A Unified Approach to Deterministic Encryption: New Constructions and a Connection to Computational Entropy*

Benjamin Fuller† Adam O'Neill‡ Leonid Reyzin§

Boston University

August 21, 2012

Abstract

We propose a general construction of deterministic encryption schemes that unifies prior work and gives novel schemes. Specifically, its instantiations provide:

- A construction from any trapdoor function that has sufficiently many hardcore bits.
- A construction that provides "bounded" multi-message security from lossy trapdoor functions.

The security proofs for these schemes are enabled by three tools that are of broader interest:

- A weaker and more precise sufficient condition for semantic security on a high-entropy message distribution. Namely, we show that to establish semantic security on a distribution M of messages, it suffices to establish indistinguishability for all conditional distribution $M|\mathsf{E}$, where E is an event of probability at least $1/4$. (Prior work required indistinguishability on *all* distributions of a given entropy.)
- A result about computational entropy of conditional distributions. Namely, we show that conditioning on an event E of probability p reduces the quality of computational entropy by a factor of p and its quantity by $\log_2 1/p$.
- A generalization of leftover hash lemma to correlated distributions.

We also extend our result about computational entropy to the average case, which is useful in reasoning about leakage-resilient cryptography: leaking λ bits of information reduces the quality of computational entropy by a factor of 2^λ and its quantity by λ.

*A short version of this work appeared at Ninth IACR Theory of Cryptography Conference, March 2012

†Email: `bfuller@cs.bu.edu`. Work done in part while the author was at MIT Lincoln Laboratory

‡Email: `amoneill@bu.edu`. Work done in part while the author was at Georgia Institute of Technology and University of Texas at Austin.

§Email: `reyzin@cs.bu.edu`.

1 Introduction

Public-key cryptosystems require randomness: indeed, if the encryption operation is deterministic, the adversary can simply use the public key to verify that the ciphertext c corresponds to its guess of the plaintext m by encrypting m. However, such an attack requires the adversary to have a reasonably likely guess for m in the first place. Recent results on deterministic public-key encryption (DE) (building on previous work in the one-time, information-theoretic symmetric-key setting [43, 18, 14]) have studied how to achieve security when the randomness comes only from m itself [3, 5, 7, 28, 8, 48]. DE has a number of practical applications, such as efficient search on encrypted data and securing legacy protocols (cf. [3]). It is also interesting from a foundational standpoint; indeed, its study has proven useful in other contexts: Bellare et al. [4] showed how it extends to a notion of "hedged" public-key encryption that reduces dependence on external randomness for probabilistic encryption more generally, and Dent et al. [13] adapted its notion of privacy to a notion of confidentiality for digital signatures.

However, our current understanding of DE is somewhat lacking. The constructions of [3, 5, 7, 28], as well as their analysis techniques, are rather disparate, and some natural questions arise from them. Namely, does the scheme of [5] inherently require using the Goldreich-Levin hardcore bit? Can it be made to work with trapdoor functions rather than permutations? Is the single-message security achieved by [5, 7, 28] an inherent limitation of standard model (i.e., non-random-oracle) schemes? In this work our main goal is to provide a *unified framework* for the construction of DE and to shed light on these questions.

1.1 Our Results

A SCHEME BASED ON TRAPDOOR FUNCTIONS. We propose a general *Encrypt-with-Hardcore* ($\mathsf{EwHCore}$) construction of DE from trapdoor functions (TDFs), which generalizes the basic idea behind the schemes of [3, 5] and leads to a unified framework for the construction of DE. Let f be a TDF with a hardcore function hc, and let $\mathcal{E}$ be any probabilistic public-key encryption algorithm. Our construction $\mathsf{EwHCore}$ encrypts an input message x as follows: it computes $y = f(x)$ and then encrypts y using $\mathcal{E}$ with $\mathsf{hc}(x)$ as the coins; that is, the encryption of x is $\mathcal{E}(f(x); \mathsf{hc}(x))$.

Intuitively, this scheme requires that the output of hc be sufficiently long to provide enough random coins for $\mathcal{E}$ (in fact, it need only be sufficiently long to be used as a seed for a psuedorandom generator), and that it not reveal any partial information about x (because $\mathcal{E}$ does not necessarily protect the privacy of its random coins). There are two nontrivial technical steps needed to make intuition precise. First, we define a condition required of hc (which we call "robustness") and show that it is sufficient for security of the resulting DE. Second, through a computational entropy argument, we show how to make *any* sufficiently long hc robust by applying a randomness extractor.

This general scheme admits a number of instantiations depending of f and hc. For example, when f is any trapdoor function and hc is a random oracle (RO), we obtain the construction of [3][1]. When f is an iterated trapdoor permutation (TDP) and hc is a collection Goldreich-Levin (GL) [24] bits extracted at each iteration, we obtain the construction of [5]. When f is a lossy trapdoor function (LTDF) [38] and hc is a pairwise-independent hash, we get a variant of the construction of [7] (which is less efficient but has a more straightforward analysis). We also obtain a variant of the construction of Hemenway et al. [28] under the same assumption as they use (see Section 5.2 for details). Note that in all but the last of these cases, the hardcore function is *already* robust (without requiring an extractor), which shows that in prior work this notion played an implicit role.

Moreover, this general scheme not only explains past constructions, but also gives us new ones. Specifically, if f is a trapdoor function with enough hardcore bits, we obtain:

[1] Technically, this construction does not even need a TDF because of the random oracle model; however, it may be prudent to use a TDF because then it seems more likely that the instantiation of the random oracle will be secure as it may be hardcore for the TDF.

- DE that works on the uniform distribution of messages;
- DE that works on any distribution of messages whose min-entropy is at most logarithmically smaller than maximum possible;
- assuming sufficient hardness distinguishing the output of hc from uniform (so in particular of inverting f), DE that works on even-lower entropy message distributions.

Prior results require more specific assumptions on the trapdoor function (such as assuming that it is a permutation or that it is lossy—both of which imply enough hardcore bits) in order to get constructions that work even just on the uniform distribution of messages. Furthermore, our results yield more efficient schemes (though sometimes under stronger assumptions) even in the permutation case, by avoiding iteration.

Notably, we obtain the *first* DE scheme without random oracles based on the hardness of syndrome decoding using the Niederreiter trapdoor function [35], which was shown to have linearly many hardcore bits by Freeman et al. [21] (and moreover to be secure under correlated products, as defined by Rosen and Segev [42]) but is not known to be lossy. (A scheme in the random oracle model follows from [3].) Additionally, the RSA [41] and Paillier [37] trapdoor permutations have linearly many hardcore bits under certain computational assumptions (the "Small Solutions RSA" [45] and "Bounded Computational Composite Residuosity" [9] assumptions respectively). Therefore, we can use these TDPs to instantiate our scheme efficiently under the same computational assumptions. Before our work, DE schemes from RSA and Paillier either required many iterations [5] or decisional assumptions that imply lossiness of these TDPs [31, 21, 7].

SECURITY FOR MULTIPLE MESSAGES: DEFINITION AND CONSTRUCTION. An important caveat is that, as in [5, 7], we can prove the above standard-model DE schemes secure only for the encryption of a *single* high-entropy plaintext, or, what was shown equivalent in [7], an unbounded number of messages drawn from a *block source* [10], where each subsequent message brings "fresh" entropy. On the other hand, the strongest and most practical security model for DE introduced by [3] considers the encryption of an unbounded number of plaintexts that have individual high entropy but may not have any conditional entropy. In order for $\mathsf{EwHCore}$ to achieve this, the hardcore function hc must also be robust on *correlated inputs*. (A general study of correlated-input security for the case of hash functions rather than hardcore functions was concurrently initiated in [26].) In particular, it follows from the techniques of [3] that a RO hash satisfies such a notion. This leads to a multi-message secure scheme in the RO model (as obtained in [3]). We thus have a large gap between what is (known to be) achievable with random oracles versus in the standard model.

To help bridge this gap, we propose a notion of "q-bounded" security for DE, where up to q high-entropy but arbitrarily correlated messages may be encrypted under the same public key (whose size may depend polynomially on q). We feel that if one is limited to the standard model, this notion is useful. Indeed, it seems that the requirement of previous results in the standard model—that messages come from a block source—may be difficult to guarantee: all that's needed to violate it is a single message that has low conditional entropy. Following [7], we also extend our security definition to unbounded multi-message security where messages are drawn from what we call a "q-block source" (essentially, a block source where each "block" consists of q messages which may be arbitrarily correlated but have individual high entropy); Theorem 4.2 of [7] extends to show that q-bounded multi-message security and unbounded multi-message security for q-block sources are equivalent for a given min-entropy.

Using our $\mathsf{EwHCore}$ construction and a generalization of the leftover hash lemma discussed below, we show q-bounded DE schemes (for long enough messages), for any polynomial q, based on LTDFs losing an $1-O(1/q)$ fraction of the input. It is known how to build such LTDFs rom the decisional Diffie-Hellman [38], d-linear [21], and decisional composite residuosity [7, 21] assumptions.

We note that a subsequent result of Wichs [49] shows that the q-bounded security we achieve in the standard model is in some sense optimal, at least using black-box reductions.

1.2 Our Tools

Our results are enabled by three tools that we believe to be of more general applicability.

A more precise condition for security of DE. We revisit the definitional equivalences for DE proven by [5] and [7]. At a high level, they showed that the semantic security style definition for DE (called PRIV) introduced in the initial work of [3], which asks that a scheme hides all public-key independent[2] functions of messages drawn from some distribution is in some sense equivalent to an indistinguishability based notion for DE, which asks that it is hard to distinguish ciphertexts of messages drawn from one of two possible distributions. Notice that while PRIV can be meaningfully said to hold for a given message distribution, IND inherently talks of *pairs* of distributions. The works of [5, 7] compensated for this by giving an equivalences in terms of *min-entropy*. That is, they showed that PRIV for all message distributions of min-entropy μ is implied by indistinguishability with respect to all pairs of plaintext distributions of min-entropy slightly less than μ.

We demonstrate a more precise equivalence that, for a *fixed* distribution $\mathbf{M}$, identifies a class of pairs of distributions such that if IND holds on those pairs, then PRIV holds on $\mathbf{M}$. By re-examining the equivalence proof of [5], we show that PRIV on $\mathbf{M}$ is implied by IND on all pairs of "slightly induced" distributions of $\mathbf{M} \mid \mathsf{E}$, where E is an arbitrary event of probability at least 1/4.

This first tool is needed to argue that "robustness" of hc is sufficient for security $\mathsf{EwHCore}$ (essentially, a robust hardcore function is one that remains hardcore on a slightly induced distribution[3]).

Conditional computational entropy. We investigate how conditioning reduces computational entropy of a random variable X. Suppose you have a distribution that has *computational* entropy (such as the pair $f(r), \mathsf{hc}(r)$ for a random r). Suppose you condition that distribution on an event E of probability p. How much computational entropy is left?

To make this question more precise, we should note that computational entropy is parameterized by quality (how distinguishable is X from a variable Z that has true entropy) and quantity (how much true entropy is there in Z).

We prove an intuitively natural result: conditioning on an event of probability p reduces the quality of metric entropy by a factor of p and the quantity of metric entropy by $\log_2 1/p$ (note that this means that the reduction in quantity and quality is the same, because the quantity of entropy is measured on logarithmic scale).

Naturally, the answer becomes so simple only once the correct notion of entropy is in place. Our result holds for $\texttt{Metric}^*$ entropy (defined in [2, 19]). This entropy is convertible (with some loss) to $\texttt{HILL}$ entropy [27, 2] using the techniques of [2], which can then be used with randomness extractors to get pseudorandom bits.

Our result improves previous bounds of Dziembowski and Pietrzak [19, Lemma 3], where the loss in the *quantity* of entropy was related to its original *quality*. The use of metric entropy simplifies the analogous result of Reingold et al. [39, Theorem 1.3] for HILL entropy. (As detailed in Appendix B, other recent work [23, Lemma 3.1], [11, Lemma 16] also addresses the question of conditional computational entropy.)

We use this result to show that randomness extractors can be used to convert a hardcore function into a robust one, through a computational entropy argument for slightly induced distributions. The result is also applicable to leakage-resilient cryptography, as demonstrated by [19]. To make the result useful in more contexts, we also provide an average-case entropy formulation, which can be helpful in situations in which not all leakage is equally informative. (For instance, in case the leakage is equal to the Hamming weight of a uniformly distributed string, sometimes—if the value of the leakage is 0 or the length of the string—the

[2]As shown in [3], the restriction to public-key independent functions is inherent here.

[3]One could alternatively define robustness as one that remains hardcore on inputs of slightly lower entropy; however, in our proofs of robustness we would then need to go through an additional argument that distributions of lower entropy are induced by distributions of higher entropy.

entropy of the string gets reduced to nothing, but most of the time it stays high.) For the information-theoretic case, it is known that leakage of λ bits reduces the average entropy by at most λ (Lemma 2.1 from [16]). We show essentially the same[4] for the computational case: if λ bits of information are leaked, then the amount of computational `Metric`* entropy decreases by at most λ and its quality decreases by at most 2^λ (again, this entropy can be converted to `HILL` entropy and be used in randomness extractors [16, 29]).

(Crooked) Leftover hash lemma for correlated distributions. We show that the leftover hash lemma (LHL) [27, Lemma 4.8], as well its generalized form [16, Lemma 2.4] and the "crooked" LHL [17]) extend in a natural way to "correlated" distributions. That is, suppose we have t random variables (sources) $X_1, \ldots, X_t$, where each X_i individually has high min-entropy but may be fully determined by the outcome of some other X_j (though we assume $X_i \neq X_j$ for all $i \neq j$). We would like to apply a hash function H such that $H(X_1), \ldots, H(X_t)$ is indistinguishable from t independent copies of the uniform distribution on the range of H (also over the choice of the key for H, which is made public). We show that this is the case assuming H is $2t$-wise independent. (The standard LHL is thus $t = 1$; previously, Kiltz et al. [32] showed this for $t = 2$.) Naturally, this requires the output size of H to be about a $1/t$ fraction of its input size, so there is enough entropy to extract.

2 Preliminaries

2.1 Notation and Background

Unless otherwise indicated, an algorithm may be randomized and must run in probabilistic polynomial-time (PPT) in its input size. An adversary is a non-uniform algorithm (or tuple of algorithms). We make the convention that the running-time of an adversary includes its program (i.e., circuit) size and the time to run any overlying experiment. The security parameter is denoted by k, and 1^k denotes the string of k ones. We often suppress dependence of variables on k for readability. A function $f\colon \mathbb{N} \to [0, 1]$ is *negligible* if $f = o(k^{-c})$ for all constants $c \geq 0$.

If A is an algorithm then $x \xleftarrow{\$} A(\ldots)$ denotes that x is assigned the output of running A on the elided inputs and a fresh random tape, while $x \leftarrow A(\ldots; r)$ denotes the same but with the random tape fixed to r. If S is a finite set then $s \xleftarrow{\$} S$ denotes that s is assigned a uniformly random element of S. We let $A(\ldots) \Rightarrow y$ denote the event that A outputs y in the above experiment. We use the abbreviation $x_1, \ldots, x_n \xleftarrow{\$} A(\ldots)$ for $x_1 \xleftarrow{\$} A(\ldots)\,;\,\ldots\,;\,x_n \xleftarrow{\$} A(\ldots)$. If A is deterministic then we drop the dollar sign above the arrow. We denote by $\{0, 1\}^*$ the set of all (binary) strings, and by $\{0, 1\}^n$ the set of strings of length n. By $x_1 \| \cdots \| x_m$ we denote an encoding of strings $x_1, \ldots, x_m$ from which $x_1, \ldots, x_m$ are uniquely recoverable. We denote by $x \oplus y$ the bitwise exclusive-or (xor) of equal-length strings x, y. For two n-bit strings x, y we denote by $\langle x, y\rangle$ the inner-product of x and y when interpreted as vectors over $GF(2)$. Vectors are denoted in boldface, for example $\mathbf{x}$. If $\mathbf{x}$ is a vector then $|\mathbf{x}|$ denotes the number of components of $\mathbf{x}$ and $\mathbf{x}[i]$ denotes its ith component, for $1 \leq i \leq |\mathbf{x}|$. For convenience, we extend algorithmic notation to operate on each vector of inputs component-wise. For example, if A is an algorithm and $\mathbf{x}, \mathbf{y}$ are vectors then $\mathbf{z} \xleftarrow{\$} A(\mathbf{x}, \mathbf{y})$ denotes that $\mathbf{z}[i] \xleftarrow{\$} A(\mathbf{x}[i], \mathbf{y}[i])$ for all $1 \leq i \leq |\mathbf{x}|$.

Let X and Y be random variables. For $t, \epsilon \geq 0$, we say that X and Y are *computationally* (t, ϵ)-*indistinguishable*, denoted $X \approx_{t,\epsilon} Y$, if $|\Pr[D(X) \Rightarrow 1] - \Pr[D(Y) \Rightarrow 1]| \leq \epsilon$ or all distinguishers D running in time at most t.

Statistical notions. Let X be a random variable on a finite set $\mathcal{X}$. We write P_X for the distribution of random variable X and $P_X(x)$ for the probability that X puts on value $x \in \mathcal{X}$, i.e., $P_X(x) = \mathrm{P}[X = x]$. Denote by $|X|$ the size of the support of X, i.e., $|X| = |\{x \text{ s.t. } P_X(x) > 0\}|$. We often identify X with P_X when there is no danger of confusion. By $x \xleftarrow{\$} X$ we denote that x is assigned a value drawn according to

[4] In case of randomized leakage, the information-theoretic result of [16, Lemma 2.2(b)] gives better bounds.

P_X. When this experiment is PPT we say that X is *efficiently sampleable*. We write $X \mid \mathsf{E}$ for the random variable X conditioned on an event E. When X is vector-valued we denote it in boldface, for example $\mathbf{X}$. For a function $f : \mathcal{X} \to \mathbb{R}$, we denote the expectation of f over X by $\mathbb{E}\, f(X) \stackrel{\text{def}}{=} \mathbb{E}_{x \in X}\, f(x) \stackrel{\text{def}}{=} \sum_{x \in \mathcal{X}} P_X(x) f(x)$.

The *max-entropy* of X is $H_0(X) = \log |X|$. The *min-entropy* of X is $\mathrm{H}_\infty(X) = -\log(\max_x P_X(x))$, the *(worst-case) conditional min-entropy* of X given Y is $\mathrm{H}_\infty(X|Y) = -\log(\max_{x,y} P_{X|Y=y}(x))$, and the *average conditional* min-entropy of X given Y [16] is $\tilde{\mathrm{H}}_\infty(X|Y) = -\log(\mathbb{E}_{y \in Y} \max_x P_{X|Y=y}(x))$. Following [3, 5], for vector-valued $\mathbf{X}$ the min-entropy is the minimum *individual* min-entropy of the components, i.e., $\mathrm{H}_\infty(\boldsymbol{X}) = -\log(\max_{\boldsymbol{x},i} P_{\mathbf{X}[i]}(\boldsymbol{x}[i]))$. The *collision probability* of X is $\mathrm{Col}(X) = \sum_x P_X(x)^2$. The *statistical distance* between random variables X and Y with the same domain is $\Delta(X,Y) = \frac{1}{2}\sum_x |P_X(x) - P_Y(x)|$. We write $X \approx_\epsilon Y$ if $\Delta(X,Y) \le \epsilon$, and when ϵ is negligible then we say X and Y are *statistically close*.

t-WISE INDEPENDENT FUNCTIONS. Let $F \colon \mathcal{K} \times D \to R$ be a function. We say that F is *t-wise independent* if for all distinct $x_1, \ldots, x_t \in D$ and all $y_1, \ldots, y_t \in R$

$$\Pr\Big[\, F(K,x_1) = y_1 \;\wedge\; \ldots \;\wedge\; F(K,x_t) = y_t \;:\; K \stackrel{\$}{\leftarrow} \mathcal{K} \,\Big] \;=\; \frac{1}{|R|^t}\,.$$

In other words, $F(K,x_1), \ldots, F(K,x_t)$ are all uniformly and independently random over R. 2-wise independence is also called *pairwise independence*.

ENTROPY AFTER INFORMATION LEAKAGE. Dodis et al. [16, Lemma 2.2] characterized the effect of auxiliary information on average min-entropy:

Lemma 2.1 [16, Lemma 2.2] Let A, B, C be random variables. Then

1. For any $\delta > 0$, the conditional entropy $H_\infty(A|B=b)$ is at least $\tilde{\mathrm{H}}_\infty(A|B) - \log(1/\delta)$ with probability at least $1 - \delta$ over the choice of b.

2. If B has at most 2^λ possible values, then $\tilde{\mathrm{H}}_\infty(A|(B,C)) \ge \tilde{\mathrm{H}}_\infty((A,B)|C) - \lambda \ge \tilde{\mathrm{H}}_\infty(A|C) - \lambda$. In particular, $\tilde{\mathrm{H}}_\infty(A|B) \ge H_\infty((A,B)) - \lambda \ge H_\infty(A) - \lambda$.

EXTRACTORS. Let χ be a finite set. A polynomial-time computable deterministic function $\mathtt{ext} : \chi \times \{0,1\}^d \to \{0,1\}^m \times \{0,1\}^d$ is a strong (k,ϵ)-*extractor* [36] if the last d outputs of bits of $\mathtt{ext}$ are equal to the last d input bits (these bits are called *seed*), and $\delta(\mathtt{ext}(X,U_d), U_m \times U_d) \le \epsilon$ for every distribution X on χ with $H_\infty(X) \ge k$. The number of extracted bits is m, and the entropy loss is $k - m$.

Average-case extractors, defined in [16, Section 2.5], are extractors extended to work with average-case, rather than unconditional, min-entropy. Vahdan [46] shows that any (k,ϵ)-extractor for $k \le \log_2 |\chi| - 1$ is also an $(m, 3\epsilon)$-average-case extractor. However, the additional loss is not always necessary. Indeed, the Leftover Hash Lemma generalizes without any loss to the average-case setting, as shown in DBLP:journals/siamcomp/DodisORS08.

Definition 2.2 Let χ_1, χ_2 be finite sets. An extractor $\mathtt{ext}$ is a (k,ϵ)-*average-case extractor* if for all pairs of random variables X, Y over χ_1, χ_2 such that $\tilde{H}_\infty(X|Y) \ge k$, we have $\delta((\mathtt{ext}(X,U_d),Y), U_m \times U_d \times Y) \le \epsilon$.

PUBLIC-KEY ENCRYPTION. A *(probabilistic) public-key encryption scheme* with plaintext-space PtSp is a triple of algorithms $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$. The key-generation algorithm $\mathcal{K}$ takes input 1^k to return a public key pk and matching secret key sk. The encryption algorithm $\mathcal{E}$ takes pk and a plaintext m to return a ciphertext; this algorithm is randomized, using randomness r. The deterministic decryption algorithm $\mathcal{D}$ takes sk and a ciphertext c to return a plaintext. We require that for all plaintexts $m \in \mathrm{PtSp}$

$$\Pr\Big[\, \mathcal{D}(sk, \mathcal{E}(pk,m)) = m \;:\; (pk,sk) \stackrel{\$}{\leftarrow} \mathcal{K}(1^k) \,\Big] \;=\; 1\,.$$

Next we define security against chosen-plaintext attack [25]. To an encryption scheme $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$, an adversary $A = (A_1, A_2)$, and $k \in \mathbb{N}$ we associate

$$\begin{array}{l}\textbf{Experiment } \mathbf{Exp}^{\text{ind-cpa}}_{\Pi,A}(k): \\ b \xleftarrow{\$} \{0,1\}\,;\ (pk, sk) \xleftarrow{\$} \mathcal{K}(1^k) \\ (m_0, m_1, \mathit{state}) \xleftarrow{\$} A_1(pk) \\ c \xleftarrow{\$} \mathcal{E}(pk, m_b) \\ d \xleftarrow{\$} A_2(pk, c, \mathit{state}) \\ \text{If } d = b \text{ return 1 else return 0}\end{array}$$

where we require A_1's output to satisfy $|m_0| = |m_1|$. Define the *IND-CPA advantage* of A against Π as

$$\mathbf{Adv}^{\text{ind-cpa}}_{\Pi,A}(k) = 2 \cdot \Pr\left[\,\mathbf{Exp}^{\text{ind-cpa}}_{\Pi,A}(k) \Rightarrow 1\,\right] - 1\,.$$

We say that Π is *IND-CPA secure* if $\mathbf{Adv}^{\text{ind-cpa}}_{\Pi,A}(\cdot)$ is negligible for any PPT adversary A.

Lossy trapdoor functions. A *lossy trapdoor function (LTDF) generator* [38] is a pair $\mathsf{LTDF} = (\mathcal{F}, \mathcal{F}')$ of algorithms. Algorithm $\mathcal{F}$ is a usual trapdoor function (TDF) generator, namely on input 1^k outputs (a description of a) function f on $\{0,1\}^n$ for $n = n(k)$ along with (a description of) its inverse f^{-1}, and algorithm $\mathcal{F}'$ outputs a (description of a) function f' on $\{0,1\}^n$. For a distinguisher D, define its *LTDF advantage* against LTDF as

$$\mathbf{Adv}^{\text{ltdf}}_{\mathsf{LTDF},D}(k) = \Pr\left[\,D(f) \Rightarrow 1 \;:\; (f, f^{-1}) \xleftarrow{\$} \mathcal{F}(1^k)\,\right] - \Pr\left[\,D(f') \Rightarrow 1 \;:\; f' \xleftarrow{\$} \mathcal{F}'(1^k)\,\right]\,.$$

We say that LTDF is *secure* if $\mathbf{Adv}^{\text{ltdf}}_{\mathsf{LTDF},D}(\cdot)$ is negligible for any PPT D. We say LTDF has *residual leakage* s if for all f' output by $\mathcal{F}'$ we have $|\mathrm{Image}(f')| \leq 2^s$. The *lossiness* of LTDF is $\ell = n - s$.

One-way and hardcore functions on non-uniform distributions. We extend the usual notion of one-wayness to vectors of inputs drawn from non-uniform and possibly correlated distributions. Let $\mathcal{F}$ be a TDF generator and $\boldsymbol{X}$ be a distribution on input vectors. To $\mathcal{F}, \boldsymbol{X}$, an inverter I, and $k \in \mathbb{N}$ we associate

$$\begin{array}{l}\textbf{Experiment } \mathbf{Exp}^{\text{owf}}_{\mathcal{F},\boldsymbol{X},I}(k): \\ (f, f^{-1}) \xleftarrow{\$} \mathcal{F} \\ \mathbf{x} \xleftarrow{\$} \boldsymbol{X} \\ x' \xleftarrow{\$} I(f, f(\mathbf{x})) \\ \text{If } \exists i \text{ such that } \mathbf{x}[i] = x' \text{ return 1 else return 0}\end{array}$$

Define the *OWF advantage* of I against $F, \boldsymbol{X}$ as

$$\mathbf{Adv}^{\text{owf}}_{\mathcal{F},\boldsymbol{X},I}(k) = \Pr\left[\,\mathbf{Exp}^{\text{owf}}_{\mathcal{F},\boldsymbol{X},I}(k) \Rightarrow 1\,\right]\,.$$

We say that $\mathcal{F}$ is *one-way* on a class of distributions on input vectors $\mathbb{X}$ if for every $\boldsymbol{X} \in \mathbb{X}$ and every PPT inverter I, $\mathbf{Adv}^{\text{owf}}_{\mathcal{F},\boldsymbol{X},I}(\cdot)$ is negligible. We extend hardcore functions (HCFs) in a similar way. Namely, to a trapdoor function generator $\mathcal{F}$, function $\mathsf{hc}\colon \{0,1\}^k \to \{0,1\}^n$, distribution on input vectors $\boldsymbol{X}$, a distinguisher D, and $k \in \mathbb{N}$ we associate

$$\begin{array}{l}\textbf{Experiment } \mathbf{Exp}^{\text{hcf}}_{\mathcal{F},\mathsf{hc},\boldsymbol{X},D}(k): \\ b \xleftarrow{\$} \{0,1\}\,;\ (f, f^{-1}) \xleftarrow{\$} \mathcal{F} \\ \mathbf{x} \xleftarrow{\$} \boldsymbol{X} \\ \mathbf{h}_0 \leftarrow \mathsf{hc}(f, \mathbf{x})\,;\ \mathbf{h}_1 \xleftarrow{\$} (\{0,1\}^n)^{\times|\mathbf{x}|} \\ d \xleftarrow{\$} D(f, f(\mathbf{x}), \mathbf{h}_b) \\ \text{If } d = b \text{ return 1 else return 0}\end{array}$$

Define the *HCF advantage* of D against $F, \mathsf{hc}, \boldsymbol{X}$ as

$$\mathbf{Adv}^{\text{hcf}}_{\mathcal{F},\mathsf{hc},\boldsymbol{X},D}(k) = 2 \cdot \Pr\left[\,\mathbf{Exp}^{\text{hcf}}_{\mathcal{F},\mathsf{hc},\boldsymbol{X},D}(k) \Rightarrow 1\,\right] - 1\,.$$

We say that hc is *hardcore* for $\mathcal{F}$ on a class of distributions on input vectors $\mathbb{X}$ if for every $\boldsymbol{X} \in \mathbb{X}$ and every PPT distinguisher D, $\mathbf{Adv}^{\mathrm{hcf}}_{\mathcal{F},\mathsf{hc},\boldsymbol{X},D}(\cdot)$ is negligible.

Note that we depart somewhat from standard treatments in that we allow a HCF to also depend on the description of the trapdoor function (via the argument f). This allows us to simplify our exposition.

Augmented trapdoor functions. It is useful to introduce the notion of an "augmented" version of a TDF, which augments the description of the latter with keying material for a HCF. More formally, let $\mathcal{F}$ be a trapdoor function generator and let H be a keyed function with keyspace $\mathcal{K}$. Define the *H-augmented version of $\mathcal{F}$*, denoted $\mathcal{F}[H]$, that on input 1^k returns $(f, K), (f^{-1}, K)$ where $(f, f^{-1}) \xleftarrow{\$} \mathcal{F}(1^k)$ and $K \xleftarrow{\$} \mathcal{K}$; evaluation is defined for $x \in \{0,1\}^k$ as $f(x)$ (i.e., evaluation just ignores K) and inversion is defined analogously.

Goldreich-Levin hardcore function. For $i \in \mathbb{N}$ define the *length-i Goldreich-Levin (GL) function* [24] $\mathcal{GL}^i \colon \{0,1\}^{i \times k} \times \{0,1\}^k \to \{0,1\}^i$ as $GL^i(M, x) = Mx$, where Mx is the matrix-vector product of M and x over $GF(2)$. We recall the following.

Theorem 2.3 (**Goldreich-Levin Theorem** [24]) Let $\mathcal{F}[\mathcal{GL}^i]$ be as defined above and let X be a distribution on inputs to $\mathcal{F}$. Let D be a distinguisher against $\mathcal{GL}^i$. Then there is a inverter I such that for all $k \in \mathbb{N}$

$$\mathbf{Adv}^{\mathrm{hcf}}_{\mathcal{F}[\mathcal{GL}^i],\mathcal{GL}^i,X,D}(k) \quad \leq \quad 2^{i+3} \cdot \mathbf{Adv}^{\mathrm{owf}}_{\mathcal{F},X,I}(k)\ . \tag{1}$$

Furthermore, the running-time of I is the time for $O(\epsilon^{-4}k^3)$ executions of D where $\epsilon = \mathbf{Adv}^{\mathrm{hcf}}_{\mathcal{F}[\mathcal{GL}^i],\mathcal{GL}^i,X,D}(k)$.

2.2 Computational Entropy

For computational entropy we define several classes of distinguishers. Let $\mathcal{D}_s^{det,\{0,1\}}$ be the set of all deterministic circuits of size s with binary output $\{0,1\}$, let $\mathcal{D}_s^{det,[0,1]}$ be the set of all deterministic circuits of size s with output in $[0,1]$, and let $\mathcal{D}_s^{rand,\{0,1\}}, \mathcal{D}_s^{rand,[0,1]}$ as the set of probabilistic circuits with output ranges $\{0,1\}$ and $[0,1]$ respectively. (We talk of circuit size rather than running-time in the context of computational entropy for consistency with the literature.) Given a circuit D, define the computational distance δ^D between X and Z as $\delta^D(X,Z) = |\mathbb{E}[D(X)] - \mathbb{E}[D(Z)]|$. While min-entropy is measured only by amount, computational min-entropy has two additional parameters: distinguisher size s and quality ϵ. Larger s and smaller ϵ mean "better" entropy.

Definition 2.4 ([27]) A distribution X has *HILL entropy* at least k, denoted $H^{\mathtt{HILL}}_{\epsilon,s}(X) \geq k$ if there exists a distribution Z where $H_\infty(Z) \geq k$, such that $\forall D \in \mathcal{D}_s^{rand,\{0,1\}}, \delta^D(X,Z) \leq \epsilon$.

An alternative notion called *Metric entropy* is often used for proofs and is obtained by switching in the order of quantifiers. Thus, a different Z can be used for each distinguisher:

Definition 2.5 ([2]) A distribution X has *Metric entropy* at least k, denoted $H^{\mathtt{Metric}}_{\epsilon,s}(X) \geq k$ if $\forall D \in \mathcal{D}_s^{rand,\{0,1\}}$ there exists a distribution Z_D with $H_\infty(Z_D) \geq k$ and $\delta^D(X, Z_D) \leq \epsilon$.

For HILL entropy, drawing D from $\mathcal{D}_s^{det,\{0,1\}}, \mathcal{D}_s^{det,[0,1]}, \mathcal{D}_s^{rand,\{0,1\}}, \mathcal{D}_s^{rand,[0,1]}$ is essentially equivalent, as shown in [19, 22]). For metric entropy, however, the choice among these four classes can make a difference. In particular, if we change the class of D in Definition 2.5 to $\mathcal{D}_s^{det,[0,1]}$, we get so-called "metric-star" entropy, denoted $H^{\mathtt{Metric}^*}_{\epsilon,s}$ (this notion was used in [19, 22]).

Equivalence (with a loss in quality) between $\mathtt{Metric}^*$ and $\mathtt{HILL}$ entropy[5] was shown by Barak, Shaltiel, and Wigderson [2, Theorem 5.2]:

[5] $\mathtt{Metric}^*$ entropy is weaker than $\mathtt{HILL}$ entropy in two ways, the distinguisher is deterministic and the distribution Z can depend on the distinguisher.

Theorem 2.6 ([2]) Let X be a discrete distribution over a finite set χ. For every $\epsilon, \epsilon_{\mathtt{HILL}} > 0, \epsilon' \geq \epsilon + \epsilon_{\mathtt{HILL}}$, k, and s, if $H^{\mathtt{Metric}^*}_{\epsilon,s}(X) \geq k$ then $H^{\mathtt{HILL}}_{\epsilon', s_{\mathtt{HILL}}}(X) \geq k$ where $s_{\mathtt{HILL}} = \Omega(\epsilon^2_{\mathtt{HILL}} s / \log|\chi|)$.

The free parameter in the above theorem, $\epsilon_{\mathtt{HILL}}$, provides a tradeoff between distinguisher size and advantage. For simplicity, we can set $\epsilon_{\mathtt{HILL}} = \sqrt[3]{\frac{\log|\chi|}{s}}$ yielding $s_{\mathtt{HILL}} = \Omega(\sqrt[3]{\frac{s}{\log|\chi|}}$ and $\epsilon' = \epsilon + \sqrt[3]{\frac{\log|\chi|}{s}}$. For typical parameters (specifically, when $\epsilon \leq (\log|\chi|/s)^{1/3}$), this setting balances the resulting ϵ' and $s_{\mathtt{HILL}}$, i.e., gives us $\epsilon' = O(1/s_{\mathtt{HILL}})$.

We show the proof of a slightly stronger version of this theorem in Theorem C.1.

Extractors can be applied to distributions with computational entropy to obtain pseudorandom, rather than random, outputs: that is, outputs that are computationally indistinguishable from, rather than statistically close to, uniformly random strings. This fact is well-known for HILL entropy. However, we have not seen it proven for Metric entropy and, although the proof is quite straightforward, we provide it here for completeness. (Since HILL entropy implies Metric entropy, this proof also works for HILL entropy.)

Theorem 2.7 Let $\mathtt{ext} : \chi \times \{0,1\}^d \to \{0,1\}^m \times \{0,1\}^d$ be a (k, ϵ_{ext})-extractor, computable by circuits of size $s_{\mathtt{ext}}$. Let X be a distribution over χ with $H^{metric}_{\epsilon_{metric},s}(X) \geq k$. Then $\forall D \in \mathcal{D}^{rand,\{0,1\}}_{s'}$, where $s' \approx s_{metric} - s_{\mathtt{ext}}$,

$$\delta^D(\mathtt{ext}(X, U_d), U_m \times U_d) \leq \epsilon_{\mathtt{ext}} + \epsilon_{metric}\,.$$

Proof: We proceed by contradiction. Suppose not, that is, $\exists D \in \mathcal{D}^{rand,\{0,1\}}_{s'}$ such that

$$\delta^D(\mathtt{ext}(X, U_d), U_m \times U_d)) > \epsilon_{\mathtt{ext}} + \epsilon_{metric}.$$

We use D to construct a distinguisher D' to distinguish X from all distributions Z where $H_\infty(Z) \geq k$, violating the metric-entropy of X. We define D' as follows: upon receiving input $\alpha \in \chi$, D' samples $\mathit{seed} \leftarrow U_d$, runs $\beta \leftarrow \mathtt{ext}(\alpha, \mathit{seed})$ and then runs $D(\beta, \mathit{seed})$ on the result. Note that $D' \in \mathcal{D}^{rand,\{0,1\}}_{s}$ where $s \approx s' + s_{\mathtt{ext}} = s_{metric}$. Thus we have the following $\forall Z$, where $H_\infty(Z) \geq k$:

$$\begin{aligned}\delta^{D'}(X, Z) &= \delta^D(\mathtt{ext}(X, U_d), \mathtt{ext}(Z, U_d)) \\ &\geq \delta^D(\mathtt{ext}(X, U_d), U_m \times U_d) - \delta^D((\mathtt{ext}(Z, U_d), U_m \times U_d) \\ &> \epsilon_{\mathtt{ext}} + \epsilon_{metric} - \epsilon_{\mathtt{ext}} = \epsilon_{metric}\end{aligned}$$

Thus D' is able to distinguish X from all Z with sufficient min-entropy. This is a contradiction. ∎

Unfortunately, the theorem does not extend to Metric* entropy, because the distinguisher D' we construct in this proof is randomized. The only way to extract from Metric* entropy that we know of is to convert $\mathtt{Metric}^*$ entropy to $\mathtt{HILL}^*$ entropy using Theorem 2.6 (which incurs some loss) and then use Theorem 2.7 (see Figure 1).

Conditional entropy has been extended to the computational case by Hsiao, Lu, Reyzin [29].

Definition 2.8 ([29]) Let (X, Y) be a pair of random variables. X has *conditional HILL entropy* at least k conditioned on Y, denoted $H^{\mathtt{HILL}}_{\epsilon,s}(X|Y) \geq k$ if there exists a collection of distributions Z_y for each $y \in Y$, giving rise to a joint distribution (Z, Y), such that $\tilde{H}_\infty(Z|Y) \geq k$ and $\forall D \in \mathcal{D}^{rand,\{0,1\}}_{s}, \delta^D((X,Y),(Z,Y)) \leq \epsilon$.

Again, we can switch the quantifiers of Z and D to obtain the definition of conditional metric entropy.

Definition 2.9 Let (X, Y) be a pair of random variables. X has *conditional Metric entropy* at least k conditioned on Y, denoted by $H^{\mathtt{Metric}}_{\epsilon,s}(X|Y) \geq k$, if $\forall D \in \mathcal{D}^{rand,\{0,1\}}_{s}$ there exists a collection of distributions Z_y for each $y \in Y$, giving rise to a joint distribution (Z, Y), such that $\tilde{H}_\infty(Z|Y) \geq k$ and $\delta^D((X,Y),(Z,Y)) \leq \epsilon$.

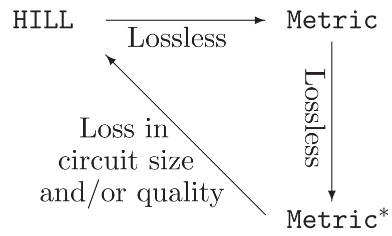

Figure 1: Known state of equivalence for HILL and Metric Entropy. It is known how to extract from HILL and Metric entropy but not Metric* entropy.

Conditional $\texttt{Metric}^*$ can be defined similarly, replacing $\mathcal{D}^{rand,\{0,1\}}$ with $\mathcal{D}^{det,[0,1]}$.

Theorem 2.6 can be extended to the conditional case with the same techniques (see [11, 22] a proof):

Theorem 2.10 Let X be a discrete discrete distribution over a finite set χ_1 and let Y be a discrete random variable over χ_2. For every $\epsilon, \epsilon_{\texttt{HILL}} > 0, \epsilon' \geq \epsilon + \epsilon_{\texttt{HILL}}, k$ and s, if $H^{\texttt{Metric}^*}_{\epsilon,s}(X|Y) \geq k$ then $H^{\texttt{HILL}}_{\epsilon',s_{\texttt{HILL}}}(X|Y) \geq k$ where $s' = \Omega(\epsilon^2_{\texttt{HILL}} s/\log|\chi_1||\chi_2|)$.

Again, it is reasonable to set $\epsilon_{\texttt{HILL}} = \sqrt[3]{\frac{\log|\chi_1||\chi_2|}{s}}$ and get $s_{\texttt{HILL}} = \Omega(\sqrt[3]{\frac{s}{\log|\chi_1||\chi_2|}}$ and $\epsilon' = \epsilon + \sqrt[3]{\frac{\log|\chi_1||\chi_2|}{s}}$.

Similar to extractors in the case of unconditional entropy, average-case extractors can be used on distributions that have conditional Metric (and therefore also on distributions that have HILL) entropy to produce psuedorandom, rather than random outputs. The proof is similar to [29, Lemma 5]. However, it is not known how to extract directly from conditional $\texttt{Metric}^*$ entropy; we first have to convert it to HILL using Theorem 2.10.

2.3 Deterministic Encryption

We say that an encryption scheme $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ is *deterministic* if $\mathcal{E}$ is deterministic.

Semantic security of DE. We recall the semantic-security style PRIV notion for DE from [3][6]. To encryption scheme $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$, an adversary $A = (A_0, A_1, A_2)$, and $k \in \mathbb{N}$ we associate

Experiment $\mathbf{Exp}^{\text{priv-1}}_{\Pi,A}(k)$:	**Experiment** $\mathbf{Exp}^{\text{priv-0}}_{\Pi,A}(k)$:
$(pk, sk) \xleftarrow{\$} \mathcal{K}(1^k)$	$(pk, sk) \xleftarrow{\$} \mathcal{K}(1^k)$
$state \xleftarrow{\$} A_0(1^k)$	$state \xleftarrow{\$} A_0(1^k)$
$(\mathbf{x}_1, t_1) \xleftarrow{\$} A_1(state)$	$(\mathbf{x}_1, t_1), (\mathbf{x}_0, t_0) \xleftarrow{\$} A_1(state)$
$\mathbf{c} \xleftarrow{\$} \mathcal{E}(pk, \mathbf{x}_1)$	$\mathbf{c} \xleftarrow{\$} \mathcal{E}(pk, \mathbf{x}_0)$
$g \xleftarrow{\$} A_2(pk, \mathbf{c}, state)$	$g \xleftarrow{\$} A_2(pk, \mathbf{c}, state)$
If $g = t_1$ Return 1 Else Return 0	If $g = t_1$ Return 1 Else Return 0

We require that there are functions $v = v(k), \ell = \ell(k)$ such that (1) $|\mathbf{x}| = v$, (2) $|\mathbf{x}[i]| = \ell$ for all $1 \leq i \leq v$, and (3) the $\mathbf{x}[i]$ are all distinct with probability 1 over $(\mathbf{x}, t) \xleftarrow{\$} A_1(state)$ for any *state* output by A_0[7]. In particular we say A *outputs vectors of size* v for v as above. Define the *PRIV advantage* of A against Π as

$$\mathbf{Adv}^{\text{priv}}_{\Pi,A}(k) = \Pr\left[\mathbf{Exp}^{\text{priv-1}}_{\Pi,A}(k) \Rightarrow 1\right] - \Pr\left[\mathbf{Exp}^{\text{priv-0}}_{\Pi,A}(k) \Rightarrow 1\right].$$

[6] More specifically, it is a "comparison-based" semantic-security style notion; this was shown equivalent to a "simulation-based" formulation in [5].

[7] In this work we only consider the definition relative to deterministic Π, so requirement (3) is without loss of generality.

Let $\mathbb{M}$ be a class of distributions on message vectors. Define $\mathbb{A}_{\mathbb{M}}$ to be the class of adversaries $\{A = (A_0, A_1, A_2)\}$ such that for each $A \in \mathbb{A}_{\mathbb{M}}$ there is a $\boldsymbol{M} \in \mathbb{M}$ for which $\mathbf{x}$ has distribution $\boldsymbol{M}$ over $(\mathbf{x}, t) \stackrel{\$}{\leftarrow} A_1(\mathit{state})$ for any *state* output by A_0. We say that Π is *PRIV secure for* $\mathbb{M}$ if $\mathbf{Adv}^{\mathrm{priv}}_{\Pi,A}(\cdot)$ is negligible for any PPT $A \in \mathbb{A}_{\mathbb{M}}$. Note that (allowing non-uniform adversaries as usual) we can without loss of generality consider only those A with "empty" A_0, since A_1 can always be hardwired with the "best" state. However, following [5] we explicitly allow state because it greatly facilitates some proofs.

INDISTINGUISHABILITY OF DE. Next we recall the indistinguishability-based formulation of security for DE given (independently) by [5, 7] (and which is adapted from [18]). To an encryption scheme $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$, an adversary $D = (D_1, D_2)$, and $k \in \mathbb{N}$ we associate

Experiment $\mathbf{Exp}^{\mathrm{ind}}_{\Pi,A}(k)$:
$(pk, sk) \stackrel{\$}{\leftarrow} \mathcal{K}(1^k)$
$b \stackrel{\$}{\leftarrow} \{0,1\}$; $(\mathbf{x}, t) \stackrel{\$}{\leftarrow} D_1(b)$
$\mathbf{c} \stackrel{\$}{\leftarrow} \mathcal{E}(pk, \mathbf{x})$
$d \stackrel{\$}{\leftarrow} D_2(pk, \mathbf{c})$
If $b = d$ return 1 else return 0

We make the analogous requirements on D_1 as on A_1 in the PRIV definition. Define the *IND advantage* of D against Π as $\mathbf{Adv}^{\mathrm{ind}}_{\Pi,D}(k) = 2 \cdot \Pr\left[\,\mathbf{Exp}^{\mathrm{ind}}_{\Pi,D}(k) \Rightarrow 1\,\right] - 1$. Let $\mathbb{M}^*$ be a class of *pairs* of distributions on message vectors. Define $\mathbb{D}_{\mathbb{M}^*}$ to be the class of adversaries $\{D = (D_1, D_2)\}$ such that for each $D \in \mathbb{D}_{\mathbb{M}^*}$, there is a pair of distributions $(\boldsymbol{M}_0, \boldsymbol{M}_1) \in \mathbb{M}^*$ such that for each $b \in \{0,1\}$ the distribution of $\mathbf{x} \stackrel{\$}{\leftarrow} D_1(b)$ is $\boldsymbol{M}_b$. We say that Π is *IND secure for* $\mathbb{M}^*$ if $\mathbf{Adv}^{\mathrm{ind}}_{\Pi,D}(\cdot)$ is negligible for any PPT $D \in \mathbb{D}_{\mathbb{M}^*}$.

3 Our Tools

3.1 A Precise Definitional Equivalence for DE

While the PRIV definition is meaningful with respect a single message distribution $\boldsymbol{M}$, the IND definition must inherently talk of *pairs* of different message distributions. Thus, in proving an equivalence between the two notions, the best we can hope to show is that PRIV security for a message distribution $\boldsymbol{M}$ is equivalent to IND security for some *class of pairs* of message distributions (depending on $\boldsymbol{M}$). However, prior works [5, 7] did not provide such a statement. Instead, they showed that PRIV security on *all* distributions of a given entropy μ is equivalent to IND security on all pairs of distributions of slightly less entropy.

INDUCED DISTRIBUTIONS. To state our result we first give some definitions relating to a notion of "induced distributions." Let X, X' be distributions (or random variables) on the same domain. For $\alpha \in \mathbb{N}$, we say that X' is an *α-induced distribution of* X if X' is a conditional distribution $X' = X \mid \mathsf{E}$ for an event E such that $\Pr[\,\mathsf{E}\,] \geq 2^{-\alpha}$. We call E the *corresponding event* to X'. We require that the pair (X, E) is efficiently samplable (where we view event E as a binary random variable).

Define $X[\alpha]$ to be the class of all α-induced distributions of X. Furthermore, let X_0, X_1 be two α-induced distributions of X with corresponding events $\mathsf{E}_0, \mathsf{E}_1$ respectively. Define $X^*[\alpha] = \{(X_0, X_1)\}$ to be the class of all pairs (X_0, X_1) for which there is a pair (X_0', X_1') of α-induced distributions of X such that X_0 (resp. X_1) is statistically close to X_0' (resp. X_1').[8]

THE EQUIVALENCE. We are now ready to state our result. The following theorem captures the "useful" direction that IND implies PRIV:

[8] We need to allow a negligible statistical distance for technical reasons; cf. Proposition A.3. (This relaxation is reminiscent of the notion of *smooth* entropy [40] by Renner and Wolf.) Since we will be interested in indistinguishability of functions of these distributions this will not make any appreciable difference, and hence we mostly ignore this issue in the remainder of the paper.

Theorem 3.1 Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be an encryption scheme. For any distribution $\boldsymbol{M}$ on message vectors, PRIV security of Π with respect to $\boldsymbol{M}$ is implied by IND security of Π with respect to $\boldsymbol{M}^*[2]$. In particular, let $A \in \mathbb{A}_{\boldsymbol{M}}$ be a PRIV adversary against Π. Then there is a IND adversary $D \in \mathbb{D}_{\boldsymbol{M}^*[2]}$ such that for all $k \in \mathbb{N}$

$$\mathbf{Adv}^{\mathrm{priv}}_{\Pi,A}(k) \quad \leq \quad 162 \cdot \mathbf{Adv}^{\mathrm{ind}}_{\Pi,D}(k) + \left(\frac{3}{4}\right)^k .$$

Furthermore, the running-time of D is the time for at most that for k executions of A (but 4 in expectation).

The theorem essentially follows from the techniques of [5]. Thus, our contribution here is not in providing any new technical tools used in proving this result but rather in extracting it from the techniques of [5]. For completeness, we give the entire proof (incorporating simplifications due to [13] that lead to better concrete security) in Appendix A.

To establish a definitional *equivalence*; that is, also show that PRIV implies IND, we need to further restrict the latter to pairs (that are statistically close to pairs) of *complementary* 2-induced distributions of $\boldsymbol{M}$ (which we did not do above for conceptual simplicity), where we call X_0, X_1 *complementary* if $\mathsf{E}_1 = \overline{\mathsf{E}_0}$. (The idea for the proof is to have the constructed PRIV adversary sample according to $\boldsymbol{M}$ and let the partial information be whether the corresponding event for the induced complementary distributions of the given IND adversary occurred or not.) We stress that this further restriction is not needed for the "useful" implication above and for our security proofs.

Why is the more precise equivalence better? This equivalence result is more precise than prior work, because it requires a weaker condition in order to show PRIV holds on a *specific* message distribution. Moreover, conceptually, viewing a lower-entropy distribution as a conditional (induced) version of a higher-entropy distribution is helpful in simplifying proofs. In particular, it allows us to use results about entropy of conditional distributions, which we explain next. Looking ahead, it also simplifies proofs for schemes based on one-wayness, because it is easy to argue that one-wayness is preserved on slightly induced distributions (the alternative would require us to go through an argument that distributions of lower entropy are induced by distributions of higher entropy).

3.2 Measuring Computational Entropy of Induced Distributions

We study how conditioning a distribution reduces its computational entropy. This result is used later in the work to show that randomness extractors can convert a hardcore function into a robust one; it also applicable to leakage-resilient cryptography. Some basic definitions and results concerning computational entropy are reviewed in Section 2.2; in particular, we will use $\mathtt{Metric}^*$ computational entropy defined there.

It is easy to see that conditioning on an event E with probability P_{E} reduces (information-theoretic) min-entropy by at most $\log P_{\mathsf{E}}$; indeed, this is shown Lemma 5.5. (Note that this statement is quite intuitive: the more surprising a leakage value is, the more it decreases the entropy.) In the following Lemma, we show that the same holds for the computational notion of $\mathtt{Metric}^*$ entropy if one considers reduction in both quantity and quality.

We actually need a slightly stronger statement in order to use Lemma 3.2 later, in the proof of Lemma 5.1: namely, we will need to make sure that the support of the indistinguishable distribution with true randomness does not increase after conditioning. We call this additional property *support preservation.*

Lemma 3.2 Let X, Y be discrete random variables. Then

$$H^{\mathtt{Metric}^*}_{\epsilon/P_Y(y),s'}(X|Y=y) \geq H^{\mathtt{Metric}^*}_{\epsilon,s}(X) - \log 1/P_Y(y)$$

where $s' \approx s$. Furthermore, the reduction is *support preserving*[9].

The use of $\texttt{Metric}^*$ entropy and an improved proof allow for a simpler and tighter formulation than results of [19, Lemma 3] and [39, Theorem 1.3] (see Appendix B for a comparison).

The proof is similar to [39]. The high level outline of the proof is: Let $\nu = H^{\texttt{Metric}^*}_{\epsilon,s}(X)$.

1. Suppose D distinguishes $X|Y = y$ from any distribution Z of min-entropy $\nu - \Delta$ with advantage ϵ'. Show that either for all Z with min-entropy $\nu - \Delta$, $\mathbb{E}[D(Z)]$ is lower than $\mathbb{E}[D(X|Y = y)]$ by at least ϵ', or for all such Z, $\mathbb{E}[D(Z)]$ is higher than $\mathbb{E}[D(X|Y = y)]$ by at least ϵ'. Assume the former without loss of generality. This initial step allows us to remove absolute values and to find a high-entropy distribution Z^+ on which $\mathbb{E}[D(Z^+)]$ is the highest.

2. Show that there exists a distinguisher D' that also has advantage ϵ' but, unlike D, outputs only 0 or 1. This is done by finding a cutoff α: if D's output is above α, it D' will output 1, and otherwise it will output 0.

3. Show that for every z outside of Z^+, D' outputs 0, and that Z^+ is essentially flat. Use these two facts to show an upper bound on $\mathbb{E}[D'(W)]$ for any W of min-entropy ν.

4. Show a lower bound on $\mathbb{E}[D'(X)]$ based the performance of D' on $X|Y = y$.

We now proceed with the full proof:

Proof: Let χ be the outcome space of X. For notational convenience, for random variables A, B we will say that $A \subseteq B$ if the support of A is a subset of the support of B. Likewise, we will say $a \in A$ to say that a is in the support of A. Fix a set $\zeta \subseteq \chi$, ζ will be used to represent the support of random variables with min-entropy. For the reduction to be support preserving, all distributions with min-entropy should have support no more than ζ.

Assume $H^{\texttt{Metric}^*}_{\epsilon,s}(X) \geq \nu$. Fix $y \in Y$; let $\epsilon' = \epsilon/P_Y(y)$ and $s' \approx s$ be some value to be precisely determined by the end of the proof. We assume for contradiction that

$$H^{\texttt{Metric}^*}_{\epsilon',s'}(X|Y = y) \geq \nu - \log 1/P_Y(y)$$

does not hold. By definition of metric entropy there exists a distinguisher $D_y \in \mathcal{D}^{det,[0,1]}_{s'}$ such that $\forall Z \subseteq \zeta$ with $H_\infty(Z) \geq \nu - \log 1/P_Y(y)$ we have

$$|\mathbb{E}[D_y(X)|Y = y] - \mathbb{E}[D_y(Z)]| > \epsilon'. \tag{2}$$

To contradict the $\texttt{Metric}^*$ entropy of X, it suffices to show there exists a distinguisher $D'_y \in \mathcal{D}^{rand,\{0,1\}}_{s}$ such that $\forall W \subseteq \zeta$ with $\mathrm{H}_\infty(W) \geq \nu$,

$$\mathbb{E}[D'_y(X)] - \mathbb{E}[D'_y(W)] = \epsilon.$$

Let $Z^- \subseteq \zeta$ and $Z^+ \subseteq \zeta$ be distributions of min-entropy $\nu - \log 1/P_Y(y)$ that are subsets of ζ minimizing $\mathbb{E}[D_y(Z^-)]$ and maximizing $\mathbb{E}[D_y(Z^+)]$ respectively. Let $\beta^- \stackrel{def}{=} \mathbb{E}[D_y(Z^-)], \beta^+ \stackrel{def}{=} \mathbb{E}[D_y(Z^+)]$ and $\beta \stackrel{def}{=} \mathbb{E}[D_y(X)|Y = y]$.

Claim 3.3 Either $\beta^- \leq \beta^+ + \epsilon' < \beta$ or $\beta < \beta^- - \epsilon' \leq \beta^+$.

[9] "Support preserving" here means the following. The definition of $\texttt{Metric}^*$ entropy of X calls for an indistinguishable from X distribution Z_D with true entropy for every distinguisher $D \in \mathcal{D}^{det,[0,1]}_{s}$. Let ζ_X be the union of supports of all Z_D. Similarly, define $\zeta_{X|Y=y}$ to be the union of supports for Z_D that cannot be distinguished by $D \in \mathcal{D}^{det,[0,1]}_{s'}$ from $X|Y = y$. Support-preserving means $\zeta_{X|Y=y} \subseteq \zeta_X$.

From (2) and the fact that Z^+, Z^- have min-entropy at least $\nu - \log 1/P_Y(y)$ it suffices to show that either $\beta^- \leq \beta^+ < \beta$ or $\beta < \beta^- \leq \beta^+$. Suppose it does not hold. Then $\beta^- < \beta < \beta^+$. Then we can define a distribution $Z \subseteq \zeta$ as a convex combination of Z^+, Z^- with $\mathbb{E}[D_y(Z)] = \beta$. Furthermore a distribution formed by taking a convex combination of distributions with min-entropy $\nu - \log 1/P_Y(y)$ has min-entropy $\nu - \log 1/P_Y(y)$ (this is easily seen by considering the maximum-probability event). Furthermore, a distribution that is a convex combination of distributions whose support is at most ζ has support at most ζ. This is a contradiction of (2).

For the rest of the proof we will assume that the first case $\beta^- < \beta^+ + \epsilon' < \beta$ holds.

Claim 3.4 There exists a point $\rho \in [0, 1]$ such that

$$\Pr[D_y(X|Y = y) > \rho] - \Pr[D_y(Z^+) > \rho] > \epsilon' . \tag{3}$$

Proof: One has that

$$\begin{aligned}
\epsilon' &< \mathbb{E}[D_y(X|Y = y)] - \mathbb{E}[D_y(Z^+)] \\
&= \int_0^1 \Pr_{x \in X|Y=y}[D_y(x) > \rho] d\rho - \int_0^1 \Pr_{z \in Z}[D_y(z) > \rho] d\rho \\
&= \int_0^1 \left(\Pr_{x \in X|Y=y}[D_y(x) > \rho] - \Pr_{z \in Z}[D_y(z) > \rho] \right) d\rho
\end{aligned}$$

Suppose no $\rho \in [0, 1]$ satisfies (3). This means $\forall \rho \in [0, 1], \Pr[D_y(X) > \rho | Y = y] - \Pr[D_y(Z^+) > \rho] \leq \epsilon'$ and thus

$$\int_0^1 \left(\Pr_{x \in X|Y=y}[D_y(x) > \rho] - \Pr_{z \in Z}[D_y(z) > \rho] \right) d\rho \leq \epsilon'.$$

This is a contradiction. ∎

Since D_y is a fixed size circuit, it outputs values of some bounded precision. Call the ordered set of possible output values $\Pi = \{p_1, ..., p_j\}$. Then, let $\alpha = \max\{p_i | p_i \leq \rho\}$. Thus, α is a fixed precision number where $\forall p_i \in \Pi, p_i > \alpha$ implies $p_i > \rho$. This means that

$$\Pr[D_y(X|Y = y) > \alpha] - \Pr[D_y(Z^+) > \alpha] > \epsilon' . \tag{4}$$

We define a distinguisher D'_y as follows:

$$D'_y(z) = \begin{cases} 0 & D_y(z) \leq \alpha \\ 1 & D_y(z) > \alpha. \end{cases} \tag{5}$$

The only difference in the size of D'_y and D_y is the addition of a comparison to α, which takes up size proportional to the number of output bits of D_y. Thus s, the size of D'_y, is approximately the same as s', the size of D_y. We define the quantities

$$\begin{aligned}
\beta_\alpha &\stackrel{def}{=} \Pr[D_y(X|Y = y) > \alpha] = \mathbb{E}[D'_y(X|Y = y)] \\
\beta_\alpha^+ &\stackrel{def}{=} \Pr[D_y(Z^+) > \alpha] = \mathbb{E}[D'_y(Z^+)].
\end{aligned}$$

Let $\gamma = \min_{z \in Z^+} D_y(z)$. Since $\beta_\alpha - \beta_\alpha^+ \geq \epsilon'$, we know that $\beta_\alpha^+ < 1$. This implies that $\gamma < \alpha$.

Claim 3.5 For all $z \in \zeta$ if $\Pr[Z^+ = z] \neq 2^{-\nu + \log 1/P_Y(y)}$, then $D_y(z) \leq \gamma < \alpha$ and therefore $D'_y(z) = 0$.

Proof: Recall that because $\mathrm{H}_\infty(Z^+) = \nu - \log 1/P_Y(y)$, for all $z \in \zeta$ we have $\Pr[Z^+ = z] \leq 2^{-\nu+\log 1/P_Y(y)}$. Thus, suppose, for contradiction that there exists a $z \in \zeta$ such that $\Pr[Z^+ = z] < 2^{-\nu+\log 1/P_Y(y)}$ and $D_y(z) > \gamma$. Choose a w with $\Pr[Z^+ = w] > 0$ such that $D_y(w) = \gamma$. Create a distribution Z' by starting with Z^+, increasing the probability of z and decreasing the probability of w by the same amount, while keeping the min-entropy guarantee. Then we have $\mathbb{E}[D_y(Z')] > \mathbb{E}[D_y(Z^+)]$ which is a contradiction to how Z^+ was chosen. ▮

Claim 3.5 implies that

$$\beta_\alpha^+ = \sum_{z \in \chi} \Pr[Z^+ = z] D'_y(z) = \sum_{z \in Z^+} 2^{-\nu+\log 1/P_Y(y)} D'_y(z) = \frac{1}{P_Y(y)} 2^{-\nu} \sum_{z \in Z^+} D'_y(z)\,.$$

Claim 3.6 For all $W \subseteq \zeta$ where $H_\infty(W) \geq \nu$, $E[D'_y(W)] \leq \beta_\alpha^+ P_Y(y)$.

Proof: Indeed,

$$\mathbb{E}[D'_y(W)] = \sum_{z \in \zeta} \Pr[W = z] D'_y(z) \leq \sum_{z \in \zeta} 2^{-\nu} D'_y(z) = 2^{-\nu} \sum_{z \in Z^+} D'_y(z) = P_Y(y)\, \mathbb{E}[D'_y(Z^+)]\,.$$

▮

Claim 3.7 $\mathbb{E}[D'_y(X)] \geq \beta_\alpha P_Y(y)$

Proof: One computes

$$\begin{aligned}
\mathbb{E}[D'_y(X)] &= \mathbb{E}[D'_y(X)|Y = y] \Pr[Y = y] + \mathbb{E}[D'_y(X)|Y \neq y] \Pr[Y \neq y] \\
&\geq \mathbb{E}[D'_y(X)|Y = y] \Pr[Y = y] \\
&= \beta_\alpha P_Y(y)
\end{aligned}$$

▮

By combining Claim 3.6, Claim 3.7, and (4) we have that for Z:

$$\mathbb{E}[D'_y(X)] - \mathbb{E}[D'_y(Z)] > \beta_\alpha P_Y(y) - \beta_\alpha^+ P_Y(y) = \epsilon' P_Y(y) = \epsilon \tag{6}$$

Thus, we have successfully distinguished the distribution X from Z. This is a contradiction. ▮

If we now consider averaging over all values of Y, we obtain the following simple formulation that expresses how much average entropy is left in X from the point of view of someone who knows Y. (This scenario naturally occurs in leakage-resilient cryptography, as exemplified in [19]).

Theorem 3.8 Let X, Y be discrete random variables. Then

$$H^{\mathtt{Metric}^*}_{\epsilon|Y|,s'}(X|Y) \geq H^{\mathtt{Metric}^*}_{\epsilon,s}(X) - \log|Y|$$

where $s' \approx s$ (recall that $|Y|$ is the size of the support of Y). The reduction is support preserving, in the same sense as in Lemma 3.2.

This statement is similar to the statement for the information-theoretic case (where the reduction is only in quantity, of course) from Lemma 2.1. In Appendix B, we compare this theorem to [11, Lemma 16] and [23, Lemma 3.1].

Unfortunately, it is not known whether `Metric`* entropy can be extracted from. Thus, for extraction we must convert the conditional `Metric`* entropy to conditional `HILL` entropy. As we discuss in Subsection 2.2, such a conversion is possible (Theorem 2.6), although it loses some quality; thus, it should be applied only when necessary (for instance, repeated conditioning is best measured in `Metric`* entropy, and then converted to `HILL` entropy once at the end). Here we provide a "HILL-to-HILL" formulation of Lemma 3.2.

Corollary 3.9 Let X be a discrete random variable over χ and let Y be a discrete random variable. Then,

$$H^{\mathtt{HILL}}_{\epsilon',s'}(X|Y=y) \geq H^{\mathtt{HILL}}_{\epsilon,s}(X) - \log 1/P_Y(y) \tag{7}$$

where $\epsilon' = \epsilon/P_Y(y) + \sqrt[3]{\frac{\log|\chi|}{s}}$, and $s' = \Omega(\sqrt[3]{s/\log|\chi|})$. The reduction is support preserving[10].

The Corollary follows by combining Lemma 3.2 and Theorem C.1, which is simply the support-preserving version of Theorem 2.6, and setting $\epsilon_{\mathtt{HILL}} = \sqrt[3]{\frac{\log|\chi|}{s}}$. A similar Corollary is available for conditioning on average-case Y (see Corollary B.4).

3.3 A (Crooked) Leftover Hash Lemma for Correlated Distributions

The following generalization of the (Crooked) LHL to correlated input distributions will be very useful to us when considering bounded multi-message security in Section 6. Since our generalization of the classical LHL is a special case of our generalization of the Crooked LHL, we just state the latter here.

Lemma 3.10 (CLHL for Correlated Sources) Let $\mathcal{H}\colon \mathcal{K} \times D \to R$ be a $2t$-wise δ-dependent function for $t > 0$ with range R, and let $f : R \to S$ be a function. Let $\mathbf{X} = (X_1, \ldots, X_t)$ where the X_i are random variables over D such that $\mathrm{H}_\infty(X_i) \geq \mu$ for all $1 \leq i \leq n$ and moreover $\Pr[\, X_i = X_j\,] = 0$ for all $1 \leq i \neq j \leq t$. Then

$$\Delta((K, f(\mathcal{H}(K,\mathbf{X}))), (K, f(\mathbf{U}))) \;\leq\; \frac{1}{2}\sqrt{|S|^t(t^2 2^{-\mu} + 3\delta)} \tag{8}$$

where $K \xleftarrow{\$} \mathcal{K}$ and $\mathbf{U} = (U_1, \ldots, U_t)$ where the U_i are all uniform and independent over R (recall that functions operate on vectors $\mathbf{X}$ and $\mathbf{U}$ component-wise).

Note that the lemma implies the corresponding generalization of the classical LHL by taking $\mathcal{H}$ to have range S and f to be the identity function.

The proof of the above lemma, which extends the proof of the Crooked LHL in [7], is in Appendix D.

Remark 3.11 We can further extend Lemma 3.10 to the case of *average conditional min-entropy* using the techniques of [16]. Such a generalization (without considering correlated sources) is similarly useful in the context of randomized encryption from lossy TDFs [38].

[10] "Support preserving" for `HILL` entropy is similar to the same notion for `Metric`* entropy explained in Lemma 3.2. It simply means that the distribution $Z_{X|Y=y}$, which is indistguishable from $X|Y=y$ according to the definition of `HILL` entropy, has no greater support than the distribution Z_X which is indistinguishable from X.

4 Deterministic Encryption from Robust Hardcore Functions

4.1 Robust Hardcore Functions

We introduce a new notion of *robustness* for hardcore functions. Intuitively, robust HCFs are those that remain one-way when the input is conditioned on an event that occurs with good probability. We expand on this below.

Definition 4.1 Let $\mathcal{F}$ be a TDF generator and let hc be a HCF such that hc is hardcore for $\mathcal{F}$ with respect to a distribution $\boldsymbol{X}$ on input vectors. For $\alpha = \alpha(k)$, we say hc is *α-robust for $\mathcal{F}$ on $\boldsymbol{X}$* if hc is also hardcore for $\mathcal{F}$ with respect to the class $\boldsymbol{X}[\alpha]$ of α-induced distributions of $\boldsymbol{X}$.

Discussion. Robustness is interesting even for the classical definition of hardcore bits, where hc is boolean and a single uniform input x is generated in the security experiment. Here robustness means that hc remains hardcore even when x is conditioned on an event that occurs with good probability. It is clear that not every hardcore bit in the classical sense is robust — note, for example, that while every bit of the input to RSA is well-known to be hardcore assuming RSA is one-way [1], they are not even 1-robust since we may condition on a particular bit of the input being a fixed value. It may also be interesting to explore robustness in contexts other than DE, such as leakage resilience [34] and computational randomness extraction (or key derivation) [33].

4.2 The Encrypt-with-Hardcore Scheme

The scheme. Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be a probabilistic encryption scheme, $\mathcal{F}$ be a TDF generator, and hc_f be a HCF. Assume that hc outputs binary strings of the same length as the random string r needed by $\mathcal{E}$. Define the associated "*Encrypt-with-Hardcore*" deterministic encryption scheme $\mathsf{EwHCore}[\Pi, \mathcal{F}, \mathsf{hc}] = (\mathcal{DK}, \mathcal{DE}, \mathcal{DD})$ with plaintext-space $\mathrm{PtSp} = \{0,1\}^k$ via

Algorithm $\mathcal{DK}(1^k)$:	Algorithm $\mathcal{DE}((pk, f), x)$:	Algorithm $\mathcal{DD}((sk, f^{-1}), c)$:
$(pk, sk) \xleftarrow{\$} \mathcal{K}(1^k)$	$r \leftarrow \mathsf{hc}_f(x)$	$y \leftarrow \mathcal{D}(sk, c)$
$(f, f^{-1}) \xleftarrow{\$} \mathcal{F}(1^k)$	$c \leftarrow \mathcal{E}(pk, f(x); r)$	$x \leftarrow f^{-1}(y)$
Return $((pk, f), (sk, f^{-1}))$	Return c	Return x

Security analysis. To gain some intuition, suppose hc is hardcore for $\mathcal{F}$ on some distribution $\boldsymbol{X}$ on input vectors. One might think that PRIV security of $\mathsf{EwHCore} = \mathsf{EwHCore}[\Pi, \mathcal{F}, \mathsf{hc}, g]$ on $\boldsymbol{X}$ then follows by IND-CPA security of Π. However, this is not true. To see this, suppose hc is a "natural" hardcore function (i.e., outputs some bits of the input). Define $\Pi' = (\mathcal{K}, \mathcal{E}', \mathcal{D}')$ to be like $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ except that the coins consumed by $\mathcal{E}'$ are extended by one bit, which $\mathcal{E}'$ outputs in the clear and $\mathcal{D}'$ ignores. That is, define $\mathcal{E}'(pk, x; r\|b) = \mathcal{E}(pk, x; r)\|b$ and $\mathcal{D}'(sk, y\|b) = \mathcal{D}(sk, y)$. Then IND-CPA security of Π' follows from that of Π, but a straightforward attack shows $\mathsf{EwHCore}$ is not PRIV on $\boldsymbol{X}$. This is how our notion of robustness comes into play.

Theorem 4.2 Suppose Π is IND-CPA secure, hc is *2-robust* for $\mathcal{F}$ on a distribution $\boldsymbol{M}$ on input vectors. Then $\mathsf{EwHCore}[\Pi, \mathcal{F}, \mathsf{hc}]$ is PRIV-secure on $\boldsymbol{M}$.

The theorem follows from combining Theorem 3.1 with the following lemma, which shows that what does follow if hc is hardcore (but not necessarily robust) is the *IND* security of $\mathsf{EwHCore}$.

Lemma 4.3 Suppose Π is IND-CPA, hc is hardcore for $\mathcal{F}$ on a distribution $\boldsymbol{M}$ on input vectors, and that g is pseudorandom. Then $\mathsf{EwHCore} = \mathsf{EwHCore}[\Pi, \mathcal{F}, \mathsf{hc}, g]$ is IND secure on $\boldsymbol{M}$. In particular, let $D \in \mathbb{D}_{\boldsymbol{M}}$

be a IND adversary against $\mathsf{EwHCore}$. Then there is an IND-CPA adversary A against Π, an adversary B against hc on $\boldsymbol{M}$ such that for all $k \in \mathbb{N}$

$$\mathbf{Adv}^{\mathrm{ind}}_{\mathsf{EwHCore},D}(k) \quad \leq \quad \mathbf{Adv}^{\mathrm{ind\text{-}cpa}}_{\Pi,A}(k) + 2 \cdot \mathbf{Adv}^{\mathrm{hcf}}_{\mathcal{F},\mathsf{hc},\boldsymbol{M},B}(k)\ . \tag{9}$$

Furthermore, the running-times of A, B are the time to run D.

Proof: Let Game G_1 correspond to the IND experiment with D against $\mathsf{EwHCore}$, and let Game G_2 be like G_1 except that the coins used to encrypt the challenge plaintext vector are truly random. For $i \in \{0,1\}$ let $B^i = (B^i_1, B^i_2)$ be the HCF adversary against $\mathcal{F}$ hc defined via

Algorithm $B^i_1(1^k)$:	**Algorithm** $B^i_2(pk, \mathbf{y}, \mathbf{h})$:
$\mathbf{x} \xleftarrow{\$} D_1(i)$	$\mathbf{c} \leftarrow \mathcal{E}(pk, \mathbf{y}; \mathbf{h})$
Return $\mathbf{x}$	$d \xleftarrow{\$} D_2(pk, \mathbf{c})$
	Return d

Then

$$\begin{aligned}
\Pr\left[\, G_1^D \Rightarrow b \,\right] &= \Pr\left[\, G_1^D \Rightarrow b \;|\; b = 1 \,\right] + \Pr\left[\, G_1^D \Rightarrow b \;|\; b = 0 \,\right] \\
&= \Pr\left[\, G_2^D \Rightarrow b \;|\; b = 1 \,\right] + \mathbf{Adv}^{\mathrm{hcf}}_{\mathcal{F},\mathsf{hc},B^1}(k) \\
&\quad + \Pr\left[\, G_2^D \Rightarrow b \;|\; b = 0 \,\right] + \mathbf{Adv}^{\mathrm{hcf}}_{\mathcal{F},\mathsf{hc},B^0}(k) \\
&\leq \Pr\left[\, G_2^D \Rightarrow b \,\right] + 2 \cdot \mathbf{Adv}^{\mathrm{hcf}}_{\mathcal{F},\mathsf{hc},B}(k)
\end{aligned}$$

where we take B to be whichever of B^0, B^1 has the larger advantage. Now define IND-CPA adversary A against Π via

Algorithm $A_1(pk)$:	**Algorithm** $A_2(pk, \mathbf{c})$:
$\mathbf{x}_0 \xleftarrow{\$} D_1(0)$	$d \xleftarrow{\$} D_2(pk, \mathbf{c})$
$\mathbf{x}_1 \xleftarrow{\$} D_1(1)$	Return d
Return $(\mathbf{x}_0, \mathbf{x}_1)$	

Then (9) follows from taking into account the definition of the advantages of D, A. ▮

A subtle point worth mentioning is where in the proof we use the fact that the Lemma 4.3 considers IND security of $\mathsf{EwHCore}$ rather than PRIV (which, as we have said, does not follow). It is in the step that uses security of the hardcore function. If we considered PRIV security, in this step the constructed HCF adversaries against $\mathcal{F}$ would need to test whether the output of the PRIV adversary against $\mathsf{EwHCore}$ is equal to a "target value" representing partial information on the input to $\mathcal{F}$, which these adversaries are not given. Indeed, this is exactly what caused complications in the original analysis of the scheme of [5], who used the PRIV notion directly.

5 Single-Message Instantiations of Encrypt-with-Hardcore

5.1 Getting Robust Hardcore Functions

Making any large hardcore function robust. We show that by applying a randomness extractor in a natural way, one can convert *any* large hardcore function in the standard sense to one that is robust (with some loss in parameters). However, while the conversion procedure is natural, proving that it works turns out to be non-trivial. We show that by applying a randomness extractor in a natural way, one can convert *any* large hardcore function in the standard sense to one that is robust (with some loss in parameters). However, while the conversion procedure is natural, proving that it works turns out to be non-trivial.

For a random variable A with support $\mathcal{A}$, define the *entropy discrepancy* of A as $\mathsf{disc}(A) = \log|\mathcal{A}| - \mathrm{H}_\infty(A) = H_0(A) - \mathrm{H}_\infty(A)$. Let $\mathcal{F}$ be a TDF generator. Let $\mathsf{disc}_k(f)$ be the entropy discrepancy of the public key f, viewed as a random variable produced by $\mathcal{F}(1^k)$. Let X be an input distribution for f and $\mathsf{hc}\colon \{0,1\}^k \to \{0,1\}^\ell$ be an HCF for f on X. Let $\mathtt{ext} : \{0,1\}^\ell \times \{0,1\}^d \to \{0,1\}^m \times \{0,1\}^d$ be a strong average-case $(\ell - \alpha - \mathsf{disc}(f) - \mathsf{disc}(X), \epsilon_{\mathtt{ext}})$-extractor for $\alpha \in \mathbb{N}$ that takes time $t_{\mathtt{ext}}$ to compute. Define a new *"extractor-augmented" HCF* $\mathsf{hc}[\mathtt{ext}]$ for $\mathcal{F}[\mathtt{ext}]$ as follows: $\mathsf{hc}[\mathtt{ext}]_s(x) = \mathtt{ext}(\mathsf{hc}(x), s)$ for all $x \in \{0,1\}^k$ and $s \in \{0,1\}^d$. (Here we view $\mathtt{ext}$ as a keyed function with the *second* argument as the key.) The following characterizes the α-robustness of $\mathsf{hc}[\mathtt{ext}]$.

Lemma 5.1 If hc is a sufficiently long hardcore function for $\mathcal{F}$ on an input distrubition X, then $\mathsf{hc}[\mathtt{ext}]$ is a hardcore function for any input distribution $X' \in X[\alpha]$. More precisely, if

$$(f, f(X), \mathsf{hc}(X)) \approx_{t,\epsilon} (f, f(X), U_\ell) \text{ , then}$$

$$(f, f(X'), \mathtt{ext}(\mathsf{hc}(X'), U_d), U_d) \approx_{t' - t_{\mathtt{ext}},, 2\epsilon' + \epsilon_{\mathtt{ext}}} (f, f(X'), U_m, U_d) \, ,$$

where in both equations f is distributed according to $\mathcal{F}(1^k)$, and $\epsilon' = \epsilon \cdot 2^\alpha + \sqrt[3]{(k + \log|\mathcal{F}| + \ell)/t}$ and $t' = \Omega(\sqrt[3]{t/(k + \log|\mathcal{F}| + \ell)})$.

We note that in order to apply this lemma, $(\ell - \alpha - \mathsf{disc}(f) - \mathsf{disc}(X))$ must be large enough in order to allow for a useful extractor. Thus, the "entropy loss" is not only α (which is expected, because it is the entropy deficiency of X'), but also $\mathsf{disc}(f) + \mathsf{disc}(X)$. Therefore, we need the starting hardcore function output length ℓ to be sufficiently large compared to the entropy discrepancies of both f and X. Fortunately, for typical trapdoor functions such as RSA, $\mathsf{disc}(f)$ is 0 because the distribution of public keys produced by the key generation method is flat. Moreover, sufficiently long ℓ can always be achieved if the starting hardcore function output is long enough to be used as a seed for a pseudorandom generator, since then it can be expanded to any polynomial length (here we are referring to running the hardcore function through a pseudorandom generator *before* applying the extractor, thus changing hc to have longer output ℓ).

Also note that when $\alpha = \log(k)$, the security loss in the reduction is polynomial (in our application we just need $\alpha = 2$). We note that the conversion procedure also works when hc is hardcore on a distribution $\boldsymbol{X}$ on input *vectors*, but we omit this since we do not know any examples of "natural" hardcore functions that are secure on correlated inputs. (Looking ahead, in Section 6 we give a direct constructions of the such hardcore function without needing the conversion procedure of Lemma 5.1.)

Proof: Let f be distributed according to the distribution of public keys produced by $\mathcal{F}(1^k)$. Slightly abusing notation, we will also denote the support of this distribution by $\mathcal{F}$. Assume that for $t, \epsilon > 0$

$$(f, f(X), \mathsf{hc}(X)) \approx_{t,\epsilon} (f, f(X), U_\ell) \, . \tag{10}$$

By definition of $\mathtt{HILL}$ entropy, $H^{\mathtt{HILL}}_{\epsilon,t}(f, f(X), \mathsf{hc}(X)) \geq \mathrm{H}_\infty(f, f(X), U_\ell) = \mathrm{H}_\infty(f) + \mathrm{H}_\infty(X) + \ell$ (using the fact that f is injective). Let ζ denote the set of all triples (f, y, r) such that $f \in \mathcal{F}$, and $y = f(x)$ for some $x \in X$. Let E be such that $X' = X \mid \mathsf{E}$; note that $\Pr[\mathsf{E}] = 2^{-\alpha}$. Applying the "HILL-to-HILL" Corollary 3.9, we know that

$$H^{\mathtt{HILL}}_{\epsilon',t'}(f, f(X), \mathsf{hc}(X) \mid \mathsf{E}) \geq H^{\mathtt{HILL}}_{\epsilon,t}(f, f(x), \mathsf{hc}(X)) - \alpha \geq \mathrm{H}_\infty(f) + \mathrm{H}_\infty(X) + \ell - \alpha \, ,$$

where $\epsilon' = \epsilon \cdot 2^\alpha + \sqrt[3]{(k + \log|\mathcal{F}| + \ell)/t}$, and $t' = \Omega(\sqrt[3]{t/(k + \log|\mathcal{F}| + \ell)})$. By Definition 2.4 of HILL entropy and the fact that Corollary 3.9 is support preserving, this implies that there exist random variables $(A, B, C) \subseteq \zeta$ such that

$$(f, f(X), \mathsf{hc}(X)) \mid \mathsf{E} \approx_{t',\epsilon'} (A, B, C) \, , \tag{11}$$

and furthermore $\mathrm{H}_\infty((A,B,C)) \geq \mathrm{H}_\infty(f) + \mathrm{H}_\infty(X) + \ell - \alpha$. Because an independent random string does not help the distinguisher,

$$(f, f(X), \mathsf{hc}(X), U_d) \mid \mathsf{E} \approx_{t',\epsilon'} (A, B, C, U_d) .$$

Because applying a deterministic function to the distributions can help the distinguisher by at most the time it takes to compute the function,

$$(f, f(X), \mathtt{ext}(\mathsf{hc}(X), U_d), U_d) \mid \mathsf{E} \approx_{t'-t_{\mathtt{ext}},\epsilon'} (A, B, \mathtt{ext}(C, U_d), U_d) . \tag{12}$$

We now claim that

$$(A, B, \mathtt{ext}(C, U_d), U_d) \approx_{\epsilon_{\mathtt{ext}}} (A, B, U_\ell, U_d) . \tag{13}$$

Indeed,

$$\begin{aligned}\tilde{\mathrm{H}}_\infty(C \mid (A,B)) &\geq \mathrm{H}_\infty(A,B,C) - \log|B| - \log|C|| \\ &\geq \mathrm{H}_\infty(A,B,C) - \log|\mathcal{F}| - \log|f(X)| \\ &\geq \ell - \alpha - \mathsf{disc}(f) - \mathsf{disc}(X) ,\end{aligned}$$

where the first inequality uses Lemma 2.1, the second inequality follows from $A \subseteq \mathcal{F}$ and $B \subseteq f(X)$, and the final inequality follows from the definition of (A,B,C), the definition of disc, and the fact that f is injective. Thus, (13) follows by security of $\mathtt{ext}$. Note that (11) implies that $(f, f(X)) \mid E \approx_{t',\epsilon'} (A,B)$, which implies

$$(A, B, U_\ell, U_d) \approx_{t',\epsilon'} (f, f(X), U_\ell, U_d) \mid \mathsf{E} . \tag{14}$$

Combining (12), (13), (14) via the triangle inequality we have

$$(f, f(X), \mathtt{ext}(\mathsf{hc}(X), U_d), U_d) \mid \mathsf{E} \approx_{t'-t_{\mathtt{ext}},2\epsilon'+\epsilon_{\mathtt{ext}}} (f, f(X), U_\ell, U_d) \mid \mathsf{E} . \tag{15}$$

Recalling that f is distributed independently of E and $X' = X|\mathsf{E}$, we get the statement of the Lemma. ▮

Remark 5.2 The conclusion of the lemma actually holds given a weaker hypothesis on the starting hardcore function. Namely, its output need not be indistinguishable from uniform but rather have high computational (HILL) entropy.

The above conversion procedure notwithstanding, we give specific examples of hardcore functions that are already robust without requiring the former. This is especially useful to view constructions from both one-wayness as in [5] and from lossiness as in [7] in a unified way: these constructions emanate from the fact that both "one-way hardness" and min-entropy are preserved on slightly induced distributions.

ROBUST GOLDREICH-LEVIN BITS FOR ANY TDF. First, we show that the Goldreich-Levin [24] hardcore function as considered in [5] is robust. Indeed, robustness of Goldreich-Levin follows from the following simple lemma, which describes how "one-way hardness" on an input distribution is preserved on induced distributions.

Lemma 5.3 Let $\mathcal{F}$ be a TDF generator. Let X be an input distribution and fix $X' \in X[\alpha]$ for $\alpha \in \mathbb{N}$. Then for any inverter I' against $\mathcal{F}$ on X' there is an inverter I against $\mathcal{F}$ on X such that for all $k \in \mathbb{N}$

$$\mathbf{Adv}^{\mathrm{owf}}_{\mathcal{F},X',I'}(k) \leq 2^\alpha \cdot \mathbf{Adv}^{\mathrm{owf}}_{\mathcal{F},X,I}(k) . \tag{16}$$

Furthermore, the running-time of I is the time to run I'.

Proof: (of Lemma 5.3) Let I' be the inverter that simply runs I on its input, and let E be the corresponding event to $\boldsymbol{X}'$. Let G be the event that $\mathbf{Exp}^{\mathrm{owf}}_{\mathcal{F},\boldsymbol{X}',I'}(k) \Rightarrow 1$. Then

$$\begin{aligned}
\mathbf{Adv}^{\mathrm{owf}}_{\mathcal{F},\boldsymbol{X}',I'}(k) &= \Pr\left[\, G \mid \mathsf{E} \,\right] \cdot \Pr\left[\, \mathsf{E} \,\right] + \Pr\left[\, G \mid \overline{\mathsf{E}} \,\right] \cdot \Pr\left[\, \overline{\mathsf{E}} \,\right] \\
&\geq \Pr\left[\, G \mid \mathsf{E} \,\right] \cdot \Pr\left[\, \mathsf{E} \,\right] \\
&= \mathbf{Adv}^{\mathrm{owf}}_{\mathcal{F},\boldsymbol{X},I}(k) \cdot 1/2^{-\alpha} ,
\end{aligned}$$

from which (16) follows by re-arranging terms. ▮

Note that when $\alpha = O(\log k)$ the reduction incurs a polynomial loss in advantage (again, in our applications we just need $\alpha = 2$). Then we have the following result.

Proposition 5.4 Let $\mathcal{F}[\mathcal{GL}^i]$ be as defined above and suppose $\mathcal{GL}^i$ is hardcore for $\mathcal{F}[\mathcal{GL}^i]$ on single-input distribution X. Then $\mathcal{GL}^i$ is $O(\log k)$-robust for $\mathcal{F}[\mathcal{GL}^i]$ on X.

Robust bits for any LTDF. Peikert and Waters [38] showed that LTDFs admit a simple, large hardcore function, namely a pairwise-independent hash function (the same argument applies also to universal hash functions or, more generally, randomness extractors). We show robustness of the latter based on the following simple lemma, which says that min-entropy of a given input distribution is preserved on sub-distributions induced by an event that occurs with good probability.

Lemma 5.5 Let X be a random variable with $\mathrm{H}_\infty(X) \geq \mu$, and let X' be a random variable where $P_{X'}$ is a an α-induced sub-distribution of P_X. Then $\mathrm{H}_\infty(X') \geq \mu - \alpha$.

Proof: (of Lemma 5.5) Suppose not, and let E be the corresponding event to X'. Then there exists an x' such that $P_{X'}(x') > 2^{-\mu+\alpha}$. But then

$$\begin{aligned}
P_X(x') &\geq \Pr\left[\, X = x' \mid \mathsf{E} \,\right] \cdot \Pr\left[\, \mathsf{E} \,\right] + \Pr\left[\, X = x' \mid \overline{\mathsf{E}} \,\right] \cdot \Pr\left[\, \overline{\mathsf{E}} \,\right] \\
&\geq \Pr\left[\, X = x' \mid \mathsf{E} \,\right] \cdot \Pr\left[\, \mathsf{E} \,\right] \\
&> 2^{-\mu+\alpha} \cdot 2^{-\alpha} \\
&= 2^{-\mu}
\end{aligned}$$

a contradiction. ▮

By combining the Generalized Leftover Hash Lemma of [16] (i.e., for the case of average min-entropy) with the "chain rule" for average conditional min-entropy (Lemma 2.1), it follows that if $\mathcal{F}$ is a lossy trapdoor function generator with residual leakage s, then a pairwise-independent hash function $\mathcal{H}\colon \mathcal{K} \times \{0,1\}^k \to \{0,1\}^r$ is hardcore for $\mathcal{F}[\mathcal{H}]$ on any single-input distribution X with min-entropy $s + r + 2(\log 1/\epsilon)$ for negligible ϵ (as compared to [38, Lemma 3.4], we simply observe that the argument does not require the input to be uniform). Then, using Lemma 5.5 we have the following.

Proposition 5.6 Let $\mathsf{LTDF} = (\mathcal{F}, \mathcal{F}')$ be a LTDF generator with residual leakage s, and let $\mathcal{H}\colon \mathcal{K} \times \{0,1\}^k \to \{0,1\}^r$ be a pairwise-independent hash function. Then $\mathcal{H}$ is a $O(\log k)$-robust hardcore function for $\mathcal{F}[\mathcal{H}]$ on any single-input distribution X with min-entropy $s + r + 2(\log 1/\epsilon)$ for negligible ϵ.

5.2 Putting It Together

Equipped with the above results, we describe instantiations of the Encrypt-with-Hardcore scheme that both explain prior constructions and produce novel ones.

Using an iterated trapdoor permutation. The prior trapdoor-permutation-based DE scheme of Bellare et al. [5] readily provides an instantiation of $\mathsf{EwHCore}$ by using an iterated trapdoor permutation as the TDF. Let $\mathcal{F}$ be a TDP and hc be a hardcore bit for $\mathcal{F}$. For $i \in \mathbb{N}$ denote by $\mathcal{F}^i$ the TDP that iterates $\mathcal{F}$ i-many times. Define the Blum-Micali-Yao (BMY) [6, 50] hardcore function for $\mathcal{F}^i$ via $\mathcal{BMY}^i[\mathsf{hc}](f, x) = \mathsf{hc}(x)\|\mathsf{hc}(f(x))\|\dots\|\mathsf{hc}(f^{i-1})$. Bellare et al. [5] used the specific choice of $\mathsf{hc} = \mathcal{GL}$ (the GL bit) in their scheme, which is explained by the fact that the latter is robust as per Proposition 5.4 and one can show that BMY iteration expands one robust hardcore bit to many (on a non-uniform distribution, the bit should be hardcore on all "permutation distributions" of the former).

However, due to our augmentation procedure to make any large hardcore function robust, we are no longer bound to any specific choice of hc. For example, we may choose hc to be a natural bit of the input in the case that the latter is hardcore. In fact, it may often be the case that $\mathcal{F}$ has many simultaneously hardcore natural bits, and therefore our construction will require fewer iterations of the TDP than the construction of [5].

Using a lossy TDF. Applying Proposition 5.6, we get an instantiation of the Encrypt-with-Hardcore scheme from lossy TDFs that is an alternative to the prior scheme of Boldyreva et al. [7] and the concurrent work of Wee [48]. Our scheme requires an LTDF with residual leakage $s \leq \mathrm{H}_\infty(X) - 2\log(1/\epsilon) - r$, where r is the number of random bits needed in $\mathcal{E}$ (or the length of a seed to a pseudorandom generator that can be used to obtain those bits). Thus the LTDF should lose a constant fraction of its input. To compare, the prior scheme of [7] encrypts under (an augmented version of) the LTDF directly and does not use the "outer" encryption scheme at all. Its analysis requires the "Crooked" LHL of Dodis and Smith [17] rather than the standard LHL but gets rid of r in the above bound leading to a better requirement on lossiness or input entropy.

Using 2-correlated product TDFs. Hemenway et al. [28] show a construction of DE from a *decisional 2-correlated product TDF*, namely where $\mathcal{F}$ has the property that $f_1(x), f_2(x)$ is indistinguishable from $f_1(x_1), f_2(x_2)$ where x_1, x_2 are sampled independently (in both cases for two independent public instances f_1, f_2 of $\mathcal{F}$). (This property is a strengthening of the notion of security under correlated products introduced in [42].) They show such a trapdoor function is a secure DE scheme for uniform messages. To obtain an instantiation of $\mathsf{EwHCore}$ under the same assumption, we can use $\mathcal{F}$ as the TDF, and an independent instance of the TDF as hc. When a randomness extractor is applied to the latter, robustness follows from Lemma 5.1, taking into account Remark 5.2.

Using any TDF with a large HCF. Our most novel instantiations in the single-message case come from considering TDFs that have a sufficiently large HCF but are not necessarily lossy or an iterated TDP. Let us first consider instantiations on the uniform message distribution (an important special case as highlighted in [5]). It was recently shown by Freeman et al. [21] that the Niederreiter TDF [35] has linearly many (simultaneous) hardcore bits under the "Syndrome Decoding Assumption (SDA)" and "Indistinguishability Assumption (IA)" as defined in [21, Section 7.2], which are already needed to show the TDF is one-way. Furthermore, the RSA [41] and Paillier [37] TDPs have linearly many hardcore bits under certain computational assumptions, namely the "Small Solutions RSA (SS-RSA) Assumption" [45] and the "Bounded Computational Composite Residuosity (BCCR) Assumption" [9] respectively. Because these hardcore functions are sufficiently long, they can be made robust via Lemma 5.1 and give us a linear number of *robust* hardcore bits—enough to use as randomness for $\mathcal{E}$ (expanded by a pseudorandom generator if necessary). (Here the "outer" encryption scheme can be instantiated under the same assumptions.) Thus, by Theorem 4.2, we obtain:

Corollary 5.7 Under SDA+IA for the Niederreiter TDF, DE for the uniform message distribution exists. Similarly, under SS-RSA the RSA TDP or BCCR for the Paillier TDP respectively, DE for the uniform message distribution exists.

In particular, the first statement provides the first DE scheme without random oracles based on the hardness of syndrome decoding. (A scheme in the random oracle model follows from [3].) Moreover, the schemes provided by the second statement are nearly as efficient as the ones obtained from lossy TDFs (since they do not use iteration), and the latter typically requires decisional assumptions (in contrast to the computational assumptions used here).

If we do not wish to rely on specific assumptions, we can also get DE from strong but general assumptions. Specifically, for general $\mathcal{F}$, we can obtain a large enough HCF by using enough GL bits and assuming the TDF is sufficiently hard to invert.[11] The GL theorem (Theorem 2.3) tells us that if $\mathcal{F}$ is s-hard on X then it has a HCF on X with output length $c' \log s$ for some constant c'. Note we can trade hardness of the TDF for greater hardness of an underlying PRG used to expand the HCF, which can be built from a one-way function *without* a trapdoor. For example, we can assume a TDF $\mathcal{F}$ that is quasi-polynomially hard to invert, which yields a GL HCF on with poly-logarithmic output length, and expand it via a PRG with sub-exponential hardness (which could be built assuming a sub-exponentially hard one-way function).

To obtain instantiations on message distributions of less than maximal entropy, we can use a technical lemma [20, Lemma 4] saying that every distribution with min-entropy α less than maximal can be viewed as an α-induced distribution of the uniform distribution, and take into account Remark 5.2. By Corollary 3.9, we know the HILL entropy of a HCF on such a distribution degrades in quantity by α and in quality polynomially in 2^α. Thus, assuming the HCF is sufficiently long and sufficiently hard to distinguish from uniform, it can still be turned into a robust HCF using Remark 5.2 For example, if $\alpha = O(\log(k))$, a standard hardness assumption suffices. We thus obtain the analogue of Corollary 5.7 for distributions whose min-entropy is logarithmically away from maximal under the same assumptions.

For any $\alpha = o(k)$, we can obtain DE for distributions of min-entropy α away from maximal by assuming sub-exponential hardness of simultaneous hardcore bits. That is, the analogue of Corollary 5.7 holds under sub-exponential hardness of the assumptions.

6 Bounded Multi-Message Security and its Instantiations

6.1 The New Notion and Variations

The new notion. The notion of q-bounded multi-message security (or just q-bounded security) for DE is quite natural, and parallels the treatment of "bounded" security in other contexts (e.g. [12]). In a nutshell, it asks for security on up to q arbitrarily correlated but high-entropy messages (where we allow the public-key size to depend on q). More formally, fix an encryption scheme $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$. For $q = q(k)$ and $\mu = \mu(k)$, let $\mathbb{M}^{q,\mu}$ be the class of distributions on message vectors $M^{\mu,q} = (M_1^{\mu,q}, \ldots, M_q^{\mu,q})$ where $\mathrm{H}_\infty(M_i^{\mu,q}) \geq \mu$ and for all $1 \leq i \leq q$ and $M_{1,q}^{\mu}, \ldots, M_{q,q}^{\mu}$ are distinct with probability 1. We say that Π is *q-bounded multi-message PRIV (resp. IND) secure for μ-sources* if it is PRIV (resp. IND) secure for $\mathbb{M}^{q,\mu}$. We note that Theorem 3.1 (combined with Lemma 5.5) tells us that PRIV on $\mathbb{M}^{q,\mu}$ is equivalent to IND on $\mathbb{M}^{q,\mu-2}$.

Unbounded multi-message security for q-block sources. We also consider unbounded multi-message security for what we call a *q-block source*, a generalization of a block-source [10] where every q-th message introduces some "fresh" entropy. More formally, fix an encryption scheme $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$. For $q = q(k)$, $n = n(k)$, and $\mu = \mu(k)$, let $\mathbb{M}^{q,n,\mu}$ be the class of distributions on message vectors $M^{q,n,\mu} = (M_1^{q,n,\mu}, \ldots, M_{qn}^{q,n,\mu})$ such that $\mathrm{H}_\infty(X_{qi+j} \mid X_1 = x_1, \ldots, X_{qi-1} = x_{qi-1}) \geq \mu$ for all $1 \leq i \leq n$, all $0 \leq j \leq q-1$, and all outcomes $x_1, \ldots, x_{qi-1}$ of $X_1, \ldots, X_{qi-1}$. We say that Π is *q-bounded multi-message PRIV (resp. IND) secure for (μ, n)-block-sources* if Π is PRIV (resp. IND) secure on $\mathbb{M}^{q,n,\mu}$. Using a similar argument to [7, Theorem 4.2], one can show equivalence of PRIV on $\mathbb{M}^{q,n,\mu}$ to IND on $\mathbb{M}^{q,n,\mu}$.

[11] For very long messages, on the uniform distribution we can actually apply any TDF block-wise to collect a large hardcore function from individual GL bits, but this does not extend to lower entropy messages.

6.2 Our Basic Scheme

Note that we cannot trivially achieve q-bounded security by running, say, q copies of a scheme secure for one message in parallel (and encrypting the i-th message under the i-th public key), since this approach would lead to a stateful scheme. The main technical tool we use to achieve the notion is Lemma 3.10. Combined with Lemma 2.1, this tells us that a $2q$-wise independent hash function is robust on correlated input distributions of sufficient min-entropy:

Proposition 6.1 For any q, let $\mathsf{LTDF} = (\mathcal{F}, \mathcal{F}')$ be an LTDF generator with input length n and residual leakage s, and let $\mathcal{H}\colon \mathcal{K} \times D \to R$ where $r = \log |R|$ be a $2q$-wise independent hash function. Then $\mathcal{H}$ is a 2-robust hardcore function for $\mathcal{F}$ on any input distribution $X = (X_1, \ldots, X_q)$ such that $\mathrm{H}_\infty(X) \geq q(s + r) + 2\log q + 2\log(1/\epsilon) - 2$ for negligible ϵ.

Thus, by Theorem 4.2 we obtain a q-bounded multi-message secure DE scheme based on lossy trapdoor functions. Note that since we require $(\mathrm{H}_\infty(X) - 2\log q - \log(1/\epsilon))/q - r \geq s$ (where r is the number of random bits needed in $\mathcal{E}$, or the length of a seed to a pseudorandom generator that can be used to obtain those bits) the lossy trapdoor function must lose a $1 - O(1/q)$ fraction of its input. The DDH-based construction of Peikert and Waters [38], the Paillier-based one of [7, 21], and the one from d-linear of [21] can all satisfy this requirement for any polynomial q.

6.3 Our Optimized Scheme

We show that by extending some ideas of [7], we obtain a more efficient DE scheme meeting q-bounded security that achieves better parameters.

Intuition and preliminaries. Intuitively, for the optimized scheme we modifying the scheme of [7] to first pre-process an input message using a $2q$-wise independent permutation (instead of pairwise as in [7]). However, there are two issues to deal with here. First, for $q > 1$ such a permutation is not known to exist (in an explicit and efficiently computable sense). Second, Lemma 3.10 applies to t-wise independent *functions* rather than permutations. (In the case $t = 2$ as considered in [7] the difference turns out to be immaterial.)

To solve the first problem, we turn to $2q$-wise "δ-dependent" permutations (as constructed in e.g. [30]). Namely, say that a permutation $H\colon \mathcal{K} \times D \to D$ is *t-wise δ-dependent* if for all distinct $x_1, \ldots, x_t \in D$

$$\Delta((H(K, x_1), \ldots, H(K, x_t)), (P_1, \ldots, P_t)) \ \leq\ \delta\ ,$$

where $K \xleftarrow{\$} \mathcal{K}$ and $P_1, \ldots, P_t$ are defined iteratively by taking P_1 to be uniform on D and, for all $2 \leq i \leq t$, taking P_i to be uniform on $R \backslash \{p_1, \ldots, p_{i-1}\}$ where $p_1, \ldots, p_{i-1}$ are the outcomes of $P_1, \ldots, P_{i-1}$ respectively.

To solve the second problem, we use the following lemma, which says that a t-wise δ-dependent permutation is a t-wise δ'-dependent function where δ' is a bit bigger than δ.

Lemma 6.2 Suppose $H\colon \mathcal{K} \times D \to D$ is a t-wise δ-dependent permutation for some $t \geq 1$. Then $\mathcal{H}$ is a t-wise δ-dependent *function* for $\delta' = \delta + t^2/|D|$.

The proof uses the fact that the distribution of $(P_1, \ldots, P_t)$ equals the distribution of $(U_1, \ldots, U_t) \mid \mathsf{DIST}$ where DIST is the event that $U_1, \ldots, U_t$ are all distinct and then applies a union bound. It will be useful to now restate Lemma 3.10 in terms of δ-dependent permutations, which follows by combining Lemma 3.10 and Lemma 6.2.

Lemma 6.3 (CLHL for Correlated Sources with Permutations) Let $\mathcal{H}\colon \mathcal{K} \times D \to D$ be a δ-dependent t-wise permutation for $t > 0$ with range R, where $\delta = t^2/|D|$. Let $f : R \to S$ be a function. Let

$\mathbf{X} = (X_1, \ldots, X_t)$ where the X_i are random variables over D such that $\mathrm{H}_\infty(X_i) \geq \mu$ for all $1 \leq i \leq n$ and moreover $\Pr[\, X_i = X_j \,] = 0$ for all $1 \leq i \neq j \leq t$. Then

$$\Delta((K, f(\mathcal{H}(K, \mathbf{X}))), (K, f(\mathbf{U}))) \;\leq\; \frac{1}{2}\sqrt{7|S|^t t^2 2^{-\mu}} \tag{17}$$

where $K \xleftarrow{\$} \mathcal{K}$ and $\mathbf{U} = (U_1, \ldots, U_t)$ where the U_i are all uniform and independent over R (recall that functions operate on vectors component-wise).

It is interesting to note here that the the bound in Equation (17) is essentially as good as the one in Equation (8). At first one might not expect this to be the case. Indeed, when the classical LHL is extended to "imperfect" hash functions [44, 15], the error probability must be taken much smaller than $1/|R|$, where R is the range of the hash function. But in Lemma 3.10 we have $\delta = t^2/|D|$, which is large compared to $1/|D|$ (where D the range of the hash function in our case as it is a permutation). The reason we can tolerate this is that it is enough for $t^2/|D|$ to be much smaller than $1/|R|$ where R is the output range of f in the lemma, which is indeed the case in applications. In other words, the Crooked LHL turns out to be more tolerant than the classical one in this respect.

THE CONSTRUCTION. We now detail our construction. Let $\mathsf{LTDF} = (\mathcal{F}, \mathcal{F}')$ be an LTDF and let $\mathcal{P}\colon \mathcal{K} \times \{0,1\}^k \to \{0,1\}^k$ be an efficiently invertible family of permutations on k bits. Define the associated deterministic encryption scheme $\Pi[\mathsf{LTDF}, \mathcal{P}] = (\mathcal{DK}, \mathcal{DE}, \mathcal{DD})$ with plaintext-space $\mathrm{PtSp} = \{0,1\}^k$ via

Algorithm $\mathcal{DK}(1^k)$:	**Algorithm** $\mathcal{DE}((f, K), x)$:	**Algorithm** $\mathcal{DD}((sk, f^{-1}), c)$:
$(f, f^{-1}) \xleftarrow{\$} \mathcal{F}(1^k)$; $K \xleftarrow{\$} \mathcal{K}$	$c \leftarrow f(\mathcal{P}(K, x))$	$x \leftarrow f^{-1}(\mathcal{P}^{-1}(K, c))$
Return $((f, K), (f^{-1}, K))$	Return c	Return x

We have the following result.

Theorem 6.4 Suppose LTDF is a lossy trapdoor function on $\{0,1\}^n$ with residual leakage s, and let $q, \epsilon > 0$. Suppose $\mathcal{P}$ is a $2q$-wise δ-dependent permutation on $\{0,1\}^n$ for $\delta = t^2/2^n$. Then for any q-message IND adversary $B \in \mathbb{D}_{\mathbb{M}^{q,\mu}}$ with min-entropy $\mu \geq qs + 2\log q + \log(1/\epsilon) + 5$, there is a LTDF distinguisher D such that for all $k \in \mathbb{N}$

$$\mathbf{Adv}^{\mathrm{ind}}_{\Pi[\mathsf{LTDF},\mathcal{P}],B}(k) \quad\leq\quad \mathbf{Adv}^{\mathrm{ltdf}}_{\mathsf{LTDF},D}(k) + \epsilon \; .$$

Furthermore, the running-time of D is the time to run B.

Proof: The first step in the proof is to switch the HCF experiment to execute not $(f, f^{-1}) \xleftarrow{\$} \mathcal{F}(1^k)$ but $f' \leftarrow \mathcal{F}'(1^k)$. We can conclude by applying Lemma 6.3 with $t = 2q$ and $\mathcal{H} = \mathcal{P}$. ▮

An efficiently invertible $2q$-wise δ-dependent permutation on $\{0,1\}^n$ for $\delta = t^2/2^n$ can be obtained from [30] using key length $nt + \log(1/\delta) = n(t+1) - 2t$.

Now, combining Theorem 6.4 with Theorem 3.1 and Lemma 5.5 (extended to message vectors rather than single-input distributions) gives us bounded multi-message PRIV (rather than IND) security for any distribution on message vectors of size q with sufficient entropy. We make explicit the following corollary.

Corollary 6.5 Suppose LTDF is a lossy trapdoor function on $\{0,1\}^n$ with residual leakage s. Then we obtain a q-bounded multi-message PRIV secure DE scheme for the class of distributions on $\{0,1\}^n$ with min-entropy $\mu \geq qs + 2\log q + 2\log(1/\epsilon) + 7$ for negligible ϵ.

Comparing to Proposition 6.1, we see that we have dropped the r in the entropy bound (indeed, there is no hardcore function here). This translates to savings on the input entropy or lossiness requirement on the trapdoor function. Namely, while we still need to lose a $1 - O(1/q)$ fraction of the input, we get rid of the factor 2 on q. We also note that we can prove that the optimized scheme meets our notion of unbounded multi-message PRIV security on q-block sources of the same entropy directly by using our precise definitional equivalence, as follows. First, its IND security on q-block sources follows by extending Lemma 3.10 to q-block sources by a hybrid argument as in the case of the original LHL [51]. Then, its PRIV security on q-block sources (of 2 bits greater entropy) follows by Theorem 3.1 after extending Lemma 5.5 to show that a 2-induced distribution of a q-block source with min-entropy μ is a q-block source with min-entropy $\mu - 2$.

Discussion. We believe that q-bounded security better illuminates the technical gap between achieving unbounded security for DE in the RO and standard models. In particular, note that for an unbounded number of arbitrarily correlated messages the information-theoretic approach we use to achieve q-bounded multi-message security breaks down (since there would not be enough randomness in the inputs to extract). Indeed, a subsequent result of Wichs [49] shows that the q-bounded security we achieve in the standard model is in some sense optimal, at least using black-box reductions.

Acknowledgements

The authors are grateful to Mihir Bellare, Alexandra Boldyreva, Kai-Min Chung, Sebastian Faust, Marc Fischlin, Serge Fehr, Péter Gács, Bhavana Kanukurthi, Fenghao Liu, Payman Mohassel, Krzysztof Pietrzak, Adam Smith, Ramarathnam Venkatesan, and Hoeteck Wee for helpful discussions, improvements to our analysis, and useful references. The work was supported, in part, by National Science Foundation awards 0546614, 0831281, 1012910, and 1012798. The work of A.O. was additionally supported by NSF CNS-0915361, CNS-0952692, NSF CAREER award 0545659, and NSF Cyber Trust award 0831184.

References

[1] Werner Alexi, Benny Chor, Oded Goldreich, and Claus-Peter Schnorr. RSA and Rabin functions: Certain parts are as hard as the whole. *SIAM J. Comput.*, 17(2), 1988.

[2] Boaz Barak, Ronen Shaltiel, and Avi Wigderson. Computational analogues of entropy. In *11th International Conference on Random Structures and Algorithms*, pages 200–215, 2003.

[3] Mihir Bellare, Alexandra Boldyreva, and Adam O'Neill. Deterministic and efficiently searchable encryption. In *CRYPTO*, pages 535–552, 2007.

[4] Mihir Bellare, Zvika Brakerski, Moni Naor, Thomas Ristenpart, Gil Segev, Hovav Shacham, and Scott Yilek. Hedged public-key encryption: How to protect against bad randomness. In *ASIACRYPT*, pages 232–249, 2009.

[5] Mihir Bellare, Marc Fischlin, Adam O'Neill, and Thomas Ristenpart. Deterministic encryption: Definitional equivalences and constructions without random oracles. In *CRYPTO*, pages 360–378, 2008.

[6] Manuel Blum and Silvio Micali. How to generate cryptographically strong sequences of pseudo-random bits. *SIAM J. Comput.*, 13(4):850–864, 1984.

[7] Alexandra Boldyreva, Serge Fehr, and Adam O'Neill. On notions of security for deterministic encryption, and efficient constructions without random oracles. In *CRYPTO*, pages 335–359, 2008.

[8] Zvika Brakerski and Gil Segev. Better security for deterministic public-key encryption: The auxiliary-input setting. In *CRYPTO*, pages 543–560, 2011.

[9] Dario Catalano, Rosario Gennaro, and Nick Howgrave-Graham. Paillier's trapdoor function hides up to $O(n)$ bits. *J. Cryptology*, 2002.

[10] Benny Chor and Oded Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM J. Comput.*, 17(2), 1988.

[11] Kai-Min Chung, Yael Tauman Kalai, Feng-Hao Liu, and Ran Raz. Memory delegation. In *CRYPTO*, pages 151–168, 2011.

[12] Ronald Cramer, Goichiro Hanaoka, Dennis Hofheinz, Hideki Imai, Eike Kiltz, Rafael Pass, Abhi Shelat, and Vinod Vaikuntanathan. Bounded CCA2-secure encryption. In *ASIACRYPT*, pages 502–518, 2007.

[13] Alexander W. Dent, Marc Fischlin, Mark Manulis, Martijn Stam, and Dominique Schröder. Confidential signatures and deterministic signcryption. In *Public Key Cryptography*, pages 462–479, 2010.

[14] Simon Pierre Desrosiers. Entropic security in quantum cryptography. *Quantum Information Processing*, 8(4):331–345, 2009.

[15] Yevgeniy Dodis, Rosario Gennaro, Johan Håstad, Hugo Krawczyk, and Tal Rabin. Randomness extraction and key derivation using the CBC, cascade and HMAC modes. In *CRYPTO*, pages 494–510, 2004.

[16] 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(1):97–139, 2008.

[17] Yevgeniy Dodis and Adam Smith. Correcting errors without leaking partial information. In *STOC*, pages 654–663, 2005.

[18] Yevgeniy Dodis and Adam Smith. Entropic security and the encryption of high entropy messages. In *TCC*, pages 556–577, 2005.

[19] Stefan Dziembowski and Krzysztof Pietrzak. Leakage-resilient cryptography. In *FOCS*, pages 293–302, 2008.

[20] Sebastian Faust, Eike Kiltz, Krzysztof Pietrzak, and Guy N. Rothblum. Leakage-resilient signatures. In *TCC*, pages 343–360, 2010.

[21] David Mandell Freeman, Oded Goldreich, Eike Kiltz, Alon Rosen, and Gil Segev. More constructions of lossy and correlation-secure trapdoor functions. In *Public Key Cryptography*, pages 279–295, 2010.

[22] Benjamin Fuller and Leonid Reyzin. Computational entropy and information leakage. Technical report, IACR Cryptology e-Print Archive, 2012. `http://eprint.iacr.org/2012/466.pdf`.

[23] Craig Gentry and Daniel Wichs. Separating succinct non-interactive arguments from all falsifiable assumptions. *STOC. ACM, New York*, pages 99–108, 2011.

[24] Oded Goldreich and Leonid A. Levin. A hard-core predicate for all one-way functions. In *STOC*, pages 25–32, 1989.

[25] Shafi Goldwasser and Silvio Micali. Probabilistic encryption. *J. Comput. Syst. Sci.*, 28(2):270–299, 1984.

[26] Vipul Goyal, Adam O'Neill, and Vanishree Rao. Correlated-input secure hash functions. In *TCC*, 2011.

[27] Johan Håstad, Russell Impagliazzo, Leonid A. Levin, and Michael Luby. A pseudorandom generator from any one-way function. *SIAM J. Comput.*, 28(4):1364–1396, 1999.

[28] Brett Hemenway, Steve Lu, and Rafail Ostrovsky. Correlated product security from any one-way function and the new notion of decisional correlated product security. Cryptology ePrint Archive, Report 2010/100, 2010. `http://eprint.iacr.org/`.

[29] Chun-Yuan Hsiao, Chi-Jen Lu, and Leonid Reyzin. Conditional computational entropy, or toward separating pseudoentropy from compressibility. In *EUROCRYPT*, pages 169–186, 2007.

[30] Eyal Kaplan, Moni Naor, and Omer Reingold. Derandomized constructions of k-wise (almost) independent permutations. *Algorithmica*, 55(1):113–133, 2009.

[31] Eike Kiltz, Adam O'Neill, and Adam Smith. Instantiability of RSA-OAEP under chosen-plaintext attack. *IACR Cryptology ePrint Archive*, 2011:559, 2011.

[32] Eike Kiltz, Krzysztof Pietrzak, Martijn Stam, and Moti Yung. A new randomness extraction paradigm for hybrid encryption. In *EUROCRYPT*, pages 590–609, 2009.

[33] Hugo Krawczyk. Cryptographic extraction and key derivation: The HKDF scheme. In *CRYPTO*, pages 631–648, 2010.

[34] Silvio Micali and Leonid Reyzin. Physically observable cryptography (extended abstract). In *TCC*, pages 278–296, 2004.

[35] Harald Niederreiter. Knapsack-type cryptosystems and algebraic coding theory. *Problems of Control and Information Theory*, 15:367–391, 1986.

[36] N. Nisan and D. Zuckerman. Randomness is linear in space. *Journal of Computer and System Sciences*, pages 43–52, 1993.

[37] Pascal Paillier. Public-key cryptosystems based on composite degree residuosity classes. In *EUROCRYPT*, pages 223–238, 1999.

[38] Chris Peikert and Brent Waters. Lossy trapdoor functions and their applications. In *STOC*, pages 187–196, 2008.

[39] Omer Reingold, Luca Trevisan, Madhur Tulsiani, and Salil Vadhan. Dense subsets of pseudorandom sets. In *2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 76–85. IEEE, 2008.

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

[41] Ronald L. Rivest, Adi Shamir, and Leonard M. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Commun. ACM*, 21(2):120–126, 1978.

[42] Alon Rosen and Gil Segev. Chosen-ciphertext security via correlated products. *SIAM J. Comput.*, 39(7):3058–3088, 2010.

[43] Alexander Russell and Hong Wang. How to fool an unbounded adversary with a short key. *IEEE Transactions on Information Theory*, 52(3):1130–1140, 2006.

[44] Aravind Srinivasan and David Zuckerman. Computing with very weak random sources. In *FOCS*, pages 264–275, 1994.

[45] Ron Steinfeld, Josef Pieprzyk, and Huaxiong Wang. On the provable security of an efficient RSA-based pseudorandom generator. In *ASIACRYPT*, pages 194–209, 2006.

[46] Salil P. Vadhan. *Pseudorandomness*. Foundations and Trends in Theoretical Computer Science. Now Publishers, 2012. To appear.

[47] J. Von Neumann. Zur theorie der gesellschaftsspiele. *Mathematische Annalen*, 100(1):295–320, 1928.

[48] Hoeteck Wee. Dual projective hashing and its applications—lossy trapdoor functions and more. In *Eurocrypt*, 2012.

[49] Daniel Wichs. Barriers in cryptography with weak, correlated and leaky sources. Cryptology ePrint Archive, Report 2012/459, 2012. `http://eprint.iacr.org/`.

[50] Andrew Chi-Chih Yao. Theory and applications of trapdoor functions (extended abstract). In *FOCS*, pages 80–91, 1982.

[51] David Zuckerman. Simulating BPP using a general weak random source. *Algorithmica*, 16(4/5):367–391, 1996.

A Proof of Theorem 3.1

Following [5], the high-level intuition for the proof is as follows. For the given distribution $\boldsymbol{M}$ on message vectors, we first show that it suffices to consider PRIV adversaries for which A_2 outputs $(\mathbf{x}, t)$ where t is *boolean*. Now, we would like to use the fact if t is easy to guess from the encryption of $\mathbf{x}$ then the encryption of $\mathbf{x}$ conditioned on (1) the output $(\mathbf{x}, t)$ of A_2 being such that $t = 1$, or (2) the output $(\mathbf{x}, t)$ of A_2 being such that $t = 0$ are easy to distinguish; indeed, these are induced distributions of $\boldsymbol{M}$ (viewing the binary t as the random variable indicating the event E). However, one of these distributions may be hard to sample from and have low entropy. Therefore, we show it additionally suffices to consider PRIV adversaries on $\boldsymbol{M}$ for which t is not just boolean but also *balanced*, meaning the probability it is 0 or 1 is about the same. Then, we can easily sample from the above-mentioned distributions by repeatedly running A. In this section, we assume PRIV adversaries have an empty A_0 and accept 1^k as input (the "best" state is hardwired) though we describe the A_0's of some adversaries for clarity.

Reduction to the boolean case. Call a PRIV adversary A *boolean* if it outputs test strings of length 1. We first show that is suffices to consider boolean PRIV adversaries (this was previously shown in both [5] and [7]).

Proposition A.1 Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be an encryption scheme and $A \in \mathbb{A}_{\boldsymbol{M}}$ be a PRIV adversary that outputs test strings of length ℓ. Then there is a boolean PRIV adversary $B \in \mathbb{A}_{\boldsymbol{M}}$ such that

$$\mathbf{Adv}^{\mathrm{priv}}_{\Pi,A}(k) \quad \leq \quad 2 \cdot \mathbf{Adv}^{\mathrm{priv}}_{\Pi,B}(k) \; .$$

Furthermore, the running-time of B is the time to run A plus $O(\ell)$.

Proof: The proof is identical to an argument in [14] for the information-theoretic setting. Adversary B works as follows:

Algorithm $B_0(1^k)$:	**Algorithm** $B_1(r)$:	**Algorithm** $B_2(pk, \mathbf{c}, r)$:
$r \xleftarrow{\$} \{0,1\}^\ell$	$(\mathbf{x}, t) \xleftarrow{\$} A_1(1^k)$	$g \xleftarrow{\$} A_2(pk, \mathbf{c})$
Return r	Return $(\mathbf{x}, \langle t, r \rangle)$	Return $\langle g, r \rangle$

For $d \in \{0,1\}$, let E_d^A denote the event $\mathbf{Exp}^{\text{priv-d}}_{\Pi,A}(k) \Rightarrow 1$ and similarly E_d^B denote $\mathbf{Exp}^{\text{priv-d}}_{\Pi,B}(k) \Rightarrow 1$. Then

$$\begin{aligned}
\mathbf{Adv}^{\text{priv}}_{\Pi,B}(k) &= \Pr\left[\,\mathsf{E}_1^B\,\right] - \Pr\left[\,\mathsf{E}_0^B\,\right] \\
&= \left(\Pr\left[\,\mathsf{E}_1^A\,\right] + \frac{1}{2}\cdot(1-\Pr\left[\,\mathsf{E}_1^A\,\right])\right) - \left(\Pr\left[\,\mathsf{E}_0^A\,\right] + \frac{1}{2}\cdot(1-\Pr\left[\,\mathsf{E}_0^A\,\right])\right) \\
&= \frac{1}{2}\cdot(\Pr\left[\,\mathsf{E}_1^A\,\right] - \Pr\left[\,\mathsf{E}_0^A\,\right]) \\
&= \frac{1}{2}\cdot \mathbf{Adv}^{\text{priv}}_{\Pi,A}(k)
\end{aligned}$$

where in the second step we use that if $t \neq g$ then $\langle t, r\rangle = \langle g, r\rangle$ with probability $1/2$ over the choice of r. The claimed running-time of B is easy to verify. ▮

Reduction to the balanced boolean case. As in [5] the next step is to show that it in fact suffices to consider boolean PRIV adersaries that are *balanced*, meaning the probability the partial information is 1 or 0 is approximately 1/2. Namely, call a boolean PRIV adversary $A = (A_0, A_1, A_2)$ *δ-balanced* [5] if for all $b \in \{0,1\}$

$$\left|\Pr\left[\,t = b \;:\; (\mathbf{x}, t) \xleftarrow{\$} A_1(\mathit{state})\,\right] - \frac{1}{2}\right| \leq \delta$$

for all *state* output by A_0 on input 1^k.

Proposition A.2 Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be an encryption scheme and $B \in \mathbb{A}_{\boldsymbol{M}}$ be a boolean PRIV adversary. Then for any $0 \leq \delta < 1/2$ there is a δ-balanced boolean PRIV adversary $C \in \mathbb{A}_{\boldsymbol{M}}$ such that

$$\mathbf{Adv}^{\text{priv}}_{\Pi,B}(k) \;\leq\; \left(\frac{2}{\delta}+1\right)^2 \cdot \mathbf{Adv}^{\text{priv}}_{\Pi,C}(k)\;.$$

Furthermore, the running-time of C is the time to run B plus $O(1/\delta)$.

Proof: As compared to [5] we give a simplified proof due to [13] (which also leads to better concrete security), where for simplicity we assume $1/\delta$ is an integer. Adversary C works as follows:

Algorithm $C_1(1^k)$:	**Algorithm** $C_2(pk, \mathbf{c})$:
$(\mathbf{x}, t) \xleftarrow{\$} B_1(1^k)$	$g \xleftarrow{\$} B_2(pk, \mathbf{c})$
$i \xleftarrow{\$} \{1, \ldots, 2(1/\delta)+1\}$	$j \xleftarrow{\$} \{1, \ldots, 2(1/\delta)+1\}$
If $i \leq 1/\delta$ then return $(\mathbf{x}, 0)$	If $j \leq 1/\delta$ then return 0
Else if $i \leq 2(1/\delta)$ then return $(\mathbf{x}, 1)$	Else if $j \leq 2(1/\delta)$ then return 1
Else return $(\mathbf{x}, t)$	Else return g

Note that C is δ-balanced, since for all $b \in \{0,1\}$

$$\left|\Pr\left[\,t = b \;:\; (\mathbf{x}, t) \xleftarrow{\$} C_1(1^k)\,\right] - \frac{1}{2}\right| \;\leq\; \frac{1}{2/\delta+1}\;.$$

As before, for $d \in \{0,1\}$, let B_d denote the event $\mathbf{Exp}^{\text{priv-d}}_{\Pi,B}(k) \Rightarrow 1$ and similarly C_d denote $\mathbf{Exp}^{\text{priv-d}}_{\Pi,C}(k) \Rightarrow 1$. We define the event E to be the event that $i = j = 2/\delta + 1$. Then

$$\begin{aligned}
\mathbf{Adv}^{\text{priv}}_{\Pi,C}(k) &= \Pr\left[\,C_1\,\right] - \Pr\left[\,C_0\,\right] \\
&= \Pr\left[\,C_1 \mid \mathsf{E}\,\right]\Pr\left[\,\mathsf{E}\,\right] - \Pr\left[\,C_0 \mid \mathsf{E}\,\right]\Pr\left[\,\mathsf{E}\,\right] + \Pr\left[\,C_1 \mid \overline{\mathsf{E}}\,\right]\Pr\left[\,\overline{\mathsf{E}}\,\right] - \Pr\left[\,C_0 \mid \overline{\mathsf{E}}\,\right]\Pr\left[\,\overline{\mathsf{E}}\,\right] \\
&= \Pr\left[\,C_1 \mid \mathsf{E}\,\right]\Pr\left[\,\mathsf{E}\,\right] - \Pr\left[\,C_0 \mid \mathsf{E}\,\right]\Pr\left[\,\mathsf{E}\,\right] + \frac{1}{2} - \frac{1}{2} \\
&= \left(\frac{1}{2/\delta+1}\right)^2 \cdot \mathbf{Adv}^{\text{priv}}_{\Pi,B}(k)\;.
\end{aligned}$$

As before, the claimed running-time of C is easy to verify. ▮

REDUCTION TO DISTRIBUTION HIDING. Similar to [5] the final component for the proof is as follows.

Proposition A.3 Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be an encryption scheme and $C \in \mathbb{A}_{\boldsymbol{M}}$ be a δ-balanced boolean PRIV adversary. Then there is an IND adversary $D \in \mathbb{D}_{\boldsymbol{M}^*[\log(1/(1/2-\delta))]}$ such that

$$\mathbf{Adv}^{\mathrm{priv}}_{\Pi,C}(k) \quad \leq \quad \mathbf{Adv}^{\mathrm{ind}}_{\Pi,D}(k) + \left(\frac{1}{2} + \delta\right)^k .$$

In particular, D samples from message distributions that are statistically $2^{\Omega(k)}$-close to complementary $\log(1/(1/2-\delta))$-induced message distributions of C. Furthermore, the running-time of D is the time for at most k executions of C.

Proof: Adversary D works as follows.

Algorithm $D_1(b)$:	**Algorithm** $D_2(pk, \mathbf{c})$:
For $i = 1$ to k do:	$g \xleftarrow{\$} B_2(pk, \mathbf{c})$
$\quad (\mathbf{x}, t) \xleftarrow{\$} B_1(1^k)$	Return g
$\quad$ If $t = b$ then return $\mathbf{x}$	
Return $\mathbf{x}$	

For the analysis, let BAD denote the event that the final return statement is executed. Let $\mathsf{CORRECT}_D$ be the event that $b = d$ when D is executed in the PRIV experiment with Π and similarly let $\mathsf{CORRECT}_B$ denote the event that $t = g$ when B is executed in the PRIV experiment with Π. Then

$$\begin{aligned}
\mathbf{Adv}^{\mathrm{priv}}_{\Pi,D}(k) &= \Pr\left[\, \mathsf{CORRECT}_D \mid b = 1 \,\right] + \Pr\left[\, \mathsf{CORRECT}_D \mid b = 0 \,\right] \\
&\geq \left(\Pr\left[\, \mathsf{CORRECT}_D \mid b = 1 \ \wedge\ \overline{\mathsf{BAD}} \,\right] + \Pr\left[\, \mathsf{CORRECT}_D \mid b = 0 \ \wedge\ \overline{\mathsf{BAD}} \,\right]\right) \cdot \Pr\left[\, \overline{\mathsf{BAD}} \,\right] \\
&= \left(\Pr\left[\, \mathsf{CORRECT}_B \mid t = 1 \,\right] + \Pr\left[\, \mathsf{CORRECT}_B \mid t = 0 \,\right]\right) \cdot \Pr\left[\, \overline{\mathsf{BAD}} \,\right] \\
&= \mathbf{Adv}^{\mathrm{priv}}_{\Pi,B}(k) \cdot \Pr\left[\, \overline{\mathsf{BAD}} \,\right] \\
&\geq \mathbf{Adv}^{\mathrm{priv}}_{\Pi,B}(k) \left(1 - \left(\frac{1}{2} + \delta\right)^k\right) \\
&\geq \mathbf{Adv}^{\mathrm{priv}}_{\Pi,B}(k) - \left(\frac{1}{2} + \delta\right)^k ,
\end{aligned}$$

where the second-to-last line uses that B is δ-balanced. The claimed running-time of D is easy to verify. It remains to argue that $D \in \mathbb{D}_{\boldsymbol{M}^*[\log(1/(1/2-\delta))]}$. Let $\boldsymbol{M}_{D,i}$ be the message distribution sampled by D_1 on input $b = i$ for $i \in \{0,1\}$ and similarly let $\boldsymbol{M}_{C,i}$ be the message distribution sampled by C_1 when $t = i$ in its output for $i \in \{0,1\}$. Observe that $\boldsymbol{M}_{C,0}$ and $\boldsymbol{M}_{C,1}$ are complementary $\log(1/(1/2-\delta))$-induced distributions of the message distribution of C, with corresponding events $t = 0$ and $t = 1$ respectively. Furthermore, we have $\boldsymbol{M}_{D,i} \mid \overline{\mathsf{BAD}} = \boldsymbol{M}_{C,i}$ for $i \in \{0,1\}$. Since $\Pr\left[\, \mathsf{BAD} \,\right] \leq (1/2 + \delta)^k$, it follows that $\boldsymbol{M}_{D,i} \mid \overline{\mathsf{BAD}}$ is statistically $2^{-\Omega(k)}$-close to $\boldsymbol{M}_{C,i}$ for $i \in \{0,1\}$, which concludes the proof.[12] ▮

Theorem 3.1 follows by combining Propositions A.1, A.2, and A.3 with $\delta = 1/4$. ▮

[12] Note that as compared to [5] our approach avoids having to analyze the min-entropy of D, which is more involved.

B Comparison to other Computational Entropy Leakage Lemmas

Previous works have considered the question of measuring conditional computational entropy under a wide array of applications and settings. Dziembowski and Pietrzak [19] show that the output of a pseudorandom generator still has entropy conditioned on functions of the seed:

Lemma B.1 [19, Lemma 3] Let $prg : \{0,1\}^n \rightarrow \{0,1\}^\nu$ and $f : \{0,1\}^n :\rightarrow \{0,1\}^\lambda$ (where $1 \leq \lambda < n < \nu$) be any functions. If prg is a (ϵ_{prg}, s)-secure pseudorandom-generator, then for any $\epsilon_1, \epsilon_2, \Delta > 0$ satisfying $\epsilon_{prg} \leq \epsilon_1\epsilon_2/2^\lambda - 2^{-\Delta}$, we have with $X \sim U_n$,

$$\Pr_{y:=f(X)}[H^{\mathtt{Metric}^*}_{\epsilon_1,s'}(prg(X)|f(X) = y) \geq \nu - \Delta] \geq 1 - \epsilon_2 \tag{18}$$

where $s' \approx s$.

Our results improve the parameters and simplify the exposition. Our result considers any random variables X, Y (not just pseudorandom X) and gives simpler statements, such as Theorem 3.8. To make the quantitative comparison, we present the following alternative formulation of our result, in the style of [19, Lemma 3]:

Lemma B.2 Let X, Y be discrete random variables with $|Y| \leq 2^\lambda$ and $H^{\mathtt{Metric}^*}_{\epsilon_{ent},s}(X) \geq \nu$, then for any $\epsilon_1, \epsilon_2, \Delta > 0$ satisfying $\epsilon_{ent} \leq \epsilon_1\epsilon_2/2^\lambda$ and $2^{-\Delta} \leq \epsilon_2/2^\lambda$,

$$\Pr_{y \in Y}[H^{\mathtt{Metric}^*}_{\epsilon_1,s'}(X|Y = y) \geq \nu - \Delta)] \geq 1 - \epsilon_2$$

where $s' \approx s$.

To compare the bounds, observe that we have removed ϵ_1 from $2^{-\Delta}$, because the constraint $\epsilon_{prg} \leq \epsilon_1\epsilon_2/2^\lambda - 2^{-\Delta}$ implies that $\epsilon_{prg} \leq \epsilon_1\epsilon_2/2^\lambda$ and $\epsilon_1\epsilon_2/2^\lambda \geq 2^{-\Delta}$.

The question has also been considered by [39] in the language of the dense model theorem. Their main result, restated in our language is:

Lemma B.3 [39, Theorem 1.3] Let X, Y be discrete random variables. Then

$$H^{\mathtt{HILL}}_{\epsilon',s'}(X|Y = y) \geq H^{\mathtt{HILL}}_{\epsilon,s}(X) - \log 1/P_Y(y) \tag{19}$$

where $\epsilon' = \Omega(\epsilon/P_Y(y))$, and $s' = s/\mathtt{poly}(P_Y(y)/\epsilon, \log 1/P_Y(y))$

Note that the quantity loss is the same as in Lemma 3.2; however, the losses in the circuit size and distinguishing advantage are different, because Lemma 3.2 separates the conditioning step and the conversion back to HILL entropy. This separation allows us set conversion parameters separately (which is needed when ϵ is smaller than $1/s$). It also allows paying for the conversion step only once in case of repeated leakage, enabling the proof of a limited chain rule for repeated conditioning (see [22, Theorem 3.6]).

Recent work concurrent with ours [23, 11] has shown results on information leakage when the starting distribution is already conditional. This is a significantly harder as the auxiliary information may shape the original distribution or its condition. Both works are able to achieve this "chain-rule" but must introduce significant restrictions. Since these works are both average-case formulations, we first present an average case formulation of Corollary 3.9:

Corollary B.4 Let X, Y be discrete random variables over χ_1, χ_2 respectively. Then

$$H^{\mathtt{HILL}}_{\epsilon',s'}(X|Y) \geq H^{\mathtt{HILL}}_{\epsilon,s}(X) - \log |Y|$$

where $\epsilon' = \epsilon|Y| + \sqrt[3]{\frac{\log|\chi_1|}{s}}$, $s' = \Omega(\sqrt[3]{\frac{s}{\log|\chi_1|}})$.

This corollary follows by Theorem 3.8, Theorem 2.10, and setting $\epsilon_{\mathtt{HILL}} = \sqrt[3]{\frac{\log|\chi_1|}{s}}$.

Gentry and Wichs consider indistinguishability with auxiliary information in their work on succinct argument systems [23, Lemma 3.1]. Their result is below (restated in our language):

Lemma B.5 [23, Lemma 3.1] Let X, Y, Z be discrete random variables with $H_\infty(Z) \geq k$ and Y ranges over $\{0,1\}^\lambda$. If $\forall D \in \mathcal{D}_s^{rand,\{0,1\}}$, $\delta^D(X,Z) \leq \epsilon$, then $\exists Y'$ such that $\forall \tilde{D} \in \mathcal{D}_{s'}^{rand,\{0,1\}}$, $\delta^{\tilde{D}}((X,Y),(Z,Y')) \leq \epsilon'$ where $\epsilon' = 2\epsilon$ and $s' = s \cdot \mathtt{poly}(\epsilon/|Y|)$.

This lemma is related to entropy as follows: X has HILL entropy k, and it can be said that, since (X,Y) is indistinguishable from (Z,Y'), that computational entropy of $X \mid Y$ is at least $\tilde{\mathrm{H}}_\infty(Z \mid Y')$, which is at least $k - \lambda$ by Lemma 2.1. Note, however, that this Lemma requires a different definition of entropy from ours, in which the condition itself may also be replaced. It is unclear the implications of this change and where it would better or worse than conditional HILL entropy. The advantage of this lemma is that it handles the case when X is already a conditional distribution (we can only handle this when the conditional distribution decomposes "nicely" into distributions for each value of the condition [22, Theorem 3.6]). The disadvantage, however, is that the lemma inherently talks about the average case Y and not a single event y. For our application in the current paper, we need to condition on a particular event y and not the distribution of events.

Chung et al. in their work on memory delegation need indistinguishability in the presence of a single bit of auxiliary information. They formulate the problem in the asymptotic setting:

Lemma B.6 [11, Lemma 16] Let k be a security parameter and n, l, t be any parameters such that $n \leq \mathtt{poly}(k), l = O(\log k)$, and $t = \omega(\log k)$. Let (X, C) be a joint distribution over $\{0,1\}^* \times \{0,1\}^*$ of $\mathtt{poly}(k)$ length. If $H^{\mathtt{HILL}}(X|C) \geq n$ w.r.t. samplable distributions, then for any distribution $B = B(X,C)$ over $\{0,1\}^l$, we have

$$H^{\mathtt{HILL}}(X|C,B) \geq n - t.$$

It is important to note that in this lemma, the "conditional HILL entropy" is different from our notion: it means indistinguishability against distributions of *worst-case* conditional min-entropy, whereas here we define conditional HILL entropy as indistinguishability against distributions of *average* min-entropy (see the precise definitions in Section 2). In addition, this lemma imposes a samplability condition that we do not.

Thus, removing limitations on the chain rule remains an interesting open problem.

C Support Preserving Extension to Theorem 2.6

Theorem C.1 Let X be a discrete distribution over a s finite set χ. For every $\epsilon, \epsilon_{\mathtt{HILL}} > 0, \epsilon' \geq \epsilon + \epsilon_{\mathtt{HILL}}, k$ and s, if $H^{\mathtt{Metric}^*}_{\epsilon,s}(X) \geq k$ then $H^{\mathtt{HILL}}_{\epsilon', s_{\mathtt{HILL}}}(X) \geq k$ where $s_{\mathtt{HILL}} = \Omega(\epsilon^2_{\mathtt{HILL}} s / \log|\chi|)$. The reduction is support preserving.[13]

Proof: This proof closely follows the proof from [2, Theorem 5.2]. For a set X (such as a set of distinguishers or distributions with a certain property) we will use $\hat{X}$ to represent the set of distributions over that set. A game is simply a function from finite sets A, B to an outcome space R, that is, $g : A \times B \to R$. We similarly define $\hat{g} : \hat{A} \times \hat{B} \to R$ as a function from distributions, $a \leftarrow \hat{A}, b \leftarrow \hat{B}$ to outcome space R.

We let ζ be the support of random variable Z that is indistinguishable from X. The proof proceeds similarly to the case where $\zeta = \chi$ [2, Theorem 5.2]. We will assume that $H^{\mathtt{HILL}}_{\epsilon', s_{\mathtt{HILL}}}(X) < k$ and seek to

[13] "Support preserving" here means the following. The definition of $\mathtt{Metric}^*$ entropy of X calls for an indistinguishable from X distribution Z_D with true entropy for every distinguisher $D \in \mathcal{D}_s^{det,[0,1]}$. The definition of $\mathtt{HILL}$ entropy of X calls for a single distribution Z that is indistinguishable from X. Support-preserving means that support of Z is no greater than the union of supports of Z_D.

show that $H^{\mathtt{Metric}^*}_{\epsilon,s}(X) < k$. Assume that $H^{\mathtt{HILL}}_{\epsilon',s_{\mathtt{HILL}}}(X) < k$. That is, $\forall Z'' \subset \zeta$ with $H_\infty(Z'') \geq k$ there exists $D \in \mathcal{D}^{det,\{0,1\}}_{s_{\mathtt{HILL}}}$ such that $\delta^D(X, Z'') \geq \epsilon'$. Recall that the definition for $H^{\mathtt{HILL}}$ is for randomized $\{0,1\}$ distinguishers, however as noted after Definition 2.4, drawing from deterministic $\{0,1\}$ distinguishers is essentially equivalent (by selecting the "best" randomness). We begin by showing a change of quantifiers similar to [2, Lemma 5.3]:

Claim C.2 Let X be a distribution over χ. Let $\mathcal{C}$ be a class that is closed under complement. If for every $Z'' \subset \zeta$ with $\mathrm{H}_\infty(Z'') \geq k$ there exists a $D \in \mathcal{C}$ such that $\delta^D(X, Z'') \geq \epsilon'$, then there is a distribution $\hat{D}$ over $\mathcal{C}$ such that for every $Z' \subset \zeta$ with $\mathrm{H}_\infty(Z') \geq k$

$$\mathop{\mathbb{E}}_{D \leftarrow \hat{D}} \left[D(X) - D(Z') \right] \geq \epsilon'$$

Proof: We use the minimax theorem of [47]:

Theorem C.3 ([47]) For every game g there is a value v such that

$$\max_{\hat{a} \in \hat{A}} \min_{b \in B} \hat{g}(\hat{a}, b) = v = \min_{\hat{b} \in \hat{B}} \max_{a \in A} \hat{g}(a, \hat{b})$$

We will use the minimax theorem to change the order of quantifiers. We define our game as follows: let $A \stackrel{def}{=} \mathcal{C}$, let $B \stackrel{def}{=} \{Z'' | \mathrm{H}_\infty(Z'') \geq k, Z'' \subseteq \zeta\}$ and let $g(D, Z) \stackrel{def}{=} [D(X) - D(Z)]$. The convex combination of distributions with min-entropy k has min-entropy at least k (this is easily seen by considering the maximum probability event), thus $\forall \hat{b} \in \hat{B}, \mathrm{H}_\infty(\hat{b}) \geq k$. Thus, both B and $\hat{B}$ are the sets of all distributions with min-entropy at least k. Then by assumption, $\forall Z'' \in \hat{B}, \exists D \in A$ such that $|D(X) - D(Z'')| \geq \epsilon'$. Because $\mathcal{C}$ is closed under complement, there must $\exists D \in A$ such that $D(X) - D(Z'') \geq \epsilon'$. Now we know that $\min_{\hat{b} \in \hat{B}} \max_{a \in A} \hat{g}(a, \hat{b}) = \min_{Z'' \in B} \max_{D \in \mathcal{C}} (D(X) - D(Z'')) \geq \epsilon'$. Then by Theorem C.3: $\max_{\hat{a} \in \hat{A}} \min_{b \in B} \hat{g}(\hat{a}, b) \geq \epsilon'$. That is, there is a distribution $\hat{D}$ over the class of distinguishers $\mathcal{C}$ such that for every $Z'' \in B$, $\mathbb{E}_{D \leftarrow \hat{D}} D(X) - D(Z'') \geq \epsilon'$. This completes the proof of the claim. ▮

Our remaining task is to approximate a distribution of distinguishers $\hat{D}$ by several distinguishers in its support where the resulting distinguisher still has advantage at least ϵ. Define $n = \log |\chi|$ and choose $t = 8n/\epsilon^2_{\mathtt{HILL}}$ samples $D_1, ..., D_t$ from $\hat{D}$ and define

$$D'_{D_1,...,D_t}(x) = 1/t \sum_{i=1}^{t} D_i(x)$$

Then by Chernoff's inequality

$$\forall x \in \chi, \Pr_{D_1,...,D_t \leftarrow \hat{D}} \left[\left| D'_{D_1,...,D_t}(x) - \mathop{\mathbb{E}}_{D \leftarrow \hat{D}}(x) \right| \geq \epsilon_{\mathtt{HILL}}/2 \right] < 2^{-2n}. \tag{20}$$

Claim C.4 There exists $D_1, ..., D_t$ such that $\forall x, \left| D'_{D_1,...,D_t}(x) - \mathop{\mathbb{E}}_{D \leftarrow \hat{D}} D(x) \right| \leq \epsilon_{\mathtt{HILL}}/2$.

Proof: Suppose not, that is $\forall D_1,...,D_t, \exists x' \in \chi, \left|D'_{D_1,\ldots,D_t} - \mathbb{E}_{D\leftarrow\hat{D}} D(x)\right| > \epsilon_{\mathtt{HILL}}/2$. For a particular, $D_1,...,D_t$ we denote x' as $x'_{D_1,\ldots,D_t}$. This implies that,

$$\Pr_{D_1,\ldots,D_t,x}\left[\left|D'_{D_1,\ldots,D_t}(x) - \mathop{\mathbb{E}}_{D\leftarrow\hat{D}}(x)\right| > \epsilon_{\mathtt{HILL}}/2\right] \geq$$

$$\Pr_{D_1,\ldots,D_t}\left[\left|D'_{D_1,\ldots,D_t}(x) - \mathop{\mathbb{E}}_{D\leftarrow\hat{D}}(x)\right| > \epsilon_{\mathtt{HILL}}/2 | X = x'_{D_1,\ldots,D_t}\right] \Pr[X = x'_{D_1,\ldots,D_t}] \geq$$

$$\Pr[X = x'_{D_1,\ldots,D_t}] \geq 1/|\chi| = 2^{-n}.$$

However, this implies that $\exists x \in \chi$ such that $\Pr_{D_1,\ldots,D_t}\left[\left|D'_{D_1,\ldots,D_t}(x) - \mathbb{E}_{D\leftarrow\hat{D}}(x)\right| > \epsilon_{\mathtt{HILL}}/2\right] \geq 1/2^{2n}$ (since there are 2^n possible x). This is a contradiction of Equation 20. ▌

Fix one such $D_1,...,D_t$. Because it holds for every x, it also holds for all distributions, and thus for the distribution X, $\left|D'_{D_1,\ldots,D_t}(X) - \mathop{\mathbb{E}}_{D\leftarrow\hat{D}} D(X)\right| \leq \epsilon_{\mathtt{HILL}}/2$, and for every distribution $Z' \subset \zeta$, $\left|D'_{D_1,\ldots,D_t}(Z') - \mathop{\mathbb{E}}_{D\leftarrow\hat{D}} D(Z')\right| \leq \epsilon_{\mathtt{HILL}}/2$. Therefore, subtracting these inequalities from the inequality of Claim C.2, and recalling that $|a| \geq |a+b+c| - |b| - |c|$, we get

$$D'_{D_1,\ldots,D_t}(X) - D'_{D_1,\ldots,D_t}(Z') \geq$$

$$\left|\mathop{\mathbb{E}}_{D\leftarrow\hat{D}}\left[D(X) - D(Z')\right]\right| - \left|D'_{D_1,\ldots,D_t}(X) - \mathop{\mathbb{E}}_{D\leftarrow\hat{D}} D(X)\right| - \left|D'_{D_1,\ldots,D_t}(Z') - \mathop{\mathbb{E}}_{D\leftarrow\hat{D}} D(Z')\right| \geq$$

$$\epsilon' - \epsilon_{\mathtt{HILL}}/2 - \epsilon_{\mathtt{HILL}}/2 \geq \qquad \epsilon .$$

Lastly, $D'_{D_1,\ldots,D_t}$ is of size

$$O(\log|\chi| s_{\mathtt{HILL}}/\epsilon^2_{\mathtt{HILL}}) = s$$

This completes the proof. ▌

D Proof of Lemma 3.10

For random variables X and Y, we define te *square of the 2-distance* between X and Y as $D(X,Y) = \sum_x \big(P_X(x) - P_Y(x)\big)^2$. Writing $\mathbf{E}_k$ for the expectation over the choice of k according to the distribution of K, it follows that

$$\begin{aligned}
\Delta\big((K, f(\mathcal{H}(K,\mathbf{X}))), (K, f(\mathbf{U}))\big) &= \mathbf{E}_k\big[\Delta\big(f(\mathcal{H}(k,\mathbf{X})), f(\mathbf{U})\big)\big] \\
&\leq \frac{1}{2}\mathbf{E}_k\left[\sqrt{|S|^t \cdot D\big(f(\mathcal{H}(k,\mathbf{X})), f(\mathbf{U})\big)}\right] \\
&\leq \frac{1}{2}\sqrt{|S|^t \cdot \mathbf{E}_k\big[D\big(f(\mathcal{H}(k,\mathbf{X}))), f(\mathbf{U})\big)\big]}
\end{aligned}$$

where the first inequality is by Cauchy-Schwarz and the second inequality is due to Jensen's inequality. We will show that

$$\mathbf{E}_k\big[D\big(f(\mathcal{H}(k,\mathbf{X})), f(\mathbf{U})\big)\big] \leq t^2 2^{-\mu} + 6t^2 2^{-r} + 3\delta ,$$

which completes the proof. Write $\mathbf{Y} = \mathcal{H}(k,\mathbf{X})$ for an arbitrary but fixed k. Then

$$\begin{aligned}
D\big(f(\mathbf{Y}), f(\mathbf{U})\big)) &= \sum_{\mathbf{s}} \big(P_{f(\mathbf{Y})}(\mathbf{s}) - P_{f(\mathbf{U})}(\mathbf{s})\big)^2 \\
&= \sum_{\mathbf{s}} P_{f(\mathbf{Y})}(\mathbf{s})^2 - 2\sum_{\mathbf{s}} P_{f(\mathbf{Y})}(\mathbf{s}) P_{f(\mathbf{U})}(\mathbf{s}) + \mathrm{Col}(f(\mathbf{U})) .
\end{aligned}$$

For a set $Z \subseteq R^t$, define $\delta_{\mathbf{r},Z}$ to be 1 if $\mathbf{r} \in Z$ and else 0. For $\mathbf{s} \in S^t$ we can write $P_{f(\mathbf{Y})}(\mathbf{s}) = \sum_{\mathbf{x}} P_{\mathbf{X}}(\mathbf{x})\delta_{\mathcal{H}(k,\mathbf{x}),f^{-1}(\mathbf{s})}$ and thus

$$\begin{aligned}\sum_{\mathbf{s}} P_{f(\mathbf{Y})}(\mathbf{s})^2 &= \sum_{\mathbf{s}}\Big(\sum_{\mathbf{x}} P_{\mathbf{X}}(\mathbf{x})\delta_{\mathcal{H}(k,\mathbf{x}),f^{-1}(\mathbf{s})}\Big)\Big(\sum_{\mathbf{x}'} P_{\mathbf{X}}(\mathbf{x}')\delta_{\mathcal{H}(k,\mathbf{x}'),f^{-1}(\mathbf{s})}\Big)\\ &= \sum_{\mathbf{s},\mathbf{x},\mathbf{x}'} P_{\mathbf{X}}(\mathbf{x})P_{\mathbf{X}}(\mathbf{x}')\delta_{\mathcal{H}(k,\mathbf{x}),f^{-1}(\mathbf{s})}\delta_{\mathcal{H}(k,\mathbf{x}'),f^{-1}(\mathbf{s})}\ ,\end{aligned}$$

so that

$$\begin{aligned}\mathbf{E}_k\Big[\sum_{\mathbf{s}} P_{f(\mathbf{Y})}(\mathbf{s})^2\Big] &= \sum_{\mathbf{s}}\sum_{\mathbf{x},\mathbf{x}'} P_{\mathbf{X}}(\mathbf{x})P_{\mathbf{X}}(\mathbf{x}')\mathbf{E}_k[\delta_{\mathcal{H}(k,\mathbf{x}),f^{-1}(\mathbf{s})}\delta_{\mathcal{H}(k,\mathbf{x}'),f^{-1}(\mathbf{s})}]\\ &= \sum_{\mathbf{s}}\sum_{\exists i,j,\ \mathbf{x}[i]=\mathbf{x}'[j]} P_{\mathbf{X}}(\mathbf{x})P_{\mathbf{X}}(\mathbf{x}')\\ &\quad+\sum_{\mathbf{s}}\sum_{\forall i,j,\ \mathbf{x}[i]\neq\mathbf{x}'[j]} P_{\mathbf{X}}(\mathbf{x})P_{\mathbf{X}}(\mathbf{x}')\mathbf{E}_k[\delta_{\mathcal{H}(k,\mathbf{x}),f^{-1}(\mathbf{s})}\delta_{\mathcal{H}(k,\mathbf{x}'),f^{-1}(\mathbf{s})}]\\ &\leq t^2 2^{-\mu} + \mathrm{Col}(f(U)) + t^2 2^{-r} + \delta\end{aligned}$$

where the first term is by a union bound over all $1 \leq i,j \leq t$ and for the remaining terms we use the δ-almost $2t$-wise independence of $\mathcal{H}$ and note that

$$E_k[\delta_{\mathcal{H}(k,\mathbf{x}),f^{-1}(\mathbf{s})}\delta_{\mathcal{H}(k,\mathbf{x}'),f^{-1}(\mathbf{s})}] = \Pr\big[\, f(\mathcal{H}(K,\mathbf{x})) = f(\mathcal{H}(K,\mathbf{x}'))\,\big]\ .$$

Similarly,

$$\begin{aligned}\sum_{\mathbf{s}} P_{f(\mathbf{Y})}(\mathbf{s})P_{f(\mathbf{U})}(\mathbf{s}) &= \sum_{\mathbf{s}}\Big(\sum_{\mathbf{x}} P_{\mathbf{X}}(\mathbf{x})\delta_{\mathcal{H}(k,\mathbf{x}),f^{-1}(\mathbf{s})}\Big)\Big(\frac{1}{|R|}\sum_{\mathbf{u}}\delta_{\mathbf{u},f^{-1}(\mathbf{s})}\Big)\\ &= \frac{1}{|R|}\sum_{\mathbf{s}}\sum_{\mathbf{u},\mathbf{x}} P_{\mathbf{X}}(\mathbf{x})\delta_{\mathcal{H}(k,\mathbf{x}),f^{-1}(\mathbf{s})}\delta_{\mathbf{u},f^{-1}(\mathbf{s})}\end{aligned}$$

so that

$$\begin{aligned}\mathbf{E}_k\Big[\sum_{\mathbf{s}} P_{f(\mathbf{Y})}(\mathbf{s})P_{f(\mathbf{U})}(\mathbf{s})\Big] &= \frac{1}{|R|}\sum_{\mathbf{s}}\sum_{\mathbf{u},\mathbf{x}} P_{\mathbf{X}}(\mathbf{x})\,\mathbf{E}_k[\delta_{\mathcal{H}(k,\mathbf{x}),f^{-1}(\mathbf{s})}\delta_{\mathbf{u},f^{-1}(\mathbf{s})}]\\ &\geq \mathrm{Col}(f(\mathbf{U})) - \delta\end{aligned}$$

using δ-almost t-wise independence of $\mathcal{H}$. By combining the above, it follows that

$$\mathbf{E}_k\big[D\big(f(\mathbf{Y}),f(\mathbf{U})\big)\big] \;\leq\; t^2 2^{-\mu} + 3\delta$$

which was to be shown. ▌