A Hardcore Lemma for Computational Indistinguishability: Security Amplification for Arbitrarily Weak PRGs with Optimal Stretch

Ueli Maurer and Stefano Tessaro

Department of Computer Science, ETH Zurich, 8092 Zurich, Switzerland
{maurer,tessaros}@inf.ethz.ch

Abstract. It is well known that two random variables X and Y with the same range can be viewed as being equal (in a well-defined sense) with probability $1 - d(X, Y)$, where $d(X, Y)$ is their statistical distance, which in turn is equal to the best distinguishing advantage for X and Y. In other words, if the best distinguishing advantage for X and Y is ϵ, then with probability $1 - \epsilon$ they are completely indistinguishable. This statement, which can be seen as an information-theoretic version of a hardcore lemma, is for example very useful for proving indistinguishability amplification results.

In this paper we prove the computational version of such a hardcore lemma, thereby extending the concept of hardcore sets from the context of computational *hardness* to the context of computational *indistinguishability*. This paradigm promises to have many applications in cryptography and complexity theory. It is proven both in a non-uniform and a uniform setting.

For example, for a weak pseudorandom generator (PRG) for which the (computational) distinguishing advantage is known to be bounded by ϵ (e.g. $\epsilon = \frac{1}{2}$), one can define an event on the seed of the PRG which has probability at least $1 - \epsilon$ and such that, conditioned on the event, the output of the PRG is essentially indistinguishable from a string with almost maximal min-entropy, namely $\log(1/(1 - \epsilon))$ less than its length. As an application, we show an optimally efficient construction for converting a weak PRG for any $\epsilon < 1$ into a strong PRG by proving that the intuitive construction of applying an extractor to an appropriate number of independent weak PRG outputs yields a strong PRG. This improves strongly over the best previous construction for security amplification of PRGs which does not work for $\epsilon \geq \frac{1}{2}$ and requires the seed of the constructed strong PRG to be very long.

1 Introduction

1.1 (Weak) Pseudorandomness

Randomness is a central resource in cryptography. In many applications, true randomness must be replaced by *pseudorandomness*, for example when it needs to be reproduced at a second location and one can only afford to transmit a short

value to be used as the seed of a so-called *pseudorandom generator* (PRG). An example are cryptographic applications where a key agreement protocol yields only a short key. More generally, PRGs are a central building block in cryptographic protocols and are used in different applications where a random functionality (e.g. a uniform random function) must be realized from a short secret key.

The concept of a PRG was first proposed by Blum and Micali [2], initiating a large body of literature dealing with various aspects of pseudorandomness: More formally, a random variable X is said to be pseudorandom if it is *computationally* indistinguishable from a uniformly distributed random variable U with the same range, i.e., no computationally bounded (i.e., polynomial time) distinguisher can tell X and U apart with better than negligible advantage. In particular, a PRG $G : \{0,1\}^k \to \{0,1\}^\ell$ (for $\ell > k$) extends a uniform random string U_k of length k into a pseudorandom string $G(U_k)$ of length ℓ.

Computational infeasibility is at the core of cryptographic security. In contrast to cryptographic primitives (like a one-way function f) assuring that a certain value (e.g. the input of f) cannot efficiently be computed, the notion of computational indistinguishability is substantially more involved. It is hence not a surprise that all constructions (cf. e.g. [6, 5, 9]) of a PRG from an arbitrary[1] one-way function f are too inefficient (in terms of the number of calls to f) to be of any practical use.

Therefore, it appears much more difficult to propose a cryptographic function that can be believed to be a PRG than one that can be believed to be a one-way function. As a consequence, a prudent approach in cryptography is to make weaker assumptions about a concrete proposal for a PRG G. One possible way[2] to achieve this is by considering a so-called *ϵ-pseudorandom generator (ϵ-PRG)*, where the best distinguishing advantage of an efficient distinguisher is not necessarily negligible, but instead bounded by some noticeable quantity ϵ, such as a constant (e.g. $\epsilon = 0.75$), or even a function in the security parameter k mildly converging to 1 (e.g. $1 - \frac{1}{p(k)}$ for some polynomial p).[3]

1.2 Security Amplification of PRGs

Security Amplification. In order to deploy some ϵ-PRG within a particular cryptographic application, we need to find an efficient construction transforming it into a fully secure PRG. This is an instance of the general problem of *security amplification*, which was first considered by Yao [17] in the context of one-way

[1] i.e. without any particular assumption on the combinatorial structure of the function

[2] An alternative approach to modeling a weak PRG is to assume its output to be computationally indistinguishable from a random variable with only moderate min-entropy. However, this approach does not capture certain failure types, such as a function G that with some substantial probability may output a constant value. In contrast, the notion of an ϵ-PRG captures this case. One of the contributions of this paper is to show a tight relation between these two approaches.

[3] An ϵ-PRG $G : \{0,1\}^k \to \{0,1\}^\ell$ is only interesting in the case $\ell > k + \log\left(\frac{1}{1-\epsilon}\right)$, as otherwise an unconditionally secure ϵ-PRG is given by the mapping $x \mapsto x \| 0^{\log(\frac{1}{1-\epsilon})}$.

functions, and has subsequently been followed by a prolific line of research considering a wide range of other cryptographic primitives.

Previous Work. The *only* known security amplification result for PRGs considers the construction $\mathsf{SUM}^G : \{0,1\}^{mk} \to \ell$ (for any $m > 1$) which outputs

$$\mathsf{SUM}^G(x) := G(x_1) \oplus \cdots \oplus G(x_m).$$

for all inputs $x = x_1 \| \ldots \| x_m \in \{0,1\}^{km}$ (with $x_1, \ldots, x_m \in \{0,1\}^k$). As pointed out in [14], Yao's XOR-lemma [17, 4] yields a direct proof of security amplification for the construction SUM, and an improved bound can be obtained using the tools from [14]. (An independent proof with a weaker bound was also given in [3].) Namely, one can show that if G is an ϵ-PRG, then SUM^G is a $(2^{m-1}\epsilon^m + \nu)$-PRG, where ν is a negligible function. Also, the result extends to the case where $\oplus$ is replaced by any quasi-group operation $\star$.

However, this construction has two major disadvantages: First, security amplification is inherently limited to the case $\epsilon < \frac{1}{2}$. For instance, the security of a PRG with a very large stretch and with one constant output bit is not amplified by the SUM construction, even if all other output bits are pseudorandom. Second, the construction is expanding only when $\ell > k \cdot m$. Note that this issue cannot be overcome by first extending the output size of the weak PRG, due to the high security loss in the extension which would yield an ϵ'-PRG with ϵ' close to one.

Our Construction. In this paper, we provide the first solution which amplifies the security of an ϵ-PRG $G : \{0,1\}^k \to \{0,1\}^\ell$ for any $\epsilon < 1$. Our construction, called *concatenate and extract* (CaE), takes input $x = x_1 \| \ldots \| x_m \| r$, where $x_1, \ldots, x_m \in \{0,1\}^k$ and $r \in \{0,1\}^d$, and outputs

$$\mathsf{CaE}^G(x) := \mathsf{Ext}(G(x_1) \,\|\, \ldots \,\|\, G(x_m), r) \,\|\, r,$$

where $\mathsf{Ext} : \{0,1\}^{m\ell} \times \{0,1\}^d \to \{0,1\}^n$ is a sufficiently good strong randomness extractor. In particular, a good instantiation (for instance using two-universal hash functions or even appropriate deterministic extractors) allows to achieve $n \approx (1-\epsilon)m \cdot \left[\ell - \log\left(\frac{1}{1-\epsilon}\right)\right]$, and we show the resulting output length $n + d$ to be optimal with respect to constructions combining m outputs of an ϵ-PRG.

We provide security proofs both in the non-uniform and in the uniform models, which follow as an application of a new characterization of computational indistinguishability that we present in this paper, and which we outline in the next section.

Finally, we point out that the idea of concatenating strings with weaker pseudorandomness guarantees and then extracting the resulting computational entropy was previously used (most notably in constructions of PRGs from one-way functions [6, 9, 5]): However, all these previous results only consider individual independent *bits* which are hard to compute (given some other part of the concatenation), whereas our result is the first to deal with the more general case of weakly pseudorandom *strings*.

1.3 A Tight Characterization of Computational Indistinguishability

Let X and Y be random variables with the same range $\mathcal{U}$. Assume that we can show that there exist events $\mathcal{A}$ and $\mathcal{B}$ defined on the choices of X and Y by some conditional probability distributions $\mathsf{P}_{\mathcal{A}|X}$ and $\mathsf{P}_{\mathcal{B}|Y}$ such that $\mathsf{P}[\mathcal{A}] \geq 1-\epsilon$, $\mathsf{P}[\mathcal{B}] \geq 1-\epsilon$, and X and Y are identically distributed when conditioned on $\mathcal{A}$ and $\mathcal{B}$, respectively. Then this implies that the advantage $\Delta^D(X,Y) := |\mathsf{P}[D(X)=1] - \mathsf{P}[D(Y)=1]|$ is upper bounded by ϵ for *every* distinguisher D. However, is the converse also true? Namely, if the best distinguishing advantage is upper bounded by ϵ, do such two events always exist?

An affirmative answer is known to exist if we maximize over *all* distinguishers: In this case, the best advantage is the *statistical distance*

$$d(X,Y) := \frac{1}{2}\sum_{u\in\mathcal{U}} |\mathsf{P}_X(u) - \mathsf{P}_Y(u)|\,,$$

and it is *always* possible to define two such events $\mathcal{A}$ and $\mathcal{B}$ by the joint probabilities $\mathsf{P}_{\mathcal{A}X}(u) = \mathsf{P}_{\mathcal{A}Y}(u) = \min\{\mathsf{P}_X(u), \mathsf{P}_Y(u)\}$. Because $d(X,Y) = 1 - \sum_{u\in\mathcal{U}} \min\{\mathsf{P}_X(u), \mathsf{P}_Y(u)\}$, it is easy to see that $\mathsf{P}[\mathcal{A}] = \mathsf{P}[\mathcal{B}] = \sum_u \mathsf{P}_{\mathcal{A}X}(u) = 1 - d(X,Y)$. This can be interpreted as saying the the random variables X and Y are equal with probability $1-\epsilon$. A generalization of this property to discrete systems was considered by Maurer, Pietrzak, and Renner [13].

However, the quantity of interest in the cryptographic setting (as for example in the definition of a PRG) is the best distinguishing advantage of a *computationally bounded* (i.e. polynomial-time) distinguisher, which in general is substantially smaller than the statistical distance $d(X,Y)$, and hence the above property is of no help in the context of computational indistinguishability.

The main technical and conceptual contribution of this paper is a computational version of the above characterization, which we prove both in the uniform and the non-uniform settings. Roughly speaking, we show that if the advantage of every computationally bounded distinguisher is bounded by ϵ (and the statistical distance may be considerably higher), there exist events $\mathcal{A}$ and $\mathcal{B}$ occurring each with probability $1-\epsilon$ such that X and Y are *computationally indistinguishable* when conditioned on $\mathcal{A}$ and $\mathcal{B}$. This can be seen as a hardcore lemma for the setting of computational indistinguishability, and hence solves, for the case of random variables, an open question stated by Myers [15].

The security of the aforementioned concatenate-and-extract approach follows then from the simple observation, due to our characterization, that the output of an ϵ-PRG can be shown to have high computational min-entropy with probability $1-\epsilon$, and hence the concatenation of sufficiently many such outputs always contains enough randomness to be extracted.

1.4 Outline of this Paper

The main part of this paper is Section 3, which is devoted to discussing the characterizations of computational indistinguishability in terms of events in both

the uniform and the non-uniform computational models. Furthermore, Section 4 is devoted to proving the soundness of the concatenate-and-extract approach for security amplification of PRGs. All tools employed throughout this paper are introduced in Section 2, where in particular we discuss the hardcore lemma in the uniform and non-uniform computational models, which is a central component of our main proofs.

2 Preliminaries

2.1 Notational Preliminaries and Computational Model

NOTATION. Recall that a function is *negligible* if it vanishes faster than the inverse of any polynomial. We use both notations poly and negl as placeholders for some polynomial and negligible function, respectively. In particular, a function $\gamma = \frac{1}{\mathsf{poly}}$ is called *noticeable*.

Throughout this paper, we use calligraphic letters $\mathcal{X}, \mathcal{Y}, \ldots$ to denote sets, upper-case letters $X, Y, \ldots$ to denote random variables, and lower-case letters $x, y, \ldots$ denote the values they take on. Moreover, $\mathsf{P}[\mathcal{A}]$ stands for the probability of the event $\mathcal{A}$, while we use the shorthands $\mathsf{P}_X(x) := \mathsf{P}[X = x]$, $\mathsf{P}_{\mathcal{A}|X}(x) := \mathsf{P}[\mathcal{A}|X = x]$, and $\mathsf{P}_{X|\mathcal{A}}(x) := \mathsf{P}[X = x|\mathcal{A}]$. Also, P_X, $\mathsf{P}_{\mathcal{A}|X}$ and $\mathsf{P}_{X|\mathcal{A}}$ are the corresponding (conditional) probability distributions, and $x \xleftarrow{\$} \mathsf{P}$ is the action of sampling a concrete value x according to the distribution P. (We use $x \xleftarrow{\$} \mathcal{X}$ in the case where P is the uniform distribution on $\mathcal{X}$.) Finally, $\mathsf{E}[X]$ is the expected value of the (real-valued) random variable X. Also, we use $\|$ to denote the concatenation of binary strings.

COMPUTATIONAL MODEL. The notation $A^{\mathcal{O}}(x, x', \ldots)$ denotes the (oracle) algorithm A which runs on inputs $x, x', \ldots$ with access to the oracle $\mathcal{O}$. In the asymptotic setting, a *uniform* algorithm A always obtains the unary representation 1^k of the current security parameter k as its first input and is said to run in time $t : \mathbb{N} \to \mathbb{N}$ (or to have *time complexity* t) if for all $k > 0$ the worst-case number of steps it takes (counting oracle queries as single steps) on first input 1^k, taken over all randomness values, all compatible additional inputs and oracles, is at most $t(k)$. In particular, we say that a family of functions $F = \{f_k\}_{k\in\mathbb{N}}$, where $f_k : \mathcal{X}_k \to \mathcal{Y}_k$ is *efficiently* (or polynomial-time) computable if there exists a uniform algorithm which for every security parameter k computes f_k. Finally, we model as usual *non-uniform* algorithms in terms of (families of) circuits $C : \{0,1\}^m \to \{0,1\}^n$ with bounded size.

For ease of notation, we do not make asymptotics explicit in this paper (in particular, we omit the input 1^k), despite the formal statements being asymptotic in nature.

2.2 Pseudorandom Generators and Randomness Extractors

DISTANCE MEASURES. The *distinguishing advantage* of the distinguisher D in distinguishing random variables X and Y with equal range $\mathcal{U}$ is

$$\Delta^D(X, Y) := |\mathsf{P}[D(X) = 1] - \mathsf{P}[D(Y) = 1]|\,,$$

whereas the *statistical distance between* X *and* Y is defined as $d(X,Y) := \frac{1}{2}\sum_{u\in\mathcal{U}} |\mathsf{P}_X(u) - \mathsf{P}_Y(u)| = \sum_{u:\mathsf{P}_X(u)\leq\mathsf{P}_Y(u)} \mathsf{P}_Y(u) - \mathsf{P}_X(u)$.

Pseudorandom Generators. An efficiently computable function $G : \{0,1\}^k \to \{0,1\}^\ell$ is a (t,ϵ)-PRG if for all distinguishers D with time complexity t we have $\Delta^D(G(U_k), U_\ell) \leq \epsilon$, where U_k and U_ℓ are uniformly distributed k- and ℓ-bit strings, respectively. (In the non-uniform setting we rather use the notation (s,ϵ)-PRG, maximizing over all circuits with size at most s.) Furthermore, we use the shorthands ϵ-PRG and PRG for (poly,ϵ)- and $(\mathsf{poly},\mathsf{negl})$-PRGs, respectively.

Randomness Extractors. A *source* $\mathcal{S}$ is a set of probability distributions, and an *ϵ-extractor for* $\mathcal{S}$ is an efficiently computable function $\mathsf{Ext} : \{0,1\}^m \times \{0,1\}^d \to \{0,1\}^n$ such that for a uniformly distributed d-bit string R, we have $d(\mathsf{Ext}(X,R), U_n) \leq \epsilon$ for all m-bit random variables X with $\mathsf{P}_X \in \mathcal{S}$ and a uniformly distributed n-bit string U_n. Furthermore, the extractor is called *strong* if the stronger condition $d\left((\mathsf{Ext}(X,R),R),(U_n,R)\right) \leq \epsilon$ holds.

Also recall that the *min-entropy of* X is $\mathsf{H}_\infty(X) := -\log\left(\max_{x\in\mathcal{X}} \mathsf{P}_X(x)\right)$. A two-parameter function $h : \{0,1\}^m \times \{0,1\}^d \to \{0,1\}^n$ is called *two-universal* if $\mathsf{P}[h(x,K) = h(x',K)] = 2^{-n}$ for any two distinct m-bit x and x' and a uniform d-bit K. An example with $d = m$ is the function $h(x,k) := (x \odot k)|_n$, where $\odot$ is the multiplication of binary strings interpreted as elements of $GF(2^m)$, and $|_n$ outputs the first n bits of a given string. Two-universal hash functions are good extractors:

Lemma 1 (Leftover Hash Lemma [1, 11]). *For any $\epsilon > 0$, every two-universal hash function $h : \{0,1\}^m \times \{0,1\}^d \to \{0,1\}^n$ is a strong ϵ-extractor for the source of m-bit random variables with min-entropy at least $n + 2\log\left(\frac{1}{\epsilon}\right)$.*

We also note that extractors with smaller seed exist for the source of random variables with guaranteed min-entropy. We refer the reader to [16] for a survey.

Deterministic Extractors. An extractor is *deterministic* if $d = 0$, i.e., no additional randomness is needed. (Note that such extractors are vacuously strong.) A class of sources allowing for deterministic extraction are so-called (m,ℓ,k)-*total-entropy independent sources* [12], consisting of random variables of the form $(X_1,\ldots,X_m)$, where $X_1,\ldots,X_m$ are *independent* ℓ-bit strings, and the total min-entropy of $(X_1,\ldots,X_m)$ is at least k.[4] In particular, the following extractor from [12] will be useful for our purposes. (Unconditional constructions requiring a higher entropy rate δ are also given in [12].)

Theorem 1 ([12]). *Under the assumption that primes with length in $[\tau, 2\tau]$ can be found in time $\mathsf{poly}(\tau)$, there is a constant η such that for all $m,\ell \in \mathbb{N}$ and all $\delta > \zeta > (m\ell)^{-\eta}$, there exists a polynomial-time computable ϵ-extractor $\mathsf{Ext} : \left(\{0,1\}^\ell\right)^m \to \{0,1\}^n$ for $(m,\ell,\delta\cdot m\ell)$-total-entropy independent sources, where $n = (\delta-\zeta)m\ell$ and $\epsilon = e^{-(m\ell)^{\Omega(1)}}$.*

[4] Note that in this case $\mathsf{H}_\infty(X_1,\ldots,X_m) = \sum_{i=1}^m \mathsf{H}_\infty(X_i)$.

2.3 Measures and the Hardcore Lemma

GUESSING ADVANTAGES. Let (X, B) be a pair of correlated random variables with joint probability distribution P_{XB}, where B is binary, and let A be an adversary taking an input in the range $\mathcal{X}$ of X and outputting a bit (i.e., A has the same form as a distinguisher): The *guessing advantage* of A in guessing B given X is

$$\mathrm{Guess}^A(B \mid X) := 2 \cdot \mathsf{P}[A(X) = B] - 1.$$

In particular, $\mathrm{Guess}^A(B \mid X) = 1$ means that A is always correct in guessing B given X, whereas it always errs if $\mathrm{Guess}^A(B \mid X) = -1$.[5]

MEASURES. A *measure* $\mathcal{M}$ on a set $\mathcal{X}$ is a mapping $\mathcal{M} : \mathcal{X} \to [0, 1]$. Intuitively, it captures the notion of a "fuzzy" characteristic function of a subset of $\mathcal{X}$. Consequently, its *size* $|\mathcal{M}|$ is defined as $\sum_{x \in \mathcal{X}} \mathcal{M}(x)$, and its *density* is $\mu(\mathcal{M}) := |\mathcal{M}|/|\mathcal{X}|$. Also one associates with each measure $\mathcal{M}$ the probability distribution $\mathsf{P}_{\mathcal{M}}$ such that $\mathsf{P}_{\mathcal{M}}(x) := \mathcal{M}(x)/|\mathcal{M}|$, and we say that a random variable M is *sampled according to* $\mathcal{M}$ if $M \xleftarrow{\$} \mathsf{P}_{\mathcal{M}}$. The following lemma shows that such random variables have high min-entropy, as long as $\mathcal{M}$ is sufficiently dense.

Lemma 2. *Let $\mathcal{M} : \mathcal{X} \to [0, 1]$ be a measure with density $\mu(\mathcal{M}) \geq \delta$, and let M be sampled according to $\mathcal{M}$. Then, $\mathsf{H}_\infty(M) \geq \log |\mathcal{X}| - \log\left(\frac{1}{\delta}\right)$.*

Proof. We have $\mathsf{P}_M(x) = \frac{\mathcal{M}(x)}{|\mathcal{M}|} \leq \frac{\mathcal{M}(x)}{\delta \cdot |\mathcal{X}|} \leq \frac{1}{\delta} \cdot \frac{1}{|\mathcal{X}|}$ due to $\mathcal{M}(x) \leq 1$, which implies $\mathsf{H}_\infty(M) = -\log \max_{x \in \mathcal{X}} \mathsf{P}_{\mathcal{M}}(x) \geq \log |\mathcal{X}| - \log(1/\delta)$. □

THE HARDCORE LEMMA. For a set $\mathcal{W}$, let $g : \mathcal{W} \to \mathcal{Y}$ be a function, and let $P : \mathcal{W} \to \{0, 1\}$ be a predicate. The so-called *hardcore lemma* shows that, roughly speaking, if $\mathrm{Guess}^A(P(W) \mid g(W)) \leq \delta$ (for W uniform in $\mathcal{W}$) for all efficient A, then for all $\gamma > 0$ there exists a measure $\mathcal{M}$ on $\mathcal{W}$ with $\mu(\mathcal{M}) \geq 1 - \delta$ such that $\mathrm{Guess}^{A'}(P(W') \mid g(W')) \leq \gamma$ for all efficient adversaries A' and for W' sampled according to $\mathcal{M}$. This result was first introduced and proven by Impagliazzo [10]. However, his original proof only ensures $\mu(\mathcal{M}) \geq \frac{1-\delta}{2}$. The following theorem, due to Holenstein [8], gives a *tight* version of the lemma for the non-uniform setting.

Theorem 2 (Non-Uniform Hardcore Lemma). *Let $g : \mathcal{W} \to \mathcal{X}$ and $P : \mathcal{W} \to \{0, 1\}$ be functions, and let $\delta, \gamma \in (0, 1)$ and $s > 0$ be given. If for all circuits C with size s we have*

$$\mathrm{Guess}^C(P(W) \mid g(W)) \leq \delta$$

for $W \xleftarrow{\$} \mathcal{W}$, then there exists a measure $\mathcal{M}$ on $\mathcal{W}$ (called the hardcore measure*) such that $\mu(\mathcal{M}) \geq 1 - \delta$ and such that all circuits C' with size $s' = \frac{s \cdot \gamma^2}{32 \log |\mathcal{W}|}$ satisfy $\mathrm{Guess}^{C'}(P(W') \mid g(W')) \leq \gamma$, where $W' \xleftarrow{\$} \mathsf{P}_{\mathcal{M}}$.*

[5] In particular, flipping the output bit of such an A yields an adversary that is always correct.

A slightly weaker statement holds in the *uniform* setting, where we can only show that for every polynomial-time adversary A' there exists a measure $\mathcal{M}$ for which $\text{Guess}^{A'}(P(W') \,|\, g(W')) \leq \gamma$ *even* if A' is allowed to query the measure $\mathcal{M}$ as an oracle[6] *before* obtaining $g(W')$. This is captured by the following theorem also due to Holenstein [8].

Theorem 3 (Uniform Hardcore Lemma). *Let $g : \mathcal{W} \to \mathcal{X}$, $P : \mathcal{W} \to \{0,1\}$, $\delta : \mathbb{N} \to [0,1]$, and $\gamma : \mathbb{N} \to [0,1]$ be functions computable in time $\mathsf{poly}(k)$, where δ and γ are both noticeable. Assume that for all polynomial-time adversaries A we have*

$$\text{Guess}^{A}(P(W) \,|\, g(W)) \leq \delta$$

for $W \xleftarrow{\$} \mathcal{W}$, then for all polynomial-time adversaries $A'^{(\cdot)}$, whose oracle queries are independent of their input[7], there exists a measure $\mathcal{M}$ on $\mathcal{W}$ with $\mu(\mathcal{M}) \geq 1 - \delta$ such that $\text{Guess}^{A',\mathcal{M}}(P(W') \,|\, g(W')) \leq \gamma$, where $W' \xleftarrow{\$} \mathsf{P}_{\mathcal{M}}$.

The independence requirement on oracle queries is due to the hardcore lemma of [8] considering a model with uniform adversaries A' which are given oracle access to $\mathcal{M}$ (with no input) and subsequently output a *circuit* for guessing $P(W')$ out of $g(W')$ (which in particular does not make queries to $\mathcal{M}$). The simpler statement of Theorem 3 follows by standard techniques.

Note that in contrast to [10, 8], and the traditional literature on the hardcore lemma, we swap the roles of δ and $1 - \delta$ in order to align our statements with the (natural) information-theoretic intuition. Also, note that both theorems have equivalent versions in terms of hardcore *sets* (i.e., where $\mathcal{M}(x) \in \{0,1\}$), yet we limit ourselves to considering the measure versions in this paper.

Efficient Sampling from Measures. Sometimes, we need to sample a random element according to a measure $\mathcal{M}$ on $\mathcal{X}$ with $\mu(\mathcal{M}) \geq \delta$ (for a noticeable δ) given only oracle access to this measure. A solution to this is to sample a random element $x \xleftarrow{\$} \mathcal{X}$ and then output x with probability $\mathcal{M}(x)$, and otherwise go to the next iteration (and abort if a maximal number of iterations k is achieved.) It is easy to see that if an output is produced, it has the right distribution, whereas the probability that no output is produced is at most $(1-\delta)^k < e^{-\delta k}$, and can hence be made smaller than any $\alpha > 0$ by choosing $k = \frac{1}{\delta} \ln\left(\frac{1}{\alpha}\right)$.

In the following, we assume that the sampling can be done perfectly, neglecting the inherent small error probability in the analysis.

3 Characterizing Computational Indistinguishability via Hardcore Theorems

3.1 Non-Uniform Case

This section considers a setting with two efficiently computable functions $E : \mathcal{U} \to \mathcal{X}$ and $F : \mathcal{V} \to \mathcal{X}$, and we define the random variables $X := E(U)$ and

[6] That is, the oracle $\mathcal{M}$ answers a query x with $\mathcal{M}(x) \in [0,1]$.

[7] In particular, they only depend on the randomness of the distinguisher and previous oracle queries.

$Y := F(V)$, where U and V are uniformly[8] distributed on $\mathcal{U}$ and $\mathcal{V}$, respectively. Note that this is the usual way to capture that X and Y are *efficiently samplable*, where typically $\mathcal{U}$ and $\mathcal{V}$ both consist of bit strings of some length.

Let us now assume that $\Delta^D(X, Y) \leq \epsilon$ for every efficient distinguisher D. In full analogy with the information-theoretic setting [13], we aim at *extending* the random experiments where $E(U)$ and $F(V)$ are sampled by adjoining, for all $\gamma > 0$, corresponding events $\mathcal{A}$ and $\mathcal{B}$ defined by conditional probability distributions $\mathsf{P}_{\mathcal{A}|U}$ and $\mathsf{P}_{\mathcal{B}|V}$, such that both events occur with probability roughly $1-\epsilon$, and, conditioned on $\mathcal{A}$ and $\mathcal{B}$, respectively, the random variables $E(U)$ and $F(V)$ can be distinguished with advantage at most γ by an efficient distinguisher. Note that for notational convenience (and in order to interpret the result as a hardcore lemma), we describe the conditional probability distributions $\mathsf{P}_{\mathcal{A}|U}$ and $\mathsf{P}_{\mathcal{B}|V}$ in terms of measures $\mathcal{M} : \mathcal{U} \to [0, 1]$ and $\mathcal{N} : \mathcal{V} \to [0, 1]$. In particular, the values $\mathcal{M}(u)$ and $\mathcal{N}(v)$ take the roles of $\mathsf{P}_{\mathcal{A}|U}(u)$ and $\mathsf{P}_{\mathcal{B}|V}(v)$, and note that $\mu(\mathcal{M}) = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \mathcal{M}(u) = \sum_{u \in \mathcal{U}} \mathsf{P}_U(u) \cdot \mathsf{P}_{\mathcal{A}|U}(u) = \mathsf{P}[\mathcal{A}]$ and hence $\mathsf{P}_M(u) = \frac{\mathcal{M}(u)}{|\mathcal{M}|} = \frac{\mathsf{P}_{\mathcal{A}|U}(u)\mathsf{P}_U(u)}{\mathsf{P}[\mathcal{A}]} = \mathsf{P}_{U|\mathcal{A}}(u)$.

This is summarized by the following theorem. We refer the reader to Section 3.2 for its proof.

Theorem 4 (Non-Uniform Indistinguishability Hardcore Lemma). *Let $E : \mathcal{U} \to \mathcal{X}$ and $F : \mathcal{V} \to \mathcal{X}$ be functions, and let $\epsilon, \gamma \in (0, 1)$ and $s > 0$ be given. If for all distinguishers D with size s we have*

$$\Delta^D(E(U), F(V)) \leq \epsilon$$

for $U \xleftarrow{\$} \mathcal{U}$ and $V \xleftarrow{\$} \mathcal{V}$, then there exist measures $\mathcal{M}$ on $\mathcal{U}$ and $\mathcal{N}$ on $\mathcal{V}$ with $\mu(\mathcal{M}) \geq 1 - \epsilon$ and $\mu(\mathcal{N}) \geq 1 - \epsilon$ such that

$$\Delta^{D'}(E(U'), F(V')) \leq \gamma,$$

for all distinguishers D' with size $s' := \frac{s \cdot \gamma^2}{128(\log |\mathcal{U}| + \log |\mathcal{V}| + 1)}$, where $U' \xleftarrow{\$} \mathsf{P}_{\mathcal{M}}$ and $V' \xleftarrow{\$} \mathsf{P}_{\mathcal{N}}$.

We stress that the measures $\mathcal{M}$ and $\mathcal{N}$ given by the theorem generally depend on γ and s.

PRGs and Computational Entropy. As an example application of Theorem 4, we instantiate the function E by an (s, ϵ)-PRG $G : \{0,1\}^k \to \{0,1\}^\ell$ (in particular $\mathcal{U} := \{0,1\}^k$ and $\mathcal{X} := \{0,1\}^\ell$), whereas F is the identity function and $\mathcal{V} = \mathcal{X} = \{0,1\}^\ell$. For any $\gamma > 0$, Theorem 4 implies that we can define an event $\mathcal{A}$ on the choice of the seed of the PRG (with $\mathsf{P}_{\mathcal{A}|U}(u) := \mathcal{M}(u)$) occurring with probability $\mathsf{P}[\mathcal{A}] = \mu(\mathcal{M}) \geq 1 - \epsilon$ such that conditioned on $\mathcal{A}$, no distinguisher with size s' can achieve advantage higher than γ in distinguishing the ℓ-bit PRG

[8] In fact, our results can naturally be generalized to the case where U and V have arbitrary distributions by considering a slightly more general version of Theorem 2 with arbitrary distributions for W.

output from an ℓ-bit string U'_ℓ sampled according to $\mathcal{N}$, which, by Lemma 2, has min-entropy at least $\ell - \log\left(\frac{1}{1-\epsilon}\right)$.

In other words, the output of every ϵ-PRG $G : \{0,1\}^k \to \{0,1\}^\ell$ exhibits (with probability $1-\epsilon$) high computational min-entropy. Note that the achieved form of computational entropy is somewhat weaker than the traditional notion of HILL min-entropy [6], where the random variable U'_ℓ is the *same* for all polynomially bounded s and all noticeable $\gamma > 0$. Still, it is strong enough to allow for the use of G's output in place of some string which has high min-entropy with probability $1 - \epsilon$.

3.2 Proof of Theorem 4

We start by defining a function $g : \mathcal{U} \times \mathcal{V} \times \{0,1\} \to \mathcal{X}$ and a predicate $P : \mathcal{U} \times \mathcal{V} \times \{0,1\} \to \{0,1\}$ such that

$$g(u,v,b) := \begin{cases} E(u) & \text{if } b = 0, \\ F(v) & \text{if } b = 1. \end{cases}$$

and $P(u,v,b) := b$ for all $u \in \mathcal{U}$ and $v \in \mathcal{V}$. It is well known that for any two random variables $\widetilde{U}$ and $\widetilde{V}$, and a distinguisher D, the distinguishing advantage $\Delta^D(E(\widetilde{U}), F(\widetilde{V}))$ can equivalently be characterized in terms of the probability that D guesses the uniform random bit B in a game where it is given $E(\widetilde{U})$ if $B = 0$ and $F(\widetilde{V})$ otherwise: In particular, we have

$$\Delta^D(E(\widetilde{U}), F(\widetilde{V})) = \left|\mathrm{Guess}^D(B \,|\, g(\widetilde{U}, \widetilde{V}, B))\right|$$

for a uniform random bit B, where $\widetilde{U}$, $\widetilde{V}$, and B are sampled independently.

We now assume towards a contradiction that for all pairs of measures $\mathcal{M}$ and $\mathcal{N}$, both with density at least $1 - \epsilon$, there exists a distinguisher D' of size at most s' with $\Delta^{D'}(E(U'), F(V')) \geq \gamma$, for $U' \overset{\$}{\leftarrow} \mathsf{P}_\mathcal{M}$ and $V' \overset{\$}{\leftarrow} \mathsf{P}_\mathcal{N}$.

We prove that, under this assumption, for all measures $\mathcal{M}$ on $\mathcal{U} \times \mathcal{V} \times \{0,1\}$ with $\mu(\mathcal{M}) \geq 1 - \epsilon$ there exists a circuit C' with size s' such that $\mathrm{Guess}^{C'}(B' \,|\, g(U', V', B')) \geq \frac{\gamma}{2}$, for $(U', V', B') \overset{\$}{\leftarrow} \mathsf{P}_\mathcal{M}$. As this contradicts the statement of the non-uniform hardcore lemma (Theorem 2) for $\frac{\gamma}{2}$ (instead of γ), this implies that there must be a circuit C with size s such that

$$\Delta^C(E(U), F(V)) \geq \mathrm{Guess}^C(B \,|\, g(U, V, B)) > \epsilon.$$

In turn, this contradicts the assumed indistinguishability of $E(U)$ and $F(V)$, concluding the proof.

Reduction to the Hardcore Lemma. In the remainder of this proof, let us assume that we are given a measure $\mathcal{M}$ on $\mathcal{U} \times \mathcal{V} \times \{0,1\}$ with $\mu(\mathcal{M}) \geq 1 - \epsilon$. We first define the measures $\mathcal{M}_0$ and $\mathcal{M}_1$ on $\mathcal{U}$ and $\mathcal{V}$, respectively, such that

$$\mathcal{M}_0(u) := \frac{1}{|\mathcal{V}|} \sum_{v \in \mathcal{V}} \mathcal{M}(u,v,0) \quad \text{and} \quad \mathcal{M}_1(v) := \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \mathcal{M}(u,v,1)$$

Furthermore, let $m_b := \sum_{u,v} \mathcal{M}(u,v,b)$ for $b \in \{0,1\}$, and let $m := |\mathcal{M}| = m_0 + m_1$. Note that in particular $\mu(\mathcal{M}_b) = \frac{m_b}{|\mathcal{U}|\cdot|\mathcal{V}|}$ and $\mu(\mathcal{M}) = \frac{m}{2\cdot|\mathcal{U}|\cdot|\mathcal{V}|}$.

We consider two cases in the following, both leading to a circuit C'.

Case $\left|\frac{m_0}{m} - \frac{1}{2}\right| > \frac{\gamma}{4}$. Assume that $\frac{m_0}{m} - \frac{1}{2} > \frac{\gamma}{4}$. (The other case is symmetric.) Then, for the circuit C' *always* outputting the bit 0,

$$\mathrm{Guess}^{C'}(B' \mid g(U',V',B')) = 2 \cdot \mathsf{P}[B'=0] - 1 = 2 \cdot \tfrac{m_0}{m} - 1 > \frac{\gamma}{2}.$$

Case $\left|\frac{m_0}{m} - \frac{1}{2}\right| \leq \frac{\gamma}{4}$. We assume that $\frac{1}{2} \geq \frac{m_0}{m} \geq \frac{1}{2}(1-\frac{\gamma}{2})$ and hence also $\frac{1}{2}(1+\frac{\gamma}{2}) \geq \frac{m_1}{m} \geq \frac{1}{2}$ (once again the other case is symmetric). This yields in particular that $(1-\gamma/2)\mu(\mathcal{M}) \leq \mu(\mathcal{M}_0) \leq \mu(\mathcal{M})$ and $\mu(\mathcal{M}) \leq \mu(\mathcal{M}_1) \leq (1+\gamma/2)\mu(\mathcal{M})$.

The goal is to define two measures $\widetilde{\mathcal{M}}_0$ on $\mathcal{U}$ and $\widetilde{\mathcal{M}}_1$ on $\mathcal{V}$, *both* with density at least $1-\epsilon$, such that a distinguisher D' achieving advantage larger than γ in distinguishing $E(\widetilde{U}')$ and $F(\widetilde{V}')$ for $\widetilde{U}' \xleftarrow{\$} \mathsf{P}_{\widetilde{\mathcal{M}}_0}$ and $\widetilde{V}' \xleftarrow{\$} \mathsf{P}_{\widetilde{\mathcal{M}}_1}$ also achieves advantage higher than $\gamma/2$ in guessing B' given $g(U',V',B')$. Ideally, we would set $\widetilde{\mathcal{M}}_b := \mathcal{M}_b$, but note that $\mu(\mathcal{M}_0) < 1-\epsilon$ possibly holds. We slightly modify $\mathcal{M}_0$ in order to satisfy this property, i.e., we define for all $u \in \mathcal{U}$ and $v \in \mathcal{V}$

$$\widetilde{\mathcal{M}}_0(u) := \frac{1-\mu(\mathcal{M})}{1-\mu(\mathcal{M}_0)} \cdot \mathcal{M}_0(u) + \frac{\mu(\mathcal{M}) - \mu(\mathcal{M}_0)}{1-\mu(\mathcal{M}_0)} \quad \text{and} \quad \widetilde{\mathcal{M}}_1(v) := \mathcal{M}_1(v).$$

(We tacitly assume $\mu(\mathcal{M}_0) < 1$, otherwise we can simply set $\widetilde{\mathcal{M}}_0 := \mathcal{M}_0$.) It is easy to verify that $\mathcal{M}_0(u) \leq \widetilde{\mathcal{M}}_0(u) \leq 1$. Moreover,

$$\mu(\widetilde{\mathcal{M}}_0) = \frac{1-\mu(\mathcal{M})}{1-\mu(\mathcal{M}_0)} \cdot \mu(\mathcal{M}_0) + \frac{\mu(\mathcal{M}) - \mu(\mathcal{M}_0)}{1-\mu(\mathcal{M}_0)} = \mu(\mathcal{M}) \geq 1-\epsilon.$$

This implies, by our assumption, that for $\widetilde{U}'$ and $\widetilde{V}'$ sampled according to $\widetilde{\mathcal{M}}_0$ and $\widetilde{\mathcal{M}}_1$ there exists D' such that

$$\mathsf{P}[D'(F(\widetilde{V}')) = 1] - \mathsf{P}[D'(E(\widetilde{U}')) = 1] > \gamma. \tag{1}$$

We now show that the advantage of D' in guessing B' given $g(U',V',B')$ is larger than $\frac{\gamma}{2}$. To this aim, we introduce the following two probability distributions P_1 and P_2, both with range $(\mathcal{U} \times \{0\}) \cup (\mathcal{V} \times \{1\})$. The former distribution is the distribution of $(g(\widetilde{U}',\widetilde{V}',B),B)$ for $\widetilde{U}' \xleftarrow{\$} \mathsf{P}_{\widetilde{\mathcal{M}}_0}$, $\widetilde{V}' \xleftarrow{\$} \mathsf{P}_{\widetilde{\mathcal{M}}_1}$, and $B \xleftarrow{\$} \{0,1\}$, that is

$$\mathsf{P}_1(u,0) := \frac{\widetilde{\mathcal{M}}_0(u)}{2|\widetilde{\mathcal{M}}_0|} \quad \text{and} \quad \mathsf{P}_1(v,1) := \frac{\widetilde{\mathcal{M}}_1(v)}{2|\widetilde{\mathcal{M}}_1|} \text{ for all } u \in \mathcal{U} \text{ and } v \in \mathcal{V}.$$

The latter is the distribution of $(g(U',V',B'),B')$ for $(U',V',B') \xleftarrow{\$} \mathsf{P}_{\mathcal{M}}$, i.e.,

$$\mathsf{P}_2(u,0) := \frac{|\mathcal{V}| \cdot \mathcal{M}_0(u)}{|\mathcal{M}|} \quad \text{and} \quad \mathsf{P}_2(v,1) := \frac{|\mathcal{U}| \cdot \mathcal{M}_1(v)}{|\mathcal{M}|} \text{ for all } u \in \mathcal{U} \text{ and } v \in \mathcal{V}.$$

We prove the following two lemmas for $(X_1,B_1) \xleftarrow{\$} \mathsf{P}_1$ and $(X_2,B_2) \xleftarrow{\$} \mathsf{P}_2$.

Lemma 3. $\mathrm{Guess}^{D'}(B' \mid g(U', V', B')) > \gamma - 2 \cdot d((X_1, B_1), (X_2, B_2))$.

Proof. Consider the distinguisher $\overline{D}$ which given a pair $(x, b) \in (\mathcal{U} \times \{0\}) \cup (\mathcal{V} \times \{1\})$ outputs 1 if $b = 0$ *and* $D'(E(x)) = 0$, or if $b = 1$ *and* $D'(F(x)) = 1$. Then, note that by (1)

$$\begin{aligned}\mathsf{P}[\overline{D}(X_1, B_1) = 1] &= \tfrac{1}{2}\left(\mathsf{P}[D'(E(\widetilde{U}')) = 0] + \mathsf{P}[D'(F(\widetilde{V}')) = 1]\right) \\ &= \tfrac{1}{2}\left(1 + \mathsf{P}[D'(F(\widetilde{V}')) = 1] - \mathsf{P}[D'(E(\widetilde{U}')) = 1]\right) > \tfrac{1}{2} + \tfrac{\gamma}{2}.\end{aligned}$$

Furthermore, $\mathsf{P}(\overline{D}(X_2, B_2) = 1] \leq \frac{1}{2} + \frac{\mathrm{Guess}^{D'}(B' \mid g(U', V', B'))}{2}$ by the definition of g. The fact that

$$\begin{aligned}\mathsf{P}[\overline{D}(X_1, B_1) = 1] - \mathsf{P}[\overline{D}(X_2, B_2) = 1] &\leq \Delta^{\overline{D}}((X_1, B_1), (X_2, B_2)) \\ &\leq d((X_1, B_1), (X_2, B_2))\end{aligned}$$

implies the lemma. □

Lemma 4. $d((X_1, B_1), (X_2, B_2)) \leq \frac{\gamma}{4}$.

Proof. For all $v \in \mathcal{V}$ we have $\mathsf{P}_1(v, 1) \leq \mathsf{P}_2(v, 1)$, since $|\mathcal{M}| \leq 2 \cdot |\mathcal{U}| \cdot |\mathcal{M}_1|$. Furthermore, for all $u \in \mathcal{U}$ we have

$$\mathsf{P}_1(u, 0) = \frac{\widetilde{\mathcal{M}}_0(u)}{2|\widetilde{\mathcal{M}}_0|} = \frac{\widetilde{\mathcal{M}}_0(u)}{2|\mathcal{U}|\mu(\mathcal{M})} \geq \frac{\mathcal{M}_0(u)}{2|\mathcal{U}|\mu(\mathcal{M})} = \frac{|\mathcal{V}| \cdot \mathcal{M}_0(u)}{|\mathcal{M}|} = \mathsf{P}_2(u, 0)$$

using the fact that $|\mathcal{M}| = 2 \cdot \mu(\mathcal{M}) \cdot |\mathcal{U}| \cdot |\mathcal{V}|$. This yields

$$\begin{aligned}d((X_1, B_1), (X_2, B_2)) &= \sum_{v \in \mathcal{V}} \mathsf{P}_2(v, 1) - \mathsf{P}_1(v, 1) \\ &= \frac{1}{2|\mathcal{V}|} \sum_{v \in \mathcal{V}} \mathcal{M}_1(v) \cdot \left(\frac{1}{\mu(\mathcal{M})} - \frac{1}{\mu(\mathcal{M}_1)}\right) \\ &= \frac{1}{2}\left(\frac{\mu(\mathcal{M}_1)}{\mu(\mathcal{M})} - 1\right) \leq \frac{\gamma}{4},\end{aligned}$$

since $\mu(\mathcal{M}_1) \leq (1 + \frac{\gamma}{2}) \cdot \mu(\mathcal{M})$. □

Therefore, we conclude the proof of Theorem 4 by combining both lemmas, which show that the advantage of $C' := D'$ is larger than $\gamma/2$, as desired.

3.3 The Uniform Case

In this section, we prove a *uniform* version of Theorem 4 in the same spirit as the uniform hardcore lemma (Theorem 3): If $E(U)$ and $F(V)$ can only be distinguished with advantage ϵ by a polynomial-time distinguisher, then for all noticeable $\gamma > 0$ and for all polynomial-time oracle distinguishers $D^{(\cdot,\cdot)}$ (making input-independent oracle queries), there exist two measures $\mathcal{M}$ and $\mathcal{N}$ on $\mathcal{U}$ and $\mathcal{V}$, each with density $1 - \epsilon$, such that $D^{\mathcal{M},\mathcal{N}}$ cannot achieve advantage better than γ in telling $E(U')$ and $F(V')$ apart, where $U' \xleftarrow{\$} \mathsf{P}_{\mathcal{M}}$ and $V' \xleftarrow{\$} \mathsf{P}_{\mathcal{N}}$.

Theorem 5 (Uniform Indistinguishability Hardcore Lemma). *Let* $E : \mathcal{U} \to \mathcal{X}$ *and* $F : \mathcal{V} \to \mathcal{X}$, $\epsilon : \mathbb{N} \to [0,1]$, *and* $\gamma : \mathbb{N} \to [0,1]$ *be functions computable in time* $\mathsf{poly}(k)$, *where* ϵ *and* γ *are both noticeable. Assume that for all polynomial-time distinguishers* D *we have*

$$\Delta^D(E(U), F(V)) \leq \epsilon$$

for $U \xleftarrow{\$} \mathcal{U}$ *and for* $V \xleftarrow{\$} \mathcal{V}$, *then for all polynomial time distinguishers* $D'^{(\cdot,\cdot)}$ *whose oracle queries are independent of their input, there exist measures* $\mathcal{M}$ *on* $\mathcal{U}$ *and* $\mathcal{N}$ *on* $\mathcal{V}$ *with* $\mu(\mathcal{M}) \geq 1-\epsilon$ *and* $\mu(\mathcal{N}) \geq 1-\epsilon$ *such that*

$$\Delta^{D'^{\mathcal{M},\mathcal{N}}}(E(U'), F(V')) \leq \gamma,$$

where $U' \xleftarrow{\$} \mathsf{P}_{\mathcal{M}}$ *and* $V' \xleftarrow{\$} \mathsf{P}_{\mathcal{N}}$.

Due to lack of space, the proof of the theorem (which follows the lines of the non-uniform case, but with extra difficulties) can be found in the full version.

4 Security Amplification of PRGs

4.1 The Concatenate-And-Extract Construction

This section presents, as an application of Theorems 4 and 5, the first construction achieving security amplification of arbitrarily weak PRGs.

Construction. Let $G : \{0,1\}^k \to \{0,1\}^\ell$ and $\mathsf{Ext} : \{0,1\}^{m\ell} \times \{0,1\}^d \to \{0,1\}^n$ be efficiently computable functions. We consider the *Concatenate-and-Extract* (CaE) construction $\mathsf{CaE}^{G,\mathsf{Ext}} : \{0,1\}^{mk+d} \to \{0,1\}^{n+d}$ such that

$$\mathsf{CaE}^{G,\mathsf{Ext}}(x_1 \,\|\, \cdots \,\|\, x_m \,\|\, r) := \mathsf{Ext}(G(x_1) \,\|\, \cdots \,\|\, G(x_m), r) \,\|\, r.$$

for all $x_1, \ldots, x_m \in \{0,1\}^k$ and $r \in \{0,1\}^d$.

Parameters and Main Security Statement. The intuition justifying the security of the CaE construction relies on the simple observation that, provided that G is an ϵ-PRG, each individual and *independent* PRG output in the concatenation $G(X_1)\| \cdots \|G(X_m)$ (for uniform $X_1, \ldots, X_m$) has computational min-entropy at least $\ell - \log\left(\frac{1}{1-\epsilon}\right)$ with probability at least $1-\epsilon$, and thus we can expect the whole concatenation to have computational min-entropy $\approx m \cdot (1-\epsilon) \cdot \left[\ell - \log\left(\frac{1}{1-\epsilon}\right)\right]$ with very high probability, which can be extracted if Ext is an appropriate extractor. Note that the resulting construction is expanding if $n/mk > 1$, and for an optimal extractor this ratio is roughly $(1-\epsilon)\left[\ell - \log\left(\frac{1}{1-\epsilon}\right)\right]/k$ (we ignore the entropy loss of the extractor for simplicity), or, turned around, our construction expands if the underlying ϵ-PRG satisfies $\ell/k > \frac{1}{1-\epsilon} + \log\left(\frac{1}{1-\epsilon}\right)/k$ holds. In particular, this value is independent of m. In Section 4.3, we show that for a large class of natural constructions

this is essentially optimal. For example, for $\epsilon = \frac{1}{2}$, the output length ℓ of the given generator G needs to be slightly larger than $2k$ in order to achieve expansion. For comparison, the SUM construction is expanding if $\ell/k > m$, where $m = \omega(k/\log(1/\epsilon))$ in order for the construction to be secure.

Also, the fact that all ℓ-bit blocks are independent allows for using *deterministic* extractors in the CaE construction, such as the one given by Theorem 1, as long as $(1-\epsilon)\left(1 - \frac{1}{\ell}\log\left(\frac{1}{1-\epsilon}\right)\right)$ is bounded from below by $(m\ell)^{-\eta}$.

The following theorem proves the soundness of the above intuition.

Theorem 6 (Strong Security Amplification for PRGs). *Let $\rho, \delta, \epsilon : \mathbb{N} \to [0,1]$ be functions, and let $G : \{0,1\}^k \to \{0,1\}^\ell$ (for $k < \ell$) be an ϵ-PRG. Furthermore, let $\mathsf{Ext} : \{0,1\}^{m\ell} \times \{0,1\}^d \to \{0,1\}^n$ be a strong δ-extractor for $\left(m, \ell, (1-\epsilon-\rho)\cdot m \cdot \left[\ell - \log\left(\frac{1}{1-\epsilon}\right)\right]\right)$-total-entropy independent sources.*

Then the function $\mathsf{CaE}^{G,\mathsf{Ext}} : \{0,1\}^{mk+d} \to \{0,1\}^{n+d}$ is a $(e^{-\rho^2 m} + \delta + \nu)$-PRG, where ν is a negligible function.

The theorem is proven by means of a *uniform reduction* using Theorem 5, and hence holds both in the uniform *and* in the non-uniform settings. However, the next paragraph gives an ad-hoc proof for the non-uniform case which follows the above simple intuition and which is also tighter than the more involved uniform reduction, which we defer to Section 4.2.

Non-Uniform Proof. In the following, let us fix $s, \gamma > 0$, and assume G is an (s,ϵ)-PRG. Also consider the $m\ell$-bit string $G(X_1)\| \ldots \|G(X_m)$, where $X_1, \ldots, X_m$ are independent uniform k-bit strings.

By Theorem 4, there exist independent events $\mathcal{A}_1, \ldots, \mathcal{A}_m$ such that $\mathcal{A}_i$ can be adjoined to X_i and $\mathsf{P}[\mathcal{A}_i] \geq 1-\epsilon$, and, conditioned on theses events, no size s' distinguisher can distinguish $G(X_i)$ from some variable U_i' with min-entropy $\mathsf{H}_\infty(U_i') \geq \ell - \log\left(\frac{1}{1-\epsilon}\right)$ with advantage larger than γ. In particular, by Hoeffding's inequality (Lemma 5), the events $\mathcal{A}_i$ occur for a subset $\mathcal{I} \subseteq \{1, \ldots, m\}$ of indices such that $|\mathcal{I}| \geq (1-\epsilon-\rho)\cdot m$, except with probability $e^{-\rho^2 m}$. In this case, for a uniform random d-bit string R, a standard hybrid argument yields that every distinguisher of size s'' (where s'' is only slightly smaller than s') can achieve advantage at most $m\gamma$ in distinguishing $\mathsf{CeE}^{G,\mathsf{Ext}}(X_1\| \ldots \|X_m\|R) = \mathsf{Ext}(G(X_1)\| \ldots \|G(X_m), R)\|R$ from the string $\mathsf{Ext}(U', R)\|R$, where U' is obtained by replacing each $G(X_i)$ with $i \in \mathcal{I}$ with the corresponding U_i'. In particular, since U' has min-entropy at least $(1-\epsilon-\rho)\cdot m \cdot \left[\ell - \log\left(\frac{1}{1-\epsilon}\right)\right]$, the variable $\mathsf{Ext}(U', R)\|R$ has distance at most δ from a uniform random $(n+d)$-bit string U_{n+d}.

Thus, using the triangle inequality and adding the three advantages, we obtain that $\mathsf{CaE}^{G,\mathsf{Ext}}$ is a $(s'', m\gamma + \delta + e^{-\rho^2 m})$-PRG. The asymptotic bound follows by applying the same argument to any polynomially bounded s and to all noticeable γ.

Distinguisher $D'^{\mathcal{M},\mathcal{N}}(z)$ *// On input* $z \in \{0,1\}^\ell$

$x_1, \ldots, x_m \xleftarrow{\$} \{0,1\}^k$, $r \xleftarrow{\$} \{0,1\}^d$
for all $i = 1, \ldots, m$ **do**
  $\mathcal{G} := \mathcal{G} \cup \{i\}$ with probability $\mathcal{M}(x_i)$
$i^* \xleftarrow{\$} \{1, \ldots, m\}$
for all $i = 1, \ldots m$ **do**
  if $i \in \mathcal{G}$ *and* $i < i^*$ **then** $y_i \xleftarrow{\$} \mathsf{P}_\mathcal{N}$ **else** $y_i := G(x_i)$
if $i^* \in \mathcal{G}$ **then**
  return $D'_{\mathcal{M},\mathcal{N}}(\cdot) := D(\mathsf{Ext}(y_1 \,\|\, \ldots \,\|\, y_{i^*-1} \,\|\, z \,\|\, y_{i^*+1} \,\|\, \ldots \,\|\, y_m, r) \,\|\, r)$
else
  return $D'_{\mathcal{M},\mathcal{N}}(\cdot) := D(\mathsf{Ext}(y_1 \,\|\, \ldots \,\|\, y_m, r) \,\|\, r)$ *// $D'_{\mathcal{M},\mathcal{N}}$ ignores its input*

Fig. 1. The distinguisher $D'^{(\cdot,\cdot)}$ in the proof of Theorem 6.

4.2 Proof of Theorem 6

Assume, towards a contradiction, that there exists a polynomial-time distinguisher D and a noticeable function η such that for infinitely many values of the security parameter k (which we omit) we have

$$\Delta^D(\mathsf{CaE}^{G,\mathsf{Ext}}(X_1 \,\|\, \cdots \,\|\, X_m \,\|\, R), U_{n+d}) > e^{-\rho^2 m} + \delta + \eta,$$

where $X_1, \ldots, X_m$ are uniformly distributed k-bit string, R is a uniformly distributed d-bit string, and U_{n+d} is a uniformly distributed $(n+d)$-bit string.

The Distinguisher $D'^{(\cdot,\cdot)}$. We give a distinguisher $D'^{(\cdot,\cdot)}$ (which is fully specified in Figure 1), which on input $z \in \{0,1\}^\ell$ and given oracle access to any two measures $\mathcal{M} : \mathcal{U} \to [0,1]$ and $\mathcal{N} : \mathcal{V} \to [0,1]$ operates as follows: First, it chooses m k-bit strings $x_1, \ldots, x_m$ independently and uniformly at random, and for each $i \in \{1, \ldots, m\}$ an independent coin is flipped (taking value 1 with probability $\mathcal{M}(x_i)$, and 0 otherwise), and if the coin flip returns 1, the position i is marked as "being in the measure". Let $\mathcal{G}$ be the set of marked positions. Subsequently, an index i^* is chosen uniformly random from $\{1, \ldots, m\}$. Then, a string $y_1 \,\|\, \ldots \,\|\, y_m \in \{0,1\}^{m\ell}$ (where $y_1, \ldots, y_m \in \{0,1\}^\ell$) is built as follows: Each y_i is set to an independent element sampled according to $\mathsf{P}_\mathcal{N}$ if $i \in \mathcal{G}$ and $i < i^*$, and in any other case it is set to $G(x_i)$. Finally, the distinguisher chooses the seed r for the extractor uniformly at random, and outputs the bit $D(\mathsf{Ext}(y_1 \,\|\, \ldots \,\|\, y_{i^*-1} \,\|\, z \,\|\, y_{i^*+1} \,\|\, \ldots \,\|\, y_m, r) \,\|\, r)$ if $i^* \in \mathcal{G}$ holds, or it outputs $D(\mathsf{Ext}(y_1 \,\|\, \ldots \,\|\, y_m, r) \,\|\, r)$ else (in particular, the input z is ignored in this latter case).

Analysis. In the following, let $\mathcal{M}$ and $\mathcal{N}$ both have density $1 - \epsilon$, let X' be sampled according to $\mathsf{P}_\mathcal{M}$, and let U' be sampled according to $\mathsf{P}_\mathcal{N}$. We compute the average advantage $\Delta^{D'}(G(X'), U')$ of the distinguisher $D' = D'^{\mathcal{M},\mathcal{N}}$.

It is convenient to use the shorthands $\mathsf{P}[D'(\cdot) \,|\, g] := \mathsf{P}[D'(X) = 1 \,|\, |\mathcal{G}| = g]$ to denote the conditional probability of D' outputting 1 on input X given

that $|\mathcal{G}| = g \in \{0, 1, \ldots, m\}$. Similarly, we denote $\mathsf{P}[D'(X) \,|\, g, i] := \mathsf{P}[D'(X) = 1 \,|\, |\mathcal{G}| = g \wedge i^* = i]$ when additionally conditioned on $i^* = i$. Then,

$$\begin{aligned}\Delta^{D'}(G(X'), U') &= |\mathsf{P}[D'(G(X')) = 1] - \mathsf{P}[D'(U') = 1]| \\ &= \left| \sum_{g=0}^{m} \mathsf{P}_{|\mathcal{G}|}(g) \left(\mathsf{P}[D'(G(X')) \,|\, g] - \mathsf{P}[D'(U') \,|\, g]\right) \right| \\ &= \left| \sum_{g=0}^{m} \mathsf{P}_{|\mathcal{G}|}(g) \cdot \frac{1}{m} \sum_{i^*=1}^{m} \left(\mathsf{P}[D'(G(X')) \,|\, g, i^*] - \mathsf{P}[D'(U') \,|\, g, i^*]\right) \right|\end{aligned}$$

By construction $\mathsf{P}[D'(G(X')) \,|\, g, i^*] = \mathsf{P}[D'(U') \,|\, g, i^* - 1]$ for $g \in \{1, \ldots, m\}$ and $i^* = \{2, \ldots, m\}$, and we hence obtain

$$\Delta^{D'}(G(X'), U') = \frac{1}{m} \left| \sum_{g=0}^{m} \mathsf{P}_{|\mathcal{G}|}(g) \cdot \left(\mathsf{P}[D'(G(X')) \,|\, g, 1] - \mathsf{P}[D'(U') \,|\, g, m]\right) \right|$$

On the one hand, we now remark that

$$\sum_{g=0}^{m} \mathsf{P}_{|\mathcal{G}|}(g) \cdot \mathsf{P}[D'(G(X')) \,|\, g, 1] = \mathsf{P}[D(\mathsf{CaE}^{G,\mathsf{Ext}}(X_1 \| \ldots \| X_m \| R)) = 1].$$

On the other hand, because $\mu(\mathcal{N}) \geq 1 - \epsilon$, whenever $g \geq (1 - \epsilon - \rho)m$ and $i^* = m$, the distribution of $y_1 \| \ldots \| y_m$ belongs to an $\left(m, \ell, (1 - \epsilon - \rho)\left(\ell - \log\left(\frac{1}{1-\epsilon}\right)\right)\right)$-total-entropy independent source, and as Ext is a δ-extractor for this source, we obtain $|\mathsf{P}[D'(U') \,|\, g, m] - \mathsf{P}[D(U_{n+d}) = 1]| \leq \delta$, whereas $\mathsf{P}[|\mathcal{G}| < (1 - \epsilon - \rho)m] < e^{-\rho^2 m}$ by Hoeffding's inequality (Lemma 5) and that fact that $\mu(\mathcal{M}) \geq 1 - \epsilon$. We can finally infer

$$\Delta^{D'}(G(X'), U') \geq \frac{\Delta^{D}(\mathsf{CaE}^{G,\mathsf{Ext}}(X_1 \| \ldots \| X_m \| R), U_{n+d}) - \delta - e^{-\rho^2 m}}{m} > \frac{\eta}{m}$$

by our assumption on D.

As the queries of D' do not depend on the inputs, and the above lower bound on its advantage holds for all measures $\mathcal{M}$ and $\mathcal{N}$ with density at least $1 - \epsilon$, the distinguisher D' contradicts Theorem 5 for $\gamma := \frac{\eta}{m}$, which is noticeable, and implies that G is *not* an ϵ-PRG, which is a contradiction.

4.3 Optimality of the Output Length

This final section discusses the optimality of the output length of the concatenate-and-extract construction with respect to the class of constructions which operate by combining a number of independent outputs from weak PRGs, and such that the corresponding security reduction is black-box. In particular, the reduction

only exploits the capability of efficiently sampling a given distribution.[9] This is formally summarized by the following definition.

Definition 1. *A* black-box (ℓ, ϵ)-indistinguishability amplifier *consists a pair of polynomial-time algorithms* (C, S) *with the following two properties:*

(i) For some functions m, d, *and* h, *the algorithm* C *implements a function family* $\left(\{0,1\}^{\ell}\right)^{m} \times \{0,1\}^{d} \to \{0,1\}^{h}$, *where the second input parameter models explicitly the* d*-bit randomness used by the algorithm* C.

(ii) Let P_X *be an arbitrary distribution on the* ℓ*-bit strings which is sampled by an algorithm* X, *let* $X_1, \ldots, X_m$ *be independent samples of* P_X, *and let* R *and* U_h *be uniformly distributed* d*- and* h*-bit strings, respectively. Then, for every distinguisher* D *such that*

$$\Delta^{D}(\mathsf{C}(X_1, \ldots, X_m, R), U_h) > \gamma$$

for infinitely many values of the security parameter and a noticeable function γ, *we have* $\Delta^{\mathsf{S}^{D,\mathsf{X}}}(X, U_\ell) > \epsilon$ *for infinitely many values of the security parameter, where* $X \xleftarrow{\$} \mathsf{P}_X$ *and* U_ℓ *is a uniform* ℓ*-bit string.*

The following theorem (proven in the full version) shows that the output length achieved by concatenate-and-extract is essentially optimal.

Theorem 7. *For all* $\ell \in \mathbb{N}$, *for all constants* $0 < \rho < \epsilon < 1$, *there exists no black-box* (ℓ, ϵ)*-indistinguishability amplifier if* $h \geq (1-\epsilon+\rho) \cdot m \cdot \left[\ell - \log\left(\frac{1}{1-\epsilon}\right)\right] + d + 1$.

Acknowledgments. We thank Russell Impagliazzo for helpful discussions. This research was partially supported by the Swiss National Science Foundation (SNF), project no. 200020-113700/1.

References

1. C. H. Bennett, G. Brassard, and J.-M. Robert, "Privacy amplification by public discussion," *SIAM Journal on Computing*, vol. 17, no. 2, pp. 210–229, 1988.
2. M. Blum and S. Micali, "How to generate cryptographically strong sequences of pseudo random bits," in *FOCS '82: Proceedings of the 23rd IEEE Annual Symposium on Foundations of Computer Science*, pp. 112–117, 1982.
3. Y. Dodis, R. Impagliazzo, R. Jaiswal, and V. Kabanets, "Security amplification for interactive cryptographic primitives," in *Theory of Cryptography — TCC 2009*, vol. 5444 of *Lecture Notes in Computer Science*, pp. 128–145, 2009.
4. O. Goldreich, N. Nisan, and A. Wigderson, "On Yao's XOR-lemma," *Electronic Colloquium on Computational Complexity (ECCC)*, vol. 2, no. 50, 1995.

[9] In particular, note that the proof itself uses black-box access to some function sampling the PRG output which is not required to be expanding. All known proofs have this form.

5. I. Haitner, D. Harnik, and O. Reingold, "On the power of the randomized iterate," in *Advances in Cryptology — CRYPTO 2006*, vol. 4117 of *Lecture Notes in Computer Science*, pp. 22–40, 2006.
6. J. Håstad, R. Impagliazzo, L. A. Levin, and M. Luby, "A pseudorandom generator from any one-way function," *SIAM Journal on Computing*, vol. 28, no. 4, pp. 1364–1396, 1999.
7. W. Hoeffding, "Probability inequalities for sums of bounded random variables," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 13–30, 1963.
8. T. Holenstein, "Key agreement from weak bit agreement," in *STOC '05: Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pp. 664–673, 2005.
9. T. Holenstein, "Pseudorandom generators from one-way functions: A simple construction for any hardness," in *Theory of Cryptography — TCC 2006*, vol. 3876 of *Lecture Notes in Computer Science*, pp. 443–461, 2006.
10. R. Impagliazzo, "Hard-core distributions for somewhat hard problems," in *FOCS '95: Proceedings of the 36th IEEE Annual Symposium on Foundations of Computer Science*, pp. 538–545, 1995.
11. R. Impagliazzo, L. A. Levin, and M. Luby, "Pseudo-random generation from one-way functions (extended abstracts)," in *STOC '89: Proceedings of the 21st Annual ACM Symposium on Theory of Computing*, pp. 12–24, 1989.
12. J. Kamp, A. Rao, S. Vadhan, and D. Zuckerman, "Deterministic extractors for small-space sources," in *STOC '06: Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, pp. 691–700, 2006.
13. U. Maurer, K. Pietrzak, and R. Renner, "Indistinguishability amplification," in *Advances in Cryptology — CRYPTO 2007*, vol. 4622 of *Lecture Notes in Computer Science*, pp. 130–149, Aug. 2007.
14. U. Maurer and S. Tessaro, "Computational indistinguishability amplification: Tight product theorems for system composition," in *Advances in Cryptology — CRYPTO 2009*, vol. 5677 of *Lecture Notes in Computer Science*, pp. 350–368, Aug. 2009.
15. S. Myers, "Efficient amplification of the security of weak pseudo-random function generators," *Journal of Cryptology*, vol. 16, pp. 1–24, 2003.
16. R. Shaltiel, "Recent developments in explicit constructions of extractors," *Bulletin of the EATCS*, vol. 77, pp. 67–95, 2002.
17. A. C. Yao, "Theory and applications of trapdoor functions," in *FOCS '82: Proceedings of the 23rd IEEE Annual Symposium on Foundations of Computer Science*, pp. 80–91, 1982.

A Tail Estimates

The following well-known result from probability theory [7] is repeatedly used throughout this paper.

Lemma 5 (Hoeffding's Inequalities). *Let $X_1, \ldots, X_m$ be independent random variables with range $[0,1]$, and let $\overline{X} := \frac{1}{m}\sum_{i=1}^{m} X_i$. Then, for all $\rho > 0$ we have*

$$\mathsf{P}[\overline{X} \geq \mathsf{E}[\overline{X}] + \rho] \leq e^{-m\rho^2} \text{ and } \mathsf{P}[\overline{X} \leq \mathsf{E}[\overline{X}] - \rho] \leq e^{-m\rho^2}.$$

In particular,

$$\mathsf{P}\left[\left|\overline{X} - \mathsf{E}[\overline{X}]\right| \geq \rho\right] \leq 2 \cdot e^{-m\rho^2}.$$