A Tight Computational Indistinguishability Bound of Product Distributions[⋆]

Nathan Geier[0000−0002−1687−6950]

Tel Aviv University, Tel Aviv, Israel
nathangeier@mail.tau.ac.il

Abstract. Assume that distributions X_0, X_1 (respectively Y_0, Y_1) are d_X (respectively d_Y) indistinguishable for circuits of a given size. It is well known that the product distributions X_0Y_0, X_1Y_1 are d_X+d_Y indistinguishable for slightly smaller circuits. However, in probability theory where unbounded adversaries are considered through statistical distance, it is folklore knowledge that in fact X_0Y_0 and X_1Y_1 are $d_X+d_Y-d_X \cdot d_Y$ indistinguishable, and also that this bound is tight.

We formulate and prove the computational analog of this tight bound. Our proof is entirely different from the proof in the statistical case, which is non-constructive. As a corollary, we show that if X and Y are d indistinguishable, then k independent copies of X and k independent copies of Y are almost $1-(1-d)^k$ indistinguishable for smaller circuits, as against $d \cdot k$ using the looser bound.

Our bounds are useful in settings where only weak (i.e. non-negligible) indistinguishability is guaranteed. We demonstrate this in the context of cryptography, showing that our bounds, coupled with the XOR Lemma, yield straightforward computational generalization to the analysis for information-theoretic amplification of weak oblivious transfer protocols.

[⋆] Supported by the European Research Council (ERC) under the European Union's Horizon Europe research and innovation programme (grant agreement No. 101042417, acronym SPP), by ISF grant 18/484, and by Len Blavatnik and the Blavatnik Family Foundation.

1 Introduction

Computational indistinguishability is a fundamental concept in computational complexity and cryptography. One of the most basic bounds in this context, which is easy to see using a simple hybrid argument, is that for distributions X_0, X_1 of distance d_X, and Y_0, Y_1 of distance d_Y, with d_{XY} denoting the distance between X_0Y_0, X_1Y_1, we have that

$$d_{XY} \leq d_X + d_Y,$$

which holds both statistically and in the computational setting holds for slightly smaller circuits. However, in probability theory where statistical distance, or equivalently, indistinguishability against unbounded attackers is considered, it is folklore knowledge [9, Lemma 2.2] that a better, tight bound holds:

$$d_{XY} \leq d_X + d_Y - d_X \cdot d_Y.$$

It is tight in the sense that for every choice of d_X, d_Y, there exist distributions X_0, X_1 with distance d_X and distributions Y_0, Y_1 with distance d_Y, such that $d_{XY} = d_X + d_Y - d_X \cdot d_Y$. The proof of this bound uses coupling [7], and is thus inherently non-constructive and not easy to generalize to the computational setting. See Subsection 1.1 for more information on dealing with coupling in the computational setting.

It is worth noting here that another very important and foundational bound that is easy to show statistically but was not easily generalized to the computational setting is the famous XOR Lemma, see [5] for a survey. Our bounds are related in spirit and some of the techniques and statement formulations presented in this paper were inspired by Levin's proof of the XOR Lemma [10], and its presentation in [5]. Further, in Section 5 we show how both bounds are needed and complement each other in order to achieve the computational generalization to the information-theoretic weak OT amplification.

We provide a direct constructive proof of the tight bound which also works in the computational setting, both uniform and non-uniform, with an additive loss of ε which can be made as small as we want, by paying in increasing the running time or circuit size with relation to $1/\varepsilon$. To be more specific, for the non-uniform case, we roughly show that

Theorem 1 (Informal). *Let X_0, X_1 be d_X indistinguishable for size s_X circuits. (Respectively Y_0, Y_1, d_Y, s_Y.) Then, for every $k \in \mathbb{N}$, we have that (X_0, Y_0) and (X_1, Y_1) are $(d_X + d_Y - d_X \cdot d_Y + \varepsilon_k)$ indistinguishable for size s_k circuits, where*

$$\varepsilon_k \leq (d_Y)^k, \qquad s_k \approx \min\{s_Y, s_X/k\}.$$

Corollary 1 (Informal). *Let D, Q be distributions that are d indistinguishable for size s circuits. Then, for every $m \in \mathbb{N}$ and ε, we have that $D^{\otimes m}, Q^{\otimes m}$ are $(1 - (1-d)^m + \varepsilon)$ indistinguishable for size $s_{m,\varepsilon}$ circuits, where*

$$s_{m,\varepsilon} \approx s(1-d)^m / \log(1/\varepsilon).$$

And we also show similar results in the uniform setting, although with worse dependency on $1/\varepsilon$. The corollary essentially states that if the computational distance between X and Y is at most d, then the computational distance between the k-product of X and the k-product of Y is upper bounded by almost $1-(1-d)^k$ for smaller circuits, as against $d \cdot k$ resulted by the looser well known bound, which in particular may be larger than 1. The proof of the corollary follows by (carefully) applying the bound of the isolated case again and again. It should be noted that the difference between the bounds is especially interesting when k is not very small compared to $1/d$. For example, if $d = 0.5$, $k = 3$, the tight bound is 0.875 while the looser bound of $1.5 \geq 1$ is trivial.

We also demonstrate how these bounds may be used in the computational setting for amplification of weak oblivious transfer protocols [2, 13], providing an alternative straightforward analysis to the fact that the information-theoretic amplification process also works computationally. In general, when considering cryptographic primitives with multiple security properties, it is common that amplifying one property may degrade another, inducing a trade-off. We expect these bounds may be used in order to achieve a larger range of parameters when amplifying a weakened version of such primitives.

Finally, an interesting observation regarding the above corollary is how the circuit size grows only logarithmically with respect to $1/\varepsilon$. We discuss it further in the context of the amplification beyond negligible problem.

1.1 Related Work

While the aforementioned coupling technique itself is non-constructive, Maurer and Tessaro [11] show how to derive a computational analog for it using Holenstein's tight version of the hardcore lemma [6]. This approach could also be used to derive the tight bound in a general way. However, we believe our direct and specific approach still holds some advantages:

- Better parameters in the non-uniform setting: In our direct approach, when building a distinguisher for D, Q from a distinguisher between $D^{\otimes m}, Q^{\otimes m}$, the circuit size is multiplied by roughly $\log(1/\varepsilon)/(1-d)^m$. In contrast, using the hardcore approach the circuit size is multiplied by roughly
$$(1/\varepsilon)^2\, m^2 \left(\log|D| + \log|Q|\right).$$
Note that the latter must always be worse as $\varepsilon < (1-d)^m$ for the bound to be meaningful.
- Simplicity and explicitness: The distinguisher given by the hardcore lemma is somewhat more involved. In contrast, here the distinguisher is rather simple and easy to understand.

It should also be mentioned that the problem of tight direct product bounds has also been studied further in the statistical setting, when additional assumptions are made about the distributions. For example, see [12, 4].

1.2 Organization

We start by introducing basic definitions and notation in Section 2. We then continue to proving the non-uniform variants and their tightness in Section 3. We show how to generalize the non-uniform variants to the uniform setting in Section 4. We then demonstrate an application of these bounds in Section 5. Finally, in Section 6, we propose a conjecture aimed to capture the XOR analog to the observation made above regarding circuit size growth with relation to the slackness.

2 Definitions

For a distribution D, denote by $D^{\otimes k}$ the distribution of k independent copies of D. For distributions X_0, X_1 over Ω, a distinguisher is a boolean $A : \Omega \to \{0,1\}$, and we let $\mathrm{adv}^+_A(X_0, X_1) \coloneqq \mathbb{E}\left[A(X_1) - A(X_0)\right]$. (The expectation is also over A if it is not deterministic.) We say that distributions X_0, X_1 are d indistinguishable for size s circuits if for any such circuit C, we have that $\mathrm{adv}^+_C(X_0, X_1) \le d$. For distributions X, Y we will denote by (X, Y) the product distribution, given by two independent samples from X and Y. We denote by $B(p)$ the Bernoulli distribution with parameter p, and more generally by $B^\ell(p)$ the distribution that is equal to 1^ℓ with probability p and otherwise 0^ℓ. For a string s, we denote by $s[i]$ the i'th bit of s. We will denote by $[m]$ the set $\{1, \ldots, m\}$. We denote by $X_{1/2}$ the distribution given by $b \leftarrow \{0,1\}, x \leftarrow X_b$. An ensemble of distributions $X = \{X_n\}$ is efficiently samplable if there exists a uniform PPT sampler that given 1^n outputs a sample from X_n.

2.1 Notation

When the same distribution is used multiple times in a single expression, e.g. $(f(D), g(D))$ for D, it should be interpreted that a single value $d \leftarrow D$ is sampled and given to both f and g, rather than two independent samples.

3 The Non-Uniform Bounds and Tightness

Let us start with the non-uniform version as it is more simple and clean. The uniform version is a generalization of the ideas presented below. Roughly speaking, we show that given a distinguisher C for $(X_0, Y_0), (X_1, Y_1)$, if $C(x, \cdot)$ is not a good enough distinguisher between Y_0, Y_1 for all values of x, then we can build an amplifier for X_0, X_1 distinguishers. We then use this amplifier to turn the trivial distinguisher that always outputs 1 into a good enough distinguisher.

Theorem 2. *Let X_0, X_1 be distributions over ℓ_X bits that are d_X indistinguishable for size s_X circuits. (Respectively $Y_0, Y_1, \ell_Y, d_Y, s_Y$.) Then, for every $k \in \mathbb{N}$, we have that (X_0, Y_0) and (X_1, Y_1) are $(d_X + d_Y - d_X \cdot d_Y + \varepsilon_k)$ indistinguishable for size s_k circuits, where*

$$\varepsilon_k \coloneqq \frac{(d_Y)^k \cdot d_X\,(1 - d_Y)}{1 - (d_Y)^k} \le (d_Y)^k, \qquad s_k \coloneqq \min\left\{ s_Y - \ell_X, \frac{s_X - 1}{k} - 5\ell_Y - 1 \right\}.$$

Remark 1. We note that our starting point, $k = 1$, matches the simple hybrid argument bound of $d_X + d_Y$ since $\varepsilon_1 = d_X \cdot d_Y$, and as k grows larger our bound gets closer and closer to the tight bound of $d_X + d_Y - d_X \cdot d_Y$, while the circuits bound grows smaller. Also note that the bound is asymmetric with respect to the circuit size bounds. This asymmetry is important for preserving a similar circuit size when applying the isolated case over and over again. See a similar argument in [5, Section 3].

Proof. Assume toward contradiction that for some circuit C of size s_k, we have that

$$\mathrm{adv}_C^+\left(\left(X_0, Y_0\right), \left(X_1, Y_1\right)\right) > \left(d_X + d_Y - d_X \cdot d_Y + \varepsilon_k\right).$$

For every fixed x, it must be that $C(x, \cdot)$ is able to distinguish between Y_0 and Y_1 by at most d_Y, otherwise we get a contradiction as the size of this circuit is $s_k + \ell_X \leq s_Y$. Then, for every candidate distinguisher A between X_0 and X_1, we have that

$$\begin{aligned}\mathrm{adv}_C^+\left(\left(X_1, Y_{A(X_1)}\right), \left(X_1, Y_1\right)\right) &\leq d_Y \cdot \Pr\left[A(X_1) = 0\right]\\ \mathrm{adv}_C^+\left(\left(X_0, Y_0\right), \left(X_0, Y_{A(X_0)}\right)\right) &\leq d_Y \cdot \Pr\left[A(X_0) = 1\right]\end{aligned}$$

where $x, y \leftarrow X_1, Y_{A(X_1)}$ is resulted by $x \leftarrow X_1$, $b \leftarrow A(x)$, $y \leftarrow Y_b$. This holds because

$$\begin{aligned}&\mathrm{adv}_C^+\left(\left(X_1, Y_{A(X_1)}\right), \left(X_1, Y_1\right)\right) = \mathbb{E}\left[C(X_1, Y_1) - C(X_1, Y_{A(X_1)})\right] =\\ &= \mathbb{E}\left[C(X_1, Y_1) - C(X_1, Y_0) | A(X_1) = 0\right] \cdot \Pr\left[A(X_1) = 0\right] +\\ &+ \mathbb{E}\left[C(X_1, Y_1) - C(X_1, Y_1) | A(X_1) = 1\right] \cdot \Pr\left[A(X_1) = 1\right] =\\ &= \mathbb{E}_{x \leftarrow X_1 | A(X_1)=0}\left[C(x, Y_1) - C(x, Y_0)\right] \cdot \Pr\left[A(X_1) = 0\right] =\\ &= \mathbb{E}_{x \leftarrow X_1 | A(X_1)=0}\left[\mathrm{adv}_{C(x,\cdot)}^+\left(Y_0, Y_1\right)\right] \cdot \Pr\left[A(X_1) = 0\right] \leq d_Y \cdot \Pr\left[A(X_1) = 0\right]\end{aligned}$$

and using a symmetric argument for the second inequality. Using that (in general)

$$\sum_{i \in [n]} \mathrm{adv}_C^+(D_i, D_{i+1}) = \mathrm{adv}_C^+(D_1, D_{n+1})$$

we conclude that

$$\begin{aligned}&\mathrm{adv}_C^+\left(\left(X_0, Y_0\right), \left(X_1, Y_1\right)\right) = \mathrm{adv}_C^+\left(\left(X_0, Y_0\right), \left(X_0, Y_{A(X_0)}\right)\right) +\\ &+ \mathrm{adv}_C^+\left(\left(X_0, Y_{A(X_0)}\right), \left(X_1, Y_{A(X_1)}\right)\right) + \mathrm{adv}_C^+\left(\left(X_1, Y_{A(X_1)}\right), \left(X_1, Y_1\right)\right)\end{aligned}$$

and thus

$$\begin{aligned}
&\mathrm{adv}^{+}_{C}\left(\left(X_0, Y_{A(X_0)}\right), \left(X_1, Y_{A(X_1)}\right)\right) = \mathrm{adv}^{+}_{C}\left(\left(X_0, Y_0\right), \left(X_1, Y_1\right)\right) - \\
&- \mathrm{adv}^{+}_{C}\left(\left(X_1, Y_{A(X_1)}\right), \left(X_1, Y_1\right)\right) - \mathrm{adv}^{+}_{C}\left(\left(X_0, Y_0\right), \left(X_0, Y_{A(X_0)}\right)\right) > \\
&> \left(d_X + d_Y - d_X \cdot d_Y + \varepsilon_k\right) - \left(d_Y \cdot \Pr\left[A(X_1) = 0\right]\right) - \left(d_Y \cdot \Pr\left[A(X_0) = 1\right]\right) = \\
&= \left(d_X - d_X \cdot d_Y + \varepsilon_k\right) + d_Y\left(1 - \Pr\left[A(X_1) = 0\right] - \Pr\left[A(X_0) = 1\right]\right) = \\
&= \left(d_X - d_X \cdot d_Y + \varepsilon_k\right) + d_Y\left(\Pr\left[A(X_1) = 1\right] - \Pr\left[A(X_0) = 1\right]\right) = \\
&= \left(d_X - d_X \cdot d_Y + \varepsilon_k\right) + d_Y\left(\mathbb{E}\left[A(X_1)\right] - \mathbb{E}\left[A(X_0)\right]\right) = \\
&= \left(d_X - d_X \cdot d_Y + \varepsilon_k\right) + d_Y \cdot \mathrm{adv}^{+}_{A}\left(X_0, X_1\right).
\end{aligned}$$

In other words, we can build a new distinguisher A' for X_0, X_1 by applying A to our input x, sampling $y \leftarrow Y_{A(x)}$ and feeding (x, y) to C, and have that

$$\mathrm{adv}^{+}_{A'}\left(X_0, X_1\right) > \left(d_X - d_X \cdot d_Y + \varepsilon_k\right) + d_Y \cdot \mathrm{adv}^{+}_{A}\left(X_0, X_1\right).$$

If we start from A_0 being the trivial distinguisher that always outputs 1 and keep repeating this process for k steps, we get that

$$\begin{aligned}
&\mathrm{adv}^{+}_{A_k}\left(X_0, X_1\right) > \left(d_X - d_X \cdot d_Y + \varepsilon_k\right) + d_Y \cdot \mathrm{adv}^{+}_{A_{k-1}}\left(X_0, X_1\right) > \\
&> \left(d_X - d_X \cdot d_Y + \varepsilon_k\right) + d_Y \cdot \left(d_X - d_X \cdot d_Y + \varepsilon_k\right) + \left(d_Y\right)^2 \cdot \mathrm{adv}^{+}_{A_{k-2}}\left(X_0, X_1\right) > \\
&> \cdots > \left(d_X - d_X \cdot d_Y + \varepsilon_k\right) \sum_{i=0}^{k-1} \left(d_Y\right)^i + \left(d_Y\right)^k \cdot \mathrm{adv}^{+}_{A_0}\left(X_0, X_1\right) = \\
&= \left(d_X - d_X \cdot d_Y + \varepsilon_k\right) \sum_{i=0}^{k-1} \left(d_Y\right)^i = \frac{\left(d_X - d_X \cdot d_Y + \varepsilon_k\right)\left(1 - \left(d_Y\right)^k\right)}{1 - d_Y} = \\
&= \frac{\left(d_X\left(1 - d_Y\right) + \frac{\left(d_Y\right)^k \cdot d_X\left(1 - d_Y\right)}{1 - \left(d_Y\right)^k}\right)\left(1 - \left(d_Y\right)^k\right)}{1 - d_Y} = \left(d_X + \frac{\left(d_Y\right)^k \cdot d_X}{1 - \left(d_Y\right)^k}\right)\left(1 - \left(d_Y\right)^k\right) = \\
&= d_X\left(1 - \left(d_Y\right)^k\right) + \left(d_Y\right)^k \cdot d_X = d_X.
\end{aligned}$$

And so, we have concluded that A_k distinguishes X_0 from X_1 with advantage better than d_X. Next, for the circuit size, in order to implement A_k we start by applying A_{k-1}, sample $y_0 \leftarrow Y_0, y_1 \leftarrow Y_1$, use a multiplexer to choose $y \leftarrow y_b$ where b is the output gate of A_{k-1}, and finally use the circuit C. Instead of sampling y_0, y_1, we can simply use non-uniformity to hard-code the best samples, at the cost of $2\ell_Y$ gates. Implementing the multiplexer can be done using $3\ell_Y + 1$ gates, with one gate computing $\neg b$ and for every $i \in [\ell_Y]$ another 3 gates to compute $y[i] = (y_0[i] \wedge \neg b) \vee (y_1[i] \wedge b)$. Overall, we conclude that $\mathrm{size}(A_k) = \mathrm{size}(A_{k-1}) + 5\ell_Y + 1 + s_k$ and therefore

$$\mathrm{size}(A_k) = \mathrm{size}(A_0) + k \cdot \left(5\ell_Y + 1 + s_k\right) \leq 1 + k \cdot \left(5\ell_Y + 1 + \left(\frac{s_X - 1}{k} - 5\ell_Y - 1\right)\right) = s_X$$

which is a contradiction to our assumption that d_X is an upper bound on the advantage of size s_X circuits distinguishing X_0 from X_1.

3.1 The N-Fold Case

Corollary 2. *Let D, Q be distributions over ℓ bits that are d indistinguishable for size s circuits. Then, for every $m \in \mathbb{N}$ and ε, we have that $D^{\otimes m}, Q^{\otimes m}$ are $(1-(1-d)^m+\varepsilon)$ indistinguishable for size $s_{m,\varepsilon}$ circuits, where*

$$s_{m,\varepsilon} = \frac{s-1}{k_{m,\varepsilon}} - 5m\ell - 1, \qquad k_{m,\varepsilon} = \left\lceil \frac{\log(d\varepsilon)}{\log(1-(1-d)^m+\varepsilon)} \right\rceil \leq \left\lceil \frac{\log(1/d\varepsilon)}{(1-d)^m-\varepsilon} \right\rceil .$$

Proof. If $\varepsilon \geq (1-d)^m$ the statement is trivially true. Otherwise, we start from D, Q and use Theorem 2 to repeatedly add copies of D, Q for $m-1$ times, using $k_{m,\varepsilon}$ set at the statement, where each time the added copy of D, Q is treated as X_0, X_1 and $D^{\otimes i}, Q^{\otimes i}$ are treated as Y_0, Y_1. Let d_i denote the bound on the advantage of i copies, then we have that $d_1 = d$ and $d_i \leq d_{i-1} + d - d_{i-1} \cdot d + (d_{i-1})^{k_{m,\varepsilon}}$. We can see by induction that $d_i \leq 1-(1-d)^i+\varepsilon$ for $i \in [m]$ as

$$\begin{aligned}
&d_i \leq d_{i-1} + d - d_{i-1} \cdot d + (d_{i-1})^{k_{m,\varepsilon}} = (1-d)d_{i-1} + d + (d_{i-1})^{k_{m,\varepsilon}} \leq \\
&\leq (1-d)\left(1-(1-d)^{i-1}+\varepsilon\right) + d + \left(1-(1-d)^{i-1}+\varepsilon\right)^{k_{m,\varepsilon}} = \\
&= 1-d-(1-d)^i+(1-d)\varepsilon + d + \left(1-(1-d)^{i-1}+\varepsilon\right)^{k_{m,\varepsilon}} = \\
&= 1-(1-d)^i+(1-d)\varepsilon + \left(1-(1-d)^{i-1}+\varepsilon\right)^{k_{m,\varepsilon}} \leq \\
&\leq 1-(1-d)^i+(1-d)\varepsilon + (1-(1-d)^m+\varepsilon)^{k_{m,\varepsilon}} \leq 1-(1-d)^i+\varepsilon
\end{aligned}$$

where in the last inequality we used the choice of $k_{m,\varepsilon}$. For the circuit size, we can easily see by induction on i that $s_{i,\varepsilon} \geq (s-1)/k_{m,\varepsilon} - 5i\ell - 1$, as we have that $s_{1,\varepsilon} = s$ and

$$\begin{aligned}
&s_{i,\varepsilon} \geq \min\left\{ s_{(i-1),\varepsilon} - \ell, \frac{s-1}{k_{m,\varepsilon}} - 5(i-1)\ell - 1 \right\} \geq \\
&\geq \min\left\{ \frac{s-1}{k_{m,\varepsilon}} - 5(i-1)\ell - 1 - \ell, \frac{s-1}{k_{m,\varepsilon}} - 5(i-1)\ell - 1 \right\} \geq \frac{s-1}{k_{m,\varepsilon}} - 5i\ell - 1.
\end{aligned}$$

3.2 Tightness

This is somewhat folklore knowledge, that we explicitly state for the sake of completeness. We show that for every choice of $d_X, d_Y, s_X, s_Y, \ell_X, \ell_Y$ there exist two pairs of distributions X_0, X_1 and Y_0, Y_1, such that X_0, X_1 are over ℓ_X bits and cannot be distinguished with advantage better than d_X by size s_X circuits (resp. for Y_0, Y_1 with ℓ_Y, d_Y, s_Y), yet (X_0, Y_0) and (X_1, X_1) can be distinguished with advantage $d_X + d_Y - d_X \cdot d_Y$ using a size 1 circuit. For the n-fold case, we show that for every choice of d, s, ℓ there exist distributions X, Y over ℓ bits with distance at most d against s-sized circuits, such that $X^{\otimes k}, Y^{\otimes k}$ can be distinguished with advantage $1-(1-d)^k$ using a circuit of size $2k-1$. We will use statistical distance in these examples, noting that the statistical distance between distributions is equal to the maximal advantage of unbounded adversaries distinguishing between them, and that the statistical distance from a constant variable is equal to the probability to differ from it.

For the isolated case, we let $X_0 \equiv 0^{\ell_X}$, $X_1 := B(d_X)^{\ell_X}$, $Y_0 \equiv 0^{\ell_Y}$, $Y_1 := B(d_Y)^{\ell_Y}$, where $B(p)^\ell$ denotes sampling from $B(p)$ and outputting ℓ copies of the result. We have that size s_X circuits can distinguish between X_0, X_1 with advantage at most d_X (resp. for Y_0, Y_1 with s_Y, d_Y) as this is the statistical distance between them. Also, it is easy to verify that the simple size 1 circuit which given (x, y) computes $x[1] \vee y[1]$ distinguishes between (X_0, Y_0) and (X_1, Y_1) with advantage $1 - (1 - d_X)(1 - d_Y) = d_X + d_Y - d_X \cdot d_Y$.

For the n-fold case, let $X \equiv 0^\ell$, $Y := B(d)^\ell$, then size s circuits can distinguish X from Y with advantage at most d. Yet, the circuit of size $2k - 1$ which given $(z_1, \ldots, z_k)$ computes $\vee_i z_i[1]$ (using a full binary tree of OR gates) distinguishes between $X^{\otimes k}$ and $Y^{\otimes k}$ with advantage $1 - (1 - d)^k$.

4 The Uniform Variant

We used non-uniformity two times in the proof of Theorem 2. The second time, which is easier to deal with, is in the circuit size analysis where we hard-coded the best samples of y_0, y_1 to each iteration of A_i. Instead, in the uniform version, we will use uniform samplers of Y_0, Y_1.

The first use of non-uniformity was when we assumed that $C(x, \cdot)$ is at most a d_Y-distinguisher between Y_0 and Y_1, for every fixed x, otherwise we can use non-uniformity to be done. More specifically, we used this assumption to get that

$$\mathrm{adv}_C^+\left(\left(X_1, Y_{A(X_1)}\right), (X_1, Y_1)\right) \le d_Y \cdot \Pr\left[A(X_1) = 0\right].$$

For the uniform case, we will relax this condition to x not being easy to hard-code, in the following sense:

$$\Pr_{x \leftarrow X_{1/2}}\left[\mathrm{adv}_{C(x,\cdot)}^+\left(Y_0, Y_1\right) > d_Y + \varepsilon_k\right] \le \varepsilon_k$$

where $X_{1/2}$ is given by $b \leftarrow \{0, 1\}, x \leftarrow X_b$. If this condition does not hold then we can efficiently compute a good x, except for negligible probability, assuming that efficient uniform samplers for X_0, X_1, Y_0, Y_1 exist. Otherwise, we will see that

$$\mathrm{adv}_C^+\left(\left(X_1, Y_{A(X_1)}\right), (X_1, Y_1)\right) \le d_Y \cdot \Pr\left[A(X_1) = 0\right] + 3\varepsilon_k$$

and so almost the same argument from the non-uniform case works, except that now we lose another small additive term. Let us state and prove this more formally:

Lemma 1. *Let $X_0 = \{X_{0,n}\}, X_1 = \{X_{1,n}\}, Y_0 = \{Y_{0,n}\}, Y_1 = \{Y_{1,n}\}$ be ensembles of efficiently samplable distributions, and $d_X(n), d_Y(n)$ be efficiently computable functions between 0 and 1. Then, for every $k \in \mathbb{N}$ and time $t(n)$ Turing* *machine M distinguishing (X_0, Y_0) from (X_1, Y_1) infinitely often with advantage at least $(d_X + d_Y - d_X \cdot d_Y + 7\varepsilon_k)$ for*

$$\varepsilon_k := \frac{(d_Y)^k \cdot d_X\,(1 - d_Y)}{1 - (d_Y)^k} \le (d_Y)^k,$$

we have that either M efficiently yields a distinguisher for Y_0, Y_1 through a hard-coding of x, in the sense that for infinitely many n's

$$\Pr_{x \leftarrow X_{1/2}}\left[\operatorname{adv}^+_{M(1^n, x, \cdot)}(Y_0, Y_1) > d_Y + \varepsilon_k\right] > \varepsilon_k,$$

or there exists a time $t \cdot \mathrm{poly}(nk)$ infinitely often distinguisher between X_0, X_1 with advantage at least d_X.

Proof. For the sake of notational ease, we will drop the asymptotic notation and replace $M(1^n)$ with C. Assume that for all but finitely many n's,

$$\Pr_{x \leftarrow X_{1/2}}\left[\operatorname{adv}^+_{C(x, \cdot)}(Y_0, Y_1) > d_Y + \varepsilon_k\right] \le \varepsilon_k.$$

Then, for every candidate distinguisher A between X_0 and X_1, for all but finitely many n's, we have that

$$\operatorname{adv}^+_C\left(\left(X_1, Y_{A(X_1)}\right), (X_1, Y_1)\right) \le d_Y \cdot \Pr\left[A(X_1) = 0\right] + 3\varepsilon_k$$
$$\operatorname{adv}^+_C\left((X_0, Y_0), \left(X_0, Y_{A(X_0)}\right)\right) \le d_Y \cdot \Pr\left[A(X_0) = 1\right] + 3\varepsilon_k$$

where $x, y \leftarrow X_1, Y_{A(X_1)}$ is resulted by $x \leftarrow X_1$, $b \leftarrow A(x)$, $y \leftarrow Y_b$. To see this, we first note that

$$\varepsilon_k \ge \Pr_{x \leftarrow X_{1/2}}\left[\operatorname{adv}^+_{C(x, \cdot)}(Y_0, Y_1) > d_Y + \varepsilon_k\right] \ge$$
$$\ge \frac{1}{2}\Pr\left[A(X_1) = 0\right] \Pr_{x \leftarrow X_1 | A(X_1) = 0}\left[\operatorname{adv}^+_{C(x, \cdot)}(Y_0, Y_1) > d_Y + \varepsilon_k\right]$$

which implies that

$$\mathbb{E}_{x \leftarrow X_1 | A(X_1) = 0}\left[\operatorname{adv}^+_{C(x, \cdot)}(Y_0, Y_1)\right] \le d_Y + \varepsilon_k + \frac{2\varepsilon_k}{\Pr\left[A(X_1) = 0\right]} \le d_Y + \frac{3\varepsilon_k}{\Pr\left[A(X_1) = 0\right]}.$$

Plugging it into the last inequality in the following, we get

$$\begin{aligned}
&\operatorname{adv}^+_C\left(\left(X_1, Y_{A(X_1)}\right), (X_1, Y_1)\right) = \mathbb{E}\left[C(X_1, Y_1) - C(X_1, Y_{A(X_1)})\right] = \\
&= \mathbb{E}\left[C(X_1, Y_1) - C(X_1, Y_0) | A(X_1) = 0\right] \cdot \Pr\left[A(X_1) = 0\right] + \\
&+ \mathbb{E}\left[C(X_1, Y_1) - C(X_1, Y_1) | A(X_1) = 1\right] \cdot \Pr\left[A(X_1) = 1\right] = \\
&= \mathbb{E}_{x \leftarrow X_1 | A(X_1) = 0}\left[C(x, Y_1) - C(x, Y_0)\right] \cdot \Pr\left[A(X_1) = 0\right] = \\
&= \mathbb{E}_{x \leftarrow X_1 | A(X_1) = 0}\left[\operatorname{adv}^+_{C(x, \cdot)}(Y_0, Y_1)\right] \cdot \Pr\left[A(X_1) = 0\right] \le d_Y \cdot \Pr\left[A(X_1) = 0\right] + 3\varepsilon_k
\end{aligned}$$

and use a symmetric argument for the second upper bound. Using that (in general)

$$\sum_{i\in[n]} \mathrm{adv}^{+}_{C}(D_i, D_{i+1}) = \mathrm{adv}^{+}_{C}(D_1, D_{n+1})$$

we conclude that

$$\begin{aligned}&\mathrm{adv}^{+}_{C}\left((X_0, Y_0), (X_1, Y_1)\right) = \mathrm{adv}^{+}_{C}\left((X_0, Y_0), \left(X_0, Y_{A(X_0)}\right)\right) + \\ &+ \mathrm{adv}^{+}_{C}\left(\left(X_0, Y_{A(X_0)}\right), \left(X_1, Y_{A(X_1)}\right)\right) + \mathrm{adv}^{+}_{C}\left(\left(X_1, Y_{A(X_1)}\right), (X_1, Y_1)\right)\end{aligned}$$

and thus

$$\begin{aligned}&\mathrm{adv}^{+}_{C}\left(\left(X_0, Y_{A(X_0)}\right), \left(X_1, Y_{A(X_1)}\right)\right) = \mathrm{adv}^{+}_{C}\left((X_0, Y_0), (X_1, Y_1)\right) - \\ &- \mathrm{adv}^{+}_{C}\left(\left(X_1, Y_{A(X_1)}\right), (X_1, Y_1)\right) - \mathrm{adv}^{+}_{C}\left((X_0, Y_0), \left(X_0, Y_{A(X_0)}\right)\right) > \\ &> (d_X + d_Y - d_X \cdot d_Y + 7\varepsilon_k) - (d_Y \cdot \Pr[A(X_1) = 0] + 3\varepsilon_k) - (d_Y \cdot \Pr[A(X_0) = 1] + 3\varepsilon_k) = \\ &= (d_X - d_X \cdot d_Y + \varepsilon_k) + d_Y(1 - \Pr[A(X_1) = 0] - \Pr[A(X_0) = 1]) = \\ &= (d_X - d_X \cdot d_Y + \varepsilon_k) + d_Y(\Pr[A(X_1) = 1] - \Pr[A(X_0) = 1]) = \\ &= (d_X - d_X \cdot d_Y + \varepsilon_k) + d_Y(\mathbb{E}[A(X_1)] - \mathbb{E}[A(X_0)]) = \\ &= (d_X - d_X \cdot d_Y + \varepsilon_k) + d_Y \cdot \mathrm{adv}^{+}_{A}(X_0, X_1).\end{aligned}$$

In other words, we can build a new distinguisher A' for X_0, X_1 by applying A to our input x, sampling $y \leftarrow Y_{A(x)}$ and feeding (x, y) to C, and have that

$$\mathrm{adv}^{+}_{A'}(X_0, X_1) > (d_X - d_X \cdot d_Y + \varepsilon_k) + d_Y \cdot \mathrm{adv}^{+}_{A}(X_0, X_1).$$

If we start from A_0 being the trivial distinguisher that always outputs 1 and keep repeating this process for k steps, we get that

$$\begin{aligned}&\mathrm{adv}^{+}_{A_k}(X_0, X_1) > (d_X - d_X \cdot d_Y + \varepsilon_k) + d_Y \cdot \mathrm{adv}^{+}_{A_{k-1}}(X_0, X_1) > \\ &> (d_X - d_X \cdot d_Y + \varepsilon_k) + d_Y \cdot (d_X - d_X \cdot d_Y + \varepsilon_k) + (d_Y)^2 \cdot \mathrm{adv}^{+}_{A_{k-2}}(X_0, X_1) > \\ &> \cdots > (d_X - d_X \cdot d_Y + \varepsilon_k) \sum_{i=0}^{k-1} (d_Y)^i + (d_Y)^k \cdot \mathrm{adv}^{+}_{A_0}(X_0, X_1) = \\ &= (d_X - d_X \cdot d_Y + \varepsilon_k) \sum_{i=0}^{k-1} (d_Y)^i = \frac{(d_X - d_X \cdot d_Y + \varepsilon_k)\left(1 - (d_Y)^k\right)}{1 - d_Y} = \\ &= \frac{\left(d_X(1 - d_Y) + \frac{(d_Y)^k \cdot d_X(1-d_Y)}{1-(d_Y)^k}\right)\left(1 - (d_Y)^k\right)}{1 - d_Y} = \left(d_X + \frac{(d_Y)^k \cdot d_X}{1 - (d_Y)^k}\right)\left(1 - (d_Y)^k\right) = \\ &= d_X\left(1 - (d_Y)^k\right) + (d_Y)^k \cdot d_X = d_X.\end{aligned}$$

And so, we have concluded that A_k distinguishes X_0 from X_1 with advantage better than d_X. In order to implement A_k we need to run C, sample Y_0, Y_1 and use a multiplexer, for k times, so we conclude that $\mathrm{time}(A_k) = t \cdot \mathrm{poly}(n, k)$.

Remark 2. In particular, we can use this lemma to show that if X_0, X_1 are d_X ind. and Y_0, Y_1 are d_Y ind. then (X_0, Y_0) and (X_1, Y_1) are $d_X + d_Y - d_X \cdot d_Y + 7\varepsilon_k$ ind. for Turing machines with running time of

$$t = \min\{t_X/\mathrm{poly}(n, k), t_Y/\mathrm{poly}(n, 1/\varepsilon_k)\},$$

which may be good enough for a constant number of uses, but does not work well beyond that, as every use costs us a division of the time bound by a polynomial. This is why we cannot prove the n-fold case immediately by repeatedly applying Lemma 1. The key idea is that we do not need to keep resampling and testing over and over again, but instead, once we find a good enough x in the i'th coordinate, we fix it for the rest of the process, or if the hard-coding of the i'th coordinate does not succeed, the above lemma states we can distinguish there.

Theorem 3. *Let $X = \{X_n\}, Y = \{Y_n\}$ be ensembles of efficiently samplable distributions that are $d(n)$ indistinguishable for time $t(n)$ Turing machines. Then, for every $m = m(n)$, we have that $X^{\otimes m}$ and $Y^{\otimes m}$ are $(1 - (1 - d)^m + 7m\varepsilon)$ indistinguishable for time $t_{m,\varepsilon}$ Turing machines, where*

$$t_{m,\varepsilon} = t/\mathrm{poly}(n, m, k_{m,\varepsilon}, 1/\varepsilon), \quad k_{m,\varepsilon} = \left\lceil \frac{\log(\varepsilon)}{\log(1 - (1-d)^m + 7m\varepsilon)} \right\rceil \leq \left\lceil \frac{\log(1/\varepsilon)}{(1-d)^m - 7m\varepsilon} \right\rceil.$$

Proof. For $i = 0, 1, \ldots, m-1$, we try to hard-code the $m - i$'th coordinate using $\mathrm{poly}(n, 1/\varepsilon)$ samples, and getting a distinguisher for $X^{\otimes m-i}, Y^{\otimes m-i}$ with advantage of at least $1 - (1-d)^{m-i} + 7(m-i)\varepsilon$ except for negligible probability (the probability that the estimate was good but not truthful to the expectation) until for some i we fail to find a good value to hard-code (if we reached $i = m-1$ and succeeded then we are done). Once we fail, we apply the isolated case of Lemma 1, which essentially states that if the hard-coding of X, Y into such circuit failed, then one can build a distinguisher for them, and we are done.

Let us be more explicit about how we sample and hard-code the $m - i$'th coordinate: We are given (except for negligible probability) good samples for the coordinates in $m - i + 1, \ldots, m$ and hard-code them into A, getting a $1 - (1-d)^{m-i} + 7(m-i)\varepsilon$ distinguisher for $X^{\otimes m-i}, Y^{\otimes m-i}$, which we view as the product of $X^{\otimes m-i-1}, Y^{\otimes m-i-1}$ with X, Y. We first note that our choice of k guarantees that $\varepsilon_k \leq \varepsilon$ for all $1 - (1-d)^{m-i} + 7(m-i)\varepsilon$. We start by trying to work under the "hard-coding" assumption that

$$\Pr_{z \leftarrow X/Y}\left[\mathrm{adv}^{+}_{A(z,\cdot)}\left(X^{\otimes m-i-1}, Y^{\otimes m-i-1}\right) > 1 - (1-d)^{m-i-1} + 7(m-i-1)\varepsilon + \varepsilon\right] > \varepsilon$$

and generate a distinguisher for $X^{\otimes m-i-1}, Y^{\otimes m-i-1}$ as follows: Keep sampling $z \leftarrow X/Y$ and estimating $\mathrm{adv}^{+}_{A(z,\cdot)}\left(X^{\otimes m-i-1}, Y^{\otimes m-i-1}\right)$ using r samples from $X^{\otimes m-i-1}/Y^{\otimes m-i-1}$, until we succeed in finding z with an estimate of at least $1 - (1-d)^{m-i-1} + 7(m-i-1)\varepsilon + 0.5\varepsilon$, then fix this good z in this coordinate and move forward, or stop after q tries if no such z has been found. Using Hoeffding's inequality, for every z, the probability that the estimate's error is greater than

$\varepsilon/2$ is at most $2e^{-r\cdot(\varepsilon/2)^2/2}$. If all estimates were $\varepsilon/2$ accurate and a good z has been drawn, the process succeeds in finding a z with advantage of at least $1-(1-d)^{m-i-1}+7(m-i-1)\varepsilon$ and we can move on, so our probability to fail at that, under the above assumption, is at most

$$q\cdot 2e^{-r\cdot\varepsilon^2/32}+(1-\varepsilon)^q \le 2e^{\log(q/2)-r\cdot\varepsilon^2/32}+e^{-q\cdot\varepsilon}\le \mathrm{neg}(n)$$

by choosing, say,

$$q=n/\varepsilon=\mathrm{poly}(n,1/\varepsilon),\quad r=64n/\varepsilon^3>(\log(q/2)+n)\cdot 32/\varepsilon^2=\mathrm{poly}(n,1/\varepsilon).$$

Hence paying with a time complexity of $t_{m,\varepsilon}\cdot\mathrm{poly}(n,1/\varepsilon)$ for every coordinate.

If we could not find a good z, we use Lemma 1: If we can distinguish $X^{\otimes m-i}, Y^{\otimes m-i}$ with advantage

$$\begin{aligned}&(1-d)\left(1-(1-d)^{m-i-1}+7(m-i-1)\varepsilon\right)+d+7\varepsilon=\\&=1-(1-d)^{m-i}+(1-d)7(m-i-1)\varepsilon+7\varepsilon\le\\&\le 1-(1-d)^{m-i}+7(m-i)\varepsilon\le \mathrm{adv}_A^{+}\left(X^{\otimes m-i},Y^{\otimes m-i}\right)\end{aligned}$$

and the assumption about finding a good z to hard-code for $X^{\otimes m-i-1}, Y^{\otimes m-i-1}$ does not hold, then we can build a d-distinguisher for X, Y in time $t_{m,\varepsilon}\cdot\mathrm{poly}(n,k)$. The probability that at some point in the process we failed to hard-code a good z at the $m-i$'th coordinate even though the assumption held is $m(n)\cdot\mathrm{neg}(n)=\mathrm{neg}(n)$.

We remark this proof is easily generalized to the case where not all pairs in the product are identical, that is, for $\bigotimes X_i$ and $\bigotimes Y_i$, with a distance bound of $(1-\prod_i(1-d_i)+7m\varepsilon)$.

5 Applications

As an application, we consider the amplification of weak oblivious transfer protocols. We briefly explain how our bounds, paired with Yao's XOR Lemma, yield a natural generalization in the computational setting to the amplification process presented in [2, Subsection 4.3]. We note that it was already shown, using The Hardcore Theorem [8, 1], that the same amplification process also works computationally [13]. Yet we find our constructive and explicit approach more natural and straightforward.

For the sake of simplicity, let us consider the amplification of error-less (p,q)-weak semi-honest 1-2 OT: The receiver with bit c is trying to learn b_c, where (b_0,b_1) is the database of the sender. We say the protocol is (p,q) weak if the view of the sender when $c=0$ is p-indistinguishable from its view when $c=1$ (equivalently, c is at most p-correlated to the view of the sender), and the view of the receiver when $b_{\overline{c}}=0$ is q-indistinguishable from its view when $b_{\overline{c}}=1$.

In [2, Subsection 4.2], two fundamental operations that will be used as building blocks in the amplification process are presented. One is an operation called

S-Reduce that amplifies indistinguishability against the sender but worsens indistinguishability against the receiver, and the other is an operation called R-Reduce that amplifies indistinguishability against the receiver but worsens indistinguishability against the sender. Both of them work using secret sharing over multiple applications of the underlying protocol, in the first the receiver's choice bit is secret shared and in the other, the sender's database. They receive a weak protocol $\mathcal{W}$ together with a parameter k and work as follows:

S-**Reduce**$(k, \mathcal{W})$

```
Inputs: $c$, $(b_0, b_1)$
The receiver splits $c$ randomly into $k$ shares $\{c_i\}_{i=1}^k$ conditioned on $\oplus_{i=1}^k c_i = c$.
The sender splits $b_0$ randomly into $k$ shares $\{b_{0i}\}_{i=1}^k$ conditioned on $\oplus_{i=1}^k b_{0i} = b_0$,
    and sets $b_{1i} = b_{0i} \oplus b_0 \oplus b_1$.
for $i = 1$ to $k$ do
    Run $\mathcal{W}$ with $c_i$, $(b_{0i}, b_{1i})$.
end for
The receiver outputs the XOR of all $k$ received bits, that is, $\oplus_{i=1}^k b_{c_i, i}$.
```

R-**Reduce**$(k, \mathcal{W})$

```
Inputs: $c$, $(b_0, b_1)$
The receiver sets $c_i = c$ for $i \in [k]$.
The sender splits $b_0$ randomly into $k$ shares $\{b_{0i}\}_{i=1}^k$ conditioned on $\oplus_{i=1}^k b_{0i} = b_0$,
    and also splits $b_1$ randomly into $k$ shares $\{b_{1i}\}_{i=1}^k$ conditioned on $\oplus_{i=1}^k b_{1i} = b_1$.
for $i = 1$ to $k$ do
    Run $\mathcal{W}$ with $c_i$, $(b_{0i}, b_{1i})$.
end for
The receiver outputs the XOR of all $k$ received bits, that is, $\oplus_{i=1}^k b_{c_i, i}$.
```

Correctness of R-Reduce is straightforward. For S-Reduce, note that in the i'th call the receiver learns $b_{0i} \oplus c_i \cdot (b_0 \oplus b_1)$. When XORing them all together over $i \in [k]$, we get $b_0 \oplus c \cdot (b_0 \oplus b_1)$ which is exactly what we needed.

For receiver-security, if $\mathcal{W}$ has receiver-security of p, we can use the XOR Lemma to deduce that S-Reduce$(k, \mathcal{W})$ has receiver-security of $p^k + \varepsilon$, because the shares $\{c_i\}_{i=1}^k$ are random and independent (over a random choice of c) and for every fixing of the sender's randomness, the i'th transcript is independent of the rest and is at most p-correlated to c_i. We can also use our own product bound to deduce that R-Reduce$(k, \mathcal{W})$ has receiver-security of $1 - (1 - p)^k + \varepsilon$, because for every fixing of the sender's randomness, the transcripts are independent conditioned on c and each one is at most p-correlated to c. Using similar arguments, it can be shown that symmetrically, sender-security amplifies to $q^k + \varepsilon$ in R-Reduce$(k, \mathcal{W})$ and weakens to $1 - (1 - q)^k + \varepsilon$ in S-Reduce$(k, \mathcal{W})$. These are exactly the same bounds used in the information-theoretic OT amplification analysis, up to the additive ε paid for each use.

The goal is to use these two operations repeatedly one after the other in order to reduce both parameters. It is already shown in [2, Lemma 4] exactly how this is done, but for the sake of completeness let us summarize the process as follows: Assume without loss of generality that $p \leq q$ (other case is symmetric). If $p \geq 0.2$, by applying R-Reduce$(2, \cdot)$ followed by S-Reduce$(2, \cdot)$, the distance between the error sum $p + q$ and 1 is multiplied by at least 1.1. Otherwise, if $q > 0.4$, by applying R-Reduce$(2, \cdot)$ the distance between $p + q$ and 1 is multiplied by at least 1.2. Otherwise, if $p + q > 0.2$, we again apply R-Reduce$(2, \cdot)$ followed by S-Reduce$(2, \cdot)$, with the guarantee that the error sum $p+q$ multiplies by a factor of at most 0.8. Finally, in the case where $p + q \leq 0.2$, we apply R-Reduce$(4, \cdot)$ followed by S-Reduce$(4, \cdot)$, and the guarantee is that the error sum is at least squared, that is, $(p' + q') \leq (p + q)^2$, so the progress downwards is quick.

To conclude, the same analysis from the information-theoretic setting holds here, up to an additive ε accumulated at each use. Let $p(n)$ be a bound on the total number of calls to the original protocol in the information-theoretic transformation, then all advantages throughout the process are $1/p(n)$-bounded away from 1, otherwise we would not be able to reduce them to negligible. By setting $\varepsilon' = \varepsilon/p(n)$, for every advantage d through the process we have $d + \varepsilon' \leq \varepsilon + (1 - \varepsilon)d$, so we can imagine, for the sake of the analysis, as if every call to either S-Reduce or R-Reduce incurs a chance of ε at failing and revealing everything, and otherwise works exactly like the information-theoretic world. Since the number of calls is polynomial, the total probability of failing is at most poly$(n) \cdot \varepsilon$ and we can make it as (polynomially) small as we want.

There is one small issue, however - the running time. In the information-theoretic process we make $\log\log(n)$ calls to S-Reduce and R-Reduce (when p and q are constants), and each such call, when using Yao's XOR Lemma or the bounds in this paper, decreases the bound on the running time by a division in a polynomial. Therefore, we need the assumption that our weak OT is secure against $n^{O(\log\log n)}$ adversaries. We remark that this issue can be overcome by choosing an increasing series of errors instead of fixing ε throughout the process. If $1-(p+q)$ is not lower bounded by a constant but by $1/\text{poly}(n)$, then we need security against $n^{O(\log n)}$ adversaries.

6 Open Questions

An issue that keeps appearing in security reductions where amplification is involved is the problem of amplification beyond negligible [3]. For example see [5, Lemma 3] and the discussion following it. Roughly speaking, in these types of reductions we can show security holds except for negligible probability but nothing concrete beyond that without increasing the running time of the reduction to be super-polynomial.

For a more specific example, let us consider Levin's proof of the XOR Lemma [10]. Informally, it is shown that if for X_0, X_1 we have that b is at most d-

correlated to X_b by s-sized circuits, then $\bigoplus_{i=1}^{t} b_i$ is at most $d^t + \varepsilon$-correlated to $X_{b_1}, \ldots, X_{b_t}$ by $s \cdot \mathrm{poly}(\varepsilon)$-sized circuits. Note the trade-off between the reduction accuracy and the circuit size bound. Another trade-off can also be seen in Theorem 2. If we only know that s is greater than any polynomial then we can push ε up to negligible but nothing concrete beyond that, otherwise the circuit size bound becomes meaningless.

As noted in [5], Rudich has observed that we cannot expect to overcome this issue in a black-box way. Further, in [3] an example is given, based on non-standard assumptions, of a weak OWF that cannot be amplified beyond negligible using the direct product transformation. That is, it may be that overcoming this issue is not just hard to prove, but can be altogether false. Nonetheless, what happens in general is still unclear, and there are still open directions of either strengthening the impossibilities by reducing assumptions, or of showing that some form of amplification beyond negligible is achievable.

Interestingly enough, when considering Corollary 2, we note how the circuit size growth is actually only logarithmic in $1/\varepsilon$, although linear in $1/(1-d)^m$. Still, if d and m are constants, then we can reduce the error exponentially well while maintaining efficiency of the circuits. This brings us to the following conjecture, aiming to formulate the XOR equivalent of the above, stated with specific parameters for simplicity:

Conjecture 1 (Informal). If b is at most 0.5-correlated to X_b by s-sized circuits, then $b_1 \oplus b_2$ is at most $0.25 + 2^{-n}$-correlated to X_{b_1}, X_{b_2} by $s/\mathrm{poly}(n)$-sized circuits.

In other words, as long as we are not trying to achieve correlation beyond negligible, we can get exponentially close efficiently. This could be seen as a first step towards a positive result.

References

1. Barak, B., Hardt, M., Kale, S.: The uniform hardcore lemma via approximate bregman projections. In: Mathieu, C. (ed.) Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2009, New York, NY, USA, January 4-6, 2009. pp. 1193–1200. SIAM (2009), http://dl.acm.org/citation.cfm?id=1496770.1496899
2. Damgård, I., Kilian, J., Salvail, L.: On the (im)possibility of basing oblivious transfer and bit commitment on weakened security assumptions. In: Stern, J. (ed.) Advances in Cryptology - EUROCRYPT '99, International Conference on the Theory and Application of Cryptographic Techniques, Prague, Czech Republic, May 2-6, 1999, Proceeding. Lecture Notes in Computer Science, vol. 1592, pp. 56–73. Springer (1999). https://doi.org/10.1007/3-540-48910-X_5, https://doi.org/10.1007/3-540-48910-X_5
3. Dodis, Y., Jain, A., Moran, T., Wichs, D.: Counterexamples to hardness amplification beyond negligible. In: Cramer, R. (ed.) Theory of Cryptography - 9th Theory of Cryptography Conference, TCC 2012, Taormina, Sicily, Italy, March 19-21, 2012. Proceedings. Lecture Notes in Computer Science, vol. 7194,

pp. 476–493. Springer (2012). https://doi.org/10.1007/978-3-642-28914-9_27, https://doi.org/10.1007/978-3-642-28914-9_27

4. Fehr, S., Vaudenay, S.: Sublinear bounds on the distinguishing advantage for multiple samples. In: Aoki, K., Kanaoka, A. (eds.) Advances in Information and Computer Security - 15th International Workshop on Security, IWSEC 2020, Fukui, Japan, September 2-4, 2020, Proceedings. Lecture Notes in Computer Science, vol. 12231, pp. 165–183. Springer (2020). https://doi.org/10.1007/978-3-030-58208-1_10, https://doi.org/10.1007/978-3-030-58208-1_10
5. Goldreich, O., Nisan, N., Wigderson, A.: On yao's xor-lemma. Electron. Colloquium Comput. Complex. **2**(50) (1995), http://eccc.hpi-web.de/eccc-reports/1995/TR95-050/index.html
6. Holenstein, T.: Key agreement from weak bit agreement. In: Gabow, H.N., Fagin, R. (eds.) Proceedings of the 37th Annual ACM Symposium on Theory of Computing, Baltimore, MD, USA, May 22-24, 2005. pp. 664–673. ACM (2005). https://doi.org/10.1145/1060590.1060689, https://doi.org/10.1145/1060590.1060689
7. Hollander, F.: Probability theory : The coupling method (2012)
8. Impagliazzo, R.: Hard-core distributions for somewhat hard problems. In: 36th Annual Symposium on Foundations of Computer Science, Milwaukee, Wisconsin, USA, 23-25 October 1995. pp. 538–545. IEEE Computer Society (1995). https://doi.org/10.1109/SFCS.1995.492584, https://doi.org/10.1109/SFCS.1995.492584
9. Kontorovich, A.: Obtaining measure concentration from markov contraction. Markov Processes and Related Fields **18**(4), 613–638 (2012)
10. Levin, L.A.: One-way functions and pseudorandom generators. Comb. **7**(4), 357–363 (1987). https://doi.org/10.1007/BF02579323, https://doi.org/10.1007/BF02579323
11. Maurer, U.M., Tessaro, S.: A hardcore lemma for computational indistinguishability: Security amplification for arbitrarily weak prgs with optimal stretch. In: Micciancio, D. (ed.) Theory of Cryptography, 7th Theory of Cryptography Conference, TCC 2010, Zurich, Switzerland, February 9-11, 2010. Proceedings. Lecture Notes in Computer Science, vol. 5978, pp. 237–254. Springer (2010). https://doi.org/10.1007/978-3-642-11799-2_15, https://doi.org/10.1007/978-3-642-11799-2_15
12. Renner, R.: On the variational distance of independently repeated experiments. CoRR **abs/cs/0509013** (2005), http://arxiv.org/abs/cs/0509013
13. Wullschleger, J.: Oblivious-transfer amplification. In: Naor, M. (ed.) Advances in Cryptology - EUROCRYPT 2007, 26th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Barcelona, Spain, May 20-24, 2007, Proceedings. Lecture Notes in Computer Science, vol. 4515, pp. 555–572. Springer (2007). https://doi.org/10.1007/978-3-540-72540-4_32, https://doi.org/10.1007/978-3-540-72540-4_32