Electronic Colloquium on Computational Complexity, Report No. 82 (2014)

# The Randomized Iterate Revisited - Almost Linear Seed Length PRGs from A Broader Class of One-way Functions

Yu Yu* Dawu Gu† Xiangxue Li‡

**Abstract**

We revisit "the randomized iterate" technique that was originally used by Goldreich, Krawczyk, and Luby (SICOMP 1993) and refined by Haitner, Harnik and Reingold (CRYPTO 2006) in constructing pseudorandom generators (PRGs) from regular one-way functions (OWFs). We abstract out a technical lemma with connections to several recent work on cryptography with imperfect randomness, which provides an arguably simpler and more modular proof for the Haitner-Harnik-Reingold PRGs from regular OWFs.

We extend the approach to a more general construction of PRGs with seed length $O(n\log n)$ from a broader class of OWFs. More specifically, consider an arbitrary one-way function $f$ whose range is divided into sets $\mathcal{Y}_1$, $\mathcal{Y}_2$, ..., $\mathcal{Y}_n$ where each $\mathcal{Y}_i \stackrel{\mathsf{def}}{=} \{y : 2^{i-1} \leq |f^{-1}(y)| < 2^i\}$. We say that the maximal preimage size of $f$ is $2^{\max}$ if $\mathcal{Y}_{\max}$ has some noticeable portion (say $n^{-c}$ for constant $c$), and $\mathcal{Y}_{\max+1}, \ldots, \mathcal{Y}_n$ only sum to a negligible fraction $\epsilon$. We construct a PRG by making $\tilde{O}(n^{2c+1})$ calls to the underlying OWF and achieve seed length $O(n\log n)$ using bounded space generators, where the only parameter required to know is $c$ (which is constant for a specific $f$ but may vary for different functions) and no knowledge is required for max and $\epsilon$. We use a proof technique that is similar to and extended from the method by Haitner, Harnik and Reingold for hardness amplification of regular weakly one-way functions.

Our construction achieves almost linear seed length for a broader class of one-way functions than previously known, where the case of regular OWFs follows as a simple corollary for $c = 0$. We show that although an arbitrary one-way function may not fall into the class of OWFs as we defined, the counterexamples must satisfy a very strong condition and thus should be somewhat artificial. Our approach takes a different route from the generic HILL-style generators (which is characterized by flattening Shannon entropy sources) where the best known construction by Vadhan and Zheng (STOC 2012) requires seed length $O(n^3)$.

**Keywords:** Foundations, Pseudorandom Generators, One-way Functions, the Randomized Iterate.

*Shanghai Jiaotong University. Email: `yuyuathk@gmail.com`.
†Shanghai Jiaotong University.
‡East China Normal University.

ISSN 1433-8092

# 1 Introduction

That one-way functions (OWFs) imply pseudorandom generators (PRGs) [10] is one of the central results upon which modern cryptography is successfully founded. The problem dates back to the early 80's when Blum, Micali [2] and Yao [15] independently observed that a PRG (often referred to as the BMY generator) can be efficiently constructed from one-way permutations (OWPs). That is, given a OWP $f$ on $n$-bit input $x$ and its hardcore predicate $h_c$ (e.g., by Goldreich and Levin [6]), a single invocation of $f$ already implies a PRG $g : x \mapsto (f(x), h_c(x))$ with a stretch[1] of $\Omega(\log\ n)$ bits and it extends to arbitrary stretch by repeated iterations (seen by a hybrid argument). Unfortunately, the BMY generator does not immediately apply to an arbitrary OWF since the output of $f$ might be of too small amount of entropy to be secure for subsequent iterations.

The sequential approach - the Randomized Iterate. Goldreich, Krawczyk, and Luby [5] extended the BMY generator by inserting a randomized operation (using $k$-wise independent hash functions) into every two applications of $f$, from which they built a PRG of seed length $O(n^3)$ assuming that the underlying OWF is known-regular[2]. Haitner, Harnik and Reingold [8] further refined the approach (for which they coined the name "*the randomized iterate*") as below:

$$x_1 \xrightarrow{f} y_1 \xrightarrow{h_1} x_2 \xrightarrow{f} y_2 \xrightarrow{h_2} \quad \cdots \quad x_k \xrightarrow{f} y_k \xrightarrow{h_{k+1}}$$

where in between every $i^{th}$ and $(i+1)^{th}$ iterations a random pairwise-independent hash function $h_i$ is applied. Haitner et al. [8] showed that, when $f$ is instantiated with any (possibly unknown) regular one-way function, it is hard to invert any $k^{th}$ iterate (i.e., recovering any $x_k$ s.t. $f(x_k) = y_k$) given $y_k$ and the description of the hash functions. This gives a PRG of seed length $O(n^2)$ by running the iterate $n+1$ times and outputting a hardcore bit at every iteration. The authors of [7] further derandomize the PRG by generating all the hash functions from bounded space generators (e.g., Nisan's generator [13]) using a seed of length $O(n \log\ n)$. Although the randomized iterate is mostly known for construction of PRGs from regular OWFs, the authors of [7] also introduced many other interesting applications such as linear seed length PRGs from any exponentially hard regular OWFs, $O(n^2)$ seed length PRGs from any exponentially hard OWFs, $O(n^7)$ seed length PRGs from any OWFs, and hardness amplification of regular weakly OWFs. Recently, Yu et al. [16] further reduced the seed length of the PRG (based on any regular OWFs) to $O(\omega(1){\cdot}n)$ for any efficiently computable $\omega(1)$.

The parallel approach - PRGs from any OWFs. Håstad, Impagliazzo, Levin and Luby (HILL) [10] presented the seminal result that pseudorandom generators can be constructed from any one-way functions. Nevertheless, they only gave a complicated (and not practically efficient) construction of PRG with seed length $\tilde{O}(n^{10})$ and sketched another one with seed length $\tilde{O}(n^8)$, which was formalized and proven in [11]. Haitner, Reingold, and Vadhan [9] introduced the notion of next-block pseudoentropy, and gave a uniform construction of seed length $\tilde{O}(n^4)$ as well as a non-uniform[3] construction of seed length $\tilde{O}(n^3)$. Vadhan and Zheng [14] further reduced the seed length of the uniform construction to $\tilde{O}(n^3)$, matching the non-uniform one of [9]. In summary, it remains open how to construct the PRGs with seed length below $\tilde{O}(n^3)$ even given polynomial-size non-uniform advice about an arbitrary one-way function or by making adaptive calls to the function.

A technical lemma. In this paper, we revisit the randomized iterate. We abstract out a technical lemma that, informally speaking, "if any algorithm wins a one-sided game (e.g., inverting a OWF) on uniformly sampled challenges only with some negligible probability, then it cannot do much better (beyond a negligible advantage) in case that the challenges are sampled from any distribution of logarithmic

[1]The stretch of a PRG refers to the difference between output and input lengths (see Definition 2.3).

[2]A function $f(x)$ is regular if every image has the same number (say $\alpha$) of preimages, and it is known- (resp., unknown-) regular if $\alpha$ is efficiently computable (resp., inefficient to approximate) from the security parameter.

[3]A "non-uniform" PRG may assume arbitrary polynomial-size advice about the underlying OWF.

collision entropy deficiency". Similar observations were made in related settings[1, 4, 3], where either the game is two-sided (e.g., indistinguishability applications) or the randomness is sampled from slightly defected min-entropy source. Plugging this lemma into [7] immediately yields a simpler proof for the key lemma of [7], namely, "any $k^{th}$ iterate (instantiated with a regular OWF) is hard-to-invert". The rationale is that $y_k$ has sufficiently high collision entropy (even conditioned on the hash functions) that is only logarithmically less than the ideal case where $y_k$ is uniform (over the range of $f$) and independent of the hash functions, which is hard to invert by the one-wayness assumption.

The main results. We consider an arbitrary OWF $f$ with range divided into sets $\mathcal{Y}_1, \ldots, \mathcal{Y}_n$, where each $\mathcal{Y}_i \stackrel{\mathsf{def}}{=} \{y : 2^{i-1} \leq |f^{-1}(y)| < 2^i\}$. We say that the maximal preimage size of $f$ is $2^{\max}$ if $\mathcal{Y}_{\max}$ has some noticeable portion ($n^{-c}$ for constant $c$), and $\mathcal{Y}_{\max+1}, \ldots, \mathcal{Y}_n$ only sum to a negligible fraction $\mathsf{negl}(n)$. It is easy to see that regular one-way functions are a special case for $c = 0$. We further show that in order to be a counterexample to the class of one-way functions defined above, the candidate function must satisfy an infinite set of conditions and should be somewhat artificial. This gives some non-trivial evidence that a natural one-way function should fall into the class of one-way functions. We give a construction that only requires the knowledge about $c$ (i.e., oblivious of max and $\mathsf{negl}$). Loosely speaking, the main idea is that conditioned on $y_k \in \mathcal{Y}_{\max}$ the collision entropy of $y_k$ given the hash functions is close to the ideal case where $f(U_n)$ hits $\mathcal{Y}_{\max}$ with noticeable probability (and is independent of the hash functions), which is hard to invert. We have by the pairwise independence (in fact, universality already suffices) of the hash functions that every $y_k \in \mathcal{Y}_{\max}$ is an independent event of probability $n^{-c}$. By a Chernoff bound, running the iterate $\Delta = n^{2c} \cdot \omega(\log n)$ times yields that with overwhelming probability there is at least one occurrence of $y_k \in \mathcal{Y}_{\max}$, which implies every $\Delta$ iterations are hard-to-invert, i.e., for any $j = \mathsf{poly}(n)$ it is hard to predict $x_{1+(j-1)\Delta}$ given $y_{j\Delta}$ and the hash functions. A PRG follows by outputting $\log n$ hardcore bits for every $\Delta$ iterations and in total making $\tilde{O}(n^{2c+1})$ calls to $f$. This requires seed length $\tilde{O}(n^{2c+2})$, and can be pushed to $O(n \cdot \log n)$ bits using bounded space generators [13, 12], ideas borrowed from [7] with more significant reductions in seed length (we reduce by factor $\tilde{O}(n^{2c+1})$ whereas [7] saves factor $\tilde{O}(n)$). Overall, our technique is similar in spirit to the hardness amplification of weakly one-way functions introduced by Haitner et al. in the same paper [7] (see full proof in its full version [8]). Roughly speaking, the idea was that for any inverting algorithm $\mathsf{A}$, a weakly one-way function has a set that $\mathsf{A}$ fails upon (the failing-set of $\mathsf{A}$), and thus sufficiently many iterations are bound to hit every such failing-set (for every inverting algorithm) to make a strongly one-way function. However, in our case the lack of a regular structure of the underlying function and the negligible fraction of $\mathcal{Y}_{\max+1}, \ldots, \mathcal{Y}_n$ further complicate the analysis (see Remark B.1 for some discussions), and we use our technical lemma to provide a neat and modular proof.

## 2 Preliminaries

Notations and definitions. We use $[n]$ to denote set $\{1, \ldots, n\}$. We use capital letters (e.g., $X$, $Y$) for random variables, standard letters (e.g., $x$, $y$) for values, and calligraphic letters (e.g., $\mathcal{Y}$, $\mathcal{S}$) for sets. $|\mathcal{S}|$ denotes the cardinality of set $\mathcal{S}$. We use shorthand $\mathcal{Y}_{[n]} \stackrel{\mathsf{def}}{=} \bigcup_{t=1}^{n} \mathcal{Y}_t$. For function $f : \{0,1\}^n \rightarrow \{0,1\}^{l(n)}$, we use shorthand $f(\{0,1\}^n) \stackrel{\mathsf{def}}{=} \{f(x) : x \in \{0,1\}^n\}$, and denote by $f^{-1}(y)$ the set of $y$'s preimages under $f$, i.e. $f^{-1}(y) \stackrel{\mathsf{def}}{=} \{x : f(x) = y\}$. We use $s \leftarrow S$ to denote sampling an element $s$ according to distribution $S$, and let $s \stackrel{\$}{\leftarrow} \mathcal{S}$ denote sampling $s$ uniformly from set $\mathcal{S}$, and $y := f(x)$ denote value assignment. We use $U_n$ and $U_{\mathcal{X}}$ to denote uniform distributions over $\{0,1\}^n$ and $\mathcal{X}$ respectively, and let $f(U_n)$ be the distribution induced by applying function $f$ to $U_n$. We use $\mathsf{CP}(X)$ to denote the collision probability of $X$, i.e., $\mathsf{CP}(X) \stackrel{\text{def}}{=} \sum_x \Pr[X = x]^2$, and denote by $\mathbf{H}_2(X) \stackrel{\text{def}}{=} -\log \mathsf{CP}(X)$ the collision entropy. We also define conditional collision entropy (and probability) of a random variable $X$ conditioned on another random variable $Z$ by

$$\mathbf{H}_2(X|Z) \stackrel{\text{def}}{=} -\log\left(\ \mathsf{CP}(X|Z)\ \right) \stackrel{\text{def}}{=} -\log\left(\ \mathbb{E}_{z\leftarrow Z}\left[\ \sum_x \Pr[X=x|\ Z=z]^2\ \right]\ \right)$$

A function $\mathsf{negl}:\mathbb{N}\rightarrow[0,1]$ is negligible if for every constant $c$ we have $\mathsf{negl}(n)<n^{-c}$ holds for all sufficiently large $n$'s, and a function $\mu:\mathbb{N}\rightarrow[0,1]$ is called noticeable if there exists constant $c$ such that $\mu(n)\geq n^{-c}$ for all sufficiently large $n$'s.

We define the *computational distance* between distribution ensembles $X \stackrel{\mathsf{def}}{=} \{X_n\}_{n\in\mathbb{N}}$ and $Y \stackrel{\mathsf{def}}{=} \{Y_n\}_{n\in\mathbb{N}}$, denoted by $\mathsf{CD}_{T(n)}(X,Y)\leq\varepsilon(n)$, if for every probabilistic distinguisher $\mathsf{D}$ of running time up to $T(n)$ it holds that

$$|\ \Pr[\mathsf{D}(1^n,X_n)=1]-\Pr[\mathsf{D}(1^n,Y_n)=1]\ |\leq\varepsilon(n)\ .$$

The *statistical distance* between $X$ and $Y$, denoted by $\mathsf{SD}(X,Y)$, is defined by

$$\mathsf{SD}(X,Y) \stackrel{\mathsf{def}}{=} \frac{1}{2}\sum_x |\Pr[X=x]-\Pr[Y=x]| = \ \mathsf{CD}_\infty(X,Y)$$

We use $\mathsf{SD}(X,Y|Z)$ (resp. $\mathsf{CD}_t(X,Y|Z)$) as shorthand for $\mathsf{SD}((X,Z),(Y,Z))$ (resp. $\mathsf{CD}_t((X,Z),(Y,Z))$).

Simplifying Assumptions and Notations. To simplify the presentation, we make the following assumptions without loss of generality. It is folklore that one-way functions can be assumed to be length-preserving (see [8] for full proofs). Throughout, most parameters are functions of the security parameter $n$ (e.g., $T(n)$, $\varepsilon(n)$, $\alpha(n)$) and we often omit $n$ when clear from the context (e.g., $T$, $\varepsilon$, $\alpha$). By notation $f:\{0,1\}^n\rightarrow\{0,1\}^l$ we refer to the ensemble of functions $\{f_n:\{0,1\}^n\rightarrow\{0,1\}^{l(n)}\}_{n\in\mathbb{N}}$. As slight abuse of notion, $\mathsf{poly}$ might be referring to the set of all polynomials or a certain polynomial, and $h$ might be either a function or its description, which will be clear from the context.

**Definition 2.1 (pairwise independent hashing)** *A family of hash functions $\mathcal{H} \stackrel{\mathsf{def}}{=} \{h\ :\ \{0,1\}^n\rightarrow\{0,1\}^m\}$ is* ***pairwise independent*** *if for any $x_1\neq x_2\in\{0,1\}^n$ and any $v\in\{0,1\}^{2m}$ it holds that $\Pr_{h\xleftarrow{\$}\mathcal{H}}[(h(x_1),h(x_2))=v]=2^{-2m}$, or equivalently, $(H(x_1),\ H(x_2))$ is i.i.d. to $U_{2m}$ where $H$ is uniform over $\mathcal{H}$. It is well known that there are efficiently computable families of pairwise independent hash functions of description length $\Theta(n+m)$.*

**Definition 2.2 (one-way functions)** *A function $f:\{0,1\}^n\rightarrow\{0,1\}^{l(n)}$ is $(T(n)$,$\varepsilon(n))$-one-way if $f$ is polynomial-time computable and for any probabilistic algorithm $\mathsf{A}$ of running time $T(n)$*

$$\Pr_{y\leftarrow f(U_n)}[\mathsf{A}(1^n,y)\in f^{-1}(y)]\ \leq\ \varepsilon(n).$$

*We say that $f$ is a (strongly) one-way function if $T(n)$ and $1/\varepsilon(n)$ are both super-polynomial in $n$.*

**Definition 2.3 (pseudorandom generators [2, 15])** *A deterministic function $g:\{0,1\}^n\rightarrow\{0,1\}^{n+s(n)}$ $(s(n)>0)$ is a $(T(n)$,$\varepsilon(n))$-secure PRG with stretch $s(n)$ if $g$ is polynomial-time computable and*

$$\mathsf{CD}_{T(n)}(\ g(1^n,U_n)\ ,\ U_{n+s(n)}\ )\ \leq\ \varepsilon(n).$$

*We say that $g$ is a pseudorandom generator if $T(n)$ and $1/\varepsilon(n)$ are both super-polynomial in $n$.*

## 2.1 Technical Tools

**Theorem 2.1 (Goldreich-Levin Theorem [6])** *Let $(X,Y)$ be a distribution ensemble over $\{\{0,1\}^n\times\{0,1\}^{\mathsf{poly}(n)}\}_{n\in\mathbb{N}}$. Assume that for any PPT algorithm $\mathsf{A}$ of running time $T(n)$ it holds that*

$$\Pr[\ \mathsf{A}(1^n,Y)=X\ ]\ \leq\ \varepsilon(n)$$

*Then, for any efficiently computable $m(n)\in O(n)$, there exists an efficient function family $\mathcal{H}_c \stackrel{\mathsf{def}}{=} \{h_c : \{0,1\}^n \rightarrow \{0,1\}^{m(n)}\}$ of description size $\Theta(n)$, such that*

$$\mathsf{CD}_{T'(n)}(\ H_c(X)\ ,\ U_{m(n)}\ \mid\ Y, H_c)\ \in\ O(2^{m(n)}\cdot(n\cdot\varepsilon)^{\frac{1}{3}})\ .$$

*where $T'(n) = T(n)\cdot(\varepsilon(n)/n)^{O(1)}$, and $H_c$ is the uniform distribution over $\mathcal{H}_c$.*

**Definition 2.4 (bounded-width layered branching program - LBP)** *An $(s,k,v)$-LBP $M$ is a finite directed acyclic graph whose nodes are partitioned into $k+1$ layers indexed by $\{1, \ldots, k+1\}$. The first layer has a single node (the source), the last layer has two nodes (sinks) labeled with 0 and 1, and each of the intermediate layers has up to $2^s$ nodes. Each node in the $i \in [k]$ layer has exactly $2^v$ outgoing labeled edges to the $(i+1)^{th}$ layer, one for every possible string $h_i \in \{0,1\}^v$.*

An equivalent (and somewhat more intuitive) model to the above is bounded space computation. That is, we assign labels to graph nodes (instead of associating them with the edges), at each $i^{th}$ layer the program performs arbitrary computation on the current node (labelled by $s$-bit string) and the current $v$-bit input $h_i$, advances (and assigns value) to a node in the $(i+1)^{th}$ layer, and repeats until it reaches the last layer to produce the final output bit.

**Theorem 2.2 (bounded-space generator [13, 12])** *Let $s(n), k(n), v(n) \in \mathbb{N}$ and $\varepsilon(n) \in (0,1)$ be polynomial-time computable functions. Then, there exist a polynomial-time computable function $q(n) \in \Theta(v(n) + (s(n) + \log(k(n)/\varepsilon(n)))\log k(n))$ and a generator $BSG : \{0,1\}^{q(n)} \rightarrow \{0,1\}^{k(n)\cdot v(n)}$ that runs in time $\mathsf{poly}(s(n), k(n), v(n), \log(1/\varepsilon(n)))$, and $\varepsilon(n)$-fools every $(s(n), k(n), v(n))$- LBP $M$, i.e.,*

$$|\ \Pr[M(U_{k(n)\cdot v(n)}) = 1]\ -\ \Pr[M(BSG(U_n)) = 1]\ |\ \leq\ \varepsilon(n)\ .$$

# 3 Pseudorandom Generators from Regular One-way Functions

## 3.1 A Technical Lemma

Before we revisit the randomize iterated PRGs from regular one-way functions, we introduce a technical lemma that simplifies the analysis in [7] and our main results in Section 5, and might be of independent interests. Informally, it states that if any one-sided game (one-way functions, MACs, and digital signatures) is $(T,\varepsilon)$-secure on uniform secret randomness, then it will be $(T,2\sqrt{2^e\cdot\varepsilon})$-secure when the randomness is sampled from any distribution with $e$ bits of collision entropy deficiency[4].

**Lemma 3.1 (one-sided game on imperfect randomness)** *For any $e \leq m \in \mathbb{N}$, let $\mathcal{W}\times\mathcal{Z}$ be any set with $|\mathcal{W}| = 2^m$, let $\mathsf{Adv} : \mathcal{W}\times\mathcal{Z} \rightarrow [0,1]$ be any (deterministic) real-valued function, let $(W, Z)$ be any joint random variables over set $\mathcal{W}\times\mathcal{Z}$ satisfying $\mathbf{H}_2(W|Z) \geq m - e$, we have*

$$\mathbb{E}[\mathsf{Adv}(W,Z)]\ \leq\ \sqrt{2^{e+2}\cdot\mathbb{E}[\mathsf{Adv}(U_{\mathcal{W}}, Z)]} \tag{1}$$

*where $U_{\mathcal{W}}$ denotes uniform distribution over $\mathcal{W}$ (independent of $Z$ and any other distributions).*

*Proof.* For any given $\delta$ define $\mathcal{S}_\delta \stackrel{\mathsf{def}}{=} \{(w,z) : \Pr[W = w|Z = z] \geq 2^{-(m-e)}/\delta\}$

$$\begin{aligned} 2^{-(m-e)} &\geq \sum_z \Pr[Z=z]\sum_w \Pr[W=w|Z=z]^2 \\ &\geq \sum_z \Pr[Z=z]\sum_{w:(w,z)\in\mathcal{S}_\delta} \Pr[W=w|Z=z]\cdot 2^{-(m-e)}/\delta \\ &\geq (2^{-(m-e)}/\delta)\cdot\Pr[(W,Z)\in\mathcal{S}_\delta]\ , \end{aligned}$$

[4] The collision entropy deficiency of a random variable $W$ over set $\mathcal{W}$ is defined as the difference between entropies of $U_{\mathcal{W}}$ and $W$, i.e., $\log|\mathcal{W}| - \mathbf{H}_2(W)$.

and thus $\Pr[(W,Z)\in\mathcal{S}_\delta]\le\delta$. It follows that

$$\begin{aligned}\mathbb{E}[\mathsf{Adv}(W,Z)] &= \sum_{(w,z)\in\mathcal{S}_\delta}\Pr[(W,Z)=(w,z)]\cdot\mathsf{Adv}(w,z) \;+\; \sum_{(w,z)\notin\mathcal{S}_\delta}\Pr[Z=z]\cdot\Pr[W=w|Z=z]\cdot\mathsf{Adv}(w,z)\\ &\le \sum_{(w,z)\in\mathcal{S}_\delta}\Pr[(W,Z)=(w,z)] \;+\; (2^e/\delta)\cdot\sum_{(w,z)\notin\mathcal{S}_\delta}\Pr[Z=z]\cdot 2^{-m}\cdot\mathsf{Adv}(w,z)\\ &\le \delta \;+\; (2^e/\delta)\cdot\mathbb{E}[\mathsf{Adv}(U_{\mathcal{W}},Z)]\;,\end{aligned}$$

and we complete the proof by setting $\delta=\sqrt{2^e\cdot\mathbb{E}[\mathsf{Adv}(U_{\mathcal{W}},Z)]}$. □

### 3.2 The Randomized Iterate

**Definition 3.1 (the randomized iterate [7, 5])** *Let $n\in\mathbb{N}$, function $f:\{0,1\}^n\to\{0,1\}^n$, and let $\mathcal{H}$ be a family of pairwise-independent length-preserving hash functions over $\{0,1\}^n$. For $k\in\mathbb{N}$, $x_1\in\{0,1\}^n$ and vector $\vec{h}^k=(h_1,\ldots,h_k)\in\mathcal{H}^k$, recursively define the $k^{th}$ randomized iterate by:*

$$y_k=f(x_k),\;\; x_{k+1}=h_k(y_k)$$

*For $k-1\le t\in\mathbb{N}$, we denote the $k^{th}$ iterate by function $f^k$, i.e., $y_k = f^k(x_1,\vec{h}^t)$, where $\vec{h}^t$ is possibly redundant as $y_k$ only depends on $\vec{h}^{k-1}$.*
*The **randomized version** refers to the case where $x_1\xleftarrow{\$}\{0,1\}^n$ and $\vec{h}^{k-1}\xleftarrow{\$}\mathcal{H}^{k-1}$.*
*The **derandomized version** refers to that $x_1\xleftarrow{\$}\{0,1\}^n$, $\vec{h}^{k-1}\leftarrow BSG(U_q)$, where $q\in\Theta(n{\cdot}\log n)$, $BSG:\{0,1\}^q\to\{0,1\}^{(k-1)\cdot\log|\mathcal{H}|}$ is a bounded-space generator[5] that $2^{-2n}$-fools every $(2n{+}1,k,\log|\mathcal{H}|)$-LBP, and $\log|\mathcal{H}|$ is the description length of $\mathcal{H}$ (e.g., $2n$ bits for concreteness).*

**Theorem 3.1 (PRGs from Regular OWFs [7])** *For $n\in\mathbb{N},k\in[n+1]$, let $f$, $\mathcal{H}$, $f^k$ and $BSG(\cdot)$ be as defined in Definition 3.1, and let $\mathcal{H}_c=\{h_c:\{0,1\}^n\to\{0,1\}\}$ be a family of Goldreich-Levin predicates, where $\mathcal{H}$ and $\mathcal{H}_c$ both have description length $\Theta(n)$. We define $G:\{0,1\}^n\times\mathcal{H}^n\times\mathcal{H}_c\to\{0,1\}^{n+1}\times\mathcal{H}^n\times\mathcal{H}_c$ and $G':\{0,1\}^n\times\{0,1\}^{q(n)}\times\mathcal{H}_c\to\{0,1\}^{n+1}\times\{0,1\}^{q(n)}\times\mathcal{H}_c$ as below:*

$$G(x_1,\vec{h}^n,h_c)=(h_c(x_1),h_c(x_2),\ldots,h_c(x_{n+1}),\vec{h}^n,h_c).$$

$$G'(x_1,u,h_c)=G(x_1,BSG(u),h_c).$$

*Assume that $f$ is a regular (length-preserving) one-way function and that $BSG(\cdot)$, $\mathcal{H}$ and $\mathcal{H}_c$ are efficient. Then, $G$ and $G'$ are pseudorandom generators.*

Proof sketch of Theorem 3.1. It suffices to prove Lemma 3.2. Namely, for any $1{\le}k\le n+1$, given $y_k$ and the hash functions (either sampled uniformly or from bounded space generators), it is hard to recover any $x_k$ satisfying $f(x_k)=y_k$. Then, Goldreich-Levin Theorem yields that each $h_c(x_k)$ is computationally unpredictable given $y_k$. Note that $y_k$ implies all the subsequent $h_c(x_{k+1})$, …, $h_c(x_{n+1})$. We complete the proof by Yao's "next (previous) bit unpredictability implies pseudorandomness" argument [15]. It thus remains to prove Lemma 3.2 below which summarizes the statements of Lemma 3.2, Lemma 3.4, Lemma 3.11 from [8], and we provide a simpler proof.

**Lemma 3.2 (the $k^{th}$ iterate is hard-to-invert)** *For any $n\in\mathbb{N},k\in[n+1]$, let $f$, $\mathcal{H}$, $f^k$ be as defined in Definition 3.1. Assume that $f$ is a $(T,\varepsilon)$ regular one-way function, i.e., for every PPT $\mathsf{A}$ and $\mathsf{A}'$ of running time $T$ it holds that*

$$\Pr\;[\;\mathsf{A}(\;f(U_n),\;\vec{H}^n)\in f^{-1}(f(U_n))\;]\;\le\;\varepsilon\quad.$$

[5] Such efficient generators exists by Theorem 2.2, setting $s(n)=2n+1$, $q(n)=\mathsf{poly}(n)$, $v(n)=\log|\mathcal{H}|$ and $\varepsilon(n)=2^{-2n}$.

$$\Pr\ [\ \mathsf{A}'(\ f(U_n),\ U_q\ ) \in f^{-1}(f(U_n)\ )\ ] \ \leq\ \varepsilon \quad .$$

*Then, for every such* $\mathsf{A}$ *and* $\mathsf{A}'$ *it holds that*

$$\Pr\ [\ \mathsf{A}(Y_k,\ \vec{H}^n) \in f^{-1}(Y_k)\ ] \ \leq\ 2\sqrt{k\cdot\varepsilon} \quad , \tag{2}$$

$$\Pr\ [\ \mathsf{A}'(Y'_k,\ U_q\ ) \in f^{-1}(Y'_k)\ ] \ \leq\ 2\sqrt{(k+1)\cdot\varepsilon} \quad , \tag{3}$$

*where* $Y_k = f^k(X_1, \vec{H}^n)$, $Y'_k = f^k(X_1, BSG(U_q))$, $X_1$ *is uniform over* $\{0,1\}^n$ *and* $\vec{H}^n$ *is uniform over* $\mathcal{H}^n$.

***A simpler proof of Lemma 3.2 via Lemma 3.1.*** To apply Lemma 3.1, let $\mathcal{W} = f(\{0,1\}^n)$, $\mathcal{Z} = \mathcal{H}^n$, let $(W,Z) = (Y_k, \vec{H}^n)$, $U_{\mathcal{W}} = f(U_n)$, and define

$$\mathsf{Adv}(y, \vec{h}^n) \stackrel{\mathsf{def}}{=} \begin{cases} 1, & \text{if } \mathsf{A}(y,\ \vec{h}^n) \in f^{-1}(y) \\ 0, & \text{if } \mathsf{A}(y,\ \vec{h}^n) \notin f^{-1}(y) \end{cases}$$

where $\mathsf{A}$ is assumed to be deterministic without loss of generality[6]. We have by Lemma 3.3 that

$$\mathbf{H}_2(Y_k \mid \vec{H}^n) \geq \mathbf{H}_2(\ f(U_n) \mid \vec{H}^n\ )\ -\ \log k$$

and thus Lemma 3.1 yields that

$$\Pr\ [\ \mathsf{A}(Y_k,\ \vec{H}^n) \in f^{-1}(Y_k)\ ] \ \leq\ 2\sqrt{k\cdot \Pr\ [\ \mathsf{A}(f(U_n),\ \vec{H}^n) \in f^{-1}(f(U_n))\ ]} \leq 2\sqrt{k\cdot\varepsilon}\ .$$

The proof for (3) is similar except for setting ( $W = Y'_k$, $Z = U_q$ ) and letting $\mathsf{Adv}(y,u) = 1$ iff $\mathsf{A}'(y,u) \in f^{-1}(y)$. We have by Lemma 3.3 that

$$\mathbf{H}_2(Y'_k \mid U_q\ ) \geq \mathbf{H}_2(\ f(U_n) \mid U_q\ )\ -\ \log{(k+1)}$$

and thus we apply Lemma 3.1 to get

$$\Pr\ [\ \mathsf{A}'(Y'_k,\ U_q\ ) \in f^{-1}(Y_k)\ ] \ \leq\ 2\sqrt{(k+1)\cdot \Pr\ [\ \mathsf{A}'(f(U_n), U_q\ ) \in f^{-1}(f(U_n))\ ]} \ \leq\ 2\sqrt{(k+1)\cdot\varepsilon}\ .$$

□

The proof of Lemma 3.3 below appeared in [7], and we include it in Appendix A for completeness.

**Lemma 3.3 (Collision Entropy [7])** *For the same assumptions and notations as in Lemma 3.2, it holds that*

$$\mathsf{CP}(\ f(U_n)\ ) \ =\ \mathsf{CP}(\ f(U_n) \mid \vec{H}^n\ ) \ =\ \mathsf{CP}(\ f(U_n) \mid BSG(U_q), U_q\ ) \ =\ \frac{1}{|f(\{0,1\}^n)|} \ , \tag{4}$$

$$\mathsf{CP}(Y_k \mid \vec{H}^n\ ) \ \leq\ \frac{k}{|f(\{0,1\}^n)|} \ , \tag{5}$$

$$\mathsf{CP}(Y'_k \mid BSG(U_q), U_q\ ) \ \leq\ \frac{k+1}{|f(\{0,1\}^n)|} \ . \tag{6}$$

[6] If $\mathsf{A}$ is probabilistic, let $\mathsf{Adv}\ (y, \vec{h}^n)$= $\Pr[\mathsf{A}(y, \vec{h}^n) \in f^{-1}(y)]$, where probability is taken over the internal coins of $\mathsf{A}$.

## 4 A New Class of OWFs vs. Regular/Arbitrary OWFs

In this section, we introduce the class of OWFs from which almost linear seed length PRGs can be constructed. We show that (almost-)regular OWFs fall into a special case (for $c = 0$) along with the argument that "if a one-way function behaves like a random function, then it is almost regular". More generally, the class covers a wider range of one-way functions (for positive $c \in \mathbb{N}$) than regular ones. We also (attempt to) characterize functions that are not captured by our definition. We show that in order not to fall into our class of one-way functions, the counterexamples must satisfy infinitely many conditions.

**Definition 4.1 (OWFs with $n^{-c}$-weighted maximal-size preimages)** *For constant c, we say that $f : \{0,1\}^n \to \{0,1\}^{l(n)}$ has an $n^{-c}$-fraction of **maximal-size preimages** if there exist (not necessarily efficient) function $\max(n) \in \mathbb{N}$ and negligible function $\epsilon(n) \in [0,1]$ such that for every $n \in \mathbb{N}$, it holds that :*

$$\Pr[f(U_n) \in \mathcal{Y}_{\max(n)}] \geq n^{-c} \tag{7}$$

$$\Pr[\ f(U_n) \in \bigcup_{j=\max(n)+1}^{n} \mathcal{Y}_j\ ] \leq \epsilon(n)\ , \tag{8}$$

*where $\mathcal{Y}_j \stackrel{\mathsf{def}}{=} \{y : 2^{j-1} \leq |f^{-1}(y)| < 2^j\}$.*
*For example, with $c = 0$ and efficiently (resp., inefficiently) computable $\max(\cdot)$ it is equivalent to assume that the underlying $f$ is a known- (resp., unknown-) regular function.*

Almost-regularity vs. the relaxed version of Definition 4.1. In fact, the construction by Haitner et al. [7] does not need strictly regular OWFs but only requires (implicit in its proof) that the logarithm of the pre-image size $\log|f^{-1}(y)|$ lies between some interval $[a, a+d]$ for some $d \in O(\log(1/\varepsilon))$, where $\varepsilon$ is the hardness parameter (as in Definition 2.2). Likewise, our proposed PRG only assumes a relaxed version of (7). That is, for some $d \in O(\log(1/\varepsilon))$ it holds that $\mathcal{Y}_{\max-d}, \mathcal{Y}_{\max-d+1}, \ldots, \mathcal{Y}_{\max}$ (instead of $\mathcal{Y}_{\max}$ alone) sum to an $n^{-c}$-fraction (see (27)). Therefore, almost regular functions fall into the (extended) class of our OWFs for $c = 0$. We give the proof under the assumption of Definition 4.1 for brevity, and sketch how to adapt the proof to a weaker assumption than (7) in Remark B.2 (see Appendix B).

Now, we use probabilistic methods to argue that almost-regularity is a good assumption in the average-case sense. That is, if the one-way function is considered as randomly drawn from the set of all (not just one-way) functions, then it is very likely to be almost-regular and thus a PRG can be efficiently constructed. The proof of Lemma 4.1 to deferred to Appendix A.

**Lemma 4.1 (A random function is almost-regular.)** *Let $\mathcal{F} = \{f : \{0,1\}^n \to \{0,1\}^m\ \}$ be the set of all functions mapping n-bit to m-bit strings. For any $0 < d < n$, we have*

- *If $m \leq n - d$, then it holds that*

$$\Pr_{f \xleftarrow{\$} \mathcal{F}}[\ \mathsf{SD}(\ f(U_n),\ U_m\ )\ \leq\ 2^{-d/4}\ ]\ \geq\ 1 - 2^{-d/4}\ .$$

- *If $m > n - d$, then we have*

$$\Pr_{f \xleftarrow{\$} \mathcal{F},\ x \xleftarrow{\$} \{0,1\}^n}[\ 1\ \leq\ |f^{-1}(f(x))| \leq\ 2^{2d+1}\ ] \geq 1 - 2^{-d}\ .$$

*Typically, we can set $d \in \omega(\log n)$ so that $f$ will be almost regular except for a negligible fraction. Note that the first bullet gives even stronger guarantee than the second one does.*

Beyond regular functions. We cannot rule out the possibility that the one-way function in consideration is far from regular, namely (using the language of Definition 4.1), an arbitrary one-way function can have non-empty sets $\mathcal{Y}_i, \ldots, \mathcal{Y}_{i+O(n)}$. Previously the best known construction [14] requires seed length $\tilde{O}(n^3)$, and our PRG achieves seed length $O(n \cdot \log n)$ provided that Definition 4.1 is respected. Below we argue that Definition 4.1 is quite generic and any function that fails to satisfy it should be somewhat artificial. As a first attempt, one may argue that if we skip all those $\mathcal{Y}_j's$ (in the descending order of $j$) that sum to negligible, the first one that is non-negligible[7] (i.e., not meeting (8)) will satisfy (7) for at least infinitely many $n$'s. In other words, it seems that our PRG construction works for any one-way function (at least for infinitely many $n$'s). This argument is unfortunately problematic as (non-)negligible is a property of a sequence of probabilities, rather than a single one. However, we will follow this intuition and provide a remedied analysis below.

**Lemma 4.2 (A necessary condition to be a counterexample.)** *Let $f : \{0,1\}^n \to \{0,1\}^{l(n)}$ be any one-way function and denote $\mathcal{Y}_j \stackrel{\mathsf{def}}{=} \{y : 2^{j-1} \le |f^{-1}(y)| < 2^j\}$, and let $\kappa = \kappa(n)$ be the number of non-empty sets $\mathcal{Y}_j$ (that comprise the range of $f$) for any given $n$, and write them as $\mathcal{Y}_{i_1}, \mathcal{Y}_{i_2}, \ldots, \mathcal{Y}_{i_\kappa}$ with $i_1 < i_2 < \ldots < i_\kappa$. For **every** $n_0 \in \mathbb{N} \cup \{0\}$, it must hold that function $\mu_{n_0}(\cdot)$ defined as*

$$\mu_{n_0}(n) \stackrel{\mathsf{def}}{=} \begin{cases} \Pr[f(U_n) \in \mathcal{Y}_{i_{\kappa(n)-n_0}}], & \text{if } \kappa(n) > n_0 \\ 0 \qquad\qquad\qquad , & \text{if } \kappa(n) \le n_0 \end{cases} \tag{9}$$

*is negligible. Otherwise (if the above condition is not met), there exists constant $c \ge 0$, $\max(n) \in \mathbb{N}$ and negligible function $\epsilon(n) \in [0,1]$ such that (8) holds (for all $n$'s) and (7) holds for infinitely many $n$'s.*

(9) is necessary and strong. The above lemma formalizes a necessary condition to constitute a counterexample to Definition 4.1 (see Appendix A for its proof). It is necessary in the sense that any one-way function that does not satisfy it must satisfy Definition 4.1 from which our PRG can be efficiently built (for at least infinitely many $n$'s). Note that the condition is actually an infinite set of conditions by requiring every $\mu_{n_0}(n)$ (for $n_0 \in \mathbb{N}$) being negligible. At the same time, it holds unconditionally that all these $\mu_{n_0}(n)$ (that correspond to the weights of all non-empty sets) must sum to unity, i.e., for every $n$ we have

$$\mu_0(n) \;+\; \mu_1(n) \;+\; \ldots + \mu_{\kappa(n)-1}(n) \;=\; 1 \quad .$$

This might look mutually exclusive to (9) as if every $\mu_{n_0}(n)$ is negligible then the above sum should be upper bounded by $\kappa(n) \cdot \mathsf{negl}(n) = \mathsf{negl}'(n)$ instead of being unity. This intuition is not right in general, as by definition a negligible function only needs to be super-polynomially small for all sufficiently large (instead of all) $n$'s. However, it is reasonable to believe that one-way functions satisfying (9) should be quite artificial.

(9) is not sufficient. Despite seeming strong, (9) is still not sufficient to make a counterexample. To show this, we give an example function that satisfies both (9) (for every $n_0 \in \mathbb{N} \cup \{0\}$) and Definition 4.1. That is, let $f$ be a one-way function where for every $n$ the non-empty sets of $f$ are

$$\mathcal{Y}_{n/3}, \mathcal{Y}_{n/3+1}, \ldots, \mathcal{Y}_{n/2} \tag{10}$$

with $\Pr[f(U_n) \in \mathcal{Y}_{n/3}] = 1 - n^{-\log n + 1}/6$, $\Pr[f(U_n) \in \mathcal{Y}_{n/3+i}] = n^{-\log n}$ for all $1 \le i \le n/6$ and thus $\kappa(n) = n/6 + 1$. It is easy to see that this function satisfies Definition 4.1 with $\max(n) = n/3$ and $\epsilon(n) = n^{-\log n + 1}/6$. In addition, for every $n_0 \in \mathbb{N} \cup \{0\}$ function $\mu_{n_0}(\cdot)$ is negligible as $\mu_{n_0}(n) = n^{-\log n}$ for all $n > 6n_0$. In summary, although an arbitrary one-way function may not fall into the class of functions given in Definition 4.1, the counterexamples must be well crafted to satisfy a very strong (yet still insufficient) condition. We leave it as an open question on finding out such a counterexample and believe that our construction should be able to deal with most natural one-way functions.

[7] Although non-negligible and noticeable are not the same, they are quite close: a non-negligible (resp., noticeable) function $\mu(\cdot)$ satisfies that there exists constant $c$ such that $\mu(n) \ge n^{-c}$ for infinitely many (resp., all large enough) $n$'s.

## 5 A More General Construction of Pseudorandom Generators

### 5.1 Overview and Definitions

In this section we construct a pseudorandom generator with seed length $O(n \log n)$ from the class of one-way functions as in Definition 4.1. We first show how to construct the PRG by running the iterate $\tilde{O}(n^{2c+1})$ times, and thus require large amount of randomness (i.e., $\tilde{O}(n^{2c+2})$ bits) to sample the hash functions. Then, we show the derandomized version where the amount of the randomness is compressed into $O(n \log n)$ bits using bounded space generators.

### 5.2 The Randomized Version: A PRG with Seed Length $\tilde{O}(n^{2c+2})$

Recall that any one-way function $f$ can be assumed to be length-preserving without loss of generality. Further, we also assume that conditioned on $f(U_n) \in \mathcal{Y}_{\max}$, $f(U_n)$ is flat over $\mathcal{Y}_{\max}$, i.e., $\forall y \in \mathcal{Y}_{\max}$ satisfies $\Pr[y = f(U_n)] = 2^{\max - n - 1}$ rather than lying in the small interval of $[2^{\max - n-1},\, 2^{\max - n})$.

**Theorem 5.1 (the randomized version)** *For $n, k \in \mathbb{N}$ and constant $c$, let $f$, $\mathcal{H}$ and $f^k$ be as in Definition 3.1, assume that $f$ is an arbitrary (length-preserving) one-way function with $n^{-c}$-fraction of maximal-size preimages, and let $\mathcal{H}_c = \{h_c : \{0,1\}^n \to \{0,1\}^{2\log n}\}$ be a family of Goldreich-Levin hardcore functions. Then, for any efficient $\alpha(n) \in \omega(1)$, $\Delta(n) = \alpha(n) \cdot \log n \cdot n^{2c}$ and $r(n) = \lceil n / \log n \rceil$, the function $g$:$\{0,1\}^n \times \mathcal{H}^{r(n)\cdot\Delta(n)-1} \times \mathcal{H}_c \to \{0,1\}^{2n} \times \mathcal{H}^{r(n)\cdot\Delta(n)-1} \times \mathcal{H}_c$ defined as*

$$g(x_1, \vec{h}^{r\cdot\Delta-1}, h_c) = (\ h_c(x_1), h_c(x_{1+\Delta}), h_c(x_{1+2\Delta}), \ldots,\ h_c(x_{1+r\cdot\Delta}), \vec{h}^{r\cdot\Delta-1}, h_c\ ) \tag{11}$$

*is a pseudorandom generator.*
*Notice that a desirable property about the technique is that a construction assuming a sufficiently large $c$ works with any one-way function whose actual parameter is less than or equal to $c$.*

*Proof.* The proof is similar to Theorem 3.1 based on Yao's hybrid argument [15]. Namely, the pseudorandomness of a sequence (with polynomially many blocks) is equivalent to that every block is pseudorandom conditioned on its suffix (or prefix). By the Goldreich-Levin Theorem and Lemma 5.1 below we know that every $h_c(x_{1+j\Delta})$ is pseudorandom conditioned on $h_c$, $y_{(j+1)\Delta}$ and $\vec{h}^{r\Delta-1}$, which efficiently implies all subsequent blocks $h_c(x_{1+(j+1)\Delta})$, …, $h_c(x_{1+r\Delta})$. This completes the proof. $\square$

**Lemma 5.1 (every $\Delta(n)$ iterations are hard-to-invert)** *For $n, k \in \mathbb{N}$ and constant $c$, let $f$, $\mathcal{H}$, $f^k$, $\alpha = \alpha(n)$, $\Delta = \Delta(n)$ and $r = r(n)$ be as defined in Theorem 5.1. More specifically, assume that $f$ is a $(T(n), \varepsilon(n))$-one-way function with $n^{-c}$-fraction of maximal-size preimages. Then, for every $j \in [r]$, and for every PPT $\mathsf{A}$ of running time $T(n) - n^{O(1)}$ (for some universal constant $O(1)$) it holds that*

$$\Pr_{x_1 \xleftarrow{\$} \{0,1\}^n,\ \vec{h}^{r\Delta-1} \xleftarrow{\$} \mathcal{H}^{r\Delta-1}} [\ \mathsf{A}(y_{j\cdot\Delta},\ \vec{h}^{r\Delta-1}) = x_{1+(j-1)\Delta}\ ] \ \in\ O(\ n^c \cdot r \cdot \Delta^2 \cdot \sqrt{\varepsilon(n)}\ )\ . \tag{12}$$

Proof sketch of Lemma 5.1 . Assume towards a contradiction that

$$\exists j^* \in [r], \exists \text{ PPT } \mathsf{A}:\ \Pr[\ \mathsf{A}(Y_{j^*\cdot\Delta},\ \vec{H}^{r\Delta-1}) = X_{1+(j^*-1)\Delta}\ ]\ \geq\ \varepsilon_{\mathsf{A}}(n) \tag{13}$$

for some non-negligible function $\varepsilon_{\mathsf{A}}(\cdot)$. Then, we build an efficient algorithm $\mathsf{M}^{\mathsf{A}}$ that invokes $\mathsf{A}$ (as in Algorithm 1) and inverts $f$ with probablity $\Omega(\varepsilon_{\mathsf{A}}^2 / n^{2c} \cdot r^2 \cdot \Delta^4)$ (as shown in Lemma 5.3), which is a contradiction to the $(T, \varepsilon)$-one-wayness of $f$ and thus completes the proof.

The proof presented here is similar to the hardness amplification of regular weakly one-way function [8], but ours is more involved (even though Lemma 3.1 already simplifies the key ingredients). We define the events $\mathcal{E}_k$ and $\mathcal{S}_k$ as in Definition 5.1, where $\mathcal{S}_k$ refers to that during the first $k$ iterates no $y_t$ ($1 \leq t \leq k$) hits the negligible fraction region (see Remark B.1 in Appendix B for the underlying intuitions), and $\mathcal{E}_k$ defines the desirable event that $y_k$ hits $\mathcal{Y}_{\max}$ (which implies the hard-to-invertness).

**Definition 5.1 (events $\mathcal{S}_k$ and $\mathcal{E}_k$)** *For any $n \in \mathbb{N}$, for any $k \le r\Delta$, define events*

$$\mathcal{S}_k \stackrel{\mathsf{def}}{=} \left( (X_1, \vec{H}^{r\Delta-1}) \in \{\ (x_1, \vec{h}^{r\Delta-1}) : \forall t \in [k] \text{ satisfies } y_t \in \mathcal{Y}_{[\max]},\ \text{ where } y_t = f^t(x_1, \vec{h}^{r\Delta-1})\ \} \right)$$

$$\mathcal{E}_k \stackrel{\mathsf{def}}{=} \left( (X_1, \vec{H}^{r\Delta-1}) \in \{\ (x_1, \vec{h}^{r\Delta-1}) : y_k \in \mathcal{Y}_{\max}\ ,\ \text{ where } y_k = f^k(x_1, \vec{h}^{r\Delta-1})\ \} \right)$$

*where $(X_1, \vec{H}^{r\Delta-1})$ is uniform distribution over $\{0,1\}^n \times \mathcal{H}^{r\Delta-1}$. We also naturally extend the definition of collision probability conditioned on $\mathcal{E}_k$ and $\mathcal{S}_k$. For example,*

$$\mathsf{CP}(\ Y_k \wedge \mathcal{E}_k \wedge \mathcal{S}_k \mid \vec{H}^{r\Delta-1}) \stackrel{\mathsf{def}}{=} \mathbb{E}_{\vec{h}^{r\Delta-1} \leftarrow \vec{H}^{r\Delta-1}} \left[ \sum_y \Pr[f^k(X_1, \vec{H}^{r\Delta-1}) = y \wedge \mathcal{E}_k \wedge \mathcal{S}_k \mid \vec{H}^{r\Delta-1} = \vec{h}^{r\Delta-1}]^2 \right]$$

$$\mathsf{CP}(\ Y_k, \vec{H}^{r\Delta-1} \mid \mathcal{E}_k \wedge \mathcal{S}_k\ ) \stackrel{\mathsf{def}}{=} \sum_{(y, \vec{h}^{r\Delta-1})} \Pr[\ (f^k(X_1, \vec{H}^{r\Delta-1}), \vec{H}^{r\Delta-1}) = (y, \vec{h}^{r\Delta-1}) \mid \mathcal{E}_k \wedge \mathcal{S}_k\ ]^2\ .$$

**Claim 5.1** *For any $n \in \mathbb{N}$, and let $\mathcal{S}_k$ and $\mathcal{E}_k$ be as defined in Definition 5.1, assume that $f$ has an $n^{-c}$-fraction of maximal-size preimages (with $\epsilon$ and* max *defined as in (7) and (8)). Then, it holds that*

$$\forall k \in [r\Delta] : \Pr[\mathcal{S}_k] \ \ge\ (1-\epsilon)^k \ \approx\ 1 - k\epsilon\ , \qquad \Pr[\mathcal{E}_k] \ \ge\ n^{-c}\ , \qquad \Pr[\mathcal{E}_k \wedge \mathcal{S}_k] \ge n^{-c}/2 \tag{14}$$

$$\forall k \in \mathbb{N} : \Pr[\mathcal{E}_{k+1} \vee \mathcal{E}_{k+2} \vee \ldots \vee \mathcal{E}_{k+\Delta}] \ \ge\ 1 - \exp^{\Delta/n^{2c}} \ \ge\ 1 - n^{-\alpha} \tag{15}$$

$$\forall k \in [r\Delta] : \mathsf{CP}(\ Y_k \ \wedge\ \mathcal{E}_k \wedge \mathcal{S}_k \mid \vec{H}^{r\Delta-1}) \ \le\ r\Delta{\cdot}2^{\max-n+1}\ ,\ \text{ where } Y_k = f^k(X_1, \vec{H}^{r\Delta-1})\ . \tag{16}$$

*Proof.* We have that $x_1$, $x_2 = h_1(y_1)$, ..., $x_{r\Delta} = h_{r\Delta-1}(y_{r\Delta-1})$ are all i.i.d. to $U_n$ due to the universality of $\mathcal{H}$. This implies that $\Pr[y_i \in \mathcal{Y}_{[\max]}] \ge 1-\epsilon$ for every $i \in [k]$ independently, and that $\mathcal{E}_1$, ... and $\mathcal{E}_{r\Delta}$ are i.i.d. events with probability at least $n^{-c}$. The former further implies $\Pr[\mathcal{S}_k] \ \ge\ (1-\epsilon)^k$ whose approximation is given by $1 - k \cdot \epsilon$ for any $k = \mathsf{poly}(n)$ and $\epsilon = \mathsf{negl}(n)$. Thus, we complete the proof for (14) by

$$\Pr[\mathcal{E}_k \wedge \mathcal{S}_k] \ \ge\ \Pr[\mathcal{E}_k] \ -\ \Pr[\neg\, \mathcal{S}_k] \ge\ n^{-c} \ -\ k \cdot \epsilon \ \ge\ n^{-c}/2\ .$$

For every $k \in \mathbb{N}$, $i \in [\Delta]$, define $\zeta_{k+i} = 1$ iff $\mathcal{E}_{k+i}$ occurs (and $\zeta_{k+i} = 0$ otherwise). It follows by a Chernoff-Hoeffding bound that

$$\forall k \in \mathbb{N} : \Pr[\ (\neg\mathcal{E}_{k+1}) \wedge \ldots \wedge (\neg\mathcal{E}_{k+\Delta})\ ] \ =\ \Pr[\ \sum_{i=1}^{\Delta} \zeta_{k+i} = 0\ ] \ \le\ \exp^{-\Delta/n^{2c}} \ \le\ n^{-\alpha}$$

which yields (15) by taking a negation. Regarding (16), consider two instances of the random iterate seeded with independent $x_1$ and $x_1'$ and a common random $\vec{h}^{r\Delta-1}$, the collision probability is upper bounded by the sum of events that the first collision occurs on points $y_1$, $y_2$, ..., $y_k \in \mathcal{Y}_{[\max]}$ respectively, where for upper bound we omit other constraints such as $\mathcal{E}_k$ and that the route after the first collision should also hit every $\mathcal{Y}_{[\max]}$ (as required by $\mathcal{S}_k$). We thus have by the pairwise independence of $\mathcal{H}$ that

$$\begin{aligned}
&\mathsf{CP}(\ Y_k \ \wedge\ \mathcal{E}_k \wedge \mathcal{S}_k \mid \vec{H}^{r\Delta-1}) \ \le\ \mathsf{CP}(\ Y_k \ \wedge\ \mathcal{S}_k \mid \vec{H}^{r\Delta-1}) \\
\le\ & \Pr_{x_1, x_1' \xleftarrow{\$} \{0,1\}^n} [f(x_1) = f(x_1') \in \mathcal{Y}_{[\max]}] \ +\ \sum_{t=2}^{k} \left( \Pr_{y_{t-1} \neq y'_{t-1},\ h_{t-1} \xleftarrow{\$} \mathcal{H}} [f(x_t) = f(x_t') \in \mathcal{Y}_{[\max]}] \right) \\
\le\ & r\Delta \sum_{y \in \mathcal{Y}_{[\max]}} \Pr[f(U_n) = y]^2 \ \le\ r\Delta \sum_{i=1}^{\max} \sum_{y \in \mathcal{Y}_i} \Pr[f(U_n) = y]{\cdot}2^{i-n} \ =\ r\Delta \sum_{i=1}^{\max} \Pr[f(U_n) \in \mathcal{Y}_i]{\cdot}2^{i-n} \\
\le\ & r\Delta{\cdot}2^{\max-n}(1 + 2^{-1} + \ldots + 2^{1-\max}) \ \le\ r\Delta{\cdot}2^{\max-n+1}\ .
\end{aligned}$$

□

**Lemma 5.2** *For any $n \in \mathbb{N}$, with the same assumptions and notations as in Theorem 5.1, Definition 4.1 and Definition 5.1, and let $j^* \in [r]$, $\mathsf{A}$, $\varepsilon_\mathsf{A}$ be as assumed in (13). Then, there exists $i^* \in [\Delta]$ such that*

$$\Pr[\ \mathsf{A}(Y_{j^*\cdot\Delta},\ \vec{H}^{r\Delta-1}) = X_{1+(j^*-1)\Delta}\ \wedge\ \mathcal{E}_{(j^*-1)\Delta+i^*} \wedge \mathcal{S}_{(j^*-1)\Delta+i^*}\ ]\ \geq\ \varepsilon_\mathsf{A}/2\Delta\ . \tag{17}$$

*Proof.* For notational convenience use shorthand $\mathcal{C}$ for the event $\mathsf{A}(Y_{j^*\cdot\Delta},\ \vec{H}^{r\Delta-1}) = X_{1+(j^*-1)\Delta}$. Then,

$$\sum_{i=1}^{\Delta}\Pr[\mathcal{C}\wedge\mathcal{E}_{(j^*-1)\Delta+i}\wedge\mathcal{S}_{(j^*-1)\Delta+i}]\ \geq\ \sum_{i=1}^{\Delta}\Pr[\mathcal{C}\wedge\mathcal{E}_{(j^*-1)\Delta+i}\wedge\mathcal{S}_{r\Delta}]\ \geq\ \Pr[\ \mathcal{C}\wedge\mathcal{S}_{r\Delta}\wedge(\bigvee_{i=1}^{\Delta}\mathcal{E}_{(j^*-1)\Delta+i})\ ]$$

$$\geq\ \Pr[\ \mathcal{C}]\ -\ \Pr[\ \neg\mathcal{S}_{r\Delta}]\ -\ \Pr[\neg(\bigvee_{i=1}^{\Delta}\mathcal{E}_{(j^*-1)\Delta+i})]\ \geq\ \varepsilon_\mathsf{A}\ -\ r\Delta\cdot\epsilon\ -\ n^{-\alpha}\ \geq\ \varepsilon_\mathsf{A}/2\ ,$$

where we use (14) and (15) in the fourth inequality above, and recall that $\epsilon$ and $n^{-\alpha}$ are both negligible in $n$. Thus, such an $i^*$ (which satisfies (17)) exists by an averaging argument. □

THE INTUITION FOR $\mathsf{M}^\mathsf{A}$. We know by Lemma 5.2 that there exist some $i^*$ and $j^*$ conditioned on which $\mathsf{A}$ inverts the iterate with non-negligible probability. If we knew which $i^*$ and $j^*$, then we simply replace $y_{(j^*-1)\Delta+i^*}$ with $f(U_n)$, simulate the iterate for the rest iterations and invoke $\mathsf{A}$ to invert $f$. Although the distribution after the replacement will not be identical to the original one, we use Lemma 3.1 to argue that the collision entropy deficiency is small enough and thus the inverting probability will not blow up by more than a polynomial factor. However, we actually do not know the values of $i^*$ and $j^*$, so we need to randomly sample $i$ and $j$ over $[\Delta]$, $[r]$ respectively. This yields the algorithm below.

---

**Algorithm 1** $\mathsf{M}^\mathsf{A}$.

**Input:** $y \in \{0,1\}^n$

Sample $j \xleftarrow{\$} [r]$, $i \xleftarrow{\$} [\Delta]$, $\vec{h}^{r\Delta-1} \xleftarrow{\$} \mathcal{H}^{r\Delta-1}$;
Let $\tilde{y}_{(j-1)\Delta+i} := y$ ;
FOR $k = (j-1)\Delta+i+1$ TO $(j-1)\Delta+\Delta$
  Compute $\tilde{x}_k := h_{k-1}(\tilde{y}_{k-1})$, $\tilde{y}_k := f(\tilde{x}_k)$;
$\tilde{x}_{(j-1)\Delta+1} \leftarrow \mathsf{A}(\tilde{y}_{j\Delta}, \vec{h}^{r\Delta-1})$;
FOR $k = (j-1)\Delta+1$ TO $(j-1)\Delta+i-1$
  Compute $\tilde{y}_k := f(\tilde{x}_k)$, $\tilde{x}_{k+1} := h_k(\tilde{y}_k)$ ;

**Output:** $\tilde{x}_{(j-1)\Delta+i}$

---

**Lemma 5.3** **($\mathsf{M}^\mathsf{A}$ inverts $f$)** *For any $n \in \mathbb{N}$, let $\mathsf{A}$ be as assumed in Lemma 5.2 and let $\mathsf{M}^\mathsf{A}$ be as defined in Algorithm 1. Then, it holds that*

$$\Pr_{y\leftarrow f(U_n);\ j\xleftarrow{\$}[r];\ i\xleftarrow{\$}[\Delta];\ \vec{h}^{r\Delta-1}\xleftarrow{\$}\mathcal{H}^{r\Delta-1}}[\ \mathsf{M}^\mathsf{A}(y;\ j,i,\ \vec{h}^{r\Delta-1}) \in f^{-1}(y)\ ]\ \geq\ \frac{\varepsilon_\mathsf{A}^2}{2^8\cdot n^{2c}\cdot r^2\cdot\Delta^4}\ .$$

*Proof.* We know by Lemma 5.2 that there exist $j^* \in [r]$ and $i^* \in [\Delta]$ satisfying (17), which implies

$$\begin{aligned}
&\Pr\ [\ \mathsf{M}^\mathsf{A}(Y_{(j-1)\Delta+i};\ j,\ i,\ \vec{H}^{r\Delta-1}) \in f^{-1}(Y_{(j-1)\Delta+i})\mid(j,i)=(j^*,i^*)\wedge\mathcal{E}_{(j-1)\Delta+i}\wedge\mathcal{S}_{(j-1)\Delta+i}\ ]\\
\geq\ &\Pr[\ \mathsf{A}(Y_{j^*\cdot\Delta},\ \vec{H}^{r\Delta-1}) = X_{1+(j^*-1)\Delta}\mid\mathcal{E}_{(j^*-1)\Delta+i^*}\wedge\mathcal{S}_{(j^*-1)\Delta+i^*}\ ]\\
\geq\ &\Pr[\ \mathsf{A}(Y_{j^*\cdot\Delta},\ \vec{H}^{r\Delta-1}) = X_{1+(j^*-1)\Delta}\ \wedge\ \mathcal{E}_{(j^*-1)\Delta+i^*}\wedge\mathcal{S}_{(j^*-1)\Delta+i^*}\ ]\ \geq\ \varepsilon_\mathsf{A}/2\Delta\ \ ,
\end{aligned}$$

where the second inequality, in abstract form, is $\Pr[\mathcal{E}_a|\mathcal{E}_b]\ \geq\ \Pr[\mathcal{E}_a|\mathcal{E}_b]\cdot\Pr[\mathcal{E}_b]\ =\ \Pr[\mathcal{E}_a\wedge\mathcal{E}_b]$. The above is not exactly what we need as conditioned on $\mathcal{E}_{(j^*-1)\Delta+i^*}\wedge\mathcal{S}_{(j^*-1)\Delta+i^*}$, the random variable

$(Y_{(j^*-1)\Delta+i^*}, \vec{H}^{r\Delta-1})$ is not uniform over $\mathcal{Y}_{\max} \times \mathcal{H}^{r\Delta-1}$. However, we show below that it has nearly full collision entropy over $\mathcal{Y}_{\max} \times \mathcal{H}^{r\Delta-1}$

$$
\begin{aligned}
& \mathsf{CP}(\ (Y_{(j^*-1)\Delta+i^*}, \vec{H}^{r\Delta-1}) \mid \mathcal{E}_{(j^*-1)\Delta+i^*} \wedge \mathcal{S}_{(j^*-1)\Delta+i^*}\ ) \\
= \quad & \mathsf{CP}(\ (Y_{(j^*-1)\Delta+i^*}, \vec{H}^{r\Delta-1})\ \wedge\ \mathcal{E}_{(j^*-1)\Delta+i^*} \wedge \mathcal{S}_{(j^*-1)\Delta+i^*})\ /\ \Pr[\mathcal{E}_{(j^*-1)\Delta+i^*} \wedge \mathcal{S}_{(j^*-1)\Delta+i^*}]^2 \\
\leq \quad & \mathsf{CP}(\ Y_{(j^*-1)\Delta+i^*}\ \wedge\ \mathcal{E}_{(j^*-1)\Delta+i^*}\ \wedge \mathcal{S}_{(j^*-1)\Delta+i^*} \mid \vec{H}^{r\Delta-1})\ \frac{1}{(n^{-2c}/4)\cdot|\mathcal{H}|^{r\Delta-1}} \\
\leq \quad & \frac{r\Delta\cdot 2^{\max-n+1}}{(n^{-2c}/4)\cdot|\mathcal{H}|^{r\Delta-1}} \ = \ \frac{8r\Delta\cdot n^{2c}}{2^{n-\max}\cdot|\mathcal{H}|^{r\Delta-1}}\ ,
\end{aligned}
$$

where equalities follow from Fact A.1 in Appendix A and the two inequalities are by (14) and (16) respectively. Taking a logarithm, we get

$$
\mathbf{H}_2(\ (Y_{(j-1)\Delta+i^*}, \vec{H}^{r\Delta-1}) \mid \mathcal{E}_{(j-1)\Delta+i^*} \wedge \mathcal{S}_{(j-1)\Delta+i^*}\ ) \geq \Big(n - \max + (r\Delta-1)\log|\mathcal{H}| - c\cdot \log n + 1\Big) - e\ ,
$$

where entropy deficiency $e \leq c\cdot\log n + \log r + \log\Delta + 4$. Note that conditioned on $f(U_n) \in \mathcal{Y}_{\max}$ the distribution $(f(U_n), \vec{H}^{r\Delta-1})$ is uniform over $\mathcal{Y}_{\max} \times \mathcal{H}^{r\Delta-1}$ with full entropy

$$
\mathbf{H}_2(\ (f(U_n), \vec{H}^{r\Delta-1}) \mid f(U_n) \in \mathcal{Y}_{\max}\ ) = \log(\frac{n^{-c}}{2^{-n+\max-1}}\cdot|\vec{H}|^{r\Delta-1}) = n - \max + (r\Delta-1)\log|\mathcal{H}| - c\cdot\log n + 1\ .
$$

To apply Lemma 3.1, let $\mathcal{W} = \mathcal{Y}_{\max} \times \mathcal{H}^{r\Delta-1}$, $\mathcal{Z} = \emptyset$, let $W$ be $(Y_{(j^*-1)\Delta+i^*}, \vec{H}^{r\Delta-1})$ conditioned on $\mathcal{E}_{(j^*-1)\Delta+i^*}$ and $\mathcal{S}_{(j^*-1)\Delta+i^*}$, and define

$$
\mathsf{Adv}(y, \vec{h}^{r\Delta-1}) \stackrel{\text{def}}{=} \begin{cases} 1, & \text{if } \mathsf{M}^{\mathsf{A}}(y;\ j^*, i^*,\ \vec{h}^{r\Delta-1}) \in f^{-1}(y) \\ 0, & \text{if } \mathsf{M}^{\mathsf{A}}(y;\ j^*, i^*,\ \vec{h}^{r\Delta-1}) \notin f^{-1}(y) \end{cases}
$$

Let $\mathcal{C}_{j^*i^*\max}$ denote the event that $(j,i) = (j^*, i^*) \wedge f(U_n) \in \mathcal{Y}_{\max}$, and we thus have

$$
\begin{aligned}
& \Pr[\mathsf{M}^{\mathsf{A}}(f(U_n);\ j, i,\ \vec{H}^{r\Delta-1}) \in f^{-1}(f(U_n))\ ] \\
\geq\ & \Pr[\mathcal{C}_{j^*i^*\max}]\cdot \Pr[\mathsf{M}^{\mathsf{A}}(f(U_n);\ j, i,\ \vec{H}^{r\Delta-1}) \in f^{-1}(f(U_n)) \mid \mathcal{C}_{j^*i^*\max}\ ] \\
\geq\ & (1/r\Delta n^c)\cdot\mathbb{E}[\ \mathsf{Adv}(U_{\mathcal{Y}_{\max}}, \vec{H}^{r\Delta-1})\ ] \\
\geq\ & (1/r\Delta n^c)\cdot\frac{\mathbb{E}[\ \mathsf{Adv}(Y_{(j^*-1)\Delta+i^*}, \vec{H}^{r\Delta-1}) \mid \mathcal{E}_{(j^*-1)\Delta+i^*} \wedge \mathcal{S}_{(j^*-1)\Delta+i^*}\ ]^2}{2^{e+2}} \\
\geq\ & (1/r\Delta n^c)\cdot\frac{\varepsilon_{\mathsf{A}}^2/4\Delta^2}{2^6\cdot n^c r\cdot\Delta} \ = \ \frac{\varepsilon_{\mathsf{A}}^2}{2^8\cdot n^{2c}\cdot r^2\cdot\Delta^4}\ ,
\end{aligned}
$$

where we apply Lemma 3.1 to complete the proof.

□

### 5.3 The Derandomized Version: A PRG with Seed Length $O(n\cdot\log\ n)$

The derandomized version uses a bounded-space generator to expand a $\Theta(n\cdot\log n)$-bit $u$ into a long string over $\mathcal{H}^{r\Delta-1}$ (rather than sampling a random element over it).

**Theorem 5.2 (the derandomized version)** *For $n, k \in \mathbb{N}$ and constant $c$, let $f$, $\mathcal{H}$, $\mathcal{H}_c$, $f^k$, $\alpha(n)$, $\Delta(n)$ and $r(n)$ be as assumed in Theorem 5.1, let $g$ be as defined in (11), let*

$$
BSG : \{0,1\}^{q(n)\in\Theta(n\cdot\log n)} \rightarrow \{0,1\}^{(\alpha(n)\cdot n^{2c+1}-1)\cdot\log|\mathcal{H}|}
$$

*be a bounded-space generator that* $2^{-2n}$*-fools every* $(2n+1, (\alpha(n)n^{2c+1}), \log|\mathcal{H}|)$*-LBP (see Footnote 5). Then, the function* $g' : \{0,1\}^n \times \{0,1\}^{q(n)} \times \mathcal{H}_c \to \{0,1\}^{2n} \times \{0,1\}^{q(n)} \times \mathcal{H}_c$ *defined as*

$$g'(x_1, u, h_c) = g(x_1, BSG(u), h_c) \tag{18}$$

*is a pseudorandom generator.*

Similar to the randomized version, it suffices to show Lemma 5.4 (the counterpart of Lemma 5.1).

**Lemma 5.4** *For the same assumptions as stated in Lemma 5.1, we have that for every* $j \in [r]$*, and for every PPT* $\mathsf{A}'$ *of running time* $T(n) - n^{O(1)}$ *(for some universal constant* $O(1)$*) it holds that*

$$\Pr_{x_1 \xleftarrow{\$} \{0,1\}^n,\ u \xleftarrow{\$} \{0,1\}^q,\ \vec{h}^{r\Delta-1} := BSG(u)} [\ \mathsf{A}'(y_{j\cdot\Delta},\ u) = x_{1+(j-1)\Delta}\ ] \ \in\ O(\ n^c \cdot r \cdot \Delta^2 \cdot \sqrt{\varepsilon(n)}\ )\ . \tag{19}$$

The proof of Lemma 5.4 follows the steps of the proof of Lemma 5.1. We define events $\mathcal{S}'_k$ and $\mathcal{E}'_k$ in Definition 5.2 (the analogues of $\mathcal{S}_k$ and $\mathcal{E}_k$). Despite that all the events (e.g., $\mathcal{E}'_1, \ldots, \mathcal{E}'_k$) are not independent due to short of randomness, we still have (20), (21) and (22) below. We defer their proofs to Appendix A due to lack of space, where for every inequality we define an LBP and argue that the advantage of the LBP on $\vec{H}^{r\Delta-1}$ and $BSG(U_q)$ is bounded by $2^{-2n}$ and thus (20), (21) and (22) follow from their respective counterparts (14), (15) and (16) by adding an additive term $2^{-2n}$.

**Definition 5.2 (events $\mathcal{S}'_k$ and $\mathcal{E}'_k$)** *For any* $n \in \mathbb{N}$*, for any* $k \le r\Delta$*, define events*

$$\mathcal{S}'_k \ \stackrel{\mathsf{def}}{=}\ \Big( (X_1, U_q) \in \{\ (x_1, u) : \forall t \in [k] \text{ satisfies } y'_t \in \mathcal{Y}_{[\max]},\ \text{ where } y'_t = f^t(x_1, BSG(u))\ \} \Big)$$

$$\mathcal{E}'_k \ \stackrel{\mathsf{def}}{=}\ \Big( (X_1, U_q) \in \{\ (x_1, u) : y'_k \in \mathcal{Y}_{\max}\ ,\ \text{ where } y'_k = f^k(x_1, BSG(u))\ \} \Big)$$

*where* $(X_1, U_q)$ *is uniform distribution over* $\{0,1\}^n \times \{0,1\}^q$*. We refer to Definition B.1 in Appendix B for the definitions of the collision probabilities in the following proofs.*

$$\forall k \in [r\Delta] : \Pr[\mathcal{S}'_k]\ \ge\ 1 - k\epsilon - 2^{-2n}\ , \qquad \Pr[\mathcal{E}'_k]\ \ge\ n^{-c} - 2^{-2n}\ , \qquad \Pr[\mathcal{E}'_k \wedge \mathcal{S}'_k] \ge n^{-c}/2 \tag{20}$$

$$\forall k \in [(r-1)\Delta] : \Pr[\mathcal{E}'_{k+1} \vee \mathcal{E}'_{k+2} \vee \ldots \vee \mathcal{E}'_{k+\Delta}]\ \ge\ 1 - n^{-\alpha} - 2^{-2n} \tag{21}$$

$$\forall k \in [r\Delta] : \mathsf{CP}(\ Y'_k\ \wedge\ \mathcal{E}'_k \wedge \mathcal{S}'_k \mid U_q)\ \le\ (r\Delta+1)\cdot 2^{\max - n + 1}, \ \text{ where }\ Y'_k = f^k(X_1, BSG(U_q)) \tag{22}$$

Proof sketch of Lemma 5.4 . Assume towards a contradiction that for some non-negligible $\varepsilon_{\mathsf{A}'}(\cdot)$ that

$$\exists j^* \in [r], \exists \text{ PPT } \mathsf{A}' :\ \Pr[\ \mathsf{A}'(Y'_{j^*\cdot\Delta},\ U_q) = X'_{1+(j^*-1)\Delta}\ ]\ \ge\ \varepsilon_{\mathsf{A}'}(n) \tag{23}$$

where for $k \in [r\Delta]$ we use notations $\vec{H}'^{r\Delta-1} = BSG(U_q)$, $Y'_k = f^k(X_1, \vec{H}'^{r\Delta-1})$ and $X'_{k+1} = H'_k(Y'_k)$. Then, we define $\mathsf{M}^{\mathsf{A}'}$ that inverts $f$ with the following probability. Since $\mathsf{M}^{\mathsf{A}'}$ is quite similar to its analogue $\mathsf{M}^{\mathsf{A}}$ we state it as Algorithm 2 in Appendix B.

$$\Pr_{y \leftarrow f(U_n);\ j \xleftarrow{\$} [r];\ i \xleftarrow{\$} [\Delta];\ u \xleftarrow{\$} \{0,1\}^q} [\ \mathsf{M}^{\mathsf{A}'}(y;\ j, i,\ u) \in f^{-1}(y)\ ]\ \in\ \Omega(\frac{\varepsilon^2_{\mathsf{A}'}}{n^{2c}\cdot r^2 \cdot \Delta^4})\ , \tag{24}$$

which is a contradiction to the one-wayness of $f$ and thus concludes Lemma 5.4.

Proof sketch of (24). Denote by $\mathcal{C}'$ the event $\mathsf{A}(Y'_{j^*\cdot\Delta},\ U_q) = X'_{1+(j^*-1)\Delta}$. Then,

$$\sum_{i=1}^{\Delta}\Pr[\mathcal{C}'\wedge\mathcal{E}'_{(j^*-1)\Delta+i}\wedge\mathcal{S}'_{(j^*-1)\Delta+i}] \ \geq\ \sum_{i=1}^{\Delta}\Pr[\mathcal{C}'\wedge\mathcal{E}'_{(j^*-1)\Delta+i}\wedge\mathcal{S}'_{r\Delta}] \ \geq\ \Pr[\ \mathcal{C}'\wedge\mathcal{S}'_{r\Delta}\wedge(\ \bigvee_{i=1}^{\Delta}\mathcal{E}'_{(j^*-1)\Delta+i})\ ]$$

$$\geq\ \Pr[\ \mathcal{C}'] \ -\ \Pr[\ \neg\mathcal{S}'_{r\Delta}] \ -\ \Pr[\neg(\bigvee_{i=1}^{\Delta}\mathcal{E}'_{(j^*-1)\Delta+i})] \ \geq\ \varepsilon_{\mathsf{A}'} \ -\ r\Delta\cdot\epsilon \ -\ n^{-\alpha} \ -\ 2^{-2n+1} \ \geq\ \varepsilon_{\mathsf{A}'}/2\ ,$$

where we use (20) and (21) in the fourth inequality. Thus, by averaging we have that

$$\exists j^*\in[r],\exists i^*\in[\Delta],\exists\ \text{PPT}\ \mathsf{A}':\ \ \Pr[\ \mathsf{A}'(Y'_{j^*\cdot\Delta},\ U_q) = X'_{1+(j^*-1)\Delta}\ ]\ \geq\ \varepsilon_{\mathsf{A}'}/2\Delta\ \ .$$

The proofs below follow the steps of Lemma 5.3. We have that (proof of (25) given in Appendix A)

$$\mathbf{H}_2(\ (Y'_{(j-1)\Delta+i^*},U_q\ )\mid\mathcal{E}'_{(j-1)\Delta+i^*}\wedge\mathcal{S}'_{(j-1)\Delta+i^*}\ )\geq\ \mathbf{H}_2(\ f(U_n),U_q\mid f(U_n)\in\mathcal{Y}_{\max}\ )\ -\ e\ , \tag{25}$$

where entropy deficiency $e\leq c\cdot\log n+\log r+\log\Delta+5$. Finally, let $\mathcal{W}=\mathcal{Y}_{\max}\times\{0,1\}^q$, $\mathcal{Z}=\emptyset$, let $W$ be $(Y'_{(j^*-1)\Delta+i^*},U_q)$ conditioned on $\mathcal{E}'_{(j^*-1)\Delta+i^*}$ and $\mathcal{S}'_{(j^*-1)\Delta+i^*}$, and define

$$\mathsf{Adv}(y,u)\stackrel{\mathsf{def}}{=}\begin{cases}1, & \text{if}\ \mathsf{M}^{\mathsf{A}'}(y;\ j^*,i^*,\ u)\in f^{-1}(y)\\ 0, & \text{if}\ \mathsf{M}^{\mathsf{A}'}(y;\ j^*,i^*,\ u)\notin f^{-1}(y)\end{cases}$$

Let $\mathcal{C}_{j^*i^*\max}$ denote the event that $(j,i)=(j^*,i^*)\wedge f(U_n)\in\mathcal{Y}_{\max}$, and we thus have

$$\begin{aligned}
&\Pr[\mathsf{M}^{\mathsf{A}'}(f(U_n);\ j,i,U_q\ )\in f^{-1}(f(U_n))\ ]\\
\geq\ &\Pr[\mathcal{C}_{j^*i^*\max}]\cdot\Pr[\mathsf{M}^{\mathsf{A}'}(f(U_n);\ j,i,U_q\ )\in f^{-1}(f(U_n))\mid\mathcal{C}_{j^*i^*\max}\ ]\\
\geq\ &(1/r\Delta n^c)\cdot\mathbb{E}[\ \mathsf{Adv}(U_{\mathcal{Y}_{\max}},U_q)\ ]\\
\geq\ &(1/r\Delta n^c)\cdot\frac{\mathbb{E}[\ \mathsf{Adv}(Y'_{(j^*-1)\Delta+i^*},U_q)\mid\mathcal{E}'_{(j^*-1)\Delta+i^*}\wedge\mathcal{S}'_{(j^*-1)\Delta+i^*}\ ]^2}{2^{e+2}}\\
\geq\ &(1/r\Delta n^c)\cdot\frac{\varepsilon^2_{\mathsf{A}'}/4\Delta^2}{2^7\cdot n^c r\cdot\Delta}\ \ =\ \ \frac{\varepsilon^2_{\mathsf{A}'}}{2^9\cdot n^{2c}\cdot r^2\cdot\Delta^4}\ \ .
\end{aligned}$$

where we apply Lemma 3.1 to complete the proof for (24).

# References


[1] Boaz Barak, Yevgeniy Dodis, Hugo Krawczyk, Olivier Pereira, Krzysztof Pietrzak, François-Xavier Standaert, and Yu Yu. Leftover hash lemma, revisited. In Phillip Rogaway, editor, *CRYPTO*, LNCS, pages 1–20. Springer, 2011.

[2] Manuel Blum and Silvio Micali. How to generate cryptographically strong sequences of pseudo random bits. In *Proceedings of the 23rd IEEE Symposium on Foundation of Computer Science*, pages 112–117, 1982.

[3] Yevgeniy Dodis, Krzysztof Pietrzak, and Daniel Wichs. Key derivation without entropy waste. In *Proceedings of the 33rd Annual International Conference on the Theory and Applications of Cryptographic Techniques (Eurocrypt 2014)*, 2014. (to appear).

[4] Yevgeniy Dodis and Yu Yu. Overcoming weak expectations. In *Proceedings of the 10th Theory of Cryptography Conference (TCC 2013)*, pages 1–22, 2013.

[5] Oded Goldreich, Hugo Krawczyk, and Michael Luby. On the existence of pseudorandom generators. *SIAM Journal on Computing*, 22(6):1163–1175, 1993.

[6] Oded Goldreich and Leonid A. Levin. A hard-core predicate for all one-way functions. In D. S. Johnson, editor, *Proceedings of the Twenty First Annual ACM Symposium on Theory of Computing*, pages 25–32, Seattle, Washington, 15–17 May 1989.

[7] Iftach Haitner, Danny Harnik, and Omer Reingold. On the power of the randomized iterate. In *Proceedings of the 26th International Cryptology Conference (CRYPTO 2006)*, pages 22–40, 2006.

[8] Iftach Haitner, Danny Harnik, and Omer Reingold. On the power of the randomized iterate. *SIAM Journal on Computing*, 40(6):1486–1528, 2011. draft of full version available at http://www.cs.tau.ac.il/~iftachh/papers/RandomizedIteate/RandomIterate.pdf.

[9] Iftach Haitner, Omer Reingold, and Salil P. Vadhan. Efficiency improvements in constructing pseudorandom generators from one-way functions. In *Proceedings of the 42nd ACM Symposium on the Theory of Computing*, pages 437–446, 2010.

[10] Johan Håstad, Russell Impagliazzo, Leonid A. Levin, and Michael Luby. Construction of pseudorandom generator from any one-way function. *SIAM Journal on Computing*, 28(4):1364–1396, 1999.

[11] Thomas Holenstein. Pseudorandom generators from one-way functions: A simple construction for any hardness. In *Proceedings of the 3rd Theory of Cryptography Conference (TCC 2006)*, 2006.

[12] Russell Impagliazzo, Noam Nisan, and Avi Wigderson. Pseudorandomness for network algorithms. In *Proceedings of the 26th ACM Symposium on the Theory of Computing*, pages 356–364, 1994.

[13] Noam Nisan. Pseudorandom generators for space-bounded computation. *Combinatorica*, 12(4):449–461, 1992.

[14] Salil P. Vadhan and Colin Jia Zheng. Characterizing pseudoentropy and simplifying pseudorandom generator constructions. In *Proceedings of the 44th ACM Symposium on the Theory of Computing*, pages 817–836, 2012.

[15] Andrew Chi-Chih Yao. Theory and applications of trapdoor functions (extended abstract). In *Proceedings of the 23rd IEEE Symposium on Foundation of Computer Science*, pages 80–91, 1982.

[16] Yu Yu, Xiangxue Li, and Jian Weng. Pseudorandom generators from regular one-way functions: New constructions with improved parameters. In *ASIACRYPT*, pages 261–279, 2013.

# A Proofs Omitted

*Proof of Lemma 3.3.* (4) follows from the assumption that $f$ is a regular function, and the fact that $f(U_n)$ is independent of any other distributions. As for (5), consider running two instances of the iterate seeded with independent $x_1$ and $x_1'$ and a common random $\vec{h}^{k-1}$, the probability of colliding on $y_k$ is upper bounded by the sum of the events that the first collision occurs on points $y_1$, …, $y_k$ respectively, where $x_1$, …, $x_k$ are all i.i.d. to uniform due to the universality of $\mathcal{H}$. It follows that

$$\mathsf{CP}(Y_k \mid \vec{H}^{k-1}\ ) \leq k \cdot \mathsf{CP}(f(U_n)) = \frac{k}{|f(\{0,1\}^n)|}\ .$$

We sketch the proof of (6) as below (see [8, Lemma 3.11] for details): consider the following $(2n, n+1, \log|\mathcal{H}|)$-LBP $M$ for the input $(x_1,x_1')$: the source node is labeled by $(y_1 = f(x_1), y_1' = f(x_1'))$, and being on node labeled by $(y_i, y'_i)$ at the $i^{th}$ layer, it takes the current layer input $h_i \in \mathcal{H}$, and computes $y_{i+1} := f(h_i(y_i))$, $y'_{i+1} := f(h_i(y'_i))$. Finally, $M$ moves to the 1-labeled node if $y_{n+1} = y'_{n+1}$ or the 0-labeled node otherwise. Note that the probability that $M$ outputs 1 is equal to that the two iterates (with inputs $x_1$ and $x'_1$ respectively, and using the same hash function $\vec{h}^n$) collide on $y_{n+1} = y'_{n+1}$. As $BSG$ $2^{-2n}$-fools every $(2n, n+1, \log|\mathcal{H}|)$-LBP (including $M$), replacing uniform $\vec{H}^{k-1}$ with $BSG(U_q)$ will not increase the collision probability by more than $2^{-2n}$, i.e.,

$$\mathsf{CP}(\ Y'_k \mid BSG(U_q)\ ) \ \leq\ \mathsf{CP}(Y_k|\vec{H}^{k-1}) + 2^{-2n} \ \leq\ \frac{k}{|f(\{0,1\}^n)|} + 2^{-2n} \ \leq\ \frac{k+1}{|f(\{0,1\}^n)|}\ .$$

and it is not hard to see that for any $\vec{h}^{k-1}$ and any $u_1, u_2 \in BSG^{-1}(\vec{h}^{k-1})$

$$\mathsf{CP}(\ Y'_k \mid U_q = u_1\ ) \ =\ \mathsf{CP}(\ Y'_k \mid U_q = u_2) \ =\ \mathsf{CP}(\ Y'_k \mid BSG(U_q) = \vec{h}^{k-1}\ )\ .$$

We complete the proof by

$$\mathsf{CP}(\ Y'_k \mid U_q\ ) \ =\ \mathsf{CP}(\ Y'_k \mid BSG(U_q)\ ) \ \leq\ \frac{k+1}{|f(\{0,1\}^n)|}\ .$$

□

*Proof of Lemma 4.1.* We see $\mathcal{F}$ as a family of universal hash functions and let $F$ be a uniform distribution over $\mathcal{F}$. For $m \leq n-d$ we have by the leftover hash lemma that

$$\mathbb{E}_{f \xleftarrow{\$} \mathcal{F}}[\ \mathsf{SD}(\ f(U_n)\ ,\ U_m\ )\ ] \ =\ \mathsf{SD}(\ F(U_n)\ ,\ U_m \mid F\ ) \ \leq\ 2^{-\frac{d}{2}}\ .$$

It follows by a Markov inequality that the above statistical distance is bounded by $2^{-d/2} \cdot 2^{d/4}$ except for a $2^{-d/4}$-fraction of $f$. We proceed to the case for $m > n-d$ to get

$$\mathsf{CP}(\ F(U_n) \mid F\ ) \ \leq\ \mathsf{CP}(U_n)\ +\ \max_{x_1 \neq x_2}\{\ \Pr[F(x_1) = F(x_2)]\ \} \ =\ 2^{-n} + 2^{-m} \ \leq\ 2^{-n+d+1}$$

We define $\mathcal{S} \stackrel{\mathsf{def}}{=} \{(y,f) : |f^{-1}(y)| > 2^{2d+1}\}$ to yield

$$\begin{aligned} 2^{-n+d+1} \quad &\geq \mathsf{CP}(\ F(U_n) \mid F\ ) \ =\ \sum_f \Pr[F=f] \sum_y \Pr[f(U_n)=y]^2 \\ \geq \quad & 2^{-n+2d+1} \cdot \sum_f \Pr[F=f] \sum_{y:(y,f)\in\mathcal{S}} \Pr[f(U_n)=y] \\ = \quad & 2^{-n+2d+1} \cdot \Pr[(F(U_n),F) \in \mathcal{S}]\ , \end{aligned}$$

and thus $\Pr[(F(U_n),F) \in \mathcal{S}] \leq 2^{-d}$. This completes the proof. Note that $|f^{-1}(y)| \geq 1$ for any $y = f(x)$.
□

*Proof of Lemma 4.2.* If (9) does not hold for every $n_0 \in \mathbb{N} \cup \{0\}$, then there must exist an $n^0$ such that $\mu_0(\cdot)$, ... $\mu_{n_0-1}(\cdot)$ are negligible and $\mu_{n_0}(\cdot)$ is non-negligible. We then define $\max(\cdot)$ as

$$\max(n) \stackrel{\mathsf{def}}{=} \begin{cases} i_{\kappa(n)-n_0}, & \text{if } \kappa(n) > n_0 \\ \perp\ , & \text{if } \kappa(n) \leq n_0 \end{cases}$$

It is easy to see that $\mathcal{Y}_{i_{\kappa(n)-n_0+1}}$, ..., $\mathcal{Y}_{i_{\kappa(n)}}$ sum to a negligible fraction in $n$ (i.e., the sum of a finite number of negligible functions $\mu_0(\cdot)$, ... $\mu_{n_0-1}(\cdot)$ results into another negligible function). Denote by $\mathcal{N}_\perp \stackrel{\mathsf{def}}{=} \{n \in \mathbb{N} \cup \{0\} : \max(n) = \perp\}$. We have by assumption that for some constant $c$ that $\mu_{n_0}(n) \geq n^{-c}$ for infinitely many $n \in \mathbb{N} \cup \{0\}$, and thus $\mu_{n_0}(n) \geq n^{-c}$ holds also for infinitely many $n \in \mathbb{N} \cup \{0\} \setminus \mathcal{N}_\perp$. This is due to $\mu_{n_0}(n) = 0$ for any $n \in \mathcal{N}_\perp$. Therefore, $\Pr[f(U_n) \in \mathcal{Y}_{\max}]$ is non-negligible, which completes the proof. □

**Fact A.1** *For any $k \in [r\Delta]$, we have*

$$\mathsf{CP}(\ (Y_k, \vec{H}^{r\Delta-1}) \mid \mathcal{E}_k \wedge \mathcal{S}_k\ ) \;=\; \frac{\mathsf{CP}(\ (Y_k, \vec{H}^{r\Delta-1})\ \wedge\ \mathcal{E}_k\ \wedge\ \mathcal{S}_k\ )}{\Pr[\mathcal{E}_k \wedge \mathcal{S}_k]^2} \;=\; \frac{\mathsf{CP}(\ Y_k\ \wedge\ \mathcal{E}_k\ \wedge\ \mathcal{S}_k \mid \vec{H}^{r\Delta-1})}{\Pr[\mathcal{E}_k \wedge \mathcal{S}_k]^2 \cdot |\mathcal{H}|^{r\Delta-1}} \tag{26}$$

*Proof of Fact A.1.* We first have that

$$\begin{aligned}
& \mathsf{CP}(\ (Y_k, \vec{H}^{r\Delta-1}) \mid \mathcal{E}_k \wedge \mathcal{S}_k\ ) \cdot \Pr[\mathcal{E}_k \wedge \mathcal{S}_k]^2 \\
= \;& \Pr[\mathcal{E}_k \wedge \mathcal{S}_k]^2 \cdot \sum_{(y, \vec{h}^{r\Delta-1})} \Pr[\ (Y_k, \vec{H}^{r\Delta-1}) = (y, \vec{h}^{r\Delta-1}) \mid \mathcal{E}_k \wedge \mathcal{S}_k\ ]^2 \\
= \;& \sum_{(y, \vec{h}^{r\Delta-1})} \big(\Pr[\ (Y_k, \vec{H}^{r\Delta-1}) = (y, \vec{h}^{r\Delta-1}) \mid \mathcal{E}_k \wedge \mathcal{S}_k\ ] \cdot \Pr[\mathcal{E}_k \wedge \mathcal{S}_k]\big)^2 \\
= \;& \sum_{(y, \vec{h}^{r\Delta-1})} \Pr[\ (Y_k, \vec{H}^{r\Delta-1}) = (y, \vec{h}^{r\Delta-1})\ \wedge\ \mathcal{E}_k \wedge \mathcal{S}_k\ ]^2 \\
= \;& \mathsf{CP}(\ (Y_k, \vec{H}^{r\Delta-1})\ \wedge\ \mathcal{E}_k\ \wedge\ \mathcal{S}_k\ )\ ,
\end{aligned}$$

and complete the proof by the following

$$\begin{aligned}
& \frac{\mathsf{CP}(\ Y_k\ \wedge\ \mathcal{E}_k\ \wedge\ \mathcal{S}_k \mid \vec{H}^{r\Delta-1}\ )}{|\mathcal{H}|^{r\Delta-1}} \\
= \;& \frac{1}{|\mathcal{H}|^{r\Delta-1}} \cdot \sum_{\vec{h}^{r\Delta-1}} \Pr[H^{r\Delta-1} = \vec{h}^{r\Delta-1}] \sum_{y} \Pr[Y_k = y\ \wedge\ \mathcal{E}_k\ \wedge\ \mathcal{S}_k\ |H^{r\Delta-1} = \vec{h}^{r\Delta-1}\ ]^2 \\
= \;& \sum_{(y, \vec{h}^{r\Delta-1})} \big(\Pr[H^{r\Delta-1} = \vec{h}^{r\Delta-1}] \cdot \Pr[Y_k = y\ \wedge\ \mathcal{E}_k\ \wedge\ \mathcal{S}_k\ |H^{r\Delta-1} = \vec{h}^{r\Delta-1}\ ]\big)^2 \\
= \;& \sum_{(y, \vec{h}^{r\Delta-1})} \Pr[(Y_k, H^{r\Delta-1}) = (y, \vec{h}^{r\Delta-1})\ \wedge\ \mathcal{E}_k\ \wedge\ \mathcal{S}_k\ ]^2 \\
= \;& \mathsf{CP}(\ (Y_k, \vec{H}^{r\Delta-1})\ \wedge\ \mathcal{E}_k\ \wedge\ \mathcal{S}_k\ )\ .
\end{aligned}$$

□

*Proof of (20).* For any $k \le r\Delta$, we will define a $(n+1, r\Delta, \log|\mathcal{H}|)$-LBP $M_1$ that on input $x_1$ (at the source node) and $\vec{h}^{r\Delta-1}$ ($h_i \in \mathcal{H}$ at each $i^{th}$ layer), outputs 1 iff every $t \in [k]$ satisfies $y_t \in \mathcal{Y}_{[\max]}$. The *BSG* $2^{-2n}$-fools $M_1$, i.e., for any $x_1 \in \{0,1\}^n$

$$|\ \Pr[M_1(x_1, \vec{H}^{r\Delta-1}) = 1] - \ \Pr[M_1(x_1, BSG(U_q)) = 1]\ | = \ |\Pr[\mathcal{S}_k \mid X_1 = x_1] - \Pr[\mathcal{S}'_k \mid X_1 = x_1]|\ \le\ 2^{-2n}$$

and thus

$$\Pr[\mathcal{S}'_k] \ge \Pr[\mathcal{S}_k] - 2^{-2n} \ge 1 - k\epsilon - 2^{-2n}\ .$$

The bounded-spaced computation of $M_1$ is as follows: the source node input is ($y_1 \in \{0,1\}^n, \mathsf{tag}_1 \in \{0,1\}$), where $y_1 = f(x_1)$ and $\mathsf{tag}_1 = 1$ iff $y_1 \in \mathcal{Y}_{[\max]}$ (or 0 otherwise). At each $i^{th}$ layer up to $i = k$, it computes $x_i := h_{i-1}(y_{i-1})$, $y_i := f(x_i)$ and sets $\mathsf{tag}_i := 1$ iff $\mathsf{tag}_{i-1} = 1$ and $y_i \in \mathcal{Y}_{[\max]}$ ($\mathsf{tag}_i := 0$ otherwise). Finally, $M_1$ produces $\mathsf{tag}_k$ as the final output.
Similarly, we define another $(n+1, r\Delta, \log|\mathcal{H}|)$-LBP $M_2$ that on input $(x_1, \vec{h}^{r\Delta-1})$, outputs 1 iff $y_k \in \mathcal{Y}_{\max}$, and thus

$$\Pr[\mathcal{E}'_k] \ge \Pr[\mathcal{E}_k] - 2^{-2n} \ge n^{-c} - 2^{-2n}\ .$$

The computation of $M_2$ is simply to compute $x_i := h_{i-1}(y_{i-1})$ and $y_i := f(x_i)$ at each $i^{th}$ iteration and to output 1 iff $y_k \in \mathcal{Y}_{\max}$. It follows that

$$\Pr[\mathcal{E}'_k \wedge \mathcal{S}'_k] \ \geq \ \Pr[\mathcal{E}'_k] \ - \ \Pr[\neg\mathcal{S}'_k] \ \geq \ n^{-c} - 2^{-2n} - (k\epsilon + 2^{-2n}) \ \geq \ n^{-c}/2 \ .$$

□

*Proof of (21).* For any $k \in [(r-1)\Delta]$, consider the following $(n+1, r\Delta, \log|\mathcal{H}|)$-LBP $M_3$: on source node input $y_1 = f(x_1)$ and layered input vector $\vec{h}^{r\Delta-1}$, it computes $x_i := h_{i-1}(y_{i-1})$, $y_i := f(x_i)$ at each $i^{th}$ layer. For iterations numbered by $(k+1) \leq i \leq (k+\Delta)$, it additionally sets $\mathsf{tag}_i = 1$ iff either $\mathsf{tag}_{i-1} = 1$ or $y_i \in \mathcal{Y}_{\max}$, where $\mathsf{tag}_k$ is initialized to 0. Finally, $M_3$ outputs $\mathsf{tag}_{k+\Delta}$. By the bounded space generator we have

$$|\ \Pr[M_3(X_1, \vec{H}^{r\Delta-1}) = 1] \ - \ \Pr[M_3(X_1, BSG(U_q) = 1]\ | = \ |\Pr[\bigvee_{i=k+1}^{k+\Delta} \mathcal{E}_i\ ] - \Pr[\bigvee_{i=k+1}^{k+\Delta} \mathcal{E}'_i\ ]\ | \ \leq \ 2^{-2n} \ ,$$

and thus by (15)

$$\Pr[\bigvee_{i=k+1}^{k+\Delta} \mathcal{E}'_i\ ] \ \geq \ \Pr[\bigvee_{i=k+1}^{k+\Delta} \mathcal{E}_i\ ] \ - \ 2^{-2n} \ \geq \ 1 - n^{-\alpha} - 2^{-2n} \ .$$

□

*Proof of (22).* For any $k \in [r\Delta]$, consider the following $(2n+1, r\Delta, \log|\mathcal{H}|)$-LBP $M_4$: on source node input $(y_1 = f(x_1), y'_1 = f(x'_1), \mathsf{tag}_1 \in \{0,1\})$, where $\mathsf{tag}_1 = 1$ iff both $y_1, y'_1 \in \mathcal{Y}_{[\max]}$. For $1 \leq i \leq k$, at each $i^{th}$ layer $M_4$ computes $y_i := f(h_{i-1}(y_{i-1}))$, $y'_i := f(h_{i-1}(y'_{i-1}))$ and sets $\mathsf{tag}_i = 1$ iff $\mathsf{tag}_{i-1} = 1 \wedge y_i \in \mathcal{Y}_{[\max]} \wedge y'_i \in \mathcal{Y}_{[\max]}$. Finally, at the $(k+1)^{th}$ layer $M_4$ outputs 1 iff $y_k = y'_k \in \mathcal{Y}_{\max}$ (in respect for event $\mathcal{E}_k/\mathcal{E}'_k$) and $\mathsf{tag}_k = 1$ (in honor of $\mathcal{S}_k/\mathcal{S}'_k$). Imagine running two iterates with random $x_1$, $x'_1$ and seeded by a common hash function from distribution either $\vec{H}^{r\Delta-1}$ or $BSG(U_q)$, we have

$$\mathsf{CP}(Y_k \ \wedge \ \mathcal{E}_k \wedge \mathcal{S}_k \mid \vec{H}^{r\Delta-1}\ ) \ = \ \Pr_{(x_1,x'_1)\leftarrow U_{2n},\ \vec{h}^{r\Delta-1}\leftarrow\vec{H}^{r\Delta-1}}[M_4(x_1, x'_1, \vec{h}^{r\Delta-1}) = 1]$$

$$\mathsf{CP}(Y'_k \ \wedge \ \mathcal{E}'_k \wedge \mathcal{S}'_k \mid BSG(U_q)\ ) \ = \ \Pr_{(x_1,x'_1)\leftarrow U_{2n},\ \vec{h}^{r\Delta-1}\leftarrow BSG(U_q)}[M_4(x_1, x'_1, \vec{h}^{r\Delta-1}) = 1]$$

and thus

$$\begin{aligned}
&|\ \mathsf{CP}(Y_k \ \wedge \ \mathcal{E}_k \wedge \mathcal{S}_k \mid \vec{H}^{r\Delta-1}\ ) \ - \ \mathsf{CP}(Y'_k \ \wedge \ \mathcal{E}'_k \wedge \mathcal{S}'_k \mid BSG(U_q)\ )\ | \\
&\leq \ \mathbb{E}_{(x_1,x'_1)\leftarrow U_{2n}}\left[\ |\ \Pr[M_4(x_1, x'_1, \vec{H}^{r\Delta-1}) = 1] \ - \ \Pr[M_4(x_1, x'_1, BSG(U_q)) = 1]|\ \right] \\
&\leq \ 2^{-2n} \quad .
\end{aligned}$$

It follows by (16) that

$$\mathsf{CP}(Y'_k \ \wedge \ \mathcal{E}'_k \wedge \mathcal{S}'_k \mid BSG(U_q)\ ) \ \leq \ \mathsf{CP}(Y_k \ \wedge \ \mathcal{E}_k \wedge \mathcal{S}_k \mid \vec{H}^{r\Delta-1}\ ) \ + \ 2^{-2n} \ \leq \ (r\Delta + 1)\cdot 2^{\max - n + 1} \quad .$$

Note that $y_k$, $\mathcal{E}'_k$ and $\mathcal{S}'_k$ depend only on $x_1$ and $\vec{h}^{r\Delta-1}$, namely, for any $\vec{h}^{k-1}$ and any $u_1, u_2 \in BSG^{-1}(\vec{h}^{k-1})$,

$$\mathsf{CP}(\ Y'_k \wedge \mathcal{E}'_k \wedge \mathcal{S}'_k \mid U_q = u_1\ ) \ = \ \mathsf{CP}(\ Y'_k \wedge \mathcal{E}'_k \wedge \mathcal{S}'_k \mid U_q = u_2) \ = \ \mathsf{CP}(\ Y'_k \wedge \mathcal{E}'_k \wedge \mathcal{S}'_k \mid BSG(U_q) = \vec{h}^{k-1}\ ) \ .$$

Therefore,

$$\mathsf{CP}(Y'_k \ \wedge \ \mathcal{E}'_k \wedge \mathcal{S}'_k \mid U_q\ ) \ = \ \mathsf{CP}(Y'_k \ \wedge \ \mathcal{E}'_k \wedge \mathcal{S}'_k \mid BSG(U_q)\ ) \ \leq \ (r\Delta + 1)\cdot 2^{\max - n + 1} \quad .$$

□

*Proof of (25).* We have that

$$
\begin{aligned}
& \mathsf{CP}(\ (Y'_{(j^*-1)\Delta+i^*}, U_q\ ) \mid \mathcal{E}'_{(j^*-1)\Delta+i^*} \wedge \mathcal{S}'_{(j^*-1)\Delta+i^*}\ ) \\
= \quad & \frac{\mathsf{CP}(\ (Y'_{(j^*-1)\Delta+i^*}, U_q\ )\ \wedge\ \mathcal{E}'_{(j^*-1)\Delta+i^*} \wedge \mathcal{S}'_{(j^*-1)\Delta+i^*})}{\Pr[\mathcal{E}'_{(j^*-1)\Delta+i^*} \wedge \mathcal{S}'_{(j^*-1)\Delta+i^*}]^2} \\
\le \quad & \mathsf{CP}(\ Y_{(j^*-1)\Delta+i^*}\ \wedge\ \mathcal{E}_{(j^*-1)\Delta+i^*}\ \wedge \mathcal{S}_{(j^*-1)\Delta+i^*} \mid U_q\ )\ \frac{1}{(n^{-2c}/4)\cdot 2^q} \\
\le \quad & \frac{(r\Delta+1)\cdot 2^{\max-n+1}}{(n^{-2c}/4)\cdot 2^q}\ \le\ \frac{16r\Delta\cdot n^{2c}}{2^{n-\max}\cdot 2^q}\ ,
\end{aligned}
$$

where the equalities are similar to that in Fact A.1 (by renaming $\vec{H}^{r\Delta-1}$ to $U_q$), and the two inequalities are due to (20) and (22) respecitvely and thus

$$\mathbf{H}_2(\ (Y'_{(j-1)\Delta+i^*}, U_q\ ) \mid \mathcal{E}'_{(j-1)\Delta+i^*} \wedge \mathcal{S}'_{(j-1)\Delta+i^*}\ ) \ge\ n - \max + q - 2c\cdot \log n - \log r - \log\Delta - 4\ .$$

The uniform distribution over $\mathcal{Y}_{\max}\times\{0,1\}^q$ has entropy

$$\mathbf{H}_2(\ (f(U_n), U_q) \mid f(U_n)\in\mathcal{Y}_{\max}\ ) = \log(\frac{n^{-c}}{2^{-n+\max-1}}\cdot 2^q) = n - \max + q - c\cdot\log n + 1\ ,$$

and thus the entropy deficiency (i.e., the difference of two entropies above) $e \le c\log n + \log r + \log\Delta + 5$.
□

## B Definitions, Explanations and Remarks

**Remark B.1 (some intuitions for $\mathcal{S}_k$)** *Throughout the proofs, we consider the (inverting, collision, etc.) probabilities conditioned on event $\mathcal{S}_k$, which requires that during the first $k$ iterations no $y_i$ ($1\le i\le k$) hits the negligible fraction. This might look redundant as $\mathcal{S}_k$ occurs with overwhelming probability (by (14)). However, our proofs crucially rely on the fact that, as stated in (16), the collision probability of $y_k$ conditioned on $\mathcal{S}_k$ is almost the same (roughly $\tilde{O}(2^{\max-n})$, omitting $\mathsf{poly}(n)$ factors) as the ideal case, i.e., the collision probability of $f(U_n)$ conditioned on $f(U_n)\in\mathcal{Y}_{\max}$ . This would not have been possible if not being conditioned on $\mathcal{S}_k$ even though $\mathcal{Y}_{\max+1}$, ..., $\mathcal{Y}_n$ only sum to a negligible function $\mathsf{negl}(n)$. To see this, consider the following simplified case for $k=1$, the collision probability of $y_1$ is equal to that of $f(U_n)$, and thus we have*

$$\frac{1}{2}\cdot\sum_{i=1}^{n} 2^{i-n}\cdot\Pr[f(U_n)\in\mathcal{Y}_i]\ \le\ \left(\mathsf{CP}(f(U_n)) = \sum_{i=1}^{n}\sum_{y\in\mathcal{Y}_i}\Pr[f(U_n)=y]^2\right)\ <\ \sum_{i=1}^{n} 2^{i-n}\cdot\Pr[f(U_n)\in\mathcal{Y}_i]$$

*Suppose that there is some $\mathcal{Y}_t$ such that $t=\max+\Omega(n)$ and $\Pr[f(U_n)\in\mathcal{Y}_t]=\mathsf{negl}(n)$, then the above collision probability is of the order $O(2^{\max-n}(n^{-c}+2^{\Omega(n)}\mathsf{negl}(n))$. By setting $\mathsf{negl}(n)=n^{-logn}$, the collision probability blows up by nearly a factor of $2^{\Omega(n)}$ than the desired case $\tilde{O}(2^{\max-n})$, and thus unable to apply Lemma 3.1. In contrast, conditioned on $\mathcal{S}_1$ the collision probability is $\tilde{O}(2^{\max-n})$.*

**Definition B.1 (Collision probabilities conditioned on $\mathcal{S}'_k$ and $\mathcal{E}'_k$)** *In the derandomized version, we will use the following conditional collision probabilities, whose definitions (quite naturally extend the standard collision probabilities) as follows:*

$$\mathsf{CP}(\ Y'_k\ \wedge\ \mathcal{E}'_k\wedge\mathcal{S}'_k \mid U_q)\ \stackrel{\mathsf{def}}{=}\ \mathbb{E}_{u\leftarrow U_q}\left[\sum_y \Pr[f^k(X_1,\vec{H'}^{r\Delta-1})=y\ \wedge\ \mathcal{E}'_k\wedge\mathcal{S}'_k \mid \vec{H'}^{r\Delta-1}=BSG(u)]^2\right]$$

$$\mathsf{CP}(\ Y'_k\ \wedge\ \mathcal{E}'_k \wedge \mathcal{S}'_k \mid BSG(U_q)\ )$$
$$\stackrel{\mathsf{def}}{=}\ \mathbb{E}_{\vec{h'}^{r\Delta-1}\leftarrow BSG(U_q)}\Big[\ \textstyle\sum_y \Pr[\ f^k(X_1,\vec{H'}^{r\Delta-1}) = y\ \wedge\ \mathcal{E}'_k\ \wedge\ \mathcal{S}'_k \mid \vec{H'}^{r\Delta-1} = \vec{h'}^{r\Delta-1}\ ]^2\ \Big]$$
$$\mathsf{CP}(\ Y'_k, U_q \mid \mathcal{E}'_k \wedge \mathcal{S}'_k\ )\ \stackrel{\mathsf{def}}{=}\ \sum_{(y,u)} \Pr[\ f^k(X_1, BSG(U_q)) = y\ \wedge\ U_q = u \mid \mathcal{E}'_k \wedge \mathcal{S}'_k\ ]^2\ .$$

---

**Algorithm 2** $\mathsf{M}^{\mathsf{A}'}$.

---

**Input:** $y \in \{0,1\}^n$

Sample $j \xleftarrow{\$} [r]$, $i \xleftarrow{\$} [\Delta]$, $u \xleftarrow{\$} \{0,1\}^q$, $\vec{h}^{r\Delta-1} := BSG(u)$;
Let $\tilde{y}_{(j-1)\Delta+i} := y$ ;
FOR $k = (j-1)\Delta + i + 1$ TO $(j-1)\Delta + \Delta$
  Compute $\tilde{x}_k := h_{k-1}(\tilde{y}_{k-1})$, $\tilde{y}_k := f(\tilde{x}_k)$;
$\tilde{x}_{(j-1)\Delta+1} \leftarrow \mathsf{A}'(\tilde{y}_{j\Delta}, u)$;
FOR $k = (j-1)\Delta + 1$ TO $(j-1)\Delta + i - 1$
  Compute $\tilde{y}_k := f(\tilde{x}_k)$, $\tilde{x}_{k+1} := h_k(\tilde{y}_k)$ ;

**Output:** $\tilde{x}_{(j-1)\Delta+i}$

---

**Remark B.2 (On weakening the condition of (7).)** *It is not hard to see from the proof that our construction only assumes a weaker condition than (7), i.e., for some constant $c \geq 0$ it holds that*

$$\Pr[\ f(U_n) \in \bigcup_{j=\max(n)-\log(1/\varepsilon)/10}^{\max(n)} \mathcal{Y}_j\ ] \geq n^{-c}\ . \tag{27}$$

*Note that there is nothing special about the constant 1/10 in (27), which can be replaced by other small constants. We sketch the idea of adapting the proof to the relaxed assumption. By averaging there exists $d \in [0, \log(1/\varepsilon)/10]$ such that $\mathcal{Y}_{\max-d}$ has weight at least $n^{-c-1}$. We thus consider the chance that $Y_j$ hits $\mathcal{Y}_{\max-d}$ (instead of $\mathcal{Y}_{\max}$ as we did in the original proof), and $O(n^{2c}\cdot\omega(\log n))$ iterations are bound to hit $\mathcal{Y}_{\max-d}$ at least once. Now we adapt the proof of Lemma 5.3. Ideally, conditioned on $f(U_n) \in \mathcal{Y}_{\max-d}$ the distribution $(f(U_n), \vec{H}^{r\Delta-1})$ is uniform over $\mathcal{Y}_{\max} \times \mathcal{H}^{r\Delta-1}$ with full entropy*

$$\mathbf{H}_2(\ (f(U_n), \vec{H}^{r\Delta-1}) \mid f(U_n) \in \mathcal{Y}_{\max-d}\ ) = \log(\frac{n^{-c-1}}{2^{-n+\max-d-1}}\cdot|\vec{H}|^{r\Delta-1}) = n-\max+d+(r\Delta-1)\log|\mathcal{H}|-O(\log n)\ .$$

*However, we actually only have that*

$$\mathbf{H}_2(\ (Y_{(j-1)\Delta+i^*}, \vec{H}^{r\Delta-1}) \mid \mathcal{E}_{(j-1)\Delta+i^*} \wedge \mathcal{S}_{(j-1)\Delta+i^*}\ ) \geq \Big(n - \max + d + (r\Delta-1)\log|\mathcal{H}| - O(\log n)\Big) - e\ ,$$

where entropy deficiency $e \leq d + O(\log n)$. Then, we apply Lemma 3.1 and the hard-to-invertness only blows up by a factor of roughly $2^e = n^{O(1)}(1/\varepsilon)^{1/10}$ than the ideal $\varepsilon$ (and taking a square root afterwards), which does not kill the iterate. Therefore, the iterate is hard to invert for every $O(n^{2c}\cdot\omega(\log n))$ iterations. The proof for the derandomized version can be adapted similarly.



ECCC ISSN 1433-8092
http://eccc.hpi-web.de