form $(M^\pi, (a_1, b_1, \ldots, b_{i-2}), r')$ for some $r' \in \{0,1\}^*$.
2. τ' is uniformly distributed in $R_{trans_{i-1}}(M^\pi)$ over the randomness of $\mathsf{Sam}^\pi_{d(n)}$, independent of all other queries.
3. $\mathsf{Sam}^\pi_{d(n)}$ answers queries only up to a depth $d(n)$, i.e. $i \leq d(n)$.

Otherwise, it outputs $\perp$. We remark that the role of r in the query is to obtain new and independent samples for each r and allow a verifier to obtain the same sample query by querying on the same r.

Our above description of the $\mathsf{Sam}^\pi_{d(n)}$-oracle is a stateful instantiation of the oracle defined in [13]. Just as in [13], for our results, we need the oracle to be stateless; [13] specify how to modify the oracle to achieve this (using "signatures"); we omit the details. It was shown by Haitner et. al that random permutations are hard to invert for polynomial time machines that query Sam oracle upto depth $o(\frac{n}{\log n})$.

Theorem 2 ([13]). *For every $\mathcal{PPT}$ machine $A^\bullet$, there exists a negligible function $\nu(\cdot)$, such that, for all n, $\Pr[A^{\pi, \mathsf{Sam}^\pi_{o(\frac{n}{\log n})}}(y) = \pi_n^{-1}(y)] \leq \nu(n)$ where the probability is taken uniformly over the randomness of $\mathsf{Sam}^\pi_{o(\frac{n}{\log n})}$, random permutation $\pi = \{\pi_n\}_{n=1}^\infty$ and $y \in \{0,1\}^n$.*

Looking ahead, in Section 4.2, we show that this result is optimal w.r.t the depth: $\mathsf{Sam}^\pi_{\frac{n}{\log n}}$ can be used to invert π.

3.2 The complexity class CF_d

We introduce a new complexity classes CF_d, which we call the "collision-finding class", that we use as part of our characterization of zero-knowledge protocols.

Definition 3 *A language $L \in \mathsf{CF}_d = \mathsf{BPP}^{\mathsf{Sam}_d}$, if there exists a $\mathcal{PPT}$ machine $M^\bullet$ such that:*

Completeness: *If $x \in L$, $M^{\pi,\mathsf{Sam}^{\pi}_d}$ outputs 1 with probability at least $\frac{2}{3}$.*
Soundness: *If $x \notin L$, $M^{\pi,\mathsf{Sam}^{\pi}_d}$ outputs 1 with probability at most $\frac{1}{3}$.*

where both the probabilities are taken uniformly over the random coins of M, the randomness of Sam and a random permutation π.

The complexity class CF_d seems to be interesting classes that lies between P and NP. Below we state some properties about this class. The formal proofs of these statements are postponed to the full version.

1. For every d, CF_d is closed under complement, (follows from the definition).
2. $\mathsf{SZK} \subseteq \mathsf{CF}_1$.
3. $\mathsf{CF}_{o(\frac{n}{\log n})}$ does not "generically" solve NP.
4. $\mathsf{CF}_{\frac{l(n)}{\log n}}$ can invert any one-way function with output length $l(n)$ on length n inputs (Theorem 3 in Section 4.2)

We leave a fuller exploration of the collision-finding class for future work. Note that, by property (4), if $\mathsf{NP} \not\subseteq \mathsf{CF}_{poly(n)}$, we have a "natural" complexity class that can inverts all one-way functions but not decide NP.

We mention that a somewhat weaker (and perhaps even more natural) definition of the collision-finding class—let us denote it $\mathsf{CF'}_d$—is defined identically, but without giving M, or Sam, access to a random permutation π. That is, in our notation $\mathsf{CF'}_d = \mathsf{BPP}^{\mathsf{Sam}^{\perp}_d}$, where $\perp$ is the all zero oracle. Clearly $\mathsf{CF'}_d \subseteq \mathsf{CF}_d$, but all the properties above continue to hold also for $\mathsf{CF'}$.

A very recent work by Haitner, Mahmoody-Ghidary and Xiao [16] takes a step towards showing that $\mathsf{CF'}_{O(1)}$ does not contain NP; they show that if the deciding machine M only makes a constant number of adaptive queries to Sam, then the language it decides is in coAM.

4 From Private Coins to Public Coins

In this section, we provide our transformation from private-coin to public-coin protocols. We provide two transformations: The first transformation—or *weak duality*—converts any private coin zero-knowledge proof into a public-coin zero-knowledge proof in the Sam-hybrid model, where the prover, verifier and simulator have oracle access to Sam. The second transformation—or *strong duality*—converts any private coin zero-knowledge argument with sublinear verifier communication complexity into a public coin zero-knowledge argument in the Sam-hybrid model. While the first transformation is *oracle efficient* (the maximum depth it queries Sam is "small"), the second transformation is *computationally efficient* (the soundness reduction is polynomial-time) and thus can be applied to arguments.

Our transformations consider zero-knowledge proofs and arguments in an oracle world. Let $\mathcal{O}$ be a set of oracles $O : \{0,1\}^* \to \{0,1\}^*$.

Definition 4 ($\mathcal{O}$-relativized Interactive Proofs) *A pair of interactive machines $\langle P^\bullet, V^\bullet \rangle$ is called an $\mathcal{O}$-relativized* interactive proof system *for a language L if machine $V^\bullet$ is polynomial-time and the following two conditions hold :*

- Completeness: *There is a negligible function $\nu(\cdot)$, such that for every n, $x \in L \cap \{0,1\}^n$,*
$$\Pr\left[\langle P^O, V^O \rangle(x) = 1\right] \geq 1 - \nu(n)$$
where the probability is taken over all the internal coin tosses of P, V and uniformly chosen $O \in \mathcal{O}$.
- Soundness: *For every machine $B^\bullet$, there exists a negligible function $\nu(\cdot)$, such that, for every $x \in \overline{L} \cap \{0,1\}^n$,*
$$\Pr\left[\langle B^O, V^O \rangle(x) = 1\right] \leq \nu(n)$$
where the probability is taken over all the internal coin tosses of V and uniformly chosen $O \in \mathcal{O}$.

If the soundness holds only against n.u.$\mathcal{PPT}$ B, then $\langle P, V \rangle$ is called an $\mathcal{O}$-relativized interactive argument system.

Definition 5 ($\mathcal{O}$-relativized black-box $\mathcal{ZK}$) *Let $\langle P^\bullet, V^\bullet \rangle$ be an $\mathcal{O}$-relativized interactive proof (argument) system for the language $L \in \mathcal{NP}$ with the witness relation R_L. We say that $\langle P^\bullet, V^{*\bullet} \rangle$ is $\mathcal{O}$-relativized computational black-box $\mathcal{ZK}$, if there exists a probabilistic expected polynomial time oracle machine $S^\bullet$ such that for every $\mathcal{PPT}$ machine $V^{*\bullet}$, and $\mathcal{PPT}$ distinguisher $D^\bullet$, there exists a negligible function $\nu(\cdot)$, such that for all n, $x \in L \cap \{0,1\}^n$, $z \in \{0,1\}^*$,*

$$\left|\Pr[D^O(S^{O,V^O(x)}(x)) = 1] - \Pr[D^O(\langle P^O, V^{*O}(z) \rangle(x)) = 1]\right| < \nu(n)$$

4.1 Weak Duality Lemma

Lemma 1 (Weak Duality). *Let $\langle P^\bullet, V^\bullet \rangle$ be a d-round fully black-box zero-knowledge proof for a language L from one-way permutations with verifier communication complexity $c(n)$ and prover communication complexity $p(n)$. Then, there exists a d-round public-coin protocol $\langle \widehat{P}^{\mathsf{Sam}_d^\bullet}, \widehat{V}^{\mathsf{Sam}_d^\bullet} \rangle$ with the verifier communication complexity $O(dc(n))$ and prover communication complexity $p(n)$ that is $(\pi, \mathsf{Sam}_d^\pi)$-relativized black-box zero-knowledge proof.*

Proof: $\widehat{V}^{\mathsf{Sam}_d^\pi}$ is a d-round public-coin verifier that sends random coins in each round. On a high-level, $\widehat{P}^{\mathsf{Sam}_d^\pi}$ is a machine that internally incorporates the code of P^π and emulates an interaction with P^π by supplying verifier messages according to the $\langle P^\pi, V^\pi \rangle$ protocol. For every verifier round in the internal emulation, $\widehat{P}^{\mathsf{Sam}_d^\pi}$ first receives random coins externally from $\widehat{V}^{\mathsf{Sam}_d^\pi}$. Using that, it samples a random message q for V^π that is "consistent" with the interaction with P^π; this is made possible using Sam_d^π. Next, it q feeds internally to P^π. Upon receiving a message a from P^π, $\widehat{P}^{\mathsf{Sam}_d^\pi}$ forwards a to $\widehat{V}^{\mathsf{Sam}_d^\pi}$ and proceeds to the

PROTOCOL $\langle \widehat{P}, \widehat{V} \rangle$

Let $\langle P^\pi, V^\pi \rangle$ be a d-round protocol with oracle access to the a permutations π. Each communication round consists of a message sent from the verifier to the prover followed by a message sent from the prover to the verifier. Without loss of generality, we assume that the verifier sends the first message and the prover sends the last message. Also, the verifier outputs its view at the end of the protocol.

1. **Common Input:** Statement $x \in L$, security parameter n.
2. **Private Input:** The statement x, for P^π and auxiliary input $z \in \{0,1\}^*$ for V^π.
3. $\widehat{P}^{\mathsf{Sam}_d^\pi}$ internally incorporates the code for P^π. Set $trans_0 = \perp$.
4. **for** $i = 1$ to d
   (a) $\widehat{V}^{\mathsf{Sam}_d^\pi}$ uniformly chooses $s_i \in \{0,1\}^{12(l_i + \log d)}$ where l_i is the length of V^π's i^{th} message.
   (b) $\widehat{V} \rightarrow \widehat{P} : s_i$
   (c) $\widehat{P}$ queries Sam_d^π on input $(V^\bullet, trans_{i-1}, s_i)$ and obtains as response $(trans_{i-1} :: q_i, r_i)$. $\widehat{P}$ runs $P^\pi(trans_{i-1} :: q_i)$ and obtains its response a_i. Set $trans_i = trans_{i-1} :: q_i :: a_i$.
   (d) $\widehat{P} \rightarrow \widehat{V} : a_i$.
5. $\widehat{V}^{\mathsf{Sam}_d^\pi}$ computes $trans_i$ for all i, by querying Sam_d^π on $(V^\bullet, trans_{i-1}, s_i)$.
   $\widehat{V}^{\mathsf{Sam}_d^\pi}$ chooses $s \in \{0,1\}^n$, queries Sam_d^π on $(V^\pi, trans_d, s)$ and obtains as response (b, r_{d+1}), ($b = 1$ means V^π accepts). It **outputs** b.

Fig. 1. Weak Duality Protocol

next round. Finally, $\widehat{V}^{\mathsf{Sam}_d^\pi}$ reconstructs the interaction emulated by P^π (again made possible using Sam_d^π) and outputs the verdict of V^π on that transcript. We remark that since V^π is a verifier for a d-round protocol, the maximum depth of a Sam_d^π query made by $\widehat{P}^{\mathsf{Sam}_d^\pi}$ and $\widehat{V}^{\mathsf{Sam}_d^\pi}$ is d.

Informally, the completeness of the protocol follows from the fact that, the internal emulation carried out by $\widehat{P}^{\mathsf{Sam}_d^\pi}$ proceeds exactly as an execution between P^π and the honest verifier V^π. The soundness and zero-knowledge of $\langle \widehat{P}, \widehat{V} \rangle$, on the other hand, holds as the transformation essentially ensures that the messages from $\widehat{V}^{\mathsf{Sam}_d^\pi}$ carry the same amount of "knowledge" as messages from V^π. This is because, in each round, $\widehat{P}^{\mathsf{Sam}_d^\pi}$ samples a fresh random tape for V^π that is consistent with the partial conversation and obtains V^π's next message by running V^π on that tape. Thus, the only extra knowledge that $\widehat{P}$ possesses in each round is the random tape sampled and (an unbounded) P^π can obtain these samples too. A formal description of the transformation is provided in Figure 1. We now proceed to prove correctness.

Claim 1 (Completeness) *For all $x \in L$, $\widehat{P}^{\mathsf{Sam}_d^\pi}$ convinces $\widehat{V}^{\mathsf{Sam}_d^\pi}$ w.p. $1 - \nu(|x|)$ where the randomness is taken over π, Sam_d^π and the internal coin tosses of $\widehat{P}$ and $\widehat{V}$, for some negligible function $\nu(\cdot)$.*

Proof: We show that, for every permutation π, the probability that $\widehat{V}^{\mathsf{Sam}_d^\pi}$ accepts is identical to the probability V^π accepts in an interaction with P (where the probability is over Sam_d^π). The completeness of $\langle \widehat{P}, \widehat{V} \rangle$ then follows from the completeness of $\langle P^\bullet, V^\bullet \rangle$.

Towards this, fix a permutation π. Consider an intermediate verifier $V'^\bullet$ that uses Sam_d^π and interacts with P^π. Informally, this verifier $V'^{\mathsf{Sam}_d^\pi}$ generates messages exactly as $\widehat{P}^{\mathsf{Sam}_d^\pi}$ does in the internal emulation with P^π. More precisely, for a partial transcript $trans_{i-1}$ at the end of round $i-1$, $V'^{\mathsf{Sam}_d^\pi}$ samples a consistent random tape r for V^π (using Sam_d^π) and runs V_r^π on $trans_{i-1}$ to generate the next verifier message q_i. At the end of the protocol, $V'^{\mathsf{Sam}_d^\pi}$, samples a random tape r' consistent on the entire transcript and outputs $V_{r'}^\pi$'s verdict on the transcript. It follows from the construction that the probability that $\widehat{V}^{\mathsf{Sam}_d^\pi}$ accepts is equal to the probability that $V'^{\mathsf{Sam}_d^\pi}$ accepts in an interaction with P^π. In Claim 1 below, we prove that the probability that P^π convinces $V'^{\mathsf{Sam}_d^\pi}$ accepts is equal to the probability that P^π convinces V^π. Therefore, combining the two facts, we get that the probability $\widehat{V}^{\mathsf{Sam}_d^\pi}$ accepts is equal to the probability V^π accepts.

Sub-Claim 1 *For every $x, z \in \{0,1\}^*$, $\pi \in \Pi$ the following distributions are identical: $D_1 = \left\{ \langle P^\pi, V^\pi(z) \rangle(x) \right\}$ and $D_2 = \left\{ \langle P^\pi, V'^{\mathsf{Sam}_d^\pi}(z) \rangle(x) \right\}$ where the distributions are generated by the internal coin tosses of P, V, V' and Sam_d^π.*

Proof: Recall that the only difference between $V'^{\mathsf{Sam}_d^\pi}$ and V^π in an interaction with P^π is that V^π selects a uniform random tape at the beginning of the execution and uses that for the entire execution, while $V'^{\mathsf{Sam}_d^\pi}$ selects a (uniformly chosen) random tape consistent with the partial transcript in each round and executes V^π on that tape. First, we observe that every verifier message in D_1 and D_2 are generated by running V^π on a particular random tape. For a transcript $trans$, let $R(trans)$ denote the set of all random tapes of V^π consistent with $trans$. We show for D_1 and D_2, separately, that for every $trans$, conditioned on the history being $trans$, the random tape used to generate the next message is uniformly distributed in $R(trans)$. This shows that the process for generating verifier messages in D_1 and D_2 are identical and that concludes the proof of the claim. For D_2, this holds directly from the definition of Sam_d^π. For D_1, we prove this fact by induction on the number of verifier messages. The base case requires that the random tape is uniformly distributed over all possible tapes; this clearly holds. Suppose that, conditioned on the transcript $trans_{i-1}$, every random tape in $R_{trans_{i-1}}$ is equally likely. Let m be a possible message for V^π, given the history is $trans_{i-1}$. Since, $R(trans_{i-1} :: m)$ are disjoint sets for different m, we have that conditioned on the transcript $trans_{i-1} :: m$, every tape in $R(trans_{i-1} :: m)$ is equally likely to be chosen. This concludes the induction step. ∎

∎

THE ALGORITHM FOR $P^{*\text{SAM}^{\pi}_{d}}$

Let $\widehat{P}^{*\mathsf{Sam}^{\pi}_{d}}$ be the cheating prover for $\langle \widehat{P}, \widehat{V} \rangle$. Denote the length of i^{th} verifier message in $\langle \widehat{P}, \widehat{V} \rangle$ by L_i. Then, $L_i = 12(l_i + \log d)$.

1. Internally incorporate $\widehat{P}^{\mathsf{Sam}^{\pi}_{d}}$. Let $trans_0 = \perp$.
2. **for** $i = 1$ to d
   (a) Receive q_i from V^{π}.
   (b) $ctr \leftarrow 0$, $found \leftarrow false$.
   (c) **while** $ctr < 2^{\frac{L_i}{3}}$ and $not(found)$.
   - Choose s_i uniformly from $\{0,1\}^{L_i}$. Let $(r_i, trans_i :: q) \leftarrow \mathsf{Sam}^{\pi}_{d}(V^{\bullet}, trans_{i-1}, s_i)$. If $q = q_i$, $found = true$.
   - $ctr \leftarrow ctr + 1$

   (d) **if** $found = false$, **abort**. Otherwise, compute $\widehat{P}^{*}$'s next message on transcript $trans_{i-1} :: q_i$. Let it be a_i. Set $trans_i = trans_{i-1} :: q_i :: a_i$.

Fig. 2. Proof of Soundness

Claim 2 (Soundness) *Let $x \in \{0,1\}^*$. If $\widehat{P}^{*\mathsf{Sam}^{\pi}_{d}}$ convinces $\widehat{V}^{\mathsf{Sam}^{\pi}_{d}}$ on x with probability p, then there exists a prover $P^{*\mathsf{Sam}^{\pi}_{d}}$ that convinces V^{π} on input x with probability at least $\frac{p}{2}$. (As usual, the probability are taken over Sam^{π}_{d} and the internal coin tosses of $\widehat{V}$ and V.)*

Proof: We prove the statement of the claim for every permutation π and over the randomness of Sam^{π}_{d}. We construct a machine $P^{*\mathsf{Sam}^{\pi}_{d}}$ that internally incorporates $\widehat{P}^{*\mathsf{Sam}^{\pi}_{d}}$ and emulates an interaction with it, while externally interacting with V^{π}. The high-level idea is to make $P^{*\mathsf{Sam}^{\pi}_{d}}$ convince V^{π} whenever $\widehat{P}^{*\mathsf{Sam}^{\pi}_{d}}$ succeeds in the internal execution. To ensure this, for every private-coin message q_i that $P^{*\mathsf{Sam}^{\pi}_{d}}$ receives externally from V^{π}, it needs to find a corresponding public-coin message s_i and feed it to $\widehat{P}^{*\mathsf{Sam}^{\pi}_{d}}$. Let $trans_{i-1}$ be the transcript of messages exchanged with V^{π} externally. Then, the message that $P^{*\mathsf{Sam}^{\pi}_{d}}$ needs to find, is a string s_i such that Sam^{π}_{d} on input $(V^{\bullet}, trans_{i-1}, s_i)$ outputs (q_i, r_i). We let $P^{*\mathsf{Sam}^{\pi}_{d}}$ sample s_i until it hits the "right" one; it cuts itself off, if it runs "too long". It then feeds it to $\widehat{P}^{*\mathsf{Sam}^{\pi}_{d}}$ internally and obtains a response a_i, which it forwards outside to V^{π}. A formal description is provided in Figure 2.

In Claim 2, we show that $P^{*\mathsf{Sam}^{\pi}_{d}}$ aborts with probability at most $\frac{1}{2}$. In Claim 3, we show that conditioned on $P^{*\mathsf{Sam}^{\pi}_{d}}$ not aborting, each verifier message fed internally to $\widehat{P}^{*\mathsf{Sam}^{\pi}_{d}}$ is uniformly distributed and thus identical to distribution of the messages received by $\widehat{P}^{*\mathsf{Sam}^{\pi}_{d}}$ in a real interaction with $\widehat{V}^{\mathsf{Sam}^{\pi}_{d}}$. Combining the two claims, we have that the probability that $P^{*\mathsf{Sam}^{\pi}_{d}}$ succeeds is at least $\frac{1}{2}\Pr[\widehat{P}^{*\mathsf{Sam}^{\pi}_{d}}$ succeeds].

Sub-Claim 2 *$P^{*\mathsf{Sam}^{\pi}_{d}}$ aborts with probability at most $\frac{1}{2}$ (with probability taken over Sam^{π}_{d}).*

Proof: We analyze the abort probability by identifying three bad events for each round and bound their probabilities separately. Then, using an union bound over the bad events for each round, we conclude that the probability of aborting is at most $\frac{1}{2}$. Let $trans_{i-1}$ be the partial transcript at the end of $i-1$ rounds. Consider the following events:

1. **$P^{*\mathsf{Sam}_d^\pi}$ picks the same sample s_i twice:** The probability that two strings in $2^{\frac{L_i}{3}}$ trials are the same is at most $\frac{2^{\frac{2L_i}{3}}}{2^{L_i}} \leq \frac{1}{2^{4(l_i+\log d)}}$ using the birthday bound.
2. **V^π sends an "unlikely" message m:** Let p_m be the probability that V^π sends m conditioned on $trans_{i-1}$ being the transcript at the end of $i-1$ rounds. We say that m is unlikely if $p_m \leq \frac{1}{2^{2l_i+2\log d}}$. Using a union bound over all m we obtain that the probability of an unlikely m being sent is at most $2^{l_i}\frac{1}{2^{2l_i+2\log d}} = \frac{1}{2^{l_i+2\log d}}$
3. **For a "likely" message m, all trials fail:** The probability of a "likely" message m occurring is at least $> \frac{1}{2^{2(l_i+\log d)}}$. Therefore, the probability that all $2^{\frac{L_i}{3}}$ trials fails is at most

$$\left(1 - \frac{1}{2^{2(l_i+\log d)}}\right)^{2^{\frac{L_i}{3}}} \leq e^{-2^{2(l_i+\log d)}}$$

If the bad events do not occur in round i, then the message m is a "likely" message and some trial succeeds, which implies that $P^{*\mathsf{Sam}_d^\pi}$ does not abort in round i. Using the union bound, we obtain that $P^{*\mathsf{Sam}_d^\pi}$ aborts with probability at most $\frac{1}{d2^{l_i}}$ in round i. Using the union bound again over all the d rounds, the probability that $P^{*\mathsf{Sam}_d^\pi}$ aborts is at most $\sum_{i=1}^{d}\frac{1}{d2^{l_i}} \leq \frac{1}{2^L}$ where L is the length of the shortest message. Thus, $P^{*\mathsf{Sam}_d^\pi}$ aborts with probability at most $\frac{1}{2}$. ∎

Sub-Claim 3 *Conditioned on $P^{*\mathsf{Sam}_d^\pi}$ not aborting, the probability that $P^{*\mathsf{Sam}_d^\pi}$ succeeds in convincing V is identical to the probability that $\widehat{P}^{*\mathsf{Sam}_d^\pi}$ succeeds in convincing $\widehat{V}^{\mathsf{Sam}_d^\pi}$.*

Proof: Recall that, in every round $P^{*\mathsf{Sam}_d^\pi}$ samples public-coin messages a fixed number of times and aborts if none of them correspond to the private-coin message received externally from V^π. We observe that, the process that decides whether the random coins sampled by $P^{*\mathsf{Sam}_d^\pi}$ are the "right" ones depends only on the randomness of Sam_d^π and, in particular, is independent of the actual public-coin message sampled by $P^{*\mathsf{Sam}_d^\pi}$. Therefore, conditioned on $P^{*\mathsf{Sam}_d^\pi}$ not aborting, the messages fed internally to $\widehat{P}^{*\mathsf{Sam}_d^\pi}$ are uniformly distributed. Since, $\langle\widehat{P},\widehat{V}\rangle$ is a public-coin protocol, we have that the distribution of messages internally fed to $\widehat{P}^{*\mathsf{Sam}_d^\pi}$ is identically distributed to the messages generated in a real interaction with $\widehat{V}^{\mathsf{Sam}_d^\pi}$, and hence the probability that the internal emulation leads to a successful interaction is identical to the probability that $\widehat{P}^{*\mathsf{Sam}_d^\pi}$ succeeds in a real interaction. Recall that the acceptance condition in the internal emulation is decided by reconstructing a $\langle P^\pi, V^\pi\rangle$ transcript, sampling a fresh

random tape consistent with the entire transcript and running V^π on that tape to obtain the verdict. By our construction, the transcript of the internal emulation with $P^{*\mathsf{Sam}_d^\pi}$ is identical to the transcript between $P^{*\mathsf{Sam}_d^\pi}$ and the external V^π. However, the random coins of the external V^π might not be the same as the ones sampled internally. Nevertheless, using the same proof as in Sub-Claim 1, it follows that conditioned on any complete transcript, the probability that the external verifier V^π and the internally emulated $\widehat{V}^{\mathsf{Sam}_d^\pi}$ accept are identical. ∎

∎

Remark 2. Note that in proof of Claim 2 we provide an algorithmic description of the cheating prover $P^{*\mathsf{Sam}_d^\pi}$ although we only need to contradict "unconditional soundness". This algortihm will be useful in proving the strong duality lemma (see Lemma 2) where consider also computationally-sound protocols.

Remark 3. The expected running time of $P^{*\mathsf{Sam}_d^\pi}$ in round i for a partial transcript $trans_{i-1}$ of the first $i-1$ rounds, is bounded by $\sum_m p_m \frac{1}{p_m} = 2^{l_i}$ where p_m is the conditional probability that V^π sent m in round i given $trans_{i-1}$. Therefore, the total expected running time of $P^{*\mathsf{Sam}_d^\pi}$ is at most $d \cdot 2^L = d2^L$ where L is the length of the longest message that V sends. If either the length of a message or the number of rounds is super-logarithmic, then the cheating prover P^* does not run in polynomial time. In the strong duality lemma, we show how to overcome this problem, as long as the verifier communication complexity is sublinear; this, however, requires querying Sam on larger depths.

Simulation: Let S^π be the simulator for $\langle P^\pi, V^\pi \rangle$. We construct a simulator $\widehat{S}^{\mathsf{Sam}_d^\pi}$ for $\langle \widehat{P}, \widehat{V} \rangle$ using S^π that has oracle access to Sam_d^π. Let $\widehat{V}^{*\mathsf{Sam}_d^\pi}$ be a malicious verifier for $\langle \widehat{P}, \widehat{V} \rangle$. On a high-level, $\widehat{S}^{\mathsf{Sam}_d^\pi}$ transforms $\widehat{V}^{*\mathsf{Sam}_d^\pi}$ to a verifier $V^{*\mathsf{Sam}_d^\pi}$ for $\langle P, V \rangle$ and simulates $V^{*\mathsf{Sam}_d^\pi}$ using S^π. The verifier $V^{*\mathsf{Sam}_d^\pi}$ with oracle access to $\widehat{V}^{*\mathsf{Sam}_d^\pi}$ and Sam_d^π proceeds as follows. In each round, on receiving a message from P^π, $V^{*\mathsf{Sam}_d^\pi}$ feeds that message to $\widehat{V}^{*\mathsf{Sam}_d^\pi}$. It obtains $\widehat{V}^{*\mathsf{Sam}_d^\pi}$'s next public-coin message r. $V^{*\mathsf{Sam}_d^\pi}$ queries Sam_d^π using r and generates the next message of V^π (i.e. generates a message following $\widehat{P}^{\mathsf{Sam}_d^\pi}$'s procedure) and forwards that to P^π. Finally, $V^{*\mathsf{Sam}_d^\pi}$ outputs what $\widehat{V}^{*\mathsf{Sam}_d^\pi}$ outputs. The simulator for $\langle \widehat{P}, \widehat{V} \rangle$, $\widehat{S}^{\mathsf{Sam}_d^\pi}$, internally incorporates S^π and verifier $V^{*\mathsf{Sam}_d^\pi}$, emulates an execution of S^π with $V^{*\mathsf{Sam}_d^\pi}$ and outputs what S^π outputs.

To show correctness of simulation, we need to show that $\langle \widehat{P}, \widehat{V} \rangle$ is $(\pi, \mathsf{Sam}_d^\pi)$-relativized zero-knowledge. Assume for contradiction, that there is a distinguisher $D^\bullet$ that can distinguish the simulation of $\widehat{V}^{*\mathsf{Sam}_d^\pi}$ by $\widehat{S}^{\mathsf{Sam}_d^\pi}$ from the real interaction for a random $(\pi, \mathsf{Sam}_d^\pi)$. More precisely, there exists a $\mathcal{PPT}$ distinguisher $D^\bullet$, polynomial $p(n)$, sequence $\{x_n, z_n\}_{n=1}^\infty$, $x_n \in L \cap \{0,1\}^n$, $z_n \in \{0,1\}^*$ such that for infinitely many n, $D^{\mathsf{Sam}_d^\pi}$ distinguishes the output of $\widehat{S}^{\mathsf{Sam}_d^\pi, \widehat{V}^{*\mathsf{Sam}_d^\pi}(x_n, z_n)}(x_n, 1^n)$ for a random $(\pi, \mathsf{Sam}_d^\pi)$ and the output of $\widehat{V}^{*\mathsf{Sam}_d^\pi}(x_n, z_n)$ in a real interaction with probability $\frac{1}{p(n)}$ (with probability taken over a random π, Sam_d^π). Using the Borel-Cantelli lemma, it follows that for measure 1 over permutations π, $D^{\mathsf{Sam}_d^\pi}$ distinguishes $\widehat{S}^{\mathsf{Sam}_d^\pi, \widehat{V}^{*\mathsf{Sam}_d^\pi}(x_n, y_n)}(x_n, 1^n)$ and

the output of $\widehat{V}^{*\mathsf{Sam}_d^\pi}(x_n, y_n)$ in a real interaction with probability $\frac{1}{n^2 p(n)}$ for infinitely many n (with probability over Sam_d^π). Fix a $\pi = \{\pi_n\}_{n=1}^\infty$ for which this happens. It follows by the construction of $V^{*\mathsf{Sam}_d^\pi}$ that the output of $\widehat{S}^{\mathsf{Sam}_d^\pi}$ on $\widehat{V}^{*\mathsf{Sam}_d^\pi}$ is identically distributed to the output of S^π on $V^{*\mathsf{Sam}_d^\pi}$. We further claim that the output of $V^{*\mathsf{Sam}_d^\pi}$ in a real interaction with P^π is identically distributed to the output of $\widehat{V}^{*\mathsf{Sam}_d^\pi}$ with $\widehat{P}^{\mathsf{Sam}_d^\pi}$ (over a random Sam_d^π). The proof of this identically follows from the proof of Claim 1. Hence, $D^{\mathsf{Sam}_d^\pi}$ distinguishes the output of S^π with $V^{*\mathsf{Sam}_d^\pi}$ from the output of $V^{*\mathsf{Sam}_d^\pi}$ in a real interaction with P^π with probability $\frac{1}{p(n)}$. Recall that, $\langle P, V\rangle$ is a fully black-box zero-knowledge based on one-way permutations, there exists a $\mathcal{PPT}$ machine $B^\bullet$, that with oracle access to $D^{\mathsf{Sam}_d^\pi}$ and $V^{*\mathsf{Sam}_d^\pi}$ inverts π (over a random Sam_d^π) for infinitely many lengths, for measure 1 over permutations π (and hence for a random π). From Theorem 2, we know that if $d \in o(\frac{n}{\log n})$, then no $\mathcal{PPT}$ machine with oracle access to Sam_d^π can invert a random one-way permutation π with more than negligible probability. Therefore, we arrive at a contradiction. This establishes that $\langle \widehat{P}, \widehat{V}\rangle$ is a $(\pi, \mathsf{Sam}_d^\pi)$-relativized black-box zero-knowledge proof. ∎

4.2 Strong Duality Lemma

Lemma 2 (Strong Duality). *Let $\langle P^\bullet, V^\bullet\rangle$ be a d-round fully black-box zero-knowledge argument for a language L from one-way permutations with verifier communication complexity $c(n)$ and prover communication complexity $p(n)$. Then, there exists a d-round public-coin protocol $\langle \widehat{P}^{\mathsf{Sam}_D^\bullet}, \widehat{V}^{\mathsf{Sam}_D^\bullet}\rangle$ with the verifier communication complexity $O(Dc(n))$ (where $D = \frac{c(n)}{\log n}$) and prover communication complexity $p(n)$ that is $(\pi, \mathsf{Sam}_D^\pi)$-relativized black-box zero-knowledge argument.*

Proof: We modify the construction and proof from the previous lemma to obtain this lemma. From Remark 3, we know that the running time of P^*, is $d2^L$ where L is the length of the longest V-message. In order to use the previous construction and obtain an efficient P^*, we need the length of every verifier message to be logarithmic. Alternatively, if we split every message into segments of length $\log n$ bits and use the random tape sampled by Sam to generate one segment of the verifier message at a time, this also makes the running time of P^* polynomial. However, now, we need only to ensure the verifier's communication complexity is $o(n)$ (as this guarantees that the maximum depth is $o(\frac{n}{\log n})$). We note that the idea of splitting messages into segments of $\log n$ bits was used in [14] but their use of this technique is not sufficient for our application. More precisely, in [14] it is only shown how Sam can be used to generate a new random tape assuming that the original random tape was also generated using the random oracle. In our application, we need to be able to find a random tape for the "external" verifier. (As observed by Haitner in a personal communication, it would seem that by using our techniques (from Lemma 1) the results of [14] could be extended to rule-out also constructions that are secure with respect to only *honest-but-curious* players.)

We describe the procedure for generating a verifier message using Sam and the rest of the proof follows identically by plugging in this procedure wherever V^π's message is required to be generated. Without loss of generality, we assume that V^π's message in the i^{th} round is a multiple of $\log n$, say $k \log n$. We describe how to sample the i^{th} round message given $trans_{i-1}$, the partial transcript for the first $i-1$ rounds and a random string s_i. We first split s_i into k equal parts, $s_i^1, \ldots, s_i^k$. Using s_i^1, we samples a random tape r_1 for V^π consistent with $trans_{i-1}$ using Sam. We then run $V_{r_1}^\pi$ to generate only the first $\log n$ bits of V^π's message in round i, say q_i^1. Next, we sample another random tape r_2 using s_i^2, but now r_2 is consistent with the extended transcript $trans_{i-1} :: q_i^1$. We run $V_{r_2}^\pi$ and obtain the next $\log n$ bits, q_i^2. In this manner we generate every $\log n$ segment up to q_i^k, each time ensuring that it is consistent with all previous segments. The depth of the maximum Sam_D^π query is total number segments counted over all verifier messages, i.e. $D = \frac{c(n)}{\log n}$ where $c(n)$ is the total verifier communication complexity. Since, $c(n) \in o(n)$, the maximum depth of a Sam query is $o(\frac{n}{\log n})$.

Completeness, Soundness and Zero-Knowledge: The proof of completeness follows exactly as before. As show for the weak-duality, we prove that if there exists a cheating prover $\widehat{P}^*$ for $\langle \widehat{P}, \widehat{V} \rangle$ that succeeds with probability p, there is a prover P^* with oracle access to Sam_D^π succeeds in $\langle P, V \rangle$ with probability $\frac{p}{2}$. The running time of P^*, computed as before, is $\frac{c(n)}{\log n} 2^{\log n} = \left(\frac{nc(n)}{\log n}\right)$ which is polynomial since $c(n) \in o(n)$. Therefore, there exists a $\mathcal{PPT}$ prover P^* with oracle access to $\mathsf{Sam}_{o(\frac{n}{\log n})}^\pi$ that cheats with probability $\frac{p}{2}$. If $\langle \widehat{P}, \widehat{V} \rangle$ is not a $(\pi, \mathsf{Sam}_D^\pi)$-relativized argument for a random permutation π, then there is a sequence $\{x_n\}_{n=1}^\infty$, $x_n \in \overline{L} \cap \{0,1\}^n$, polynomial p such that P^* succeeds in convincing on V on x_n with probability $\frac{1}{p(n)}$ over a random permutation π for infinitely many n. Applying the Borel-Cantelli lemma, we again have that with measure 1 over permutations π, P^* cheats for infinitely many n. Using the fully black-box property, we have that for measure 1 over permutations π, there exists an adversary A that inverts π (and hence, for a random π); this violates Theorem 2 and we arrive at a contradiction. This completes the proof of soundness. To prove zero-knowledge, we use the same simulator from the weak duality, with the exception that it treats the verifier messages in $\log n$-bit segments. The rest of the proof follows as before. ∎

We mention that the proof of the strong duality transformation shows that Theorem 2 (due to [13]) is optimal.

Theorem 3. *Let $f^\bullet$ be a function that on inputs of length n has output length $l(n)$. Then, for any π, there exists an oracle $\mathcal{PPT}$ machine $A^\bullet$, such that $A^{\mathsf{Sam}_{\frac{l(n)}{\log n}}^\pi}$ inverts f^π.*

Proof: First, we construct a protocol a 1-round protocol $\langle P^\bullet, V^\bullet \rangle$ for the empty language as follows: On input 1^n, V^π computes $y = f^\pi(r)$, where r is its random tape and sends y to P^π. P^π sends a string x' to V^π. V^π accepts

if $f^\pi(x') = y$. Next, we apply the strong-duality transformation to the protocol $\langle P^\bullet, V^\bullet \rangle$ and obtain $\langle \widehat{P}, \widehat{V} \rangle$. In $\langle \widehat{P}, \widehat{V} \rangle$, $\widehat{P}^{\mathsf{Sam}^\pi_{l(n)/\log n}}$ on receiving a random string s from $\widehat{V}^{\mathsf{Sam}^\pi_{l(n)/\log n}}$, queries $\mathsf{Sam}^\pi_{l(n)/\log n}$ with input s and obtains a random tape r for V^π. $\widehat{P}^{\mathsf{Sam}^\pi_{l(n)/\log n}}$ runs V^π_r and obtains $f^\pi(r)$. $\widehat{V}^{\mathsf{Sam}^\pi_{l(n)/\log n}}$ accepts at the end, if $\widehat{P}^{\mathsf{Sam}^\pi_{l(n)/\log n}}$ can send r' such that $f^\pi(r') = f^\pi(r)$. Notice that $\widehat{P}^{\mathsf{Sam}^\pi_{l(n)/\log n}}$ knows r from $\mathsf{Sam}^\pi_{l(n)/\log n}$'s response, and can just forward r directly to $\widehat{V}^{\mathsf{Sam}^\pi_{l(n)/\log n}}$. Therefore, there is a cheating prover for $\langle \widehat{P}, \widehat{V} \rangle$, that succeeds with probability 1. From the proof of soundness of the strong duality lemma, we know how to construct a cheating prover $P^{*\mathsf{Sam}^\pi_{l(n)/\log n}}$ that convinces V^π with probability at least $\frac{1}{2}$. This means that $P^{*\mathsf{Sam}^\pi_{l(n)/\log n}}$ inverts f^π with probability at least $\frac{1}{2}$. The maximum depth of a query by $P^{*\mathsf{Sam}^\pi_{l(n)/\log n}}$ is $\frac{l(n)}{\log n}$. ∎

We call a function *compressing* if it on inputs of length n has output length $o(n)$.

Corollary 1 *There exists no fully black-box construction of a compressing one-way function from one-way permutations.*

Proof: From Theorem 3, we have an adversary that inverts f^π with oracle access to $\mathsf{Sam}^\pi_{o(\frac{n}{\log n})}$. By the fully black-box property, we have an adversary A with oracle access to $\mathsf{Sam}^\pi_{o(\frac{n}{\log n})}$ that inverts π. Since this holds for every π, we arrive at a contradiction to Theorem 2. ∎

5 Black-box Lower Bounds for Zero Knowledge

All our black-box lower bounds follow by combining the weak or strong duality lemma with known lower bounds for public-coin protocols.

5.1 Lower Bounds Zero-Knowledge Proofs and Arguments

Goldreich-Krawczyk [8] show that only languages L in BPP have black-box constant-round public-coin zero-knowledge proofs. We remark that the proof of GK uses the simulator as a black-box to decide the language L, and relativizes. We therefore have:

Theorem 4 (Implicit in [8]). *Let $\langle P^\bullet, V^\bullet \rangle$ be a $\mathcal{O}$-relativized constant-round public-coin zero-knowledge proof for a language L with a black-box simulator S. Then, there exists a $\mathcal{PPT}$ machine $M^\bullet$, such that M^{V^O, S^O} decides L with probability $\frac{2}{3}$ when the probability is taken over a uniformly chosen $O \in \mathcal{O}$.*

Combining this theorem with the weak-duality lemma, we obtain the following corollary.

Corollary 2 (Constant-round Zero-Knowledge Proofs) *For any constant d, only languages L in CF_d have d-round fully black-box zero-knowledge proofs from one-way permutations.*

Proof: Let $\langle P, V \rangle$ be fully black-box zero-knowledge proof based on one-way permutations. Applying the weak duality lemma, we obtain a protocol $\langle \widehat{P}, \widehat{V} \rangle$ that is public-coin protocol where the prover, verifier and the simulator have access to Sam_d^π, that is $(\pi, \mathsf{Sam}_d^\pi)$-relativized black-box zero-knowledge proof. Using Theorem 4, we have that $L \in \mathsf{CF}_d$. ∎

We remark that if $\mathsf{NP} \not\subseteq \mathsf{CF}_{O(1)}$, then the corollary is tight; Goldreich, Micali and Wigderson [10] present a fully black-box construction of an $\omega(1)$-round protocol for NP based on one-way functions. On the other hand, Goldreich and Kahan [7], present a fully black-box $O(1)$-round zero-knowledge proofs for all of NP using claw-free permutations.[1]

Remark 4. A very recent work by Gordon, Wee, Xiao and Yerukhimovich [12] strengthens Corollary 2 by removing the usage of the random oracle π, and thus placing the class of languages having $O(1)$-round fully black-box zero-knowledge proofs from one-way permutations in $\mathsf{CF}'_{O(1)}$ (see Section 3.2). By relying on the recent work of [16], they obtain as a corollary that only languages in coAM have constant-round fully black-box zero-knowledge proofs from one-way permutations where the black-box simulator only makes a "constant number of adaptive queries" (where adaptive queries are defined in an appropriate way).

Using the strong-duality transformation, we obtain an analogous result for zero-knowledge arguments as well.

Corollary 3 (Constant-round Zero-Knowledge Arguments) *For any constant d, only languages L in $\mathsf{CF}_{o(\frac{n}{\log n})}$ have d-round fully black-box computational zero-knowledge argument based on one-way permutations where the total verifier communication complexity $c(n)$ is sub-linear (i.e. $o(n)$).*

Proof: Applying the strong-duality lemma, there exists a protocol $\langle \widehat{P}, \widehat{V} \rangle$, that is public-coin protocol where the prover, verifier and the simulator have access to $\mathsf{Sam}^\pi_{o(\frac{n}{\log n})}$ that is $(\pi, \mathsf{Sam}^\pi_{o(\frac{n}{\log n})})$-relativized black-box zero-knowledge argument. Thus, using Theorem 4, we have that $L \in \mathsf{CF}_{o(\frac{n}{\log n})}$. ∎

If $\mathsf{NP} \not\subseteq \mathsf{CF}_{o(\frac{n}{\log n})}$, then the corollary is essentially tight. Feige and Shamir [5] and Pass and Wee [22] present an $O(1)$-round zero-knowledge arguments based on one-way functions. While, the former construction relies on one-way functions in a non black-box way, the latter is a fully black-box construction. Nevertheless, both the constructions require superlinear verifier communication complexity. On the other hand, efficient zero-knowledge arguments due to Kilian [17] have poly-logarithmic communication complexity, but are fully black-box based only on collision-resistant hash functions.

[1] Goldreich-Kahan use claw-free permutations to construct constant-round statistically-hiding commitments. However, these can be constructed under the potentially weaker assumption of collision-resistant hash functions [3, 15]. Therefore, there also exists constant-round black-box zero-knowledge proofs for all of NP based on collision-resistant hash functions.

5.2 Lower Bounds for Concurrent Zero Knowledge

The notion of concurrent zero-knowledge introduced by Dwork, Naor and Sahai [4], considers the execution of zero-knowledge in a concurrent setting. That is, a single adversary participates as a verifier in many concurrent executions (see [23] for a formal definition and discussion). Analogous strong and weak duality transformation for concurrent zero-knowledge proofs and arguments follow directly by the proof of Lemma 1 and 2. We now turn to prove our lower bounds.

Recently, Pass, Tseng and Wikström in [21] prove that only languages in BPP have public-coin black-box concurrent zero-knowledge proofs or arguments. As the result of Goldreich-Krawczyk [8], this proof uses the simulator as a black-box to decide the language L, and relativizes. We therefore have:

Theorem 5 (Implicit in [21]). *Let $\langle P^\bullet, V^\bullet \rangle$ be a $\mathcal{O}$-relativized public-coin concurrent zero-knowledge proof (or argument) for a language L with a black-box simulator S (and negligible soundness error). Then, there exists a $\mathcal{PPT}$ machine $M^\bullet$, such that M^{V^O, S^O} decides L with probability $\frac{2}{3}$ when the probability is taken over a uniformly chosen $O \in \mathcal{O}$ and the internal coin tosses of M.*

As corollary of the strong and weak duality transformation for concurrent zero-knowledge, we obtain the following.

Corollary 4 *A language L has a $o(\frac{n}{\log n})$-round fully black-box concurrent zero-knowledge proof (or argument with $o(n)$ verifier communication complexity) based on one-way permutations, then $L \in \mathsf{CF}_{o(\frac{n}{\log n})}$.*

This result is tight if $\mathsf{NP} \not\subseteq \mathsf{CF}_{o(\frac{n}{\log n})}$; Prabhakaran, Rosen and Sahai [20] provide a fully black-box constructions of $\omega(\log n)$-round concurrent zero-knowledge proofs, or arguments with polylogarithmic communication complexity, based on collision-resistant hash functions; Pass and Wee [22] provide an $O(n)$-round fully black-box argument based on one-way functions.

References

1. L. Babai and S. Moran. Arthur-Merlin games: A randomized proof system, and a hierarchy of complexity classes. *JCSS*, Vol. 36, pages 254–276, 1988.
2. B. Barak. How to go Beyond the Black-Box Simulation Barrier. In *42nd FOCS*, pages 106–115, 2001.
3. I. Damgård, T. Pedersen and B. Pfitzmann. On the Existence of Statistically Hiding Bit Commitment Schemes and Fail-Stop Signatures. In *Crypto93*, Springer-Verlag LNCS Vol. 773, pages 250–265, 1993.
4. C. Dwork, M. Naor and A. Sahai. Concurrent Zero-Knowledge. In *30th STOC*, pages 409–418, 1998.
5. U. Feige and A. Shamir. Zero Knowledge Proofs of Knowledge in Two Rounds. In *Crypto89*, Springer LNCS 435, pages. 526–544, 1989.
6. O. Goldreich. *Foundation of Cryptography – Basic Tools*. Cambridge University Press, 2001.

7. O. Goldreich and A. Kahan. How to Construct Constant-Round Zero-Knowledge Proof Systems for NP. *Journal of Cryptology*, Vol. 9, No. 2, pages 167–189, 1996.
8. O. Goldreich and H. Krawczyk. On the Composition of Zero-Knowledge Proof Systems. *SIAM Jour. on Computing*, Vol. 25(1), pages 169–192, 1996.
9. S. Goldwasser, S. Micali and C. Rackoff. The Knowledge Complexity of Interactive Proof Systems. *SIAM Jour. on Computing*, Vol. 18(1), pp. 186–208, 1989.
10. O. Goldreich, S. Micali and A. Wigderson. Proofs that Yield Nothing But Their Validity or All Languages in NP Have Zero-Knowledge Proof Systems. *JACM*, Vol. 38(1), pp. 691–729, 1991.
11. S. Goldwasser and M. Sipser. *Private Coins versus Public Coins in Interactive Proof Systems*. In *18th STOC*, pages 59–68, 1986.
12. S. D. Gordon and H. Wee and D. Xiao and A. Yerukhimovich. On the Round Complexity of Zero-Knowledge Proofs Based on One-Way Permutations. *Manuscript*, 2009.
13. I. Haitner, J. Hoch, O. Reingold and G. Segev. *Finding Collisions in Interactive Protocols - A Tight Lower Bound on the Round Complexity of Statistically-Hiding Commitments*. In *48th FOCS*, pages 669–679, 2007.
14. I. Haitner, J. Hoch and G. Segev. *A Linear Lower Bound on the Communication Complexity of Single-Server Private Information Retrieval*. In *5th TCC*, pages 445–464, 2008.
15. S. Halevi and S. Micali. Practical and Provably-Secure Commitment Schemes from Collision-Free Hashing. In *Crypto96*, Springer LNCS 1109, pages 201–215, 1996.
16. I. Haitner and M. Mahmoody-Ghidary, and D. Xiao. A constant-round public-coin protocol for sampling with size, and applications. Technical Report TR-867-09, Princeton University, 2009.
17. J. Kilian. A Note on Efficient Zero-Knowledge Proofs and Arguments. In *24th STOC*, pages 723–732, 1992.
18. J. Kilian and E. Petrank. Concurrent and resettable zero-knowledge in poly-logarithmic rounds. In *33rd STOC*, pages 560–569, 2001.
19. S. J. Ong and S. Vadhan. An Equivalence between Zero Knowledge and Commitments. In *5th TCC*, 482–500, 2008.
20. M. Prabhakaran, A. Rosen and A. Sahai. Concurrent zero-Knowledge with logarithmic round complexity. In *43rd FOCS*, pages 366–375, 2002.
21. R. Pass and W. Tseng and D. Wikström. On the Composition of Public-Coin Zero-Knowledge Protocols. In *CYPTO*, Springer LNCS 5677, pages 160–176, 2009.
22. R. Pass and H. Wee. *Black-box constructions of two-party primitives from one-way functions*. In *6th TCC*, 403–418, 2009.
23. A. Rosen. *Concurrent Zero-Knowledge*. Springer, 2006.
24. O. Reingold, L. Trevisan and S. Vadhan. *Notions of reducibility between cryptographic primitives*. In *1st TCC*, pp. 1–20, 2004.
25. D. Simon. *Finding collisions on a one-way street: Can secure hash functions be based on general assumptions?*. In *Eurocrypt*, pages 334–345, 1998.
26. S. Vadhan. *On Transformations of Interactive Proofs that Preserve Prover's Complexity*. In *32nd STOC*, pages 200–207, 2000.