Condensed Unpredictability

Maciej Skórski[1⋆], Alexander Golovnev[2], and Krzysztof Pietrzak[3⋆⋆]

[1] University of Warsaw maciej.skorski@gmail.com
[2] New York University alexgolovnev@gmail.com
[3] IST Austria pietrzak@ist.ac.at

Abstract. We consider the task of deriving a key with high HILL entropy (i.e., being computationally indistinguishable from a key with high min-entropy) from an unpredictable source.
Previous to this work, the only known way to transform unpredictability into a key that was ϵ indistinguishable from having min-entropy was via pseudorandomness, for example by Goldreich-Levin (GL) hardcore bits. This approach has the inherent limitation that from a source with k bits of unpredictability entropy one can derive a key of length (and thus HILL entropy) at most $k - 2\log(1/\epsilon)$ bits. In many settings, e.g. when dealing with biometric data, such a $2\log(1/\epsilon)$ bit entropy loss in not an option.
Our main technical contribution is a theorem that states that in the high entropy regime, unpredictability implies HILL entropy. Concretely, any variable K with $|K| - d$ bits of unpredictability entropy has the same amount of so called metric entropy (against real-valued, deterministic distinguishers), which is known to imply the same amount of HILL entropy. The loss in circuit size in this argument is exponential in the entropy gap d, and thus this result only applies for small d (i.e., where the size of distinguishers considered is exponential in d).
To overcome the above restriction, we investigate if it's possible to first "condense" unpredictability entropy and make the entropy gap small. We show that any source with k bits of unpredictability can be condensed into a source of length k with $k - 3$ bits of unpredictability entropy. Our condenser simply "abuses" the GL construction and derives a k bit key from a source with k bits of unpredicatibily. The original GL theorem implies nothing when extracting that many bits, but we show that in this regime, GL still behaves like a "condenser" for unpredictability. This result comes with two caveats (1) the loss in circuit size is exponential in k and (2) we require that the source we start with has *no* HILL entropy (equivalently, one can efficiently check if a guess is correct). We leave it as an intriguing open problem to overcome these restrictions or to prove they're inherent.

1 Introduction

Key-derivation considers the following fundamental problem: Given a joint distribution (X, Z) where $X|Z$ (which is short for "X conditioned on Z") is guaranteed to have some kind of entropy, derive a "good" key $K = h(X, S)$ from

⋆ Research supported by the WELCOME/2010-4/2 grant.
⋆⋆ Research supported by ERC starting grant (259668-PSPC).

X by means of some efficient key-derivation function h, possibly using public randomness S.

In practice, one often uses a cryptographic hash function like SHA3 as the key derivation function $h(.)$ [6, 17], and then simply assumes that $h(.)$ behaves like a random oracle [2].

In this paper we continue the investigation of key-derivation with provable security guarantees, where we don't make any computational assumption about $h(.)$. This problem is fairly well understood for sources $X|Z$ that have high min-entropy (we'll formally define all the entropy notions used in 2 below), or are computationally indistinguishable from having so (in this case, we say $X|Z$ has high HILL entropy). In the case where $X|Z$ has k bits of min-entropy, we can either use a strong extractor to derive a $k - 2\log\epsilon^{-1}$ key that is ϵ-close to uniform, or a condenser to get a k bit key which is ϵ-close to a variable with $k - \log\log\epsilon^{-1}$ bits of min-entropy. Using extractors/condensers like this also works for HILL entropy, except that now we only get computational guarantees (pseudorandom/high HILL entropy) on the derived key.

Often one has to derive a key from a source $X|Z$ which has no HILL entropy at all. The weakest assumption we can make on $X|Z$ for any kind of key-derivation to be possible, is that X is hard to predict given Z. This has been formalized in [15] by saying that $X|Z$ has k bits of unpredictability entropy, denoted $H^{\mathsf{unp}}_s(X|Z) \geqslant k$, if no circuit of size s can predict X given Z with advantage $\geqslant 2^{-k}$ (to be more general, we allow an additional parameter $\delta \geqslant 0$, and $H^{\mathsf{unp}}_{\delta,s}(X|Z) \geqslant k$ holds if (X, Z) is δ-close to some distribution (Y, Z) with $H^{\mathsf{unp}}_s(Y|Z) \geqslant k$). We will also consider a more restricted notion, where we say that $X|Z$ has k bits of *list*-unpredictability entropy, denoted $H^{*\mathsf{unp}}_s(X|Z) \geqslant k$, if it has k bits of unpredictability entropy relative to an oracle Eq which can be used to verify the correct guess (Eq outputs 1 on input X, and 0 otherwise).[4] We'll discuss this notion in more detail below. For now, let us just mention that for the important special case where it's easy to verify if a guess for X is correct (say, because we condition on $Z = f(X)$ for some one-way function[5] f), the oracle Eq does not help, and thus unpredictability and list-unpredictability coincide. The results proven in this paper imply that from a source $X|Z$ with k bits of list-unpredictability entropy, it's possible to extract a k bit key with $k - 3$ bits of HILL entropy

Proposition 1. *Consider a joint distribution (X, Z) over $\{0,1\}^n \times \{0,1\}^m$ where*

$$H^{*\mathsf{unp}}_{s,\gamma}(X|Z) \geqslant k \tag{1}$$

[4] We chose this name as having access to Eq is equivalent to being allowed to output a list of guesses. This is very similar to the well known concept of list-decoding.

[5] To be precise, this only holds for *injective* one-way functions. One can generalise list-unpredictability and let Eq output 1 on some set $\mathcal{X}$, and the adversary wins if she outputs any $X \in \mathcal{X}$. Our results (in particular Theorem 1) also hold for this more general notion, which captures general one-way functions by letting $\mathcal{X} = f^{-1}(f(X))$ be the set of all preimages of $Z = f(X)$.

Let $S \in \{0,1\}^{n \times k}$ *be uniformly random and* $K = X^T S \in \{0,1\}^k$*, then the unpredictability entropy of* K *is*

$$H^{\mathsf{unp}}_{s/2^{2k}poly(m,n),\gamma}(K|Z,S) \geqslant k-3 \tag{2}$$

and the HILL entropy of K *is*

$$H^{\mathsf{HILL}}_{t,\epsilon+\gamma}(K|Z,S) \geqslant k-3 \tag{3}$$

with[6] $t = s \cdot \frac{\epsilon^7}{2^{2k}poly(m,n)}$.

Proposition 1 follows from two results we prove in this paper.

First, in Section 4 we prove Theorem 1 which shows how to "abuse" Goldreich-Levin hardcore bits by generating a k bit key $K = X^T S$ from a source $X|Z$ with k bits of list-unpredictability. The Goldreich-Levin theorem [12] implies nothing about the pseudorandomness of $K|(Z,S)$ when extracting that many bits. Instead, we prove that GL is a good "condenser" for unpredictability entropy: if $X|Z$ has k bits of list-unpredictability entropy, then $K|(Z,S)$ has $k-3$ bits of unpredictability entropy (note that we start with list-unpredictability, but only end up with "normal" unpredictability entropy). This result is used in the first step in Proposition 1, showing that (1) implies (2).

Second, in Section 5 we prove our main result, Theorem 2 which states that any source $X|Z$ which has $|X|-d$ bits of unpredictability entropy, has the same amount of HILL entropy (technically, we show that it implies the same amount of metric entropy against deterministic real-valued distinguishers. This notion implies the same amount of HILL entropy as shown by Barak et al. [1]). The security loss in this argument is exponential in the entropy gap d. Thus, if d is very large, this argument is useless, but if we first condense unpredictability as just explained, we have a gap of only $d = 3$. This result is used in the second step in Proposition 1, showing that (2) implies (3). In the two sections below we discuss two shortcomings of Theorem 1 which we hope can be overcome in future work.[7]

On the dependency on 2^k in Theorem 1. As outlined above, our first result is Theorem 1, which shows how to condense a source with k bits of list-unpredictability into a k bit key having $k-3$ bits of unpredictability entropy. The loss in circuit size is $2^{2k}poly(m,n)$, and it's not clear if the dependency on 2^k

[6] We denote with $poly(m,n)$ some fixed polynomial in (n,m), but it can denote different polynomial throughout the paper. In particular, the *poly* here is not the same as in (2) as it hides several extra terms.

[7] After announcing this result at a workshop, we learned that Colin Jia Zheng proved a weaker version of this result. Theorem 4.18 in this PhD thesis, which is available via http://dash.harvard.edu/handle/1/11745716 also states that k bits of unpredictability imply k bits of HILL entropy. Like in our case, the loss in circuit size in his proof is polynomial in ϵ^{-1}, but it's also exponential in n (the length of X), whereas our loss is only exponential in the entropy gap $\Delta = n-k$.

is necessary here, or if one can replace the dependency on 2^k with a dependency on $poly(\epsilon^{-1})$ at the price of an extra ϵ term in the distinguishing advantage. In many settings $\log(\epsilon^{-1})$ is in the order of k, in which case the above difference is not too important. This is for example the case when considering a k bit key for a symmetric primitive like a block-cipher, where one typically assumes the hardness of the cipher to be exponential in the key-length (and thus, if we want ϵ to be in the same order, we have $\log(\epsilon^{-1}) = \Theta(k)$). In other settings, k can be superlinear in $\log(\epsilon^{-1})$, e.g., if the the high entropy string is used to generate an RSA key.

List vs. normal Unpredictability. Our Theorem 1 shows how to condense a source where $X|Z$ has k bits of *list*-unpredictability entropy into a k bit string with $k-3$ bits unpredictability entropy. It's an open question to which extent it's necessary to assume *list*-unpredictability here, maybe "normal" unpredictability is already sufficient? Note that list-unpredictability is a lower bound for unpredictability as one always can ignore the Eq oracle, i.e., $H^{\mathsf{unp}}_{\epsilon,s}(X|Z) \geqslant H^{*\mathsf{unp}}_{\epsilon,s}(X|Z)$, and in general, list-unpredictability can be much smaller than unpredictability entropy.[8] Interestingly, we can derive a k bit key with almost k bits of HILL entropy from a source $X|Z$ which k bits unpredictability entropy $H^{\mathsf{unp}}_{\epsilon,s}(X|Z) \geqslant k$ in two extreme cases, namely, if either

1. if $X|Z$ has basically no HILL entropy (even against small circuits).
2. or when $X|Z$ has (almost) k bits of (high quality) HILL entropy.

In case 1. we observe that if $H^{\mathsf{HILL}}_{\epsilon,t}(X|Z) \approx 0$ for some $t \ll s$, or equivalently, given Z we can efficiently distinguish X from any $X' \neq X$, then the Eq oracle used in the definition of list-unpredictability can be efficiently emulated, which means it's redundant, and thus $X|Z$ has the same amount of list-unpredictability and unpredictability entropy, $H^{\mathsf{unp}}_{s,\epsilon}(X|Z) \approx H^{*\mathsf{unp}}_{s',\epsilon'}(X|Z)$ for $(\epsilon', s') \approx (\epsilon, s)$. Thus, we can use Theorem 1 to derive a k bit key with $k - O(1)$ bits of HILL entropy in this case. In case 2., we can simply use any condenser for min-entropy to get a key with HILL entropy $k - \log\log\epsilon^{-1}$p As condensing almost all the unpredictability entropy into HILL entropy is possible in the two extreme cases where $X|Z$ has either no or a lot of HILL entropy, it seems conceivable that it's also possible in all the in-between cases (i.e., without making any additional assumptions about $X|Z$ at all).

GL vs. Condensing. Let us stress as this point that, because of the two issues discussed above, our result does not always allow to generate more bits with high HILL entropy than just using the Goldreich-Levin theorem. Assuming k bits of unpredictability we get $k - 3$ of HILL, whereas GL will only give $k - 2\log(1/\epsilon)$. But as currently our reduction has a quantitatively larger loss in circuit size than the GL theorem, in order to get HILL entropy of the same quality

[8] E.g., let X by uniform over $\{0,1\}^n$ and Z arbitrary, but independent of X, then for $s = \exp(n)$ we have $H^{\mathsf{unp}}_s(X|Z) = n$ but $H^{*\mathsf{unp}}_s(X|Z) = 0$ as we can simply invoke Eq on all $\{0,1\}^n$ until X is found.

(i.e., secure against (s, δ) adversaries for some fixed (s, δ)) we must consider the unpredictability entropy of the source $X|Z$ against more powerful adversaries than if we're about to use GL. And in general, the amount of unpredictability (or any other computational) entropy of $X|Z$ can decrease as we consider more powerful adversaries.

2 Entropy Notions

In this section we formally define the different entropy notions considered in this paper. We denote with $\mathcal{D}_s^{rand,\{0,1\}}$ the set of all *probabilistic* circuits of size s with *boolean* output, and $\mathcal{D}_s^{rand,[0,1]}$ denotes the set of all *probabilistic* circuits with *real-valued* output in the range $[0, 1]$. The analogous *deterministic* circuits are denoted $\mathcal{D}_s^{det,\{0,1\}}$ and $\mathcal{D}_s^{det,[0,1]}$. We use $X \sim_{\epsilon,s} Y$ to denote computational indistinguishability of variables X and Y, formally[9]

$$X \sim_{\epsilon,s} Y \iff \forall \mathsf{C} \in \mathcal{D}_s^{rand,\{0,1\}} \ : \ |\Pr[\mathsf{C}(X) = 1] - \Pr[\mathsf{C}(Y) = 1]| \leqslant \epsilon \quad (4)$$

$X \sim_\epsilon Y$ denotes that X and Y have statistical distance ϵ, i.e., $X \sim_{\epsilon,\infty} Y$, and with $X \sim Y$ we denote that they're identically distributed. With U_n we denote the uniform distribution over $\{0, 1\}^n$.

Definition 1. *The* **min-entropy** *of a random variable X with support $\mathcal{X}$ is*

$$H_\infty(X) = -\log_2 \max_{x \in \mathcal{X}} \Pr[X = x]$$

For a pair (X, Z) of random variables, the **average min-entropy** *of X conditioned on Z is*

$$\widetilde{H}_\infty(X|Z) \ = \ -\log_2 \mathop{\mathbb{E}}_{z \leftarrow Z} \max_x \Pr[X = x|Z = z] \ = \ -\log_2 \mathop{\mathbb{E}}_{z \leftarrow Z} 2^{-H_\infty(X|Z=z)}$$

HILL entropy is a computational variant of min-entropy, where X (conditioned on Z) has k bits of HILL entropy, if it cannot be distinguished from some Y that (conditioned on Z) has k bits of min-entropy, formally

Definition 2 ([14], [15]). *A random variable X has* **HILL entropy** *k, denoted by $H^{\mathsf{HILL}}_{\varepsilon,s}(X) \geq k$, if there exists a distribution Y satisfying $H_\infty(Y) \geq k$ and $X \sim_{\varepsilon,s} Y$.*

Let (X, Z) be a joint distribution of random variables. Then X has **conditional HILL entropy** *k conditioned on Z, denoted by $H^{\mathsf{HILL}}_{\varepsilon,s}(X|Z) \geq k$, if there exists a joint distribution (Y, Z) such that $\widetilde{H}_\infty(Y|Z) \geq k$ and $(X, Z) \sim_{\varepsilon,s} (Y, Z)$.*

[9] Let us mention that the choice of the distinguisher class in (4) irrelevant (up to a small additive difference in circuit size), we can replace $\mathcal{D}_s^{rand,\{0,1\}}$ with any of the three other distinguisher classes.

Barak, Sahaltiel and Wigderson [1] define the notion of metric entropy, which is defined like HILL, but the quantifiers are exchanged. That is, instead of asking for a single distribution (Y, Z) that fools all distinguishers, we only ask that for every distinguisher D, there exists such a distribution. For reasons discussed in Section 2, in the definition below we make the class of distinguishers considered explicit.

Definition 3 ([1], [10]). *Let (X, Z) be a joint distribution of random variables. Then X has* **conditional metric entropy** *k conditioned on Z (against probabilistic boolean distinguishers), denoted by $H^{\mathsf{Metric},rand,\{0,1\}}_{\varepsilon,s}(X|Z) \geq k$, if for every $\mathsf{D} \in \mathcal{D}^{rand,\{0,1\}}_s$ there exists a joint distribution (Y, Z) such that $\widetilde{H}_\infty(Y|Z) \geq k$ and*

$$|\Pr[\mathsf{D}(X, Z) = 1] - \Pr[\mathsf{D}(Y, Z) = 1]| \leqslant \epsilon$$

More generally, for $class \in \{rand, det\}, range \in \{[0, 1], \{0, 1\}\}$, $H^{\mathsf{Metric},class,range}_{\varepsilon,s}(X|Z) \geq k$ if for every $\mathsf{D} \in \mathcal{D}^{class,range}_s$ such a (Y, Z) exists.

Like HILL entropy, also unpredictability entropy, which we'll define next, can be seen as a computational variant of min-entropy. Here we don't require indistinguishability as for HILL entropy, but only that the variable is hard to predict.

Definition 4 ([15]). *X has* **unpredictability entropy** *k conditioned on Z, denoted by $H^{\mathsf{unp}}_{\epsilon,s}(X|Z) \geq k$, if (X, Z) is (ϵ, s) indistinguishable from some (Y, Z), where no probabilistic circuit of size s can predict Y given Z with probability better than 2^{-k}, i.e., $H^{\mathsf{unp}}_{s,\epsilon}(X|Z) \geq k$ if and only if*

$$\exists (Y, Z), (X, Z) \sim_{\varepsilon,s} (Y, Z)\ \forall \mathsf{C}, |\mathsf{C}| \leqslant s\ :\ \Pr_{(y,z)\leftarrow(Y,Z)}[\mathsf{C}(z) = y] \leqslant 2^{-k} \qquad (5)$$

We also define a notion called "list-unpredictability", denoted $H^{*\mathsf{unp}}_{\epsilon,s}(X|Z) \geq k$, which holds if $H^{\mathsf{unp}}_{\epsilon,s}(X|Z) \geq k$ as in (5), but where C additionally gets oracle access to a function $\mathsf{Eq}(.)$ which outputs 1 on input y and 0 otherwise. So, C can efficiently test if some candidate guess for y is correct.[10]

Remark 1 (The ϵ parameter). The ϵ parameter in the definition above is not really necessary, following [16], we added it so we can have a "smooth" notion, which is easier to compare to HILL or smooth min-entropy. If $\epsilon = 0$, we'll simply omit it, then the definition simplifies to

$$H^{\mathsf{unp}}_s(X|Z) \geq k \iff \Pr_{(x,z)\leftarrow(X,Z)}[\mathsf{C}(z) = x] \leqslant 2^{-k}$$

Let us also mention that unpredictability entropy is only interesting if the conditional part Z is not empty as (already for s that is linear in the length of X)

[10] We name this notion "list-unpredictability" as we get the same notion when instead of giving C oracle access to $\mathsf{Eq}(.)$, we allow $\mathsf{C}(z)$ to output a list of guesses for y, not just one value, and require that $\Pr_{(y,z)\leftarrow(Y,Z)}[y \in \mathsf{C}(z)] \leqslant 2^{-k}$. This notion is inspired by the well known notion of list-decoding.

we have $H_s^{\mathsf{unp}}(X) = H_\infty(X)$ which can be seen by considering the circuit C (that gets no input as Z is empty) which simply outputs the constant x maximizing $\Pr[X = x]$.

Metric vs. HILL. We will use a lemma which states that deterministic real-valued metric entropy implies the same amount of HILL entropy (albeit, with some loss in quality). This lemma has been proven by [1] for the unconditional case, i.e., when Z in the lemma below is empty, it has been observed by [4, 10] that the proof also holds in the conditional case as stated below

Lemma 1 ([1, 4, 10]). *For any joint distribution* $(X, Z) \in \{0,1\}^n \times \{0,1\}^m$ *and any* ϵ, δ, k, s

$$H^{\mathsf{Metric},det,[0,1]}_{\epsilon,s}(X|Z) \geqslant k \quad \Rightarrow \quad H^{\mathsf{HILL}}_{\epsilon+\delta,s\cdot\delta^2/(m+n)}(X|Z) \geqslant k$$

Note that in Definition 2 of HILL entropy, we only consider security against probabilistic boolean distinguishers (as $\sim_{\epsilon,s}$ was defined this way), whereas in Definiton 3 of metric entropy we make the class of distinguishers explicit. The reason for this is that in the definition of HILL entropy the class of distinguishers considered is irrelevant (except for a small additive degradation in circuit size, cf. [10, Lemma 2.1]).[11] Unlike for HILL, for metric entropy the choice of the distinguisher class does matter. In particular, deterministic boolean metric entropy $H^{\mathsf{Metric},det,\{0,1\}}_{\epsilon,s}(X|Y) \geqslant k$ is only known to imply deterministic real-valued metric entropy $H^{\mathsf{Metric},det,[0,1]}_{\epsilon+\delta,s}(X|Y) \geqslant k - \log(\delta^{-1})$, i.e., we must allow for a $\delta > 0$ loss in distinguishing advantage, and this will at the same time result in a loss of $\log(\delta^{-1})$ in the amount of entropy. For this reason, it is crucial that in Theorem 2 we show that unpredictability entropy implies deterministic *real-valued* metric entropy, so we can then apply Lemma 1 to get the same amount of HILL entropy. Dealing with real-valued distinguishers is the main source of technical difficulty in the proof of the Theorem 2, proving the analogous statement for deterministic *boolean* distinguishers is much simpler.

3 Known Results on Provably Secure Key-Derivation

We say that a cryptographic scheme has security α, if no adversary (from some class of adversaries like all polynomial size circuits) can win some security game with advantage $\geqslant \alpha$ if the scheme is instantiated with a uniformly random string.[12] Below we will distinguish between *unpredictability* applications, where the advantage bounds the probability of winning some security game (a typical

[11] This easily follows from the fact that in the definition (4) of computational indistinguishability the choice of the distinguisher class is irrelevant.

[12] We'll call this string "key". Though in many settings (in particular when keys are not simply uniform random strings, like in public-key crypto) this string is not used as a key directly, but one rather should think of it as the randomness used to sample the actual keys.

example are digital signature schemes, where the game captures the existential unforgeability under chosen message attacks), and *indistinguishability* applications, where the advantage bounds the distinguishing advantage from some ideal object (a typical example is the security definition of pseudorandom generators or functions).

3.1 Key-Derivation from Min-Entropy

Strong Extractors. Let (X, Z) be a source where $\widetilde{H}_\infty(X|Z) \geqslant k$, or equivalently, no adversary can guess X given Z with probability better than 2^{-k} (cf. Def. 1). Consider the case where we want to derive a key $K = h(X, S)$ that is statistically close to uniform given (Z, S). For example, X could be some physical source (like statistics from keystrokes) from which we want to generate almost uniform randomness. Here Z models potential side-information the adversary might have on X. This setting is very well understood, and such a key can be derived using a strong extractor as defined below.

Definition 5 ([18], [5]). *A function* $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^\ell$ *is an average-case* (k, ϵ)*-strong extractor if for every distribution* (X, Z) *over* $\{0,1\}^n \times \{0,1\}^m$ *with* $\widetilde{H}_\infty(X|Z) \geqslant k$ *and* $S \sim U_d$*, the distribution* $(\mathsf{Ext}(X, S), S, Z)$ *has statistical distance* ϵ *to* (U_ℓ, S, Z)*.*

Extractors Ext as above exist with $\ell = k - 2\log(1/\epsilon)$ [14]. Thus, from any (X, Z) where $\widetilde{H}_\infty(X|Z) \geqslant k$ we can extract a key $K = \mathsf{Ext}(X, S)$ of length $k - 2\log(1/\epsilon)$ that is ϵ close to uniform [14]. The entropy gap $2\log(1/\epsilon)$ is optimal by the so called "RT-bound" [19], even if we assume the source is efficiently samplable [7].

If instead of using a uniform ℓ bit key for an α secure scheme, we use a key that is ϵ close to uniform, the scheme will still be at least $\beta = \alpha + \epsilon$ secure. In order to get security β that is of the same order as α, we thus must set $\epsilon \approx \alpha$. When the available amount k of min-entropy is small, for example when dealing with biometric data [3, 5], a loss of $2\log(1/\epsilon)$ bits (that's 160 bits for a typical security level $\epsilon = 2^{-80}$) is often unacceptable.

Condensers. The above bound is basically tight for many *indistinguishability* applications like pseudorandom generators or pseudorandom functions.[13] Fortunately, for many applications a close to uniform key is not necessary, and a key $|K|$ with min-entropy $|K| - \Delta$ for some small Δ is basically as good as a uniform one. This is the case for all *unpredictability* applications, which includes OWFs, digital-signatures and MACs.[14] It's not hard to show that if the scheme

[13] For example, consider a pseudorandom function $\mathsf{F} : \{0,1\}^k \times \{0,1\}^a \to \{0,1\}$ and a key K that is uniform over all keys where $\mathsf{F}(K, 0) = 0$, this distribution is $\epsilon \approx 1/2$ close to uniform and has min-entropy $\approx |K| - 1$, but the security breaks completely as one can distinguish $\mathsf{F}(U_k, .)$ from $\mathsf{F}(K, .)$ with advantage $\beta \approx 1/2$ (by quering on input 0, and outputting 1 iff the output is 0).

[14] [8] identify an interesting class of applications called "square-friendly", this class contains all unpredictability applications, and some indistinguishability applications

is α secure with a uniform key it remains at least $\beta = \alpha 2^{\Delta}$ secure (against the same class of attackers) if instantiated with any key K that has $|K| - \Delta$ bits of min-entropy.[15] Thus, for unpredictability applications we don't have to extract an almost uniform key, but "condensing" X into a key with $|K| - \Delta$ bits of min-entropy for some small Δ is enough.

[7] show that a $(\log \epsilon + 1)$-wise independent hash function $\mathsf{Cond} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^\ell$ is a condenser with the following parameters. For any (X, Z) where $\widetilde{H}_\infty(X|Z) \geqslant \ell$, for a random seed S (used to sample a $(\log \epsilon + 1)$-wise independent hash function), the distribution $(\mathsf{Cond}(X, S), S)$ is ϵ close to a distribution (Y, S) where $\widetilde{H}_\infty(Y|Z) \geqslant \ell - \log\log(1/\epsilon)$. Using such an ℓ bit key (condensed from a source with ℓ bits min-entropy) for an unpredictability application that is α secure (when using a uniform ℓ bit key), we get security $\beta \leqslant \alpha 2^{\log\log(1/\epsilon)} + \epsilon$, which setting $\epsilon = \alpha$ gives $\beta \leqslant \alpha(1 + \log(1/\alpha))$ security, thus, security degrades only by a logarithmic factor.

3.2 Key-Derivation from Computational Entropy

HILL Entropy. As already discussed in the introduction, often we want to derive a key from a distribution (X, Z) where there's no "real" min-entropy at all $\widetilde{H}_\infty(X|Z) = 0$. This is for example the case when Z is the transcript (that can be observed by an adversary) of a key-exchange protocol like Diffie-Hellman, where the agreed value $X = g^{ab}$ is determined by the transcript $Z = (g^a, g^b)$ [11, 17]. Another setting where this can be the case is in the context of side-channel attacks, where the leakage Z from a device can completely determine its internal state X. If $X|Z$ has k bits of HILL entropy, i.e., is computationally indistinguishable from having min-entropy k (cf. Def. 2) we can derive keys exactly as described above assuming $X|Z$ had k bits of min-entropy. In particular, if $X|Z$ has $|K| + 2\log(1/\epsilon)$ bits of HILL entropy for some negligible ϵ, we can derive a key K that is pseudorandom, and if $X|Z$ has $|K| + \log\log(1/\epsilon)$ bits of HILL entropy, we can derive a key that is almost as good as a uniform one for any unpredictability application.

Unpredictability Entropy. Clearly, the minimal assumption we must make on a distribution $(X, Z) \in \{0,1\}^n \times \{0,1\}^m$ for any key derivation to be possible at all is that X is hard to compute given Z, that is, $X|Z$ must have some unpredictability entropy as in Definition 4. Goldreich and Levin [12] show how to generate

like weak PRFs (which are PRFs that can only be queried on random inputs). This class of applications remains somewhat secure even for a small entropy gap Δ: For $\Delta = 1$ the security is $\beta \approx \sqrt{\alpha}$. This is worse that the $\beta = 2\alpha$ for unpredictability applications, but much better than the complete loss of security $\beta \approx 1/2$ required for some indistinguishability apps like (standard) PRFs.

[15] Assume some adversary breaks the scheme, say, forges a signature, with advantage β if the key comes from the distribution K. If we sample a uniform key instead, it will have the same distribution as K conditioned on an event that holds with probability $2^{-\Delta}$, and thus this adversary will still break the scheme with probability $\beta/2^{\Delta}$.

pseudorandom bits from such a source. In particular, the Goldreich-Levin theorem implies that if $X|Z$ has at least $2\log\epsilon^{-1}$ bits of list-unpredictability, then the inner product R^TX of X with a random vector R is ϵ indistinguishable from uniformly random (the loss in circuit size is $poly(n,m)/\epsilon^4$). Using the chain rule for unpredictability entropy,[16] we can generate an $\ell = k - 2\log\epsilon^{-1}$ bit long pseudorandom string that is $\ell\epsilon$ indistinguishable (the extra ℓ factor comes from taking the union bound over all bits) from uniform.

Thus, we can turn k bits of list-unpredictability into $k - 2\log\epsilon^{-1}$ bits of pseudorandom bits (and thus also that much HILL entropy) with quality roughly ϵ. The question whether it's possible to generate significantly more than $k - 2\log\epsilon^{-1}$ of HILL entropy from a source with k bits of (list-)unpredictability seems to have never been addressed in the literature before. The reason might be that one usually is interested in generating pseudorandom bits (not just HILL entropy), and for this, the $2\log\epsilon^{-1}$ entropy loss is inherent. The observation that for many applications high HILL entropy is basically as good as pseudorandomness is more recent, and recently gained attention by its usefulness in the context of leakage-resilient cryptography [8, 9].

In this paper we prove that it's in fact possible to turn almost all list-unpredictability into HILL entropy.

4 Condensing Unpredictability

Let $X|Z$ have k bits of list-unpredictability, and assume we start extracting Goldreich-Levin hardcore bits $A_1, A_2, \ldots$ by taking inner products $A_i = R_i^TX$ for random R_i. The first extracted bits $A_1, A_2, \ldots$ will be pseudorandom (given the R_i and Z), but with every extracted bit, the list-unpredictability can also decrease by one bit. As the GL theorem requires at least $2\log\epsilon^{-1}$ bits of list-unpredictability to extract an ϵ secure pseudorandom bit, we must stop after $k - 2\log\epsilon^{-1}$ bits. In particular, the more we extract, the worse the pseudorandomness of the extracted string becomes. Unlike the original GL theorem, in our Theorem 1 we only argue about the unpredictability of the extracted string, and unpredictability entropy has the nice property that it can never decrease, i.e., predicting $A_1, \ldots, A_{i+1}$ is always at least as hard as predicting $A_1, \ldots, A_i$. Thus, despite the fact that once i approaches k it becomes easier and easier to predict A_i (given $A_1, \ldots, A_{i-1}, Z$ and the R_i's)[17] this hardness will still add up to $k - O(1)$ bits of unpredictability entropy.

The proof is by contradiction, we assume that $A_1, \ldots, A_k$ can be predicted with advantage 2^{-k+3} (i.e., does not have $k-3$ bits of unpredictability), and then

[16] Which states that if $X|Z$ has k bits of list-unpredictability, then for any (A,R) where R is independent of (X,Z), $X|(Z,A,R)$ has $k - |A|$ bits of list-unpredictability entropy. In particular, extracting ℓ inner product bits, decreases the list-unpredictability by at most ℓ.

[17] The only thing we know about the last extracted bit A_k is that it cannot be predicted with advantage $\geqslant 0.75$, more generally, A_{k-j} cannot be predicted with advantage $1/2 + 1/2^{j+2}$.

use such a predictor to predict X with advantage $> 2^{-k}$, contradicting the k bit list-unpredictability of $X|Z$. If $A_1, \ldots, A_k$ can be predicted as above, then there must be an index j s.t. A_j can be predicted with good probability conditioned on $A_1, \ldots, A_{j-1}$ being correctly predicted. We then can use the Goldreich-Levin theorem, which tells us how to find X given such a predictor. Unfortunately, j can be close to k, and to apply the GL theorem, we first need to find the right values for $A_1, \ldots, A_{j-1}$ on which we condition, and also can only use the predictor's guess for A_j if it was correct on the first $j-1$ bits. We have no better strategy for this than trying all possible values, and this is the reason why the loss in circuit size in Theorem 1 depends on 2^k.

In our proof, instead of using the Goldreich-Levin theorem, we will actually use a more fine-grained variant due to Hast which allows to distinguish between errors and erasures, this will give a much better quantitative bound.

Theorem 1 (Condensing Upredictability Entropy). *Consider any distribution (X, Z) over $\{0,1\}^n \times \{0,1\}^m$ where*

$$H^{*\mathsf{unp}}_{\epsilon,s}(X|Z) \geqslant k$$

then for a random $R \leftarrow \{0,1\}^{k \times n}$

$$H^{\mathsf{unp}}_{\epsilon,t}(R.X|Z,R) \geqslant k - \Delta$$

where[18] $t = \frac{s}{2^{2k}\,\mathrm{poly}(m,n)}$, $\Delta = 3$

5 High Unpredictability implies Metric Entropy

In this section we state our main results, showing that k bits of unpredictability entropy imply the same amount of HILL entropy, with a loss exponential in the "entropy gap".

Theorem 2 (Unpredictability Entropy Implies HILL Entropy). *For any distribution (X, Z) over $\{0,1\}^n \times \{0,1\}^m$, if $X|Z$ has unpredictability entropy*

$$H^{\mathsf{unp}}_{\gamma,s}(X|Z) \geqslant k \tag{6}$$

then, with $\Delta = n - k$ denoting the entropy gap, $X|Z$ has (real valued, deterministic) metric entropy

$$H^{\mathsf{Metric},det,[0,1]}_{\epsilon+\gamma,t}(X|Z) \geqslant k \quad \textit{for} \quad t = \Omega\left(s \cdot \frac{\epsilon^5}{2^{5\Delta}\log^2\left(2^{\Delta}\epsilon^{-1}\right)}\right) \tag{7}$$

By Lemma 1 this further implies that $X|Z$ has, for any $\delta > 0$, HILL entropy

$$H^{\mathsf{HILL}}_{\epsilon+\delta+\gamma,\Omega(t\delta^2/(n+m))}(X|Z) \geqslant k$$

which for $\epsilon = \delta = \gamma$ is $H^{\mathsf{HILL}}_{3\epsilon,\Omega(s\cdot\epsilon^7/2^{5\Delta}(n+m)\log^2(2^{\Delta}\epsilon^{-1}))}(X|Z) \geqslant k$

[18] We can set Δ to be any constant > 1 here, but choosing a smaller Δ would imply a smaller t.

References

1. Barak, B., Shaltiel, R., Wigderson, A.: Computational Analogues of Entropy. In: Arora, S., Jansen, K., Rolim, J.D.P., Sahai, A. (eds.) RANDOM-APPROX 03. LNCS, vol. 2764, pp. 200–215. Springer (2003)
2. Bellare, M., Rogaway, P.: Random oracles are practical: A paradigm for designing efficient protocols. In: Ashby, V. (ed.) ACM CCS 93. pp. 62–73. ACM Press (Nov 1993)
3. Boyen, X., Dodis, Y., Katz, J., Ostrovsky, R., Smith, A.: Secure remote authentication using biometric data. In: Cramer, R. (ed.) EUROCRYPT 2005. LNCS, vol. 3494, pp. 147–163. Springer (May 2005)
4. Chung, K.M., Kalai, Y.T., Liu, F.H., Raz, R.: Memory delegation. In: Rogaway, P. (ed.) CRYPTO 2011. LNCS, vol. 6841, pp. 151–168. Springer (Aug 2011)
5. Dodis, Y., Ostrovsky, R., Reyzin, L., Smith, A.: Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data. SIAM Journal on Computing 38(1), 97–139 (2008)
6. Dodis, Y., Gennaro, R., Håstad, J., Krawczyk, H., Rabin, T.: Randomness extraction and key derivation using the CBC, cascade and HMAC modes. In: Franklin, M. (ed.) CRYPTO 2004. LNCS, vol. 3152, pp. 494–510. Springer (Aug 2004)
7. Dodis, Y., Pietrzak, K., Wichs, D.: Key derivation without entropy waste. In: EUROCRYPT 14. LNCS, Springer (2014)
8. Dodis, Y., Yu, Y.: Overcoming weak expectations. In: Sahai, A. (ed.) TCC 2013. LNCS, vol. 7785, pp. 1–22. Springer (Mar 2013)
9. Dziembowski, S., Pietrzak, K.: Leakage-resilient cryptography. In: 49th FOCS. pp. 293–302. IEEE Computer Society Press (Oct 2008)
10. Fuller, B., Reyzin, L.: Computational entropy and information leakage. Cryptology ePrint Archive, Report 2012/466 (2012), http://eprint.iacr.org/
11. Gennaro, R., Krawczyk, H., Rabin, T.: Secure Hashed Diffie-Hellman over non-DDH groups. In: Cachin, C., Camenisch, J. (eds.) EUROCRYPT 2004. LNCS, vol. 3027, pp. 361–381. Springer (May 2004)
12. Goldreich, O., Levin, L.A.: A hard-core predicate for all one-way functions. In: 21st ACM STOC. pp. 25–32. ACM Press (May 1989)
13. Hast, G.: Nearly one-sided tests and the Goldreich-Levin predicate. In: Biham, E. (ed.) EUROCRYPT 2003. LNCS, vol. 2656, pp. 195–210. Springer (May 2003)
14. Håstad, J., Impagliazzo, R., Levin, L.A., Luby, M.: A pseudorandom generator from any one-way function. SIAM Journal on Computing 28(4), 1364–1396 (1999)
15. Hsiao, C.Y., Lu, C.J., Reyzin, L.: Conditional Computational Entropy, or Toward Separating Pseudoentropy from Compressibility. In: Naor, M. (ed.) EUROCRYPT 07. LNCS, vol. 4515, pp. 169–186. Springer (2007)
16. Hsiao, C.Y., Lu, C.J., Reyzin, L.: Conditional computational entropy, or toward separating pseudoentropy from compressibility. In: Naor, M. (ed.) EUROCRYPT 2007. LNCS, vol. 4515, pp. 169–186. Springer (May 2007)
17. Krawczyk, H.: Cryptographic extraction and key derivation: The HKDF scheme. In: Rabin, T. (ed.) CRYPTO 2010. LNCS, vol. 6223, pp. 631–648. Springer (Aug 2010)
18. Nisan, N., Zuckerman, D.: More deterministic simulation in logspace. In: 25th ACM STOC. pp. 235–244. ACM Press (May 1993)
19. Radhakrishnan, J., Ta-Shma, A.: Bounds for dispersers, extractors, and depth-two superconcentrators. SIAM J. Discrete Math. 13(1), 2–24 (2000)