# On Yao's XOR-Lemma

Oded Goldreich, Noam Nisan, and Avi Wigderson

**Abstract.** A fundamental lemma of Yao states that computational weak-unpredictability of Boolean predicates is amplified when the results of several independent instances are XOR together. We survey two known proofs of Yao's Lemma and present a third alternative proof. The third proof proceeds by first proving that a function constructed by *concatenating* the values of the original function on several independent instances is much more unpredictable, with respect to specified complexity bounds, than the original function. This statement turns out to be easier to prove than the XOR-Lemma. Using a result of Goldreich and Levin (1989) and some elementary observation, we derive the XOR-Lemma.

**Keywords:** Yao's XOR Lemma, Direct Product Lemma, One-Way Functions, Hard-Core Predicates, Hard-Core Regions.

An early version of this survey appeared as TR95-050 of *ECCC*, and was revised several times (with the latest revision posted in January 1999). Since the first publication of this survey, Yao's XOR Lemma has been the subject of intensive research. The current revision contains a short review of this research (see Section 7), but the main text (i.e., Sections 1–6) is *not* updated according to these subsequent discoveries. The current version also include a new appendix (Appendix B), which discusses a variant of the XOR Lemma, called the Selective XOR Lemma.

## 1 Introduction

A fundamental lemma of Yao states that computational weak-unpredictability of Boolean predicates is amplified when the results of several independent instances are XOR together. Indeed, this is analogously to the information theoretic wire-tape channel Theorem (cf., Wyner), but the computational analogue is significanly more complex.

Loosely speaking, by weak-unpredictability we mean that any efficient algorithm will fail to guess the value of the function with probability beyond a stated bound, where the probability is taken over all possible inputs (say, with uniform probability distribution). In particular, the lemma known as Yao's XOR Lemma asserts that if the predicate $f$ is weakly-unpredictable (within some complexity bound), then for sufficiently large $t$ (which depends on the bound) the predicate $F(x_1, ..., x_t) \stackrel{\text{def}}{=} \oplus_{i=1}^{t} f(x_i)$ is almost unpredictable within a related complexity bound (i.e., algorithms of this complexity cannot do substantially better than flip a coin for the answer).

O. Goldreich et al.: Studies in Complexity and Cryptography, LNCS 6650, pp. 273–301, 2011.

Yao stated the XOR Lemma in the context of one-way functions, where the predicate $f$ is the composition of an easy to compute Boolean predicate and the inverse of the one-way function (i.e., $f(x) = b(g^{-1}(x))$, where $g$ is a 1-1 one-way function and $b$ is an easy to compute predicate). Clearly, this is a special case of the setting described above. Yet, the XOR Lemma is sometimes used within the more general setting (under the false assumption that proofs for this setting have appeared in the literature). Furthermore, in contrary to common beliefs, the lemma itself has not appeared in Yao's original paper "Theory and Applications of Trapdoor Functions" [17] (but rather in oral presentations of his work).

A proof of Yao's XOR Lemma has first appeared in Levin's paper [12]. Levin's proof is for the context of one-way functions and is carried through in a uniform model of complexity. The presentation of this proof in [12] is very succinct and does not decouple the basic approach from difficulties arising from the uniform-complexity model. In Section 3, we show that Levin's basic approach suffices for the general case (mentioned above) provided it is stated in terms of non-uniform complexity. The proof also extends to a uniform-complexity setting, provided that some sampling condition (which is satisfied in the context of one-way functions) holds. We do not know whether the XOR Lemma holds in the uniform-complexity model in case this sampling condition is not satisfied.

Recently, Impagliazzo has shown that, in the non-uniform model, any weakly-unpredictable predicate has a "hard-core"[1] on which it is almost unpredictable [7]. Using this result, Impagliazzo has presented an alternative proof for the general case of the XOR-Lemma (within the non-uniform model). We present this proof in Section 4.

A third proof for the general case of the XOR-Lemma is presented in Section 5. This proof proceeds by first proving that a function constructed by *concatenating* the values of the predicate on several independent instances is much more unpredictable, with respect to specified complexity bounds, than the original predicate. Loosely speaking, it is hard to predict the value of the function with probability substantially higher than $\delta^t$, where $\delta$ is a bound on the probability of predicting the predicate and $t$ is the number of instances concatenated. Not surprisingly, this statement turns out to be easier to prove than the XOR-Lemma. Using a result of Goldreich and Levin [5] and some elementary observation, we derive the XOR-Lemma.

We remark that Levin's proof yields a stronger quantitative statement of the XOR Lemma than the other two proofs. In fact, the quantitative statement provided by Levin's proof is almost optimal. Both Levin's proof and our proof can be transformed to the uniform-complexity provided some natural sampling condition holds. We do not know how to transform Impagliazzo's proof to the uniform-complexity setting, even under this condition.

[1] Here the term 'hard-core' means a subset of the predicate's domain. This meaning is certainly different from the usage of the term 'hard-core' in [5], where it means a strongly-unpredicatable predicate associated with a one-way function.

A different perspective on the concatenating problem considered above is presented in Section 6, where we consider the conditional entropy of the function's value given the result of a computation (rather than the probability that the two agree).

## 2 Formal Setting

We present a general framework, and view the context of one-way functions as a specail case. The general framework is presented in term of non-uniform complexity, but uniformity conditions can be added in.

### 2.1 The Basic Setting

The basic framework consists of a Boolean predicate $f : \{0,1\}^* \to \{0,1\}$ and a non-uniform complexity class such as $\mathcal{P}$/poly. Specifically, we consider all families of polynomial-size circuits and for each family, $\{C_n\}$, we consider the probability that it correctly computes $f$, where the probability is taken over all $n$-bit inputs with uniform probability distribution. Alternatively, one may consider the most successful $n$-bit input circuit among all circuits of a given size. This way we obtain a bound on unpredictability of $f$ with respect to a specific complexity class.

In the sequel, it will be more convenient to redefine $f$ as mapping bit string into $\{\pm 1\}$ and to consider the correlation of a circuit (outputting a value in $\{\pm 1\}$) with the value of the function (i.e., redefine $f(x) \stackrel{\text{def}}{=} (-1)^{f(x)}$).[2] Using this notation allows to replace $\text{Prob}[C(X) = f(X)]$ by $(1 + \text{E}[C(X) \cdot f(X)])/2$, by noting that $\text{E}[C(X) \cdot f(X)] = \text{Prob}[C(X) = f(X)] - \text{Prob}[C(X) \neq f(X)]$.

We also generalize the treatment to arbitrary distributions over the set of $n$-bit long inputs (rather than uniform ones) and to "probabilistic" predicates (or processes) that on input $x$ return some distribution on $\{\pm 1\}$; that is, for a fixed $x$, we let $f(x)$ be a random variable distributed over $\{\pm 1\}$ (rather than a fixed value). One motivation for this generalization is that it allows us to treat as a special case 'hard predicates' of one-way functions, when the functions are not necessarily 1-1.

**Definition 1** (algorithmic correlation): *Let $P$ be a randomized process/ algorithm that maps bit strings into values in $\{\pm 1\}$ and let $\mathbf{X} \stackrel{\text{def}}{=} \{X_n\}$ be a probability ensemble such that, for each $n$, the random variable $X_n$ is distributed over $\{0,1\}^n$. The* correlation *of a circuit family $\mathbf{C} = \{C_n\}$ with $P$ over $\mathbf{X}$ is defined as $c : \mathbb{N} \to \mathbb{R}$ such that*

$$c(n) \stackrel{\text{def}}{=} \text{E}[C_n(X_n) \cdot P(X_n)],$$

[2] This suggestion, of replacing the standard $\{0,1\}$ by $\{\pm 1\}$ and using correlations rather than probabilities, is due to Levin. It is indeed amazing how this simple change of notation simplifies both the statements and the proofs.

*where the expectation is taken over the random variable $X_n$ (and the process $P$). We say that a complexity class* (i.e., a set of circuit families) *has* correlation *at most* $c(\cdot)$ *with* $P$ *over* $\mathbf{X}$ *if, for every circuit family* $\mathbf{C}$ *in this class, the correlation of* $\mathbf{C}$ *with* $P$ *over* $\mathbf{X}$ *is bounded by* $c(\cdot)$.

The foregoing definition may be used to discuss both uniform and non-uniform complexity classes. In the next subsection we relate the Definition 1 to the standard treatment of unpredictability within the context of one-way functions.

### 2.2 The Context of One-Way Functions

For sake of simplicity, we consider only length-preserving functions (i.e., functions $f : \{0,1\}^* \rightarrow \{0,1\}^*$ satisfying $|f(x)| = |x|$ for all $x$). A one-way function $f : \{0,1\}^* \rightarrow \{0,1\}^*$ is a function that is easy to compute but hard to invert. Namely, there exists a polynomial-time algorithm for computing $f$, but for any probabilistic polynomial-time[3] algorithm $A$, the probability that $A(f(x))$ is a preimage of $f(x)$ is negligible (i.e., smaller than $1/p(|x|)$ for any positive polynomial $p$), where the probability is taken uniformly over all $x \in \{0,1\}^n$ and all possible internal coin tosses of algorithm $A$.

Let $b : \{0,1\}^* \rightarrow \{\pm 1\}$ be an easy to compute predicate and let $\delta : \mathbb{N} \rightarrow \mathbb{R}$. The predicate $b$ is said to be at most $\delta$-correlated to $f$ in polynomial-time if for any probabilistic polynomial-time algorithm $G$, the expected correlation of $G(f(x))$ and $b(x)$, is at most $\delta(n)$ (for all but finitely many $n$'s). (Again, the probability space is uniform over all $x \in \{0,1\}^n$ and all possible internal coin tosses of the algorithm.) Thus, although $b$ is easy to evaluate (i.e., the mapping $x \mapsto b(x)$ is polynomial-time computable), it is hard to predict $b(x)$ from $f(x)$, for a random $x$.

Let us relate the latter notion to Definition 1. Suppose, first, that $f$ is 1-1. Then, saying that $b$ is at most $\delta$-correlated to $f$ in polynomial-time is equivalent to saying that the class of (probabilistic) polynomial-time algorithms has correlation at most $\delta(\cdot)$ with the predicate $P(x) \stackrel{\text{def}}{=} b(f^{-1}(x))$, over the uniform distribution. Note that if $f$ is polynomial-time computable and $b$ is at most $(1 - (1/\text{poly}))$-correlated to $f$ in polynomial-time, then $f$ must be one-way (because otherwise $b(x)$ can be correlated too well by first obtaining $f^{-1}(x)$ and then evaluating $b$),

The treatment can be extended to arbitrary one-way functions, which are not necessarily 1-1. Let $f$ be such a function and $b$ a predicate that is at most $\delta$-correlated to $f$ (by polynomial-time algorithms). Define the probability ensemble $\mathbf{X} = \{X_n\}$ by letting $X_n = f(r)$, where $r$ is uniformly selected in $\{0,1\}^n$, and define the randomized process $P(x)$ by uniformly selecting $r \in f^{-1}(x)$ and outputting $b(r)$. Now, it follows that the class of (probabilistic) polynomial-time algorithms has correlation at most $\delta(\cdot)$ with the predicate $P$ over $\mathbf{X}$.

[3] Here we adopt the standard definition of one-way function; however, our treatment applies also to the general definition where inverting is infeasible with respect to a specified time bound and success probability.

### 2.3 Getting Random Examples

An important issue regarding the general setting, is whether it is possible to obtain random examples of the distribution $(X_n, P(X_n))$. Indeed, random examples are needed in all known proofs of the XOR Lemma (i.e., they are used in the algorithms deriving a contradiction to the difficulty of correlating the basic predicate).[4] Other than this aspect (i.e., the use of random examples), two of the three proofs can be adapted to the uniform-complexity setting (see Section 2.5).

Note that in the context of one-way functions such random examples can be generated by a probabilistic polynomial-time algorithm. Specifically, although the corresponding $P$ is assumed not to be polynomial-time computable, it is easy to generate randomly pairs $(x, P(x))$ for $x \leftarrow X_n$. (This is done, by uniformly selecting $r \in \{0,1\}^n$, and outputting the pair $(f(r), b(r)) = (f(r), P(f(r)))$.) Thus, we can prove the XOR Lemma in the (uniform-complexity) context of one-way functions.

We also note that the effect of random examples can be easily simulated by non-uniform polynomial-size circuits (i.e., random examples can be hard-wired into the circuit). Thus, we can prove the XOR Lemma in the general non-uniform complexity setting.

### 2.4 Three (Non-uniform) Forms of the XOR Lemma

Following the description in the introduction (and Yao's expositions), the basic form of the XOR Lemma states that the tractable algorithmic correlation of the XOR-predicate $P^{(t)}(x_1, ..., x_t) \stackrel{\text{def}}{=} \prod_{i=1}^{t} P(x_i)$ decays exponentially with $t$ (upto a negligible fraction). Namely:

**Lemma 1** (XOR Lemma – Yao's version): *Let $P$ and $\mathbf{X} = \{X_n\}$ be as in Definition 1. For every function $t:\mathbb{N}\to\mathbb{N}$, define the predicate*

$$P^{(t)}(x_1, ..., x_{t(n)}) \stackrel{\text{def}}{=} \prod_{i=1}^{t(n)} P(x_i) \ ,$$

*where $x_1, ..., x_{t(n)} \in \{0,1\}^n$, and let $\mathbf{X}^{(t)} \stackrel{\text{def}}{=} \{X_n^{(t)}\}$ be a probability ensemble such that $X_n^{(t)}$ consists of $t(n)$ independent copies of $X_n$.*

(hypothesis) *Let $s:\mathbb{N}\to\mathbb{N}$ be a size function, and $\delta:\mathbb{N}\to[-1,+1]$ be a function that is* bounded-away-from-1 (i.e., $|\delta(n)| < 1 - \frac{1}{p(n)}$, for some polynomial $p$ and all sufficiently large $n$'s). *Suppose that $\delta$ is an upper bound on the correlation of families of $s(\cdot)$-size circuits with $P$ over $\mathbf{X}$.*

(conclusion) *Then, there exists a bounded-away-from-1 function $\delta':\mathbb{N}\to[-1,+1]$ and a polynomial $p$ such that, for every function $t:\mathbb{N}\to\mathbb{N}$ and every function $\epsilon:\mathbb{N}\to[0,1]$, the function*

$$\delta^{(t)}(n) \stackrel{\text{def}}{=} p(n) \cdot \delta'(n)^{t(n)} + \epsilon(n)$$

[4] This assertion refers to what was known at the time this survey was written. As noted in Section 7, the situation regarding this issue has changed recently.

*is an upper bound on the correlation of families of* $s'(\cdot)$*-size circuits with* $P^{(t)}$ *over* $\mathbf{X}^{(t)}$*, where*

$$s'(t(n)\cdot n) \stackrel{\text{def}}{=} \text{poly}\left(\frac{\epsilon(n)}{n}\right)\cdot s(n) - \text{poly}(n\cdot t(n)).$$

All three proofs presented below establish Lemma 1. The later two proofs do so for various values of $\delta'$ and $p$; that is, in Impagliazzo's proof (see Section 4) $\delta'(n) = \frac{1+\delta(n)}{2} + o(1-\delta(n))$ and $p(n) = 2$, whereas in our proof (see Section 5) $\delta'(n) = \sqrt[3]{\frac{1+\delta(n)}{2}}$ and $p(n) = o(n)$. Levin's proof (see Section 3) does even better; it establishes the following:

**Lemma 2** (XOR Lemma – Levin's version): *Yao's version holds with* $\delta' = \delta$ *and* $p = 1$.

Lemma 2 still contains some slackness; specifically, the closest one wants to get to the "obvious" bound of $\delta^{(t)}(n) = \delta(n)^{t(n)}$, the more one losses in terms of the complexity bounds (i.e., bounds on circuit size).[5] In particular, if one wishes to have $s'(t(n)\cdot n) = \frac{s(n)}{\text{poly}(n)}$, then one can only get a result for $\epsilon(n) = 1/\text{poly}(n)$ (i.e., get $\delta^{(t)}(n) = \delta(n)^{t(n)} + 1/p(n)$, for any polynomial $p$). We do not know how to remove this slackness. We even do not know if it can be reduced "a little" as follows.

**Lemma 3** (XOR Lemma – dream version – a conjecture): *For some fixed negligible function* $\mu$ (e.g., $\mu(n) \stackrel{\text{def}}{=} 2^{-n}$ or even $\mu(n) \stackrel{\text{def}}{=} 2^{-(\log_2 n)^2}$)*, Yao's version holds with* $\delta^{(t)}(n) = \delta'(n)^{t(n)} + \mu(n)$*, and* $s'(t(n)\cdot n) = \frac{s(n)}{\text{poly}(n)}$.

Steven Rudich has observed that the Dream Version does not hold in a relativized world. Specifically, his argument proceeds as follows. Fix $\mu$ as in the Dream Version and set $t$ such that $\delta^{(t)} < 2\mu(n)$. Consider an oracle that for every $(x_1, ..., x_{t(n)}) \in (\{0,1\}^n)^{t(n)}$ and for a $2\mu(n)$ fraction of the $r$'s in $\{0,1\}^n$, answers the query $(x_1, ..., x_{t(n)}, r)$ with $(P(x_1), ..., P(x_t))$, otherwise the oracle answers with a special symbol. These $r$'s may be selected at random (thus constructing a random oracle). The hypothesis of the lemma may hold relative to this oracle, but the conclusion cannot possibly hold. Put differently, one can argue that there is no (polynomial-time) "black-box" reduction of the task of correlating $P$ (by at least $\delta$) to the task of correlating $P^{(t)}$ (by at least $\mu$). The reason being that the polynomial-time machine (effecting this reduction) cannot distinguish a black-box of negligible correlation (i.e., correlation $2\mu$) from a black-box of zero correlation.

### 2.5 Uniform Forms of the XOR Lemma

So far, we have stated three forms of the XOR Lemma in terms of non-uniform complexity. Analogous statements in terms of uniform complexity can be made

[5] I.e., $\delta^{(t)}(n) = \delta'(n)^{t(n)} + \epsilon(n)$ is achieved for $s'(t(n)\cdot n) = \text{poly}(\epsilon(n)/n)\cdot s(n)$.

as well. These statements relate to the time required to construct the circuits in the hypothesis and those in the conclusion. For example, one may refer to circuit families, $\{C_n\}$, for which, given $n$, the circuit $C_n$ can be constructed in poly($|C_n|$)-time. In addition, all functions referred to in the statement of the lemma (i.e., $s, t : \mathbb{N} \to \mathbb{N}$, $\delta : \mathbb{N} \to [-1, +1]$ and $\epsilon : \mathbb{N} \to [-1, +1]$) need to be computable within corresponding time bounds. Such analogues of the two first versions can be proven, provided that one can construct random examples of the distribution $(X_n, P(X_n))$ within the stated (uniform) complexity bounds (and in particular in polynomial-time). See Section 2.3 as well as comments in the subsequent sections.

# 3 Levin's Proof

The key ingredient in Levin's proof is the following lemma, which provides an accurate account of the decrease of the computational correlation in the case that two predicates are xor-ed together. It should be stressed that the statement of the lemma is intentionally asymmetric with respect to the two predicates.

**Lemma 4** (Isolation Lemma): *Let $P_1$ and $P_2$ be two predicates, $l : \mathbb{N} \to \mathbb{N}$ be a length function, and $P(x) \stackrel{\text{def}}{=} P_1(y) \cdot P_2(z)$ where $x = yz$ and $|y| = l(|x|)$. Let $\mathbf{X} = \{X_n\}$ be a probability ensemble such that the first $l(n)$ bits of $X_n$ are statistically independent of the rest, and let $\mathbf{Y} = \{Y_{l(n)}\}$ (resp., $\mathbf{Z} = \{Z_{n-l(n)}\}$) denote the projection of $\mathbf{X}$ on the first $l(\cdot)$ bits (resp., last $n - l(n)$ bits).*

(hypothesis) *Suppose that $\delta_1(\cdot)$ is an upper bound on the correlation of families of $s_1(\cdot)$-size circuits with $P_1$ over $\mathbf{Y}$, and that $\delta_2(\cdot)$ is an upper bound on the correlation of families of $s_2(\cdot)$-size circuits with $P_2$ over $\mathbf{Z}$.*

(conclusion) *Then, for every function $\epsilon : \mathbb{N} \to \mathbb{R}$, the function*

$$\delta(n) \stackrel{\text{def}}{=} \delta_1(l(n)) \cdot \delta_2(n - l(n)) + \epsilon(n)$$

*is an upper bound on the correlation of families of $s(\cdot)$-size circuits with $P$ over $\mathbf{X}$, where*

$$s(n) \stackrel{\text{def}}{=} \min\left\{ \frac{s_1(l(n))}{\text{poly}(n/\epsilon(n))} \;,\; s_2(n - l(n)) - n \right\}$$

The lemma is asymmetric with respect to the dependency of $s(\cdot)$ on the $s_i$'s. The fact that $s(\cdot)$ maybe almost equal to $s_2(\cdot)$ plays a central role in deriving the XOR Lemma from the Isolation Lemma.

## 3.1 Proof of the Isolation Lemma

Assume, towards the contradiction, that a circuit family $\mathbf{C}$ (of size $s(\cdot)$) has correlation greater than $\delta(\cdot)$ with $P$ over $\mathbf{X}$. Thus, denoting by $Y_l$ (resp., $Z_m$)

the projection of $X_n$ on the first $l \stackrel{\text{def}}{=} l(n)$ bits (resp., last $m \stackrel{\text{def}}{=} n - l(n)$ bits), we get

$$\begin{aligned}\delta(n) &< \mathrm{E}[C_n(X_n) \cdot P(X_n)] \\ &= \mathrm{E}[C_n(Y_l, Z_m) \cdot P_1(Y_l) \cdot P_2(Z_m)] \\ &= \mathrm{E}[P_1(Y_l) \cdot \mathrm{E}[C_n(Y_l, Z_m) \cdot P_2(Z_m)]]\end{aligned}$$

where, in the last expression, the outer expectation is over $Y_l$ and the inner one is over $Z_m$. For every fixed $y \in \{0,1\}^l$, let

$$T(y) \stackrel{\text{def}}{=} \mathrm{E}[C_n(y, Z_m) \cdot P_2(Z_m)]. \tag{1}$$

Then, by the foregoing,

$$\mathrm{E}[T(Y_l) \cdot P_1(Y_l)] > \delta(n). \tag{2}$$

We shall see that Eq. (2) either contradicts the hypothesis concerning $P_2$ (see Claim 4.1) or contradicts the hypothesis concerning $P_1$ (by a slightly more involved argument).

Claim 4.1: For all but finitely many $n$'s and every $y \in \{0,1\}^l$

$$|T(y)| \leq \delta_2(m).$$

Proof: Otherwise, fixing a $y$ contradicting the claim, we get a circuit $C'_m(z) \stackrel{\text{def}}{=} C_n(y, z)$ of size $s(n) + l < s_2(m)$, having greater correlation with $P_2$ than that allowed by the lemma's hypothesis. □

By Claim 4.1, the value $T(y)/\delta_2(m)$ lies in the interval $[-1, +1]$; while, on the other hand (by Eq. (2)), it (i.e., $T(\cdot)/\delta_2(m)$) has good correlation with $P_1$. In the rest of the argument we "transform" the function $T$ into a circuit which contradicts the hypothesis concerning $P_1$. Suppose for a moment, that one could compute $T(y)$, on input $y$. Then, one would get an algorithm with output in $[-1, +1]$ that has correlation at least $\delta(n)/\delta_2(m) > \delta_1(l)$ with $P_1$ over $Y_l$, which is almost in contradiction to the hypothesis of the lemma.[6] The same holds if one can approximate $T(y)$ "well enough" using circuits of size $s_1(l)$. Indeed, the lemma follows by observing that such an approximation is possible. Namely:

Claim 4.2: For every $n$, $l = l(n)$, $m = n - l$, $q = \mathrm{poly}(n/\epsilon(n))$ and $y \in \{0,1\}^l$, let

$$\tilde{T}(y) \stackrel{\text{def}}{=} \frac{1}{q} \sum_{i=1}^{q} C_n(y, z_i) \cdot \sigma_i$$

where $(z_1, \sigma_1), ..., (z_q, \sigma_q)$ is a sequence of $q$ independent samples from the distribution $(Z_m, P_2(Z_m))$. Then,

$$\mathrm{Prob}[|T(y) - \tilde{T}(y)| > \epsilon(n)] < 2^{-l(n)}$$

[6] See discussion below; the issue is that the output is in the interval $[-1, +1]$ rather than being a binary value in $\{\pm 1\}$.

Proof: Immediate by the definition of $T(y)$ and application of Chernoff bound.□

Claim 4.2 suggests an approximation algorithm (for the function $T$), where we assume that the algorithm is given as auxiliary input a sequence of samples from the distribution $(Z_m, P_2(Z_m))$. (The algorithm merely computes the average of $C_n(y, z_i) \cdot \sigma_i$ over the sample sequence $(z_1, \sigma_1), ..., (z_q, \sigma_q)$.)

If such a sample sequence can be generated efficiently, by a uniform algorithm (as in the context of one-way functions), then we are done. Otherwise, we use non-uniformity to obtain a fixed sequence that is good for all possible $y$'s. (Such a sequence does exist since with positive probability, a randomly selected sequence, from the above distribution, is good for all $2^{l(n)}$ possible $y$'s.) Thus, there *exists* a circuit of size $\text{poly}(n/\epsilon(n)) \cdot s(n)$ that, on input $y \in \{0,1\}^{l(n)}$, outputs a value $(T(y) \pm \epsilon(n))/\delta_2(m)$.

We note that this output is at least $\frac{\delta(n)}{\delta_2(m)} - \frac{\epsilon(n)}{\delta_2(m)} = \delta_1(l)$ correlated with $P_1$, which almost contradicts the hypothesis of the lemma. The only problem is that the resulting circuit has output in the interval $[-1, +1]$ instead of a binary output in $\{\pm 1\}$. This problem is easily corrected by modifying the circuit so that on output $r \in [-1, +1]$ it outputs $+1$ with probability $(1 + r)/2$ and $-1$ otherwise. Noting that this modification preserves the correlation of the circuit, we derive a contradiction to the hypothesis concerning $P_1$. ■

### 3.2 Proof of Lemma 2

The stronger version of the XOR Lemma (i.e., Lemma 2) follows by a (careful) successive application of the Isolation Lemma. Loosely speaking, we write $P^{(t)}(x_1, x_2, ..., x_{t(n)}) = P(x_1) \cdot P^{(t-1)}(x_2, ..., x_{t(n)})$, assume that $P^{(t-1)}$ is hard to correlate as claimed, and apply the Isolation Lemma to $P \cdot P^{(t-1)}$. This way, the lower bound on the size of circuits correlating $P^{(t)}$ is related to the lower bound assumed for circuits correlating the original $P$, since the lower bound derived for $P^{(t-1)}$ is larger and is almost preserved by the Isolation Lemma (losing only an additive term!).

### 3.3 Remarks Concerning the Uniform Complexity Setting

A uniform-complexity analogue of Lemma 2 can be proven provided that one can construct random examples of the distribution $(X_n, P(X_n))$ within the stated (uniform) complexity bounds. To this end, one should state and prove a uniform-complexity version of the Isolation Lemma, which also assumes that example from both distributions (i.e., $(Y_l, P_1(Y_l))$ and $(Z_m, P_2(Z_m))$)[7] can be generated within the relevant time complexity; certainly, sampleability in probabilistic polynomial-time suffices. Furthermore, in order to derive the XOR Lemma it is important to prove a strong statement regarding the relationship between the time required to construct the circuits referred to in the lemma. Namely:

**Lemma 5** (Isolation Lemma – uniform complexity version): *Let* $P_1, P_2, l, P$, $\mathbf{X}, \mathbf{Y}$ *and* $\mathbf{Z}$ *be as in Lemma 4.*

[7] Actually, it suffices to be able to sample the distributions $Y_l$ and $(Z_m, P_2(Z_m))$.

(hypothesis) *Suppose that* $\delta_1(\cdot)$ *(resp.,* $\delta_2$*) is an upper bound on the correlation of* $t_1(\cdot)$*-time-constructible families of* $s_1(\cdot)$*-size (resp.,* $t_2(\cdot)$*-time-constructible families of* $s_2(\cdot)$*-size) circuits with* $P_1$ *over* $\mathbf{Y}$ *(resp.,* $P_2$ *over* $\mathbf{Z}$*). Furthermore, suppose that one can generate in polynomial-time a random sample from the distribution* $(Y_l, Z_m, P_2(Z_m))$*.*

(conclusion) *Then, for every function* $\epsilon{:}\mathbb{N}\to\mathbb{R}$*, the function*

$$\delta(n) \stackrel{\text{def}}{=} \delta_1(l(n)) \cdot \delta_2(n - l(n)) + \epsilon(n)$$

*is an upper bound on the correlation of* $t(\cdot)$*-time-constructible families of* $s(\cdot)$*-size circuits with* $P$ *over* $\mathbf{X}$*, where*

$$s(n) \stackrel{\text{def}}{=} \min\left\{\frac{s_1(l(n))}{\text{poly}(n/\epsilon(n))} \;\; , \; s_2(n-l(n)) - n\right\}$$

$$t(n) \stackrel{\text{def}}{=} \min\left\{t_1(l(n)) \;\; , \; t_2(n-l(n))\right\} - \text{poly}(n/\epsilon(n)) \cdot s(n).$$

The uniform-complexity version of the Isolation Lemma is proven by adapting the proof of Lemma 4 as follows. First, a weaker version of Claim 4.1 is stated, asserting that (for all but finitely many $n$'s) it holds that

$$\text{Prob}[|T(Y_l)| > \delta_2(m) + \epsilon'(n)] < \epsilon'(n),$$

where $\epsilon'(n) \stackrel{\text{def}}{=} \epsilon(n)/3$. The new claim is valid, since otherwise, one can find in $\text{poly}(n/\epsilon(n))$-time a $y$ violating it; to this end we need to sample $Y_l$ and, for each sample $y$, approximate the value of $T(y)$ (by using $\text{poly}(n/\epsilon(n))$ samples of $(Z_m, P_2(Z_m))$). Once a good $y$ is found, we incorporate it in the construction of $C_n$, obtaining a circuit that contradicts the hypothesis concerning $P_2$. (We stress that we have presented an efficient algorithm for constructing a circuit for $P_2$, given an algorithm that constructs the circuit $C_n$. Furthermore, the running time of our algorithm is the sum of the time required to construct $C_n$ and the time required for sampling $(Z_m, P_2(Z_m))$ sufficiently many times and for evaluating $C_n$ on sufficiently many instances.)

Clearly, Claim 4.2 remains unchanged (except for the replacing $\epsilon(n)$ by $\epsilon'$). Using the hypothesis that samples from $(Z_m, P_2(Z_m))$ can be efficiently generated, we can construct a circuit for correlating $P_1$ within time $t(n)+\text{poly}(n/\epsilon(n))\cdot(n+s(n))$. This circuit is merely an approximater of the function $T$, which operates by averaging (as in Claim 4.2); this circuit is constructed by first constructing $C_n$, generating $\text{poly}(n/\epsilon(n))$ samples of $(Z_m, P_2(Z_m))$ and incorporating them in corresponding copies of $C_n$ – thus justifying the above time and size bounds. However, unlike in the non-uniform case, we are not guaranteed that $|T(y)|$ is bounded above (by $\delta_2(m) + \epsilon'(n)$) for all $y$'s. Yet, if we modify our circuit to do nothing whenever its estimate violates the bound, we loss at most $\epsilon'(n)$ of the correlation and we can proceed as in the non-uniform case.

**Proving a uniform complexity version of Lemma 2:** As in the non-uniform case, the (strong form of the) XOR Lemma follows by a (careful) successive application of the Isolation Lemma. Again, we write $P^{(\tau)}(x_1, x_2, ..., x_{\tau(n)}) = P(x_1) \cdot P^{(\tau-1)}(x_1, ..., x_{\tau(n)-1})$, assume that $P^{(\tau-1)}$ is hard to correlate as claimed, and apply the Isolation Lemma to $P \cdot P^{(\tau-1)}$. This way, the lower bounds on circuits correlating $P^{(\tau)}$ is related to the lower bound assumed for circuits correlating the original $P$ and is almost the bound derived for $P^{(\tau-1)}$ (losing only an additive terms!). This almost concludes the proof, except that we have implicitly assumed that we know the value of $\tau$ for which the XOR Lemma first fails; this value is needed in order to construct the circuit violating the hypothesis for the original $P$. In the non-uniform case this value of $\tau$ can be incorporated into the circuit, but in the uniform-complexity case we need to find it. This is not a big problem as they are only polynomially many possible values and we can test each of them within the allowed time complexity.

## 4 Impagliazzo's Proof

The key ingredient in Impagliazzo's proof is the notion of a hard-core region of a weakly-unpredictable predicate and a lemma that asserts that every weakly-unpredictable predicate has a hard-core region of substantial size.

**Definition 2** (hard-core region of a predicate): *Let $f : \{0,1\}^* \to \{0,1\}$ be a Boolean predicate, $s : \mathbb{N} \to \mathbb{N}$ be a size function, and $\epsilon : \mathbb{N} \to [0,1]$ be a function.*

- *We say that a sequence of sets, $\mathbf{S} = \{S_n \subseteq \{0,1\}^n\}$, is a* hard-core (region) *of $f$ with respect to $s(\cdot)$-size circuits families and advantage $\epsilon(\cdot)$ if for every $n$ and every circuit $C_n$ of size at most $s(n)$, it holds that*

$$\mathrm{Prob}[C_n(X_n) = f(X_n)] \leq \frac{1}{2} + \epsilon(n)$$

*where $X_n$ is a random variable uniformly distributed on $S_n$.*

- *We say that $f$ has a* hard-core (region) *of density $\rho(\cdot)$ with respect to $s(\cdot)$-size circuits families and advantage $\epsilon(\cdot)$ if there exists a sequence of sets $\mathbf{S} = \{S_n \subseteq \{0,1\}^n\}$ such that $\mathbf{S}$ is a hard-core of $f$ with respect to the above and $|S_n| \geq \rho(n) \cdot 2^n$.*

We stress that the usage of the term 'hard-core' in the above definition (and in the rest of this section) is different from the usage of this term in [5]. Observe that every strongly-unpredictable predicate has a hard-core of density 1 (i.e., the entire domain itself). Impagliazzo proves that also weakly-unpredicatabe predicates have hard-core sets that have density related to the amount of unpredictability. Namely:

**Lemma 6** (existence of hard-core regions for unpredictable predicates): *Let $f : \{0,1\}^* \to \{0,1\}$ be a Boolean predicate, $s : \mathbb{N} \to \mathbb{N}$ be a size function, and $\rho : \mathbb{N} \to [0,1]$ be a noticeable function* (i.e., $\rho(n) > 1/\mathrm{poly}(n)$)*, such that for every $n$ and every circuit $C_n$ of size at most $s(n)$ it holds that*

$$\mathrm{Prob}[C_n(U_n) = f(U_n)] \leq 1 - \rho(n),$$

*where $U_n$ is a random variable uniformly distributed on $\{0,1\}^n$. Then, for every function $\epsilon:\mathbb{N}\to[0,1]$, the function $f$ has a hard-core of density $\rho'(\cdot)$ with respect to $s'(\cdot)$-size circuits families and advantage $\epsilon(\cdot)$, where $\rho'(n) \stackrel{\rm def}{=} (1-o(1))\cdot\rho(n)$ and $s'(n) \stackrel{\rm def}{=} s(n)/\mathrm{poly}(n/\epsilon(n))$.*

The proof of Lemma 6 is given in Appendix A. Using Lemma 6, we derive a proof of the XOR-Lemma, for the special case of uniform distribution.

Suppose that $\delta(\cdot)$ is a bound on the correlation of $s(\cdot)$-circuits with $f$ over the uniform distribution. Then, it follows that such circuits cannot guess the value of $f$ better than with probability $p(n) \stackrel{\rm def}{=} \frac{1+\delta(n)}{2}$ and the existence of a hard-core $\mathbf{S} = \{S_n\}$ (w.r.t. $s'(n)$-circuits and $\epsilon(n)$-advantage) with density $\rho'(n) \stackrel{\rm def}{=} (1-o(1))\cdot(1-p(n))$ follows. Clearly,

$$\rho'(n) = (1-o(1))\cdot\frac{1-\delta(n)}{2} > \frac{1}{3}\cdot(1-\delta(n)).$$

Now, suppose that in contradiction to the XOR Lemma, the predicate $F^{(t)}$ defined as $F^{(t)}(x_1,...,x_t) \stackrel{\rm def}{=} \oplus_i f(x_i)$ can be correlated by "small" circuits with correlation greater than $c'(n) \stackrel{\rm def}{=} 2\cdot(\frac{2+\delta(n)}{3})^t + \epsilon(n)$. In other words, such circuits can guess $F^{(t)}$ with success probability at least $\frac{1}{2}+\frac{1}{2}\cdot c'(n)$. However, the probability that none of the $t$ arguments to $F^{(t)}$ falls in the hard-core is at most $(1-\rho'(n))^t$. Thus, conditioned on the event that at least one argument falls in the hard-core $\mathbf{S}$, the circuit guess $F^{(t)}$ correctly with probability at least

$$\frac{1}{2}+\frac{1}{2}\cdot c'(n) - (1-\rho'(n))^t > \frac{1}{2}+\frac{\epsilon(n)}{2}\,.$$

Note, however, that this does not seem to yield an immediate contradition to the definition of a hard-core of $f$, yet we shall see that such a contradiction can be derived.

For every non-empty $I \subseteq \{1,...,t\}$, we consider the event, denoted $E_I$, that represents the case that the arguments to $F^{(t)}$ that fall in the hard-core of $f$ are exactly those with index in $I$. We have just shown that, conditioned on the union of these events, the circuit guesses the predicate $F^{(t)}$ correctly with probability at least $\frac{1}{2}+\frac{\epsilon(n)}{2}$. Thus, there exists an (non-empty) $I$ such that, conditioned on $E_I$, the circuit guesses $F^{(t)}$ correctly with probability at least $\frac{1}{2}+\frac{\epsilon(n)}{2}$. Let $i \in I$ be arbitrary. By another averaging argument, we fix all inputs to the circuit except the $i^{\rm th}$ input and obtain a circuit that guesses $f$ correctly with probability at least $\frac{1}{2}+\frac{\epsilon(n)}{2}$. (For these fixed $x_j$'s, $j \neq i$, the circuit incorporates also the value of $\oplus_{j\neq i} f(x_j)$.) This contradicts the hypothesis that $\mathbf{S}$ is a hard-core.

*Generalization.* We have just established the validity of the Lemma 1 for the case of the uniform probability ensemble and parameters $p(n)=2$ and $\delta'(n)=\frac{2+\delta(n)}{3}$. The bound for $\delta'$ can be improved to $\delta'(n)=\frac{1+\delta(n)}{2}+o(1-\delta(n))$. The argument extends to arbitrary probability ensembles. To this end one needs to properly generalize Definition 2 and prove a generalization of Lemma 6; for details the interested reader is referred to Appendix A.

# 5 Going through the Direct Product Problem

The third proof of the XOR Lemma proceeds in two steps. First it is shown that the success probability of feasible algorithms that try to predict the values of a predicate on several unrelated arguments decreases exponentially with the number of arguments. This statement is a generalization of another theorem due to Yao [17], hereafter called the *Concatenation Lemma.* Invoking a result of Goldreich and Levin [5], the XOR-Lemma follows.

## 5.1 The Concatenation Lemma

(This lemma is currently called the *Direct Product Theorem.*)

**Lemma 7** (concatenation lemma): *Let* $P$*,* $\mathbf{X} = \{X_n\}$*,* $s : \mathbb{N} \to \mathbb{N}$*, and* $\delta : \mathbb{N} \to [-1, +1]$ *be as in Lemma 1. For every function* $t : \mathbb{N} \to \mathbb{N}$*, define the function* $F^{(t)}(x_1, ..., x_{t(n)}) \overset{\text{def}}{=} (P(x_1), ..., P(x_{t(n)}))$*, where* $x_1, ..., x_{t(n)} \in \{0,1\}^n$*, and the probability ensemble* $\mathbf{X}^{(t)} = \{X_n^{(t)}\}$*, where* $X_n^{(t)}$ *consists of* $t(n)$ *independent copies of* $X_n$*.*

(hypothesis) *Suppose that* $\delta$ *is an upper bound on the correlation of families of* $s(\cdot)$*-size circuits with* $P$ *over* $\mathbf{X}$*. Namely, suppose that for every* $n$ *and for every* $s(n)$*-size circuit* $C$*, it holds that*

$$\text{Prob}[C(X_n) = P(X_n)] \leq p(n) \overset{\text{def}}{=} \frac{1 + \delta(n)}{2} .$$

(conclusion) *Then, for every function* $\epsilon : \mathbb{N} \to [0, +1]$*, for every* $n$ *and for every* $\text{poly}(\frac{\epsilon(n)}{n}) \cdot s(n)$*-size circuit* $C'$*, it holds that*

$$\text{Prob}[C'(X_n^{(t)}) = F^{(t)}(X_n^{(t)})] \leq p(n)^{t(n)} + \epsilon(n).$$

*Remark.* Nisan et. al. [14] have used the XOR-Lemma in order to derive the Concatenation Lemma. Our feeling is that the Concatenation Lemma is more "basic" than the XOR Lemma, and thus that their strategy is not very natural.[8] In fact, this feeling was our motivation for trying to find a "direct" proof for the Concatenation Lemma. Extrapolating from the situation regarding the two original lemmata of Yao (i.e., the XOR Lemma and the Concatenation Lemma w.r.t. one-way functions),[9] we believed that such a proof (for the Concatenation

[8] This assertion is supported by a recent work of Viola and Wigderson, which provides a very simple proof that, in the general setting, the XOR Lemma implies the Concatenation Lemma [16, Prop. 1.4].

[9] Yao's original XOR Lemma (resp., Concatenation Lemma) refers to the setting of one-way functions. In this setting, the basic predicate $P$ is a composition of an easy to compute predicate $b$ and the inverse of a 1-1 one-way function $f$; i.e., $P(x) \overset{\text{def}}{=} b(f^{-1}(x))$. For years, the first author has considered the proof of the XOR Lemma (even for this setting) too complicated to be presented in class; whereas, a proof of the Concatenation Lemma (for this setting) has appeared in his classnotes [1] (see also [2]).

Lemma) should be easy to find. Indeed, we consider the following proof of Concatenation Lemma much simpler than the proofs of the XOR Lemma (given in previous sections).

*A tight two-argument version.* Lemma 7 is derived from the following Lemma 8 (which is a tight two-argument version of Lemma 7) analogously to the way that Lemma 2 was derived from Lemma 4; that is, we write $F^{(t)}(x_1, x_2, ..., x_{t(n)}) = (P(x_1), F^{(t-1)}(x_2, ..., x_{t(n)}))$, assume that $F^{(t-1)}$ is hard to guess as claimed, and apply the Concatenation Lemma to $(P, F^{(t-1)})$. This way, the lower bound on circuits guessing $F^{(t)}$ is related to the lower bound assumed for circuits guessing the original $P$ and is almost the bound derived for $F^{(t-1)}$ (losing only an additive term!). It is thus left to prove the following two-argument version.

**Lemma 8** (two argument version of concatenation lemma): *Let $F_1$ and $F_2$ be two functions, $l : \mathbb{N} \to \mathbb{N}$ be a length function, and $F(x) \stackrel{\rm def}{=} (F_1(y), F_2(z))$ where $x = yz$ and $|y| = l(|x|)$. Let $\mathbf{X} = \{X_n\}$, $\mathbf{Y} = \{Y_{l(n)}\}$ and $\mathbf{Z} = \{Z_{n-l(n)}\}$ be probability ensembles as in Lemma 4* (i.e., $X_n = (Y_{l(n)}, Z_{n-l(n)})$).

(hypothesis) *Suppose that $p_1(\cdot)$ is an upper bound on the probability that families of $s_1(\cdot)$-size circuits guess $F_1$ over $\mathbf{Y}$. Namely, for every such circuit family $\mathbf{C} = \{C_l\}$ it holds that*

$$\mathrm{Prob}[C_l(Y_l) = F_1(Y_l)] \le p_1(l).$$

*Likewise, suppose that $p_2(\cdot)$ is an upper bound on the probability that families of $s_2(\cdot)$-size circuits guess $F_2$ over $\mathbf{Z}$.*

(conclusion) *Then, for every function $\epsilon : \mathbb{N} \to \mathbb{R}$, the function $p(n) \stackrel{\rm def}{=} p_1(l(n)) \cdot p_2(n - l(n)) + \epsilon(n)$ is an upper bound on the probability that families of $s(\cdot)$-size circuits guess $F$ over $\mathbf{X}$, where*

$$s(n) \stackrel{\rm def}{=} \min\left\{ \frac{s_1(l(n))}{\mathrm{poly}(n/\epsilon(n))} \;,\; s_2(n - l(n)) - n \right\}.$$

**Proof:** Let $\mathbf{C} = \{C_n\}$ be a family of $s(\cdot)$-size circuits. Fix an arbitrary $n$, and write $C = C_n$, $\epsilon = \epsilon(n)$, $l = l(n)$, $m = n - l(n)$, $Y = Y_l$ and $Z = Z_m$. Abusing notation, we let $C_1(x, y)$ denote the first component of $C(x, y)$ (i.e., the guess for $F_1(x)$) and likewise $C_2(x, y)$ is $C$'s guess for $F_2(y)$. It is instructive to write the success probability of $C$ as follows:

$$\begin{aligned}\mathrm{Prob}[C(Y, Z) = F(Y, Z)] &= \mathrm{Prob}[C_2(Y, Z) = F_2(Z)] \\ &\quad \cdot \mathrm{Prob}[C_1(Y, Z) = F_1(Y) \mid C_2(Y, Z) = F_2(Z)]\end{aligned}$$

The basic idea is that using the hypothesis regarding $F_2$ allows to bound the first factor by $p_2(m)$, whereas the hypothesis regarding $F_1$ allows to bound the second factor by approximately $p_1(l)$. The basic idea for the latter step is that a sufficiently large sample of $(Z, F_2(Z))$, which may be hard-wired into the circuit, allows to use the conditional probability space (in such a circuit), provided the condition holds with noticeable probability. The last caveat motivates a separate treatment for $y$'s with noticeable $\mathrm{Prob}[C_2(y,Z)=F_2(Z)]$ and for the rest.

We call $y$ good if $\mathrm{Prob}[C_2(y,Z)=F_2(Z)] \geq \epsilon/2$ and bad otherwise. Let $G$ be the set of good $y$'s. Then, using $\mathrm{Prob}[C(Y,Z)=F(Y,Z)] < \epsilon/2$ for every bad $y$, we upper bound the success probability of $C$ as follows

$$\begin{aligned}\mathrm{Prob}[C(Y,Z)=F(Y,Z)] &= \mathrm{Prob}[C(Y,Z)=F(Y,Z) \ \& \ Y\in G] \\ &\quad + \mathrm{Prob}[C(Y,Z)=F(Y,Z) \ \& \ Y\notin G] \\ &< \mathrm{Prob}[C(Y,Z)=F(Y,Z) \ \& \ Y\in G] + \frac{\epsilon}{2}.\end{aligned}$$

Thus, using $p(n) = p_1(l) \cdot p_2(m) + \epsilon$, it remains to prove that

$$\mathrm{Prob}[C(Y,Z)=F(Y,Z) \ \& \ Y\in G] \leq p_1(l) \cdot p_2(m) + \epsilon/2. \tag{3}$$

We proceed according to the foregoing outline. We first show that $\mathrm{Prob}[C_2(Y,Z)=F_2(Z)]$ cannot be too large, as otherwise the hypothesis concerning $F_2$ is violate. Actually, we prove the following

Claim 8.1: For every $y$, it holds that

$$\mathrm{Prob}[C_2(y,Z)=F_2(Z)] \leq p_2(m).$$

Proof: Otherwise, using any $y \in \{0,1\}^l$ such that $\mathrm{Prob}[C_2(y,Z)=F_2(Z)] > p_2(m)$, we get a circuit $C'(z) \stackrel{\mathrm{def}}{=} C_2(y,z)$ that contradicts the lemma's hypothesis concerning $F_2$. □

Next, we use Claim 8.1 in order to relate the success probability of $C$ to the success probability of small circuits for $F_1$.

Claim 8.2: There exists a circuit $C'$ of size $s_1(l)$ such that

$$\mathrm{Prob}[C'(Y)=F_1(Y)] \ \geq \ \frac{\mathrm{Prob}[C(Y,Z)=F(Y,Z) \ \& \ Y\in G]}{p_2(m)} \ - \ \frac{\epsilon}{2}.$$

Proof: The circuit $C'$ is constructed as suggested in the foregoing outline. Specifically, we take a $\mathrm{poly}(n/\epsilon)$-large sample, denoted $S$, from the distribution $(Z, F_2(Z))$ and let $C'(y) \stackrel{\mathrm{def}}{=} C_1(y,z)$, where $(z,\beta)$ is a uniformly selected among the elements of $S$ for which $C_2(y,z)=\beta$ holds. Details follow.

Let $S$ be a sequence of $t \stackrel{\mathrm{def}}{=} \mathrm{poly}(n/\epsilon)$ pairs, generated by taking $t$ independent samples from the distribution $(Z, F_2(Z))$. We stress that we do not assume here that such a sample can be produced by an efficient (uniform) algorithm (but, jumping ahead, we remark that such a sequence can be fixed non-uniformly). For each $y \in G \subseteq \{0,1\}^l$, we denote by $S_y$ the set of pairs $(z,\beta) \in S$ for which $C_2(y,z)=\beta$. Note that $S_y$ is a random sample for the residual probability space defined by $(Z, F_2(Z))$ conditioned on $C_2(y,Z) = F_2(Z)$. Also, with overwhelmingly high probability, $|S_y| = \Omega(l/\epsilon^2)$ (since $y \in G$ implies

$\text{Prob}[C_2(y,Z)=F_2(Z)] \geq \epsilon/2)$. Thus, with overwhelming probability (i.e., probability greater than $1-2^{-l}$), taken over the choices of $S$, the sample $S_y$ provides a good approximation to the conditional probability space, and in particular

$$\frac{|\{(z,\beta)\in S_y : C_1(y,z)=F_1(y)\}|}{|S_y|} \geq \text{Prob}[C_1(y,Z)=F_1(y) \,|\, C_2(y,Z)=F_2(Z)]-\frac{\epsilon}{2} \quad (4)$$

Thus, with positive probability, Eq. (4) holds for all $y \in G \subseteq \{0,1\}^l$. The circuit $C'$ guessing $F_1$ is now defined as follows. A set $S = \{z_i, \beta_i\}$ satisfying Eq. (4) for all good $y$'s is "hard-wired" into the circuit $C'$. (In particular, $S_y$ is not empty for any good $y$.) On input $y$, the circuit $C'$ first determines the set $S_y$, by running $C$ for $t$ times and checking, for each $i = 1, ..., t$, whether $C_2(y, z_i) = \beta_i$. In case $S_y$ is empty, the circuit returns an arbitrary value. Otherwise, the circuit selects uniformly a pair $(z, \beta) \in S_y$ and outputs $C_1(y, z)$. (This latter random choice can be eliminated by a standard averaging argument.) Using the definition of $C'$ and Eq. (4), we get

$$\begin{aligned}
&\text{Prob}[C'(Y)=F_1(Y)]\\
&\geq \sum_{y\in G} \text{Prob}[Y=y] \cdot \text{Prob}[C'(y)=F_1(y)]\\
&= \sum_{y\in G} \text{Prob}[Y=y] \cdot \frac{|\{(z,\beta) \in S_y : C_1(y,z)=F_1(y)\}|}{|S_y|}\\
&\geq \sum_{y\in G} \text{Prob}[Y=y] \cdot \left(\text{Prob}[C_1(y,Z)=F_1(y) \,|\, C_2(y,Z)=F_2(Z)] - \frac{\epsilon}{2}\right)\\
&\geq \left(\sum_{y\in G} \text{Prob}[Y=y] \cdot \frac{\text{Prob}[C(y,Z)=F(y,Z)]}{\text{Prob}[C_2(y,Z)=F_2(Z)]}\right) - \frac{\epsilon}{2}.
\end{aligned}$$

Next, using Claim 8.1, we get

$$\text{Prob}[C'(Y)=F_1(Y)] \geq \left(\sum_{y\in G} \text{Prob}[Y=y] \cdot \frac{\text{Prob}[C(y,Z)=F(y,Z)]}{p_2(m)}\right) - \frac{\epsilon}{2}$$

and the claim follows. □

Now, by the lemma's hypothesis concerning $F_1$, we have $\text{Prob}[C'(Y)=F_1(Y)] \leq p_1(l)$, and so using Claim 8.2 we get

$$\begin{aligned}
\text{Prob}[Y \in G \;\&\; C(Y,Z)=F(Y,Z)] &\leq (p_1(l) + \epsilon/2) \cdot p_2(m)\\
&\leq p_1(l) \cdot p_2(m) + \epsilon/2.
\end{aligned}$$

This proves Eq. (3) and the lemma follows. ■

### 5.2 Deriving the XOR Lemma from the Concatenation Lemma

Using the techniques of Goldreich and Levin [5], we obtain the following result.

**Lemma 9** (hard-core predicate of unpredictable functions): *Let $F:\{0,1\}^* \to \{0,1\}^*$, $p:\mathbb{N}\to[0,1]$, and $s:\mathbb{N}\to\mathbb{N}$, and let $\mathbf{X}=\{X_n\}$ be as in Definition 1. For $\alpha,\beta\in\{0,1\}^\ell$, we denote by $IP_2(\alpha,\beta)$ the inner-product mod 2 of $\alpha$ and $\beta$, viewed as binary vectors of length $\ell$.*

(hypothesis) *Suppose that, for every $n$ and for every $s(n)$-size circuit $C$, it holds that*

$$\mathrm{Prob}[C(X_n)=F(X_n)] \le p(n).$$

(conclusion) *Then, for some constant $c>0$, for every $n$ and for every* $\mathrm{poly}(\frac{p(n)}{n})\cdot$ *$s(n)$-size circuit $C'$, it holds that*

$$\mathrm{Prob}[C'(X_n,U_\ell)=IP_2(F(X_n),U_\ell)] \le \frac{1}{2}+c\cdot\sqrt[3]{n^2\cdot p(n)},$$

*where $U_\ell$ denotes the uniform distribution over* $\{0,1\}^\ell$, *with $\ell \stackrel{\rm def}{=} |F(X_n)|$.* (That is, $C'$ has correlation at most $2c\sqrt[3]{n^2 p(n)}$ with $\mathrm{IP}_2$ over $(F(X_n),U_\ell)$.)

**Proof Sketch:** Let $q(n) \stackrel{\rm def}{=} c\,\sqrt[3]{n^2\,p(n)}$. Suppose that $C'$ contradicts the conclusion of the lemma. Then, there exists a set $S$ such that $\mathrm{Prob}[X_n\in S]\ge q(n)$ and for every $x\in S$ the probability that $C'(x,U_\ell)=\mathrm{IP}_2(F(x),U_\ell)$ is at least $\frac{1}{2}+\frac{q(n)}{2}$, where the probability is taken over $U_\ell$ (while $x$ is fixed). Employing the techniques of [5][10], we obtain a randomized circuit $C$ (of size at most a $\mathrm{poly}(n/p(n))$ factor larger than $C'$) such that, for every $x\in S$, it holds that $\mathrm{Prob}[C(X_n)=F(X_n)]\ge c'\cdot(q(n)/n)^2$ (where the constant $c'>0$ is determined in the proof of [5] according to Chebishev's Inequality).[11] Thus, $C$ satisfies

$$\begin{aligned}\mathrm{Prob}[C(X_n)=F(X_n)] &\ge \mathrm{Prob}[C(X_n)=F(X_n)\wedge X_n\in S]\\ &= \mathrm{Prob}[X_n\in S]\cdot\mathrm{Prob}[C(X_n)=F(X_n)|X_n\in S]\\ &\ge q(n)\cdot\left(c'\cdot(q(n)/n)^2\right)=p(n)\end{aligned}$$

in contradiction to the hypothesis. The lemma follows. ■

*Conclusion.* Combining the Concatenation Lemma (Lemma 7) with Lemma 9 we establish the validity of Lemma 1 for the third time; this time with respect to the parameters $p(n)=cn^{2/3}=o(n)$ and $\delta'(n)=\sqrt[3]{\frac{1+\delta(n)}{2}}$. Details follow.

Starting with a predicate for which $\delta$ is a correlation bound and using Lemma 7, we get a function that is hard to guess with probability substantially higher than

[10] See alternative expositions in either [4, Sec. 7.1.3] or [3, Sec. 2.5.2].

[11] The algorithm in [5] will actually retrieve all values $\alpha\in\{0,1\}^\ell$ for which the correlation of $C'(x,U_\ell)$ and $\mathrm{IP}_2(\alpha,U_\ell)$ is at least $q(n)$. With overwhelming probability it outputs a list of $O((n/q(n))^2)$ strings containing all the values just mentioned and thus uniformly selecting one of the values in the list yields $F(x)$ with probability at least $1/O((n/q(n))^2)$.

$(\frac{1+\delta(n)}{2})^{t(n)}$. Applying Lemma 9 establishes that given $(x_1, ..., x_{t(n)})$ and a uniformly chosen subset $S \subseteq \{1, 2, ..., t(n)\}$ it is hard to correlate $\oplus_{i\in S} P(x_i)$ better than with correlation

$$O\left(\sqrt[3]{n^2 \cdot \left(\frac{1+\delta(n)}{2}\right)^{t(n)}}\right) = o(n) \cdot \left(\sqrt[3]{\frac{1+\delta(n)}{2}}\right)^{t(n)}.$$

This is almost what we need, but not quite (what we need is a statement concerning $S = \{1, ..., t(n)\}$). The gap is easily bridged by some standard "padding" trick. For example, by using a sequence of fixed pairs $(z_i, \sigma_i)$, such that $\sigma_i = P(z_i)$, we reduce the computation of $\oplus_{i\in S} P(x_i)$ to the computation of $\oplus_{i=1}^{t(n)} P(y_i)$ by setting $y_i = x_i$ if $i \in S$ and $y_i = z_i$ otherwise. (See Appendix B for more details.) Thus, Lemma 1 follows (with the stated parameters).

### 5.3 Remarks Concerning the Uniform Complexity Setting

A uniform-complexity analogue of the foregoing proof can be carried out provided that one can construct random examples of the distribution $(X_n, P(X_n))$ within the stated (uniform) complexity bounds (and in particular in polynomial-time). Actually, this condition is required only for the proof of the Concatenation Lemma. Thus we confine ourselves to presenting a uniform-complexity version of the Concatenation Lemma.

**Lemma 10** (Concatenation Lemma – uniform complexity version): *Let $P, \mathbf{X}$, $s, \delta, t$ and $F^{(t)}$ be as in Lemma 7.*

(hypothesis) *Suppose that $\delta(\cdot)$ is an upper bound on the correlation of $T(\cdot)$-time-constructible families of $s(\cdot)$-size circuits with $P$ over $\mathbf{X}$. Furthermore, suppose that one can generate in polynomial-time a random sample from the distribution $(X_n, P(X_n))$.*

(conclusion) *Then, for every function $\epsilon : \mathbb{N} \to [0, +1]$, the function $q(n) \stackrel{\text{def}}{=} p(n)^{t(n)} + \epsilon(n)$ is an upper bound on the correlation of $T'(\cdot)$-time-constructible families of $s'(\cdot)$-size circuits with $F$ over $\mathbf{X}^{(t)}$, where $T'(t(n) \cdot n) = \text{poly}(\epsilon(n)/n) \cdot T(n)$ and $s'(t(n) \cdot n) = \text{poly}(\epsilon(n)/n) \cdot s(n)$.*

The uniform-complexity version of the Concatenation Lemma is proven by adapting the proof of Lemma 7 as follows. Firstly, we observe that it suffices to prove an appropriate (uniform-complexity) version of Lemma 8. This is done by first proving a weaker version of Claim 8.1 that asserts that for all but at most an $\epsilon(n)/8$ measure of the $y$'s (under $Y$), it holds that

$$\text{Prob}[C_2(y, Z) = F_2(Z)] \leq p_2(m) + \epsilon(n)/8.$$

This holds because otherwise one may sample $Y$ with the aim of finding a $y$ such that $\text{Prob}[C_2(y, Z) = F_2(Z)] > p_2(m)$ holds, and then use this $y$ to construct (uniformly!) a circuit that contradicts the hypothesis concerning $F_2$. Next, we prove a weaker version of Claim 8.2 by observing that, for a uniformly selected pair sequence $S$, with overwhelmingly high probability (and not only with positive probability), Eq. (4) holds for all good $y \in \{0,1\}^l$. Thus, if we generate $S$ by taking random samples from the distribution $(Z_m, F_2(Z_m))$, then with overwhelmingly high probability we end-up with a circuit as required by the modified claim. (The modified claim has $p_2(m) + \epsilon/8$ in the denominator (rather than $p_2(m)$) as well as an extra additive term of $\epsilon/8$.) Using the hypothesis concerning $F_1$, we are done as in the non-uniform case.

## 6 A Different Perspective: The Entropy Angle

The XOR Lemma and the Concatenation Lemma are special cases of the so-called "direct sum conjecture" asserting that computational difficulty increases when many independent instances of the problem are to be solved. In both cases the "direct sum conjecture" is postulated by considering insufficient resources and bounding the probability that these tasks can be performed within these resources, as a function of the number of instances. In this section we suggest an analogous analysis based on entropy rather than probability. Specifically, we consider the amount of information remaining in the task (e.g., of computing $f(x)$) when given the result of a computation (e.g., $C(x)$). This analysis turns out to be much easier.

**Proposition 11.** *Let $f$ be a predicate, $X$ be a random variable and $\mathcal{C}$ be a class of circuits so that for every circuit $C \in \mathcal{C}$*

$$\mathrm{H}(f(X)|C(X)) \geq \epsilon,$$

*where* H *denotes the* (conditional) *binary entropy function. Furthermore, suppose that, for every circuit $C \in \mathcal{C}$, fixing any of the inputs of $C$ yields a circuit also in $\mathcal{C}$. Then, for every circuit $C \in \mathcal{C}$, it holds that*

$$\mathrm{H}(f(X^{(1)}), ..., f(X^{(t)})|C(X^{(1)}, ..., X^{(t)})) \geq t \cdot \epsilon,$$

*where the $X^{(i)}$'s are independently distributed copies of $X$.*

We stress that the class $\mathcal{C}$ in Proposition 11 may contain circuits with several Boolean outputs. Furthermore, for a meaningful conclusion, the class $\mathcal{C}$ must contain circuits with $t$ outputs (otherwise, for a circuit $C$ with much fewer outputs, the conditional entropy $\mathrm{H}(f(x_1), ..., f(x_t)|C(x_1, ..., x_t))$ is large merely due to information theoretical reasons). On the other hand, the more outputs the circuits in $\mathcal{C}$ have, the stronger the hypothesis of Proposition 11 is. In particular, the number of outputs must be smaller that $|X|$ otherwise the value of the circuit $C(x) = x$ determines $f(x)$ (i.e., $\mathrm{H}(f(x)|x) = 0$). Thus, a natural instantiation of Proposition 11 is for a family of small (e.g., poly-size) circuits each having $t$ outputs.

**Proof:** By definition of conditional entropy, we have for every $C \in \mathcal{C}$,

$$\begin{aligned}&\mathrm{H}(f(X^{(1)}), ..., f(X^{(t)})|C(X^{(1)}, ..., X^{(t)}))\\ &= \sum_{i=1}^{t} \mathrm{H}(f(X^{(i)})|C(X^{(1)}, ..., X^{(t)}), f(X^{(1)}), ..., f(X^{(i-1)}))\\ &\geq \sum_{i=1}^{t} \mathrm{H}(f(X^{(i)})|C(X^{(1)}, ..., X^{(t)}), X^{(1)}, ..., X^{(i-1)}).\end{aligned}$$

Now, for each $i$, we show that

$$\mathrm{H}(f(X^{(i)})|C(X^{(1)}, ..., X^{(t)}), X^{(1)}, ..., X^{(i-1)}) \geq \epsilon.$$

We consider all possible settings of all variables, except $X^{(i)}$, and bound the conditional entropy under this setting (which does not effect $X^{(i)}$). The fixed values $X^{(j)} = x_j$ can be eliminated from the entropy condition and incorporated into the circuit. However, fixing some of the inputs in the circuit $C$ yields a circuit also in $\mathcal{C}$ and so we can apply the proposition's hypothesis and get

$$\mathrm{H}(f(X^{(i)})|C(x_1, ..., x_{i-1}, X^{(i)}, x_{i+1}, ..., x_t)) \geq \epsilon.$$

The proposition follows. ■

**Proposition 11 vs the Concatenation Lemma.** We compare the hypotheses and conclusions of these two results.

***The hypotheses.*** The hypothesis in Proposition 11 is related to the hypotheses in the Concatenation Lemma. Clearly, an entropy lower bound (on a single bit) translates to some unpredictability bound on this bit. (This does not hold for many bits as can be seen below.) The other direction (i.e., unpredictability implies a lower bound on the conditional entropy) is obvious for a single bit.

***The conclusions.*** For $t = O(\log n)$ the conclusion of Proposition 11 is implied by the conclusion of the Concatenation Lemma, but for sufficiently large $t$ the conclusion of Proposition 11 does not imply the conclusion of Concatenation Lemma. Details follow.

1. To show that, for $t = O(\log n)$, the conclusion of the Concatenation Lemma implies the conclusion of Proposition 11, suppose that for a small circuit $C$ it holds that $h \stackrel{\text{def}}{=} \mathrm{H}(f(X^{(1)}), ..., f(X^{(t)})|C(X^{(1)}, ..., X^{(t)})) = o(t)$. Then, for every value of $C$, denoted $v$, there exists a string $w = w(v)$ such that $\mathrm{Prob}[f(X^{(1)}), ..., f(X^{(t)}) = w | C(X^{(1)}, ..., X^{(t)}) = v] \geq 2^{-h}$. Hardwiring these $2^t$ strings $w(\cdot)$ into $C$, we obtain a small circuit that predicts $f(X^{(1)}), ..., f(X^{(t)})$ with probability at least $2^{-h} = 2^{-o(t)}$, in contradiction to the conclusion of the Concatenation Lemma.
2. To show that the conclusion of Proposition 11 does not imply the conclusion of the Concatenation Lemma, consider the possibility of a small

(randomized) circuit $C$ that with probability $1-\epsilon$ correctly determines all the $f$ values (i.e., $\text{Prob}[C(X^{(1)},...,X^{(t)})=f(X^{(1)}),...,f(X^{(t)})] = 1-\epsilon$), and yields no information (e.g., outputs a special fail symbol) otherwise. Then, although $C$ has success probability $1-\epsilon$, the conditional entropy is $(1-\epsilon)\cdot 0+\epsilon\cdot t$ (assuming that $\text{Prob}[f(X)=1]=1/2$).

## 7 Subsequent Work

Since the first publication of this survey, Yao's XOR Lemma has been the subject of intensive research. Here we only outline three themes that were pursued, while referring the interested reader to [10] and the references therein.

*Derandomization.* A central motivation for Impagliazzo's work [7,8] has been the desire to present "derandomized versions" of the XOR Lemma; that is, predicates that use their input in order to define a sequence of related instances, and take the XOR of the original predicate on these instances.[12] The potential benefit in such a construction is that the hardness of the resulting predicate is related to shorter inputs (i.e., the seed of a generator of a $t$-long sequence of $n$-bit long strings, rather than the $tn$-bit long sequence itself). Indeed, Impagliazzo's work [7,8] presented such a construction (based on a pairwise independent generator), and left the question of providing a "full derandomization" (that uses a seed of length $O(n)$ to generate $t$ instances) to subsequent work. The goal was achieved by Impagliazzo and Wigderson [11] by using a generator that combines Impagliazzo's generator [7,8] with a new generator, which in turn combines an expander walk generator with the Nisan-Wigderson generator [15].

*Avoiding the use of random examples.* As pointed out in Section 2.3, all proofs presented in this survey make an essential use of random examples. For more than a decade, this feature stood in the way of a general uniform version of the XOR Lemma (i.e., all uniform proofs assumed access to such random examples). This barrier was lifted by Impagliazzo, Jaiswal, and Kabanets [9], which culminated in comprehensive treatment of [10]. The latter work provides simplified, optimized, and derandomized versions of the XOR and Concatenation Lemmas.[13] The key idea is to use the hypothetical solver of the concatenated problem in order to obtain a sequence of random examples that are all good with noticeable probability. An instance of the original problem is then solved by hiding it in a random sequence that has a fair intersection with the initial sequence of random examples. The interested reader is referred to [10] for a

[12] That is, the predicate consists of an "instance generator" and multiple applications of the original predicate, $P$. Specifically, on input an $s$-bit long seed, denoted $y$, the generator produces a $t$-long sequence of $n$-bit long strings (i.e., $(x_1,...,x_t)\leftarrow G(y)$), and the value of the new predicate is defined as the XOR of the values of $P$ on these $t$ strings (i.e., $\oplus_{i=1}^{t}P(x_i)$).

[13] The focus of [10] is actually on the Concatenation Lemma, which is currently called the Direct Product Theorem. See next paragraph regarding the relation to the XOR Lemma.

mature description of this idea (and its sources of inspirarion) as well as for a discussion of the relation this problem (i.e., proofs of the Concatenation Lemma) and list-decoding of the direct product code.

*The relation between the XOR and Concatenation Lemmas.* In Section 5 we advocated deriving the XOR Lemma from the Concatenation Lemma, and this suggestion was adopted in several works (including [9,10]). Our intuition that the Concatenation Lemma is simpler than the XOR Lemma is supported by a recent work of Viola and Wigderson, which provides a very simple proof that, in the general setting, the XOR Lemma implies the Concatenation Lemma [16, Prop. 1.4]. We mention that the both directions of the equivalence between the Concatenation Lemma and the XOR Lemma pass through an intermediate lemma called the Selective XOR Lemma (see [4, Exer. 7.17]). For further discussion see Appendix B.

**Acknowledgement.** We wish to thank Mike Saks for useful discussions regarding Levin's proof of the XOR Lemma. We also thank Salil Vadhan and Ronen Shaltiel for pointing out errors in previous versions, and for suggesting ways to fix these errors.

## References

1. Goldreich, O.: Foundation of Cryptography – Class Notes. Computer Science Department, Technion, Haifa, Israel (Spring 1989)
2. Goldreich, O.: Foundation of Cryptography – Fragments of a Book, Available from ECCC (February 1995)
3. Goldreich, O.: Foundation of Cryptography: Basic Tools. Cambridge University Press, Cambridge (2001)
4. Goldreich, O.: Computational Complexity: A Conceptual Perspective. Cambridge University Press, Cambridge (2008)
5. Goldreich, O., Levin, L.A.: A Hard-Core Predicate for all One-Way Functions. In: 21st STOC, pp. 25–32 (1989)
6. Håstad, J., Impagliazzo, R., Levin, L.A., Luby, M.: A Pseudorandom Generator from any One-way Function. SICOMP 28(4), 1364–1396 (1999); Combines papers of Impagliazzo et al. (21st STOC, 1989) and Håstad (22nd STOC, 1990)
7. Impagliazzo, R.: See [8], which appeared after our first posting (1994) (manuscript)
8. Impagliazzo, R.: Hard-core Distributions for Somewhat Hard Problems. In: 36th FOCS, pp. 538–545 (1995); This is a later version of [7]
9. Impagliazzo, R., Jaiswal, R., Kabanets, V.: Approximately List-Decoding Direct Product Codes and Uniform Hardness Amplification. In: 47th FOCS, pp. 187–196 (2006)
10. Impagliazzo, R., Jaiswal, R., Kabanets, V., Wigderson, A.: Uniform Direct Product Theorems: Simplified, Optimized, and Derandomized. SIAM J. Comput. 39(4), 1637–1665 (2010); Preliminary version in 40th STOC (2008)
11. Impagliazzo, R., Wigderson, A.: P=BPP if E requires exponential circuits: Derandomizing the XOR Lemma. In: 29th STOC, pp. 220–229 (1997)
12. Levin, L.A.: One-Way Functions and Pseudorandom Generators. Combinatorica 7(4), 357–363 (1987)
13. Levin, L.A.: Average Case Complete Problems. SICOMP 15, 285–286 (1986)

14. Nisan, N., Rudich, S., Saks, M.: Products and Help Bits in Decision Trees. In: 35th FOCS, pp. 318–329 (1994)
15. Nisan, N., Wigderson, A.: Hardness vs Randomness. JCSS 49(2), 149–167 (1994)
16. Viola, E., Wigderson, A.: Norms, XOR Lemmas, and Lower Bounds for Polynomials and Protocols. Theory of Computing 4(1), 137–168 (2008); Preliminary version in IEEE Conf. on Comput. Complex. (2007)
17. Yao, A.C.: Theory and Application of Trapdoor Functions. In: 23rd FOCS, pp. 80–91 (1982)

## Appendix A: Proof of a Generalization of Lemma 6

We first generalize Impagliazzo's treatment to the case of non-uniform distributions; Impagliazzo's treatment is regained by letting $\mathbf{X}$ be the uniform probability ensemble.

**Definition 3** (hard-core of a predicate relative to a distribution): *Let $f:\{0,1\}^* \to \{0,1\}$ be a Boolean predicate, $s:\mathbb{N}\to\mathbb{N}$ be a size function, $\epsilon:\mathbb{N}\to[0,1]$ be a function, and $\mathbf{X}=\{X_n\}$ be a probability ensemble.*

- *We say that a sequence of sets, $\mathbf{S}=\{S_n\subseteq\{0,1\}^n\}$, is a* hard-core of $f$ relative to $\mathbf{X}$ *with respect to $s(\cdot)$-size circuits families and advantage $\epsilon(\cdot)$ if for every $n$ and every circuit $C_n$ of size at most $s(n)$, it holds that*

$$\mathrm{Prob}[C_n(X_n)=f(X_n)|X_n\in S_n] \leq \frac{1}{2}+\epsilon(n).$$

- *We say that $f$ has a* hard-core *of density $\rho(\cdot)$ relative to $\mathbf{X}$ with respect to $s(\cdot)$-size circuits families and advantage $\epsilon(\cdot)$ if there exists a sequence of sets $\mathbf{S}=\{S_n\subseteq\{0,1\}^n\}$ such that $\mathbf{S}$ is a hard-core of $f$ relative to $\mathbf{X}$ with respect to the above and* $\mathrm{Prob}[X_n\in S_n]\geq\rho(n)$.

**Lemma 12** (generalization of Lemma 6): *Let $f:\{0,1\}^*\to\{0,1\}$ be a Boolean predicate, $s:\mathbb{N}\to\mathbb{N}$ be a size function, $\mathbf{X}=\{X_n\}$ be a probability ensemble, and $\rho:\mathbb{N}\to[0,1]$ be a noticeable function such that for every $n$ and every circuit $C_n$ of size at most $s(n)$, it holds that*

$$\mathrm{Prob}[C_n(X_n)=f(X_n)]\leq 1-\rho(n).$$

*Then, for every function $\epsilon:\mathbb{N}\to[0,1]$, the function $f$ has a hard-core of density $\rho'(\cdot)$ relative to $\mathbf{X}$ with respect to $s'(\cdot)$-size circuits families and advantage $\epsilon(\cdot)$, where $\rho'(n)\overset{\mathrm{def}}{=}(1-o(1))\cdot\rho(n)$ and $s'(n)\overset{\mathrm{def}}{=}s(n)/\mathrm{poly}(n/\epsilon(n))$.*

**Proof:** We start by proving a weaker statement; namely, that $\mathbf{X}$ "dominates" an ensemble $\mathbf{Y}$ under which the function $f$ is strongly unpredictable. Our notion of domination originates in a different work of Levin [13]. Specifically, referring to a fixed function $\rho$, we define domination as assigning probability mass that is at least a $\rho$ fraction of the mass assigned by the dominated ensemble; namely:

Definition: *Fixing the function $\rho$ for the rest of the proof*, we say that the ensemble $\mathbf{X} = \{X_n\}$ dominates the ensemble $\mathbf{Y} = \{Y_n\}$ if for every string $\alpha$,

$$\mathrm{Prob}[X_n = \alpha] \geq \rho(|\alpha|) \cdot \mathrm{Prob}[Y_n = \alpha].$$

In this case we also say that $\mathbf{Y}$ is dominated by $\mathbf{X}$. We say that $\mathbf{Y}$ is critically dominated by $\mathbf{X}$ if for every string $\alpha$ either $\mathrm{Prob}[Y_n = \alpha] = (1/\rho(|\alpha|)) \cdot \mathrm{Prob}[X_n = \alpha]$ or $\mathrm{Prob}[Y_n = \alpha] = 0$. (Actually, to avoid trivial difficulties, we allow at most one string $\alpha \in \{0,1\}^n$ such that $0 < \mathrm{Prob}[Y_n = \alpha] < (1/\rho(|\alpha|)) \cdot \mathrm{Prob}[X_n = \alpha]$.)

The notions of domination and critical domination play central roles in the following proof, which consists of two parts. In the first part (cf., Claim 12.1), we prove the existence of a ensemble dominated by $\mathbf{X}$ such that $f$ is strongly unpredictable under this ensemble. In the second part (cf., Claims 12.2 and 12.3), we essentially prove that the existence of such a dominated ensemble implies the existence of an ensemble that is *critically* dominated by $\mathbf{X}$ such that $f$ is strongly unpredictable under the latter ensemble. However, such a critically dominated ensemble defines a hard-core of $f$ relative to $\mathbf{X}$, and the lemma follows. Before starting, we make the following simplifying assumptions (used in Claim 12.3).

Simplifying assumptions: Without loss of generality, the following two conditions hold:

1. $\log_2 s(n) \leq n$.
   (Otherwise the hypothesis of the lemma cannot hold.)
2. $\mathrm{Prob}[X_n = x] < \mathrm{poly}(n)/s(n)$, for all $x$'s.
   (This assumption is justified since $x$'s violating this condition cannot contribute to the hardness of $f$ with respect to $X_n$, because one can incorporate all these $s(n)/poly(n)$ many violating $x$'s with their corresponding $f(x)$'s into the circuit).

Claim 12.1: Under the hypothesis of the lemma it holds that there exists a probability ensemble $\mathbf{Y} = \{Y_n\}$ such that $\mathbf{Y}$ is dominated by $\mathbf{X}$ and, for every $s'(n)$-circuit $C_n$, it holds that

$$\mathrm{Prob}[C_n(Y_n) = f(Y_n)] \leq \frac{1}{2} + \frac{\epsilon(n)}{2}. \tag{5}$$

Proof:[14] We start by assuming, towards the contradiction, that for every distribution $Y_n$ that is dominated by $X_n$ there exists an $s'(n)$-size circuits $C_n$ such that $\mathrm{Prob}[C_n(Y_n) = f(Y_n)] > 0.5 + \epsilon'(n)$, where $\epsilon'(n) = \epsilon(n)/2$. One key observation is that there is a correspondence between the set of all distributions that are each dominated by $X_n$ and the set of all the convex combinations of critically dominated (by $X_n$) distributions; that is, each dominated distribution is a convex combinations of critically dominated distributions and vice versa. Thus, considering an enumeration $Y_n^{(1)}, ..., Y_n^{(t)}$ of all the critically dominated (by $X_n$)

[14] The current text was revised following the revision in [4, Sec. 7.2.2.1].

distributions, we conclude that, for every distribution (or convex combination) $\pi$ on $[t]$, there exists an $s'(n)$-size circuits $C_n$ such that

$$\sum_{i=1}^{t} \pi(i) \cdot \text{Prob}[C_n(Y_n^{(i)}) = f(Y_n^{(i)})] > 0.5 + \epsilon'(n). \tag{6}$$

Now, consider a finite game between two players, where the first player selects a critically dominated (by $X_n$) distribution, and the second player selects an $s'(n)$-size circuit and obtains a payoff as determined by the corresponding success probability; that is, if the first player selects the $i^{\text{th}}$ critically dominated distribution and the second player selects the circuit $C$, then the payoff equals $\text{Prob}[C(Y_n^{(i)}) = f(Y_n^{(i)})]$. Taking this perspective Eq. (6) means that, for any randomized strategy for the first player, there exists a deterministic strategy for the second player yielding average payoff greater than $0.5 + \epsilon'(n)$. The Min-Max Principle asserts that, in such a case, there exists a randomized strategy for the second player that yields average payoff greater than $0.5 + \epsilon'(n)$ no matter what strategy is employed by the first player. This means that there exists a distribution, denoted $D_n$, on $s'(n)$-size circuits such that for every $i$ it holds that

$$\text{Prob}[D_n(Y_n^{(i)}) = f(Y_n^{(i)})] > 0.5 + \epsilon'(n), \tag{7}$$

where the probability refers both to the choice of the circuit $D_n$ and to the random variable $Y_n^{(i)}$. Let $B_n = \{x : \text{Prob}[D_n(x) = f(x)] \leq 0.5 + \epsilon'(n)\}$. Then, $\text{Prob}[X_n \in B_n] < \rho(n)$, because otherwise we reach a contradiction to Eq. (7) by defining $Y_n$ such that $\text{Prob}[Y_n = x] = \text{Prob}[X_n = x]/\text{Prob}[X_n \in B_n]$ if $x \in B_n$ and $\text{Prob}[Y_n = x] = 0$ otherwise.[15] By employing standard amplification to $D_n$, we obtain a distribution $D'_n$ over $\text{poly}(n/\epsilon'(n)) \cdot s'(n)$-size circuits such that for every $x \in \{0,1\}^n \setminus B_n$ it holds that $\text{Prob}[D'_n(x) = f(x)] > 1 - 2^{-n}$. It follows that there exists an $s(n)$-sized circuit $C_n$ such that $C_n(x) = f(x)$ for every $x \in \{0,1\}^n \setminus B_n$, which implies that $\text{Prob}[C_n(X_n) = f(X_n)] \geq \text{Prob}[X_n \in \{0,1\}^n \setminus B_n] > 1 - \rho(n)$, in contradiction to the theorem's hypothesis. The claim follows. □

*From a dominated ensemble to a hard-core.* In the rest of the proof, we fix an arbitrary ensemble, denoted $\mathbf{Y} = \{Y_n\}$ satisfying Claim 12.1. Using this ensemble, which is dominated by $\mathbf{X}$, we prove the validity of the lemma (i.e., the existence of a hard-core) by a probabilistic argument. Specifically, we consider the following probabilistic construction.

Probabilistic construction: We define a random set $R_n \subseteq \{0,1\}^n$ by selecting each string $x \in \{0,1\}^n$ to be in $R_n$ with probability

$$p(x) \stackrel{\text{def}}{=} \frac{\rho(n) \cdot \text{Prob}[Y_n = x]}{\text{Prob}[X_n = x]} \leq 1 \tag{8}$$

[15] Note that $Y_n$ is dominated by $X_n$, whereas by the hypothesis $\text{Prob}[D_n(Y_n) = f(Y_n)] \leq 0.5 + \epsilon'(n)$. Using the fact that any dominated distribution is a convex combination of critically dominated distributions, it follows that $\text{Prob}[D_n(Y_n^{(i)}) = f(Y_n^{(i)})] \leq 0.5 + \epsilon'(n)$ holds for some critically dominated $Y_n^{(i)}$.

independently of the choices made for all other strings. Note that the inequality holds because $\mathbf{X}$ dominates $\mathbf{Y}$.

First we show that, with overwhelmingly high probability over the choive of $R_n$, it holds that $\text{Prob}[X_n \in R_n] \approx \rho(n)$.

Claim 12.2: Let $\alpha > 0$ and suppose that $\text{Prob}[X_n = x] \leq \rho(n) \cdot \alpha^2/\text{poly}(n)$, for every $x$. Then, for all but at most a $2^{-\text{poly}(n)}$ measure of the choices of $R_n$, it holds that

$$|\text{Prob}[X_n \in R_n] - \rho(n)| < \alpha \cdot \rho(n).$$

Proof: For every $x \in \{0,1\}^n$, let $w_x \stackrel{\text{def}}{=} \text{Prob}[X_n = x]$. We define random variables $\zeta_x = \zeta_x(R_n)$, over the probability space defined by the random choices of $R_n$, such that $\zeta_x$ indicate whether $x \in R_n$; that is, the $\zeta_x$'s are independent of one another, and $\text{Prob}[\zeta_x = 1] = p(x)$ (and $\zeta_x = 0$ otherwise). Thus, for every possible choice of $R_n$, it holds that

$$\text{Prob}[X_n \in R_n] = \sum_x \zeta_x(R_n) \cdot w_x$$

and consequently we are interested in the behaviour of the sum $\sum_x w_x \zeta_x$ as a random variable (over the probability space of all possible choices of $R_n$). Taking expactation (over the possible choices of $R_n$), we get

$$\begin{aligned} \text{E}\left[\sum_x w_x \zeta_x\right] &= \sum_x p(x) \cdot w_x \\ &= \sum_x \frac{\rho(n) \cdot \text{Prob}[Y_n = x]}{\text{Prob}[X_n = x]} \cdot \text{Prob}[X_n = x] \\ &= \rho(n). \end{aligned}$$

Now, using Chernoff bound, we get

$$\text{Prob}\left[\left|\sum_x w_x \zeta_x - \rho(n)\right| > \alpha \cdot \rho(n)\right] < \exp\left(-\Omega\left(\frac{\alpha^2 \rho(n)}{\max_x\{w_x\}}\right)\right).$$

Finally, using the claim's hypotheses $w_x \leq \alpha^2 \cdot \rho(n)/\text{poly}(n)$ (for all $x$'s), the latter expression is bounded by $\exp(-\text{poly}(n))$, and the claim follows. □

Finally, we show that $R_n$ is likely to be a hard-core of $f$ realtive to $\mathbf{X}$ (w.r.t. sufficiently small circuits).

Claim 12.3:[16] For all but at most a $2^{-\text{poly}(n)}$ measure of the choices of $R_n$, it holds that every circuit $C_n$ of size $s'(n)$ satisfies

$$\text{Prob}[C_n(X_n) = f(X_n) | X_n \in R_n] < \frac{1}{2} + \epsilon(n).$$

Proof: We define the same random variables $\zeta_x = \zeta_x(R_n)$ as in the proof of Claim 12.2; that is, $\zeta_x(R_n) = 1$ if $x \in R_n$ and $\zeta_x(R_n) = 0$ otherwise. Also, as

[16] The current statement and its proof were somewhat revised.

before, $w_x \overset{\text{def}}{=} \text{Prob}[X_n = x]$, for every $x \in \{0,1\}^n$. Fixing any circuit $C_n$, let $C$ be the set of inputs on which $C_n$ correctly computes $f$; namely,

$$C \overset{\text{def}}{=} \{x : C_n(x) = f(x)\}. \tag{9}$$

For every choice of $R_n$, we are interested in the probability

$$\text{Prob}[X_n \in C | X_n \in R_n] = \frac{\text{Prob}[X_n \in C \wedge X_n \in R_n]}{\text{Prob}[X_n \in R_n]} . \tag{10}$$

We first determine the expected value of the numerator of Eq. (10), where the expactation is taken over the possible choices of $R_n$. We rewrite the numerator as $\sum_{x \in C} \zeta_x(R_n) \cdot w_x$, and lower bound it as follows

$$\begin{aligned} \text{E}\left[\sum_{x \in C} \zeta_x \cdot w_x\right] &= \sum_{x \in C} p(x) \cdot w_x \\ &= \sum_{x \in C} \frac{\rho(n) \cdot \text{Prob}[Y_n = x]}{\text{Prob}[X_n = x]} \cdot \text{Prob}[X_n = x] \\ &= \rho(n) \cdot \text{Prob}[Y_n \in C] \\ &\leq \rho(n) \cdot \left(\frac{1}{2} + \frac{\epsilon(n)}{2}\right) , \end{aligned}$$

where the last inequality is due to the hypothesis regarding $Y_n$. Next, we use a (multiplicative) Chernoff bound, and get

$$\begin{aligned} \text{Prob}\left[\sum_{x \in C} w_x \zeta_x > \left(\frac{1}{2} + \frac{2\epsilon(n)}{3}\right) \cdot \rho(n)\right] &< \exp\left(-\Omega\left(\frac{\epsilon(n)^2 \rho(n)}{\max_x\{w_x\}}\right)\right) \\ &< \exp\left(-\Omega\left(\frac{\epsilon(n)^2 s(n) \log_2 s(n)}{\text{poly}(n)}\right)\right) , \end{aligned}$$

where the last inequality uses the simplifying assumptions regarding the $w_x$'s and $s(n)$ (i.e., $w_x < \text{poly}(n)/s(n)$ and $\log_2 s(n) \leq n$). Thus, for all but at most a $\exp(-\text{poly}(n) \cdot s'(n) \log_2 s'(n))$ measure of the $R_n$'s, the numerator of Eq. (10) is at most $(\frac{1}{2} + \frac{2\epsilon(n)}{3}) \cdot \rho(n)$. This holds for each possible circuit of size $s'(n)$. Applying the union bound to the set of all $2^{s'(n)(O(1)+2\log_2 s'(n))}$ possible circuits of size $s'(n)$, we conclude that the probability that for some of these circuits the numerator of Eq. (10) is greater than $(\frac{1}{2} + \frac{2\epsilon(n)}{3}) \cdot \rho(n)$ is at most $\exp(-\text{poly}(n))$, where the probability is taken over the choice of $R_n$. Using Claim 12.2, we conclude that, for a similar measure of $R_n$'s, the denumerator of Eq. (10) is at least $(1 - \frac{\epsilon(n)}{3}) \cdot \rho(n)$. The claim follows. □

*Conclusion.* The lemma now follows by combining the foregoing three claims. Claim 12.1 provides us with a suitable $\mathbf{Y}$ for which we apply the probabilistic construction, whereas Claims 12.2 and 12.3 establish the existence of a set $R_n$ such that both

$$\text{Prob}[X_n \in R_n] > (1 - o(1)) \cdot \rho(n)$$

and

$$\text{Prob}[C_n(X_n) = f(X_n) | X_n \in R_n] < \frac{1}{2} + \epsilon(n)$$

holds for all possible circuits, $C_n$, of size $s'(n)$. The lemma follows. ■

## Appendix B: On the Selective XOR Lemma

Following [4, Exer. 7.17], we explicitly introduce a variant of the XOR Lemma, called the *Selective XOR Lemma.* Recall that the standard XOR Lemma refers to the predicate $P^{(t)}(x_1, ..., x_{t(n)}) \stackrel{\text{def}}{=} \prod_{i=1}^{t(n)} P(x_i)$, where $P$ is the original predicate and $x_i \in \{0,1\}^n$ for every $i$. Instead, the Selective XOR Lemma refers to the predicate $Q^{(t)}(x_1, ..., x_{t(n)}, S) \stackrel{\text{def}}{=} \prod_{i \in S} P(x_i)$, where the $x_i$'s are as before and $S \subseteq \{1, ..., t(n)\}$ is represented as an $t(n)$-bit long string. Thus, we have the following variant of Lemma 1.

**Lemma 13** (Selective XOR Lemma): *Let $P$ and $\mathbf{X} = \{X_n\}$ be as in Definition 1. For every function $t : \mathbb{N} \to \mathbb{N}$, define the predicate*

$$Q^{(t)}(x_1, ..., x_{t(n)}, S) \stackrel{\text{def}}{=} \prod_{i \in S} P(x_i) \ ,$$

*where $x_1, ..., x_{t(n)} \in \{0,1\}^n$ and $S \subseteq \{1, ..., t(n)\}$. Let $\mathbf{Y}^{(t)} \stackrel{\text{def}}{=} \{(X_n^{(t)}, U_{t(n)})\}$, where $X_n^{(t)}$ is as in Lemma 1 and $U_{t(n)}$ be a random variable that is independently and uniformly distributed over $\{0,1\}^{t(n)}$.*

(hypothesis) *As in Lemma 1; that is, suppose that for some function $s : \mathbb{N} \to \mathbb{N}$ and some bounded-away-from-1 function $\delta : \mathbb{N} \to [-1, +1]$, it holds that $\delta$ is an upper bound on the correlation of families of $s(\cdot)$-size circuits with $P$ over $\mathbf{X}$.*

(conclusion) *Analogously to Lemma 1, there exists a bounded-away-from-1 function $\delta' : \mathbb{N} \to [-1, +1]$ and a polynomial $p$ such that, for every function $t : \mathbb{N} \to \mathbb{N}$ and every function $\epsilon : \mathbb{N} \to [0, 1]$, the function*

$$\delta^{(t)}(n) \stackrel{\text{def}}{=} p(n) \cdot \delta'(n)^{t(n)} + \epsilon(n)$$

*is an upper bound on the correlation of families of $s'(\cdot)$-size circuits with $Q^{(t)}$ over $\mathbf{Y}^{(t)}$, where*

$$s'(t(n) \cdot n) \stackrel{\text{def}}{=} \text{poly}\left(\frac{\epsilon(n)}{n}\right) \cdot s(n) - \text{poly}(n \cdot t(n)).$$

In this appendix we discuss the relation of the Selective XOR Lemma to the XOR Lemma and to the Concatenation Lemma.

*The Selective XOR Lemma vs the Concatenation Lemma.* As shown in Section 5.2, the Concatenation Lemma implies the Selective XOR Lemma (by using Lemma 9). The opposite implication was recently shown in [16, Prop. 1.4]. The proof boils down to showing that any algorithm that computes the concatenation of the $t$ values, can be used to correlate the selective XOR as follows: On input $(x_1, ..., x_t, S)$, we obtain (from the algorithm) a guess $(b_1, ..., b_t)$ for $(P(x_1), ..., P(x_t))$, and output $b(S) \stackrel{\text{def}}{=} \prod_{i \in S} b_i$. Note that if $(b_1, ..., b_t) = (P(x_1), ..., P(x_t))$, then our answer $b(S)$ is correct for any $S$, whereas if $(b_1, ..., b_t) \neq (P(x_1), ..., P(x_t))$, then $\text{Prob}_S[b(S) = \prod_{i \in S} P(x_i)] = 1/2$. Thus, if the algorithm is correct with probability $p$, then our answer has correlation $p$ with $Q^{(t)}$.

*The Selective XOR Lemma implies the XOR Lemma.* This implication was sketched in Section 5.2, and we provide more details next. We show how to use an algorithm that correlates $P^{(t)}$ in order to correlate $Q^{(t)}$. We shall use $t$ random examples, denoted $(z_1, P(z_1)), ..., (z_t, P(z_t))$. On input $(x_1, ..., x_t, S)$, we set $x'_i = x_i$ if $i \in S$ and $x'_i = z_i$ otherwise, obtain (from the algorithm) a guess $b$ for $P^{(t)}(x'_1, ..., x'_t)$, and output $b \cdot \prod_{i \in [t] \setminus S} P(z_i)$. Thus, our answer is correct if and only if $b = P^{(t)}(x'_1, ..., x'_t)$, because $P^{(t)}(x'_1, ..., x'_t)$ equals $Q^{(t)}(x_1, ..., x_t, S) \cdot \prod_{i \in [t] \setminus S} P(z_i)$.

*The XOR Lemma implies the Selective XOR Lemma.* Following [16, Prop. 1.4], we show how to use an algorithm that correlates $Q^{(3t)}$ in order to correlate $P^{(t)}$. Here we shall use $3t$ random examples, denoted $(z_1, P(z_1)), ..., (z_{3t}, P(z_{3t}))$. On input $(x_1, ..., x_t)$, we select at random a subset $S \subseteq \{1, ..., 3t\}$, and let $i_1, ..., i_t$ be arbitrary $t$ distinct elements of $S$ (assuming that $|S| \geq t$). Next, we set $x'_{i_j} = x_j$ for every $j = 1, .., t$, and set $x'_i = z_i$ for every $i \in S'$, where $S' \stackrel{\text{def}}{=} \{1, ..., 3t\} \setminus \{i_j : j = 1, ..., t\}$. We obtain (from the algorithm) a guess $b$ for $Q^{(3t)}(x'_1, ..., x'_{3t}, S)$, and output $b \cdot \prod_{i \in S \setminus S'} P(z_i)$. Thus, our answer is correct if and only if $b = Q^{(3t)}(x'_1, ..., x'_{3t}, S)$, because $Q^{(3t)}(x'_1, ..., x'_{3t}, S)$ equals $P^{(t)}(x_1, ..., x_t) \cdot \prod_{i \in S \setminus S'} P(z_i)$. Note that this works assuming that $|S| \geq t$, which holds with probability $1 - 2^{-\Omega(t)}$. Thus, our correlation with $P^{(t)}$ is lower bounded by $p - 2^{-\Omega(t)}$, where $p$ is the correlation of the given algorithm with $Q^{(3t)}$.