Is it possible to improve Yao's XOR lemma using reductions that exploit the efficiency of their oracle?

Ronen Shaltiel*

December 22, 2022

Abstract

Yao's XOR lemma states that for every function $f : \{0,1\}^k \to \{0,1\}$, if f has hardness $2/3$ for $P/poly$ (meaning that for every circuit C in $P/poly$, $\Pr[C(X) = f(X)] \leq 2/3$ on a uniform input X), then the task of computing $f(X_1) \oplus \ldots \oplus f(X_t)$ for sufficiently large t, has hardness $\frac{1}{2} + \epsilon$ for $P/poly$.

Known proofs of this lemma cannot achieve $\epsilon = \frac{1}{k^{\omega(1)}}$, and even for $\epsilon = \frac{1}{k}$, we do not know how to replace $P/poly$ by $\mathrm{AC}^0[\mathrm{PARITY}]$ (the class of constant depth circuits with the gates $\{\mathrm{AND}, \mathrm{OR}, \mathrm{NOT}, \mathrm{PARITY}\}$ of unbounded fan-in).

Grinberg, Shaltiel and Viola (FOCS 2018) (building on a sequence of earlier works) showed that these limitations cannot be circumvented by *black-box reductions*. Namely, by reductions $\mathrm{Red}^{(\cdot)}$ that given oracle access to a function D that violates the conclusion of Yao's XOR lemma, implement a circuit that violates the assumption of Yao's XOR lemma.

There are a few known reductions in the related literature on worst-case to average case reductions that are *non-black box*. Specifically, the reductions of Gutfreund, Shaltiel and Ta-Shma (Computational Complexity 2007) and Hirahara (FOCS 2018)) are "class reductions" that are only guaranteed to succeed when given oracle access to an oracle D from some efficient class of algorithms. These works seem to circumvent some black-box impossibility results.

In this paper we extend the previous limitations of Grinberg, Shaltiel and Viola to several types of class reductions, giving evidence that class reductions cannot yield the desired improvements in Yao's XOR lemma. To the best of our knowledge, this is the first limitation on reductions for hardness amplification that applies to class reductions.

Our technique imitates the previous lower bounds for black-box reductions, replacing the inefficient oracle used in that proof, with an efficient one that is based on limited independence, and developing tools to deal with the technical difficulties that arise following this replacement.

*Department of computer science, University of Haifa. E-mail: ronen@cs.haifa.ac.il. This research was supported by ISF grant 1628/17.

1 Introduction

Yao's XOR Lemma is a fundamental and celebrated result in complexity theory, that is extensively studied (from various aspects) and has found many applications. See [GNW11] for a survey article.

Definition 1.1 (The XOR function)**.** *Given a function $f : \{0,1\}^k \to \{0,1\}$, and a number t, we define $f^{\oplus t} : \{0,1\}^{t\cdot k} \to \{0,1\}$, as follows: Given $y \in \{0,1\}^{t\cdot k}$, we view y as $(y_1,\ldots,y_t) \in (\{0,1\}^k)^t$, and define:*

$$f^{\oplus t}(y) = f(y_1) \oplus \ldots \oplus f(y_t)$$

Let U_k denote the uniform distribution on k bit strings. Loosely speaking, Yao's XOR lemma says that if a function f is "mildly hard on average" on input $X \leftarrow U_k$, then as t increases, computing $f^{\oplus t}$ on input $Y \leftarrow U_{tk}$, becomes "very hard on average".

Lemma 1.2 (Yao's XOR lemma, for poly-size circuits)**.** *For every $f : \{0,1\}^k \to \{0,1\}$, and $t \leq \mathrm{poly}(k)$ such that $t = \omega(\log k)$:*

If, for every $\mathrm{poly}(k)$ size circuit C, $\Pr_{X\leftarrow U_k}[C(X) = f(X)] < \frac{2}{3}$,

Then, for every constant c, and every $\mathrm{poly}(k)$ size circuit D, $\Pr_{Y\leftarrow U_{tk}}[D(Y) = f^{\oplus t}(Y)] < \frac{1}{2} + \frac{1}{k^c}$.[1]

One weakness of Yao's XOR lemma, is that it cannot be used to conclude a statement in which the "hardness on average" $\frac{1}{2} + \frac{1}{k^c}$ is replaced by $\frac{1}{2} + \frac{1}{k^{\omega(1)}}$. This holds, even if the number of repetitions t is increased from slightly larger than $\log k$ (as is the case in Lemma 1.2) to the maximal choice of $t = \mathrm{poly}(k)$. Specifically, the following question is wide open:

Open problem 1.3 (Yao's XOR lemma for subpolynomial error?)**.** *Is it true that for every $f : \{0,1\}^k \to \{0,1\}$, taking $t = \omega(\log k)$ (or even the maximal choice of $t = \mathrm{poly}(k)$) it holds that:*

If, for every $\mathrm{poly}(k)$ size circuit C, $\Pr_{X\leftarrow U_k}[C(X) = f(X)] < \frac{2}{3}$,

Then, for every $\mathrm{poly}(k)$ size circuit D, $\Pr_{Y\leftarrow U_{tk}}[D(Y) = f^{\oplus t}(Y)] < \frac{1}{2} + \frac{1}{k^{\omega(1)}}$.

Another weakness of Yao's XOR Lemma is that known proofs fail to prove Yao's XOR lemma when replacing $P/poly$ with many interesting constant depth circuit classes. An especially frustrating case is the class $\mathrm{AC}^0[\mathrm{PARITY}]$ of poly-size constant depth circuits over the gates $\{\mathrm{AND},\mathrm{OR},\mathrm{NOT},\mathrm{PARITY}\}$ of unbounded fan-in. There are known lower bounds showing explicit functions that have hardness $\frac{2}{3}$ for $\mathrm{AC}^0[\mathrm{PARITY}]$ (or even $\frac{1}{2} + o(1)$ hardness) for circuits of depth d and size $2^{k^{\Omega(1/d)}}$ [Raz87, Smo87]. However, we do not have lower bounds for $\mathrm{AC}^0[\mathrm{PARITY}]$ that achieve hardness $\frac{1}{2} + \frac{1}{k}$ for a function on k bits, and the best bound by Razborov [Raz87] achieves hardness $\frac{1}{2} + \frac{1}{\sqrt{k}}$.

This is a twenty five year old barrier that prevents us from "using the hybrid argument" when constructing pseudorandom generators for $\mathrm{AC}^0[\mathrm{PARITY}]$ (and related classes). This barrier limits the best known pseudorandom generators for $\mathrm{AC}^0[\mathrm{PARITY}]$ by [FSUV13] to very poor seeds. (See [FSUV13] for a discussion of this limitation). Specifically, the following question is wide open:

Open problem 1.4 (Yao's XOR lemma for constant depth circuits?)**.** *Let G be the set of gates $\{\mathrm{AND},\mathrm{OR},\mathrm{NOT},\mathrm{PARITY}\}$ of unbounded fan-in. Is it true that for every $f : \{0,1\}^k \to \{0,1\}$, taking $t = \omega(\log k)$ (or even the maximal choice of $t = \mathrm{poly}(k)$) it holds that:*

[1] Naturally, in order to make this asymptotic statement precise, one needs to consider an infinite sequence of functions $\{f_k\}$ with growing input length (so that terms like "poly-size", "$\omega(\log k)$", and "constant" are well defined). We allow ourselves to be imprecise, as a more general, and quantitatively precise statement of Yao's XOR lemma is given below in Lemma 1.5.

If, for every $\mathrm{poly}(k)$ *size, constant-depth circuit* C *with gates in* G, $\Pr_{X \leftarrow U_k}[C(X) = f(X)] < \frac{2}{3}$,

Then, for every $\mathrm{poly}(k)$ *size, constant-depth circuit* D, *with gates in* G, $\Pr_{Y \leftarrow U_{tk}}[D(Y) = f^{\oplus t}(Y)] < \frac{1}{2} + \frac{1}{tk}$.

1.1 Proofs of Yao's Lemma as (nonuniform) black-box reductions

Before discussing the best known proofs of Yao's XOR Lemma, let us state the lemma more precisely, in a more general and quantitative form. The next formulation is achieved using Impagliazzo's proof of Yao's XOR lemma [Imp95, GNW11] together with the quantitative improvement of Klivans and Servedio [KS03] of Impagliazzo's hard-core lemma [Imp95].

Lemma 1.5 (Yao's XOR lemma, general version)**.** *There exist a constant* c, *and a polynomial* p, *such that for every* $f : \{0,1\}^k \to \{0,1\}$, *every* $\epsilon > 0$, *every* $0 < \delta < \frac{1}{2}$, *and every* $t \geq c \cdot \frac{\log(1/\epsilon) \cdot \log\log\frac{1}{\epsilon}}{\delta}$, *setting* $q = c \cdot \frac{\log(1/\delta)}{\epsilon^2}$, *we have that:*

If, for every circuit C *of size* $s \geq p(t,k,q)$, $\Pr_{X \leftarrow U_k}[C(X) = f(X)] < 1 - \delta$,

Then, for every circuit D *of size* $s' = \frac{s}{q}$, $\Pr_{Y \leftarrow U_{tk}}[D(Y) = f^{\oplus t}(Y)] < \frac{1}{2} + \epsilon$.

Remark 1.6 (Other tradeoffs in the XOR lemma)**.** *Lemma 1.5 is not stated explicitly in [KS03]. However, it immediately follows by "plugging in" the hard-core lemma of Klivans and Servedio [KS03] in Impagliazzo's proof of the XOR lemma [Imp95]. To the best of our knowledge, this version is the best known in the parameter* q *(which is a focus of this paper). However, it uses* t *that is larger by a* $\log \cdot \log(1/\epsilon)$ *multiplicative factor, compared to other proofs. We remark that other proofs (see [GNW11] for a survey) can achieve smaller* t, *namely,* $t = c \cdot \frac{\log(1/\epsilon)}{\delta}$ *for some constant* c, *at the cost of using larger values for* q. *One possible tradeoff is obtained by plugging the "min-max proof" of the hard-core lemma (attributed to Nisan in [Imp95]) in Impagliazzo's proof [Imp95] which gives* $t = c \cdot \frac{\log(1/\epsilon)}{\delta}$ *for* $q = c \cdot \frac{\log(1/(\epsilon\delta)}{\epsilon^2}$.

The special case of Lemma 1.2 is obtained by taking s to be a polynomial in k, and $\delta = \frac{1}{3}$. In order to reduce the number of live parameters, we recommend that the reader focuses on these choices on a first reading. We point out that s' (which is the size of D) is smaller by a factor of $q = \Omega(\frac{1}{\epsilon^2})$, than s (which is the size of C). This implies that $s' \leq O(\epsilon^2 \cdot s)$, implying that $\epsilon \geq \Omega(\frac{1}{\sqrt{s}})$, and it is impossible to get $\epsilon < \frac{1}{s}$ with current proofs. (This is a more quantitative way to state the phenomenon in open problem 1.3).

All known proofs of Yao's XOR lemma work by *reduction*. That is, the proof shows a reduction that transforms a circuit D such that

$$\Pr_{Y \leftarrow U_{tk}}[D(Y) = f^{\oplus t}(Y)] \geq \frac{1}{2} + \epsilon,$$

into a circuit C such that

$$\Pr_{X \leftarrow U_k}[C(X) = f(X)] \geq 1 - \delta.$$

We will use the concept of an "oracle circuit" to give a formal definition of a reduction.

Definition 1.7 (Oracle circuit)**.** *An oracle circuit* $R^{(\cdot)}$ *is a circuit that in addition to standard gates also has additional "oracle gates" (possibly of large fan-in* m*). For a function* $D : \{0,1\}^m \to \{0,1\}$, *the function* C^D *is defined by instantiating the oracle gates with* D.

All known proofs of Yao's XOR lemma are "nonuniform black-box reductions", meaning that they provide a reduction (namely an oracle circuit $\mathrm{Red}^{(\cdot)}(x,\alpha)$ where x is an input, and α is an "advice string") and the circuit C is obtained by $C(x) = \mathrm{Red}^{D}(x,\alpha)$ where α is a "nonuniform advice string" that may depend on f and D.[2] This is made precise in the next definition.

Definition 1.8 (Nonuniform black-box reduction for Yao's XOR lemma)**.** *Let $\epsilon,\delta > 0$, and let k,t,a be integers. A $(\frac{1}{2}+\epsilon) \to (1-\delta)$* black-box reduction for Yao's XOR lemma *(with input length k, t repetitions and advice length a) is an oracle circuit $\mathrm{Red}^{(\cdot)}(x,\alpha)$, where $x \in \{0,1\}^k$ and $\alpha \in \{0,1\}^a$, such that for every $f : \{0,1\}^k \to \{0,1\}$, the following holds:*

For every function $D : \{0,1\}^{tk} \to \{0,1\}$, such that $\Pr_{Y \leftarrow U_{tk}}[D(Y) = f^{\oplus t}(Y)] \geq \frac{1}{2}+\epsilon$,

there exists $\alpha \in \{0,1\}^a$, such that $\Pr_{X \leftarrow U_k}[\mathrm{Red}^{D}(X,\alpha) = f(X)] \geq 1-\delta$.

The version of Yao's XOR lemma stated in Lemma 1.5, follows by showing the following reduction:

Lemma 1.9 (Known black-box reductions for Yao's XOR lemma)**.** *There exist a constant c, and a polynomial p, such that for every integer k, every $\epsilon,\delta > 0$ such that $1-\delta > \frac{1}{2}+\epsilon$, and every $t \geq c \cdot \frac{\log(1/\epsilon)\cdot\log\log\frac{1}{\epsilon}}{\delta}$, there is a $(\frac{1}{2}+\epsilon) \to (1-\delta)$ black-box reduction $\mathrm{Red}^{(\cdot)}(x,\alpha)$ for Yao's XOR lemma with input length k, t repetitions, and advice length a such that:*

- Red *makes at most $q = c \cdot \frac{\log(1/\delta)}{\epsilon^2}$ queries to its oracle (meaning that on every input, and for every function D, the number of "oracle calls" made to D is at most q).*
- Red *is an oracle circuit of size $r = p(t,k,q)$, (and in particular, $a \leq r$).*
- Red *is an oracle circuit of constant depth d over the gates $\{$AND,OR,NOT$\}$ of unbounded fan-in and also uses a single majority gate over q bits.*

In order to understand the limitations that prevent known proofs from solving the aforementioned open problems, it is instructive to see how Lemma 1.5 follows from Lemma 1.9. Specifically, assume (for contradiction) that Lemma 1.5 does not hold and let D be a circuit of size s' that is violating the conclusion. By Lemma 1.9, there exists $\alpha \in \{0,1\}^a$, such that the circuit $C(x) = \mathrm{Red}^{D}(x,\alpha)$ computes $f(X)$ with success $1-\delta$ on $X \leftarrow U_k$. The size of C is bounded by $s = r + a + q \cdot s' \geq q \cdot s'$, and the obtained circuit C has depth at least d, and needs to compute majority on q bits. Summing up:

- The number of queries q made by the reduction is a lower bound on $\frac{s}{s'}$, meaning that $s' \leq \frac{s}{q}$ and as the known reductions have $q \geq \frac{1}{\epsilon^2}$ we cannot expect $\epsilon < \frac{1}{\sqrt{s}}$, and cannot solve open problem 1.3.
- The fact that the best known reductions require a majority gate on $q \geq \frac{1}{\epsilon}$ inputs, means that we need to assume hardness against a class that can perform this computation. For $\epsilon = 1/k$, Razborov's lower bound [Raz87] (see also [OSS19]) shows that for every depth d', majority on k bits, cannot be computed by circuits of depth d' and size $2^{k^{\Omega(1/d')}}$ over the gates $\{$AND,OR,NOT,PARITY$\}$, explaining why current reductions cannot solve open problem 1.4.

[2] There is a formal connection between "black box hardness amplification" and list-decodable error correcting code [STV01], see for example the discussion in [SV10, GSV18]. Using this connection, it is known that black-box reductions for Yao's XOR lemma, must be nonuniform and use an advice string if $1-\delta > \frac{1}{2}+\epsilon$ and $\epsilon < \frac{1}{4}$.

Limitations on black-box reductions. A sequence of works [Vio06, SV10, GR08, AS14, AASY16] culminating in [GSV18], shows that known black-box reductions for Yao's XOR lemma must suffer from the limitations above: They require $q = \Omega(\frac{\log(1/\delta)}{\epsilon^2})$ queries, and require computing majority on input length $\Omega(\frac{1}{\epsilon})$.[3]

1.2 Class reductions

On a closer examination, black-box reductions seem to be an overkill for the task of proving Yao's XOR lemma. For proving Yao's XOR lemma, we don't need $\mathrm{Red}^{(\cdot)}$ to succeed given oracle access to *every* function D such that $\Pr_{Y \leftarrow U_{tk}}[D(Y) = f^{\oplus t}(Y)] \geq \frac{1}{2} + \epsilon$. It is sufficient that Red succeeds only given oracle access to functions D that are *efficiently computable* and belong to the class $\mathcal{D}$ of circuits with size s' (if we're in the setup of open problem 1.3) and size s' with constant depth (if we're in the setup of open problem 1.4).

This motivates a notion of *class reduction* (suggested for example in [GT07]) in which reductions are only required to succeed if given oracle access to a function D that belongs to some class $\mathcal{D}$ of "efficient circuits", and do not need to succeed when given oracle access to a function D that does not belong to $\mathcal{D}$. The definition of *class $\mathcal{D}$ reduction* below is identical to definition 1.8 with the single exception (that is underlined for emphasis) being that we only require the reduction to succeed when given oracle access to a function D from the class $\mathcal{D}$.

Definition 1.10 (Nonuniform class reduction for Yao's XOR lemma)**.** *Let $\epsilon, \delta > 0$ and let k, t, a be integers, and let $\mathcal{D}$ be some class of functions $D : \{0,1\}^{tk} \to \{0,1\}$. A $(\frac{1}{2} + \epsilon) \to (1 - \delta)$ class $\mathcal{D}$ reduction for Yao's XOR lemma (with input length k, t repetitions and advice length a) is an oracle circuit $\mathrm{Red}^{(\cdot)}(x, \alpha)$, where $x \in \{0,1\}^k$ and $\alpha \in \{0,1\}^a$, such that for every $f : \{0,1\}^k \to \{0,1\}$, the following holds:*

For every function $D : \{0,1\}^{tk} \to \{0,1\}$ in the class $\mathcal{D}$, such that $\Pr_{Y \leftarrow U_{tk}}[D(Y) = f^{\oplus t}(Y)] \geq \frac{1}{2} + \epsilon$, there exists $\alpha \in \{0,1\}^a$, such that $\Pr_{X \leftarrow U_k}[\mathrm{Red}^D(X, \alpha) = f(X)] \geq 1 - \delta$.

Note that a black-box reduction is a special case of a class reduction where $\mathcal{D}$ is the class of all boolean functions on tk bits. This raises the following questions:

1. Are there reductions in the literature that are class reductions but not black-box reductions?
2. Can class reductions circumvent the limitations on black-box reductions and solve open problem 1.3 or open problem 1.4?

The answer to the first question is affirmative in the sense that there are at least two examples that we are aware of, where a worst-case to average case amplification is proven by a reduction that is not black-box. Furthermore, in both cases, the reduction is a class reduction, and there is strong evidence that it cannot be made black-box.

The first example is a worst-case hardness to average case hardness tradeoff for SAT (with respect to a distribution sampled in quasipolynomial time) by Gutfreund, Shaltiel and Ta-Shma [GST07] (see also a related work [Ats06, Gut06, GT07]). The correctness of the reduction of [GST07] relies on the efficiency of the oracle and the term "class reduction" was suggested by Gutfreund and Ta-Shma [GT07]. It was also

[3] More formally, saying that $\mathrm{Red}^{(\cdot)}$ "requires computing majority on input length $\Omega(1/\epsilon)$" means that every such reduction $\mathrm{Red}^{(\cdot)}$ can be transformed into a circuit (with no oracle) of roughly the same size and depth as $\mathrm{Red}^{(\cdot)}$ for computing the majority function on inputs of length $\Omega(\frac{1}{\epsilon})$.

argued in [GT07] that limitations on black-box reductions proven by Bogdanov and Trevisan [BT06] can be extended to the scenario studied in [GST07], and show that if the class reduction of [GST07] could be made black-box, then co-NP has nondeterministic circuits of quasipolynomial size.

Another example is Hirahara's recent worst-case to average case reductions for variants of MCSP and MINKT [Hir18]. These reductions are non-black-box, in the sense that their correctness relies on the efficiency of their oracle. The aforementioned work of Bogdanov and Trevisan [BT06] shows that if these reductions can be made black-box, then these problems are in co-NP/poly, which is not known, and is false, if these problems are NP-complete. See [Hir18] for an elaborate discussion of consequences of the existence of such black-box reductions.

1.3 Limitations on class reductions for Yao's XOR lemma

In this paper we give evidence that the answer to the second question above is negative. We extend the aforementioned limitations of [GSV18] on black-box reductions for Yao's XOR lemma to *class reductions* for any $\mathcal{D}$ of that contains circuits that have polynomial size and constant depth over the gates $\{\text{AND,OR,NOT,PARITY}\}$ with unbounded fan-in. There are however some caveats that are discussed in Section 1.3.1.

To the best of our knowledge, this is the first example of proving limitations on class reductions in this setup. Our results are stated formally in the next theorem.[4]

Theorem 1.11 (Limitations on class reductions for Yao's XOR lemma)**.** *There exist constants $\delta_0 > 0$, $\nu > 0$, $d_0 > 1$ and a polynomial p such that: Let $\mathrm{Red}^{(\cdot)}(x, \alpha)$ be a $(\frac{1}{2} + \epsilon) \to (1 - \delta)$ class $\mathcal{D}$ reduction for Yao's XOR lemma, with input length k, t repetitions and advice length a. Assume that:*

- $\mathrm{Red}^{(\cdot)}$ *is a size r oracle circuit, that makes at most q queries.*
- *The class $\mathcal{D}$ contains circuits of size $p(r)$ and depth d_0 over the gates $\{\text{AND,OR,NOT,PARITY}\}$ of unbounded fan-in.*
- $k, t, a, \frac{1}{\epsilon}, \frac{1}{\delta} \leq r \leq 2^{\nu \cdot k}$ *and* $\delta \leq \delta_0$.

Then the following holds:

- $\mathrm{Red}^{(\cdot)}$ *requires many queries, specifically:* $q = \Omega(\frac{\log(1/\delta)}{\epsilon^2})$.
- $\mathrm{Red}^{(\cdot)}$ *requires majority, specifically: if in addition to the size restriction on* Red*, we also have that* $\mathrm{Red}^{(\cdot)}$ *is an oracle circuit of depth d over a set of gates G that contains the gates $\{\text{AND,OR,NOT,PARITY}\}$ of unbounded fan-in, then the majority function over $\Omega(\frac{1}{\epsilon})$ bits can be computed by a circuit of size $\mathrm{poly}(r)$ and depth $O(d)$ over the set of gates G.*

The case that $\delta = \frac{1}{3}$. Theorem 1.11 is stated for $\delta \leq \delta_0$ for some unspecified constant $\delta_0 > 0$. In Open problems 1.3,1.4 we took $\delta = \frac{1}{3}$, which does not seem to be covered by Theorem 1.11. Nevertheless, we will now observe that Theorem 1.11 has implications also for $\delta = \frac{1}{3}$.

By known proofs of the XOR lemma, and in particular by Lemma 1.9, there exists a $\frac{2}{3} \to (1 - \delta_0)$ black-box reduction for Yao's XOR lemma with constant t, q and a. Moreover, this black-box reduction

[4] We remark that any circuit of size r over the gates $\{\text{AND,OR,NOT,PARITY}\}$ with unbounded fan-in, cannot use fan-in larger than r, and therefore can be simulated by a circuit of size $O(r^2)$ over the standard gates $\{\text{AND,OR,NOT}\}$ with bounded fan-in. This allows us to state our results in a way that captures both circuits of small depth (using gates with unbounded fan-in) and circuits that use the standard gates with bounded fan-in.

is an oracle circuit of size $\mathrm{poly}(k)$ and constant depth over the gates $\{\text{AND,OR,NOT}\}$ of unbounded fan-in. (We stress that a majority gate is not needed, as the majority gate is over $q = O(1)$ bits).

This implies that a $(\frac{1}{2}+\epsilon) \to \frac{2}{3}$ class $\mathcal{D}$ reduction for Yao's XOR lemma implies a $(\frac{1}{2}+\epsilon) \to (1-\delta_0)$ class $\mathcal{D}'$ reduction for Yao's XOR lemma, with constant losses in t, a, q, and where the class $\mathcal{D}'$ is slightly weaker than $\mathcal{D}$ in terms of the relevant measures of circuit size and depth. Altogether, this implies that Theorem 1.11 also applies for $\delta = \frac{1}{3}$.

1.3.1 What kind of reductions are ruled out by this result?

Theorem 1.11 achieves exactly the same limitations on *class reductions* for Yao's XOR lemma as the limitations of [GSV18] for *black-box reductions*. This is achieved for any class $\mathcal{D}$ that contains small circuits with constant depth over the gates $\{\text{AND,OR,NOT,PARITY}\}$ of unbounded fan-in, which is exactly the classes that come up if one wants to use class reductions to solve open problems 1.3 and 1.4. (As we explain in detail below in Section 1.4, the results of [GSV18] are more general and apply to *every form* of hardness amplification, rather than just XOR lemmas).

In order to explain what kind of reductions are ruled out, let us consider three kinds of potential reductions:

1. Nonuniform black-box reductions (as defined in Definition 1.8. (This is the type of reductions covered in [GSV18]).
2. Class $\mathcal{D}$ reductions where the reductions $\mathrm{Red}^{(\cdot)}$ is "less powerful" than the class $\mathcal{D}$. (This is the type of reductions covered in Theorem 1.11).
3. Class $\mathcal{D}$ reductions where the reduction $\mathrm{Red}^{(\cdot)}$ is "more powerful" than the class $\mathcal{D}$. (This type of reductions are not covered in Theorem 1.11, as we explian below).

The distinction between the second and third items can be seen in Theorem 1.11, and specifically, in the requirement that $\mathcal{D}$ contains circuits of size larger than the size of the reduction $\mathrm{Red}^{(\cdot)}$. This is crucially used in the proof of the theorem.[5]

In both kinds of class reductions (both the second and third items) one obtains an XOR lemma where the hardness assumption is against circuits of the form Red^D for every $D \in \mathcal{D}$, and the conclusion is against $\mathcal{D}$. Our results apply for the second kind of reductions, and in particular imply that:

- If $\mathcal{D}$ is the class of circuits of size $s(k)$, and $\mathrm{Red}^{(\cdot)}$ is a $\mathcal{D}$ class reduction of size $r(k)$, and if furthermore $r(k) \leq s(k)/p(k)$ (for a universal polynomial $p(k)$), then by Theorem 1.11, $\mathrm{Red}^{(\cdot)}$ has to make at least $q = \Omega(\frac{\log(1/\delta)}{\epsilon^2})$ queries. In particular, to obtain an XOR-lemma with hardness $\frac{1}{2}+\epsilon(k)$ against circuits of size $s(k)$ it is necessary to assume hardness against circuits of size

$$s'(k) \geq r(k) + q(k) \cdot s(k) \geq \frac{s(k)}{\epsilon(k)^2}.$$

  This implies that when assuming hardness against circuits of size $s'(k)$ one can not get $\epsilon(k) \leq \frac{1}{\sqrt{s'(k)}}$.

- If in addition to the requirements in the previous item, $\mathrm{Red}^{(\cdot)}$ is a depth d circuit over the gates $\{\text{AND,OR,NOT,PARITY}\}$ of unbounded fan-in, then majority on inputs of length $\Omega(\frac{1}{\epsilon(k)})$ can be computed by circuits of depth $O(d)$ and size $\mathrm{poly}(r(k))$ the gates $\{\text{AND,OR,NOT,PARITY}\}$ over the gates

[5] We remark that curiously, we do not need to require that $\mathcal{D}$ contains circuits of depth larger than the depth of $\mathrm{Red}^{(\cdot)}$

$\{$AND,OR,NOT,PARITY$\}$ of unbounded fan-in. As $s'(k) \geq r(k)$ this means that we need to assume hardness against circuits of size

$$s'(k) = 2^{\Omega((\frac{1}{\epsilon(k)})^{\Omega(\frac{1}{d})})}.$$

In particular, if $\epsilon(k) \leq \frac{1}{k}$, it follows that we must assume hardness against circuits of size

$$s'(k) = 2^{\Omega(k)^{\Omega(\frac{1}{d})}}.$$

However, for both conclusions we need to require that the reduction is less powerful than the class $\mathcal{D}$, and specifically, that $r(k) \leq s(k)/p(k)$. We do not know how to remove this requirement, and leave this as an open problem, see Section 7.

In the context of Open problem 1.3, the fact that we do not rule out reductions that are more powerful than the class does not tule out an approach in which for every polynomial p_1, there exists a larger polynomial p_2 such that there is a $(\frac{1}{2} + \frac{1}{k^{\omega(1)}}) \to \frac{2}{3}$ class $\mathcal{D}_{p_1}$ reduction Red_{p_1} of size $p_2(k)$ for the class $\mathcal{D}_{p_1}$ of circuits of size $p_1(k)$ (but not for the class of circuits of size $p(p_2(k))$ where p is the polynomial in Theorem 1.11). This is sufficient for solving open problem 1.3 and is not ruled out by our impossibility results.

An optimistic view is that this may point us to the kind of reductions we need to design, in order to solve the aforementioned open problems. We remark however that the aforementioned reduction by Hirahara [Hir18] *does not* need to assume that the oracle is weaker than the reduction. (The reduction of [GST07] involves a more complicated scenario where there is also a third entity which is the samplable distribution, and so, it is arguable whether the reduction is more powerful than the oracle).

Perspective. Limitations for black-box reductions are extensively studied in various settings in complexity theory and cryptography. In order to prove impossibility results on black-box reduction, it is sufficient to show the existence of an oracle D (*that does not need to be efficient*) on which the reduction cannot succeed.

Many impossibility results and limitations in the literature strongly utilize the ability to choose an oracle D that is not efficient. One notable example is the aforementioned results of Bogdanov and Trevisan [BT06] (that build on earlier work of Feigenbaum and Fortnow [FF93]). Indeed, this is why these limitations do not apply to class reductions like the aforementioned results [GST07, Hir18].

This work puts an emphasis on whether or not the oracle D that one designs when showing a black-box impossibility result, can be made efficient, and demonstrates that achieving this, has the additional benefit of also ruling out class reductions.

1.4 Extensions to some other forms of hardness amplification

Theorem 1.11 is weaker than the results of [GSV18] in the sense that the limitations of [GSV18] apply not only to Yao's XOR lemma, but to *any* hardness amplification technique. More precisely, in the results of [GSV18] one can replace $f^{\oplus t}$ by any other function f' over $n = 2^{o(k)}$ bits, with the same limitations. Our approach cannot give such a general result. We can extend our result to some specific forms of f' that we now explain.

A construction map is an oracle circuit $\mathrm{Con}^{(\cdot)}$ where for every $f : \{0,1\}^k \to \{0,1\}$, the function $f' = \mathrm{Con}^f$ is a function $f' : \{0,1\}^n \to \{0,1\}$. Yao's XOR lemma is the special case where for some parameter t, $n = kt$ and $f' = \mathrm{Con}^f = f^{\oplus t}$. In the definition below, we extend the notion of class reductions to this more general setup. (Once again, the difference between this notion and nonuniform black-box reductions is that we only require the conclusion for D in the class $\mathcal{D}$).

Definition 1.12 (Nonuniform class reduction for hardness amplification)**.** *Let $\epsilon, \delta > 0$ and let k, n, a be integers. Let $\mathrm{Con}^{(\cdot)}$ be an oracle circuit such that for $f : \{0,1\}^k \to \{0,1\}$, $\mathrm{Con}^f : \{0,1\}^n \to \{0,1\}$. Let $\mathcal{D}$ be some class of functions $D : \{0,1\}^n \to \{0,1\}$. A $(\frac{1}{2} + \epsilon) \to (1 - \delta)$* class $\mathcal{D}$ reduction for Con *(with input length k, output length n and advice length a) is an oracle circuit $\mathrm{Red}^{(\cdot)}(x, \alpha)$, where $x \in \{0,1\}^k$ and $\alpha \in \{0,1\}^a$, such that for every $f : \{0,1\}^k \to \{0,1\}$, the following holds:*

For every function $D : \{0,1\}^n \to \{0,1\}$ in the class $\mathcal{D}$, such that $\Pr_{Y \leftarrow U_{tk}}[D(Y) = \mathrm{Con}^f(Y)] \geq \frac{1}{2} + \epsilon$, there exists $\alpha \in \{0,1\}^a$, such that $\Pr_{X \leftarrow U_k}[\mathrm{Red}^D(X, \alpha) = f(X)] \geq 1 - \delta$.

We can extend Theorem 1.11 to several choices of construction maps. Note that the lower bounds of [GSV18] on nonuniform black-box reductions apply to *every* construction map.

Extension to efficient hardness amplification constructions. Our results extend to the case that $\mathrm{Con}^{(\cdot)}$ is "sufficiently efficient". More precisely, to the case where $\mathrm{Con}^{(\cdot)}$ is a constant depth and size $\mathrm{poly}(r)$ over the gates $\{$AND,OR,NOT,PARITY$\}$ of unbounded fan-in.

Theorem 1.13 (Limitations on class reductions for efficient hardness amplification)**.** *There exist constants $\delta_0 > 0$, $\nu > 0$, $d_0 > 1$ and a polynomial p such that: Let $\mathrm{Red}^{(\cdot)}(x, \alpha)$ be a $(\frac{1}{2} + \epsilon) \to (1 - \delta)$ class $\mathcal{D}$ reduction for Con, with input length k, output length n and advice length a. Assume that:*

- *$\mathrm{Red}^{(\cdot)}$ is a size r oracle circuit, that makes at most q queries.*
- *The class $\mathcal{D}$ contains circuits of size $p(r)$ and depth d_0 over the gates $\{$AND,OR,NOT,PARITY$\}$ of unbounded fan-in.*
- *$k, n, a, \frac{1}{\epsilon}, \frac{1}{\delta} \leq r \leq 2^{\nu \cdot k}$ and $\delta \leq \delta_0$.*
- *$\mathrm{Con}^{(\cdot)}$ is a size $\mathrm{poly}(r)$ circuit.*

Then the following holds:

- *$\mathrm{Red}^{(\cdot)}$ requires many queries, specifically: $q = \Omega(\frac{\log(1/\delta)}{\epsilon^2})$.*
- *$\mathrm{Red}^{(\cdot)}$ requires majority, specifically: if in addition to the size restriction on $\mathrm{Con}, \mathrm{Red}$, we also have that $\mathrm{Con}^{(\cdot)}, \mathrm{Red}^{(\cdot)}$ are oracle circuits of depth e, d (respectively) over a set of gates G that contains the gates $\{$AND,OR,NOT,PARITY$\}$ of unbounded fan-in, then the majority function over $\Omega(\frac{1}{\epsilon})$ bits can be computed by a circuit of size $\mathrm{poly}(r)$ and depth $O(d + e)$ over the set of gates G.*

In fact, Theorem 1.13 follows by the same proof as Theorem 1.11, as the only property of $f^{\oplus t}$ used in the proof, is that it can be efficiently computed in $\mathrm{AC}^0[2]$. See Remark 3.1.

Extension to hardness amplification based on sufficiently explicit linear codes. A common construction map in hardness amplification is the construction map $\mathrm{Con}^f = f'$ where we think of $f : \{0,1\}^k \to \{0,1\}$ as a 2^k bit long truth table, encode f by an error-correcting code, and let f' be the function whose truth table is this encoding. (It was observed in [STV01] that there is a formal connection between black-box reductions for hardness amplification, and list-decodable error correcting codes, see for example [GSV18] for a discussion). Note that this typically yields a construction map that has size exponential in k (rather than polynomial in k) and is not considered efficient.

Nevertheless, using ideas from [Vio06], our results also extend to the case of $\delta = 2^{-2k}$ (which captures worst-case to average case hardness amplification) for functions f' over $n = 2^{o(k)}$ bits, such that:

$$f'(y) = \sum_{x \in \{0,1\}^k} f(x) \cdot g(x, y),$$

where the sum is taken in the field $\mathbb{F}_2$, and $g : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}$ can be computed by circuits of size $\mathrm{poly}(r)$ and depth d over the gates $\{$AND,OR,NOT,PARITY$\}$ of unbounded fan-in.

This definition of f' corresponds to "hardness amplification by a linear code". More specifically, we can view g as a matrix A of order $2^k \times 2^n$ over $\mathbb{F}_2$ by $A_{x,y} = g(x,y)$, and view the truth tables of the functions f and f' as vectors over $\mathbb{F}_2$ of dimension 2^k and 2^n, respectively. In this interpretation, the definition of f' above, says that $f' = f \cdot A$, for a matrix A in which the entry $A_{x,y}$ can be efficiently computed given x, y.

Many worst-case to average-case hardness amplification results in the literature choose f' so that the truth table f' is obtained by applying a *linear* error correcting code on the truth table of f. Our results apply to this scenario, with the weaker conclusion that $q = \Omega(\frac{\log r}{\epsilon^2})$, and the same conclusion for the case of majority.

Theorem 1.14 (Limitations on class reductions for hardness amplification by linear maps)**.** *There exist constants $\nu > 0$, $d_0 > 1$ and a polynomial p such that: Let $\mathrm{Red}^{(\cdot)}(x,\alpha)$ be a $(\frac{1}{2}+\epsilon) \to (1-2^{-2k})$ class $\mathcal{D}$ reduction for Con, with input length k, output length n and advice length a. Assume that:*

- *$\mathrm{Red}^{(\cdot)}$ is a size r oracle circuit, that makes at most q queries.*
- *The class $\mathcal{D}$ contains circuits of size $p(r)$ and depth d_0 over the gates $\{$AND,OR,NOT,PARITY$\}$ of unbounded fan-in.*
- *$k, n, a, \frac{1}{\epsilon} \leq r \leq 2^{\nu \cdot k}$.*
- *The construction map $\mathrm{Con}^{(\cdot)}$ is defined by*

$$\mathrm{Con}^f = \sum_{x \in \{0,1\}^k} f(x) \cdot g(x,y),$$

  where the sum is taken in the field $\mathbb{F}_2$, and $g : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}$ can be computed by circuits of size $\mathrm{poly}(r)$ and depth d over the gates $\{$AND,OR,NOT,PARITY$\}$ of unbounded fan-in.

Then the following holds:

- *$\mathrm{Red}^{(\cdot)}$ requires many queries, specifically: $q = \Omega(\frac{\log r}{\epsilon^2})$.*
- *$\mathrm{Red}^{(\cdot)}$ requires majority, specifically: if in addition to the size restriction on $\mathrm{Con}, \mathrm{Red}$, we also have that $\mathrm{Red}^{(\cdot)}$ are oracle circuits of depth d (respectively) over a set of gates G that contains the gates $\{$AND,OR,NOT,PARITY$\}$ of unbounded fan-in, then the majority function over $\Omega(\frac{1}{\epsilon})$ bits can be computed by a circuit of size $\mathrm{poly}(r)$ and depth $O(d)$ over the set of gates G.*

1.5 Some more related work

It is beyond the scope of this paper to survey the vast literature on Yao's XOR lemma and hardness amplification. The reader is referred to [GNW11] for a survey on Yao's XOR lemma, and to [Vio06, SV10, GSV18] for detailed discussions on the more general problem of hardness amplification.

A significant advantage of Yao's XOR lemma (over some other suggested methods of hardness amplification) is that the "construction" $f' = f^{\oplus t}$ can be computed very efficiently, when given oracle access to f. A line of work (that is orthogonal to studying the complexity of *reductions* for hardness amplification) is interested in the complexity of *constructions* yielding hardness amplification. This line of work is mostly interested in starting from worst-case hard functions (which correspond to $\delta < 2^{-k}$) and aims to design (or prove impossibility results for) efficient constructions $\mathrm{Con}^{(\cdot)}$ for which one can prove that if f has hardness $1-\delta$, then $f' = \mathrm{Con}^f$ has hardness $\frac{1}{2}+\epsilon$. (See e.g., [TV07, Vio03, LTW08, GV08] for further discussion).

In this orthogonal line of work, there are examples of *constructions* which are non-black-box, and utilize specific properties of the function f (for example that $f \in$ NP or that f is a low degree polynomial). This is a different form of "non-black-box" than the one studied in this paper, and it is interesting to combine the two orthogonal directions.

There is a large body of work on proving black-box impossibility results in cryptography. This study was initiated by Impagliazzo and Rufich [IR89] and is concerned both with issues that are related to black-box constructions and to black-box reductions. See for example the discussion in Reingold, Trevisan and Vadhan [RTV04] for a taxonomy of various notions.

As we explain below, the approach taken in this paper in order to extend the previous black-box lower bounds of Grinberg, Shaltiel and Viola [GSV18], is inspired by a classical work of Goldreich and Krawczyk [GK96] who show black-box impossibility results on "black-box simulation for zero knowledge proofs". Specifically, it is this work that introduced the use of ℓ-wise independent oracles that is used in this work and is explained in the next section.

2 Technique and a road map for proof

Our results are obtained by carefully examining the argument of the black-box impossibility result of [GSV18], replacing the inefficient oracle with an efficient one, and handling the technical difficulties arising from this modification.

In this section we survey our technique, and discuss similarities and differences to [GSV18]. We give a roadmap of the proof of Theorem 1.11, stating the key technical lemmas that are used to prove Theorem 1.11. These technical lemmas are proven in subsequent sections.

We assume the setup of Theorem 1.11. Specifically, let $\mathrm{Red}^{(\cdot)}(x,\alpha)$ be a $(\frac{1}{2}+\epsilon) \to (1-\delta)$ class $\mathcal{D}$ reduction for Yao's XOR lemma, with input length k, t repetitions and advice length a, which makes at most q queries and satisfies the requirements of the theorem. We will assume w.l.o.g. that $\mathrm{Red}^{(\cdot)}$ does not make the same query twice. Let r be the size of Red and let d (which is not necessarily a constant) be the depth of Red. Our goal is to show that $\mathrm{Red}^{(\cdot)}$ requires many queries, and that $\mathrm{Red}^{(\cdot)}$ requires majority. We will assume that $\frac{1}{\epsilon}$ is a power of two.

Let $f : \{0,1\}^k \to \{0,1\}$ be some function that we choose later, and let $n = tk$ be the input length of $f^{\oplus t}$. We start by surveying the approach of of the previous papers (which only handle *black-box* reductions rather than *class* reductions).

2.1 The approach of [Vio06, SV10]

We first introduce the following notation.

Definition 2.1 (Random sequences/functions)**.** *For a number $0 \leq p \leq 1$, and an integer q, we define a distribution Noise^q_p over $\{0,1\}^q$ which consists of q i.i.d. bit variables $\mathsf{Noise}^q_p(1),\ldots,\mathsf{Noise}^p_q(q)$ where each of them has probability p to be one. This notation also allows us to view Noise^q_p as a distribution over functions from $[q]$ to $\{0,1\}$.*

Following [Vio06, SV10] (and as done in later works [GR08, GSV18]) our plan is to show that a $(\frac{1}{2}+\epsilon) \to (1-\delta)$ reduction $\mathrm{Red}^{(\cdot)}(x,\alpha)$ that makes q queries, can be transformed into a (distribution) over circuits $T : \{0,1\}^q \to \{0,1\}$ with no oracle (that have roughly the same size and depth as Red) that distinguishes $\mathsf{Noise}^q_{1/2-2\epsilon}$ from $\mathsf{Noise}^q_{1/2}$. We will prove the following lemma (which we call the "zoom lemma").

Lemma 2.2 (Zoom lemma)**.** *If $\epsilon = a/b$ where $a \leq b \leq r$ are integers, and b is a power of two, then under the assumption of Theorem 1.11, for every $x \in \{0,1\}^k$, there exists a circuit T_x over q bits, with size $\mathrm{poly}(r)$ and depth $O(d)$ over the set of gates G, such that:*

- $\Pr_{X \leftarrow U_k}[T_X(\mathsf{Noise}^q_{1/2-2\epsilon}) = 1] \geq 1 - 2\delta$.
- $\Pr_{X \leftarrow U_k}[T_X(\mathsf{Noise}^q_{1/2}) = 1] \leq \frac{1}{2} + \frac{1}{200}$.

Shaltiel and Viola [SV10] (see also [LSS+19]) showed that Theorem 1.11 follows from Lemma 2.2. This is formally stated and explained in Section 5.[6]

In the remainder of this section, we prove Lemma 2.2 modulo some other lemmas and claims, that are stated in this section, and proven in later sections of the paper.

2.2 The oracle used for black-box reductions

Lemmas that are similar to the zoom lemma are at the heart of earlier results [SV10, GR08, GSV18] on *black-box* reductions, and we would like to imitate the argument working with *class* reductions. Let us start by explaining the oracle used in previous works.

Specifically, let us set $N = 2^n$ and identify the set $[N]$ with the set $\{0,1\}^n$ (so that we can think of Noise^N_p as a function $\mathsf{Noise}^N_p : \{0,1\}^n \to \{0,1\}$). We consider the following two (distributions over) oracles $D : \{0,1\}^n \to \{0,1\}$.

- $D_{1/2-2\epsilon}(y) = f^{\oplus t}(y) \oplus \mathsf{Noise}^N_{1/2-2\epsilon}(y)$
- $D_{1/2}(y) = f^{\oplus t}(y) \oplus \mathsf{Noise}^N_{1/2}(y)$.

Definition 2.3. *We say that a function $D : \{0,1\}^n \to \{0,1\}$ is* useful*, if there exists an $\alpha \in \{0,1\}^a$ such that $\Pr_{X \leftarrow U_k}[\mathrm{Red}^D(X, \alpha) = f^{\oplus t}(X)] \geq 1 - \delta$.*

In the oracle $D_{1/2}$, the noise $\mathsf{Noise}^N_{1/2}(y)$ is uniform and completely masks out the information in $f^{\oplus t}(y)$. Intuitively, this means that the oracle $D_{1/2}$ isn't useful for the reduction. On the other hand, a Chernoff bound shows that w.h.p. over choosing $h \leftarrow \mathsf{Noise}^N_{1/2-2\epsilon}$, we have that $|\{y \in \{0,1\}^n : h(y) = 1\}| \leq (\frac{1}{2} - \epsilon) \cdot N$. This gives that w.h.p. over choosing $D \leftarrow D_{1/2-2\epsilon}$, we have that $\Pr_{Y \leftarrow U_n}[D(Y) = f^{\oplus t}(Y)] \geq \frac{1}{2} + \epsilon$.

If Red is a $(\frac{1}{2} + \epsilon) \to (1 - \delta)$ *black-box* reduction, then by definition, this implies that every such good D is useful. The proof of [Vio06, SV10] then proceeds to transform a *black-box* reduction Red into the circuits T_x required from Lemma 2.2. We will elaborate on this argument shortly.

Our plan is to imitate this argument when Red is not necessarily a *black-box reduction*, and is only guaranteed to be a *class* reduction. Using this weaker assumption, we are not guaranteed that w.h.p. $D \leftarrow D_{1/2-2\epsilon}$ is useful. This is because we are not guaranteed that w.h.p. $D \leftarrow D_{1/2-2\epsilon}$ belongs to the class $\mathcal{D}$, and the reduction does not need to succeed if $D \notin \mathcal{D}$.

[6] On an intuitive level, the connection between the consequence of the zoom lemma and the consequence of Theorem 1.11 is that the "best way" to distinguish $\mathsf{Noise}^q_{1/2-2\epsilon}$ from $\mathsf{Noise}^q_{1/2}$ is to check whether the fraction of ones is below or above $\frac{1}{2} - \epsilon$. This is similar in spirit to majority over inputs of length $\Omega(1/\epsilon)$, and it can be shown that majority on length $\Omega(1/\epsilon)$ can be reduced to this task. A Chernoff bound shows that $q = O(\frac{\log(1/\delta)}{\epsilon^2})$ is sufficient to distinguish between the two distributions with confidence $1 - \delta$, and it can be shown that such a confidence requires $q = \Omega(\frac{\log(1/\delta)}{\epsilon^2})$.

2.3 Using limited independence to obtain efficient oracles

We would like to make the oracle $D_{1/2-2\epsilon}$ efficiently computable by small circuits, so that it belongs to $\mathcal{D}$. This presents two difficulties:

1. $f^{\oplus t}$ is harder to compute than f (and f is assumed to be hard).
2. $\mathsf{Noise}^N_{1/2-2\epsilon}$ is a random function, and w.h.p. requires circuits of exponential size.

In order to circumvent the first problem we use an idea from [Vio06] and will choose the function $f : \{0,1\}^k \to \{0,1\}$ to be a specific function with properties stated in the next lemma (in which the reader should recall that r is the size of the oracle circuit $\mathrm{Red}^{(\cdot)}$).

Lemma 2.4. *There exist constants c_1 such that for every constant c_2, there exists a function $f : \{0,1\}^k \to \{0,1\}$ such that:*

- *For every circuit $B : \{0,1\}^k \to \{0,1\}$ of size r^{c_2}, $\Pr_{X \leftarrow U_k}[B(X) = f(X)] \leq \frac{1}{2} + \frac{1}{200}$.*
- *f can be computed by a DNF of size $r^{c_1 \cdot c_2}$.*

Proof. By a standard counting argument, there exists a constant c_1 such that for every constant c_2, setting $m = c_1 \cdot c_2 \cdot \log r$, there exists a function $g : \{0,1\}^m \to \{0,1\}$ such that for every circuit B of size $2^{m/c_1} = r^{c_2}$, $\Pr_{X \leftarrow U_m}[B(X) = g(X)] \leq \frac{1}{2} + \frac{1}{200}$. Recall that the conditions of Theorem 1.11 allow us to choose a constant $\nu > 0$, such that the condition $r \leq 2^{\nu \cdot k}$ holds. By choosing $\nu > 0$ to be sufficiently small as a function of c_1 and c_2, we can get that $m \leq k$. The function $f : \{0,1\}^k \to \{0,1\}$ is the function that given $x \in \{0,1\}^k$ applies g on the first m bits of x. □

We choose f by the lemma, where c_2 is a constant that we choose later. With this choice, and using the assumption that $t \leq r$ we have that:

Corollary 2.5. *The function $f^{\oplus t}$ can be computed by circuits of size $\mathrm{poly}(r)$ and constant depth over the gates $\{$AND,OR,NOT,PARITY$\}$ of unbounded fan-in.*

Remark 2.6 (Replacing $f^{\oplus t}$ by a different target function f')**.** *Corollary 2.5 is the only place in the proof where we use specific properties of $f^{\oplus t}$. The corollary holds for every function $f' : \{0,1\}^n \to \{0,1\}$ for which there exists an oracle circuit $\mathrm{Con}^{(\cdot)}$ of size $\mathrm{poly}(r)$ and constant depth over the gates $\{$AND,OR,NOT,PARITY$\}$ of unbounded fan-in, such that $f' = \mathrm{Con}^f$. This means that our results hold for every such function f'. Furthermore, if Con does not have constant depth, then Corollary 2.5 gives a size bound on f', and this is sufficient to show the lower bound on number of queries with respect to the class $\mathcal{D}$ of circuits of size $p(r)$.*

Corollary 2.5 takes care of the first difficulty above. It says that $f^{\oplus t}$ can be computed by circuits in the class $\mathcal{D}$. We would like to replace $\mathsf{Noise}^N_{1/2-2\epsilon}$ by a (distribution) over efficient circuits in $\mathcal{D}$. Our approach is to replace $\mathsf{Noise}^N_{1/2-2\epsilon}$ (which consists of N independent bits) by a distribution which is ℓ-wise independent, for $\ell = \mathrm{poly}(r)$.

Definition 2.7 (ℓ-wise independence with bias p)**.** *A sequence $R_1, \ldots, R_N$ of bit random variables is ℓ-*wise independent with bias *p, if $R_1, \ldots, R_N$ are ℓ-wise independent, and for every $i \in [N]$, $\Pr[R_i = 1] = p$.*

Gutfreund and Viola [GV04] gave a construction of an "ℓ-wise independent generator" where computing individual output bits of the generator can be done by a constant depth circuit with parity gates. We now state this result.

Theorem 2.8. *[GV04, Theorem 10] Let $N = 2^n$ be sufficiently large. For every integer $\ell \leq N$, there is a function $G : \{0,1\}^r \to \{0,1\}^N$, such that:*

- $r = \mathrm{poly}(n, \ell)$.
- *The N random variables obtained by in the experiment $G(U_r)$ are ℓ-wise independent with bias $1/2$.*
- *The function $g(x, i) = G_i(x)$ can be computed by a circuit of size $\mathrm{poly}(n, \ell)$ and depth $O(1)$ over the gates $\{$AND,OR,NOT,PARITY$\}$ of unbounded fan-in.*[7]

It is standard to interpret a generator G as a collection of "ℓ-wise independent hash functions". That is, for every $x \in \{0,1\}^r$, we define $h_x : \{0,1\}^n \to \{0,1\}$, by identifying $[N]$ with $\{0,1\}^n$, and defining $h_x(i) = g(x, i)$. Let $H^\ell_{1/2}$ denote the distribution of choosing a random $x \leftarrow U_r$ and obtaining the function $h = h_x$. It follows that the random variables $h(1), \ldots, h(N)$ are ℓ-wise independent with bias $p = 1/2$. In the corollary below, we observe that it is easy to extend this result to other values of p (as we will be interested in $p = 1/2 - 2\epsilon$).

Corollary 2.9. *Let $N = 2^n$ be sufficiently large. For every integers $\ell \leq N$ and $a \leq b$, such that b is a power of 2, setting $p = \frac{a}{b}$, there exists a distribution H^ℓ_p over circuits $h : \{0,1\}^n \to \{0,1\}$ of size $\mathrm{poly}(n, \ell, b)$ and depth $O(1)$ (over the gates $\{$AND,OR,NOT,PARITY$\}$ of unbounded fan-in) such that the distribution obtained by choosing $h \leftarrow H^\ell_p$ and considering $(h(1), \ldots, h(N))$, is ℓ-wise independent with bias p.*

Proof. Given N, ℓ and $p = \frac{a}{b}$, we apply Theorem 2.8 with $N' = N \cdot \log b$. We think of G as $G' : \{0,1\}^r \to (\{0,1\}^{\log b})^N$ and for every $x \in \{0,1\}^r$, we define $h_x : \{0,1\}^n \to \{0,1\}$ by $h(i) = 1$ if the binary representation of $G'_i(x)$ is at most a. Note that for every $x \in \{0,1\}^r$, h_x requires computing $\log b$ instantiations of $G_i(x)$, and comparing binary strings of length $\log b$. It follows that for every $x \in \{0,1\}^r$, h_x and can be computed by circuits of constant depth and size $\mathrm{poly}(n, \ell, b)$ (or in fact, even $\mathrm{poly}(n, \ell, \log b)$ if we are more careful). We also obviously have that the distribution obtained by choosing $h \leftarrow H^\ell_p$ and considering $(h(1), \ldots, h(N))$, is ℓ-wise independent with bias p. □

Let $\ell = p_0(r)$ for a polynomial p_0 that we will specify later. We define the following two (distributions over) oracles $D : \{0,1\}^n \to \{0,1\}$, in which we replace the independent bits of Noise^N by ℓ-wise independent bits:[8]

- $D^\ell_{1/2-2\epsilon}(y) = f^{\oplus t}(y) \oplus H^\ell_{1/2-2\epsilon}(y)$
- $D^\ell_{1/2}(y) = f^{\oplus t}(y) \oplus H^\ell_{1/2}(y)$.

We now have that every D in the support of $D^\ell_{1/2-2\epsilon}$ has size $r^{c_1 \cdot c_2} + \mathrm{poly}(n, \ell, 1/\epsilon)$ which can be bounded by $p(r)$ for some polynomial p. Furthermore, each such D has constant depth over the set of gates $\{$AND,OR,NOT,PARITY$\}$. This gives that every such D is sufficiently efficient, and belongs to the class $\mathcal{D}$.

[7] We remark that the result of Gutfreund and Viola [GV04] is stronger, giving a tighter bound on r, and a uniform circuit for computing $g(x, i)$.

[8] Replacing fully independent oracles by limited independence oracles, and arguing that black-box procedures with few queries cannot tell the difference, is a common approach in proving black-box impossibility results, originating from the work of Goldreich and Krawczyk [GK96]. It should be noted that even when ignoring the issue of class reductions, and focusing on black-box reductions, we are considering reductions which are *nonuniform*. Nonuniform reductions get an advice string α that depends on the choice of the oracle. Loosely speaking, this may give them information about the "seed" used to generate the limited independence oracle. This creates technical difficulties that do not occur when reductions are uniform.

This will allow us to imitate the argument for black box reductions. Specifically, by Chebyshev's inequality, with probability at least $1 - \frac{1}{\epsilon^2 2^n} \geq \frac{1}{2}$ over choosing $h \leftarrow H^{\ell}_{1/2-2\epsilon}$, we have that:

$$|\{y \in \{0,1\}^n : h(y) = 1\}| \leq (\frac{1}{2} - \epsilon) \cdot N.$$

This means that with probability at least half over choosing $D \leftarrow D^{\ell}_{1/2-2\epsilon}$, we have that:

$$\Pr_{Y \leftarrow U_n}[D(Y) = f^{\oplus t}(Y)] \geq \frac{1}{2} + \epsilon.$$

As Red is a $(\frac{1}{2} + \epsilon) \to (1 - \delta)$ class $\mathcal{D}$ reduction, and every such good D belongs to $\mathcal{D}$, we get that:

Claim 2.10. $\Pr_{h \leftarrow H^{\ell}_{1/2-2\epsilon}}[(f^{\oplus t} \oplus h)$ *is useful*$] \geq \frac{1}{2}$.

2.4 A more general fixed set lemma

We will now proceed in a similar manner to [SV10, GSV18]. Specifically, let Advice be a function that given a useful D, produces an advice string α such that $\Pr_{X \leftarrow U_k}[\mathrm{Red}^D(x, \alpha) = f^{\oplus t}(y)] \geq 1 - \delta$ (such an α exists by definition). For every $\alpha \in \{0,1\}^a$, let A_α be the event

$$A_\alpha = \left\{h : \{0,1\}^n \to \{0,1\} : (f^{\oplus t} \oplus h) \text{ is useful, and } \mathsf{Advice}(f^{\oplus t} \oplus h) = \alpha\right\}.$$

By averaging over the 2^a advice strings we obtain that:

Claim 2.11. *There exists* $\alpha' \in \{0,1\}^a$ *s.t.* $\Pr_{h \leftarrow H^{\ell}_{1/2-2\epsilon}}[h \in A_{\alpha'}] \geq \frac{1}{2} \cdot 2^{-a} = 2^{-(a+1)}$.

Let $R = H^{\ell}_{1/2-2\epsilon}$ and $Z = (R|R \in A_{\alpha'})$, following [SV10, GSV18] we would like to argue that (in some sense to be explained below) for every $x \in \{0,1\}^k$, $\mathrm{Red}^{(\cdot)}(x, \alpha')$ does not distinguish between the oracle $f^{\oplus t} \oplus Z$ (in which bits can be correlated in complicated ways) and the oracle $f^{\oplus t} \oplus R$ (in which bits are ℓ-wise independent).

At this point we have already chosen α' and f. For every $x \in \{0,1\}^k$, and $h : \{0,1\}^n \to \{0,1\}$, we can think of the computation $\mathrm{Red}^{f^{\oplus t} \oplus h}(x, \alpha')$ as a decision tree $P_x(h)$ (that depends on x and the fixed choices of α' and f) that makes q queries to (the $N = 2^n$ bit long truth table of) h.

With this intuition in mind, we will prove the following lemma (which generalizes a "fixed set lemma" proven in [GSV18] for the special case where the random variables $R_1, \ldots, R_N$ are independent).

Lemma 2.12 (A more general fixed set lemma)**.** *Let N, a and q be integers, and let $R = (R_1, \ldots, R_N)$ be some distribution, let $A \subseteq \{0,1\}^N$ be an event such that $\Pr[R \in A] \geq 2^{-a}$, and let $Z = (R|R \in A)$. For every $\eta > 0$, there exists a set $B \subseteq [N]$ of size $b \leq O(a \cdot q/\eta)$, and $v \in \{0,1\}^B$ in the support of Z_B, such that for $R' = (R|R_B = v)$ and $Z' = (Z|Z_B = v) = (R|R_B = v, R \in A)$, and every q-query decision tree P, $P(R')$ and $P(Z')$ are η-close.*[9]

Remark 2.13 (Proof sketch and comparison to the proof technique of [GSV18])**.** *The proof works by showing that if there exists a q-query decision tree that distinguishes R from Z, then by fixing the variables on some path of that tree, one obtains a distribution R' such that $\Pr[R' \in A] \geq \Pr[R \in A] \cdot (1 + \eta)$. We apply this argument iteratively (using R' as R) until there does not exist a q query decision tree that distinguishes*

[9] Two distributions X, Y over the same domain S are η-close if for every $A \subset S$, $|\Pr[X \in A] - \Pr[Y \in A]| \leq \eta$.

R from Z. In each iteration, $\Pr[R \in A]$ increases by a factor of $1+\eta$, and as this probability cannot be larger than one, this process has to stop after $O(a/\eta)$ steps. By then, we have fixed no more than $O(qa/\eta)$ of the variables. The full proof of Lemma 2.12 appears in Section 4.

This iterative approach is very similar to the one used in [GSV18]. Two proofs are given in [GSV18] (the second one is attributed to an anonymous referee). Indeed, the iterative approach is used in both proofs in [GSV18]. Namely, the proofs show that if there exists a q-query decision tree that distinguishes R from Z, then by fixing the variables on some path of that tree, we make progress on some "progress measure", and that there is a bound on the number of times that this can happen.

Both proofs in [GSV18] rely on the assumption that R has "maximal entropy". More precisely, under the condition assumed in [GSV18] (that $R_1, \ldots, R_N$ are independent) it follows that w.l.o.g. one can assume that each R_i is uniformly distributed. This means that $Z = (R|R \in A)$ is uniformly distributed on A, and in particular has high entropy, and large support size. These two measures are used (respectively) in the two proofs presented in [GSV18].

Here, we are interested in dealing with a distribution R that does not have high entropy, and observe that it is possible to imitate the argument of [GSV18] while using $\Pr[R \in A]$ as a progress measure. This allows the argument to go through, even when the initial R is a low entropy distribution, in which the individual variables are correlated.

We now continue with the proof of Lemma 2.2. We apply Lemma 2.12 on $R = H^{\ell}_{1/2-2\epsilon}$ and the event A'_{α}, using $\eta = \delta$, and let Z, R', Z', B, v and b be as in the lemma. It follows that for every $x \in \{0,1\}^k$, the random variables $P_x(R') = \mathrm{Red}^{f^{\oplus t} \oplus R'}(x, \alpha')$ and $P_x(Z') = \mathrm{Red}^{f^{\oplus t} \oplus Z'}(x, \alpha')$ are δ-close.

As this holds for every fixed $x \in \{0,1\}^k$, this also holds for an independently chosen $X \leftarrow U_k$, and we obtain that:

$$\Pr[\mathrm{Red}^{f^{\oplus t} \oplus R'}(X, \alpha') = f(X)] \geq \Pr[\mathrm{Red}^{f^{\oplus t} \oplus Z'}(X, \alpha') = f(X)] - \delta.$$

(More formally, in the probabilities above, $X \leftarrow U_k$ is a random variable that is independent of the choice of R' in the right hand side, and independent of the choice of Z' in the left hand side).

The support of Z' is contained in $A_{\alpha'}$, and so, for every h in the support of Z', $(f^{\oplus t} \oplus h)$ is useful (with the advice string α') and we get that:

$$\Pr[\mathrm{Red}^{f^{\oplus t} \oplus Z'}(X, \alpha') = f(X)] \geq 1 - \delta. \tag{1}$$

Combining this with the previous inequality, gives that:

$$\Pr[\mathrm{Red}^{f^{\oplus t} \oplus R'}(X, \alpha') = f(X)] \geq 1 - 2\delta. \tag{2}$$

The advantage of (2) over (1) is that we have replaced Z' (in which the bits of $Z'([N] \setminus B)$ can be correlated in complicated ways) with R', where $R'([N] \setminus B)$ is $(\ell - b)$-wise independent. This will allow us to relate this oracle to $\mathsf{Noise}^q_{1/2-2\epsilon}$ and prove the zoom lemma.

2.5 Constructing the circuits for the zoom lemma

We would like to show that a reduction $\mathrm{Red}^{(\cdot)}$ that makes at most q queries, can be used to obtain the circuits guaranteed in the zoom lemma. For this purpose, we define the following oracle circuit.

Definition 2.14. *We define an oracle circuit $E^{(\cdot)}(x)$ as follows: On input x and oracle h, $E^h(x)$ simulates $\mathrm{Red}^{(\cdot)}(x, \alpha')$. Whenever* Red *makes a query y to its oracle, R acts as follows: if $y \notin B$, then R makes the query y to h, and returns $f^{\oplus t}(y) \oplus h(y)$ to* Red*. If $y \in B$, then R returns $f^{\oplus t}(y) \oplus v(y)$ to* Red*. The output of R is the output of* Red *at the end of this simulation.*

We will now show that the oracle circuit $E^{(\cdot)}$ gives rise to the circuits T_x required to prove Lemma 2.2.

More specifically, recall that we have not yet chosen the constant c_2, and the polynomial p_0 (which determines ℓ). In the next lemma we show that by choosing c_2 and p_0, we can get that on an independent and uniformly chosen $X \leftarrow U_k$, $E^{(\cdot)}(X)$ distinguishes between the case that its oracle function h is chosen from a distribution $H^q_{1/2-2\epsilon}$ (that is q-wise independent with bias $1/2 - 2\epsilon$) and the case where the oracle function h is chosen from a distribution $H^q_{1/2}$ (that is q-wise independent with bias $1/2$).

Lemma 2.15. *By choosing the constant c_2 and the polynomial p_0 to be sufficiently large, we get that:*

- $\Pr[E^{H^q_{1/2-2\epsilon}}(X) = f(X)] \geq 1 - 2\delta$.
- $\Pr[E^{H^q_{1/2}}(X) = f(X)] \leq \frac{1}{2} + \frac{1}{200}$.
- *For every $x \in \{0,1\}^k$, there exists a circuit $T_x : \{0,1\}^q \to \{0,1\}$ of size $\mathrm{poly}(r)$ and depth $O(d)$ over the gates G, such that for every $0 \leq p \leq 1$, $T_x(\mathsf{Noise}^q_p) = E^{H^q_p}(x)$.*

The proof of Lemma 2.15 appears in Section 3, and is similar in spirit to earlier work [SV10, GR08, GSV18]. It is in fact significantly simpler, as in this paper, we have the additional advantage that $f^{\oplus t}$ has circuits of size $\mathrm{poly}(r)$ and constant depth.
Note that Lemma 2.15 immediately implies Lemma 2.2.

Organization of the paper

Our goal is to prove the main theorem (Theorem 1.11). We start by concluding the proof of the zoom lemma (Lemma 2.2). In Section 2 we showed that the zoom lemma (Lemma 2.2), follows once we prove Lemma 2.15 and the more general fixed set lemma (Lemma 2.12). We prove Lemma 2.15 in Section 3. We prove Lemma 2.12 in Section 4.

Having established the zoom lemma (Lemma 2.2), in Section 5 we state and survey the results of [SV10] showing that the zoom lemma (Lemma 2.2) implies the main theorem (Theorem 1.11).

In Section 6 we explain how to extend the argument to sufficiently explicit linear codes.

3 Proof of Lemma 2.15

In this section we prove Lemma 2.15. Note that by the assumptions of Theorem 1.11 we have that $q, \frac{1}{\delta} \leq r$, and the number of queries q that Red makes, satisfies $q \leq r$. This gives that $b = O(qa/\delta)$ is bounded by some polynomial in r. We are allowed to choose the polynomial p_0 to be sufficiently large so that $\ell = p_0(r)$ satisfies $(\ell - b) \geq r \geq q$. This gives that the $N - b$ coordinates of $R'([N] \setminus B)$ are $(\ell - b)$-wise independent (because R' was obtained by fixing b indices of R which is ℓ-wise independent).

Recall that we are also assuming w.l.o.g. that $\mathrm{Red}^{(\cdot)}$ does not make the same query twice. Throughout the proof, we will use the observation that when an oracle procedure $P^{(\cdot)}$ makes q queries to an oracle function chosen from a distribution H^q_p that is q-wise independent with bias p, then the distribution of answers that P sees is distributed like Noise^q_p. (This is immediate if P makes nonadaptive queries, but also holds if P makes adaptive queries, as every sequence $a_1, \ldots, a_q \in \{0,1\}$ of answers, is obtained with probability $p^{|\{i:a_i=1\}|} \cdot (1-p)^{|\{i:a_i=0\}|}$).

We have that $R'([N] \setminus B)$ are q-wise independent with bias $1/2 - 2\epsilon$, and that E answers queries in B using v. By the observation above, this gives that for every $x \in \{0,1\}^k$, the queries and answers

that Red makes in the evaluation of $E^{H^q_{1/2-2\epsilon}}(x)$ are distributed exactly like the queries and answers of $\mathrm{Red}^{f^{\oplus t}\oplus R'}(x,\alpha')$, meaning that for every $x\in\{0,1\}^k$:

$$\Pr[E^{H^q_{1/2-2\epsilon}}(x)=f(x)]=\Pr[\mathrm{Red}^{f^{\oplus t}\oplus R'}(x,\alpha')=f(x)].$$

This immediately means that for an independent $X\leftarrow U_k$:

$$\Pr[E^{H^q_{1/2-2\epsilon}}(X)=f(X)]=\Pr[\mathrm{Red}^{f^{\oplus t}\oplus R'}(X,\alpha')=f(X)].$$

We have already seen in (2) that:

$$\Pr[\mathrm{Red}^{f^{\oplus t}\oplus R'}(X,\alpha')=f(X)]\geq 1-2\delta,$$

and this gives the first item.

For the second item, we note that if $E^{H^q_{1/2}}$ makes a query $y\notin B$, then it obtains a uniform coin, and the coins obtained on different queries are independent. This follows because here the bias is $1/2$, and so when E xors the uniform bit answer that it gets from the oracle $H^q_{1/2}$ with the fixed $f^{\oplus t}(y)$, it obtains a random coin.

Recall that on queries $y\in B$, E answers the queries without consulting the oracle. we can simulate $E^{H^q_{1/2}}(x)$ by a randomized circuit $\bar{C}$ that on input x, simulates E and answers queries $y\notin B$ by random coins. (In particular, note that $\bar{C}$ does not need to compute $f^{\oplus t}(y)$ for $y\notin B$.

It follows that for every $x\in\{0,1\}^k$:

$$\Pr[E^{H^q_{1/2}}(x)=f(x)]=\Pr[\bar{C}(x)=f(x)].$$

This immediately means that for an independent $X\leftarrow U_k$:

$$\Pr[E^{H^q_{1/2}}(X)=f(X)]=\Pr[\bar{C}(X)=f(X)].$$

By an averaging argument, there exists some fixing for the random coins of $\bar{C}$ such that the obtained (deterministic) circuit C satisfies $Pr[C(X)=f(X)]\geq\Pr[\bar{C}(X)=f(X)]$. The circuit C is hardwired with this choice of random coins, and with α', B, v, and $f^{\oplus t}(B)$. (Again, a crucial observation is that C does not need to compute $f^{\oplus t}$ for $y\notin B$). Overall, this is a circuit of size r^c for some constant c, and by choosing the constant c_2 to be a larger constant, and using Lemma 2.4, we have that:

$$\Pr[E^{H^q_{1/2}}(X)=f(X)]\leq\Pr[C(X)=f(X)]\leq\frac{1}{2}+\frac{1}{200}.$$

This proves the second item.

For the third item, we recall that for every p, and for every $x\in\{0,1\}^k$, the distribution of the q answers that $E^{H^q_p}$ obtains from its oracle is distributed like Noise^q_p. This means that for every x, we can construct a circuit T_x (with no oracle) that on input from No^q_p simulates $E^{H^q_p}(x)$, using its i'th input bit to answer the i'th query of E. The circuit T_x is hardwired with α', B, v and $f^{\oplus t}(B)$. Unlike the circuit C from the second item, T_x does need to compute $f^{\oplus t}$ on each of the q queries. This is because for $p=1/2-2\epsilon$, the answer from H^q_p does not mask out the corresponding output of $f^{\oplus t}$.

However, by Corollary 2.5, the function $f^{\oplus t}$ can be computed by circuits of size $\mathrm{poly}(r)$ and constant depth over the gates $\{\text{AND,OR,NOT,PARITY}\}$ of unbounded fan-in. Therefore, we can implement T_x by a circuit of size $\mathrm{poly}(r)$ (this time the polynomial is larger than r^{c_2}) and depth $O(d)$ over the gates

$\{$AND,OR,NOT,PARITY$\}$ of unbounded fan-in. The increase in depth occurs because on every oracle call of Red, T_x may have to compute $f^{\oplus t}$ (which takes constant depth according to Corollary 2.5).[10]

Remark 3.1 (Extension to the case of Theorem 1.13)**.** *In Theorem 1.13 we stated a version of Theorem 1.11 for the case where $f' = \mathrm{Con}^f$ and Con is efficient.*

The only place in the proof of Theorem 1.11 where we use the specific choice of $f' = f^{\oplus t}$ is within the proof of Lemma 2.15. More specifically, in the final step of the proof of Lemma 2.15, in which we used Corollary 2.5 in order to argue that T_x is a circuit of size $\mathrm{poly}(r)$ and depth $O(d)$ over the gates $\{$AND,OR,NOT,PARITY$\}$ of unbounded fan-in.

However, note that we get the same conclusion also in the case that $f' = \mathrm{Con}^f$ for an oracle circuit $\mathrm{Con}^{(\cdot)}$ that has size $\mathrm{poly}(r)$ and depth e (as assumed in Theorem 1.13). With this modification the proof of Theorem 1.11 also yields a proof of Theorem 1.13 as from here the proof of Theorem 1.13 can proceed unchanged.

4 Proof of the general fixed-set lemma

In this section we prove Lemma 2.12. The proof will iteratively apply the following lemma.

Lemma 4.1. *Let $R = (R_1, \ldots, R_N)$ be a distribution, let $A \subseteq \{0,1\}^N$ be an event, and let $Z = (R|R \in A)$. If there exists a q-query decision tree P such that $|\Pr[P(R) = 1] - \Pr[P(Z) = 1]| > \eta$ then there exists $Q \subset [N]$ of size q and $v \in \{0,1\}^Q$ in the support of Z_Q, such that*

$$\Pr[R \in A | R(Q) = v] > (1 + \eta) \cdot \Pr[R \in A].$$

Proof. Let P be a q-query decision tree, and assume w.l.o.g. (by complementing P if necessary) that $\Pr[P(Z) = 1] - \Pr[P(R) = 1] > \eta$. A path in the decision tree corresponds to a subset $Q \subset [N]$ of the q variables queried on the path, and a string $v \in \{0,1\}^q$ of the answers. For every such path, let $\mathit{path}_{Q,v} : \{0,1\}^N \to \{0,1\}$ be the function that evaluates to 1 on input $r = (r_1, \ldots, r_N)$ if $r(Q) = v$ (meaning that the tree P takes the path (Q, v) on input r). Let S be the set of all pairs (Q, v) corresponding to paths of P that answer 1. The path taken by a decision tree is unique, and therefore, for any distribution R on $\{0,1\}^N$, we have that:

$$\Pr[P(R) = 1] = \sum_{(Q,v) \in S} \Pr[\mathit{path}_{Q,v}(R) = 1].$$

Claim 4.2. *There exists a path $(Q, v) \in S$ such that:*

$$\Pr[\mathit{path}_{Q,v}(Z) = 1] > (1 + \eta) \cdot \Pr[\mathit{path}_{Q,v}(R) = 1].$$

Proof. (of claim)

[10] Our proof of Lemma 2.15 relies on the fact that $f^{\oplus t}$ has small constant depth circuits. This allows us to simplify the argument used by some of the previous work [SV10, GR08, GSV18] which wasn't allowed to assume that the target function $f' = f^{\oplus t}$ can be computed by a small constant depth circuit. The proofs in [SV10, GR08, GSV18] need to resort to different arguments (and this creates additional difficulties if Red makes adaptive calls to its oracle, meaning that the queries that $\mathrm{Red}^{(\cdot)}(x, \alpha')$ makes are not a function of x and α', and may also depend on previous answers). However, using a clever hybrid argument of [GR08] and additional ideas explained in [GSV18], it is possible to conclude that T_x has depth $O(d)$ *without* relying on the fact that f' is computable by constant depth circuits.

This is because otherwise:

$$\begin{aligned}\Pr[P(Z)=1] &= \sum_{(Q,v)\in S} \Pr[\mathit{path}_{Q,v}(Z)=1] \\ &\leq (1+\eta)\cdot \sum_{(Q,v)\in S} \Pr[\mathit{path}_{Q,v}(R)=1] \\ &= (1+\eta)\cdot \Pr[P(R)=1] \\ &\leq \Pr[P(R)=1]+\eta.\end{aligned}$$

□

In particular, the events $\{Z(Q)=v\}$ and $\{R(Q)=v\}$ occur with positive probability, and it follows that:

$$\begin{aligned}\Pr[R\in A|R(Q)=v] &= \frac{\Pr[R\in A\cap R(Q)=v]}{\Pr[R(Q)=v]} \\ &= \frac{\Pr[R\in A]\cdot\Pr[R(Q)=v|R\in A]}{\Pr[\mathit{path}_{Q,v}(R)=1]} \\ &= \frac{\Pr[R\in A]\cdot\Pr[Z(Q)=v]}{\Pr[\mathit{path}_{Q,v}(R)=1]} \\ &= \frac{\Pr[R\in A]\cdot\Pr[\mathit{path}_{Q,v}(Z)=1]}{\Pr[\mathit{path}_{Q,v}(R)=1]} \\ &> (1+\eta)\cdot\Pr[R\in A].\end{aligned}$$

□

We are now ready to prove Lemma 2.12.

Proof. (of Lemma 2.12) We consider the following iterative process: At step i, we have:

- A distribution $R^{(i)}$ over $\{0,1\}^N$.
- A set $B^{(i)}\subseteq[N]$.
- $v^{(i)}\in\{0,1\}^{B^{(i)}}$.

We will assume that the following invariant is satisfied:

- $B^{(i)}$ is of size $i\cdot q$.
- $\Pr[R^{(i)}\in A]\geq 2^{-a}\cdot(1+\eta)^i$.
- $R^{(i)}(B^{(i)})=v^{(i)}$ (with probability one).

Note that the assumption in the lemma fulfills this invariant for $i=0$ with $R^{(0)}=R.$ and $B^{(0)}=\emptyset$.

At step i, we define $\bar{R}=R^{(i)}([N]\setminus B^{(i)})$. As $R^{(i)}$ is fixed on $B^{(i)}$, we can think of A as an event that only observes the indices in $[N]\setminus B^{(i)}$. More formally, there is an event $\bar{A}\subseteq\{0,1\}^{[N]\setminus B^{(i)}}$ such that $R^{(i)}\in A$ iff $\bar{R}\in\bar{A}$, and

$$\Pr[\bar{R}\in\bar{A}]=\Pr[R^{(i)}\in A]\geq 2^{-a}\cdot(1+\eta)^i.$$

Let $\bar{Z}=(\bar{R}|\bar{R}\in\bar{A})$. If the conclusion of Lemma 2.12 does not hold with respect to $B^{(i)},v^{(i)}$, then there exists a q-query decision tree P that distinguishes $R^{(i)}$ from $(R^{(i)}|R^{(i)}\in A)$, with advantage η, and as the

two distributions agree on the queries in $B^{(i)}$, we conclude that P distinguishes $\bar{R}$ from $\bar{Z}$ with the same advanatge. We apply Lemma 4.1 on $\bar{R}$ and $\bar{A}$, and conclude that there exists $Q \subseteq [N] \setminus B^{(i)}$ and $v \in \{0,1\}^Q$ such that

$$\Pr[\bar{R} \in \bar{A} | \bar{R}(Q) = v] > (1+\eta) \cdot \Pr[\bar{R} \in \bar{A}] \geq 2^{-a} \cdot (1+\eta)^{i+1}.$$

We set:

- $B^{(i+1)} = B^{(i)} \cup Q$.
- $v^{(i+1)}$ to be the "concatenation of $v^{(i)}$ and v". More precisely, for $y \in B^{(i)}$, $v_y^{(i+1)} = v_y^{(i)}$ and for $y \in Q$, $v_y^{(i+1)} = v_y$.
- $R^{(i+1)} = (R^{(i)} | R^{(i)}(Q) = v)$. (Note that by definition $B^{(i)} \cap Q = \emptyset$).

We now observe that the invariant is maintained in step $i+1$. Specifically:

- $|B^{(i+1)}| = |B^{(i)}| + q = i \cdot q + q = (i+1) \cdot q$.
- $\Pr[R^{(i+1)} \in A] = \Pr[R^{(i)} \in A | R^{(i)}(Q) = v] = Pr[\bar{R} \in \bar{A} | \bar{R}(Q) = v] \geq 2^{-a} \cdot (1+\eta)^{i+1}$.
- By definition, $R^{(i+1)}(B^{(i+1)}) = v^{(i+1)}$ with probability one.

Therefore, if this process fails to deliver the lemma after i steps, then the invariant is maintained at the end of step i, and in particular, $\Pr[R^{(i)} \in A] \geq 2^{-a} \cdot (1+\eta)^i$. However, this is impossible for $i > \frac{a}{\log(1+\eta)} = \Theta(a/\eta)$, and so, this process has to deliver the lemma within this number of steps. We obtain that the lemma follows with $b = |B| \leq O(\frac{qa}{\eta})$. □

5 Showing that the main theorem follows from the zoom lemma

In this section we show that Theorem 1.11 follows from Lemma 2.2. This follows by the earlier work of Shaltiel and Viola [Vio06, SV10] which we now explain.

5.1 Consequences distinguishing noise $\frac{1}{2}$ from $\frac{1}{2} - \epsilon$

The next lemma due to [Vio06, SV10] shows that distinguishing between $\mathsf{Noise}^q_{1/2-\epsilon}$ and $\mathsf{Noise}^q_{1/2}$ requires many queries.

Lemma 5.1. *[Vio06], see formulation in [SV10, Lemma 6.1] For every $\epsilon, \delta > 0$, such that $\delta < 0.4$, if $T : \{0,1\}^q \to \{0,1\}$ satisfies:*

- $\Pr[T(\mathsf{Noise}^q_{1/2-\epsilon}) = 1] \geq 1 - \delta$.
- $\Pr[T(\mathsf{Noise}^q_{1/2}) = 1] \leq 0.51$.

Then, $q = \Omega(\frac{\log \frac{1}{\delta}}{\epsilon^2})$.

The next lemma essentially shows that distinguishing between $\mathsf{Noise}^q_{1/2-\epsilon}$ and $\mathsf{Noise}^q_{1/2}$ requires majority on $\Omega(1/\epsilon)$ bits. A technicality is that for this conclusion, it is not sufficient to distinguish $\mathsf{Noise}^q_{1/2-\epsilon}$ from $\mathsf{Noise}^q_{1/2}$, and one needs circuits that distinguish $\mathsf{Noise}^q_{1/2-j\cdot\epsilon}$ from $\mathsf{Noise}^q_{1/2}$ for every $j \in [\frac{1}{2\epsilon}]$.[11]

[11] This complication is in some sense necessary, as the approach of [SV10] cannot give the conclusion of Lemma 5.2 under the simpler condition of Lemma 5.1. See discussion in [SV10].

Lemma 5.2 ([Vio06, SV10]). *For every $\epsilon, \delta > 0$, such that $\delta < 0.4$, and $\frac{1}{2\epsilon}$ is an integer, if $T_1, \ldots, T_{\frac{1}{2\epsilon}}$ are circuits over q bits, with size $s \geq q$ and depth d (over some set of gates G that includes the standard set $\{$AND,OR,NOT$\}$ with unbounded fan-in) and for every $j \in [\frac{1}{2\epsilon}]$, we have that:*

- $\Pr[T_j(\mathsf{Noise}^q_{1/2-j\cdot\epsilon}) = 1] \geq 1 - \delta$.
- $\Pr[T_j(\mathsf{Noise}^q_{1/2}) = 1] \leq 0.51$.

Then, there exists a circuit A that computes the majority function over $\frac{1}{\epsilon}$ bits, and A has size $\mathrm{poly}(s, \frac{1}{\epsilon})$ and depth $d + O(1)$ over the same set of gates G.

This lemma (in fact, with better parameters) follows from the argument of [SV10, Section 5]. For completeness, we provide a proof below:

Proof of Lemma 5.2. Let $\ell = \frac{1}{\epsilon}$, and note that ℓ is even. For $x \in \{0,1\}^\ell$, let $\mathsf{wt}(x)$ denote the fraction of ones in x. Our goal is to compute the majority function on length ℓ (that is to distinguish $x \in \{0,1\}^\ell$ with $\mathsf{wt}(x) \geq 1/2$ from $x \in \{0,1\}^\ell$ with $\mathsf{wt}(x) < 1/2$.

For every $x \in \{0,1\}^\ell$, consider the experiment in which we choose $i_1, \ldots, i_q \leftarrow [\ell]$ independently, and construct the q bit string $x_{i_1} \circ, \ldots, \circ x_{i_q}$. Note that this string is distributed according to $\mathsf{Noise}^q_{\mathsf{wt}(x)}$ and that given x, this distribution can be prepared by a (distribution over) constant depth circuits of size $O(q) = O(s)$.

By composing this with T_j, we obtain that for every $j \in [\frac{1}{2\epsilon}]$ there is a (distribution over) circuits depth $d + O(1)$ and size $\mathrm{poly}(s, \ell)$ that distinguishes strings $x \in \{0,1\}^\ell$ with $\mathsf{wt}(x) = 1/2$ from strings $x \in \{0,1\}^\ell$ with $\mathsf{wt}(x) = 1/2 - j \cdot \epsilon$.

It then follows that for every $j \in [\frac{1}{2\epsilon}]$, this distribution can be replaced with a single deterministic depth $d + O(1)$ circuit of size $\mathrm{poly}(s, \ell)$ that performs the same task.[12]

If we take the "AND" of the $\frac{1}{2\epsilon}$ circuits, we obtain a depth $d + O(1)$, size $\mathrm{poly}(s, \ell)$ A circuit that distinguishes $x \in \{0,1\}^\ell$ with $\mathsf{wt}(x) < 1/2$, from $x \in \{0,1\}^\ell$ from $x \in \{0,1\}^\ell$ with $\mathsf{wt}(x) = \frac{1}{2}$.

Finally, in order to compute majority, for every $0 \leq i \leq \frac{\ell}{2}$, we consider the circuit $A_i(x)$ which replaces the first i bits of x with zeros, and applies A_i. It is immediate that if we take the "OR" of all A_i for $0 \leq i \leq \frac{\ell}{2}$, we obtain a depth $d + O(1)$, size $\mathrm{poly}(s, \ell)$ circuit that distinguishes $x \in \{0,1\}^\ell$ with $\mathsf{wt}(x) \geq 1/2$ from $x \in \{0,1\}^\ell$ with $\mathsf{wt}(x) < 1/2$, and computes the majority function on ℓ bits. □

5.2 Finishing up

We now prove that Theorem 1.11 follows from Lemma 2.2. As the consequences of Theorem 1.11 are stated in terms of $\Omega(\frac{1}{\epsilon})$, we are allowed to modify ϵ by a constant factor. Consequently, w.l.o.g. we can assume that $\frac{1}{\epsilon}$ is a power of two. Moreover, we set $\epsilon' = \epsilon/2$ (and note that $1/\epsilon'$ is also a power of two) and w.l.o.g. we are allowed to assume that we have a $(\frac{1}{2} + \epsilon') \to (1 - \delta)$ class $\mathcal{D}$ reduction for Yao's XOR lemma.

We first observe that a $(\frac{1}{2} + \epsilon') \to (1 - \delta)$ class $\mathcal{D}$ reduction for Yao's XOR lemma, is in particular a $(\frac{1}{2} + j \cdot \epsilon') \to (1 - \delta)$ class $\mathcal{D}$ reduction for Yao's XOR lemma, for every $j \in [\frac{1}{2\epsilon'}]$.

This means that for every integer $j \in [\frac{1}{2\epsilon'}]$, we can apply Lemma 2.2 choosing the parameter ϵ of the lemma to be $j \cdot \epsilon'$. Note that as $\frac{1}{\epsilon'}$ is a power of two, for every $j \in [\frac{1}{2\epsilon'}]$, $j\epsilon'$ can be expressed as $\frac{a}{b}$ where

[12] This follows as we can reduce the error to $2^{-2\ell}$, by taking $O(\ell)$ independent copies from the distribution, computing the *approximate majority* of the outputs, and taking a union bound over the 2^ℓ inputs (a-la Adleman's proof that BPP is in P/poly). The point is that this can be done in constant depth because the gap between $1 - \delta$ and 0.51 is a constant, and approximate majority to within a constant can be computed in AC^0 by the celebrated result of Ajtai [Ajt83]. (In fact, it is sufficient that the gap between the acceptance and rejection probability of T_j is $\Omega(\frac{1}{\log \ell})$ to perform this amplification, and the lemma follows also with such a gap).

$b = \frac{1}{\epsilon'}$, and we meet the conditions of Lemma 2.2. We conclude that for every $j \in [\frac{1}{2\epsilon'}]$, and $x \in \{0,1\}^k$ there exists a circuit T_x^j over q bits with size $\mathrm{poly}(r)$ and depth $O(d)$ over the set of gates G such that:

- $\Pr_{X \leftarrow U_k}[T_X^j(\mathsf{Noise}^q_{1/2-2j\epsilon'}) = 1] \geq 1 - 2\delta$.
- $\Pr_{X \leftarrow U_k}[T_X^j(\mathsf{Noise}^q_{1/2}) = 1] \leq \frac{1}{2} + \frac{1}{200} = 0.505$.

Applying Markov's inequality to the second item, we obtain that there exists a constant $\beta > 0$ such that for every j, for a β fraction of $x \in \{0,1\}^k$,

$$\Pr[T_x^j(\mathsf{Noise}^q_{1/2}) = 1] \leq 0.51.$$

Applying Markov's inequality to the first item, we obtain that for every $j \in [\frac{2}{\epsilon'}]$, for a $1 - \beta/2$ fraction of $x \in \{0,1\}^k$,

$$\Pr[T_x^j(\mathsf{Noise}^q_{1/2-2j\epsilon'}) = 1] \geq 1 - \frac{2\delta}{\beta/2} = 1 - \frac{4 \cdot \delta}{\beta} \geq 1 - \frac{4 \cdot \delta_0}{\beta},$$

where the last inequality is because in Theorem 1.11 there is an assumption that $\delta \leq \delta_0$ for some δ_0 (and note that at this point we have not yet chosen δ_0).

Recalling that $\epsilon' = \epsilon/2$, this can be rephrased as saying that: for every $j \in [\frac{2}{\epsilon}]$:

- For a $1 - \beta/2$ fraction of $x \in \{0,1\}^k$,

$$\Pr[T_x^j(\mathsf{Noise}^q_{1/2-j\epsilon}) = 1] \geq 1 - \frac{4 \cdot \delta_0}{\beta}.$$

- For a β fraction of $x \in \{0,1\}^k$,

$$\Pr[T_x^j(\mathsf{Noise}^q_{1/2}) = 1] \leq 0.51.$$

Together, these two consequences give that for every $j \in [\frac{2}{\epsilon}]$, there exists $x \in \{0,1\}^k$ that satisfies both inequalities. Theorem 1.11 now follows directly from Lemma 5.1 and Lemma 5.2, by choosing the constant $\delta_0 > 0$ to be sufficiently small so that $\frac{4 \cdot \delta_0}{\beta} < 0.4$ as required in the two lemmas.[13]

6 Extending the argument to sufficiently explicit linear codes

We now prove Theorem 1.14. Recall that this is an extension of Theorem 1.11 that assumes that $\delta = 2^{-2k}$ and uses a construction map $f' = \mathrm{Con}^f$ where instead of $f' = f^{\oplus t}$ the target function f' that is used is a function $f' : \{0,1\}^n \to \{0,1\}$ defined by:

$$f'(y) = \sum_{x \in \{0,1\}^k} f(x) \cdot g(x,y),$$

where the sum is taken in the field $\mathbb{F}_2$, and $g : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}$ can be computed by circuits of size $\mathrm{poly}(r)$ and depth d over the set G of gates.

[13] The Markov argument above is wasteful, and leads to a rather small constant $\delta_0 > 0$. We remark that with a more careful argument (that was used in [SV10]) we could have chosen δ_0 to be any constant smaller than $\frac{1}{2}$, and even allow it to approach $\frac{1}{2}$. We chose not to do this in order to make the proof simpler and more modular.

This argument is based on a trick by Viola [Vio06] that we can incorporate into our framework. We will modify the proof of Lemma 2.2 so that it holds in this setting, the modified version of Theorem 1.11 will follow from Lemma 2.2 just as before.

We start by replacing the function f of Lemma 2.4 with a sightly different function:

Lemma 6.1. *There exist constants c_1 such that for every constant c_2, there exists a function $f : \{0,1\}^k \to \{0,1\}$ such that there exists $m \leq c_1 \cdot c_2 \cdot \log r$ such that:*

- *For every circuit $B : \{0,1\}^k \to \{0,1\}$ of size r^{c_2}, $\Pr_{X \leftarrow U_m}[B(X \circ 0^{k-m}) = f(X \circ 0^{k-m})] \leq \frac{1}{2} + \frac{1}{200}$.*
- *f can be computed by a DNF of size $r^{c_1 \cdot c_2}$.*

Proof. We repeat the proof of Lemma 2.4, but this time we take $f(x)$ to be the following function: Let x' denote the first m bits of x and x'' denote the remaining $k-m$ bits. We define $f(x)$ to be $g(x')$ if $x'' = 0^{k-m}$ and zero otherwise. □

On a random $X \leftarrow U_k$, $\Pr[f(X) = 0] \geq 1 - 2^{k-m} = 1 - 2^{-\Omega(k)}$. Therefore, it is very easy to compute f with success probability $1 - 2^{-\Omega(k)}$ by simply answering zero. However, by Lemma 6.1 it is hard for circuits of size r^{c_2} to compute f with success probability 1, or equivalently success probability $1 - \delta$ for $\delta = 2^{-2k}$. This is why this approach can only succeed for very small δ.

With this choice, we can get a corollary that is analogous to Corollary 2.5.

Corollary 6.2. *The function f' can be computed by circuits of size $\mathrm{poly}(r)$ and constant depth over the gates $\{$AND,OR,NOT,PARITY$\}$ of unbounded fan-in.*

Proof. The function f' is defined by:

$$f'(y) = \sum_{x \in \{0,1\}^k} f(x) \cdot g(x,y).$$

The sum ranges over 2^k choices of x. However, for our function f, except for $\mathrm{poly}(r)$ of these x (the ones for which the second part of x is $k-m$ zeros) all the remaining x have $f(x) = 0$. This, together with the requirement on g, gives the required result. □

The proof proceeds as in Section 2, with the following modifications:

- For $\delta = 2^{-2k} < 2^{-k}$, Red is a $(\frac{1}{2} + \epsilon) \to 1$ reduction. This means in particular that if D is useful, then there exists $\alpha \in \{0,1\}^a$ such that $\Pr_{X \leftarrow U_m}[\mathrm{Red}^D(X \circ 0^{k-m}, \alpha) = f(X \circ 0^{k-m})] = 1$.
- We set $\delta' = \frac{1}{r}$ and will replace some occurrences of δ in the earlier argument by δ'. This is done because the choice of $\delta = 2^{-2k}$ does not satisfy the requirement that $\frac{1}{\delta} \leq r$ made in Theorem 1.11. Specifically, the requirement that $\frac{1}{\delta} \leq r$ was used to argue that when we apply Lemma 2.12 with $\eta = \delta$, the size of the set B (which is polynomial in $\frac{1}{\eta}$) is polynomial in r. In order to obtain a set B of size $\mathrm{poly}(r)$ we will now choose $\eta = \delta'$.
- In Section 2.4 we argued that for an independent $X \leftarrow U_k$:

  $$\Pr[\mathrm{Red}^{f^{\oplus t} \oplus R'}(X, \alpha') = f(X)] \geq \Pr[\mathrm{Red}^{f^{\oplus t} \oplus Z'}(X, \alpha') = f(X)] - \delta.$$

  With our modifications we get that for an independent $X \leftarrow U_m$:

  $$\Pr[\mathrm{Red}^{f' \oplus R'}(X \circ 0^{k-m}, \alpha') = f(X \circ 0^{k-m})] \geq \Pr[\mathrm{Red}^{f^{\oplus t} \oplus Z'}(X \circ 0^{k-m}, \alpha') = f(X \circ 0^{k-m})] - \delta'.$$

- This allows us to continue the argument, replacing occurrences of $X \leftarrow U_k$ by $X \circ 0^{k-m}$ for $X \leftarrow U_m$, and occurrences of δ by δ'.
- When we finish the proof and obtain a lower bound of $q = \Omega(\frac{\log(1/\delta')}{\epsilon^2}) = \Omega(\frac{\log r}{\epsilon^2})$ as required. The result on majority is not affected by replacing δ by δ'.

7 Conclusion and open problems

Class reductions are known to bypass some limitations on black-box reductions (as explained in Section 1.2). This work demonstrates that it is sometimes possible to extend limitations on black-box reductions to some class reductions. Studying the power of class reductions may promote our understanding of how to bypass limitations on black-box reductions. We now mention some more specific open problems:

- Unlike the results of [GSV18], our results do not hold for any construction of target functions f' from f. Is it possible to extend our results to this general setting?
- In Theorem 1.11, the class $\mathcal{D}$ contains circuits that are polynomially larger than the size of the reduction. Is it possible to extend our limitations on class reductions with respect to a classes $\mathcal{D}$ of circuits smaller than the circuit size of the reduction?
- Yao's XOR lemma states that *for every* function f, if f is somewhat hard, then $f^{\oplus t}$ is very hard. It makes sense to focus on some specific choice for a somewhat hard function f and prove and improved result for this specific function. If we prove such an assertion by reduction, we can allow the reduction to be tailored to the specific function f, and do no need to require that the reduction performs on any function f, but only on the chosen one. This type of reductions was termed "function specific" by Artemenko and Shaltiel [AS14], who proved limitations on nonuniform black-box functions specific reductions. It is interesting to understand whether function specific class reductions can circumvent the limitations proven in this paper. We remark that our proof technique indeed relies on the fact that the reduction is not function specific, and must work for *any* function f. This allows us to choose f with specific properties that are useful for our argument.

Acknowledgement

A preliminary version of this paper appeared in RANDOM 2020 [Sha20]. We are grateful to Emanuele Viola for very helpful discussions, and to anonymous referees for excellent comments and suggestions.

References

[AASY16] Benny Applebaum, Sergei Artemenko, Ronen Shaltiel, and Guang Yang. Incompressible functions, relative-error extractors, and the power of nondeterministic reductions. *Computational Complexity*, 25(2):349–418, 2016.

[AGGM06] Adi Akavia, Oded Goldreich, Shafi Goldwasser, and Dana Moshkovitz. On basing one-way functions on NP-hardness. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, pages 701–710, 2006.

[Ajt83] Miklós Ajtai. $\sum^1_1$-formulae on finite structures. *Ann. Pure Appl. Log.*, 24(1):1–48, 1983.

[AS14] Sergei Artemenko and Ronen Shaltiel. Lower bounds on the query complexity of non-uniform and adaptive reductions showing hardness amplification. *Computational Complexity*, 23(1):43–83, 2014.

[Ats06] Albert Atserias. Distinguishing SAT from polynomial-size circuits, through black-box queries. In *21st Annual IEEE Conference on Computational Complexity*, pages 88–95, 2006.

[BT06] Andrej Bogdanov and Luca Trevisan. On worst-case to average-case reductions for NP problems. *SIAM J. Comput.*, 36(4):1119–1159, 2006.

[FF93] Joan Feigenbaum and Lance Fortnow. Random-self-reducibility of complete sets. *SIAM J. Comput.*, 22(5):994–1005, 1993.

[FSUV13] Bill Fefferman, Ronen Shaltiel, Christopher Umans, and Emanuele Viola. On beating the hybrid argument. *Theory of Computing*, 9:809–843, 2013.

[GK96] Oded Goldreich and Hugo Krawczyk. On the composition of zero-knowledge proof systems. *SIAM J. Comput.*, 25(1):169–192, 1996.

[GNW11] Oded Goldreich, Noam Nisan, and Avi Wigderson. On yao's XOR-lemma. In Oded Goldreich, editor, *Studies in Complexity and Cryptography. Miscellanea on the Interplay between Randomness and Computation*, volume 6650 of *Lecture Notes in Computer Science*, pages 273–301. Springer, 2011.

[GR08] Dan Gutfreund and Guy N. Rothblum. The complexity of local list decoding. In *Approximation, Randomization and Combinatorial Optimization. Algorithms and Techniques, 11th International Workshop, APPROX, and 12th International Workshop, RANDOM*, volume 5171 of *Lecture Notes in Computer Science*, pages 455–468, 2008.

[GST07] Dan Gutfreund, Ronen Shaltiel, and Amnon Ta-Shma. If NP languages are hard on the worst-case, then it is easy to find their hard instances. *Computational Complexity*, 16(4):412–441, 2007.

[GSV18] Aryeh Grinberg, Ronen Shaltiel, and Emanuele Viola. Indistinguishability by adaptive procedures with advice, and lower bounds on hardness amplification proofs. In *59th IEEE Annual Symposium on Foundations of Computer Science*, pages 956–966, 2018.

[GT07] Dan Gutfreund and Amnon Ta-Shma. Worst-case to average-case reductions revisited. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, 10th International Workshop, APPROX, and 11th International Workshop, RANDOM*, volume 4627 of *Lecture Notes in Computer Science*, pages 569–583, 2007.

[Gut06] Dan Gutfreund. Worst-case vs. algorithmic average-case complexity in the polynomial-time hierarchy. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, 9th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, APPROX and 10th International Workshop on Randomization and Computation, RANDOM*, volume 4110 of *Lecture Notes in Computer Science*, pages 386–397, 2006.

[GV04] Dan Gutfreund and Emanuele Viola. Fooling parity tests with parity gates. In *Approximation, Randomization, and Combinatorial Optimization, Algorithms and Techniques, 7th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, APPROX, and 8th International Workshop on Randomization and Computation, RANDOM*, volume 3122 of *Lecture Notes in Computer Science*, pages 381–392, 2004.

[GV08] Dan Gutfreund and Salil P. Vadhan. Limitations of hardness vs. randomness under uniform reductions. In *Approximation, Randomization and Combinatorial Optimization. Algorithms and Techniques, 11th International Workshop, APPROX, and 12th International Workshop, RANDOM*, volume 5171 of *Lecture Notes in Computer Science*, pages 469–482, 2008.

[Hir18] Shuichi Hirahara. Non-black-box worst-case to average-case reductions within NP. In *59th IEEE Annual Symposium on Foundations of Computer Science*, pages 247–258, 2018.

[Imp95] Russell Impagliazzo. Hard-core distributions for somewhat hard problems. In *36th Annual Symposium on Foundations of Computer Science*, pages 538–545, 1995.

[IR89] Russell Impagliazzo and Steven Rudich. Limits on the provable consequences of one-way permutations. In *Proceedings of the 21st Annual ACM Symposium on Theory of Computing*, pages 44–61, 1989.

[IW97] Russell Impagliazzo and Avi Wigderson. $P = BPP$ if E requires exponential circuits: Derandomizing the XOR lemma. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on the Theory of Computing*, pages 220–229, 1997.

[KS03] Adam R. Klivans and Rocco A. Servedio. Boosting and hard-core set construction. *Machine Learning*, 51(3):217–238, 2003.

[LSS+19] Nutan Limaye, Karteek Sreenivasaiah, Srikanth Srinivasan, Utkarsh Tripathi, and S. Venkitesh. A fixed-depth size-hierarchy theorem for $\mathrm{AC}^0[\oplus]$ via the coin problem. In *Proceedings of the 51st Annual ACM Symposium on Theory of Computing*, pages 442–453, 2019.

[LTW08] Chi-Jen Lu, Shi-Chun Tsai, and Hsin-Lung Wu. On the complexity of hardness amplification. *IEEE Trans. Information Theory*, 54(10):4575–4586, 2008.

[OSS19] Igor Carboni Oliveira, Rahul Santhanam, and Srikanth Srinivasan. Parity helps to compute majority. In *34th Computational Complexity Conference*, volume 137, pages 23:1–23:17, 2019.

[Raz87] Alexander Razborov. Lower bounds on the dimension of schemes of bounded depth in a complete basis containing the logical addition function. *Akademiya Nauk SSSR. Matematicheskie Zametki*, 41(4):598–607, 1987. English translation in Mathematical Notes of the Academy of Sci. of the USSR, 41(4):333-338, 1987.

[RTV04] Omer Reingold, Luca Trevisan, and Salil P. Vadhan. Notions of reducibility between cryptographic primitives. In *Theory of Cryptography, First Theory of Cryptography Conference, TCC*, volume 2951 of *Lecture Notes in Computer Science*, pages 1–20, 2004.

[Sha20] Ronen Shaltiel. Is it possible to improve Yao's XOR lemma using reductions that exploit the efficiency of their oracle? In Jaroslaw Byrka and Raghu Meka, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM*

2020, August 17-19, 2020, Virtual Conference, volume 176 of *LIPIcs*, pages 10:1–10:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020.

[Smo87] Roman Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity. In *Proceedings of the 19th Annual ACM Symposium on Theory of Computing*, pages 77–82, 1987.

[STV01] Madhu Sudan, Luca Trevisan, and Salil P. Vadhan. Pseudorandom generators without the XOR lemma. *J. Comput. Syst. Sci.*, 62(2):236–266, 2001.

[SV10] Ronen Shaltiel and Emanuele Viola. Hardness amplification proofs require majority. *SIAM J. Comput.*, 39(7):3122–3154, 2010.

[TV07] Luca Trevisan and Salil P. Vadhan. Pseudorandomness and average-case complexity via uniform reductions. *Computational Complexity*, 16(4):331–364, 2007.

[Vio03] Emanuele Viola. Hardness vs. randomness within alternating time. In *18th Annual IEEE Conference on Computational Complexity*, page 53, 2003.

[Vio06] Emanuele Viola. *The Complexity of Hardness Amplification and Derandomization*. PhD thesis, Harvard University, 2006.