Reducing Complexity Assumptions for Statistically-Hiding Commitment*

Iftach Haitner† Omer Horvitz‡ Jonathan Katz‡ Chiu-Yuen Koo‡
Ruggero Morselli‡ Ronen Shaltiel§

August 7, 2007

Abstract

We revisit the following question: *what are the minimal assumptions needed to construct statistically-hiding commitment schemes*? Naor et al. show how to construct such schemes based on any one-way permutation. We improve upon this by showing a construction based on any *approximable preimage-size* one-way function. These are one-way functions for which it is possible to efficiently approximate the number of pre-images of a given output. A special case is the class of *regular* one-way functions where all points in the image of the function have the same (known) number of pre-images.

We also prove two additional results related to statistically-hiding commitment. First, we prove a (folklore) *parallel composition theorem* showing, roughly speaking, that the statistical hiding property of any such commitment scheme is amplified exponentially when multiple independent parallel executions of the scheme are carried out. Second, we show a *compiler* which transforms any commitment scheme which is statistically hiding against an honest-but-curious receiver into one which is statistically hiding even against a malicious receiver.

1 Introduction

A central focus of modern cryptography has been to investigate the weakest possible assumptions under which various cryptographic primitives exist. This direction of research has been quite fruitful, and minimal assumptions are known for a wide variety of primitives: e.g., it has been shown that one-way functions imply (and are implied by) pseudorandom generators, pseudorandom functions, symmetric-key encryption/message authentication, statistically-binding commitment, and digital signatures [16, 17, 29, 30, 33, 35, 40]. In other cases, black-box separation results exist which indicate the difficulty of constructing "strong" cryptographic protocols (say, key-exchange) from "weak" building blocks (say, one-way permutations; see [31]).

*A preliminary version of this paper appeared at Eurocrypt 2005 [25].

†Department of Computer Science, Weizmann Institute of Science. email: `iftach.haitner@weizmann.ac.il`. Research supported by US-Israel Binational Science Foundation grant 2002246.

‡Department of Computer Science, University of Maryland. email: {`horvitz,jkatz,cykoo,ruggero`}`@cs.umd.edu`. Research of O.H. supported by U.S. Army Research Office award DAAD19-01-1-0494. Research of J.K. supported by NSF CAREER award #0447075.

§Department of Computer Science, University of Haifa. email: `ronen@cs.haifa.ac.il`. Research supported by US-Israel Binational Science Foundation grant 2004329.

In this work, we focus on constructing *statistically-hiding commitment schemes*. Informally, a commitment scheme defines a two-phase interactive protocol between a sender $\mathcal{S}$ and a receiver $\mathcal{R}$; after the *commitment phase*, $\mathcal{S}$ is uniquely bound to (at most) one value which is not yet revealed to $\mathcal{R}$, and in the *decommitment phase* $\mathcal{R}$ finally learns this value. The two security properties hinted at in this informal description are known as *binding* (namely, that $\mathcal{S}$ is bound to at most one value after the commitment phase) and *hiding* (namely, that $\mathcal{R}$ does not learn the value to which $\mathcal{S}$ commits before the decommitment phase). In a statistically-hiding commitment scheme the hiding property holds *even against all-powerful receivers* (i.e., hiding holds information-theoretically), while the binding property is required to hold only for computationally-bounded (say, polynomial-time) senders.

Statistically-hiding commitment schemes have been used as a building block in constructions of statistical zero-knowledge arguments [7, 34] and coin-tossing protocols [32]. They are also advantageous when used within protocols in which certain commitments are never revealed; in this case, one can argue that computational binding suffices since it need only be infeasible to violate the binding property *during the period of time the protocol is run*, whereas statistical hiding has the advantage of ensuring that committed values remain hidden *forever* (i.e., regardless of how much time the receiver invests after completion of the protocol). Indeed, this is part of the motivation for statistical zero-knowledge as well. For further discussion, the reader is referred to [34, 37, 38].

Perfectly-hiding[1] commitment schemes were first shown to exist based on specific number-theoretic assumptions [6, 7] or, more generally, based on any collection of claw-free permutations [10, 24] with an efficiently-recognizable index set [19] (see [19] for a weaker variant of statistically-hiding commitment which suffices for some applications and for which an efficiently-recognizable index set is not needed). Naor et al. [34], building on Ostrovsky et al. [37, 38], showed a construction of a perfectly-hiding commitment scheme based on any one-way permutation. Statistically-hiding commitment schemes can also be constructed from collision-resistant hash functions [35, 11, 28]; see [41] for assumptions implying the existence of the latter.

1.1 Our Results

1.1.1 Main Result

We show how to construct a statistically-hiding commitment scheme given any *approximable pre-image-size* one-way function. Informally, this is a one-way function f satisfying the additional property that, given any y in the image of f, the value $|\{x : f(x) = y\}|$ (i.e., the number of pre-images of y) can be efficiently estimated. An interesting special case is an *approximately-regular* one-way function for which every point in the image of f has roughly the same number of pre-images. (We still require that it be feasible to approximate the number of pre-images.) A variety of conjectured one-way functions are in fact regular; we refer the reader to [20] for examples.

Our result may be viewed as an example of the paradigm in which a sequence of works constructs a given primitive from ever-weaker assumptions; e.g., in the cases of pseudorandom generators and universal one-way hash functions/signature schemes (see [14, Chap. 2] and [15, Chap. 6]), constructions were first based on specific, number-theoretic assumptions [5, 24], and then the minimal assumptions were gradually reduced to trapdoor permutations [3], one-way per-

[1] Very informally, in a statistically-hiding commitment scheme the receiver learns at most a negligible amount of information about the sender's committed value, whereas in a perfectly-hiding commitment scheme the receiver learns *nothing*. Note that any perfectly-hiding scheme is also statistically-hiding.

mutations [5, 21, 35, 42], regular one-way functions [20, 12], and (finally) one-way functions [29, 40]. This work has similarly served as a step toward resolving the question of the minimal assumptions required for statistically-hiding commitment; see Section 1.3.

1.1.2 Additional Results

We also provide two additional results of independent interest that may be useful for future constructions of statistically-hiding commitment schemes. Before describing these results, we review the standard definition of statistical hiding. Say a commitment scheme $(\mathcal{S}, \mathcal{R})$ is *ρ-hiding against $\mathcal{R}^*$* if the distribution over the view of the (malicious) receiver $\mathcal{R}^*$ when the sender $\mathcal{S}$ commits to '0' is within statistical difference ρ from the distribution over the view of $\mathcal{R}^*$ when $\mathcal{S}$ commits to '1'. The standard definition of statistical hiding requires that for all (even all-powerful) $\mathcal{R}^*$, the commitment scheme should be ε-hiding against $\mathcal{R}^*$ for some negligible function ε. One way of relaxing this is to require only that the scheme be $(1 - \frac{1}{\mathsf{poly}})$-hiding (for all $\mathcal{R}^*$). An alternate relaxation is to require only that the scheme be ε-hiding against the honest receiver $\mathcal{R}$ (this corresponds to the classical cryptographic notion of an *honest-but-curious* adversarial entity). In all cases, we require binding to hold with all but negligible probability for any polynomial-time sender.

We show that a scheme satisfying either of the relaxations above suffices to construct a scheme secure in the standard sense, with minimal increase in the round complexity. Specifically:

1. We prove a *parallel repetition theorem* for statistically-hiding commitment. Given commitment scheme $(\mathcal{S}, \mathcal{R})$, consider the scheme $(\mathcal{S}^q, \mathcal{R}^q)$ in which commitment to a bit b is done as follows: $\mathcal{S}^q$ chooses random bits $b_1, \ldots, b_q$ subject to the constraint $\bigoplus_i b_i = b$, and then runs q parallel executions of $\mathcal{S}$ using input bit b_i in the i^{th} execution. We show that if the initial scheme $(\mathcal{S}, \mathcal{R})$ is ρ-hiding, then the derived scheme $(\mathcal{S}^q, \mathcal{R}^q)$ is ρ^q-hiding. A corollary is that the existence of a $(1 - \frac{1}{\mathsf{poly}})$-hiding scheme implies the existence of a statistically-hiding commitment scheme *using the same number of rounds*.

   Parallel repetition fails in many settings (e.g., [2, 39]), and so the above result should not be taken for granted. The result is trivial to prove for the case of an honest-but-curious receiver, but (as when analyzing the effect of parallel repetition on the soundness of interactive proofs [13, Appendix C]) is more difficult to prove for the case when a malicious receiver may correlate its actions in the different parallel executions.

2. We show a general *compiler* that converts any commitment scheme that is statistically-hiding for an honest-but-curious receiver into one that is statistically-hiding for a malicious receiver. If the initial scheme is ρ-hiding for an honest-but-curious receiver, we can obtain a scheme that is $(\rho + \frac{1}{\mathsf{poly}})$-hiding (for any given polynomial poly) using only a constant number of additional rounds. (Applying the previous result, we can then obtain a scheme that is ε-hiding — for a malicious receiver — without any further increase in the round complexity.) Our compiler requires only the existence of one-way functions, which are implied anyway by the commitment scheme we start with.

1.2 Overview of Our Techniques

Our construction is based on the protocol of Naor et al. [34], which is shown by those authors to be perfectly hiding (and computationally binding) when based on any one-way permutation. It is natural to ask what happens when their protocol is instantiated with some other function

$f : \{0,1\}^n \to \{0,1\}^\ell$. Our first observation, that can also be derived from subsequent work in this area [36, 26], is that it is implicit in the main proof of [34] that their protocol is computationally binding as long as f cannot be efficiently inverted with respect to the uniform distribution U_ℓ over its range (formally, we mean that no efficient algorithm can find an x such that $f(x) = y$, for uniformly-chosen y, with non-negligible probability; see Definition 2.8). We call a function with this property *one-way over its range*. We stress that a function with this property is not necessarily one-way in the standard sense: the constant function $f(x) = 0^\ell$ is not one-way, but is trivially one-way over its range since the probability that a uniformly-selected $y \in \{0,1\}^\ell$ lies in the image of f (that is, the probability that $y = 0^\ell$) is negligible.

As our first main technical result, we show that the protocol of Naor et al. is "weakly hiding" when based on a function $f : \{0,1\}^n \to \{0,1\}^\ell$ that is *balanced*; i.e., for which the probability that $f(U_n) = y$ is "close" to $2^{-\ell}$ for "most" elements $y \in \{0,1\}^\ell$. (In the formal definition we allow some elements to have probability outside this range as long as both the number of such elements and their total weight are small; see Definition 3.1.)

Taken together, the above show that statistically-hiding commitment is implied by the existence of a function f that is both balanced and one-way over its range.[2] We then show how to construct such functions based on any approximately-regular one-way function. Inspired by [29, 40], we use pairwise-independent hashing to achieve this goal. Restricting our attention here to functions f that are *regular*, we define $f'(h, x) = (h, h(f(x)))$ where h is selected from a family of pairwise-wise independent hash functions. Intuitively, if the output length of h is "small enough" (relative to the regularity parameter[3] of f) then f' will be sufficiently balanced, while if the output length of h is "large enough" then f' will be one-way over its range. We show that it is possible to set the output length "in the middle" and obtain a function that is one-way over its range and is somewhat balanced. Such a function translates into a bit-commitment protocol that is only somewhat hiding, but the hiding property can then be amplified by repetition. We stress that our construction requires that the regularity parameter of f is known.

The same construction works also if f is only approximately regular. It is also fairly easy to show how to convert any approximable pre-image-size one-way function into a one-way function that is approximately regular.

1.3 Subsequent Work

Subsequent to the initial publication of this paper, Nguyen et al. [36] showed how to construct statistical zero-knowledge arguments for $\mathcal{NP}$ based on any one-way function; as mentioned in their paper, some of the intuition for their construction builds on the work described here. Recall that statistical zero-knowledge arguments are one of the primary applications of statistically-hiding commitment schemes; interestingly, Nguyen et al. construct the former without using the latter. Haitner and Reingold [27], using tools developed in [36], gave a construction of a statistically-hiding commitment scheme from any one-way function. This settles the question of the minimal assumptions needed to construct such commitment schemes.

Other work of Haitner and Reingold [26] (done subsequent to the present paper) provides a new analysis of the protocol from [34], from which they derive our main result more directly (cf. [26,

[2] The "almost-everywhere one-to-one" one-way functions of [18] are not balanced (since their image is a negligible fraction of their range) and thus do not suffice for our purposes. The same also holds for the functions constructed in [29] and [14, Sect. 3.5].

[3] That is, the number of pre-images of each value in the image of f.

Theorem A.10]).

1.4 Outline of the Paper

We begin by reviewing some preliminaries and establishing some notation in Section 2. In that section, we also note that any approximable pre-image size one-way function can be converted into an approximately-regular one-way function. In Section 3, we formally define the notion of "balanced" functions described informally earlier, and show that any balanced function that is one-way over its range can be used to construct a statistically-hiding commitment scheme. Our task is thus reduced to constructing such a function starting from any approximately-regular one-way function, and we tackle this in Section 4. This completes the proof of our main result.

In Section 5 we prove a parallel composition theorem for statistically-hiding commitment, and in Section 6 we show a compiler converting any commitment scheme statistically-hiding for an honest-but-curious receiver into one that is statistically-hiding for a malicious receiver.

2 Preliminaries

Throughout this paper, we let k denote the security parameter. We use 'log' to denote logarithms base 2, and 'ln' to denote the natural logarithm. For a function $f : \{0,1\}^n \to \{0,1\}^\ell$, we define $\mathsf{image}(f) \stackrel{\text{def}}{=} \{f(x) \mid x \in \{0,1\}^n\}$.

2.1 Probability Distributions and Entropy

If X is a distribution over a finite set $\mathcal{X}$, the *support* of X (denoted $\mathsf{supp}(X)$) consists of those elements having non-zero probability under X. The *min-entropy* of X is defined as:

$$H_\infty(X) \stackrel{\text{def}}{=} \min_{x \in \mathsf{supp}(X)} \log\left(\frac{1}{\Pr_X[x]}\right).$$

The *collision probability* of X is defined as:

$$CP(X) \stackrel{\text{def}}{=} \sum_{x \in \mathsf{supp}(X)} (\Pr_X[x])^2.$$

It is convenient to think of collision probability as a notion of entropy and for this purpose we normalize it as follows: The 2-*entropy* of X is defined as:

$$H_2(X) \stackrel{\text{def}}{=} \log \frac{1}{CP(X)}.$$

If X_1 and X_2 are two distributions over a finite set $\mathcal{X}$, their statistical difference, written $\mathsf{SD}(X_1, X_2)$, is defined as:

$$\mathsf{SD}(X_1, X_2) \stackrel{\text{def}}{=} \frac{1}{2} \sum_{x \in \mathcal{X}} \left|\Pr_{X_1}[x] - \Pr_{X_2}[x]\right|.$$

Two distribution ensembles $\mathcal{X}_1 = \{X_1(k)\}_{k \in \mathbb{N}}$ and $\mathcal{X}_2 = \{X_2(k)\}_{k \in \mathbb{N}}$ have statistical difference ρ (for ρ a function of k) if $\mathsf{SD}(X_1(k), X_2(k)) \leq \rho(k)$ for all k large enough. If ρ is negligible, we say the ensembles are *statistically indistinguishable*.

We let U_n denote the uniform distribution over $\{0,1\}^n$. For a function $f : \{0,1\}^n \to \{0,1\}^\ell$, we let $f(U_n)$ denote the distribution over $\{0,1\}^\ell$ induced by choosing x uniformly and outputting $f(x)$.

2.2 Commitment Schemes

An interactive bit commitment scheme is defined via a triple of PPT algorithms $(\mathcal{S}, \mathcal{R}, \mathcal{V})$. Looking ahead, $\mathcal{S}$ and $\mathcal{R}$ will interact during what is called a *commitment phase*, while $\mathcal{V}$ will be used during the (non-interactive) *decommitment phase*. Formally:

- $\mathcal{S}$ (the *sender*) is an interactive Turing machine (ITM) which receives as initial input the security parameter 1^k and a bit b. Following its interaction, it outputs some information decom (the *decommitment*).

- $\mathcal{R}$ (the *receiver*) is an ITM which receives the security parameter 1^k as initial input. Following its interaction, it outputs some state information com.

- $\mathcal{V}$ (acting as a receiver, in the decommitment phase) is a deterministic algorithm which receives as input state information com and a decommitment decom; it outputs either a bit b or the distinguished value $\perp$.

Denote by $(\mathsf{decom} \mid \mathsf{com}) \leftarrow \langle \mathcal{S}(1^k, b), \mathcal{R}(1^k) \rangle$ the experiment in which $\mathcal{S}$ and $\mathcal{R}$ interact (using the given inputs and uniformly random coins), and then $\mathcal{S}$ outputs decom while $\mathcal{R}$ outputs com. It is required that for all k, all b, and every pair $(\mathsf{decom} \mid \mathsf{com})$ that may be output by $\langle \mathcal{S}(1^k, b), \mathcal{R}(1^k) \rangle$, it is the case that $\mathcal{V}(\mathsf{com}, \mathsf{decom}) = b$.

The security of a commitment scheme can be defined in two complementary ways, protecting against either an all-powerful sender or an all-powerful receiver. Since we are interested in the case of statistically-hiding commitment (i.e., the latter case), we only provide the definition for this case.

Definition 2.1. Commitment scheme $(\mathcal{S}, \mathcal{R}, \mathcal{V})$ is *ρ-hiding* (for ρ a function of k) if the following holds: Given a deterministic ITM $\mathcal{R}^*$, let $\mathsf{view}_{\langle \mathcal{S}(b), \mathcal{R}^* \rangle}(k)$ denote the distribution on the view of $\mathcal{R}^*$ when interacting with $\mathcal{S}(1^k, b)$ (this view simply consists of the sequence of messages it receives from $\mathcal{S}$), where this distribution is taken over the random coins of $\mathcal{S}$. Then we require that for any (even all-powerful) $\mathcal{R}^*$ the ensembles $\{\mathsf{view}_{\langle \mathcal{S}(0), \mathcal{R}^* \rangle}(k)\}$ and $\{\mathsf{view}_{\langle \mathcal{S}(1), \mathcal{R}^* \rangle}(k)\}$ have statistical difference at most $\rho(k)$.

A commitment scheme is *statistically hiding* if it is ρ-hiding for negligible ρ. A 0-hiding scheme is called *perfectly hiding*.

Assuming $\mathcal{R}^*$ to be deterministic is without loss of generality since $\mathcal{R}^*$ may be all-powerful.

Definition 2.2. Commitment scheme $(\mathcal{S}, \mathcal{R}, \mathcal{V})$ is *computationally-binding* if the following is negligible for all PPT $\mathcal{S}^*$:

$$\Pr\left[((\mathsf{decom}, \mathsf{decom}') \mid \mathsf{com}) \leftarrow \langle \mathcal{S}^*(1^k), \mathcal{R}(1^k) \rangle : \begin{array}{c} \mathcal{V}(\mathsf{com}, \mathsf{decom}), \mathcal{V}(\mathsf{com}, \mathsf{decom}') \in \{0,1\} \wedge \\ \mathcal{V}(\mathsf{com}, \mathsf{decom}) \neq \mathcal{V}(\mathsf{com}, \mathsf{decom}') \end{array} \right],$$

where the probability is taken over the random coins of both $\mathcal{S}^*$ and $\mathcal{R}$.

Given the above, we now define a statistically-hiding commitment scheme:

Definition 2.3. Commitment scheme $(\mathcal{S}, \mathcal{R}, \mathcal{V})$ is *ρ-secure* (resp., *statistically secure*, *perfectly secure*) if it is computationally binding and ρ-hiding (resp., statistically hiding, perfectly hiding).

2.3 One-Way Function Families and Variants

All function families $\mathcal{F} = \{f_k : \{0,1\}^{n(k)} \to \{0,1\}^{\ell(k)}\}_{k\in\mathbb{N}}$ in this paper will have $n, \ell = \mathsf{poly}(k)$, and n, f_k computable in time polynomial in k. We say $\mathcal{F}$ is *one-way* if, for all PPT algorithms A, the following is negligible (in k):

$$\Pr[x \leftarrow \{0,1\}^{n(k)}; y = f_k(x); x' \leftarrow A(1^k, y) : f_k(x') = y].$$

We also consider some additional properties of function families:

- $\mathcal{F}$ is $r(k)$**-regular** if for every k and every $x \in \{0,1\}^{n(k)}$ we have

  $$\left|\{x' \in \{0,1\}^{n(k)} : f_k(x') = f_k(x)\}\right| = 2^{r(k)}.$$

  $\mathcal{F}$ is $r(k)$**-known regular** if, in addition, $r(k)$ is computable in time polynomial in k.

- $\mathcal{F}$ is $(r(k), p(k))$**-approximately-regular** if for every k and every $x \in \{0,1\}^{n(k)}$ we have

  $$\frac{1}{p(k)} \cdot 2^{r(k)} \le \left|\{x' \in \{0,1\}^{n(k)} : f_k(x') = f_k(x)\}\right| \le p(k) \cdot 2^{r(k)},$$

  and $r(k), p(k)$ are computable in time polynomial in k. We will be interested in the case where $p(k)$ is upper-bounded by a polynomial in k.

  Note that if f is (r,p)-approximately-regular, then the min-entropy of $D = f(U_n)$ satisfies

  $$n - r - \log p \le H_\infty(D) \le n - r + \log p.$$

- $\mathcal{F}$ is $p(k)$**-approximable pre-image-size** if for every k and every $x \in \{0,1\}^{n(k)}$ we have

  $$\frac{1}{p(k)} \cdot 2^{D(f_k(x))} \le \left|\{x' \in \{0,1\}^{n(k)} : f_k(x') = f_k(x)\}\right| \le p(k) \cdot 2^{D(f_k(x))},$$

  and p, D are computable in time polynomial in k. As in the approximately-regular case, we will be interested in the case where $p(k)$ is upper-bounded by a polynomial in k.

For simplicity, we sometimes drop the explicit dependence on k when clear and write, e.g., $f : \{0,1\}^n \to \{0,1\}^\ell$ rather than $f_k : \{0,1\}^{n(k)} \to \{0,1\}^{\ell(k)}$.

By the following lemma, to prove our main result it will be sufficient for us to construct a statistically-hiding commitment scheme starting from any approximately-regular one-way function.

Lemma 2.4. *Let $\mathcal{F} = \{f_k : \{0,1\}^{n(k)} \to \{0,1\}^{\ell(k)}\}_{k\in\mathbb{N}}$ be a $p(k)$-approximable pre-image-size one-way function family. Let $\mathcal{F}' = \{f'_k : \{0,1\}^{2n(k)} \to \{0,1\}^{n(k)+\ell(k)}\}_{k\in\mathbb{N}}$, where*

$$f'_k(x\|z) \stackrel{\text{def}}{=} f_k(x) \parallel z[1 \ldots D(f_k(x))] \parallel 0^{n-D(f_k(x))}.$$

(In the above, "$\|$" denotes concatenation and $z[1 \ldots r]$ denotes the first r bits of z.) Then $\mathcal{F}'$ is an (n,p)-approximately-regular one-way function family.

Proof. The one-wayness of $\mathcal{F}'$ is evident. The number of pre-images of an element $y\|\bar{z}\|0^{n-D(y)}$ satisfies

$$\begin{aligned}\left|(f'_k)^{-1}\left(y \parallel \bar{z} \parallel 0^{n-D(y)}\right)\right| &= \left|f_k^{-1}(y)\right| \cdot 2^{n-D(y)} \\ &\in \left[\frac{2^{D(y)}}{p(k)} \cdot 2^{n-D(y)},\ p(k) \cdot 2^{D(y)} \cdot 2^{n-D(y)}\right] \\ &= \left[\frac{2^n}{p(k)},\ p(k) \cdot 2^n\right].\end{aligned}$$

□

2.4 Hash functions and the Leftover Hash Lemma

Let $\mathcal{H} = \{H_k\}_{k\in\mathbb{N}}$ be a collection of function families, where each H_k is a family of functions mapping strings of length $\ell(k)$ to strings of length $v(k)$. We assume that the size of each H_k is a power of 2, and that we can identify each binary string of some appropriate length $s(k)$ with a unique function $h \in H_k$. (In particular, choosing random $h \in H_k$ is identified with choosing a random string of length $s(k)$.) Following [8], we say that $\mathcal{H}$ is an $n(k)$-*universal hash family* (i.e., an $n(k)$-wise independent hash family) if for each k, any distinct $x_1, \ldots, x_{n(k)} \in \{0,1\}^{\ell(k)}$, and any $y_1, \ldots, y_{n(k)} \in \{0,1\}^{v(k)}$ we have:

$$\Pr_{h\leftarrow H_k}[h(x_1) = y_1 \wedge \cdots \wedge h(x_{n(k)}) = y_{n(k)}] = 2^{-v(k)\cdot n(k)}.$$

Put another way, for any fixed, distinct $x_1, \ldots, x_{n(k)}$, the random variables $h(x_1), \ldots, h(x_{n(k)})$ (where h is chosen uniformly from H_k) are n-wise independent. Constructions of $n(k)$-universal hash families with $s(k) = O(n(k) \cdot \max(\ell(k), v(k)))$ are known [1, 9]. Simpler constructions exist for $n = 2$, and these are sufficient for achieving our results. We will rely on the following result:

Lemma 2.5 (Leftover hash lemma [29]). *Let Y be a distribution over $\{0,1\}^\ell$ with $H_\infty(Y) \geq q$ and let $\varepsilon > 0$. Let H be a family of 2^s functions mapping strings of length ℓ to strings of length $t \leq q - 2\log(1/\varepsilon) - 2$ and assume that H is a 2-universal family of hash functions. Consider the distribution P obtained by choosing uniformly a function h from the family H and y according to Y and outputting $(h\|h(y))$. Then:*

1. *$H_2(P) \geq s + t - \varepsilon^2$.*
2. *The statistical distance between P and U_{s+t} is at most ε.*

We remark that the first item above implies the second item. In the proof we sometimes need the first item whereas in other cases the weaker conclusion of the second item suffices.

2.5 Interactive Hashing

Interactive hashing was introduced by Ostrovsky, et al. [37, 38], and used by Naor, et al. [34] to construct a statistically-hiding (actually, perfectly-hiding) commitment scheme based on any one-way permutation family. We review interactive hashing, as well as the resulting commitment scheme, below. In what follows, we let $x \cdot y$ denote $\sum_{i=1}^{\ell} x_i y_i \bmod 2$ for $x, y \in \{0,1\}^\ell$.

Construction 2.6 (Interactive hashing). *The protocol is defined by algorithms $\mathcal{S}$ and $\mathcal{R}$, where $\mathcal{S}$ begins with an ℓ-bit value y (with ℓ known to $\mathcal{R}$), and proceeds as follows:*

1. *The parties interact in $\ell - 1$ stages. In stage i (for $i = 1, \ldots, \ell - 1$), $\mathcal{R}$ chooses $r_i \in \{0,1\}^{\ell - i}$ uniformly at random and sends the "query" $q_i = 0^{i-1}1r_i$ to $\mathcal{S}$ (in case $\mathcal{R}$ aborts, $\mathcal{S}$ simply takes q_i to be some default value); in response, $\mathcal{S}$ sends $c_i = q_i \cdot y$.*

2. *At the conclusion of the above, there are exactly two strings $y_0, y_1 \in \{0,1\}^\ell$ such that $q_i \cdot y_b = c_i$ for $1 \le i \le \ell - 1$; let y_0 denote the lexicographically smaller of the two. Both parties compute (y_0, y_1), and $\mathcal{S}$ sets v such that $y = y_v$.*

The output of the protocol is defined to be (y_0, y_1, v) for $\mathcal{S}$ and (y_0, y_1) for $\mathcal{R}$. We denote by $IH(y)$ an execution of the interactive hashing protocol, where $\mathcal{S}$ begins with input y.

The above was used in [34] to construct a perfectly-secure commitment scheme based on one-way permutations via the following approach:

Construction 2.7. *Let $\mathcal{F} = \{f_k : \{0,1\}^{n(k)} \to \{0,1\}^{\ell(k)}\}$ be a function family. The commitment scheme $(\mathcal{S}, \mathcal{R}, \mathcal{V})$ is defined as follows: $\mathcal{S}(1^k, b)$ chooses $x \in \{0,1\}^{n(k)}$ uniformly at random, computes $y = f_k(x)$, and then executes $IH(y)$ with $\mathcal{R}$; this protocol results in output (y_0, y_1, v) for $\mathcal{S}$ and (y_0, y_1) for $\mathcal{R}$. The commitment phase concludes by having $\mathcal{S}$ send $\hat{v} = v \oplus b$ to $\mathcal{R}$. Finally, $\mathcal{S}$ outputs $\mathsf{decom} = x$ while $\mathcal{R}$ outputs state $\mathsf{com} = (y_0, y_1, \hat{v})$.*

In the decommitment phase, $\mathcal{V}((y_0, y_1, \hat{v}), x)$ proceeds as follows: if $f_k(x) = y_0$, output $\hat{v}$; if $f_k(x) = y_1$, output $\hat{v} \oplus 1$; otherwise, output $\perp$.

It is relatively easy to observe that the above protocol is perfectly hiding if $\ell = n$ and $\mathcal{F}$ is a permutation family (regardless of whether $\mathcal{F}$ is one-way). The main result of [34] was to prove that the above is *computationally binding* when $\mathcal{F}$ is a *one-way* permutation family. In fact, careful examination of their proof shows the above commitment scheme is computationally binding under a *weaker* condition on $\mathcal{F}$; it suffices for $\mathcal{F}$ to be *one-way over its range*, defined as follows:

Definition 2.8. Let $\mathcal{F} = \{f_k : \{0,1\}^{n(k)} \to \{0,1\}^{\ell(k)}\}_{k \in \mathbb{N}}$ be a function family. We say $\mathcal{F}$ is *one-way over its range* if, for all PPT A, the following is negligible (in k):

$$\Pr[y \leftarrow \{0,1\}^{\ell(k)}; x \leftarrow A(1^k, y) : f_k(x) = y].$$

We stress that, in contrast to the definition in the case of a (standard) one-way function, here y is chosen uniformly in the range of f_k rather than according to $f(U_n)$.

The following result was implicit in [34], and a proof can be easily derived from [36, Theorem 4.4] or [26, Theorem 4.2]:

Theorem 2.9. *If $\mathcal{F}$ is one-way over its range, then Construction 2.7 is computationally binding.*

Proof. Let W be a binary relation and let $\mathcal{S}^*$ be a polynomial-time algorithm that, after playing the role of $\mathcal{S}$ in Construction 2.6, outputs with probability ε two elements x_0 and x_1 for which (x_0, y_0) and (x_1, y_1) are in W. By [26, Theorem 4.2], there exists an efficient oracle algorithm $A^{\mathcal{S}^*}$ such that $\Pr_{y \leftarrow \{0,1\}^\ell}[(A^{\mathcal{S}^*}(y), y) \in W] = \Omega(\mathsf{poly}(\varepsilon))$. Taking $W = \{(x, f(x)) : x \in \{0,1\}^{n(k)}\}$, it follows that Construction 2.7 is computationally binding if f is one-way over its range. $\square$

3 Statistical Hiding from Balanced Functions

In this section we define a notion of "balance" and show that if a function family $\mathcal{F}$ is sufficiently balanced then Construction 2.7 yields a protocol that is "somewhat hiding." Roughly speaking, a distribution D on $\{0,1\}^\ell$ is balanced if D is "close" to uniform "most" of the time. A function $f : \{0,1\}^n \to \{0,1\}^\ell$ is then defined to be balanced if the distribution $f(U_n)$ is balanced. Formally:

Definition 3.1. Distribution D on $\{0,1\}^\ell$ is (α, δ)-*balanced* if there is a set $\mathsf{Bad} \subset \{0,1\}^\ell$ such that:

1. $|\mathsf{Bad}| \leq \alpha \cdot 2^\ell$.
2. $\Pr_{y \leftarrow D}[y \in \mathsf{Bad}] \leq \alpha$.
3. For every $y_0 \notin \mathsf{Bad}$, $\left|\Pr_{y \leftarrow D}[y = y_0] - \frac{1}{2^\ell}\right| \leq \frac{\delta}{2^\ell}$ (we will always have $\delta < 1$).

Function $f : \{0,1\}^n \to \{0,1\}^\ell$ is (α, δ)-balanced if the distribution $f(U_n)$ is (α, δ)-balanced. Function family $\mathcal{F} = \{f_k : \{0,1\}^{n(k)} \to \{0,1\}^{\ell(k)}\}$ is (α, δ)-balanced if, for all k large enough, f_k is $(\alpha(k), \delta(k))$-balanced.

Our main result of this section is the following:

Theorem 3.2. *If $\mathcal{F} = \{f_k : \{0,1\}^{n(k)} \to \{0,1\}^{\ell(k)}\}$ is an (α, δ)-balanced function family, then Construction 2.7 is ρ-hiding for $\rho = 2\alpha + \delta + \alpha\delta$.*

Proof. Fix k large enough so that f_k is $(\alpha(k), \delta(k))$-balanced; from now on we simply write f, α, δ, ρ without explicitly indicating their dependence on k. For a given execution of the scheme, let τ denote the initial transcript resulting from the interactive hashing sub-protocol; thus, the view of $\mathcal{R}^*$ consists of τ and the bit $\hat{v}$ sent in the final round. Given a particular (deterministic) $\mathcal{R}^*$, we write $\mathsf{Exp}(b)$ to denote the experiment in which $\mathcal{S}$ chooses a uniform random tape and then executes the protocol with $\mathcal{R}^*$ using this random tape and the bit b, resulting in view $(\tau, \hat{v})$ for $\mathcal{R}^*$. Note that the distribution on τ is identical in $\mathsf{Exp}(0)$ and $\mathsf{Exp}(1)$, since the first phase of the commitment scheme is independent of b.

Below, we define a "good" set of initial transcripts Good, and show that:

Claim 3.3. $\Pr_{\mathsf{Exp}(0)}[\tau \in \mathsf{Good}] = \Pr_{\mathsf{Exp}(1)}[\tau \in \mathsf{Good}] \geq 1 - \alpha(2+\delta)$. *Since this probability is independent of the bit b being committed to, we simply write $\Pr_{\mathsf{Exp}}[\tau \in \mathsf{Good}]$ for this probability.*

Claim 3.4. *The following holds for all $\tau^* \in \mathsf{Good}$ and $\hat{v}^* \in \{0,1\}$:*

$$\left| \Pr_{\mathsf{Exp}(0)}[\hat{v} = \hat{v}^* \mid \tau = \tau^*] - \Pr_{\mathsf{Exp}(1)}[\hat{v} = \hat{v}^* \mid \tau = \tau^*] \right| \leq \delta.$$

These claims suffice to prove the theorem, since the statistical difference between the view of $\mathcal{R}^*$ when the sender commits to 0 (i.e., $b = 0$) and the view of $\mathcal{R}^*$ when the sender commits to 1 (i.e.,

$b = 1$) may be bounded as:

$$\begin{aligned}
&\frac{1}{2}\sum_{\tau^*,\hat{v}^*}\left|\Pr_{\mathsf{Exp}(0)}[(\tau,\hat{v})=(\tau^*,\hat{v}^*)] \;-\; \Pr_{\mathsf{Exp}(1)}[(\tau,\hat{v})=(\tau^*,\hat{v}^*)]\right| \\
&\quad= \;\frac{1}{2}\sum_{\tau^*,\hat{v}^*}\left|\Pr_{\mathsf{Exp}(0)}[\tau=\tau^*]\Pr_{\mathsf{Exp}(0)}[\hat{v}=\hat{v}^*|\tau=\tau^*] \;-\; \Pr_{\mathsf{Exp}(1)}[\tau=\tau^*]\Pr_{\mathsf{Exp}(1)}[\hat{v}=\hat{v}^*|\tau=\tau^*]\right| \\
&\quad\leq \;\Pr_{\mathsf{Exp}}[\tau\notin\mathsf{Good}] + \frac{1}{2}\sum_{\tau^*\in\mathsf{Good},\hat{v}^*}\Pr_{\mathsf{Exp}}[\tau=\tau^*]\cdot\left|\Pr_{\mathsf{Exp}(0)}[\hat{v}=\hat{v}^*|\tau=\tau^*] \;-\; \Pr_{\mathsf{Exp}(1)}[\hat{v}=\hat{v}^*|\tau=\tau^*]\right| \\
&\quad\leq \;\alpha(2+\delta) + \frac{1}{2}\sum_{\tau^*\in\mathsf{Good};\hat{v}^*}\Pr[\tau=\tau^*]\cdot\delta \quad\leq\quad \alpha(2+\delta)+\delta.
\end{aligned}$$

We now prove the two stated claims. Let $\mathsf{Bad} \subset \{0,1\}^\ell$ be a subset whose existence is guaranteed by Definition 3.1 (using the fact that f is balanced). Recall that the initial transcript τ defines two strings $y_0^\tau, y_1^\tau \in \{0,1\}^\ell$ (cf. Construction 2.6). We say $\tau \in \mathsf{Good}$ if and only if $y_0^\tau, y_1^\tau \notin \mathsf{Bad}$.

We first bound the probability that $y_v = y$ is in Bad (recall that y_v is the value that the sender starts with; cf. Construction 2.6). Since f is (α,δ)-balanced and y is distributed according to $f(U_n)$ (cf. Construction 2.7), it follows immediately that $y_v \in \mathsf{Bad}$ with probability at most α.

Next, we bound the probability that $y_v \notin \mathsf{Bad}$ but $y_{\bar{v}} \in \mathsf{Bad}$. Since $\mathcal{R}^*$ is deterministic, we have that $y_{\bar{v}}$ is uniquely determined by y_v. Let ϕ be the function mapping the sender's chosen value y_v to the second value $y_{\bar{v}}$ resulting from the interactive hashing protocol. Let $\mathsf{MapToBad} \stackrel{\text{def}}{=} \phi^{-1}(\mathsf{Bad})$. Observe that if $\phi(y) = y'$ then $\phi(y') = y$; this is because, for either of these choices, the sender responds with the exact same answer to each of the receiver's queries during the interactive hashing sub-protocol. It follows that ϕ is a permutation and $|\mathsf{MapToBad}| = |\mathsf{Bad}|$. We have:

$$\begin{aligned}
\Pr[y_v \notin \mathsf{Bad} \wedge y_{\bar{v}} \in \mathsf{Bad}] &= \Pr[y_v \in \mathsf{MapToBad} \setminus \mathsf{Bad}] \\
&= \sum_{y^*\in\mathsf{MapToBad}\setminus\mathsf{Bad}} \Pr[y_v = y^*] \\
&\leq \sum_{y^*\in\mathsf{MapToBad}\setminus\mathsf{Bad}} \frac{1+\delta}{2^\ell},
\end{aligned}$$

using condition 3 of Definition 3.1 and the fact that $y^* \notin \mathsf{Bad}$. Continuing:

$$\begin{aligned}
\sum_{y^*\in\mathsf{MapToBad}\setminus\mathsf{Bad}} \frac{1+\delta}{2^\ell} &= |\mathsf{MapToBad}\setminus\mathsf{Bad}|\cdot\frac{1+\delta}{2^\ell} \\
&\leq |\mathsf{MapToBad}|\cdot\frac{1+\delta}{2^\ell} \\
&= |\mathsf{Bad}|\cdot\frac{1+\delta}{2^\ell} \;\leq\; \alpha\cdot(1+\delta),
\end{aligned} \tag{1}$$

using condition 1 of Definition 3.1. It follows that $\tau \notin \mathsf{Good}$ with probability at most $\alpha\cdot(2+\delta)$, completing the proof of the first claim.

We proceed to prove the second claim. Let $P(\tilde{y}) \stackrel{\text{def}}{=} \Pr_{x\in\{0,1\}^n}[f(x)=\tilde{y}]$. For any τ^* and any

$\hat{v}^* \in \{0,1\}$ we have

$$\begin{aligned}
\Pr_{\mathsf{Exp}(b)}[\hat{v} = \hat{v}^* \mid \tau = \tau^*] &= \Pr_{\mathsf{Exp}(b)}[v = \hat{v}^* \oplus b \mid \tau = \tau^*] \\
&= \Pr_{\mathsf{Exp}(b)}[y = y^{\tau^*}_{\hat{v}^* \oplus b} \mid \tau = \tau^*] \\
&= \frac{P(y^{\tau^*}_{\hat{v}^* \oplus b})}{P(y^{\tau^*}_0) + P(y^{\tau^*}_1)}.
\end{aligned}$$

If $\tau^* \in \mathsf{Good}$, then $y_0^{\tau^*}, y_1^{\tau^*} \notin \mathsf{Bad}$ and so $P(y_0^{\tau^*}), P(y_1^{\tau^*})$ lie in the range $[(1-\delta)2^{-\ell}, (1+\delta)2^{-\ell}]$. It follows that when $\tau^* \in \mathsf{Good}$ the following holds for any $\hat{v}^* \in \{0,1\}$:

$$\left| \Pr_{\mathsf{Exp}(0)}[\hat{v} = \hat{v}^* \mid \tau = \tau^*] - \Pr_{\mathsf{Exp}(1)}[\hat{v} = \hat{v}^* \mid \tau = \tau^*] \right| = \frac{\left|P(y_0^{\tau^*}) - P(y_1^{\tau^*})\right|}{P(y_0^{\tau^*}) + P(y_1^{\tau^*})} \leq \delta.$$

(The last inequality follows from the fact that the second-to-last expression is maximized by taking $P(y_0^{\tau^*}) = (1+\delta)2^{-\ell}$ and $P(y_1^{\tau^*}) = (1-\delta)2^{-\ell}$.) This proves the claim and completes the proof of Theorem 3.2. □

Combining the above and Theorem 2.9 we obtain:

Corollary 3.5. *If $\mathcal{F}$ is (α, δ)-balanced function family and one-way over its range, then Construction 2.7 is a $(2\alpha + \delta + \alpha\delta)$-secure commitment scheme.*

We see that if $2\alpha + \delta + \alpha\delta \leq 1 - \frac{1}{\mathsf{poly}(k)}$, then we obtain a "weakly hiding" commitment scheme. This statistical difference can be amplified to give a statistically-hiding scheme (i.e., an ε-hiding scheme for negligible ε) using polynomially-many sequential repetitions (an appropriate sequential composition theorem is easy to prove). In Section 5 we prove a parallel composition theorem which also enables amplification of the statistical hiding property using polynomially-many repetitions but without increasing the round complexity.

In the following section, we show how to construct an $\mathcal{F}$ with the required properties starting from any $\mathcal{F}$ which is one-way and approximately regular. Applying the observation at the end of Section 2.3, we thus obtain a construction of a statistically-hiding commitment scheme from any approximable pre-image-size one-way function.

4 Starting from Approximately-Regular One-Way Functions

As discussed at the very end of the previous section, we show here that given an $(r(k), \mathsf{poly}(k))$-approximately-regular one-way function family $\mathcal{F}$ it is possible to construct a $(1/k, 1/k)$-balanced function family $\mathcal{F}'$ that is also one-way over its range. The construction is as follows:

Construction 4.1. *Let $\mathcal{F} = \{f_k : \{0,1\}^{n(k)} \to \{0,1\}^{\ell(k)}\}_{k \in \mathbb{N}}$ be a family of functions, and let $\mathcal{H} = \{H_k\}$ be a 2-universal hash family where each $h \in H_k$ maps strings of length $\ell(k)$ to strings of length $t(k)$, and each such h can be described using $s(k)$ bits. Define*

$$\mathcal{F}' = \left\{ f'_k : \{0,1\}^{s(k)+n(k)} \to \{0,1\}^{s(k)+t(k)} \right\}_{k \in \mathbb{N}}$$

via $f'_k(h, x) = (h, h(f_k(x)))$.

The main result of this section is the following.

Theorem 4.2. *Let $\mathcal{F}$ be an $(r(k), p(k))$-approximately-regular one-way function family with $p(k) = k^{O(1)}$, let $\mathcal{F}'$ be as in Construction 4.1 for $t(k) = n(k) - r(k) - \log p(k) - 8\log k$. Then, for all sufficiently-large k $\mathcal{F}'$ is $(1/k, 1/k)$-balanced and one-way over its range.*

Note that applying Theorems 2.9 and 3.2 we immediately obtain a bit-commitment scheme that is computationally binding and ρ-hiding for $\rho = O(1/k)$. The hiding property of such a scheme can be amplified to give a statistically-hiding scheme (one with negligible ρ) by repetition. In particular, we show in Section 5) that this can be achieved by parallel repetition. The main theorem of this paper immediately follows:

Theorem 4.3. *If there exists an approximately-regular one-way function family then there exists a statistically-hiding commitment scheme.*

We remark that by changing the choice of $t(k)$ in Theorem 4.2 we could get a function that is more balanced. More specifically, for any constant $d > 0$ we can get a function that is one-way over its range and $(1/k^d, 1/k^d)$-balanced by replacing the constant 8 in the choice of $t(k)$ with the constant $8d$. However, once d becomes super constant, this choice of $t(k)$ does not seem to give a function that is one-way over its range. Therefore to avoid adding additional parameters, we restrict our attention to $d = 1$.

In the remainder of this section We prove Theorem 4.2. In Section 4.1 we show that $\mathcal{F}'$ is $(1/k, 1/k)$-balanced, and in Section 4.2 we prove that it is one-way over its range.

To avoid visual clutter we will omit the dependence on k in the notation and write f, n, ℓ, r, p instead of $f_k, n(k), \ell(k), r(k), p(k)$. We first observe that the almost regularity of f implies that $f(U_n)$ has high min-entropy. Specifically:

$$H_\infty(f(U_n)) \geq n - r - \log p. \tag{2}$$

Note that Construction 2.6 applies the leftover hash lemma to the distribution $Y = f(U_n)$ and therefore (using the fact that $t \leq (n - r - \log p) - 8\log k$) we have that the conditions of Lemma 2.5 are satisfied. We can thus conclude that:

1. $f'(U_{n+s})$ has statistical distance at most $1/k^3$ from U_{t+s}.
2. $H_2(f'(U_{n+s})) \geq t + s - 1/k \geq t + s - 1$.

We will use the first item to show that f' is balanced and the second item to show that f' is one-way over its range.

4.1 Showing that $\mathcal{F}'$ is Balanced

We begin by showing that $\mathcal{F}'$ is $(1/k, 1/k)$-balanced. We start by observing that if a distribution is close to uniform then it is somewhat balanced.

Lemma 4.4. *Let D be a distribution over strings of length v such that $SD(D, U_v) \leq \varepsilon$. Then for every $0 < \delta < 1$, distribution D is $(\frac{3\varepsilon}{\delta}, \delta)$-balanced.*

Proof. We define the following two sets:

$$\mathsf{Bad}^+ = \left\{ y_0 \in \{0,1\}^v : \Pr_{y \leftarrow D}[y = y_0] \geq (1+\delta)2^{-v} \right\}$$

$$\mathsf{Bad}^- = \left\{ y_0 \in \{0,1\}^v : \Pr_{y \leftarrow D}[y = y_0] \leq (1-\delta)2^{-v} \right\}.$$

Let $\mathsf{Bad} = \mathsf{Bad}^+ \cup \mathsf{Bad}^-$. Note that by construction we have that Bad satisfies the third item in Definition 3.1. It remains to show that Bad is small and has low probability with respect to D. We will do this for each of Bad^+ and Bad^- separately. As the argument is similar in both cases we will only do it for Bad^+.

By the definition of Bad^+ we have that:

$$\Pr_{y \leftarrow D}[y \in \mathsf{Bad}^+] - \Pr_{y \leftarrow U_v}[y \in \mathsf{Bad}^+] \geq |\mathsf{Bad}^+| \cdot \delta 2^{-v}.$$

As $SD(D, U_v) \leq \varepsilon$ we have that $|\mathsf{Bad}^+| \cdot \delta 2^{-v} \leq \varepsilon$ which implies that $|\mathsf{Bad}^+| \leq \left(\frac{\varepsilon}{\delta}\right) \cdot 2^v$. The same bound can be derived on Bad^- which gives that $|\mathsf{Bad}| \leq \left(\frac{2\varepsilon}{\delta}\right) \cdot 2^v$. This concludes the proof of the first item in Definition 3.1. For the second item we recall once more that $SD(D, U_v) \leq \varepsilon$ and thus:

$$\Pr_{y \leftarrow D}[y \in \mathsf{Bad}^+] \leq \Pr_{y \leftarrow U_v}[y \in \mathsf{Bad}^+] + \varepsilon \leq 2\varepsilon/\delta + \varepsilon \leq 3\varepsilon/\delta\,,$$

where the first inequality follows from the estimate on the size of Bad and the second follows because $\delta < 1$. This concludes the proof of the second item. $\square$

We have that $f'(U_{n+s})$ has statistical distance $1/k^3$ from U_{s+t}. Using Lemma 4.4 and setting $\delta = 1/k$ we conclude that $f'(U_{n+s})$ is (α, δ)-balanced for $\alpha = 3k/k^3 \leq 1/k$ as required.

Using hash functions with higher independence: While we can show that f' is (α, δ)-balanced for $\alpha = \delta = 1/k$ we are not able to achieve α and δ that are negligible in k (which would allow us to construct a statistically-hiding commitment scheme directly, without repetition). Had we chosen $t = n - r - \log p - \omega(\log k)$ we could have achieved this goal. However, for this choice of t we cannot show that f' is one-way over its range. We can show that if the family of hash functions is $3k$-universal then f' is $(\alpha = 2^{-k}, \delta = k^{-\Omega(1)})$-balanced. However, we do not know how to also achieve negligible δ in this case.

4.2 Showing that $\mathcal{F}'$ is One-Way Over Its Range

We now show that $\mathcal{F}'$ is one-way over its range (assuming $\mathcal{F}$ is one-way in the standard sense). Before giving the proof let us explain why this does not follow directly from the fact that $f'(U_{n+s})$ is statistically close to the uniform distribution. The reader may safely skip the following paragraph and go directly to the formal proof if he wishes.

Observe that if the output length t is "too small" compared to the min-entropy of $Y = f(U_n)$, then f' may not be one-way over its range. More precisely, as the min-entropy of Y is $n - r - \log p$ we have to choose $t \geq (n - r - \log p) - c \log k$ for some constant c. On the other hand, once c is fixed the distribution $f'(U_{n+s})$ may only be of distance $\approx k^{-c/2}$ from uniform. (Note that this distance is not negligible.) Therefore we cannot handle an adversary that inverts f' with probability, say,

k^{-2c}. To overcome this difficulty we use the stronger property that $f'(U_{n+s})$ has high 2-entropy. As we see below this characterization will in some sense allow us to handle events of probability smaller than k^{-c}.

We now proceed with the proof. We start by proving a technical lemma that shows that a distribution with high 2-entropy cannot have too many elements that have very high weight.

Lemma 4.5. *Let D be a distribution such that $H_2(D) \geq q$. Then for every $\varepsilon > 0$ there exists a set* Bad *such that:*

1. $\Pr_{y \leftarrow D}[y \in \mathsf{Bad}] \leq \varepsilon$.
2. *For any $y_0 \notin \mathsf{Bad}$,* $\Pr_{y \leftarrow D}[y = y_0] \leq 2^{-(q - \log(1/\varepsilon))}$.

Proof. Consider the random variable $R(y) = \Pr_D[y]$ (that is the random variable given by the probability distribution itself). Note that

$$CP(D) = \sum_{y \in \mathsf{Supp}(D)} \Pr_D[y]^2 = \sum_{y \in \mathsf{Supp}(D)} \Pr_D[y] \cdot R(y) = \mathsf{E}_D[R].$$

By Markov's inequality we have that $\Pr_D[R \geq \mathsf{E}_D(R)/\varepsilon] \leq \varepsilon$. Let $\mathsf{Bad} \stackrel{\text{def}}{=} \{y : R(y) \geq \mathsf{E}_D(R)/\varepsilon\}$. We have that $\Pr_{y \leftarrow D}[y \in \mathsf{Bad}] \leq \varepsilon$. For $y_0 \notin \mathsf{Bad}$ we have that

$$\Pr_{y \leftarrow D}[y = y_0] \leq CP(D)/\varepsilon = 2^{-(q - \log(1/\varepsilon))},$$

completing the proof. □

We now show that f' is one-way (in the standard sense). We will later use this to argue that f' is one-way over its range.

Lemma 4.6. *Function $\mathcal{F}'$ is one-way.*

Before giving the formal proof let us give a high level overview. The reader can safely skip the next paragraph and go directly to the formal proof if he wishes.

Suppose that an adversary A inverts f' (on the distribution $f'(U_{n+s})$) with probability k^{-d}. We now try to invert f (when given y from the distribution $f(U_n)$) by choosing h uniformly at random, computing $h \| h(y)$, and then running A. With noticeable probability, the adversary A will return a pair (h', x') with $h' = h$ and $h(f(x')) = h(y)$. However, this does not necessarily mean that $f(x') = y$, as it could be the case that $f(x') = y'$ such that $y' \neq y$ yet $h(y') = h(y)$. In this case we will not succeed on y but we will succeed when given y'. For simplicity let us assume that f is regular. Thus all y in the image of f have the same weight under $f(U_n)$. In this case, for the strategy above to succeed with noticeable probability we require that, for "most" y, only a polynomial number of y' satisfy $h(y) = h(y')$. This is where we utilize Lemma 4.5, which says that most outputs of f' do not have too many pre-images under h. The formal argument follows.

Proof (of Lemma 4.6). Assume toward a contradiction that there exists a PPT A' that achieves noticeable advantage in inverting $\mathcal{F}'$. That is, there exists a constant d such that for infinitely many k:

$$\Pr[h \leftarrow H_k; x \leftarrow \{0,1\}^{n(k)}; z = h(f_k(x)); x' \leftarrow A'(1^k, h, z) : h(f_k(x')) = z] \geq k^{-d}. \qquad (3)$$

(The above implicitly assumes that $A'(1^k, h, z)$ would never output (h', x') with $h' \neq h$; this is without loss of generality since A' can always be modified accordingly without decreasing its advantage.) Fix an arbitrary (large enough) k for which A' achieves such an advantage. To avoid visual clutter, we write f and H in place of f_k, H_k from now on. Construct a PPT adversary A (attempting to invert $\mathcal{F}$) as follows:

$\underline{A(1^k, y)}$
Choose $h \in H$ at random, and set $z = h(y)$
Run $A'(1^k, h, z)$ and obtain output x'
Output x'

We will show that A inverts f_k with noticeable probability, contradicting the one-wayness of $\mathcal{F}$. Note that the distribution over the inputs of A' in the above experiment is identical to the distribution over the inputs of A' in Equation (3). To simplify the notation in the remainder of this proof all probabilities refer to the probability space where x is chosen uniformly from $\{0,1\}^n$, the hash function h is chosen uniformly from H, and uniform random coins are chosen for A'. By a standard averaging argument, there is a set G_1 of pairs (h_0, z_0) such that:

1. $\Pr[f'(h,x) \in G_1] \geq 1/2k^d$.
2. For any $(h_0, z_0) \in G_1$, $\Pr[h_0(f(A'(1^k, h_0, z_0))) = z_0] \geq 1/2k^d$. Note that here the probability is only over the coin tosses of A' (as h_0 and z_0 are constants).

Recall that we have that $H_2(f'(U_{n+s})) \geq s + t - 1$. Set $\varepsilon = 1/4k^d$ and apply Lemma 4.5. We conclude that there is a set Bad such that:

1. $\Pr[f'(h,x) \in \mathsf{Bad}] \leq 1/4k^d$.
2. For every $(h_0, z_0) \notin \mathsf{Bad}$, $\Pr[f'(h,x) = (h_0, z_0)] \leq 2^{-(s+t-d\log k - O(1))}$.

By the fact that f is (r,p)-almost regular we have that for every $y_0 \in \mathsf{image}(f)$, $\Pr[f(x) = y_0] \geq 2^{-(n-r+\log p)}$. We now argue that, since each such y_0 has "large" weight, then for $(h_0, z_0) \notin \mathsf{Bad}$ there cannot be too many y_0's such that $h_0(y_0) = z_0$. More precisely, for every $(h_0, z_0) \notin \mathsf{Bad}$ and every $y_0 \in \mathsf{image}(f)$ such that $h_0(y_0) = z_0$,

$$
\begin{aligned}
Weight_{h_0,y_0}(y) &\stackrel{\text{def}}{=} \Pr[f(x) = y_0 | h(f(x)) = z_0 \wedge h = h_0] \\
&= \frac{\Pr[f(x) = y_0 \wedge h(f(x)) = z_0 \wedge h = h_0]}{\Pr[h(f(x)) = z_0 \wedge h = h_0]} \\
&= \frac{\Pr[f(x) = y_0 \wedge h = h_0]}{\Pr[h(f(x)) = z_0 \wedge h = h_0]} \\
&= \frac{\Pr[f(x) = y_0] \cdot \Pr[h = h_0]}{\Pr[h(f(x)) = z_0 \wedge h = h_0]} \\
&\geq \frac{2^{-(n-r+\log p)} \cdot 2^{-s}}{2^{-(s+t-d\log k - O(1))}} \\
&\geq 2^{-(2\log p + 8\log k + d\log k + O(1))} \\
&\geq k^{-O(1)}.
\end{aligned}
$$

We now show that A has noticeable probability of inverting f. Set $G = G_1 \setminus \mathsf{Bad}$. Note that:

$$\Pr[f'(h,x) \in G] \geq 1/2k^d - 1/4k^d = 1/4k^d.$$

For every $(h_0, z_0) \in G$, we have that $(h_0, z_0) \in G_1$ and therefore:

$$\Pr[h_0(f(A'(1^k, h_0, z_0))) = z_0 | h = h_0 \wedge h(f(x)) = z_0] = \Pr[h_0(f(A'(1^k, h_0, z_0))) = z_0] \geq 1/2k^d.$$

In words, for every $(h_0, z_0) \in G$ a $1/2k^d$ fraction of coin tosses of A' results in A' outputting a string x' such that $h_0(f(x')) = z_0$. For each such set of coins $coins_0$, the string x' is fixed (as a function of $h_0, z_0, coins_0$) and we define $y_0 \stackrel{\text{def}}{=} f(x')$. As $(h_0, z_0) \notin \mathsf{Bad}$ it follows that:

$$\Pr[f(x) = y_0 | h(f(x)) = z_0 \wedge h = h_0] = Weight_{h_0, y_0}(y) \geq k^{-O(1)}.$$

Thus, with noticeable probability all the following events happen simultaneously:

1. $f'(h,x) \in G$.
2. A' select a coin toss on which its output x' satisfies $h(f(x')) = h(f(x))$.
3. $f(x) = f(x')$.

When all these events happen then $A(1^k, f(x))$ indeed outputs a string x' such that $f(x') = f(x)$; i.e., A inverts f. Thus, we obtain a contradiction. □

We now complete the proof and show that f' is one-way over its range.

Lemma 4.7. *Function $\mathcal{F}'$ is one-way over its range.*

Proof. For every output (h_0, z_0) of f' there must be a $y_0 \in \mathsf{image}(f)$ such that $h_0(y_0) = z_0$. We have that $\mathcal{F}$ is $(r(k), p(k))$-approximately-regular and therefore y_0 has at least $2^{r(k)}/p(k)$ preimages under f. It follows that (h_0, z_0) have at least that many preimages under f'. Furthermore, for a different pair $(h_1, z_1) \neq (h_0, z_0)$ the set of preimages of (h_0, z_0) under f' does not intersect the preimages of (h_0, z_0). It follows that for every set $S \subseteq \mathsf{image}(f')$, $|f'^{-1}(S)| \geq \frac{2^{r(k)} \cdot |S|}{p(k)}$. Therefore,

$$\begin{aligned}
\Pr[h \leftarrow H; x \leftarrow \{0,1\}^{n(k)} : f'(h,x) \in S] &\geq \frac{\frac{2^{r(k)} \cdot |S|}{p(k)}}{2^{n(k)+s(k)}} \\
&\geq \frac{1}{2^{2 \log p(k) + 8 \log k}} \cdot \frac{|S|}{2^{t(k)+s(k)}} \\
&= \frac{1}{p(k)^2 \cdot k^8} \cdot \Pr[h \leftarrow H; z \leftarrow \{0,1\}^{t(k)} : (h,z) \in S],
\end{aligned}$$

where in the last inequality we used the fact that $t(k) = n(k) - r(k) - \log p(k) - 8 \log k$. Consider any PPT algorithm A'' inverting $\mathcal{F}'$ "over its range" with non-negligible probability $\varepsilon(k)$. By a straightforward averaging argument, there exists a set $S \subseteq \mathsf{image}(f')$ of relative size at least $\varepsilon(k)/2$ and such that, for any $y \in S$, $\Pr[A''(y) \in f'^{-1}(y)] \geq \varepsilon(k)/2$. Thus, by Eq.(4), A'' inverts f' (in the usual sense) with probability at least $\frac{\varepsilon(k)^2}{\mathsf{poly}(k)}$, contradicting the one-wayness of $\mathcal{F}'$. This completes the proof that $\mathcal{F}'$ is one-way over its range. □

5 Parallel Repetition of Commitments

In this section, we prove a parallel repetition theorem for the case of statistically-hiding commitment. We first define formally the notion of parallel repetition we consider:

Construction 5.1 (Parallel Repetition). *Let $(\mathcal{S}, \mathcal{R}, \mathcal{V})$ be a commitment scheme and $q = \mathsf{poly}(k)$. Construct commitment scheme $(\mathcal{S}_q, \mathcal{R}_q, \mathcal{V}_q)$ as follows.*

- *On input a bit b, $\mathcal{S}_q$ chooses q bits $b_1, \ldots, b_q$ uniformly at random subject to $b = \bigoplus_{i=1}^q b_i$. It then runs (in parallel) q instances of $\mathcal{S}$, where the i^{th} instance commits to b_i. The output of $\mathcal{S}_q$ is $\mathsf{decom} = (\mathsf{decom}_1, \ldots, \mathsf{decom}_q)$, where decom_i is the output of the i^{th} instance of $\mathcal{S}$.*
- *$\mathcal{R}_q$ runs (in parallel) q instances of $\mathcal{R}$. The output of $\mathcal{R}_q$ is $\mathsf{com} = (\mathsf{com}_1, \ldots, \mathsf{com}_q)$, where com_i is the output of the i^{th} instance of $\mathcal{R}$.*
- *$\mathcal{V}_q$, on input $\mathsf{com} = (\mathsf{com}_1, \ldots, \mathsf{com}_q)$ and $\mathsf{decom} = (\mathsf{decom}_1, \ldots, \mathsf{decom}_q)$, computes $b_i = \mathcal{V}(\mathsf{com}_i, \mathsf{decom}_i)$ for all i. If $b_i = \perp$ for any i, $\mathcal{V}_q$ outputs $\perp$; otherwise, it outputs $\bigoplus_{i=1}^q b_i$.*

We are now ready to state the result.

Theorem 5.2. *Let $(\mathcal{S}, \mathcal{R}, \mathcal{V})$ and $(\mathcal{S}_q, \mathcal{R}_q, \mathcal{V}_q)$ be as in Construction 5.1. If $(\mathcal{S}, \mathcal{R}, \mathcal{V})$ is a ρ-secure commitment scheme, then $(\mathcal{S}_q, \mathcal{R}_q, \mathcal{V}_q)$ is a ρ^q-secure commitment scheme.*

We stress that the initial scheme $(\mathcal{S}, \mathcal{R}, \mathcal{V})$ is computationally binding in the sense of Definition 2.2; that is, a cheating polynomial-time sender has only *negligible* probability of correctly opening a commitment to two different messages. In other words, we use parallel repetition here only to strengthen the *hiding* property (and not to strengthen the binding property). Nevertheless, a straightforward hybrid argument (omitted here) shows that binding is not affected: i.e., if $(\mathcal{S}, \mathcal{R}, \mathcal{V})$ is computationally binding then so is $(\mathcal{S}_q, \mathcal{R}_q, \mathcal{V}_q)$. The interesting part of the theorem is that if $(\mathcal{S}, \mathcal{R}, \mathcal{V})$ is ρ-hiding then $(\mathcal{S}_q, \mathcal{R}_q, \mathcal{V}_q)$ is ρ^q-hiding. Although seemingly obvious, it is not easy to prove: the difficulty is that the views of the receiver in the different instances of the basic scheme $(\mathcal{S}, \mathcal{R}, \mathcal{V})$ are not necessarily independent, since a malicious receiver can correlate its messages in each of these executions.[4]

In our proof, we rely on the ideas used to prove an analogous parallel repetition theorem for reducing the *soundness error* in interactive proof systems [13, Appendix C]. Our result does not appear to follow directly from that result; rather, we use similar techniques. (Note that in both settings the adversary — i.e., the verifier in the case of interactive proofs, and the receiver here — is all-powerful.) Theorem 5.2 follows immediately from the following lemma.

Lemma 5.3. *Let $\mathsf{Com}_1 = (\mathcal{S}_1, \mathcal{R}_1, \mathcal{V}_1)$ and $\mathsf{Com}_2 = (\mathcal{S}_2, \mathcal{R}_2, \mathcal{V}_2)$ be two commitment schemes, and construct $\mathsf{Com} = (\mathcal{S}, \mathcal{R}, \mathcal{V})$ by parallel composition of these schemes (in the obvious way, as in Construction 5.1). If Com_1 is ρ_1-hiding and Com_2 is ρ_2-hiding, then Com is $(\rho_1\rho_2)$-hiding.*

Proof. Assume each of the component commitment schemes has an r-round commitment phase (this can always be ensured by sending "dummy messages" as needed), and assume without loss of generality that $\mathcal{S}_1$ (resp., $\mathcal{S}_2$) sends the first and last message. Inspired by the proof of [13, Lemma C.1], we employ the notion of a *game tree* T, defined for any commitment scheme as follows:

[4]This difficulty goes away if either (1) we use *sequential* repetition; or (2) the receiver is *honest-but-curious*. In either of these cases, a parallel composition theorem is easy to prove.

- The root of T is a node at level 0 denoted ε. This corresponds to the beginning of an execution of the scheme.
- Each node v at an *even* level ℓ corresponds to a point when the *honest* sender makes a move. This node has children at level $\ell + 1$ corresponding to all possible (legal) messages of the sender (i.e., for all possible random coins and either possible input bit).
- Each node v at an *odd* level $\ell < r$ corresponds to a point when the malicious receiver makes a move. This node has children at level $\ell + 1$ corresponding to all possible messages of the receiver.

We identify a node v with the partial view of the receiver up to that point. As a special case, if v is a leaf then v corresponds to a possible view of the receiver when the commitment scheme is run to completion. For a node v, we let $\mathsf{ch}(v)$ denote the set of children of v.

Let T_1, T_2, T denote the game trees for $\mathsf{Com}_1, \mathsf{Com}_2, \mathsf{Com}$, respectively. Let $\mathcal{C}_1$ denote the size of the space of random coins for $\mathcal{S}_1$ (i.e., if $\mathcal{S}_1$ uses s_1 random coins, then $\mathcal{C}_1 = 2^{s_1}$), and define $\mathcal{C}_2, \mathcal{C}$ analogously. For a partial transcript v_1 of an execution of Com_1, let $C_1(v_1; b)$ denote the size of the set of random coins for $\mathcal{S}_1$ consistent with input bit b and v_1; $C_2(v_2; b)$ and $C(v; b)$ are defined analogously.

We now define a *value* val for each node in a game tree. We focus on game tree T corresponding to Com, but the value $\mathsf{val}_1, \mathsf{val}_2$ of a node in T_1, T_2 is defined analogously. The value of a node in T is defined inductively:

- If v is a leaf, then $\mathsf{val}(v) \stackrel{\text{def}}{=} \frac{|C(v;0) - C(v;1)|}{2\mathcal{C}}$.
- If v is an even node, then $\mathsf{val}(v) \stackrel{\text{def}}{=} \sum_{w \in \mathsf{ch}(v)} \mathsf{val}(w)$.
- If v is an odd node, then $\mathsf{val}(v) \stackrel{\text{def}}{=} \max_{w \in \mathsf{ch}(v)} \mathsf{val}(w)$.

If v is a partial transcript and $\widehat{\mathsf{view}}$ is a full transcript, we say that $\widehat{\mathsf{view}}$ *is consistent with* v if v is a prefix of $\widehat{\mathsf{view}}$. It is not hard to see that for any (unbounded) receiver $\mathcal{R}^*$ and every node v:

$$\mathsf{val}(v) \geq \frac{1}{2} \cdot \sum_{\widehat{\mathsf{view}} \text{ consistent with } v} \left| \Pr[\mathsf{view}_{\langle \mathcal{S}(0), \mathcal{R}^* \rangle} = \widehat{\mathsf{view}}] - \Pr[\mathsf{view}_{\langle \mathcal{S}(1), \mathcal{R}^* \rangle} = \widehat{\mathsf{view}}] \right|.$$

Furthermore, there exists an unbounded receiver $\mathcal{R}^*$ for which equality holds. By the assumption of the lemma, then, we have $\mathsf{val}_1(\varepsilon) \leq \rho_1$ and $\mathsf{val}_2(\varepsilon) \leq \rho_2$. Moreover, we can prove the theorem by showing that $\mathsf{val}(\varepsilon) \leq \rho_1 \rho_2$.

Note that every node (i.e., partial view) v in T corresponds in the natural way to a tuple $(v_1, v_2) \in T_1 \times T_2$ (in particular, ε corresponds to $(\varepsilon, \varepsilon)$). From now on, we write $\mathsf{val}(v_1, v_2)$ in place of $\mathsf{val}(v)$. The desired bound on $\mathsf{val}(\varepsilon, \varepsilon)$ is immediate from the following, more general claim

Claim 5.4. *For all v_1, v_2:*

$$\mathsf{val}(v_1, v_2) = \mathsf{val}_1(v_1) \cdot \mathsf{val}_2(v_2).$$

Proof. We prove this by induction on the level of the tree, starting from the bottom. The base case occurs when $v = (v_1, v_2)$ is a leaf in T. By construction of Com, we have $\mathcal{C} = 2 \cdot \mathcal{C}_1 \cdot \mathcal{C}_2$ (the sender $\mathcal{S}$ uses an additional random bit to determine the inputs to $\mathcal{S}_1, \mathcal{S}_2$) and furthermore

$$C(v; 0) = C_1(v_1; 0) \cdot C_2(v_2; 0) + C_1(v_1; 1) \cdot C_2(v_2; 1)$$

and

$$C(v;1) = C_1(v_1;0)\cdot C_2(v_2;1) + C_1(v_1;1)\cdot C_2(v_2;0).$$

So:

$$\begin{aligned}
\frac{|C(v;0)-C(v;1)|}{2\mathcal{C}} &= \frac{|C_1(v_1;0)-C_1(v_1;1)|\cdot|C_2(v_2;0)-C_2(v_2;1)|}{2\cdot 2\cdot \mathcal{C}_1,\cdot\mathcal{C}_2}\\
&= \frac{|C_1(v_1;0)-C_1(v_1;1)|}{2\mathcal{C}_1}\cdot\frac{|C_2(v_2;0)-C_2(v_2;1)|}{2\mathcal{C}_2}\\
&= \mathsf{val}_1(v_1)\cdot\mathsf{val}_2(v_2).
\end{aligned}$$

Now, assume the claim is true for all nodes at level $\ell+1$ and consider a node $v=(v_1,v_2)$ at level ℓ. There are two case. If ℓ is odd, then

$$\begin{aligned}
\mathsf{val}(v) &= \max_{w\in\mathsf{ch}(v)}\mathsf{val}(w)\\
&= \max_{w_1\in\mathsf{ch}(v_1),w_2\in\mathsf{ch}(v_2)}\mathsf{val}(w_1,w_2)\\
&= \max_{w_1\in\mathsf{ch}(v_1),w_2\in\mathsf{ch}(v_2)}\mathsf{val}_1(w_1)\cdot\mathsf{val}_2(w_2)\\
&= \left(\max_{w_1\in\mathsf{ch}(v_1)}\mathsf{val}_1(w_1)\right)\cdot\left(\max_{w_2\in\mathsf{ch}(v_2)}\mathsf{val}_2(w_2)\right)\\
&= \mathsf{val}_1(v_1)\cdot\mathsf{val}_2(v_2).
\end{aligned}$$

If, on the other hand, ℓ is even:

$$\begin{aligned}
\mathsf{val}(v) &= \sum_{w\in\mathsf{ch}(v)}\mathsf{val}(w)\\
&= \sum_{w_1\in\mathsf{ch}(v_1),w_2\in\mathsf{ch}(v_2)}\mathsf{val}_1(w_1)\cdot\mathsf{val}_2(w_2)\\
&= \left(\sum_{w_1\in\mathsf{ch}(v_1)}\mathsf{val}_1(w_1)\right)\cdot\left(\sum_{w_2\in\mathsf{ch}(v_2)}\mathsf{val}_2(w_2)\right)\\
&= \mathsf{val}_1(v_1)\cdot\mathsf{val}_2(v_2).
\end{aligned}$$

This completes the proof of the claim, and hence the lemma. □

□

6 Honest-but-Curious vs. Malicious Receivers

In this section, we demonstrate a *compiler* that converts any commitment scheme that is statistically-hiding against an *honest-but-curious* (i.e., semi-honest) receiver into a commitment scheme that is similarly hiding even against a *malicious* receiver. More formally, given a scheme that is ρ-hiding against an honest but-curious receiver (see Definition 6.2) we obtain a scheme that is $(\rho+\frac{1}{e(k)}+\mathsf{negl})$-hiding against a malicious receiver using $O(\log e(k)/\log k)$ additional rounds.[5] Moreover, if the initial scheme is computationally binding then so is the derived scheme. The compiler requires only the existence of one-way functions, and we therefore obtain the following corollary:

[5] In the preliminary version of this work [25], this was only claimed for the particular case of $e(k)=k^{\omega(1)}$.

Corollary 6.1. *The existence of an $r(k)$-round commitment scheme that is computationally binding and $(1 - \frac{1}{\mathsf{poly}})$-hiding against an honest-but-curious receiver implies the existence of a statistically-hiding commitment scheme that requires $r(k) + O(1)$ rounds.*

Proof. Say the initial commitment scheme is $(1 - \frac{1}{p(k)})$-hiding against an honest-but-curious receiver for some polynomial p. The existence of such a scheme implies the existence of a one-way function, and we can therefore apply our compiler to obtain a scheme that is $(1 - \frac{1}{2p(k)} + \mathsf{negl})$-hiding against a *malicious* receiver and that requires $r(k) + \log p(k)/\log k = r(k) + O(1)$ rounds. Using parallel repetition and applying Theorem 5.2 gives the states result. $\square$

We remark that given a *constant-round* zero-knowledge (ZK) proof system (with negligible soundness error) for $\mathcal{NP}$, it would be possible to obtain a compiler that transforms a scheme that is ρ-hiding against an honest but-curious receiver to a scheme that is $(\rho + \mathsf{negl})$-hiding against a malicious receiver and uses only $O(1)$ additional rounds. Currently, however, the best known constructions of ZK proof systems with negligible soundness error based on one-way functions require $\omega(1)$ rounds [23]. In particular, the constant-round construction of [19] requires a statistically-hiding commitment scheme — the very primitive we are trying to construct!

For completeness, we provide a definition of security against an honest-but-curious receiver; such a definition follows easily from Definitions 2.1 and 2.3.

Definition 6.2. Commitment scheme $(\mathcal{S}, \mathcal{R}, \mathcal{V})$ is *ρ-hiding against an honest-but-curious receiver* if the following holds: Let $\mathsf{view}_{\langle \mathcal{S}(b), \mathcal{R}\rangle}(k)$ be the view of an *honest* receiver $\mathcal{R}$. Then we require that the ensembles $\{\mathsf{view}_{\langle \mathcal{S}(0), \mathcal{R}\rangle}(k)\}$ and $\{\mathsf{view}_{\langle \mathcal{S}(1), \mathcal{R}\rangle}(k)\}$ have statistical difference at most ρ.

Since we consider the view of the honest receiver $\mathcal{R}$, we must allow $\mathcal{R}$ to be probabilistic. The view of $\mathcal{R}$ consists of its random coins as well as the messages sent by $\mathcal{S}$.

Our compiler uses a coin-tossing protocol and zero-knowledge proofs (in a way similar to [22]) to "force" honest behavior on the part of the receiver. However, we do not require "simulatable" coin-tossing (as in [4, 22, 32]) or ZK proofs of correctness following each round (as in [22]); instead, we show that a weaker variant of coin-tossing along with a single ZK proof at the end suffice. (The latter in particular is essential for obtaining a round-efficient compiler.)

Informally, our compiler proceeds as follows: the receiver first uses a statistically-*binding* commitment scheme (e.g., [33]) to commit to a sufficiently-long string r_1, and the sender responds with a string r_2 of the same length. The sender and receiver then execute the original protocol, with the receiver using $r_1 \oplus r_2$ as its random tape and the sender committing to a *random* bit b'. At the conclusion of the original protocol, the receiver uses a ZK proof (with soundness error $1/e(k)$) to show that each of the messages it sent during the course of the protocol is consistent with the messages sent by $\mathcal{S}$ as well as the random tape $r_1 \oplus r_2$ (we stress that r_1 is never revealed to $\mathcal{S}$). Finally (assuming $\mathcal{S}$ accepts the proof), $\mathcal{S}$ concludes the protocol by sending $b' \oplus b$ (where b is the bit that $\mathcal{S}$ wants to commit to).

Before giving a formal description and proof of security, some brief remarks are in order. First, one-way functions are sufficient for both statistically-binding commitment [33] as well as zero-knowledge proofs for all of $\mathcal{NP}$ [23, 33] with round complexity $O(\lceil \log e(k)/\log k \rceil)$. Second, since we have the receiver provide a ZK proof of correctness only at the conclusion of the protocol we must take into account the fact that the receiver may cheat during the course of the protocol, learn some information about the bit committed to by $\mathcal{S}$, and then abort. To prevent this, we have $\mathcal{S}$ commit to a *random* bit b' in the initial phase of the protocol (i.e., before the ZK proof); the only portion of the transcript that depends on the input bit of $\mathcal{S}$ is sent *after* the receiver successfully proves correctness of its actions. We now provide a detailed description of our compiler.

Construction 6.3. *Let* $(\mathcal{S}, \mathcal{R}, \mathcal{V})$ *be a commitment scheme, and* $e(k)$ *a function. Construct commitment scheme* $(\widehat{\mathcal{S}}, \widehat{\mathcal{R}}, \widehat{\mathcal{V}})$*, where the sender* $\widehat{\mathcal{S}}$ *with input bit* b *interacts with the receiver* $\widehat{\mathcal{R}}$ *as follows:*

1. *Let* $\ell = \ell(k)$ *denote the length of the random tape used by* $\mathcal{R}$*. Then* $\widehat{\mathcal{R}}$ *uses a (possibly interactive) statistically-binding commitment scheme to commit to a random string* $r_1 \in \{0,1\}^\ell$*. Let* com *denote the resulting commitment (known to both* $\widehat{\mathcal{S}}$ *and* $\widehat{\mathcal{R}}$*) and let* decom *be the decommitment (known to* $\widehat{\mathcal{R}}$*).*

2. $\widehat{\mathcal{S}}$ *sends a random string* $r_2 \in \{0,1\}^\ell$*. This defines a string* $r \stackrel{\text{def}}{=} r_1 \oplus r_2$ *which is known to* $\widehat{\mathcal{R}}$ *(but not to* $\widehat{\mathcal{S}}$*).*

3. $\widehat{\mathcal{S}}$ *chooses a random bit* b'*, and then* $\widehat{\mathcal{S}}$ *and* $\widehat{\mathcal{R}}$ *run protocols* $\mathcal{S}(b')$ *and* $\mathcal{R}$*, respectively, where the latter is run using random tape* r*. At the conclusion of this stage,* $\widehat{\mathcal{S}}$ *obtains* decom' *as output from* $\mathcal{S}$*, while* $\widehat{\mathcal{R}}$ *obtains* com' *as output from* $\mathcal{R}$*.*

4. $\widehat{\mathcal{R}}$ *provides a ZK proof (with soundness error* $1/e(k)$*) that it acted correctly throughout the previous stage. Formally,* $\widehat{\mathcal{R}}$ *proves that there exists* (decom, r_1) *such that* com *is a commitment to* r_1 *(with decommitment* decom*) and all the messages sent by* $\widehat{\mathcal{R}}$ *in the previous stage are consistent with the messages sent by* $\widehat{\mathcal{S}}$ *and the random tape* $r = r_1 \oplus r_2$*.*

5. *If* $\widehat{\mathcal{S}}$ *rejects the proof given by* $\widehat{\mathcal{R}}$*, it aborts. Otherwise,* $\widehat{\mathcal{S}}$ *sends* $\hat{b} = b \oplus b'$ *and outputs* decom'*; the receiver* $\widehat{\mathcal{R}}$ *outputs* $(\mathsf{com}', \hat{b})$*.*

In the decommitment phase, $\widehat{\mathcal{V}}$ *proceeds as follows: it runs* $\mathcal{V}(\mathsf{com}', \mathsf{decom}')$ *to obtain a bit* b' *(if the output of* $\mathcal{V}$ *is* $\perp$*, then* $\widehat{\mathcal{V}}$ *outputs* $\perp$ *as well), and then outputs* $\hat{b} \oplus b'$*.*

Theorem 6.4. *If commitment scheme* $(\mathcal{S}, \mathcal{R}, \mathcal{V})$ *is computationally binding and* ρ*-hiding against an honest-but-curious receiver, then commitment scheme* $(\widehat{\mathcal{S}}, \widehat{\mathcal{R}}, \widehat{\mathcal{V}})$ *as generated by the above compiler is computationally binding and* $(\rho + \frac{1}{e(k)} + \mathsf{negl})$*-hiding (even against a malicious receiver).*

In proving the theorem, we consider the hiding and binding properties individually.

Claim 6.5. *If* $(\mathcal{S}, \mathcal{R}, \mathcal{V})$ *is* ρ*-hiding against an honest-but-curious receiver, then* $(\widehat{\mathcal{S}}, \widehat{\mathcal{R}}, \widehat{\mathcal{V}})$ *as in Construction 6.3 is* $(\rho + \frac{1}{e(k)} + \mathsf{negl})$*-hiding.*

Proof. Let $\widehat{\mathcal{R}}^*$ denote an unbounded malicious receiver who interacts with $\widehat{\mathcal{S}}$, and assume that $\widehat{\mathcal{R}}^*$ is deterministic without loss of generality. We say a transcript of an execution of $\widehat{\mathcal{S}}$ with $\widehat{\mathcal{R}}^*$ is *non-aborting* if $\widehat{\mathcal{S}}$ sends the final bit in the protocol; i.e., $\widehat{\mathcal{R}}^*$ gave a successful ZK proof that it acted correctly. (We say it is *aborting* otherwise.) We say a transcript is *good* if (1) the commitment com in the transcript indeed commits $\widehat{\mathcal{R}}^*$ to a single value r_1; and (2) $\widehat{\mathcal{R}}^*$ indeed acted correctly in its execution of the sub-routine $\mathcal{R}$; that is, each message sent by $\widehat{\mathcal{R}}^*$ in this transcript is consistent with $r_1 \oplus r_2$ (note that r_1 is uniquely defined by (1), and r_2 is explicit in the transcript). (We say that it is *bad* otherwise.) Note that the probability of a transcript that is bad and non-aborting is at most $1/e(k) + \mathsf{negl}$.

The statistical difference between distributions $\mathsf{view}_{\langle \widehat{\mathcal{S}}(0),\widehat{\mathcal{R}}^*\rangle}(k)$ and $\mathsf{view}_{\langle \widehat{\mathcal{S}}(1),\widehat{\mathcal{R}}^*\rangle}(k)$ is

$$\mathsf{SD}^*(k) \stackrel{\text{def}}{=} \frac{1}{2} \cdot \sum_{\mathsf{view}} \left|\Pr^*_{b=0}[\mathsf{view}] - \Pr^*_{b=1}[\mathsf{view}]\right|, \tag{4}$$

where $\Pr^*_{b=0}[\cdot]$ denotes the probability taken over coins of the sender when its input bit is 0, and the case of $b = 1$ is defined analogously. When view is aborting, $\Pr^*_{b=0}[\mathsf{view}] = \Pr^*_{b=1}[\mathsf{view}]$ (since only the final message depends on the input bit). On the other hand, as we have already noted, the probability of obtaining a bad and non-aborting view is at most $1/e(k) + \mathsf{negl}$. Defining

$$\overline{\mathsf{SD}}^* \stackrel{\text{def}}{=} \frac{1}{2} \cdot \sum_{\substack{\mathsf{view}\text{ non-aborting} \\ \text{and good}}} \left|\Pr^*_{b=0}[\mathsf{view}] - \Pr^*_{b=1}[\mathsf{view}]\right|,$$

we see that $|\mathsf{SD}^*(k) - \overline{\mathsf{SD}}^*| \leq 1/e(k) + \mathsf{negl}$.

Any non-aborting view can be parsed as an initial portion view' and the bit $\hat{b}$ sent in the final round. Note

$$\begin{aligned}
\overline{\mathsf{SD}}^*(k) &= \frac{1}{2} \cdot \sum_{\substack{\mathsf{view}'\text{ non-aborting} \\ \text{and good}}} \sum_{\hat{b}\in\{0,1\}} \left|\Pr^*_{b=0}[\mathsf{view}'\|\hat{b}] - \Pr^*_{b=1}[\mathsf{view}'\|\hat{b}]\right| \\
&= \frac{1}{2} \cdot \sum_{\substack{\mathsf{view}'\text{ non-aborting} \\ \text{and good}}} \left|\Pr^*_{b'=0}[\mathsf{view}'] - \Pr^*_{b'=1}[\mathsf{view}']\right|. \qquad (5)
\end{aligned}$$

(The notions of 'non-aborting' and 'good' depend only on the initial portion, and so are meaningfully defined above.) Above, $\Pr_{b'=0}[\mathsf{view}']$ denotes the probability of view' conditioned on the random bit b' of the sender being equal to 0 (with the case $b' = 1$ defined analogously). Note that $\Pr_{b'=0}[\mathsf{view}']$ is independent of the input bit b.

We show the existence of a randomized (but not polynomial-time) procedure ψ, outputting either a partial transcript or $\perp$, with the following property. Let $D(b)$ be the distribution defined by $\psi\left(\mathsf{view}_{\langle \mathcal{S}(b),\mathcal{R}\rangle}(k)\right)$ (i.e., ψ applied to a randomly-generated view of $\mathcal{R}$ interacting with $\mathcal{S}(b)$), and let $\Pr_{D(b)}[\mathsf{view}']$ be the probability of view' with respect to distribution $D(b)$. Then if view' is non-aborting and good

$$\Pr_{D(b)}[\mathsf{view}'] = \Pr_{b'=b}[\mathsf{view}'], \tag{6}$$

while for any other view' we have $\Pr_{D(b)}[\mathsf{view}'] = 0$. That is, any view output by ψ is good and non-aborting (and is furthermore output with probability as in Eq. (6)), and ψ outputs $\perp$ otherwise. The statistical difference between $D(0)$ and $D(1)$ is exactly given by Eq. (5); on the other hand, since the statistical difference between $\mathsf{view}_{\langle \mathcal{S}(b),\mathcal{R}\rangle}(k)$ and $\mathsf{view}_{\langle \mathcal{S}(1),\mathcal{R}\rangle}(k)$ is at most ρ (as is guaranteed by the ρ-hiding of $(\mathcal{S},\mathcal{R},\mathcal{V})$ against an honest-but-curious receiver), this statistical difference can be at most ρ. We conclude that $\overline{\mathsf{SD}}^*(k) \leq \rho(k)$, and so $\mathsf{SD}^*(k) \leq \rho(k) + 1/e(k) + \mathsf{negl}$.

It remains to show ψ. Procedure ψ, on input a tuple $(m_1, \ldots, m_i, r)$ (where $m_1, \ldots, m_i$ denote the messages of the sender $\mathcal{S}$ and r denotes the random coins used by honest-but-curious $\mathcal{R}$), proceeds as follows:

1. Begin interacting with $\widehat{\mathcal{R}}^*$, simulating an execution of $\widehat{\mathcal{S}}$.
2. When ψ obtains a commitment com from $\widehat{\mathcal{R}}^*$, it computes (in exponential time) a unique string r_1 consistent with this commitment. In case no such r_1 exists or multiple such r_1 exist, ψ outputs $\perp$. (In the first case, the ZK proof of $\widehat{\mathcal{R}}^*$ will fail except with probability at most $1/e(k)$; the second case occurs with negligible probability by statistical binding of the commitment scheme.)
3. ψ sends the string $r_2 = r \oplus r_1$ to $\widehat{\mathcal{R}}^*$.
4. ψ interacts with $\widehat{\mathcal{R}}^*$ by sending messages $m_1, \ldots, m_r$ in response to the messages of $\widehat{\mathcal{R}}^*$. If $\widehat{\mathcal{R}}^*$ ever sends a message inconsistent with random tape r and these messages, ψ outputs $\perp$.
5. Finally, ψ acts as a verifier in an execution of the ZK proof with $\widehat{\mathcal{R}}^*$. If the proof succeeds, then ψ outputs the entire view view' to this point. If the proof fails, ψ outputs $\perp$.

It is immediate that ψ never outputs a view' that is bad or aborting. It is also not hard to see that for any view' that is non-aborting and good, the probability that view' is output by $\psi\left(\mathsf{view}_{\langle \mathcal{S}(b),\mathcal{R}\rangle}(k)\right)$ is exactly $\Pr^*_{b'=b}[\mathsf{view}']$. The claim follows. ☐

We next consider the binding property.

Claim 6.6. *If* $\Pi = (\mathcal{S}, \mathcal{R}, \mathcal{V})$ *is computationally binding, then* $\Pi^* = (\widehat{\mathcal{S}}, \widehat{\mathcal{R}}, \widehat{\mathcal{V}})$ *as in Construction 6.3 is computationally binding as well.*

Proof. Given a PPT sender $\hat{\mathcal{S}}^*$ violating the binding property of Π^* with non-negligible probability, we construct a PPT sender $\hat{\mathcal{S}}$ violating the binding property of Π with non-negligible probability. $\hat{\mathcal{S}}$ is defined as follows:

1. $\hat{\mathcal{S}}$ interacts with an honest receiver $\mathcal{R}$, and runs a copy of $\hat{\mathcal{S}}^*$ internally. $\hat{\mathcal{S}}$ begins by sending a random commitment to the string $r_1 = 0^\ell$ to $\hat{\mathcal{S}}^*$, who responds with a string $r_2 \in \{0,1\}^\ell$.
2. $\hat{\mathcal{S}}$ then relays messages faithfully between $\mathcal{R}$ and $\hat{\mathcal{S}}^*$. At the conclusion of this phase, no more messages are sent to $\mathcal{R}$.
3. Finally, $\hat{\mathcal{S}}$ simulates a ZK proof of correct behavior with $\hat{\mathcal{S}}^*$ acting as the potentially-dishonest verifier. ($\hat{\mathcal{S}}^*$ then sends a final bit, which $\hat{\mathcal{S}}$ ignores.)
4. If $\hat{\mathcal{S}}^*$ outputs valid decommitments to two different bits, then $\hat{\mathcal{S}}$ does so as well.

To complete the proof, we argue that the probability that $\hat{\hat{\mathcal{S}}}$ outputs two valid decommitments in its interaction with $\hat{\mathcal{S}}$, above, is negligibly-close to the probability that $\hat{\hat{\mathcal{S}}}$ outputs two valid decommitments when interacting with an honest receiver $\widehat{\mathcal{R}}$. This is relatively straightforward to show via a hybrid argument, and we only sketch the proof. Consider a sequence of experiments, and let $\Pr_i[\mathsf{NoBind}]$ denote the probability that $\hat{\mathcal{S}}^*$ outputs two valid decommitments in experiment i):

Experiment 0. This is the original experiment, where $\hat{\mathcal{S}}^*$ interacts with $\widehat{\mathcal{R}}$.

Experiment 1. Here, we have $\widehat{\mathcal{R}}$ act exactly as in Experiment 0, except that it simulates the final ZK proof of correct behavior. By the ZK property of the proof system (against computationally-bounded verifiers), $|\Pr_0[\mathsf{NoBind}] - \Pr_1[\mathsf{NoBind}]|$ is negligible.

Experiment 2. Now, we have $\widehat{\mathcal{R}}$ act as in the previous experiment, except that its initial commitment is to 0^ℓ rather than to a random $r_1 \in \{0,1\}^\ell$. (However, it uses random tape $r_1 \oplus r_2$ in computing its messages to send.) Computational hiding of the commitment scheme implies that $|\Pr_2[\mathsf{NoBind}] - \Pr_1[\mathsf{NoBind}]|$ is negligible.

Experiment 2 corresponds exactly to an interaction of $\hat{\mathcal{S}}^*$ with $\hat{\mathcal{S}}$. $\square$

Acknowledgments

We are grateful to Yan Zong Ding, Virgil Gligor, Oded Goldreich, Danny Harnik, Omer Reingold, and Alon Rosen for helpful conversations and for reading preliminary versions of this manuscript. We thank the anonymous referees for helpful comments that improved the presentation.

References

[1] N. Alon, L. Babai, and A. Itai. A fast and simple randomized parallel algorithm for the maximal independent set problem. *Journal of Algorithms*, 7(4):567–583, 1986.

[2] M. Bellare, R. Impagliazzo, and M. Naor. Does parallel repetition lower the error in computationally sound protocols? In *38th FOCS*, pages 374–383, 1997.

[3] M. Bellare and S. Micali. How to sign given any trapdoor permutation. *J. ACM*, 39(1):214–233, 1992.

[4] M. Blum. Coin flipping by telephone. In *Advances in Cryptology — CRYPTO '81*, pages 11–15, 1981.

[5] M. Blum and S. Micali. How to generate cryptographically-strong sequences of pseudorandom bits. *SIAM J. Computing*, 13(4):850–864, 1984.

[6] J. Boyar, S. Kurtz, and M. Krentel. Discrete logarithm implementation of perfect zero-knowledge blobs. *Journal of Cryptology*, 2(2):63–76, 1990.

[7] G. Brassard, D. Chaum, and C. Crépeau. Minimum disclosure proofs of knowledge. *J. Computer and System Sciences*, 37(2):156–189, 1988.

[8] J. Carter and M. Wegman. Universal classes of hash functions. *J. Computer and System Sciences*, 18(2):143–154, 1979.

[9] B. Chor and O. Goldreich. On the power of two-point based sampling. *Journal of Complexity*, 5(1):96–106, 1989.

[10] I. Damgård. Collision free hash functions and public key signature schemes. In *Eurocrypt '87*, volume 304 of *LNCS*, pages 203–216. Springer, 1988.

[11] I. Damgård, M. Pedersen, and B. Pfitzmann. On the existence of statistically hiding bit commitment schemes and fail-stop signatures. *Journal of Cryptology*, 10(3):163–194, 1997.

[12] A. De Santis and M. Yung. On the design of provably-secure cryptographic hash functions. In *EUROCRYPT: Advances in Cryptology: Proceedings of EUROCRYPT*, pages 412–431, 1990.

[13] O. Goldreich. *Modern Cryptography, Probabilistic Proofs, and Pseudorandomness*. Springer-Verlag, 1999.

[14] O. Goldreich. *Foundations of Cryptography, Vol. 1: Basic Tools*. Cambridge University Press, 2001.

[15] O. Goldreich. *Foundations of Cryptography, Vol. 2: Basic Applications*. Cambridge University Press, 2004.

[16] O. Goldreich, S. Goldwasser, and S. Micali. On the cryptographic applications of random functions. In *Advances in Cryptology — CRYPTO '84*, pages 276–288, 1985.

[17] O. Goldreich, S. Goldwasser, and S. Micali. How to construct random functions. *J. ACM*, 33(4):792–807, 1986.

[18] O. Goldreich, R. Impagliazzo, L. Levin, R. Venkatesan, and D. Zuckerman. Security preserving amplification of hardness. In *31st FOCS*, pages 169–178, 1990.

[19] O. Goldreich and A. Kahan. How to construct constant-round zero-knowledge proof systems for $\mathcal{NP}$. *Journal of Cryptology*, 9(3):167–190, 1996.

[20] O. Goldreich, H. Krawczyk, and M. Luby. On the existence of pseudorandom generators. *SIAM J. Computing*, 22(6):1163–1175, 1993.

[21] O. Goldreich and L. Levin. Hard-core predicates for any one-way function. In *21st STOC*, pages 25–32, 1989.

[22] O. Goldreich, S. Micali, and A. Widgerson. How to play any mental game — a completeness theorem for protocols with honest majority. In *19th STOC*, pages 218–229, 1987.

[23] O. Goldreich, S. Micali, and A. Widgerson. Proofs that yield nothing but their validity or all languages in $\mathcal{NP}$ have zero-knowledge proof systems. *J. ACM*, 38(1):691–729, 1991.

[24] S. Goldwasser, S. Micali, and R. Rivest. A digital signature scheme secure against adaptive chosen-message attacks. *SIAM J. on Computing*, 17(2):281–308, 1988.

[25] I. Haitner, O. Horvitz, J. Katz, C. Koo, R. Morselli, and R. Shaltiel. Reducing complexity assumptions for statistically-hiding commitment. In *Advances in Cryptology — Eurocrypt 2005*, pages 58–77, 2005.

[26] I. Haitner and O. Reingold. A new interactive hashing theorem. In *22nd Computational Complexity Conference*, 2007. Draft of full version appears in www.wisdom.weizmann.ac.il/ iftachh/papers/IneractiveHashing.pdf.

[27] I. Haitner and O. Reingold. Statistically-hiding commitment from any one-way function. In *39th STOC*, 2007.

[28] S. Halevi and S. Micali. Practical and provably-secure commitment schemes from collision-free hashing. In *Advances in Cryptology — CRYPTO '96*, pages 201–215, 1996.

[29] J. Håstad, R. Impagliazzo, L. Levin, and M. Luby. A pseudorandom generator from any one-way function. *SIAM J. Comput.*, 28(4):1364–1396, 1999.

[30] R. Impagliazzo and M. Luby. One-way functions are essential for complexity-based cryptography. In *30th FOCS*, pages 230–235, 1989.

[31] R. Impagliazzo and S. Rudich. Limits on the provable consequences of one-way permutations. In *21st STOC*, pages 44–61, 1989.

[32] Y. Lindell. Parallel coin-tossing and constant-round secure two-party computation. *Journal of Cryptology*, 16(3):143–184, 2003.

[33] M. Naor. Bit commitment using pseudorandomness. *Journal of Cryptology*, 4(2):151–158, 1991.

[34] M. Naor, R. Ostrovsky, R. Venkatesan, and M. Yung. Perfect zero-knowledge arguments for $\mathcal{NP}$ using any one-way permutation. *J. Crypto.*, 11(2):87–108, 1998.

[35] M. Naor and M. Yung. Universal one-way hash functions and their cryptographic applications. In *21st STOC*, pages 33–43, 1989.

[36] M. Nguyen, S. Ong, and S. Vadhan. Statistical zero-knowledge arguments for $\mathcal{NP}$ from any one-way function. In *39th FOCS*, 2006.

[37] R. Ostrovsky, R. Venkatesan, and M. Yung. Secure commitment against a powerful adversary. In *STACS*, pages 439–448, 1992.

[38] R. Ostrovsky, R. Venkatesan, and M. Yung. Fair games against an all-powerful adversary. In *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, volume 13, pages 418–429, 1993.

[39] R. Raz. A parallel repetition theorem. *SIAM J. Computing*, 27(3):763–803, 1998.

[40] J. Rompel. One-way functions are necessary and sufficient for secure signatures. In *22nd STOC*, pages 387–394, 1990.

[41] A. Russell. Necessary and sufficient conditions for collision-free hashing. *J. Cryptology*, 8(2):87–100, 1995.

[42] A. Yao. Theory and application of trapdoor functions. In *23rd FOCS*, pages 80–91, 1982.