On the Complexity of Breaking Pseudoentropy*

Maciej Skorski**

IST Austria
maciej.skorski@gmail.com

Abstract. Pseudoentropy has found a lot of important applications to cryptography and complexity theory. In this paper we focus on the foundational problem that has not been investigated so far, namely by how much pseudoentropy (the amount seen by computationally bounded attackers) differs from its information-theoretic counterpart (seen by unbounded observers), given certain limits on attacker's computational power?

We provide the following answer for HILL pseudoentropy, which exhibits a *threshold behavior* around the size exponential in the entropy amount:

- If the attacker size (s) and advantage (ϵ) satisfy $s \gg 2^k \epsilon^{-2}$ where k is the claimed amount of pseudoentropy, then the pseudoentropy boils down to the information-theoretic smooth entropy
- If $s \ll 2^k \epsilon^2$ then pseudoentropy could be arbitrarily bigger than the information-theoretic smooth entropy

Besides answering the posted question, we show an elegant application of our result to the complexity theory, namely that it implies the classical result on the existence of functions hard to approximate (due to Pippenger). In our approach we utilize non-constructive techniques: the duality of linear programming and the probabilistic method.

Keywords: nonuniform attacks, pseudoentropy, smooth entropy, hardness of boolean functions

1 Introduction

Pseudoentropy has recently attracted a lot of attention because of applications to complexity theory [RTTV08], leakage-resilient cryptography [DP08, Pie09], deterministic encryption [FOR15], memory delegation [CKLR11], randomness extraction [HLR07], key derivation, [SGP15] constructing pseudorandom number generators [VZ12, YLW13] or black-box separations [GW11].

* Accepted to *Theory and Applications of Models of Computation 2017*.
** Supported by the European Research Council Consolidator Grant (682815-TOCNeT).

What differs between pseudoentropy and information-theoretic entropy notions is the parametrization by adversarial resources. That is, pseudoentropy not only has *quantity* k but also *quality*, which is typically described by the attacker size s and the advantage ϵ achieved in the security game.

Despite many works on pseudoentropy applications, not much is known about relationships between k, s and ϵ for a given distribution X, in particular parameter settings that make pseudoentropy non-trivial (bigger than the information-theoretic entropy). Concrete numbers can be conjectured for some applications under assumptions about computational hardness, for example for outputs of pesudorandom generators, or keys obtained by the Diffie-Hellman protocol. Yet in many cases, like key derivation where pseudoentropy can model "weak" sources, one simply assumes pseudoentropy of certain (strong enough) quality.

Without understanding relationships between s, ϵ and k it is not clear how demanding or even non-trivial is the use of pseudoentropy in applications. This is precisely the issue we are going to address in this work.

1.1 Problem statement

In this paper we are interested in separating pseudoentropy (entropy seen by bounded attackers) from its information-theoretic counterpart (measured against unbounded attackers).

An n-bit random variable X is said to have k bits of pseudoentropy[1] against attackers of size s and advantage ϵ if for some distribution Y of min-entropy k, no circuit of size s can distinguish it from Y with advantage bigger than ϵ (see Section 2.4) [2]. Note that the notion is parametrized by the adversarial specific size s and advantage ϵ. In particular the amount decreases when s gets bigger and ϵ gets smaller (it is harder to fool adversaries with bigger resources). When s is unbounded, pseudoentropy equals the information-theoretic smooth min-entropy (see Section 2.4).

To better understand possibilities and limitations of using pseudoentropy, it is natural to ask in what parameter regimes pseudoentropy provides non-trivial computational security, that is when we have a real gain in the entropy amount comparing to the information-theoretic case.

> **Q**: How much computational power is needed to boil pseudoentropy down to information-theoretic smooth entropy?

[1] We consider here the most popular notion of HILL pseudoentropy

[2] This matches the definition of pseudorandomness when k is the length of X.

1.2 Our Contribution

Nonuniform attacks against pseudoentropy Our result exhibit a *threshold phenomana*. Intuitively, with enough computational power (say size 2^n for n-bit random variables[3]) the notion of pseudoentropy is no more stronger than the corresponding information-theoretic entropy notion. We *estimate* the value of this threshold on the circuit size s, so that above there is no computational gain and below there exists nontrivial pseudoentropy. There result is somewhat surprising because: (a) the threshold doesn't depend on the length but the entropy amount and (b) the threshold depends also on the square of the advantage

Theorem (Informal) (Breaking pseudoentropy with enough computational power). *For any k, and any s, ϵ satisfying*

$$s \gg 2^k \epsilon^{-2}$$

and for every distribution of min-entropy k, unbounded attackers and attackers of size s see the same entropy amount.

Theorem (Informal) (Matching lower bound). *For any k, and any s, ϵ satisfying*

$$s \ll 2^k \epsilon^2$$

there exists a distribution X such that

(a) (bounded attackers see k bits) pseudoentropy of X against circuits of size s and advantage ϵ is k
(b) (k bits for unbounded attackers see less than k bits) information-theoretic entropy of X is k

A short overview of our results is given in Table 1 below.

regime	result	techniques	reference
$s \gg 2^k \epsilon^{-2}$	same entropy for attackers of size s as for $s = \infty$	LP duality distinguisher optimization	Theorem 1
$s \ll 2^k \epsilon^2$	arbitrary gap in the amount for size s and $s = \infty$	probabilistic method concentration bounds	Theorem 2

Table 1: Overview of our results. The analyzed setting is k bits of pseudoentropy against size circuits of size s and advantage ϵ.

[3] As this complexity is enough to compute every boolean functions

Proof outline and our tools

Breaking pseudoentropy We outline the proof of the first result below

1. We first consider somewhat weaker pseudoentropy notion, called Metric entropy, where the order of quantifiers is reversed. That is, for any distinguisher D there has to be some Y of min-entropy k which is close to X under that particular test D, that is $\mathbb{E}\,\mathsf{D}(X) \approx \mathbb{E}\,\mathsf{D}(Y)$.
2. We prove that this weaker pseudoentropy notion collapses when $s \gg 2^k$, by "compressing" distinguishers down to size 2^k. The intuitive reason for that is that we can always manipulate Y so that it has "small" support (only $O(2^k)$ elements), and if an attacker wants to maximize the advantage $|\,\mathbb{E}\,\mathsf{D}(X) - \mathbb{E}\,\mathsf{D}(Y)|$, the best strategy is to hardcode the elements x such that $\Pr[Y = x] > \Pr[X = x]$ which is a subset of the support of Y and can be implemented in size $\tilde{O}(2^k)$.
3. We use a generic transformation due to Barak at al. [BSW03, Sko15] to go back to our standard entropy notion. The transformation losses $\tilde{O}(\epsilon^2)$ in size and is based on the duality of linear programming.

This way we obtain that pseudoentropy with parameters (s, ϵ) becomes the same as the amount seen by unbounded attackers when $s = \tilde{O}\left(2^k \epsilon^{-2}\right)$. The details are explained in the proof of Theorem 1.

Matching lower bounds The proof of the second result goes as follows

1. We take a random subset $\mathcal{X} \subset \{0,1\}^k$ of size $k - c$, where c will be the gap between what bounded and unbounded attackers can see. The distribution X is the uniform distribution over $\mathcal{X}$ plus a "random shift" of an ϵ-fraction of the probability mass.
2. We argue that the ϵ-smooth entropy is still roughly k, because we have shifted only that fraction of the total probability mass. This is handled by a result of independent interest, stating that "almost" smooth distributions cannot be further smoothened (see Corollary 2)
3. We argue that the distribution X is pseudorandom provided that the class of test functions is small enough. This fact is proved by applying concentration bounds twice, once to handle the random shift and for the second time to handle the choice of $\mathcal{X}$. Intuitively, the advantage of bounded attackers is much smaller than ϵ because they are "fooled" by the random shift of a part of the probability mass. In turn, the entropy amount seen by bounded attackers is much bigger than $k - c$ because $\mathcal{X}$ is a random subset of $\{0,1\}^k$.

Putting this all together we get a strict separation: not only the amount of entropy is bigger, but also the advantage is smaller. The necessary assumption to make it work is that the class of distinguishers is much smaller than $2^{2^{k-c}\epsilon^2}$ members. For the details see the proof of Theorem 2.

1.3 Related works

Pseudorandomness exists unconditionally The classical textbook results [Gol06] shows that pseudorandomness exists unconditionally, which can be seen as a separation between pseudorandomness and min-entropy.

Our Theorem 2 is stronger as we separate pseudoentropy from smooth min-entropy (and cannot derive it from the mentioned result). From a technical point of view, the main difference is the extra random mass fluctuation (Step 1 in the above explanation), which needs to be later handled by bit more subtle probability tools (we use concentration inequalities for random variables with local dependence due to Janson).

Complexity of non-uniform attacks against PRGs De, Trevisan and Tulsani studied the complexity of nonuniform attacks against pseudorandom generators [**?**]. Their results are specialized to outputs of PRGs and are constructive, whereas our results apply to any random variable (unfortunately don't offer non-trivial results for the case of PRGs).

1.4 Applications

Hard-to-approximate boolean functions Our Theorem 2 implies the classical result [Pip76] which states that for any n and $\delta \in (0,1)$, there exist δ-hard functions[4] for size $s = \tilde{\Omega}\left(2^n(1-\delta)^2\right)$. For details, see Section 5.1.

1.5 Organization

We start with explaining basic concepts and notions in Section 2. In Section 3 we prove useful auxiliary facts about smooth min-entropy. In Section 4 we give proofs of our main results. In Section 5 we discuss applications to the complexity of approximating boolean functions.

[4] f is δ hard for size s if every circuit of size s fails to predict f w.p. at least $\frac{1+\delta}{2}$.

2 Preliminaries

2.1 Model of computations

Our results hold in the non-uniform model. We consider general classes of distinguishers, denoted by $\mathcal{D}$, which are families of functions from n bits to real values. When discussing complexity applications, we restrict $\mathcal{D}$ to classes of circuits of certain size s, with boolean or real-valued outputs.

2.2 Basic notions

Definition 1 (Statistical distance). *The statistical distance of two random variables X, Y taking values in the same finite set is defined as $\mathsf{SD}(X,Y) = \frac{1}{2}\sum_x |\Pr[X = x] - \Pr[Y = x]|$. Equivalently, $\mathsf{SD}(X,Y) = \max_{\mathsf{D}} |\mathbb{E}\,\mathsf{D}(X) - \mathbb{E}\,\mathsf{D}(Y)|$ where D runs over all boolean functions.*

2.3 Information-theoretic entropies

Definition 2 (Min-entropy). *We say that X has k bits of min-entropy if $\min_x \log \frac{1}{\Pr[X=x]} = k$.*

Definition 3 (Smooth min-entropy [RW05]). *We say that X has k bits of ϵ-smooth min-entropy, denoted by $\mathbf{H}^{\epsilon}_{\infty}(X) \geqslant k$, if X is ϵ-close in the statistical distance to some Y of min-entropy k.*

Remark 1. Smoothing entropy allows for increasing the entropy by shifting a part of the probability mass, to make the distribution look "more flat" or "more smooth".

2.4 Pseudoentropy

In what follows, X denotes an arbitrary n-bit random variable.

Definition 4 (HILL pseudoentropy [HILL88]). *We say that X has k bits of* HILL pseudoentropy *against a distinguisher class $\mathcal{D}$ and advantage ϵ, denoted by*

$$\mathbf{H}^{\mathrm{HILL}}_{s,\epsilon}(X) \geqslant k$$

if there is a random variable Y of min-entropy at least k that ϵ-fools any $\mathsf{D} \in \mathcal{D}$, that is for every $\mathsf{D} \in \mathcal{D}$ we have such that $|\mathbb{E}\,\mathsf{D}(X) - \mathbb{E}\,\mathsf{D}(Y)| \leqslant \epsilon$.

Definition 5 (Metric Pseudoentropy [BSW03]). *We say that X has k bits of* metric pseudoentropy *against a distinguisher class $\mathcal{D}$ and advantage ϵ, denoted by*

$$\mathbf{H}^{\text{Metric}}_{s,\epsilon}(X) \geqslant k$$

if for any $\mathsf{D} \in \mathcal{D}$ there is a random variable Y of min-entropy at least k that ϵ-fools this particular D that is such that $|\mathbb{E}\,\mathsf{D}(X) - \mathbb{E}\,\mathsf{D}(Y)| \leqslant \epsilon$.

Metric entropy is a convenient relaxation of HILL entropy, more suitable to work with in many cases. The important fact below shows that both notions are equivalent up to some loss in the circuit size.

Lemma 1 (Metric-to-HILL Transformation [BSW03, Sko15]). *If $\mathbf{H}^{\text{Metric}}_{s,\epsilon}(X) \geqslant k$ then $\mathbf{H}^{\text{HILL}}_{s',\epsilon'}(X) \geqslant k$ where $\epsilon' = 2\epsilon$ and $s' \approx s\epsilon^2/n$.*

Remark 2 (Abbreviations and equivalences for circuit classes). In the specific setting where $\mathcal{D}$ consists of deterministic boolean or deterministic real-valued circuits of size s we will slightly abuse the notation and write $\mathbf{H}^{\text{Metric}}_{s,\epsilon}(X) = \mathbf{H}^{\text{Metric}}_{\mathcal{D},\epsilon}(X)$. This is justified by the fact that for metric entropy deterministic real-valued circuits of size s give the same amount as deterministic boolean circuits of size $s' \approx s$ [FOR15]. In turn, for HILL entropy, deterministic boolean, deterministic randomized and deterministic real-valued circuits are equivalent with no entropy loss and with roughly same sizes [FOR15], so we also simply write $\mathbf{H}^{\text{HILL}}_{s,\epsilon}(X) = \mathbf{H}^{\text{HILL}}_{\mathcal{D},\epsilon}(X)$.

2.5 Relationships between entropy, smooth entropy, and computational entropy

The following proposition states that for extreme parameter regimes (unbounded attackers or zero advantage), pseudoentropy collapses to the information-theoretic notion of smooth-entropy (we skip the easy proof).

Proposition 1. *Let X be any n-bit random variable. Then we have*

(a) (Unbounded attackers) If $s = \infty$[5] *then*

$$\mathbf{H}^{\text{Metric}}_{s,\epsilon}(X) = \mathbf{H}^{\text{HILL}}_{s,\epsilon}(X) = \mathbf{H}^{\epsilon}_{\infty}(X) > \mathbf{H}_{\infty}(X)$$

.

[5] If the domain consists of n-bit strings, it is enough to assume $s > 2^n$ as every function over n bits has complexity at most 2^n

(b) (No smoothing) If $\epsilon = 0$ then for any s

$$\mathbf{H}_{s,\epsilon}^{\text{Metric}}(X) = \mathbf{H}_{s,\epsilon}^{\text{HILL}}(X) = \mathbf{H}_{\infty}^{\epsilon}(X) = \mathbf{H}_{\infty}(X)$$

.

(c) (General) For any s, ϵ

$$\mathbf{H}_{s,\epsilon}^{\text{Metric}}(X) \geqslant \mathbf{H}_{s,\epsilon}^{\text{HILL}}(X) \geqslant \mathbf{H}_{\infty}^{\epsilon}(X) \geqslant \mathbf{H}_{\infty}(X)$$

.

2.6 Concentration inequalities

The following lemma is a corollary from the famous concentration bound due to Jason, which exploits local dependencies

Lemma 2 (Concentration bounds for local dependencies [Jan04]). *Let $X_1, \ldots, X_n$ be random variables taking values in $[a, b]$, such that every X_i is not independent of at most Δ other variables $X_{i'}$. Let $\mu = n^{-1} \sum_{i=1}^{n} \mathbb{E}X_i$. Then*

$$\Pr\left[n^{-1}\sum_{i=1}^{n} X_i \geqslant \mu + \delta\right] \leqslant \exp\left(-\frac{2n\delta^2}{(a-b)^2(\Delta+1)}\right).$$

In particular, for $\Delta = 0$ we obtain the following bound

Corollary 1 (Hoeffding's Inequality [Hoe63]). *Let $X_1, \ldots, X_n$ be independent random variables taking values in $[a, b]$. Let $\mu = n^{-1} \sum_{i=1}^{n} \mathbb{E}X_i$. Then*

$$\Pr\left[n^{-1}\sum_{i=1}^{n} X_i \geqslant \mu + \delta\right] \leqslant \exp\left(-\frac{2n\delta^2}{(a-b)^2}\right).$$

Remark 3 (Hoeffding's Inequality for sampling without repetitions). The above inequality applies also the the setting where X_i are random samples taken from the same distribution without repetitions [Ser74].

3 Auxiliary Facts

3.1 Auxiliary results on smooth Renyi entropy

In the lemma below we show that smoothing doesn't help to increase entropy for flat distributions.

Lemma 3 (Flat distributions cannot be smoothened). *Let X be an n-bit random variable. Suppose that the distribution of X is flat and $\mathbf{H}_\infty(X) = k$. Then $\mathbf{H}^\epsilon_\infty(X) \leqslant k + \log\left(\frac{1}{1-\epsilon}\right)$ for every $\epsilon \in (0,1)$.*

Proof. Let X' be any distribution of min-entropy at least $k' > k + \log\left(\frac{1}{1-\epsilon}\right)$. Consider the distinguisher D which outputs $\mathsf{D}(x) = 1$ if $x \in \mathrm{supp}(X)$ and $\mathsf{D}(x) = 0$ otherwise. Note that $\mathbb{E}\mathsf{D}(X) = 1$ and $\mathbb{E}\mathsf{D}(X') = \frac{\mathrm{supp}(X)}{2^{k'}} < 1 - \epsilon$. Therefore $\mathbb{E}\mathsf{D}(X) - \mathbb{E}\mathsf{D}(X')$ and thus the statistical distance between X and X' is bigger ϵ.

Corollary 2 (Almost-flat distributions cannot be smoothened). *Suppose that X is ϵ_1-close to some X' being flat over 2^k elements. Then $\mathbf{H}^{\epsilon_2}_\infty(X) \leqslant k + \log\left(\frac{1}{1-\epsilon_1-\epsilon_2}\right)$ for any $\epsilon_1, \epsilon_2 > 0$ such that $\epsilon_1 + \epsilon_2 < 1$.*

Proof. Suppose not, then there exists X'' that is ϵ_2-close to X an has min-entropy at least $k' > k + \log\left(\frac{1}{1-\epsilon_1-\epsilon_2}\right)$. In particular, X'' is ϵ-close to X', where $\epsilon = \epsilon_1 + \epsilon_2$. Since X' is flat, Lemma 3 implies that the min-entropy of X'' is at most $k + \log\left(\frac{1}{1-\epsilon}\right)$, which is a contradiction.

4 Main Results

4.1 Complexity of breaking pseudoentropy

The following result specifies the attacker size for which pseudoentropy provides no computational security.

Theorem 1 (Breaking pseudoentropy is exponentially easy in the amount). *For any n bit random variable X, if $\mathbf{H}^\epsilon_\infty(X) = k$ then also $\mathbf{H}^{\mathrm{HILL}}_{s,\epsilon}(X) = k$ for $s > n^2 2^k \epsilon^{-2}$.*

The proof follows the steps explained in Section 1.2 and is given in Appendix A.

4.2 Matching lower bounds

Theorem 2 (Breaking pseudoentropy can be exponentially hard in the amount). *Let $\mathcal{S} \subset \{0,1\}^n$ be a set of cardinality 2^k, $\epsilon' \in (0,1)$ be arbitrary, and let $\mathcal{D}$ be a class of functions from $\mathcal{S}$ to $[0,1]$ such that*

$$|\mathcal{D}| < 2^{-2} \cdot 2^{2^{k-C-1}\epsilon'^2}.$$

Then for any $\epsilon < \frac{1}{4}$ there exists a random variable X on $\mathcal{S}$ such that

(a) $\mathbf{H}^{\mathrm{HILL}}_{\mathcal{D},\epsilon'}(X) = k$

(b) $\mathbf{H}^{\epsilon}_{\infty}(X) = k - C + \log\left(\frac{1}{1-2\epsilon}\right)$

Moreover, we have the following symmetry*: the probability mass function of X takes only two values on two subsets of $\mathcal{S}$ of equal size.*

Remark 4 (Doubly-strong separation: by the amount and the advantage). Note that the interesting setting of the parameters in the theorem above is when $\epsilon' \ll \epsilon$ so that not only we have a gap in the entropy amount, but even for much bigger advantage for unbounded distinguishers.

The proof follows the steps explained in Section 1.2 and appears in Appendix B.

5 Applications

5.1 Complexity of hard boolean functions

For any function f and a distribution μ on the domain of f by $\mathsf{Guess}^{\mathsf{D}}(f,\mu)$ we denote the probability of guessing f by a function D when the input is sampled according to μ, that is $\mathsf{Guess}^{\mathsf{D}}(f,\mu) = \Pr_{x\sim\mu}[\mathsf{D}(x) = f(x)]$. We say that f on n bits is δ-hard[6] for size s if $\mathsf{Guess}^{\mathsf{D}}(f,\mu) < 1 - \frac{\delta}{2}$ for every circuit D of size s and uniform μ (we also write $\mathsf{Guess}^{\mathsf{D}}(f) < 1 - \frac{\delta}{2}$).

The corollary bellow is the classical result on the complexity of hard functions. Our result is optimal up to a factor linear in n (note that for large n, the value of n is negligible comparing to 2^n. Also, most interesting settings are with $\delta \approx 1$ with a negligible gap, and we get the optimal dependency on $1-\delta$.).

Corollary 3 (Functions hard to approximate by boolean circuits). *For any n and $\delta \in (0,1)$ there exists an n-bit function which is δ-hard for all n-bit circuits of size $s = \Omega\left(2^n(1-\delta)^2\right)$.*

[6] We use the convention for which $\delta = 1$ corresponds to completely unpredictable function. Some works substitute $1-\delta$ in place of δ.

Proof (of Corollary 3). Let $\mathsf{D}'(x) = 2\mathsf{D}(x) - 1$. Denote for shortness $\mathsf{Adv}^{\mathsf{D}}(X,Y) = \mathbb{E}\,\mathsf{D}(X) - \mathbb{E}\,\mathsf{D}(Y)$. Observe that for any X, Y we have

$$\begin{aligned}
&\mathsf{Adv}^{\mathsf{D}}(X,Y)\\
&\quad= \mathbb{E}\,\mathsf{D}(X) - \mathbb{E}\,\mathsf{D}(Y)\\
&\quad= \frac{1}{2}\sum_x (2\mathsf{D}(x) - 1)\,(\Pr[X = x] - \Pr[Y = x])\\
&\quad= \mathsf{SD}(X,Y)\,\mathbb{E}_{x\sim\mu}\,\mathsf{D}'(x)\cdot \mathsf{sign}\,(\mathbf{P}_X(x) - \mathbf{P}_Y(x))\\
&\quad= \mathsf{SD}(X,Y)\left(\Pr_{x\sim\mu}\left[\mathsf{D}'(x) = f(x)\right] - \Pr_{x\sim\mathbf{P}_X-\mathbf{P}_Y}\left[\mathsf{D}'(x) \neq f(x)\right]\right)\\
&\quad= \mathsf{SD}(X,Y)\left(2\Pr_{x\sim\mu}\left[\mathsf{D}'(x) = f(x)\right] - 1\right)\\
&\quad= \mathsf{SD}(X,Y)\cdot\left(2\mathsf{Guess}^{\mathsf{D}}\,(f,\mu) - 1\right)
\end{aligned}$$

where $f(x) = \mathsf{sign}\,(\mathbf{P}_X(x) - \mathbf{P}_Y(x))$ and $\mu(x) = \frac{|\mathbf{P}_X(x) - \mathbf{P}_Y(x)|}{2\mathsf{SD}(X,Y)}$ (note that $\sum_x \mu(x) = 1$). Let us apply Theorem 2 to $k = n$, $\epsilon = \frac{1}{8}$, $\epsilon' = (1-\delta)\epsilon$ and $\mathcal{D}$ being the class of deterministic circuits of size s. Let Y be the indistinguishable distribution from the definition of HILL entropy. Since in our case Y is uniform, the function f is well-defined and moreover $\mathsf{SD}(X,Y) \geqslant \epsilon$ by (b). Thus

$$\mathsf{Adv}^{\mathsf{D}}(X,Y) > \epsilon\cdot\left(2\mathsf{Guess}^{\mathsf{D}}\,(f,\mu) - 1\right)$$

Moreover, $|\mathbf{P}_X(x) - \mathbf{P}_Y(x)|$ is constant by construction. Therefore μ is uniform and we obtain

$$\mathsf{Adv}^{\mathsf{D}}(X,Y) > \epsilon\cdot\left(2\mathsf{Guess}^{\mathsf{D}}\,(f) - 1\right)$$

Now $\mathsf{Adv}^{\mathsf{D}}(X,Y) < \epsilon(1-\delta)$ implies $\mathsf{Guess}^{\mathsf{D}}\,(f) < 1 - \frac{\delta}{2}$ for any D, which means that f is $1-\delta$-hard for size s (here we use the fact that there are exponentially many circuits of size s, so that $2^{O(s)} < 2^{2^{k-O(1)}(1-\delta)^2}$ and the assumption on the class size is satisfied).

.

References

BSW03. Boaz Barak, Ronen Shaltiel, and Avi Wigderson, *Computational analogues of entropy.*, RANDOM-APPROX, Lecture Notes in Computer Science, vol. 2764, Springer, 2003, pp. 200–215.

CKLR11. Kai-Min Chung, Yael Tauman Kalai, Feng-Hao Liu, and Ran Raz, *Memory delegation*, Advances in Cryptology - CRYPTO 2011 - 31st Annual Cryptology Conference, Santa Barbara, CA, USA, August 14-18, 2011. Proceedings, 2011, pp. 151–168.

DP08. Stefan Dziembowski and Krzysztof Pietrzak, *Leakage-resilient cryptography in the standard model*, IACR Cryptology ePrint Archive **2008** (2008), 240.

FOR15. Benjamin Fuller, Adam O'neill, and Leonid Reyzin, *A unified approach to deterministic encryption: New constructions and a connection to computational entropy*, J. Cryptol. **28** (2015), no. 3, 671–717.

Gol06. Oded Goldreich, *Foundations of cryptography: Volume 1*, Cambridge University Press, New York, NY, USA, 2006.

GW11. Craig Gentry and Daniel Wichs, *Separating succinct non-interactive arguments from all falsifiable assumptions*, Proceedings of the 43rd ACM Symposium on Theory of Computing, STOC 2011, San Jose, CA, USA, 6-8 June 2011, 2011, pp. 99–108.

HILL88. Johan Håstad, Russell Impagliazzo, Leonid A. Levin, and Michael Luby, *Pseudo-random generation from one-way functions*, PROC. 20TH STOC, 1988, pp. 12–24.

HLR07. Chun-Yuan Hsiao, Chi-Jen Lu, and Leonid Reyzin, *Conditional computational entropy, or toward separating pseudoentropy from compressibility*, Advances in Cryptology - EUROCRYPT 2007, 2007, pp. 169–186.

Hoe63. Wassily Hoeffding, *Probability inequalities for sums of bounded random variables*, Journal of the American Statistical Association **58** (1963), no. 301, 13–30.

Jan04. Svante Janson, *Large deviations for sums of partly dependent random variables*, Random Struct. Algorithms **24** (2004), no. 3, 234–248.

Pie09. Krzysztof Pietrzak, *A leakage-resilient mode of operation*, pp. 462–482, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

Pip76. Nicholas Pippenger, *Information theory and the complexity of boolean functions*, Mathematical systems theory **10** (1976), no. 1, 129–167.

RTTV08. Omer Reingold, Luca Trevisan, Madhur Tulsiani, and Salil Vadhan, *Dense subsets of pseudorandom sets*, Proceedings of the 2008 49th Annual IEEE Symposium on Foundations of Computer Science (Washington, DC, USA), FOCS '08, IEEE Computer Society, 2008, pp. 76–85.

RW05. Renato Renner and Stefan Wolf, *Simple and tight bounds for information reconciliation and privacy amplification*, ASIACRYPT'05, Springer-Verlag, 2005, pp. 199–216.

Ser74. R. J. Serfling, *Probability inequalities for the sum in sampling without replacement*, Ann. Statist. **2** (1974), no. 1, 39–48.

SGP15. Maciej Skorski, Alexander Golovnev, and Krzysztof Pietrzak, *Condensed unpredictability*, Automata, Languages, and Programming - 42nd International Colloquium, ICALP 2015, 2015, pp. 1046–1057.

Sko15. Maciej Skorski, *Metric pseudoentropy: Characterizations, transformations and applications*, Information Theoretic Security - 8th International Conference, ICITS 2015, 2015, pp. 105–122.

VZ12. Salil Vadhan and Colin Jia Zheng, *Characterizing pseudoentropy and simplifying pseudorandom generator constructions*, Proceedings of the 44th symposium on Theory of Computing (New York, NY, USA), STOC '12, ACM, 2012, pp. 817–836.

YLW13. Yu Yu, Xiangxue Li, and Jian Weng, *Pseudorandom generators from regular one-way functions: New constructions with improved parameters*, Advances in Cryptology - ASIACRYPT 2013, 2013, pp. 261–279.

A Proof of Theorem 1

Proof. We start with proving a weaker result, namely that for Metric pseudoentropy (weaker notion) the threshold equals 2^k.

Lemma 4 (The complexity of breaking Metric pseudoentropy). *If* $\mathbf{H}^{\epsilon}_{\infty}(X) = k$ *then also* $\mathbf{H}^{\mathrm{Metric}}_{s,\epsilon}(X) = k$ *for* $s > n2^k$.

Proof (Proof of Lemma 4). We will show the following claim which, by Proposition 1, implies the statement.

Claim. If $s > n2^k$ and $s' = \infty$ then $\mathbf{H}^{\mathrm{Metric}}_{s,\epsilon}(X) = \mathbf{H}^{\mathrm{Metric}}_{s',\epsilon}(X)$

Proof (Proof of Claim). It suffices to show only $\mathbf{H}^{\mathrm{Metric}}_{s,\epsilon}(X) \leqslant \mathbf{H}^{\mathrm{Metric}}_{s',\epsilon}(X)$ as the other implication is trivial. Our strategy is to show that any distinguisher D that negates the definition of Metric entropy can be implemented in size 2^k.

Suppose that $\mathbf{H}^{\mathrm{Metric}}_{s',\epsilon}(X) < k$. This means that for some D of size s' and all Y of min-entropy at least k we have $|\mathbb{E}\,\mathsf{D}(X) - \mathbb{E}\,\mathsf{D}(Y)| \geqslant \epsilon$. Since the set of all Y of min-entropy at least k is convex, the range of the expression $|\mathbb{E}\,\mathsf{D}(X) - \mathbb{E}\,\mathsf{D}(Y)|$ is an interval, so we either have always $\mathbb{E}\,\mathsf{D}(X) - \mathbb{E}\,\mathsf{D}(Y) > \epsilon$ or $\mathbb{E}\,\mathsf{D}(X) - \mathbb{E}\,\mathsf{D}(Y) < -\epsilon$. Without loosing generality assume the first possibility (otherwise we proceed the same way with the negation $\mathsf{D}'(x) = 1 - \mathsf{D}(x)$). Thus

$$\mathbb{E}\,\mathsf{D}(X) - \mathbb{E}\,\mathsf{D}(Y) > \epsilon \quad \text{for all } n \text{ bit } Y \text{ of min-entropy } k$$

where by Remark 2 we can assume that D is boolean. In particular, the set $\{x : \mathsf{D}(x) = 1\}$ cannot have more than 2^k elements, as otherwise we would put Y being uniform over x such that $\mathsf{D}(x) = 1$ and get $\mathbb{E}\,\mathsf{D}(X) - 1 > \epsilon > 0$ which contradicts the fact that D is boolean. But if D is boolean and outputs 1 at most 2^k times, can be implemented in size $n2^k$, by hardcoding this set and outputting 0 everywhere else. This means precisely that $\mathbf{H}^{\mathrm{Metric}}_{s,\epsilon}(X) < k$. Now by Proposition 1 we see that also $\mathbf{H}^{\epsilon}_{\infty}(X) < k$ which proves that $\mathbf{H}^{\mathrm{Metric}}_{s,\epsilon}(X) \leqslant \mathbf{H}^{\epsilon}_{\infty}(X)$ finishes the proof of the claim.

Having proven Lemma 4, we obtain the statement for HILL pseudoentropy by applying the transformation from Lemma 1.

B Proof of Theorem 2

Proof (Proof of Theorem 2). Let $\mathcal{X}$ be a random subset of $\mathcal{S}$ of cardinality 2^{k-C}. Let $x_1, \dots, x_{2^{k-C}}$ be the all elements of $\mathcal{X}$ enumerated according the lexicographic order. Define the following random variables $\xi(x)$

$$\xi(x) = \begin{cases} \text{random element from } \{-1, 1\}, & x = x_{2i-1} \text{ for some } i \\ -x_{2i-1}, & x = x_{2i} \text{ for some } i \end{cases} \tag{1}$$

for any x such that $x \in \mathcal{X}$. Once the choice of $\xi(x)$ is fixed, consider the distribution

$$\Pr[X = x] = \begin{cases} 2^{-k} + 2\epsilon \cdot 2^{-k} \cdot \xi(x) & x \in \mathcal{X} \\ 0, & x \notin \mathcal{X} \end{cases} \tag{2}$$

The rest of the proof splits into the following two claims:

Claim (X has small smooth min-entropy). For any choice of X and $\epsilon(x)$, we have $\mathbf{H}^{\epsilon}_{\infty}(X) \leqslant k - C + \log\left(\frac{1}{1-2\epsilon}\right)$.

Claim (X has large metric pseudo-entropy). We have $\mathbf{H}^{\text{Metric}}_{\mathcal{D},\epsilon}(X) = k$.

Proof (Small smooth min-entropy). Note that by Equation (2) the distribution of X is ϵ-close to the uniform distribution over $\mathcal{X}$. By Corollary 2 (note that k is replaced by $\log|\mathcal{X}| = k - C$), this means that that the ϵ-smooth min-entropy of X is at most $k - C + \log\left(\frac{1}{1-2\epsilon}\right)$.

Proof (Large metric entropy). Note that for any D we have

$$\begin{aligned} \mathbb{E}\,\mathsf{D}(X) &= \sum_{x \in \mathcal{X}} \mathsf{D}(x)\left(2^{-k} + \xi(x)2^{-k} \cdot 2\epsilon\right) \\ &= \mathbb{E}\,\mathsf{D}(U_{\mathcal{X}}) + 2^{-k} \cdot 2\epsilon \cdot \sum_{x \in \mathcal{X}} \mathsf{D}(x)\xi(x) \end{aligned}$$

In the next step we observe that the random variables $\xi(x)$ have the degree of dependence $\Delta = 1$. Indeed, by the construction in Equation (2), for any fixed x the random variables $\xi(x')$ are independent of $\xi(x)$ except at most one value of x'. Now, by Lemma 2 applied to the random variables $\mathsf{D}(x)\xi(x)$ we obtain

$$\Pr\left[2^{-k}\sum_{x \in \mathcal{X}} \mathsf{D}(x)\xi(x) > \delta\right] \leqslant \exp\left(-2^{k-1}\delta^2\right)$$

for any $\delta > 0$, where the probability is over $\xi(x)$ after fixing the choice of the set $\mathcal{X}$ for $z \in \{0,1\}^m$. In other words, we have

$$\Pr_{\xi(x)}\left[\mathbb{E}\,\mathsf{D}(X) \leqslant \mathbb{E}\,\mathsf{D}(U_{\mathcal{X}}) + 2\delta\epsilon\right] \tag{3}$$

with probability $1 - \exp\left(2^{k-1}\delta^2\right)$ for any fixed choice of sets $\mathcal{X}$.

In the last step, we observe that since the choice of the sets $\mathcal{X}$ is random, we have $\mathbb{E}\,\mathsf{D}(U_{\mathcal{X}}) \approx \mathbb{E}\,\mathsf{D}(U_{\mathcal{S}})$ with high probability. Indeed, by the Hoeffding bound for samples taken without repetitions (see Remark 3)

$$\Pr_{\mathcal{X}}\left[\mathbb{E}\,\mathsf{D}(U_{\mathcal{X}}) \leqslant \mathbb{E}\,\mathsf{D}(U) + 2\delta\epsilon\right] \geqslant 1 - \exp(-2^{k-C+3}\delta^2\epsilon^2) \tag{4}$$

By combining Equation (4) and Equation (3) for any D and any $\epsilon < \frac{1}{4}$ we obtain

$$\Pr_{\mathcal{X},\xi(x)}\left[\mathbb{E}\,\mathsf{D}(X) \leqslant \mathbb{E}\,\mathsf{D}(U_{\mathcal{S}}) + 4\delta\epsilon\right] \geqslant 1 - 2\exp(-2^{k-C+3}\delta^2\epsilon^2). \tag{5}$$

Replacing δ with $\delta/4$ and applying the union bound over $\mathcal{D}$ we see that

$$\Pr_{\mathcal{X},\xi(x)}\left[\forall \mathsf{D} \in \mathcal{D}:\ \mathbb{E}\,\mathsf{D}(X) \leqslant \mathbb{E}\,\mathsf{D}(U_{\mathcal{S}}) + \delta\epsilon\right] \geqslant 1 - 2|\mathcal{D}|\exp(-2^{k-C-1}\delta^2\epsilon^2).$$

and thus we have a distribution X such that

$$\forall \mathsf{D} \in \mathcal{D}:\ \mathbb{E}\,\mathsf{D}(X) \leqslant \mathbb{E}\,\mathsf{D}(U_{\mathcal{S}}) + \delta\epsilon \tag{6}$$

as long as

$$2|\mathcal{D}| < 2^{2^{k-C-1}\delta^2\epsilon^2}. \tag{7}$$

Finally, note that by adding to the class $\mathcal{D}$ all negations (functions $\mathsf{D}'(x) = 1 - \mathsf{D}(x)$) we have $\mathbb{E}\,\mathsf{D}(X) \leqslant \mathbb{E}\,\mathsf{D}(U_{\mathcal{S}}) + \delta\epsilon$ as well as $\mathbb{E}\,\mathsf{D}(X) \geqslant \mathbb{E}\,\mathsf{D}(U_{\mathcal{S}}) - \delta\epsilon$, for every $\mathsf{D} \in \mathcal{D}$. In particular, we have

$$\forall \mathsf{D} \in \mathcal{D}: |\mathbb{E}\,\mathsf{D}(X) - \mathbb{E}\,\mathsf{D}(U_{\mathcal{S}})| < \delta\epsilon \tag{8}$$

as long as

$$4|\mathcal{D}| < 2^{2^{k-C-1}\delta^2\epsilon^2}. \tag{9}$$

It remains to observe that for every $\mathcal{X}$ the probability mass function of X takes two values on two halves of $\mathcal{X}$.