A Tight Transformation between HILL and Metric Conditional Pseudoentropy[⋆]

Maciej Skorski[⋆⋆]
maciej.skorski@gmail.com

Cryptology and Data Security Group, University of Warsaw

Abstract. HILL Entropy and Metric Entropy are generalizations of the information-theoretic notion of min-entropy to the realistic setting where adversaries are computationally bounded.
The notion of HILL Entropy appeared in the breakthrough construction of a PRG from any one-way function (Håstad et al.), and has become the most important and most widely used variant of computational entropy.
In turn, Metric Entropy defined as a relaxation of HILL Entropy, has been proven to be much easier to handle, in particular in the context of computational generalizations of the Green-Tao-Ziegler Dense Model Theorem which find applications in leakage-resilient cryptography, memory delegation or deterministic encryption.
Fortunately, Metric Entropy can be converted, with some loss in quality, to HILL Entropy as shown by Barak, Shaltiel and Wigderson.
In this paper we improve their result, reducing the loss in quality of entropy. Our bound is tight and, interestingly, independent of size of the probability space. As an interesting example of application we derive the computational dense model theorem with best possible parameters.
Our approach is based on the theory of convex approximation in L^p-spaces.

Keywords: Pseudoentropy, Dense Model Theorem, Convex Approximation

1 Introduction

1.1 Computational Entropy

The Idea of Computational Entropy. The notion of entropy, as a measure of randomness, is a fundamental concept in information-theory. The need for computational versions of entropy comes from the fact that security definitions based on classical information theoretic entropy notions are quite often

[⋆] This work appears as a part of the paper “Metric Pseudoentropy: Characterizations, Transformations and Applications ”, to appear at ICITS 2015. Preliminary versions of this work appeared in the Proceedings of Student Research Forum Papers and Posters at SOFSEM 2015.

[⋆⋆] This work was partly supported by the WELCOME/2010-4/2 grant founded within the framework of the EU Innovative Economy Operational Programme.

too strong for "practical" purposes, where resources of adversaries are limited. The distribution which is not perfectly random might look random from the computational point of view, when finding the real difference is simply inefficient. The metrics used to quantify the amount and quality of pseudorandomness are called computational entropies. Computational notions of entropy find important applications in leakage-resilient cryptography [DP08], deterministic encryption [FOR12], memory delegation [CKLR11], computational complexity [RTTV08] and foundations of cryptography [HRV10].

HILL Entropy and Metric Entropy. The most popular approach to extend information-theoretic notions of entropy into computational case, is based on the notion of computational indistinguishability. We discuss it briefly below.

Given a class of $[0,1]$-valued functions $\mathcal{D}$ and a parameter ϵ we say that two n-bit random variables X and Y are $(\mathcal{D},\epsilon)$-indistinguishable if no function D from $\mathcal{D}$ can distinguish X and Y with the advantage better than ϵ[1]. Formally

$$\mathbf{E}\,\mathrm{D}(X) - \mathbf{E}\,\mathrm{D}(Y) \leqslant \epsilon \quad \text{for all } \mathrm{D} \in \mathcal{D}$$

If $\mathcal{D}$ is the class of all circuits of size at most s we slightly abbreviate this notation and say that X and Y are (s,ϵ) indistinguishable.

In theoretical computer science and especially in cryptography the most popular notion of entropy is min-entropy (in contrast to information-theory where one uses extensively Shannon entropy), because it measures randomness in terms of "hardness" of predictability. The min entropy of X given (possibly) Z is at least 2^{-k} if one cannot predict X given $Z = z$, on average[2] over z, better than with probability 2^{-k}. Formally, for X and Z one defines the average min-entropy [DORS08] of X given Z as follows

$$\widetilde{\mathbf{H}}_{\infty}\left(X|Z\right) \geqslant k \text{ iff } \mathop{\mathbf{E}}_{z\leftarrow Z}\left[\max_{x} \Pr[X = x|Z = z]\right] \leqslant 2^{-k}.$$

Based on the concept of indistinguishability and min-entropy, one defines the computational HILL entropy [HILL99] of X as the maximum amount of min-entropy in a random variable Y (taking values in the same set as X) which is indistinguishable from X. This idea was extended to the conditional case (in the presence of side information) in [HLR07]. We can state the definition as follows:

> $\mathbf{H}_{\infty}^{\mathrm{HILL},(s,\epsilon)}\left(X|Z\right) \geqslant k$ if and only if there exists Y jointly distributed with Z of average min-entropy (given Z) at least k such that $|\,\mathbf{E}\,\mathrm{D}(X,Z) - \mathbf{E}\,\mathrm{D}(Y,Z)| \leqslant \epsilon$ for all circuits D of size at most s.

Note that this captures the standard notion of a pseudorandom distribution (for $k = n$). By *switching the order of quantifiers* and restricting distinguishers to

[1] It can be shown, that in the experiment when one needs to guess whether a given sample is either from X or from Y, by the use of D one gets advantage $|\,\mathbf{E}\,\mathrm{D}(X) - \mathbf{E}\,\mathrm{D}(Y)|$ over a random guess.

[2] Sometimes one uses the stronger notion, called the worst-case min-entropy, when we require the same upper bound on guessing probability for every auxiliary input z.

deterministic circuits one obtains a slightly weaker version, called computational metric entropy [BSW03, DP08]

> $\mathbf{H}_{\infty}^{\mathrm{M},\det[0,1],(s,\epsilon)}(X|Z) \geqslant k$ iff for every *deterministic* $[0,1]$-*valued* circuits D of size at most s there exists Y jointly distributed with Z of average min-entropy (given Z) at least k such that $|\mathbf{E}\,\mathrm{D}(X,Z) - \mathbf{E}\,\mathrm{D}(Y,Z)| \leqslant \epsilon$.

In both definitions the parameters s, ϵ quantify the quality of pseudorandomness: the bigger s and the smaller ϵ, the higher quality is. Metric entropy is known to be equivalent to HILL entropy with the same amount and some loss in quality parameters [BSW03, CKLR11].

Advantages of metric entropy and applications. There are very good reasons to introduce and study metric entropy: quite often it is much easier to prove a statement for metric-entropy and then pass to the HILL version. Actually, this strategy is unavoidable for the standard proof technique which uses the min-max theorem to switch the order of players in a game. Therefore, many facts on HILL entropy uses metric entropy explicitly or implicitly [VZ12, FOR12, CKLR11, RTTV08, DP08, BSW03]. Perhaps the most spectacular example is the efficient version of the Dense Model Theorem [RTTV08, DP08], being the key ingredient of the famous result of Tao and Ziegler on primes in arithmetic progressions [TZ08]. The efficient version, which found many interesting applications in complexity theory, was originally proved using the idea of metric computational entropy in [RTTV08] and independently in [DP08]. A much simpler proof achieving significant improvements in quality was given later in [FR12]. It uses only very basic properties of Metric entropy!

Conversions between HILL and metric entropy. The following result, due to Barak, Shaltiel and Widgerson[3], states that metric and HILL computational entropy are equivalent up to some loss in quality

Theorem 1 (From Metric to HILL Entropy [BSW03]). *Let X and Z be, respectively, n-bit and m-bit correlated random variables. Then*

$$\mathbf{H}^{\mathrm{HILL},(s',\epsilon')}(X|Z) \geqslant \mathbf{H}^{\mathrm{M},\det[0,1],(s,\epsilon)}(X|Z)$$

where $s' = \Omega\left(s \cdot \delta^2/(n+m)\right)$ and $\epsilon' = \epsilon + \delta$ for arbitrary $\delta \in (0,1)$.

Remark 1. Since we have $\mathbf{H}^{\mathrm{HILL},(s,\epsilon)}(X|Z) \leqslant \mathbf{H}^{\mathrm{M},\det[0,1],(s,\epsilon)}(X|Z)$, the conversion in the other direction is lossless.

1.2 Our Contribution

Our result. We improve Theorem 1 in the following way:

[3] In [BSW03] it was stated only for unconditional entropy, however the generalization below is straightforward, see for example [CKLR11] and [FOR12]

Theorem 2 (Dimension-independent transformation between metric and HILL entropy). *For any n-bit random variable X and a correlated random variable Z we have*

$$\mathbf{H}^{\mathrm{HILL},(s',\epsilon')}\left(X|Z\right) \geqslant \mathbf{H}^{\mathrm{M},\det[0,1],(s,\epsilon)}\left(X|Z\right)$$

where $\delta \in (0,1)$ is an arbitrary parameter, $s' = \Omega\left(s \cdot \delta^2/(\Delta+1)\right)$, $\epsilon' = \epsilon + \delta$ and $\Delta = n - k$ is the entropy deficiency.

Remark 2 (Tightness of our bound). Our bound is tight up to a constant factor because it does not depend on the dimension of Z and, in the uncoditional case, implies best possible Dense Model Theorem, as show in Section 4.

In comparison to Theorem 1 we replace the factor $n + m$ by $\Delta + 1$. Our result shows that the conversion does not depended on the dimension of the domain of X and Z but *only* on the entropy deficiency Δ.

Applications. At first glance, our result might seem to be not really better than Theorem 1 as it does not offer significant improvement in the asymptotic setting. However, there are some applications where is has significant advantages. First, it might be of some interest in case when m is much longer than n (see for instance the equivalence between HILL and unpredictability entropy of X given Z for short X [VZ12], or models that captures possibly long leakages [DDF14]) or when the deficiency Δ is very small. As a concrete application, we show that our result implies the efficient dense model theorem with the best possible parameters due to Zhang [Zha11], which is not the case of the bound in Theorem 1. The proof, based on basic properties of metric entropy, is super-simple.

Our techniques. Our results might be interesting because of the novel proof technique: instead of using Chernoff Bounds for approximating convex hulls with *uniformly* small error as in [BSW03], we show that it is enough to do the approximation with respect to the p-th norm induced by some appropriately chosen measure, and optimize the value of p. There is a lot of research focused on achieving better rates of convex approximations in L^p-spaces for some restricted class of functions. In case of the metric-to-HILL transformation (or a similar result) it might be possible to obtain some further improvements for restricting classes of adversaries.

1.3 Organization of the paper

In Section 2 we explain basic notions and provide necessary definitions. The proof of our main technical result together with an improved Metric-to-HILL transformation appears in Section 3. In Section 4 we demonstrate an interesting application: a simple alternative proof of the Dense Model Theorem which achieves the best possible parameters.

2 Preliminaries

Probabilities, measures and integrals. By μ_X or $\mathbf{P}_X$ we denote the probability mass function (distribution) of X, that is $\mu_X(x) = \mathbf{P}_X(x) = \Pr[X = x]$ for all x. A measure ν on a finite set Ω is a function $\mu : \Omega \to \mathbb{R}^+ \cup \{0\}$. For notation convenience, we use the signs of sums and integrals interchangeably. The integral of a function D on E with respect to a measure ν is defined as $\int_E D\mathrm{d}\nu = \sum_{x\in E} D(x)\nu(x)$. For the integral over the entire domain we omit the subscript E.

L^p spaces. Given a finite set Ω and a measure μ on Ω one defines the p-th norm of a real-valued function D defined on Ω as $\|D\|_p = \int_\Omega D\mathrm{d}\mu$

Convex combinations. Given a set of real-valued functions $\mathcal{C}$ defined on the same domain, by $\mathrm{conv}_t(\mathcal{C})$ we denote the set of all convex combinations of length at most t of members of $\mathcal{C}$. That is,

$$\mathcal{C}_t = \left\{ \sum_{i=1}^{t} \alpha_i D_i : \ \sum_{i=1}^{t} \alpha_i = 1,\ \alpha_i \geqslant 0 \text{ for } i = 1,\ldots,t,\ D_i \in \mathcal{C} \text{ for } i = 1,\ldots,t \right\}$$

Computational Entropy notions.

Definition 1 (Conditional HILL Pseudoentropy [HLR07]). *Let X, Z be a joint distribution with the following property: there exists Y of conditional min-entropy at least k given Z such that for all circuits D of size at most s we have $|\mathbf{E}\,\mathrm{D}(X,Z) - \mathbf{E}\,\mathrm{D}(Y,Z)| \leqslant \epsilon$. Then we say that X given Z has k bits of HILL min-entropy of quality (s,ϵ) and denote by $\mathbf{H}_\infty^{\mathrm{HILL},(s,\epsilon)}(X|Z) \geqslant k$.*

Remark 3 (HILL entropy against different circuits classes). For conditional HILL entropy all kinds of circuits: deterministic boolean, deterministic real valued and randomized boolean (for the same size s), are equivalent [FR12].

Definition 2 (Conditional Metric Pseudoentropy [DP08]). *Let X, Z be a joint distribution with the following property: for every deterministic boolean (respectively: deterministic real valued or boolean randomized) circuit D of size at most s there exists Y of (conditional min entropy at least k given Z such that $|\mathbf{E}\,\mathrm{D}(X,Z;Y,Z) - \mathbf{E}\,\mathrm{D}(Y,Z)| \leqslant \epsilon$. Then we say that X given Z has k bits of deterministic (respectively: deterministic real valued or boolean randomized) metric min-entropy of quality (s,ϵ) and denote by $\mathbf{H}_\infty^{\mathrm{M},\det\{0,1\},(s,\epsilon)}(X|Z)$ (respectively: $\mathbf{H}_\infty^{\mathrm{M},\det[0,1],(s,\epsilon)}(X|Z)$ and $\mathbf{H}_\infty^{\mathrm{M},\mathrm{rand}\{0,1\},(s,\epsilon)}(X|Z)$).*

3 Main Result

In this section we prove our main technical result which inmediately implies Theorem 2. It is a constrained version of the standard approximation result.

Lemma 1 (Approximating long convex combinations with respect to all high-min-entropy distributions). *Let X be an n-bit random variable, be Z be a correlated m-bit random variable, and let $\mathcal{C}$ be a class of $[0,1]$-valued function on $\{0,1\}^n \times \{0,1\}^m$. Let $D \in \mathrm{conv}(\mathcal{C})$. Then for $\ell = 49(n+1-k)/\delta^2$ there exists $D_\ell \in \mathrm{conv}_\ell(\mathcal{C})$ such that*

$$\mathbf{E}\left|\mathrm{D}(X) - \mathrm{D}_\ell(X)\right| \leqslant \delta \tag{1}$$

and simultaneously

$$\mathbf{E}\left|\mathrm{D}(X) - \mathrm{D}_\ell(Y)\right| \leqslant \delta \tag{2}$$

for every distribution Y jointly distributed with Z such that $\mathbf{H}_\infty(Y|Z) \geqslant k$.

Corollary 1. *Lemma 1 implies Theorem 2*

Proof (of Corollary 1). If $\mathbf{H}_\infty^{\mathrm{HILL},(s',\epsilon')}(X|Z) < k$ then for every Y satisfying $\widetilde{\mathbf{H}}_\infty(Y|Z) \geqslant k$ we find D of size at most s' such that $|\,\mathbf{E}\,\mathrm{D}(X,Z) - \mathbf{E}\,\mathrm{D}(Y,Z)| \geqslant \epsilon'$. Replacing D by D^c of necessary we can assume that $\mathbf{E}\,\mathrm{D}(X,Z) - \mathbf{E}\,\mathrm{D}(Y,Z) \geqslant \epsilon$ for some D of size $s'+1$. By applying the min-max theorem we get that there exists a convex combination D' of circuits of size at most $s'+1$ such that

$$\mathbf{E}\,\mathrm{D}(X,Z) - \mathbf{E}\,\mathrm{D}(Y,Z) \geqslant \epsilon' \quad \forall Y:\ \widetilde{\mathbf{H}}_\infty(Y|Z) \geqslant k$$

That combination might be very long. But applying Lemma 1 we can approximate it by a combination D' of at most $O\left((n+1-k)/\delta^2\right)$ circuits of size $s'+1$ in such a way that the expectations with respect to X,Z and Y,Z differs at most by $\delta/2$. This way we obtain

$$\mathbf{E}\,\mathrm{D}'(X,Z) - \mathbf{E}\,\mathrm{D}'(Y,Z) \geqslant \epsilon' - 2\cdot\delta/2 \quad \forall Y:\ \widetilde{\mathbf{H}}_\infty(Y|Z) \geqslant k$$

which finishes the proof. □

Now we prove our main approximation result

Proof (of Lemma 1). Consider the space of all functions on $\{0,1\}^{n+m}$. We start by the following trivial observation

Claim 1. It suffices to show that for some $\mathrm{D}' \in \mathrm{conv}_\ell(\mathcal{C})$ we have $\int|\mathrm{D}-\mathrm{D}'|\cdot \mathrm{d}(\mu_X + \mu_Y) \leqslant \delta$ for all Y such that $\widetilde{\mathbf{H}}_\infty(Y|Z) \geqslant k$.

By applying the Hölder Inequality, we immediately get

Claim 2. For every functions D, D' and every $p, q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$ we have

$$\int|\mathrm{D}-\mathrm{D}'|\cdot\mathrm{d}(\mu_X+\mu_Y) \leqslant \|\mathrm{D}-\mathrm{D}'\|_p \cdot \left\|\frac{\mu_{X,Z}+\mu_{Y,Z}}{\mu}\right\|_q \tag{3}$$

Now we give estimates on both factors on the right hand side of Equation (3).

Claim 3. If $q \in [1,2]$ then for any Y such that $\widetilde{\mathbf{H}}_\infty\left(Y|Z\right) \geqslant k$ we have

$$\left\| \frac{\mu_{X,Z} + \mu_{Y,Z}}{\mu} \right\|_q \leqslant \left(2^q + 2^{(q-1)(n+1-k)} \right)^{1/q} \tag{4}$$

Proof (Of Claim 3). Recall the well-known inequality

Proposition 1. *If $a, b > 0 and$ $q \geqslant 1$ then $(a+b)^q \leqslant 2^{q-1}(a^q + b^q)$.*

From Proposition 1 it follows now that

$$\left\| \frac{\mu_{X,Z} + \mu_{Y,Z}}{\mu} \right\|_q \leqslant 2^{q-1} \left(\left\| \frac{\mu_{X,Z}}{\mu} \right\|_q + \left\| \frac{\mu_{Y,Z}}{\mu} \right\|_q \right) \tag{5}$$

We shall estimate two terms in Equation (4) separately. Since $\mu_{X,Z}(x,z) < \mu_{X,Z}(x,z) + \mu_{U,Z}(x,z) = \mu(x,z)$ for all x, z we have

$$\left\| \frac{\mu_{X,Z}}{\mu} \right\|_q < \int 1 \mathrm{d}\mu = 2 \tag{6}$$

To bound the second term note that the functional $\mu_{Y,Z} \rightarrow \left\| \frac{\mu_{X,Z} + \mu_{Y,Z}}{\mu} \right\|_q$ is convex as a function of $\mu_{Y,Z}$ (being a composition of an affine function and the p-th norm). Therefore, the maximum among all distributions Y, Z satisfying $\widetilde{\mathbf{H}}_\infty\left(Y|Z\right) \geqslant k$, which form a convex set, is attained at an extreme point. This means that the maximum is attained for a distribution (Y^*, Z) such that the distribution $Y^*|_{Z=z}$ is flat for every z and the conditional min-entropy of Y given Z is exactly k. Since $\mu(x,z) = \mu_U(x)\mu_Z(z)$ and $\mu_{Y^*,Z}(x,z) = \mu_{Y^*|_{Z=z}}(x)\mu_Z(z)$ we obtain

$$\begin{aligned}
\left\| \frac{\mu_{Y,Z}}{\mu} \right\|_q^q &\leqslant \int \left(\frac{\mu_{Y^*,Z}}{\mu} \right)^q \mathrm{d}\mu \\
&= \int \left(\int \left(\frac{\mu_{Y^*_{Z=z}}}{\mu_U} \right)^q \mathrm{d}\mu_U \right) \mathrm{d}\mu_Z \\
&= \int \left(2^{(q-1)(n - \mathbf{H}_\infty\left(Y^*|Z=z\right)} \right) \mathrm{d}\mu_Z \\
&= 2^{(q-1)n} \int 2^{-(q-1)\,\mathbf{H}_\infty\left(Y^*|Z=z\right)} \mathrm{d}\mu_Z
\end{aligned}$$

By applying the Jensen Inequality to the function $u \rightarrow u^{q-1}$ (which is concave by the assumption on q) we get

$$\begin{aligned}
\left\| \frac{\mu_{Y,Z}}{\mu} \right\|_q^q &\leqslant 2^{(q-1)n} \left(\int 2^{-\mathbf{H}_\infty\left(Y^*|Z=z\right)} \mathrm{d}\mu_Z \right)^{q-1} \\
&\leqslant 2^{(q-1)n} \left(2^{-\widetilde{\mathbf{H}}_\infty\left(Y|Z\right)} \right)^{q-1} = 2^{(q-1)(n-k)}
\end{aligned} \tag{7}$$

Plugin Equation (7) and Equation (6) into Equation (5) yields

$$\left\|\frac{\mu_{X,Z}+\mu_{Y,Z}}{\mu}\right\|_q^q \leqslant 2^{q-1}\left(2+2^{(q-1)(n-k)}\right) = 2^q + 2^{(q-1)(n+1-k)}.$$

and Equation (4) follows. □

Claim 4. Suppose that $p \geqslant 2$. Then for any $\mathrm{D} \in \mathrm{conv}(\mathcal{C})$ and $\ell \geqslant 1$ there exists $\mathrm{D}_\ell \in \mathrm{conv}_\ell(D)$ such that $\|\mathrm{D}-\mathrm{D}_\ell\|_p < 1.74\sqrt{p/\ell}$.

Proof. The proof relies on the following approximation result on rates of convex approximation, which generalizes the famous Maurey-Johnes-Barron Theorem.

Lemma 2 (Convex approximation in L^p spaces [DDGS97]). *Let E be an L^p space with $1 \leqslant p < +\infty$. Suppose that $S \subset E$, $f \in \mathrm{conv}(S)$ and let $K > 0$ be such that for all $g \in S$ we have $\|g-f\|_p \leqslant K$. Then for any ℓ we have*

$$\min_{s\in\mathrm{conv}_\ell(S)} \|f-s\|_p \leqslant \frac{KC_p}{\ell^{1-\frac{1}{t}}}$$

where $t = \min(2,p)$ and $C_p = 1$ if $1 \leqslant p \leqslant 2$, $C_p = \sqrt{2}[\Gamma((p+1)/2)/\sqrt{\pi}]^{1/p}$ for $2 < p < +\infty$.

Remark 4. The constant C_p can be estimated using the following approximation for the gamma function [Mor11], valid for $x \geqslant 1$:

$$\sqrt{\pi}(x/\mathrm{e})^x\sqrt{2x+0.33} < \Gamma(x+1) < \sqrt{\pi}(x/\mathrm{e})^x\sqrt{2x+0.36}$$

From this we find that $C_p < 0.87\sqrt{p}$ for all $p > 2$.

The claim follows by setting E to be the space of $[0,1]$-valued functions on $\{0,1\}^n \times \{0,1\}^m$ and $K = \int 1\mathrm{d}\mu = 2$. □

By Claim 3 and Claim 4 combined with Claim 2 and Claim 1 it suffices to find $p \geqslant 2$ (which automatically ensures $q \in [1,2]$) and ℓ such that

$$1.74\sqrt{p/\ell}\cdot\left(2^q + 2^{(q-1)(n+1-k)}\right)^{1/q} \leqslant \delta.$$

If $k \geqslant n-1$ then we put $p = q = 2$. Then it suffices to ensure that $1.74\sqrt{2/\ell}(2^2 + 2^2)^{1/2} \leqslant \delta$ which is equivalent to $6.96\sqrt{\ell} \leqslant \delta$. Suppose that $k \leqslant n-1$. By the inequality $(a+b)^r \leqslant a^r + b^r$ valid for $a, b > 0$ and $0 < r \leqslant 1$, we see that it suffices if $1.74\sqrt{p/\ell}\left(2+2^{(n+1-k)/p)}\right) \leqslant \delta$. For $p = n+1-k$ we obtain $6.96\sqrt{\ell} \leqslant \delta$. This finishes the proof. □

4 Application to the Dense Model Theorem

Dense Model Theorem. Given a pair of two distributions W and V over the same finite domain we say that W is δ-dense in V if and only if $\Pr[W = x] \leqslant \Pr[V = x]/\delta$[4]. The dense model theorem [TZ08], specialized to boolean distinguishers, can be formulated as follows:

Theorem 3 (Dense Model Theorem.). *Let $\mathcal{D}'$ be a class of n-bit boolean functions, R be uniform over $\{0,1\}^n$, X be an n-bit random variable and let X' be δ-dense in X. If X and R are $(\mathcal{D}, \epsilon)$-indistinguishable then there exists a distribution R' which is δ-dense in R such that X' and R' are $(\mathcal{D}', \epsilon')$-indistinguishable, where $\epsilon' = (\epsilon/\delta)^{O(1)}$ and $\mathcal{D}$ consists of all functions of the form $g(\mathrm{D}_1, \ldots, \mathrm{D}_\ell)$ where $\mathrm{D}_i \in \mathcal{D}'$, $\ell = \mathrm{poly}(1/\delta, 1/\epsilon)$ and g is some function.*

Informally, this statement reads as follows: if a distribution X' is dense in a pseudorandom distribution X, then X' must be indistinguishable from a distribution dense in the uniform distribution. Note that the indistinguishability parameters for X' are worse than for X: to achieve $(\mathcal{D}', \epsilon')$-indistinguishably we need to start with ϵ smaller than ϵ' and a class $\mathcal{D}$ sufficiently more complicated than $\mathcal{D}'$. Note also that for this statement to be computationally meaningful we need g to be efficient.

Applications of the Dense Model Theorem. Efficient versions of the Dense Model Theorem have found applications in differential privacy [MPRV09], pseudoentropy and leakage-resilient cryptography [DP08, CKLR11], graph decompositions [RTTV08], and some further applications in additive combinatorics [GW11]. We refer the reader to [Tre11] for a survey.

Comparison of different formulations. Below we compare the different versions of the Dense Model Theorem. We note that some of them are conclusions from more general statements about pseudoentropy and that the proof technique which utilizes properties of pseudoentropy is due to [DP08].

Author	Technique	Function g	ℓ as complexity of $\mathcal{D}'$ w.r.t $\mathcal{D}$	ϵ' vs ϵ
[TZ08]	Complicated	Inefficient	$\ell = \mathsf{poly}(1/(\epsilon/\delta), \log(1/\delta))$	$\epsilon' = O(\epsilon/\delta)$
[RTTV08, Gow08]	Min-Max Theorem	Linear threshold	$\ell = O(\log(1/\epsilon)/(\epsilon/\delta)^2)$	$\epsilon' = O(\epsilon/\delta)$
[FR12], [DP08]	Metric Entropy	Linear threshold	$\ell = O(n/(\epsilon/\delta)^2)$	$\epsilon' = O(\epsilon/\delta)$
[Zha11]	Boosting	Linear threshold	$\ell = O(\log(1/\delta)/(\epsilon/\delta)^2)$	$\epsilon' = O(\epsilon/\delta)$
This paper	Metric Entropy	Linear threshold	$\ell = O(\log(1/\delta)/(\epsilon/\delta)^2)$	$\epsilon' = O(\epsilon/\delta)$

Table 1: The quantitative comparison of versions of the Dense Model Theorem

Below we show how to derive from our Lemma 1 the optimal Dense Model Theorem due to Zhang.

[4] The term "δ-dense" comes from the fact that V can be written as a convex combination of W with weight δ and some other distribution with weight $1 - \delta$

Corollary 2. *Dense Model Theorem (Theorem 3) holds with $\epsilon' = O(\epsilon/\delta)$, g being a linear threshold and $\ell = O(\log(1/\delta)/(\epsilon/\delta)^2$.*

Proof. We show how to reduce the formulation of the Dense Model Theorem to the statement about HILL entropy. We start by the following observation:

Claim 5. X' is δ-dense in X if and only if X' can be written as $X|A$ for some event A of probability δ.

Proof. of Claim Consider a random variable $A \in \{0,1\}$ jointly distributed with X as follows: $\Pr[X = x, A = 1] = \delta \Pr[X']$. By the assumption on X and X' we have $\Pr[X = x, A = 1] \leqslant 1$ and thus this distribution is well defined, in particular we have $\Pr[A = 1] = \delta$ and $\Pr[X|A = 1] = \Pr[X']$. In the other hand if we have $X' \overset{d}{=} X|A$ then $\Pr[X' = x] = \Pr[X = x, A]/\Pr[A] \leqslant \Pr[X = x] \leqslant \Pr[X = x]/\Pr[A]$ and hence X' is $\Pr[A]$-dense in X. □

The second fact we need is the so called leakage lemma for metric-entropy

Lemma 3 ([FOR12], reformulated). *Let X be a random variable, A be an event of probability δ, and let $\mathcal{D}$ be a class of $[0,1]$-valued functions. Suppose that there exists D such that $\mathbf{E}\,\mathrm{D}(X|A) - \mathbf{E}\,\mathrm{D}(Y) \geqslant \epsilon'$ for all Y of min-entropy at least $k - \log(1/\Pr[A])$ and $\epsilon' = \epsilon/\Pr[A]$. Then there exists a a function D' being a threshold of some $\mathrm{D} \in \mathcal{D}$ (or its complement) such that $\mathbf{E}\,\mathrm{D}'(X) - \mathbf{E}\,\mathrm{D}'(Y) \geqslant \epsilon$.*

The name "leakage lemma" is because for $s' \approx s$ the lemma implies

$$\mathbf{H}_{\infty}^{\mathrm{M},\det[0,1],(s,\epsilon)}(X|A) \geqslant \mathbf{H}_{\infty}^{\mathrm{M},\mathcal{D},(s'\epsilon/\Pr[A])}(X) - \log(1/\Pr[A]).$$

Now we are ready to give the proof. Suppose contrary, that the Dense Model Theorem is not true with the claimed parameters. Then for some event A of probability δ, some ϵ' and every distribution Y of min-entropy $n - \log(1/\delta)$ (which is equivalent to be δ-dense in the uniform distribution) there exists $\mathrm{D} \in \mathcal{D}$ or $D \in \mathbf{1} - \mathrm{D} \in \mathcal{D}$ such that

$$\mathbf{E}\,D(X|A) - \mathbf{E}\,D(Y) \geqslant \epsilon'$$

By applying the min-max theorem we get that there exists a long convex combination $\bar{\mathrm{D}}$ of functions from $\mathcal{D} \cup (\mathbf{1} - \mathcal{D})$ such that

$$\mathbf{E}\,\bar{\mathrm{D}}(X|A) - \mathbf{E}\,\bar{D}(Y) \geqslant \epsilon' \quad \forall Y: \ \mathbf{H}_{\infty}(Y) \geqslant n - \log(1/\delta).$$

Now we use our Lemma 1, with the class $\mathcal{D} \cup (\mathbf{1} - \mathcal{D})$ and δ replaced by $\epsilon'/3$, approximating $\bar{\mathrm{D}}$ by a convex combination D' of length $\ell = O\left(\log(1/\delta)/\epsilon'^2\right)$. Then we get

$$\mathbf{E}\,\mathrm{D}'(X|A) - \mathbf{E}\,\mathrm{D}'(Y) \geqslant \epsilon' \quad \forall Y: \ \mathbf{H}_{\infty}(Y) \geqslant n - \log(1/\delta).$$

Note that D' is a linear threshold of ℓ functions from $\mathcal{D}$. By Lemma 3 we replace D' by D'' which is again a linear threshold of ℓ functions from $\mathcal{D}$ and satisfies

$$\mathbf{E}\,\mathrm{D}''(X) - \mathbf{E}\,\mathrm{D}''(Y) \geqslant \epsilon' \quad \forall Y: \ \mathbf{H}_{\infty}(Y) \geqslant n.$$

Hence, we get a contradiction. □

5 Conclusion

In this paper we improve the transformation between conditional Metric and HILL entropy, by replacing the dimension factor by the entropy deficiency. This result immediately implies the best efficient version of the Dense Model Theorem.

6 Acknowledgments

I would like to thank Krzysztof Pietrzak for helpful discussions.

References

BSW03. Boaz Barak, Ronen Shaltiel, and Avi Wigderson, *Computational analogues of entropy.*, RANDOM-APPROX (Sanjeev Arora, Klaus Jansen, Jos D. P. Rolim, and Amit Sahai, eds.), 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 (Phillip Rogaway, ed.), Lecture Notes in Computer Science, vol. 6841, Springer, 2011, pp. 151–168.

DDF14. Alexandre Duc, Stefan Dziembowski, and Sebastian Faust, *Unifying leakage models: From probing attacks to noisy leakage*, Advances in Cryptology - EUROCRYPT 2014 - 33rd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Copenhagen, Denmark, May 11-15, 2014. Proceedings (Phong Q. Nguyen and Elisabeth Oswald, eds.), Lecture Notes in Computer Science, vol. 8441, Springer, 2014, pp. 423–440.

DDGS97. M.J. Donahue, C. Darken, L. Gurvits, and E. Sontag, *Rates of convex approximation in non-hilbert spaces*, Constructive Approximation **13** (1997), no. 2, 187–220 (English).

DORS08. Yevgeniy Dodis, Rafail Ostrovsky, Leonid Reyzin, and Adam Smith, *Fuzzy extractors: How to generate strong keys from biometrics and other noisy data*, SIAM J. Comput. **38** (2008), no. 1, 97–139.

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

FOR12. Benjamin Fuller, Adam O'Neill, and Leonid Reyzin, *A unified approach to deterministic encryption: New constructions and a connection to computational entropy*, Theory of Cryptography - 9th Theory of Cryptography Conference, TCC 2012, Taormina, Sicily, Italy, March 19-21, 2012. Proceedings (Ronald Cramer, ed.), Lecture Notes in Computer Science, vol. 7194, Springer, 2012, pp. 582–599.

FR12. Benjamin Fuller and Leonid Reyzin, *Computational entropy and information leakage*, Cryptology ePrint Archive, Report 2012/466, 2012, http://eprint.iacr.org/.

Gow08. W. T. Gowers, *Decompositions, approximate structure, transference, and the Hahn-Banach theorem*, ArXiv e-prints (2008).

GW11. W.T. Gowers and J. Wolf, *Linear forms and higher-degree uniformity for functions on $\mathbb{F}_p^n$*, Geometric and Functional Analysis **21** (2011), no. 1, 36–69 (English).

HILL99. Johan Hastad, Russell Impagliazzo, Leonid A. Levin, and Michael Luby, *A pseudorandom generator from any one-way function*, SIAM J. Comput. **28** (1999), no. 4, 1364–1396.

HLR07. Chun-Yuan Hsiao, Chi-Jen Lu, and Leonid Reyzin, *Conditional computational entropy, or toward separating pseudoentropy from compressibility*, Proceedings of the 26th annual international conference on Advances in Cryptology (Berlin, Heidelberg), EUROCRYPT '07, Springer-Verlag, 2007, pp. 169–186.

HRV10. Iftach Haitner, Omer Reingold, and Salil Vadhan, *Efficiency improvements in constructing pseudorandom generators from one-way functions*, Proceedings of the 42nd ACM symposium on Theory of computing (New York, NY, USA), STOC '10, ACM, 2010, pp. 437–446.

Mor11. Christinel Mortici, Journal of Mathematical Inequalities **5** (2011), no. 4, 611–614.

MPRV09. Ilya Mironov, Omkant Pandey, Omer Reingold, and Salil Vadhan, *Computational differential privacy*, Proceedings of the 29th Annual International Cryptology Conference on Advances in Cryptology (Berlin, Heidelberg), CRYPTO '09, Springer-Verlag, 2009, pp. 126–142.

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.

Tre11. Luca Trevisan, *Dense model theorems and their applications*, Proceedings of the 8th Conference on Theory of Cryptography (Berlin, Heidelberg), TCC'11, Springer-Verlag, 2011, pp. 55–57.

TZ08. Terence Tao and Tamar Ziegler, *The primes contain arbitrarily long polynomial progressions*, Acta Mathematica **201** (2008), no. 2, 213–305 (English).

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.

Zha11. Jiapeng Zhang, *On the query complexity for showing dense model*, Electronic Colloquium on Computational Complexity (ECCC) **18** (2011), 38.