Some Notions of Entropy for Cryptography*

Leonid Reyzin
Boston University Computer Science
http://www.cs.bu.edu/~reyzin

June 2, 2011

Abstract

This paper presents a brief and (necessarily) incomplete survey of some notions of entropy that have been recently used in the analysis of cryptographic constructions. It focuses on min-entropy and its extensions to the cases when the adversary has correlated information and/or is computationally bounded. It also presents results that can be used to bound such entropy and apply it to the analysis of cryptographic constructions.

1 Information-Theoretic Case

In many contexts, particularly in security-related ones, the ability to guess the value of a random variable (in a single attempt) is an important measures of the variable's quality. This ability is captured by the following notion.

Definition 1. A random variable X has **min-entropy** k, denoted $H_\infty(X) = k$, if

$$\max_x \Pr[X = x] = 2^{-k}.$$

Randomness extractors were defined to work with any distribution that has min-entropy [NZ96]. Moreover, strong extractors (whose outputs are nearly uniform even the presence of the seed) produce outputs that have, with high probability over the choice of seed, almost maximal min-entropy.

Lemma 1 ([CKOR10])**.** *If $\mathsf{Ext} : N \times I \to \{0,1\}^\ell$ is a (k, ε)-strong extractor with inputs from a set N and seeds from a distribution I, and X is a random variable taking values in N with $\mathbf{H}_\infty(X) \geq k$, then $\mathbf{H}_\infty(\mathsf{Ext}(X; i)) \geq \ell - 1$ with probability at least $1 - 2^\ell\varepsilon$ over the choice of the seed i.*

A less demanding notion is sometimes more suitable and allows for better analysis of constructions, because one can "pretend" to work with a very close distribution Y that has more min-entropy:

Definition 2 ([RW04])**.** A random variable X has **ε-smooth min-entropy** k if

$$\max_{Y:\ \mathbf{SD}(X,Y)\leq\varepsilon} \mathbf{H}_\infty(Y) = k$$

(here, $\mathbf{SD}(X, Y)$ is the usual statistical distance, defined as $\max_T \Pr[X \in T] - \Pr[Y \in T]$).

*A slightly updated and corrected version of [Rey11]

Quite often, the adversary has some additional information Z that is correlated with X. Conditional min-entropy $\mathbf{H}_\infty(X|Z)$ is defined in [RW05] as $-\log \max_{x,z} \Pr(X = x \mid Z = z) = \min_z \mathbf{H}_\infty(X \mid Z = z)$ (an ε-smooth version is also defined in [RW05, Section 1.3] by eliminating bad portions of (X, Z) that occur with probability at most ε). Again, a less restrictive notion is sometimes more suitable:

Definition 3 ([DORS08, Section 2.4]). Let (X, Z) be a pair of random variables. The **average min-entropy** of X conditioned on Z is

$$\tilde{H}_\infty(X|Z) \stackrel{\text{def}}{=} -\log \mathop{\mathbf{E}}_{z \leftarrow Z} \max_x \Pr[X = x|Z = z] = -\log[\mathop{\mathbf{E}}_{z \leftarrow Z}(2^{-H_\infty(X|Z=z)})] .$$

Average min-entropy, like min-entropy, is simply the logarithm of the probability that the adversary (this time, given the value of Z) will guess the value of X in a single attempt. Again, an ε-smooth variant of it can be defined (a comparison of ε-smooth, conditional, and average min-entropy notions is given in [DORS08, Appendix B]).

Average min-entropy exhibits some properties that agree with our intuition: conditioning on Z that has b bits of information reduces the entropy of X by at most b.

Lemma 2 ([DORS08, Lemma 2.2b]). *$\widetilde{\mathbf{H}}_\infty(X \mid Z) \geq \mathbf{H}_\infty(X, Z) - b$, where 2^b is the number of elements in Z (more generally, $\widetilde{\mathbf{H}}_\infty(X \mid Z_1, Z_2) \geq \widetilde{\mathbf{H}}_\infty(X, Z_1 \mid Z_2) - b$, where 2^b is the number of elements in Z_2).*

Randomness extractors, which were originally analyzed for distribution of min-entropy, can also be used on distributions that have average min-entropy, with essentially the same results. A (k, ε)-average-case extractor is defined in [DORS08, Section 2.5] as a function that takes in a sample from a distribution X such that $\widetilde{\mathbf{H}}_\infty(X \mid Z) \geq k$ and a random seed, and produces an output that is ε-close to uniform even in the presence of the correlated value from Z and the seed. In some cases (for instance, in universal-hashing-based extractors), a (k, ε)-extractor is also a (k, ε)-average-case extractor [DORS08, Lemma 2.4]; in all but the most pathological cases, a (k, ε)-extractor is a $(k, 3\varepsilon)$-average-case extractor [Vad11]. The following lemma shows that outputs extracted by average-case extractors will themselves have average min-entropy.

Lemma 3 ([KR09, Lemma 1]). *If $\mathsf{Ext} : N \times I \to \{0, 1\}^\ell$ is a (k, ε)-average-case extractor with inputs from a set N and seeds from a distribution I, and (X, Z) is a pair of random variables with X taking values in N and $\widetilde{\mathbf{H}}_\infty(X|Z) \geq k$, then $\widetilde{\mathbf{H}}_\infty(\mathsf{Ext}(X; I) \mid Z, I) \geq \min\left(\ell, \log \frac{1}{\varepsilon}\right) - 1$.*

Average min-entropy often allows for simpler statements and analyses; for example, the security of information-theoretic MACs with nonuninform keys can be analyzed using the average min-entropy of the keys (see [KR09, Proposition 1]). However, average min-entropy can be converted to min-entropy when needed.

Lemma 4 ([DORS08, Lemma 2.2a]). *For any $\delta > 0$, $\mathbf{H}_\infty(X|Z = z)$ is at least $\widetilde{\mathbf{H}}_\infty(X|Z) - \log(1/\delta)$ with probability at least $1 - \delta$ over the choice of z.*

This style of analysis—using average min-entropy wherever possible and converting it to min-entropy when needed—was used, for example, in [KR09], [CKOR10], to analyze complex interactive protocols involving extractors and MACs.

2 Computational Case

It is natural to say that if a distribution cannot be distinguished by a resource-bounded adversary from one that has entropy, then it has computational entropy. For example, pseudorandom distributions have this property.

Definition 4 ([HILL99, BSW03]). A distribution X has **HILL entropy** at least k, denoted by $H^{\mathsf{HILL}}_{\varepsilon,s}(X) \geq k$, if there exists a distribution Y such that $H_\infty(Y) \geq k$ and no circuit of size s can distinguish X and Y with advantage more than ε.

(Here and below, unless otherwise specified, distinguishers are randomized and output a single bit.)

A conditional notion can be defined similarly.

Definition 5 ([HLR07, Section 2]). X has **conditional HILL entropy** at least k conditioned on Z, denoted $H^{\mathsf{HILL}}_{\varepsilon,s}(X|Z) \geq k$, if there exists a collection of distributions Y_z (for $z \in Z$) giving rise to a joint distribution (Y, Z), such that the average min-entropy $\tilde{H}_\infty(Y|Z) \geq k$ and no circuit of size s can distinguish (X, Z) and (Y, Z) with advantage more than ε.

However, there are many variations of the computational definitions, and which one is "right" is unclear. For example, [GW11, Lemma 3.1] allow one to change not only X, but also Z, as long as the change is computationally indistinguishable.

As another example, [BSW03], following [Yao82], proposed an alternative way to measure computational entropy: by measuring compressibility of the string by efficient algorithms. It was further converted to conditional entropy in [HLR07].

Definition 6 ([HLR07, Section 2]). X has **Yao entropy** at least k conditioned on Z, denoted by $H^{\mathsf{Yao}}_{\varepsilon,s}(X|Z) \geq k$, if for every pair of circuits c, d of total size s with the outputs of c having length ℓ,

$$\Pr_{(x,z)\leftarrow(X,Z)}[d(c(x, z), z) = x] \leq 2^{\ell-k} + \varepsilon.$$

It was shown in [HLR07, Theorem 4] that the two notions (which are equivalent in the information-theoretic case) are actually different in the computational setting: Yao entropy may be higher than HILL (but never lower), and measuring Yao entropy rather than HILL entropy may allow one to extract more pseudorandom bits from a distribution.

Another seemingly natural computational analog of min-entropy is "unpredictability" entropy, because it also measures the chances of correctly guessing X in a single try.

Definition 7 ([HLR07, Section 5]). X has **unpredictability entropy** at least k conditioned on Z, denoted by $H^{\mathsf{unp}}_{\varepsilon,s}(X|Z) \geq k$, if there exists a collection of distributions Y_z (for $z \in Z$), giving rise to a joint distribution (Y, Z), such that no circuit of size s can distinguish (X, Z) and (Y, Z) with advantage more than ε, and for all circuits C of size s,

$$\Pr[C(Z) = Y] \leq 2^{-k}.$$

As shown in [HLR07, Section 5], unpredictability entropy can be higher than HILL entropy but never higher than Yao entropy. We know that extractors work with conditional HILL entropy to produce pseudorandom outputs; some extractors ("reconstructive" ones) also work with conditional compressibility and unpredictability entropies.

Understanding how conditioning on information leakage Z impacts the entropy of X is particularly difficult. It would be highly desirable to have an analog of the simple statement of Lemma 2 to simplify the analysis of protocols in a variety of scenarios, particularly in leakage-resilient cryptography. The following result, for both average-case and worst-case entropy, is relatively simple to state. However, it is for a notion of entropy that is a lot less natural: $\mathtt{Metric}^*$ entropy, which differs from HILL entropy in two respects: there can be a different distribution Y for each distinguishing circuit of size s, and the circuit, instead outputting 1 with some probability p and 0 with probability $1-p$, deterministically outputs a value p in the interval $[0,1]$.

Theorem 1 ([FR11])**.** *Define P_z as $\Pr[Z=z]$. Assume Z has 2^b elements. Then*

$$H^{\mathtt{Metric}^*}_{\varepsilon/P_z,s'}(X|Z=z) \geq H^{\mathtt{Metric}^*}_{\varepsilon,s}(X) - \log 1/P_z$$

and

$$H^{\mathtt{Metric}^*}_{\varepsilon 2^b,s'}(X|Z) \geq H^{\mathtt{Metric}^*}_{\varepsilon,s}(X) - b\,,$$

where $s' \approx s$.

A weaker version of this statement appeared in [DP08]. Fortunately, $\mathtt{Metric}^*$ entropy can be converted, with some relatively small loss in s and ε, to HILL entropy ([BSW03, Theorem 5.2],[FR11]). A similar statement, but with the conversion to HILL entropy already performed, appeared in [RTTV08].

An alternative statement, in which the circuit size (rather than the distinguishability ε) loses a factor polynomial in 2^b, is implied by [GW11, Lemma 3.1] and Lemma 2. Again, the statement is not with respect to HILL conditional entropy of Definition 5, but rather with respect to a relaxed notion that I will denote here $\mathsf{HILL\text{-}relaxed}$. It is the same as conditional HILL, except we are allowed to change not just X, but the entire pair (X,Z) to an indistinguishable pair (Y,W).

Theorem 2 ([GW11])**.** *Assume elements of Z are length-b bit strings (or, more generally, can be enumerated in time $\mathrm{poly}(2^b)$). Then*

$$H^{\mathsf{HILL\text{-}relaxed}}_{2\varepsilon,s'/\mathrm{poly}(\varepsilon,2^b)}(X|Z) \geq H^{\mathsf{HILL}}_{\varepsilon,s}(X) - b\,.$$

This theorem extends to the case when the initial entropy of X is *conditional* $\mathsf{HILL\text{-}relaxed}$ (conditioned on some Z_1), similarly to the more general case of Lemma 2.

References

[BSW03] Boaz Barak, Ronen Shaltiel, and Avi Wigderson. Computational analogues of entropy. In Sanjeev Arora, Klaus Jansen, José D. P. Rolim, and Amit Sahai, editors, *RANDOM-APPROX 2003*, volume 2764 of *LNCS*, pages 200–215. Springer, 2003.

[CKOR10] Nishanth Chandran, Bhavana Kanukurthi, Rafail Ostrovsky, and Leonid Reyzin. Privacy amplification with asymptotically optimal entropy loss. In Leonard J. Schulman, editor, *STOC*, pages 785–794. ACM, 2010. Full version available from `http://www.cs.bu.edu/fac/reyzin/`.

[DORS08] Yevgeniy Dodis, Rafail Ostrovsky, Leonid Reyzin, and Adam Smith. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. *SIAM Journal on Computing*, 38(1):97–139, 2008. arXiv:cs/0602007.

[DP08] Stefan Dziembowski and Krzysztof Pietrzak. Leakage-resilient cryptography. In Ravi [Rav08], pages 293–302.

[FR11] Benjamin Fuller and Leonid Reyzin. Computational entropy and information leakage. Available from `http://www.cs.bu.edu/fac/reyzin`, 2011.

[GW11] Craig Gentry and Daniel Wichs. Separating succinct non-interactive arguments from all falsifiable assumptions. In Salil Vadhan, editor, *STOC*. ACM, 2011.

[HILL99] J. Hrastad, R. Impagliazzo, L.A. Levin, and M. Luby. Construction of pseudorandom generator from any one-way function. *SIAM Journal on Computing*, 28(4):1364–1396, 1999.

[HLR07] Chun-Yuan Hsiao, Chi-Jen Lu, and Leonid Reyzin. Conditional computational entropy, or toward separating pseudoentropy from compressibility. In Moni Naor, editor, *Advances in Cryptology—EUROCRYPT 2007*, volume 4515 of *LNCS*, pages 169–186. Springer, 2007.

[KR09] Bhavana Kanukurthi and Leonid Reyzin. Key agreement from close secrets over unsecured channels. In Antoine Joux, editor, *Advances in Cryptology—EUROCRYPT 2009*, volume 5479 of *LNCS*. Springer, 2009. Full version available at `http://eprint.iacr.org/2008/494`.

[NZ96] Noam Nisan and David Zuckerman. Randomness is linear in space. *Journal of Computer and System Sciences*, 52(1):43–53, 1996.

[Rav08] R. Ravi, editor. *49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008, October 25-28, 2008, Philadelphia, PA, USA*. IEEE Computer Society, 2008.

[Rey11] Leonid Reyzin. Some notions of entropy for cryptography - (invited talk). In Serge Fehr, editor, *ICITS*, volume 6673 of *Lecture Notes in Computer Science*, pages 138–142. Springer, 2011.

[RTTV08] Omer Reingold, Luca Trevisan, Madhur Tulsiani, and Salil P. Vadhan. Dense subsets of pseudorandom sets. In Ravi [Rav08], pages 76–85.

[RW04] Renato Renner and Stefan Wolf. Smooth Rényi entropy and applications. In *Proceedings of IEEE International Symposium on Information Theory*, page 233, June 2004.

[RW05] Renato Renner and Stefan Wolf. Simple and tight bounds for information reconciliation and privacy amplification. In Bimal Roy, editor, *Advances in Cryptology—ASIACRYPT 2005*, LNCS, Chennai, India, 4–8 December 2005. Springer-Verlag.

[Vad11] Salil Vadhan. Pseudorandomness: Draft survey/monograph. `http://people.seas.harvard.edu/~salil/pseudorandomness/`, 2011.

[Yao82] A. C. Yao. Theory and applications of trapdoor functions. In *23rd Annual Symposium on Foundations of Computer Science*, pages 80–91, Chicago, Illinois, 3–5 November 1982. IEEE.