# Optimal Overcoming Weak Expectations*

Maciej Skorski

IST Austria

**Abstract.** Barak et al. (CRYPTO'11) initiated the study of so called *square-friendly applications* which offer good security for keys with entropy deficiency (weak keys), for this reason being important for key derivation. The state of the art of security bounds was established by Dodis and Yu (TCC'13), by modelling "weak" keys as distributions of high collision entropy. In this paper we answer the question what is the *minimum* requirement on weak keys to be "good" for these applications. The answer gives an elegant operational meaning to the notion of *smooth collision entropy*. Namely, smooth collision entropy is both sufficient and necessary (with essentially the same entropy parameters) to guarantee the security of square-friendly applications under weak keys. This characterization is a consequence of constrained optimization techniques.

**Keywords:** key derivation, square-friendly applications, weak expectations, smooth collision entropy

## 1 Introduction

### 1.1 Square-friendly applications

Most cryptographic objects require randomness for seeding, salting or derivation of secure keys. Since ideal randomness is hardly accessible, one has to retrieve it from *imperfect sources* (such as biometrics or other noisy data). As a result, the random bit string passed to a cryptographic application (which we refer to as *key* without loss of generality) is not truly random, but may have bias or entropy deficiency. If this bias or deficiency is small enough, one might hope to achieve security comparable as with the perfectly random key. The security of cryptographic applications is formally defined by quantifying the advantage that an attacker (with certain computational resources) may achieve, averaged over the distribution of the key. Therefore, to understand the impact of weak randomness on the security, one needs to compare the theoretical security under the uniform distribution $U$ over $\{0,1\}^m$

$$\epsilon_{ideal} = \mathbb{E}_{r \leftarrow U} \mathsf{Adv}^{\mathsf{A}}(r)$$

with the real security under the actual key distribution $R$ over $\{0,1\}^m$

$$\epsilon_{\text{real}} = \mathbb{E}_{r \leftarrow R} \mathsf{Adv}^{\mathsf{A}}(r),$$

* The paper is available at eprint.

where by $\mathsf{Adv}^{\mathsf{A}}(r)$ we denote $\mathsf{A}$'s advantage conditioned on the key being $r$.

From a technical point of view, the problem boils down to comparing the expected advantage under the uniform distribution and under the weak key $R$. As observed by Barak et al. [BDKPP+11], and made more explicit by Dodis and Yu [DY13], real and ideal security can be related by the following inequality

$$\epsilon_{\text{real}} \leqslant \epsilon_{\text{ideal}} + \sqrt{\mathbb{Var}\left[\mathsf{Adv}^{\mathsf{A}}(U)\right]} \cdot \sqrt{2^d - 1} \tag{1}$$

where $d$ is the gap between the collision entropy of $R$ and the (full) entropy of $U$ (referred to as *deficiency*). In general, the advantage is trivially bounded by 1. As a consequence we can get $\epsilon_{\text{real}} \approx \epsilon_{\text{ideal}}$ if $d \approx \epsilon^2_{\text{ideal}}$[1]. Retrieving an $m$-bit key with defficiency that small is possible by the use of *randomness extractors* [IZ89], however they require at least $m + 2\log(1/\epsilon_{\text{ideal}})$ bits of entropy in the source (as shown by the so called RT-bound [RT00]). This essentially means the loss of $L \approx 2\log(1/\epsilon_{\text{ideal}})$ bits of entropy[2], as the necessary entropy exceeds the key length by $L$. This loss may be a serious problem in settings where available entropy is limited, e.g. in the case of biometric sources.

Fortunately, for a broad class of cryptographic applications the variance term can be shown to be much smaller. In particular, if we know that

$$\mathbb{Var}\left[\mathsf{Adv}^{\mathsf{A}}(U)\right] \leqslant \sigma \tag{2}$$

for some $\sigma \ll 1$, then from Equation (1) we obtain

$$\epsilon_{\text{real}} \leqslant \epsilon_{\text{ideal}} + \sqrt{\sigma} \cdot \sqrt{2^d - 1} \tag{3}$$

and thus $\epsilon_{\text{real}} \approx \epsilon_{\text{ideal}}$ when $d \approx \frac{\epsilon^2_{\text{ideal}}}{\sigma}$, which reduces the entropy loss (in the use of an extractor) to $L \approx 2\log(1/\epsilon_{\text{ideal}}) - \log(1/\sigma)$. Moreover, with the use of *randomness condensers* one can achieve security $\epsilon_{\text{real}} \approx \epsilon_{\text{ideal}} + O\left(\sqrt{2^d\sigma}\right)$ from any source having $m - d$ bits of entropy[3]. For $\sigma \ll 1$ and appropriate $d$ this gives meaningful security with entropy smaller than the key length.

Cryptographic applications where there exists a non-trivial bound on $\sigma$ in are called *square-friendly applications*. Concrete examples include stateless chosen plaintext attack (CPA) secure encryption or weak pseudo-random functions [DY13]. Specifically, for these cases $\sigma = O(\epsilon_{\text{ideal}})$ and the entropy loss for extraction reduces to $L \approx \log(1/\epsilon_{\text{ideal}})$ (by half) whereas condensing gives meaningful results for deficiency $d < \log(1/\epsilon_{\text{ideal}})$.

Since entropy is generaly a limited and precious resource, square-friendly applications are of special importance to key derivation, as they allow for better

[1] We have $\sqrt{2^d - 1} = \sqrt{O(d)} = O(\sqrt{d})$ if $d = O(1)$. If $\mathbb{Var}\left[\mathsf{Adv}^{\mathsf{A}}(U)\right]$ then the formula gives $\epsilon_{\text{real}} = \epsilon_{\text{ideal}} + O(\sqrt{d})$ and hence $\epsilon_{\text{real}} = \Theta(\epsilon_{\text{ideal}})$ when $d = O(\epsilon^2_{\text{ideal}})$.

[2] Matched by the extractor built from universal hash families [ILL89]

[3] Seeded condensers get an $m$-bit string of entropy $m - d$ out of a source of entropy $m - d + O(1)$. This can be achieved with highly-independent hash families [DPW14]

tradeoffs between entropy and security [BDKPP+11; YS13]. They also allow for improving (provable) security of schemes utilizing square-friendly objects as building blocks, such as leakage-resilient stream ciphers [YS13; JP14].

### 1.2 Problem statement

The original motivation for studying square-friendly applications was *saving entropy* [BDKPP+11; DY13]. In this paper we keep exploring this research direction, asking a more fundamental question

**Q**: What key distributions $R$ guarantee a given security level (of $\epsilon$)?

The previous works [BDKPP+11; DY13] require keys to be of sufficiently high entropy, where the entropy notion is collision entropy (less restrictive than min-entropy). While this assumption may work well in theory, classical entropy notions are not robust against even small biases that may easily occur when evaluating entropy in practice - and hence may not capture the security properties accurately. For example, entropy estimation requires addressing statistical errors [TBKM16] which lower the estimated entropy. Bias may be also introduced by measurements or other noise processes, resulting similarly in underestimation (hence a waste) of entropy. To illustrate this with a concrete example, consider a random variable $R$ that takes a fixed value $x_0$ with probability $\epsilon$ and is uniform over $m$ bits with probability $1-\epsilon$. While $R$ is practically indistinguishable from the uniform distribution for small $\epsilon$ (say $\epsilon \approx 2^{-80}$) and thus equally good from a cryptographic point of view, it has only $\log(1/\epsilon) \ll m$ bits of min-entropy. Thus, motivated by the idea of saving entropy, we restate our problem as

**Q'**: what's the weakest entropy in $R$ to guarantee security of $\epsilon$?

A similar problem was studied in the context of randomness extraction, and the so called *smooth min-entropy* was introduced [RW05] as the notion that *optimally* characterize the number of extractable uniform bits. Smooth entropy essentially removes small bias before evaluating the number of entropy bits. In this paper, we use a conceptually similar notion to *optimally characterize* security in the context of key derivation. It turns out that for our setting (square-friendly applications), the *optimal* entropy notion is the *smoothed* version of *collision entropy*.

### 1.3 Results and techniques

*The optimal entropy notion is smooth collision entropy* We start by observing that the entropy notion used to define $d$ in Equation (1) can be relaxed to *smooth collision entropy*. Namely we get $\epsilon_{\text{real}} \leqslant \epsilon_{\text{ideal}} + O\left(\sqrt{\sigma} \cdot \sqrt{2^d - 1}\right)$ when $R$ is only $\epsilon$-close (in total variation) to a distribution of entropy deficiency $d$, for $\epsilon = O(\sqrt{\sigma} \cdot \sqrt{2^d - 1})$. Since the square-root term typically dominates $\epsilon_{\text{ideal}}$, this bound is comparable to the one in Equation (3). While this observation is easy, our main contribution is the proof of the converse part. Recall that an application

is said to be $\sigma$-square-secure against an attacker $\mathsf{A}$ when $\mathbb{E}\left(\mathsf{Adv}^{\mathsf{A}}(r)\right)^2 \leqslant \sigma$, which in particular implies Equation (2) (see Section 2 for formal definitions). The main result of this paper is the following theorem

**Theorem (Optimal keys for square-friendly applications).** *Then the following holds*

*(a) Smooth collision-entropy is sufficient: Suppose that an $m$-bit distribution $R$ is $O(\epsilon)$-close to a distribution with collision entropy deficiency $d$, where $2^d - 1 = O\left(\sigma\epsilon^{-2}\right)$. Then for any application $P$ and attacker $\mathsf{A}$ against whom $P$ is $\sigma$-square secure, the application $P$ is $\epsilon_{\text{ideal}} + O(\epsilon)$ secure under $R$.*

*(b) Smooth collision-entropy is necessary: Suppose that for a key distribution $R$, for every application $P$ and attacker $\mathsf{A}$ such that $P$ is $\sigma$-square secure against $\mathsf{A}$, the application $P$ is $\epsilon_{\text{ideal}} + \epsilon$-secure under $R$. Then $R$ is $O(\epsilon)$-close to a distribution with collision entropy deficiency $d$, where $2^d - 1 = O\left(\sigma\epsilon^{-2}\right)$.*

The proof if the easy part (a) appears in Section 3.1. We prove the non-trivial part (b) in Section 3.3; its proof is based on a more general result about comparing keys for square-friendly applications, which we discuss in the next paragraph.

We note that this theorem gives an *operational meaning* to smooth collision entropy. Formally, $\epsilon$-smooth collision entropy is defined as the maximum entropy of a distribution within distance $\epsilon$ (see Section 2). Thus, the condition on $R$ can be expressed as "$R$ has $|R| - d$ bits of $O(\epsilon)$-smooth collision entropy". In view of the above theorem, smooth collision entropy is both necessary and sufficient for a key to be "good" for square-friendly applications, with essentially no gap between quantitative parameters in both statements. To our knowledge, this is the first such functional characterization for smooth collision entropy.

*Characterizing equally secure keys* In order to prove our main result (the necessary part), we solve a more general problem of characterizing keys that provide comparably security for square-friendly applications. Intuitively, if two key distributions $X, Y$ provide a similar level of security, their shapes should be related by a *small perturbation in probability mass*. Our result below provides a quantitative and optimal characterization of this sort (note the characterization is somewhat more complicated that one might expect).

**Theorem (Characterizing keys with similar security).** *Suppose that for two key distributions $X, Y$ over $m$ bits, some number $\epsilon > 0$ and every pair $(P, \mathsf{A})$ such that $P$ is $\sigma$-square secure against $\mathsf{A}$, the security $\epsilon_X$ under the key $X$ and the security $\epsilon_Y$ under the key $Y$ satisfy*

$$\epsilon_Y \leqslant \epsilon_X + \epsilon.$$

*Then there exists a distribution $Z$ such that*

*(a) $Y$ and $Z$ are $O(\epsilon)$-close in the $\ell_1$ distance*

*(b) $Z$ and $X$ are $O(2^{-\frac{m}{2}}\sigma^{-\frac{1}{2}}\epsilon)$-close in the $\ell_2$ distance*

*Conversely, for $Z$, $P$ and $\mathsf{A}$ as above we have $\epsilon_Y \leqslant \epsilon_X + O(\epsilon)$.*

This theorem follows from results proved in Section 3.2. We illustrate it in the diagram in Figure 1, and note that the previous theorem is obtained by the application to the setting when $Y = R$ and $X = U$.

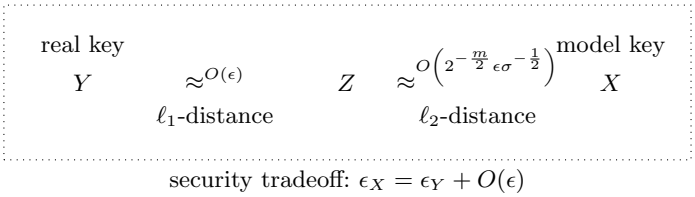


Fig. 1: Relation between key distributions $X$ and $Y$ that provide similar security for square-secure applications. A transformation from $X$ to $Y$ involves probability mass "smoothing" first in $\ell_1$-distance and next in $\ell_2$-distance.

Not discussing the technicalities in full detail, we would like to provide some intuitions and more insight into proof techniques. The problem boils down to explaining why the difference of the expectations $\mathbb{E}\mathsf{D}(X) - \mathbb{E}\mathsf{D}(Y)$ is small when $\mathsf{D}$ is of small variance ($\mathsf{D}$ runs over different attacker advantage profiles). The first reason is that there is a small bias between $X$ and $Y$. This is addressed by passing from $Y$ to $Z$, which is essentially smoothing in the $\ell_1$ distance. However, there is another more subtle reason - when the bias may be substantially bigger but it changes over particular inputs adaptively to $\mathsf{D}$, so that the difference is small because of the variance of $\mathsf{D}$. To smooth it out, we use the $\ell_2$-distance, passing from $Z$ to $X$. With the help of the tools from convex optimization we prove that these "pathologies" are actually the only ones. We illustrate the proof technique in Figure 2 below.

### 1.4 Organization

We explain notions and definitions in Section 2. Our main result is proved in Section 3. In Section 4 we conclude our work.

## 2 Preliminaries

*Basic conventions* By $P_X$ we denote the probability mass function of a random variable $X$, that is $P_X(x) = \Pr[X = x]$. All logarithms are taken at base 2

*Distance metrics*

**Definition 1 ($\ell_1$-distance).** *The $\ell_1$ distance between two vectors $p, q \in \mathbb{R}^N$ is defined by*

$$d_1(p; q) = \sum_i |p_i - q_i|.$$

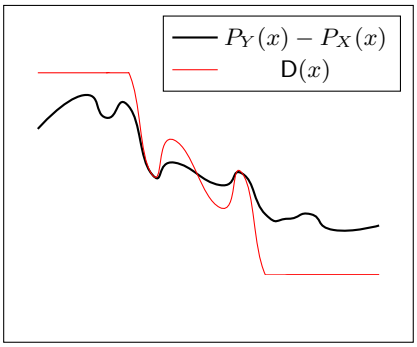


Fig. 2: Our technique. To maximize the difference $\mathbb{E}\mathsf{D}(X) - \mathbb{E}\mathsf{D}(Y)$ (when $\mathsf{D}$ is constrained by its variance), biggest values of $\mathsf{D}$ correspond to the biggest bias $\Delta(x) = P_X(x) - P_Y(x)$, while moderate values of $\mathsf{D}$ are proportional to moderate bias $\Delta(x)$. This behavior follows by the techniques of constrained optimization (KKT conditions).

**Definition 2 (statistical distance).** *The statistical distance between (distributions of) two random variables $X, Y$ taking values in the same finite set equals*

$$\mathsf{SD}(X;Y) = \frac{1}{2} d_1(P_X; P_Y) = \frac{1}{2} \sum_x |P_X(x) - P_Y(x)|$$

**Definition 3 ($\ell_2$-distance).** *The $\ell_2$ distance between two vectors $p, q \in \mathbb{R}^N$ is defined by*

$$d_2(p;q) = \sqrt{\sum_i (p_i - q_i)^2}.$$

*Entropy measures*

**Definition 4 (collision entropy).** *The* collision entropy *of a random variable $X$ is defined as*

$$H_2(X) = -\log \sum_x \Pr[X = x]^2$$

**Definition 5 (smooth collision entropy).** *For $\epsilon \geqslant 0$, the* $\epsilon$-smooth collision entropy *of a random variable $X$ is defined as*

$$H_2^\epsilon(X) = \max_{X' : \mathsf{SD}(X';X) \leqslant \epsilon} H_2(X').$$

*Security of cryptographic applications* We consider cryptographic application (primitive, protocol, scheme) $P$ where the security is defined by an *interactive game* between an attacker $\mathsf{A}$ and a challenger $\mathsf{C}$ (both probabilistic). The challenger uses a key $r$ sampled either from the uniform distribution $U_m$ (the "ideal"

setting) or a distribution $R$ (the "real" setting). The probability (over the internal coins of the attacker and challenger) that the attacker wins the game when the key is $r$ is denoted by $\Pr[\mathsf{A} \text{ wins } |r]$. Now we can formally define security

**Definition 6 (advantage).** *For a fixed cryptographic application, the advantage of an attacker* $\mathsf{A}$ *is defined as*

$$\mathsf{Adv}^{\mathsf{A}}(r) = \Pr[\mathsf{A} \textit{ wins } |r] - c$$

*where "*$\mathsf{A}$ *wins* $|r$*" denotes the event that an attacker wins the security game conditioned on the key used being* $r$*. For indistinguishability games* $c = \frac{1}{2}$ *and for unpredictability games* $c = 0$*.*

**Definition 7 (square-security [DY13]).** *An application is* $\sigma$*-square secure against a class of attackers* $\mathcal{A}$ *if for every* $\mathsf{A} \in \mathcal{A}$ *it holds that*

$$\mathbb{E}_{r \leftarrow U} \left(\mathsf{Adv}^{\mathsf{A}}(r)\right)^2 \leqslant \sigma.$$

*In particular,* $\sigma$*-square security implies*

$$\mathbb{V}\mathrm{ar}\left[\mathsf{Adv}^{\mathsf{A}}(U)\right] \leqslant \sigma.$$

## 3 Main results

### 3.1 Smooth entropy is sufficient for security of weak keys

We start with the following lemma, which slightly generalizes Lemma 2.2 from [BDKPP+11].

**Lemma 1.** *Let* $\mathcal{X}$ *be an* $N$*-element set and* $X$ *be a random variable over* $\mathcal{X}$*. Then for any* $\mathsf{D} : \mathcal{X} \to [-1, 1]$ *we have*

$$\mathbb{E}\mathsf{D}(X) \leqslant \mathbb{E}\mathsf{D}(U) + \sqrt{\mathbb{V}\mathrm{ar}\mathsf{D}(U)} \cdot \sqrt{2^d - 1} + \epsilon \tag{4}$$

*where* $U$ *is uniform over* $\mathcal{X}$ *and*

$$d = \log N - H_2^{\epsilon}(X) \tag{5}$$

*is the* smooth collision entropy deficiency*.*

*Proof.* By the definition of smooth collision entropy, there exist $X'$ such that $H_2(X') = \log N - d$ and $\mathsf{SD}(X'; X) \leqslant \epsilon$. Note that for any constant $c$ we have

$$\mathbb{E}\mathsf{D}(X') - \mathbb{E}\mathsf{D}(U) = \sum_x (\mathsf{D}(x) - c) \cdot (P_{X'}(x) - P_U(x))$$

By applying the Cauchy-Schwarz inequality we obtain

$$\mathbb{E}\mathsf{D}(X') \leqslant \mathbb{E}\mathsf{D}(U) + \sqrt{\mathbb{E}_{x \leftarrow U}(\mathsf{D}(x) - c)^2} \cdot \sqrt{2^d - 1}.$$

Maximizing over the choice of $c$ and using the definition of variance yields

$$\mathbb{E}\mathsf{D}(X') \leqslant \mathbb{E}\mathsf{D}(U) + \sqrt{\mathbb{V}\mathbb{a}\mathbb{r}\mathsf{D}(U)} \cdot \sqrt{2^d - 1}.$$

Since $X$ and $X'$ are $\epsilon$-close, we have $\mathbb{E}\mathsf{D}(X) \leqslant \mathbb{E}\mathsf{D}(X') + \epsilon$ which finishes the proof.

As a corollary we immediately obtain that smooth collision entropy can replace the plain collision entropy for square-friendly applications.

**Corollary 1 (Security under smooth colision entropy).** *Fix an application that needs an $m$-bit key. Let $X$ be a key distribution and $d = m - H_2^\epsilon(X)$. Then for any attacker $\mathsf{A}$ such that the application is $\sigma$-square secure we have*

$$\epsilon_{\text{real}} \leqslant \epsilon_{\text{ideal}} + \sqrt{\sigma} \cdot \sqrt{2^d - 1} + \epsilon.$$

*In particular, when $\epsilon = O\left(\sqrt{\sigma} \cdot \sqrt{2^d - 1}\right)$ one has*

$$\epsilon_{\text{real}} \leqslant \epsilon_{\text{ideal}} + O\left(\sqrt{\sigma} \cdot \sqrt{2^d - 1}\right).$$

### 3.2 Characterizing equally secure weak keys

**Theorem 1 (Best key smoothing for square-friendly applications).** *Let $\mathcal{X}$ be an $N$-element set and $X, Y$ be random variables over $\mathcal{X}$ such that*

$$\mathbb{E}\mathsf{D}(Y) \leqslant \mathbb{E}\mathsf{D}(X) + \epsilon \tag{6}$$

*for all $\mathsf{D} : \mathcal{X} \to [-1, 1]$ such that*

$$\mathbb{V}\mathbb{a}\mathbb{r}\mathsf{D}(U) \leqslant \sigma \tag{7}$$

*Then for some random variable $Y'$ over $\mathcal{X}$ the following holds true:*

*(a) $Y$ and $Z$ are $4\epsilon$-close in $\ell_1$-distance*
*(b) $Z$ and $X$ are $\sqrt{\frac{2\epsilon^2}{N\sigma}}$-close in $\ell_2$-distance*

*Remark 1 (Variance vs expected square).* Note that $\mathbb{E}\mathsf{D}(Y) \leqslant \mathbb{E}\mathsf{D}(X) + \epsilon$ holds for $\mathsf{D}$ if and only if it holds for $\mathsf{D}'(x) \stackrel{def}{=} \mathsf{D}(x) + c$ where $c$ is a constant. Thus the condition in Equation (7) can be replaced by $\mathbb{E}_{r \leftarrow U}\mathsf{D}(r)^2 \leqslant \sigma$.

The result is optimal up to a constant factor in $O(\epsilon)$, as shown below. We skip the proof which is merely an extension of the proof of Lemma 1.

**Corollary 2 (The characterization is optimal up to constants).** *Let $X, Y, Z$ be random variables over $\mathcal{X}$ such that conditions (a) and (b) in Theorem 1. Then*

$$\mathbb{E}\mathsf{D}(Y) \leqslant \mathbb{E}\mathsf{D}(X) + (2 + \sqrt{2})\epsilon$$

*for all $\mathsf{D}$ such that $\mathbb{V}\mathbb{a}\mathbb{r}(\mathsf{D}) \leqslant \sigma$.*

*Proof (of Theorem 1).* Define $\Delta(x) = P_Y(x) - P_X(x)$. Let $\mathcal{X}^+$ contain inputs $x$ that correspond to the first $2^n\sigma$ greatest values of $\Delta(x)$. Similarly, let $\mathcal{X}^-$ contain inputs $x$ that correspond to the first $2^n\sigma$ smallest values of $\Delta(x)$. Let $\mathcal{X}_0 = \mathcal{X} \setminus \mathcal{X}^+ \cup \mathcal{X}^-$.

Ideas from convex optimization We can consider Equation (6) and Equation (7) as the optimization program of maximizing $\mathbb{E}\mathsf{D}(X) - \mathbb{E}\mathsf{D}(Y)$ under the variance assumption. By Remark 1 we know that the variance constraint can be replaced by the second moment constraint. By applying the necessary condition for the constrained optimization programs - Karush-Kuhn-Tucker Theorem (see for instance [BL05]) - the maximizing $\mathsf{D}$ must be of the following form

$$\mathsf{D}(x) = \max(\min(a \cdot \Delta(x) + b, 1), 1).$$

where $a$ and $b$ are some constants. In the further analysis we will apply the assumptions of the theorem to functions $\mathsf{D}$ fitting his shape.

Heavy weights Define $\mathsf{D}(x) = \operatorname{sgn}(\Delta(x))$ when $x \in \mathcal{X}^+$ and $\mathsf{D}(x) = 0$ otherwise. Note that $\mathsf{D}$ satisfies Equation (7). By Equation (6) and the definition of $\mathsf{D}$ we obtain

$$\epsilon \geqslant \mathbb{E}\mathsf{D}(Y) - \mathbb{E}\mathsf{D}(X) = \sum_{x \in \mathcal{X}^+} |\Delta(x)| \tag{8}$$

Since the smallest element in a set is not bigger than the average, we obtain

$$\min_{x \in \mathcal{X}^+} \Delta(x) \leqslant \frac{\epsilon}{N\sigma}. \tag{9}$$

Small weights Define $\mathsf{D}(x) = \operatorname{sgn}(\Delta(x))$ when $x \in \mathcal{X}^-$ and $\mathsf{D}(x) = 0$ otherwise. Clearly $\mathsf{D}$ satisfies Equation (7). By Equation (6) and the definition of $\mathsf{D}$ we obtain

$$\epsilon \geqslant \mathbb{E}\mathsf{D}(Y) - \mathbb{E}\mathsf{D}(X) = \sum_{x \in \mathcal{X}^-} |\Delta(x)|. \tag{10}$$

Estimating the maximum element in a set by the average, and combining this with the triangle inequality we obtain

$$\begin{aligned} \max_{x \in \mathcal{X}^-} \Delta(x) &\geqslant \frac{\sum_{x \in \mathcal{X}^-} \Delta(x)}{\#\mathcal{X}^-} \\ &\geqslant -\frac{\sum_{x \in \mathcal{X}^-} |\Delta(x)|}{\#\mathcal{X}^-} \\ &\geqslant -\frac{\epsilon}{N\sigma}. \end{aligned} \tag{11}$$

Moderate weights Define $\mathsf{D}(x)$ in the following way

$$\mathsf{D}(x) = \begin{cases} \Delta(x) \cdot \frac{N\sigma}{\epsilon}, x \in \mathcal{X}_0 \\ 0, x \notin \mathcal{X}_0 \end{cases}$$

Note that $|\mathsf{D}(x)| \leqslant 1$ by Equation (11) and Equation (9). Note that $\mathsf{D}$ satisfies Equation (7). By Equation (6) and the definition of $\mathsf{D}$ we obtain

$$\epsilon \geqslant \mathbb{E}\mathsf{D}(Y) - \mathbb{E}\mathsf{D}(X) = \frac{N\sigma}{\epsilon} \sum_{x \in \mathcal{X}_0} \Delta(x)^2, \tag{12}$$

and therefore we conclude that

$$\sum_{x \in \mathcal{X}_0} \Delta(x)^2 \leqslant \frac{\epsilon^2}{N\sigma}. \tag{13}$$

Smoothing. Let $Z$ be a random variable with the following distribution

$$P_Z(x) = \begin{cases} P_Y(x) - \Delta(x) + \Delta\,, & x \in \mathcal{X}^- \cup \mathcal{X}^+ \\ P_Y(x)\,, & x \in \mathcal{X}_0. \end{cases}$$

where

$$\Delta = \frac{\sum_{x \in \mathcal{X}^- \cup \mathcal{X}^+} \Delta(x)}{\#\mathcal{X}^- + \#\mathcal{X}^+}$$

is chosen so that $P_Z$ is the probability measure. By Equation (8) and Equation (10) we obtain

$$d_1(P_Y; P_Z) = \sum_{x \in \mathcal{X}^- \cup \mathcal{X}^+} |\Delta(x) - \Delta| \leqslant 4\epsilon.$$

Moreover

$$(d_2(P_Z; P_X))^2 =^{(a)} \sum_{x \in \cup \mathcal{X}^- \cup \mathcal{X}^+} \Delta^2 =^{(b)} \frac{\left(\sum_{x \in \mathcal{X}^- \cup \mathcal{X}^+} \Delta(x)\right)^2}{\#\mathcal{X}^- + \#\mathcal{X}^+} \leqslant \frac{2\epsilon^2}{N\sigma}.$$

where (a) follows by the definition of $\Delta$ and (b) follows by Equations (8) and (10) and the fact that $\#\mathcal{X}^- = \#\mathcal{X}^+ = N\sigma$. □

### 3.3 Characterizing good weak keys

**Theorem 2.** *Let $\mathcal{X}$ be an $N$-element set, $Y$ be a random variable over $\mathcal{X}$ and $U$ be uniform over $\mathcal{X}$. Suppose that*

$$\mathbb{E}\mathsf{D}(Y) \leqslant \mathbb{E}\mathsf{D}(U) + \epsilon$$

*for all $\mathsf{D} : \mathcal{X} \to [-1, 1]$ such that $\mathbb{V}\mathrm{ar}\mathsf{D}(U) \leqslant \sigma$. Then*

$$H_2^{2\epsilon}(Y) \geqslant \log N - d$$

*where*

$$2^d - 1 = \frac{2\epsilon^2}{\sigma}$$

From Theorem 2, applied to $\mathsf{D}$ being all square-secure advantage profiles[4], we immediately obtain the following corollary

**Corollary 3 (Smooth colision entropy is necessary for security under weak keys).** *Let $X$ be an $m$-bit key. Suppose that for every application $P$ and every attacker $\mathsf{A}$ such that $P$ is $\sigma$-secure against $\mathsf{A}$, under the key $X$ one has*

$$\epsilon_{\text{real}} \leqslant \epsilon_{\text{ideal}} + \epsilon$$

*Then $H_2^{2\epsilon}(X) \geqslant m - d$ where*

$$2^d - 1 = O\left(\epsilon^2 \sigma^{-1}\right)$$

*Proof (Proof of Theorem 2).* By Theorem 1 and Remark 1 there is $Z$ which is $4\epsilon$-close to $X$ in $d_1$ and $\epsilon'$-close to $U$ in $d_2$ where $(\epsilon')^2 = \frac{2\epsilon^2}{N\sigma}$. We obtain

$$\begin{aligned}
\sum_x P_Z(x)^2 - \frac{1}{N} &=^{(a)} \sum_x P_Z(x)^2 + 2\sum_x P_U(x)P_Z(x) - \sum_x P_U(x)^2 \\
&=^{(b)} \sum_x \left(P_Z(x) - P_U(x)\right)^2 \\
&\leqslant^{(c)} \frac{2\epsilon^2}{N\sigma}
\end{aligned}$$

where (a) follows because $\sum_x P_Z(x) = 1$ and $P_U(x) = \frac{1}{N}$, (b) is the elementary identity and (c) follows by the assumption on $Z$. This means that

$$H_2(Z) = \log N - \log\left(1 + \frac{2\epsilon^2}{\sigma}\right).$$

and, since $\mathsf{SD}(Y; Z) \leqslant 2\epsilon$ we get

$$H_2^{2\epsilon}(Y) \geqslant \log N - \log\left(1 + \frac{2\epsilon^2}{\sigma}\right).$$

which completes the proof.

## 4 Conclusion

In this paper we showed a functional characterization of *smooth colision entropy* in the context of key derivation. Namely, it provides optimal bounds for the security of square-friendly applications fed with weak keys. Our result is complementary to the previous works of Barak at al. [BDKPP+11] and Dodis, Yu [DY13].

[4] The advantage for idistinguishability games has the range $\left[-\frac{1}{2}, \frac{1}{2}\right]$, and for unpredictability it is $[0, 1]$. Therefore, we need to scale $\mathsf{D}$ (by an affine transform) to the range $[-1, 1]$ when applying Theorem 2. This slightly changes constants under $O(\cdot)$.

## References


[BDKPP+11] B. Barak, Y. Dodis, H. Krawczyk, O. Pereira, K. Pietrzak, F. Standaert, and Y. Yu. "Leftover Hash Lemma, Revisited". In: *Advances in Cryptology - CRYPTO 2011 - 31st Annual Cryptology Conference, Santa Barbara, CA, USA, August 14-18, 2011. Proceedings*. 2011, pp. 1–20.

[BL05] J. Borwein and A. Lewis. *Convex Analysis and Nonlinear Optimization: Theory and Examples*. CMS Books in Mathematics. Springer New York, 2005.

[DPW14] Y. Dodis, K. Pietrzak, and D. Wichs. "Key Derivation without Entropy Waste". In: *Advances in Cryptology - EUROCRYPT 2014 - 33rd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Copenhagen, Denmark, May 11-15, 2014. Proceedings*. 2014, pp. 93–110.

[DY13] Y. Dodis and Y. Yu. "Overcoming Weak Expectations". In: *Theory of Cryptography - 10th Theory of Cryptography Conference, TCC 2013, Tokyo, Japan, March 3-6, 2013. Proceedings*. 2013, pp. 1–22.

[ILL89] R. Impagliazzo, L. A. Levin, and M. Luby. "Pseudo-random Generation from one-way functions (Extended Abstracts)". In: *Proceedings of the 21st Annual ACM Symposium on Theory of Computing, May 14-17, 1989, Seattle, Washigton, USA*. 1989, pp. 12–24.

[IZ89] R. Impagliazzo and D. Zuckerman. "How to Recycle Random Bits". In: *30th Annual Symposium on Foundations of Computer Science, Research Triangle Park, North Carolina, USA, 30 October - 1 November 1989*. 1989, pp. 248–253.

[JP14] D. Jetchev and K. Pietrzak. "How to Fake Auxiliary Input". In: *Theory of Cryptography - 11th Theory of Cryptography Conference, TCC 2014, San Diego, CA, USA, February 24-26, 2014. Proceedings*. 2014, pp. 566–590.

[RT00] J. Radhakrishnan and A. Ta-Shma. "Bounds for Dispersers, Extractors, and Depth-Two Superconcentrators". In: *SIAM J. Discrete Math.* 13.1 (2000), pp. 2–24.

[RW05] R. Renner and S. Wolf. "Simple and Tight Bounds for Information Reconciliation and Privacy Amplification". In: *Advances in Cryptology - ASIACRYPT 2005, 11th International Conference on the Theory and Application of Cryptology and Information Security, Chennai, India, December 4-8, 2005, Proceedings*. 2005, pp. 199–216.

[TBKM16] M. S. Turan, E. Barker, J. Kelsey, and K. McKay. "NIST DRAFT Special Publication 800-90B Recommendation for the Entropy Sources Used for Random Bit Generation". In: http://csrc.nist.gov/publications/drafts/800-90/sp800-90b_second_draft.pdf. 2016.

[YS13] Y. Yu and F. Standaert. "Practical Leakage-Resilient Pseudorandom Objects with Minimum Public Randomness". In: *Topics in Cryptology - CT-RSA 2013 - The Cryptographers' Track at the RSA Conference 2013, San Francisco,CA, USA, February 25-March 1, 2013. Proceedings*. 2013, pp. 223–238.