Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 126 (2010)

Query Complexity in Errorless Hardness Amplification

Thomas Watson*

December 5, 2010

Abstract

An errorless circuit for a boolean function is one that outputs the correct answer or "don't know" on each input (and never outputs the wrong answer). The goal of errorless hardness amplification is to show that if f has no size s errorless circuit that outputs "don't know" on at most a δ fraction of inputs, then some f' related to f has no size s' errorless circuit that outputs "don't know" on at most a $1-\epsilon$ fraction of inputs. Thus the hardness is "amplified" from δ to $1-\epsilon$. Unfortunately, this amplification comes at the cost of a loss in circuit size. This is because such results are proven by reductions which show that any size s' errorless circuit for f' that outputs "don't know" on at most a $1-\epsilon$ fraction of inputs could be used to construct a size s errorless circuit for f that outputs "don't know" on at most a δ fraction of inputs. If the reduction makes q queries to the hypothesized errorless circuit for f', then plugging in a size s' circuit yields a circuit of size $\geq qs'$, and thus we must have $s' \leq s/q$. Hence it is desirable to keep the query complexity to a minimum.

The first results on errorless hardness amplification were obtained by Bogdanov and Safra. They achieved query complexity $O\big((\frac{1}{\delta}\log\frac{1}{\epsilon})^2 \cdot \frac{1}{\epsilon}\log\frac{1}{\delta}\big)$ when f' is the XOR of several independent copies of f. We improve the query complexity (and hence the loss in circuit size) to $O\big(\frac{1}{\epsilon}\log\frac{1}{\delta}\big)$, which is optimal up to constant factors for nonadaptive black-box errorless hardness amplification.

Bogdanov and Safra also proved a result that allows for errorless hardness amplification within NP. They achieved query complexity $O\big(k^3 \cdot \frac{1}{\epsilon^2}\log\frac{1}{\delta}\big)$ when f' consists of any monotone function applied to the outputs of k independent copies of f, provided the monotone function satisfies a certain combinatorial property parameterized by δ and ϵ. We improve the query complexity to $O\big(\frac{k}{t} \cdot \frac{1}{\epsilon}\log\frac{1}{\delta}\big)$, where $t \geq 1$ is a certain parameter of the monotone function.

As a side result, we prove a lower bound on the advice complexity of black-box reductions for errorless hardness amplification.

1 Introduction

Traditionally, an algorithm for solving a computational problem is required to be correct on all inputs and is judged in terms of its efficiency (the amount of computational resources it uses). One criticism of this model is that it is too strict: In practice, an algorithm only needs to be correct on "real-world" inputs and not on contrived worst-case inputs. To address this issue within the framework of complexity theory, researchers developed the theory of average-case complexity (starting with the work of Levin [13]). In this theory, an algorithm is judged in terms of both its

*Computer Science Division, University of California, Berkeley. Supported by a National Science Foundation Graduate Research Fellowship.

ISSN 1433-8092

efficiency and the fraction of inputs on which it fails to solve the problem correctly. The topic of this paper is the relationship between these two measures of the quality of an algorithm.

There are two standard settings for average-case complexity. In the original setting proposed by Levin [13], one only considers *errorless algorithms*, which are required to output the correct answer or "don't know" on each input.[1] An errorless algorithm is judged in terms of both its efficiency and the fraction of inputs on which it outputs "don't know". We refer to this setting as *errorless average-case complexity*. In the other setting, one considers arbitrary algorithms which may output the wrong answer rather than just "don't know" on an input. We refer to this setting as *non-errorless average-case complexity*. Errorless average-case complexity is an intermediate setting between worst-case complexity and non-errorless average-case complexity.

We first discuss non-errorless average-case complexity. A boolean function is said to be *mildly average-case hard* if no efficient algorithm can compute it on almost all inputs. Applications such as derandomization and cryptography require functions that are *strongly average-case hard*, meaning that no efficient algorithm can compute the function on noticeably more than half the inputs. This motivates hardness amplification, which is the problem of transforming a mildly average-case hard function into a strongly average-case hard function. A classic result in this area is the XOR Lemma [14, 9, 5, 11], which states that the XOR of sufficiently many independent copies of a mildly average-case hard function is strongly average-case hard, provided the model of efficient algorithms is small circuits.

However, the XOR Lemma (as well as the numerous subsequent results on hardness amplification) incurs an unfortunate loss in circuit size. Suppose the original function f is mildly average-case hard in the sense that no size s circuit succeeds on at least a $1-\delta$ fraction of inputs, and we wish for the new function f' to be strongly average-case hard in the sense that no size s' circuit succeeds on at least a $1/2+\epsilon$ fraction of inputs. Then we would like s' to be as large as possible, but the XOR Lemma requires that s' is actually smaller than s. This is because such results are proven by reductions which show that if f' is not strongly average-case hard, then a circuit witnessing this could be used to construct a circuit witnessing that f is not mildly average-case hard. If the reduction makes q queries to the hypothesized circuit, then plugging in a size s' circuit yields a circuit of size $\geq qs'$, and thus we must have $s' \leq s/q$. Hence the query complexity q governs the loss in circuit size. For the XOR Lemma, the query complexity is well-understood. The proof due to Impagliazzo [9] and Klivans and Servedio [12] shows that $q = O\left(\frac{1}{\epsilon^2}\log\frac{1}{\delta}\right)$ queries are sufficient, and Shaltiel and Viola [18] showed that in a certain sense, $q = \Omega\left(\frac{1}{\epsilon^2}\log\frac{1}{\delta}\right)$ queries are necessary.

Bogdanov and Safra [1] initiated the study of hardness amplification in Levin's original setting of errorless average-case complexity. A boolean function is said to be *mildly errorless average-case hard* if no efficient errorless algorithm (say, size s circuit) can compute it on almost all inputs (say, a $1-\delta$ fraction). A function is said to be *strongly errorless average-case hard* if no efficient errorless algorithm (say, size s' circuit) can compute it on a noticeable fraction of inputs (say, an ϵ fraction). Note that in the non-errorless setting, computing a boolean function on half the inputs is trivial (using constant 0 or constant 1), but in the errorless setting, computing a boolean function on even a small fraction of inputs is nontrivial. The goal of errorless hardness amplification is to transform a mildly errorless average-case hard function f into a strongly errorless average-case hard function f'. Such results suffer from a loss in circuit size for the same reason as in the non-errorless setting.

[1] Actually, Levin proposed considering algorithms that are correct on all inputs but which are efficient "on average" with respect to a random input. Under a suitable formalization, such algorithms are equivalent to errorless algorithms that may fail on a small fraction of inputs but are efficient on all inputs.

Bogdanov and Safra [1] showed that $q = O\big((\frac{1}{\delta}\log\frac{1}{\epsilon})^2 \cdot \frac{1}{\epsilon}\log\frac{1}{\delta}\big)$ queries are sufficient when f' is the XOR of several independent copies of f. The result of Shaltiel and Viola [18] can be modified without difficulty to show that in a certain sense, $q = \Omega\big(\frac{1}{\epsilon}\log\frac{1}{\delta}\big)$ queries are necessary. We close the gap by showing that $q = O\big(\frac{1}{\epsilon}\log\frac{1}{\delta}\big)$ queries are sufficient.

Another natural goal for hardness amplification is to guarantee that if f represents an NP language at some input length, then f' also represents an NP language at some input length. In the non-errorless setting this goal has been studied in numerous works [17, 19, 8, 20, 2, 15, 4], and in the errorless setting this goal has been studied by Bogdanov and Safra [1]. We significantly improve the query complexity of the Bogdanov-Safra result.

1.1 The Errorless XOR Lemma

Given $f : \{0,1\}^n \to \{0,1\}$ we define $f^{\oplus k} : \{0,1\}^{n\times k} \to \{0,1\}$ as follows: $f^{\oplus k}(x_1,\ldots,x_k) = f(x_1) \oplus \cdots \oplus f(x_k)$.

Definition 1 (Errorless Average-Case Hardness). *We say a circuit $A : \{0,1\}^n \to \{0,1,\perp\}$ is a δ-errorless circuit for $f : \{0,1\}^n \to \{0,1\}$ if*

(i) $A(x) \in \{f(x), \perp\}$ for all $x \in \{0,1\}^n$, and

(ii) $\Pr_x[A(x) = \perp] \le \delta$ where $x \in \{0,1\}^n$ is chosen uniformly at random.

We say f is (s,δ)-hard if it has no δ-errorless circuit of size $\le s$.

Theorem 1 (Query-Optimal Errorless XOR Lemma). *If f is (s,δ)-hard then $f' = f^{\oplus k}$ is $(s', 1-\epsilon)$-hard where $s' = s/\big(\frac{4}{\epsilon}\ln\frac{2}{\delta}\big)$, provided $k \ge \frac{16}{\delta}\ln\frac{2}{\epsilon}$.*

We prove Theorem 1 in Section 2.[2] Bogdanov and Safra [1] proved a version of Theorem 1 where $s' = s/\big(k^2 \cdot \frac{2}{\epsilon}\ln\frac{2}{\delta}\big)$, provided $k \ge \frac{2}{\delta}\ln\frac{2}{\epsilon}$.[3] Even for the best value of k, they only achieve $O\big((\frac{1}{\delta}\log\frac{1}{\epsilon})^2 \cdot \frac{1}{\epsilon}\log\frac{1}{\delta}\big)$ query complexity. Also, our bound on the query complexity does not depend on k.

We prove Theorem 1 by a reduction similar to the one used in [1]. Our contribution is a new, tight analysis of the reduction. The crux of the reduction is a randomized procedure that solves f errorlessly (meaning that for each input x it may output $f(x)$ with some probability and $\perp$ with some probability, but it never outputs $\overline{f(x)}$) while making one query to a hypothesized $(1-\epsilon)$-errorless circuit A' for f'. Suppose for some $\beta > 0$ we knew that $\le \delta/2$ fraction of inputs x are bad in the sense that the probability the procedure outputs $f(x)$ is $< \beta$. Then by amplifying the success probability on the good inputs and hard-wiring the randomness appropriately, we obtain a δ-errorless circuit A for f, via a reduction with query complexity $O\big(\frac{1}{\beta}\log\frac{1}{\delta}\big)$. The heart of our improvement over the Bogdanov-Safra proof is in arguing that we can take $\beta = \epsilon/4$. To prove this, we suppose the fraction of bad inputs is $> \delta/2$ and prove that then A' must compute f' on $< \epsilon$ fraction of inputs. The procedure outputs $f(x)$ if and only if the query is an input on which A' computes f'; furthermore the distribution of this query $(x_1,\ldots,x_k)$ is obtained by setting $x_i = x$

[2] In the statement of Theorem 1, it would be more accurate to say $s' = s/\big(\frac{4}{\epsilon}\ln\frac{2}{\delta}\big) - O(1)$ to account for the trivial circuitry needed to combine the results of the $\frac{4}{\epsilon}\ln\frac{2}{\delta}$ queries the reduction makes. Throughout this paper, we ignore such details. We also ignore details arising from the fact that numbers such as $\frac{4}{\epsilon}\ln\frac{2}{\delta}$ might not be integers.

[3] Actually, their proof gives $s' = s/\big(k^2 \cdot \frac{2}{\epsilon}n\big)$, provided $k \ge \frac{1}{\delta}\ln\frac{2}{\epsilon}$, but a minor tweak to their proof yields the stated result.

for a uniformly random i and picking $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k$ uniformly at random. Consider the following two distributions on queries to A': the uniform distribution, and the distribution obtained by picking a random bad x and running the procedure on input x. We know A' computes f' with probability $< \beta = \epsilon/4$ under the latter distribution, and we wish to show that A' computes f' with probability $< \epsilon$ under the former. For this, we show the two distributions are "close" in the sense that the probability of any event under the former is less than twice the probability under the latter plus $\epsilon/2$. The argument involves a dichotomy: Since we assume a large fraction of x's are bad, a uniform query is unlikely to have few bad coordinates. Assuming there are many bad coordinates, we can essentially pretend there is one bad coordinate and then argue that we have overcounted the probability by a lot. This is the intuition for the ideas behind the proof of Theorem 1.

It can be shown that $\Omega\big(\frac{1}{\epsilon}\log\frac{1}{\delta}\big)$ queries are needed by any nonadaptive black-box reduction achieving errorless hardness amplification, regardless of how f' is constructed from f (see Section 1.3 for the precise statement). Since our proof of Theorem 1 (and the Bogdanov-Safra proof) is by a nonadaptive black-box reduction, this shows that Theorem 1 is optimal in a sense. Shaltiel and Viola [18] gave a general technique for lower bounding the query complexity of nonadaptive black-box reductions, and they noted that their technique applies to non-errorless hardness amplification (including the XOR Lemma and the Direct Product Lemma) and to constructions of pseudorandom generators from average-case hard functions. Similarly, we observe that their technique applies to errorless hardness amplification.

1.2 Monotone Errorless Amplification

Consider the problem of errorless hardness amplification within NP. That is, if f is computable in nondeterministic polynomial time, then we want f' to also be computable in nondeterministic polynomial time. Taking $f' = f^{\oplus k}$ does not guarantee this. We instead consider more general constructions of the form $f' = C \circ f^k$ where $C : \{0,1\}^k \to \{0,1\}$, and $f^k : \{0,1\}^{n\times k} \to \{0,1\}^k$ is defined as $f^k(x_1, \ldots, x_k) = \big(f(x_1), \ldots, f(x_k)\big)$. In the setting of the XOR Lemma, the combiner function C is the k-bit parity function. If C is monotone (that is, $C(y_1, \ldots, y_k) \le C(z_1, \ldots, z_k)$ whenever $y_i \le z_i$ for all $i \in [k]$) and f and C are both computable in nondeterministic polynomial time, then f' is guaranteed to be computable in nondeterministic polynomial time. This approach dates back to [17, 8].

Bogdanov and Safra [1] showed that this construction yields errorless hardness amplification provided the monotone combiner function C satisfies a certain combinatorial property. To describe this property, we need some definitions from [1] (though we use somewhat different notation). Fix $b \in \{0,1\}$. Given a monotone function $C : \{0,1\}^k \to \{0,1\}$ and a string $y \in \{0,1\}^k$, we say that coordinate $i \in [k]$ is b-*sensitive* if flipping the ith bit of y causes the value of C to flip from b to $\overline{b}$, and we let $\sigma(C, y, b)$ denote the set of b-sensitive coordinates. That is,

$$\sigma(C, y, b) \;=\; \big\{ i \in [k] \;:\; C(y) = b \text{ and } C(y \oplus e_i) = \overline{b} \big\}.$$

Note that if $C(y) = \overline{b}$ then $\sigma(C, y, b) = \emptyset$ and if $C(y) = b$ then by the monotonicity of C, $\sigma(C, y, b)$ only contains coordinates i such that $y_i = b$. For $p \in [0,1]$, we use $y \sim_p \{0,1\}^k$ to denote that y is sampled from the p-biased distribution, that is, each bit is independently set to 1 with probability p.

Definition 2. *For $b \in \{0,1\}$, a function $C : \{0,1\}^k \to \{0,1\}$ is a (t,ρ,p,b)-amplifier if C is monotone and*

$$\Pr_{y \sim_p \{0,1\}^k}\Big[\big|\sigma(C,y,b)\big| \geq t\Big] \;\geq\; 1-\rho.$$

Note that a monotone function $C : \{0,1\}^k \to \{0,1\}$ is a (t,ρ,p,b)-amplifier if and only if its monotone complement $C^\dagger : \{0,1\}^k \to \{0,1\}$ is a $(t,\rho,1-p,\overline{b})$-amplifier, where $C^\dagger$ is defined as

$$C^\dagger(y_1,\ldots,y_k) \;=\; \overline{C(\overline{y_1},\ldots,\overline{y_k})}.$$

For reasons discussed in [1], it is necessary to consider the following one-sided version of Definition 1.

Definition 3 (One-Sided Errorless Average-Case Hardness). *For $b \in \{0,1\}$, we say a circuit $A : \{0,1\}^n \to \{0,1,\perp\}$ is a (δ,b)-errorless circuit for $f : \{0,1\}^n \to \{0,1\}$ if*

(i) $A(x) \in \{f(x),\perp\}$ for all $x \in \{0,1\}^n$, and

(ii) $\Pr_x[A(x) = \perp] \leq \delta$ where $x \in f^{-1}(b)$ is chosen uniformly at random.

We say f is (s,δ,b)-hard if it has no (δ,b)-errorless circuit of size $\leq s$.

Note that if f is (s,δ)-hard then f is either $(s/2,\delta,0)$-hard or $(s/2,\delta,1)$-hard.

Theorem 2 (Monotone Errorless Amplification Lemma). *For $b \in \{0,1\}$, if f is (s,δ,b)-hard and $C : \{0,1\}^k \to \{0,1\}$ is a (t,ρ,p,b)-amplifier then $f' = C \circ f^k$ is $(s',1-\epsilon)$-hard where $s' = s/\big(\frac{k}{t}\cdot\frac{4}{\epsilon}\ln\frac{2}{\delta}\big)$, provided $t \geq \frac{16}{\delta}\ln\frac{4}{\epsilon}$, $\rho \leq \epsilon/4$, and $p = \Pr_x\big[f(x)=1\big]$.*

We prove Theorem 2 in Section 3. Bogdanov and Safra [1] proved a version of Theorem 2 where $s' = s/\big(k^3\cdot\frac{64}{\epsilon^2}\ln\frac{2}{\delta}\big)$, provided $t \geq \frac{4}{\delta}\ln\frac{8}{\epsilon}$ and $\rho \leq \epsilon/2$. Their argument involves considering the subcubes of $\{0,1\}^{n\times k}$ given by $f^k(x_1,\ldots,x_k) = y$ for each y individually and then combining the results for the different subcubes using a nontrivial probabilistic argument. We show how to give a direct argument that handles all the subcubes simultaneously. This idea alone actually simplifies the proof and reduces the query complexity to $O\big(k^2\cdot\frac{1}{\epsilon}\log\frac{1}{\delta}\big)$. Combining this idea with the ideas from our analysis in the proof of Theorem 1 allows us to further reduce the query complexity to $O\big(\frac{k}{t}\cdot\frac{1}{\epsilon}\log\frac{1}{\delta}\big)$. We believe this bound on the query complexity cannot be improved without exploiting some non-obvious structural property of (t,ρ,p,b)-amplifiers; however, we could not come up with a compelling way to formalize this.

Bogdanov and Safra [1] showed how to construct good amplifiers (with large t and small ρ) and how to use Theorem 2 with these amplifiers to do uniform and nonuniform errorless hardness amplification within NP.

1.3 Black-Box Lower Bounds

We give lower bounds on the query complexity and advice complexity of black-box errorless hardness amplification proofs. We allow ourselves to identify strings with functions; for example, we identify $\{0,1\}^{2^n}$ with the set of all functions from $\{0,1\}^n$ to $\{0,1\}$.

Definition 4. *An $(n, n', \delta, \epsilon, \alpha)$-black-box errorless hardness amplification is a pair (Enc, Dec) with $Enc : \{0,1\}^{2^n} \to \{0,1\}^{2^{n'}}$ and $Dec : \{0,1,\perp\}^{2^{n'}} \times \{0,1\}^{\alpha} \to \{0,1,\perp\}^{2^n}$, such that for all $f \in \{0,1\}^{2^n}$ and $A' \in \{0,1,\perp\}^{2^{n'}}$ there exists an $a \in \{0,1\}^{\alpha}$ such that the following holds, where $f' = Enc(f)$ and $A = Dec(A', a)$. If*

(i) $A'(x') \in \{f'(x'), \perp\}$ *for all* $x' \in \{0,1\}^{n'}$, *and*

(ii) $\Pr_{x'}[A'(x') = \perp] \leq 1 - \epsilon$,

then

(i) $A(x) \in \{f(x), \perp\}$ *for all* $x \in \{0,1\}^n$, *and*

(ii) $\Pr_x[A(x) = \perp] \leq \delta$.

We say it is q-query nonadaptive if there is an algorithm that takes (x, a) as input, nonadaptively makes q queries to A', and outputs $A(x)$.

In fact, our lower bounds hold even for the following weaker type of reduction.

Definition 5. *An $(n, n', \delta, \epsilon, \alpha)$-black-box non-errorless to errorless hardness amplification is a pair (Enc, Dec) with $Enc : \{0,1\}^{2^n} \to \{0,1\}^{2^{n'}}$ and $Dec : \{0,1,\perp\}^{2^{n'}} \times \{0,1\}^{\alpha} \to \{0,1\}^{2^n}$, such that for all $f \in \{0,1\}^{2^n}$ and $A' \in \{0,1,\perp\}^{2^{n'}}$ there exists an $a \in \{0,1\}^{\alpha}$ such that the following holds, where $f' = Enc(f)$ and $A = Dec(A', a)$. If*

(i) $A'(x') \in \{f'(x'), \perp\}$ *for all* $x' \in \{0,1\}^{n'}$, *and*

(ii) $\Pr_{x'}[A'(x') = \perp] \leq 1 - \epsilon$,

then $\Pr_x[A(x) \neq f(x)] \leq \delta$. *We say it is q-query nonadaptive if there is an algorithm that takes (x, a) as input, nonadaptively makes q queries to A', and outputs $A(x)$.*

In the proof of Theorem 1 we show that there exists a q-query nonadaptive $(n, n', \delta, \epsilon, \alpha)$-black-box errorless hardness amplification where $q = \frac{4}{\epsilon} \ln \frac{2}{\delta}$, $n' = kn$, and $\alpha = \left(\log_2 k + (k-1)n + 1\right) \cdot q$, provided $k \geq \frac{16}{\delta} \ln \frac{2}{\epsilon}$.

Theorem 3. *There exists a universal constant $c > 1$ such that the following holds. If there exists a q-query nonadaptive $(n, n', \delta, \epsilon, \alpha)$-black-box non-errorless to errorless hardness amplification then $q \geq \frac{1}{c} \cdot \frac{1}{\epsilon} \ln \frac{1}{\delta}$, provided $n \geq c$, $n' \geq c$, $2^{-n/c} \leq \delta \leq 1/3$, $2^{-n/c} \leq \epsilon \leq 1/3$, and $\alpha \leq 2^{n/c}$.*

Shaltiel and Viola [18] proved a similar result for the fully non-errorless setting (with the conclusion $q \geq \frac{1}{c} \cdot \frac{1}{\epsilon^2} \ln \frac{1}{\delta}$). Their proof can be adapted to our setting as follows. Where they use noise that flips a bit with probability $1/2 - \epsilon$, instead use noise that masks the bit with $\perp$ with probability $1 - \epsilon$ and reveals the correct bit with probability ϵ. Where they use noise that flips a bit with probability $1/2$, instead use "noise" that masks the bit with $\perp$ with probability 1. The rest of their proof goes through with some minor changes but without major changes.

Theorem 4. *If there exists an $(n, n', \delta, \epsilon, \alpha)$-black-box non-errorless to errorless hardness amplification then $2^{\alpha} \geq \log_3 \frac{1}{\epsilon}$, provided $\delta < 1/8$ and $\epsilon \geq 1/1.01^{1.01^{2^n}}$.*

We prove Theorem 4 in Section 4. In the fully non-errorless setting, lower bounds on advice complexity correspond to lower bounds on list size for approximately list-decoding error-correcting codes from flipped bits. Such a lower bound was given in [16] (see also [7]). In the non-errorless to errorless setting, lower bounds on advice complexity correspond to lower bounds on list size for approximately list-decoding error-correcting codes from erasures. For unique decoding, such lower bounds were given in [6, 3]. Our proof of Theorem 4 is simpler and cleaner than the proofs of the latter results (at the cost of achieving worse constants), and it handles approximate decoding. Also, the presentation in [6], which is geared toward coding theorists, views the rate, list size, and fraction of erasures as constants for an infinite family of codes. Our presentation is geared toward complexity theorists, who are interested more generally in the asymptotic relationships among all the parameters.

It is an open problem to prove some sort of uniform version of the Errorless XOR Lemma. Impagliazzo et al. [10] proved a sort of uniform version of the (non-errorless) XOR Lemma, but their techniques do not seem to apply to the errorless setting.

1.4 Preliminaries

We use the following standard Chernoff bound several times.

Theorem 5. *If $X_1, \ldots, X_\tau$ are fully independent indicator random variables each with expectation π, then* $\Pr\left[\sum_{j=1}^{\tau} X_j < \pi\tau/2\right] < e^{-\pi\tau/8}$.

2 Proof of Theorem 1

We prove the contrapositive. Suppose f' is not $(s', 1-\epsilon)$-hard and thus there is a circuit A' of size $\leq s'$ such that

(i) $A'(x_1, \ldots, x_k) \in \left\{f'(x_1, \ldots, x_k), \perp\right\}$ for all $x_1, \ldots, x_k$, and

(ii) $\Pr_{x_1,\ldots,x_k}\left[A'(x_1, \ldots, x_k) = \perp\right] \leq 1 - \epsilon$.

We give a nonuniform reduction that makes $\frac{4}{\epsilon}\ln\frac{2}{\delta}$ nonadaptive queries to A' and combines the results with some trivial computation, yielding a circuit A that witnesses that f is not (s, δ)-hard. To start out, we give a randomized algorithm (Algorithm 1) that solves f errorlessly using oracle access to A' and oracle access to f. The oracle queries to f only depend on the randomness (and not on the input), and later we will hard-wire a particular choice of randomness to get a circuit without oracle access to f.

Define the good set

$$G \;=\; \left\{x \in \{0,1\}^n \;:\; \Pr_{i,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_k}\left[A'(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_k) \neq \perp\right] \geq \epsilon/4\right\},$$

and define the bad set $B = \{0,1\}^n \backslash G$. That is, G is the set of inputs for which each iteration of the loop has at least an $\epsilon/4$ probability of producing output.

Claim 1. $|B| \leq (\delta/2) \cdot 2^n$.

Input: $x \in \{0,1\}^n$
Output: $f(x)$ or $\perp$

repeat $\frac{4}{\epsilon}\ln\frac{2}{\delta}$ times
  pick $i \in [k]$ and $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k \in \{0,1\}^n$ uniformly at random
  if $A'(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_k) \neq \perp$ then halt and output

$$f(x_1) \oplus \cdots \oplus f(x_{i-1}) \oplus A'(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_k) \oplus f(x_{i+1}) \oplus \cdots \oplus f(x_k)$$

end
halt and output $\perp$

Algorithm 1: Reduction for Theorem 1

Proof. Suppose for contradiction that $|B| > (\delta/2) \cdot 2^n$. Let $\gamma = |B|/2^{n+1}$. We define the event

$$W \;=\; \Big\{(x_1, \ldots, x_k) \in \{0,1\}^{n\times k} \;:\; \big|\{i \;:\; x_i \in B\}\big| \geq \gamma k\Big\}.$$

That is, W is the event that at least a γ fraction of coordinates are bad. We have

$$\begin{aligned}\Pr_{x_1,\ldots,x_k}\big[A'(x_1, \ldots, x_k) \neq \perp\big] \;\leq\; &\Pr_{x_1,\ldots,x_k}\big[(x_1, \ldots, x_k) \notin W\big] + \\ &\Pr_{x_1,\ldots,x_k}\big[A'(x_1, \ldots, x_k) \neq \perp \text{ and } (x_1, \ldots, x_k) \in W\big].\end{aligned}$$

We show that both terms on the right side are $< \epsilon/2$, thus contradicting property (ii) of A'.

Bounding the first term. Applying Theorem 5 with X_i as the indicator variable for $x_i \in B$, and with $\tau = k$ and $\pi = |B|/2^n$, we have

$$\Pr_{x_1,\ldots,x_k}\big[(x_1, \ldots, x_k) \notin W\big] \;<\; e^{-k\cdot|B|/2^{n+3}} \;<\; e^{-k\delta/16} \;\leq\; \epsilon/2$$

where the middle inequality follows by our assumption on $|B|$ and the last inequality follows by $k \geq \frac{16}{\delta}\ln\frac{2}{\epsilon}$.

Bounding the second term. For each $S \subseteq [k]$ we define the event

$$W_S \;=\; \big\{(x_1, \ldots, x_k) \in \{0,1\}^{n\times k} \;:\; \forall i \;\; x_i \in B \Leftrightarrow i \in S\big\}.$$

Note that the W_S's are disjoint and

$$W \;=\; \bigcup_{S \,:\, |S| \geq \gamma k} W_S.$$

We have

$$\begin{aligned}&\Pr_{x_1,\ldots,x_k}\big[A'(x_1, \ldots, x_k) \neq \perp \text{ and } (x_1, \ldots, x_k) \in W\big] \\ =\; &\sum_{S \,:\, |S| \geq \gamma k} \Pr_{x_1,\ldots,x_k}\big[A'(x_1, \ldots, x_k) \neq \perp \text{ and } (x_1, \ldots, x_k) \in W_S\big]\end{aligned}$$

$$
\begin{aligned}
&\leq \frac{1}{\gamma k} \sum_{S \subseteq [k]} |S| \cdot \Pr_{x_1,\ldots,x_k} \left[A'(x_1,\ldots,x_k) \neq \perp \text{ and } (x_1,\ldots,x_k) \in W_S \right] \\
&= \frac{1}{\gamma k} \sum_{i \in [k]} \sum_{S \ni i} \Pr_{x_1,\ldots,x_k} \left[A'(x_1,\ldots,x_k) \neq \perp \text{ and } (x_1,\ldots,x_k) \in W_S \right] \\
&= \frac{1}{\gamma k} \sum_{i \in [k]} \Pr_{x_1,\ldots,x_k} \left[A'(x_1,\ldots,x_k) \neq \perp \text{ and } x_i \in B \right] \\
&= \frac{1}{\gamma k} \sum_{i \in [k]} \sum_{x \in B} \Pr_{x_1,\ldots,x_k} \left[A'(x_1,\ldots,x_k) \neq \perp \;\middle|\; x_i = x \right] \cdot \Pr_{x_1,\ldots,x_k} [x_i = x] \\
&= \frac{1}{\gamma k 2^n} \sum_{x \in B} \sum_{i \in [k]} \Pr_{x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_k} \left[A'(x_1,\ldots,x_{i-1},x,x_{i+1},\ldots,x_k) \neq \perp \right] \\
&= \frac{1}{\gamma k 2^n} \sum_{x \in B} k \cdot \Pr_{i,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_k} \left[A'(x_1,\ldots,x_{i-1},x,x_{i+1},\ldots,x_k) \neq \perp \right] \\
&< \frac{1}{\gamma k 2^n} \sum_{x \in B} k \cdot \epsilon/4 \\
&= \frac{\epsilon/4}{\gamma 2^n} \cdot |B| \\
&= \epsilon/2
\end{aligned}
$$

where the second and fifth lines follow by the disjointness of the W_S's, and the remaining lines follow by simple rearrangements. □

The rest of the proof of Theorem 1 is similar to the argument from [1]. First we note that for all $x \in \{0,1\}^n$ and all choices of randomness, Algorithm 1 does indeed output either $f(x)$ or $\perp$. This follows trivially from the fact that if $A'(x_1,\ldots,x_{i-1},x,x_{i+1},\ldots,x_k) \neq \perp$ then

$$
A'(x_1,\ldots,x_{i-1},x,x_{i+1},\ldots,x_k) \;=\; f(x_1) \oplus \cdots \oplus f(x_{i-1}) \oplus f(x) \oplus f(x_{i+1}) \oplus \cdots \oplus f(x_k)
$$

by property (i) of A'. Next we observe that for each $x \in G$, we have

$$
\Pr_{\text{randomness}} \left[\text{Algorithm 1 outputs } \perp \right] \;\leq\; \left(1 - \epsilon/4\right)^{\frac{4}{\epsilon} \ln \frac{2}{\delta}} \;\leq\; \delta/2.
$$

Therefore

$$
\begin{aligned}
\Pr_{x,\text{ randomness}} \left[\text{Algorithm 1 outputs } \perp \right] \;&\leq\; \Pr_x \left[x \in B \right] + \\
&\quad \mathop{\mathrm{E}}_x \left[\Pr_{\text{randomness}} \left[\text{Algorithm 1 outputs } \perp \right] \;\middle|\; x \in G \right] \\
&\leq\; \delta/2 + \mathop{\mathrm{E}}_x \left[\delta/2 \;\middle|\; x \in G \right] \\
&=\; \delta
\end{aligned}
$$

where the second inequality follows by Claim 1 and by the above observation. It follows that there exists a setting of the randomness such that

(i) Algorithm 1 outputs $f(x)$ or $\perp$ for all x, and

(ii) $\Pr_x\left[\text{Algorithm 1 outputs } \perp\right] \leq \delta$.

To get a circuit A that witnesses that f is not (s,δ)-hard, just hard-wire the randomness and the values of $f(x_1)\oplus\cdots\oplus f(x_{i-1})\oplus f(x_{i+1})\oplus\cdots\oplus f(x_k)$ needed for this choice of randomness, and plug in the hypothesized circuit A'. Since A' has size $\leq s'$ and Algorithm 1 makes $\frac{4}{\epsilon}\ln\frac{2}{\delta}$ queries to A', A has size $\leq s'\cdot\frac{4}{\epsilon}\ln\frac{2}{\delta} = s$. Note that Algorithm 1 can trivially be implemented with *nonadaptive* access to A'.

3 Proof of Theorem 2

We prove the contrapositive. Suppose f' is not $(s', 1-\epsilon)$-hard and thus there is a circuit A' of size $\leq s'$ such that

(i) $A'(x_1,\ldots,x_k) \in \left\{f'(x_1,\ldots,x_k), \perp\right\}$ for all $x_1,\ldots,x_k$, and

(ii) $\Pr_{x_1,\ldots,x_k}\left[A'(x_1,\ldots,x_k) = \perp\right] \leq 1-\epsilon$.

We give a nonuniform reduction that makes $\frac{k}{t}\cdot\frac{4}{\epsilon}\ln\frac{2}{\delta}$ nonadaptive queries to A' and combines the results with some trivial computation, yielding a circuit A that witnesses that f is not (s,δ,b)-hard. To start out, we give a randomized algorithm (Algorithm 2) that solves f errorlessly using oracle access to A' and oracle access to f and $\sigma(C,\cdot,b)$. The oracle queries to f and $\sigma(C,\cdot,b)$ only depend on the randomness (and not on the input), and later we will hard-wire a particular choice of randomness to get a circuit without oracle access to f or $\sigma(C,\cdot,b)$.

Input: $x \in \{0,1\}^n$
Output: $f(x)$ or $\perp$

1 **repeat** $\frac{k}{t}\cdot\frac{4}{\epsilon}\ln\frac{2}{\delta}$ times
2 pick $i \in [k]$ and $x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_k \in \{0,1\}^n$ uniformly at random
3 if $A'(x_1,\ldots,x_{i-1},x,x_{i+1},\ldots,x_k) \neq \perp$ and $i \in \sigma(C,y,b)$ where

$$y \;=\; \big(f(x_1),\ldots,f(x_{i-1}),b,f(x_{i+1}),\ldots,f(x_k)\big)$$

  then halt and output $A'(x_1,\ldots,x_{i-1},x,x_{i+1},\ldots,x_k)$
4 **end**
5 halt and output $\perp$

Algorithm 2: Reduction for Theorem 2

Define the good set

$$G \;=\; \Bigg\{x \in f^{-1}(b) \;:\; \Pr_{i,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_k}\left[A'(x_1,\ldots,x_{i-1},x,x_{i+1},\ldots,x_k) \neq \perp \text{ and } i \in \sigma(C,y,b)\right] \geq \epsilon t/4k\Bigg\}$$

where y is as in line 3 of Algorithm 2, and define the bad set $B = f^{-1}(b)\backslash G$. That is, G is the set of inputs in $f^{-1}(b)$ for which each iteration of the loop has at least an $\epsilon t/4k$ probability of producing output.

Claim 2. $|B| \le (\delta/2) \cdot \left|f^{-1}(b)\right|$.

Proof. Suppose for contradiction that $|B| > (\delta/2) \cdot \left|f^{-1}(b)\right|$. Let $\gamma = |B|/2\left|f^{-1}(b)\right|$. We define the event

$$W \;=\; \left\{(x_1,\ldots,x_k) \in \{0,1\}^{n\times k} \;:\; \left|\left\{i \;:\; x_i \in B \text{ and } i \in \sigma\big(C, f^k(x_1,\ldots,x_k), b\big)\right\}\right| \ge \gamma t\right\}.$$

That is, W is the event that at least a $\gamma t/k$ fraction of coordinates are both bad and b-sensitive. We have

$$\begin{aligned}\Pr_{x_1,\ldots,x_k}\left[A'(x_1,\ldots,x_k) \neq \perp\right] \;\le\; &\Pr_{x_1,\ldots,x_k}\left[(x_1,\ldots,x_k) \notin W\right] + \\ &\Pr_{x_1,\ldots,x_k}\left[A'(x_1,\ldots,x_k) \neq \perp \text{ and } (x_1,\ldots,x_k) \in W\right].\end{aligned}$$

We show that both terms on the right side are $< \epsilon/2$, thus contradicting property (ii) of A'.

Bounding the first term. We have

$$\begin{aligned}\Pr_{x_1,\ldots,x_k}\left[(x_1,\ldots,x_k) \notin W\right] \;\le\; &\Pr_{x_1,\ldots,x_k}\left[\left|\sigma\big(C, f^k(x_1,\ldots,x_k), b\big)\right| < t\right] + \\ &\Pr_{x_1,\ldots,x_k}\left[(x_1,\ldots,x_k) \notin W \;\middle|\; \left|\sigma\big(C, f^k(x_1,\ldots,x_k), b\big)\right| \ge t\right].\end{aligned}$$

To show that this is $< \epsilon/2$, we show that the first of the two terms on the right side is $\le \epsilon/4$ and the second is $< \epsilon/4$. Since C is a (t,ρ,p,b)-amplifier, we have

$$\Pr_{x_1,\ldots,x_k}\left[\left|\sigma\big(C, f^k(x_1,\ldots,x_k), b\big)\right| < t\right] \;=\; \Pr_{y\sim_p\{0,1\}^k}\left[\left|\sigma(C,y,b)\right| < t\right] \;\le\; \rho \;\le\; \epsilon/4.$$

We have

$$\begin{aligned}&\Pr_{x_1,\ldots,x_k}\left[(x_1,\ldots,x_k) \notin W \;\middle|\; \left|\sigma\big(C, f^k(x_1,\ldots,x_k), b\big)\right| \ge t\right] \\ =\; &\mathop{\mathrm{E}}_{y\sim_p\{0,1\}^k}\left[\Pr_{x_1,\ldots,x_k}\left[(x_1,\ldots,x_k) \notin W \;\middle|\; f^k(x_1,\ldots,x_k) = y\right] \;\middle|\; \left|\sigma(C,y,b)\right| \ge t\right].\end{aligned}$$

Fix any $y \in \{0,1\}^k$ such that $\left|\sigma(C,y,b)\right| \ge t$, and for now let us abbreviate $\sigma(C,y,b)$ as σ. Then we have

$$\begin{aligned}&\Pr_{x_1,\ldots,x_k}\left[(x_1,\ldots,x_k) \notin W \;\middle|\; f^k(x_1,\ldots,x_k) = y\right] \\ =\; &\mathop{\mathrm{E}}_{(x_i)_{i\notin\sigma}}\left[\Pr_{(x_i)_{i\in\sigma}}\left[(x_1,\ldots,x_k) \notin W \;\middle|\; f(x_i) = b \;\;\forall i \in \sigma\right] \;\middle|\; f(x_i) = y_i \;\;\forall i \notin \sigma\right]\end{aligned}$$

since $\sigma \subseteq \big\{i \;:\; y_i = b\big\}$. Now fix any $(x_i)_{i\notin\sigma}$ such that $f(x_i) = y_i$ for all $i \notin \sigma$. Then we have

$$\begin{aligned}&\Pr_{(x_i)_{i\in\sigma}}\left[(x_1,\ldots,x_k) \notin W \;\middle|\; f(x_i) = b \;\;\forall i \in \sigma\right] \\ =\; &\Pr_{(x_i)_{i\in\sigma}}\left[\left|\big\{i \in \sigma \;:\; x_i \in B\big\}\right| < \gamma t \;\middle|\; f(x_i) = b \;\;\forall i \in \sigma\right]\end{aligned}$$

$$\leq \Pr_{(x_i)_{i\in\sigma}}\Big[\big|\{i\in\sigma \ : \ x_i\in B\}\big| < \gamma\cdot|\sigma| \ \Big|\ f(x_i)=b \ \ \forall i\in\sigma\Big]$$

where the inequality follows by $t\leq|\sigma|$. Applying Theorem 5 with X_j as the indicator variable for $x_i\in B$ where i is the jth value in σ and x_i is chosen uniformly from $f^{-1}(b)$, and with $\tau=|\sigma|$ and $\pi=|B|/\big|f^{-1}(b)\big|$, we have that the latter quantity is less than

$$e^{-|\sigma|\cdot|B|/8|f^{-1}(b)|} \ < \ e^{-|\sigma|\cdot\delta/16} \ \leq \ e^{-t\delta/16} \ \leq \ \epsilon/4$$

where the first inequality follows by our assumption on $|B|$, the middle inequality follows by $|\sigma|\geq t$, and the last inequality follows by $t\geq\frac{16}{\delta}\ln\frac{4}{\epsilon}$. This establishes that

$$\Pr_{x_1,\ldots,x_k}\Big[(x_1,\ldots,x_k)\notin W \ \Big|\ \big|\sigma\big(C,f^k(x_1,\ldots,x_k),b\big)\big|\geq t\Big] \ < \ \epsilon/4.$$

Bounding the second term. This is similar to the corresponding part of the analysis in the proof of Theorem 1. For each $S\subseteq[k]$ we define the event

$$W_S \ = \ \Big\{(x_1,\ldots,x_k)\in\{0,1\}^{n\times k} \ : \ \forall i \ \ \Big(x_i\in B \text{ and } i\in\sigma\big(C,f^k(x_1,\ldots,x_k),b\big)\Big)\Leftrightarrow i\in S\Big\}.$$

Note that the W_S's are disjoint and

$$W \ = \ \bigcup_{S \ : \ |S|\geq\gamma t} W_S.$$

Using the shorthand y as in line 3 of Algorithm 2, we have

$$\begin{aligned}
&\Pr_{x_1,\ldots,x_k}\big[A'(x_1,\ldots,x_k)\neq\perp \text{ and } (x_1,\ldots,x_k)\in W\big]\\
=\ &\sum_{S \ : \ |S|\geq\gamma t}\Pr_{x_1,\ldots,x_k}\big[A'(x_1,\ldots,x_k)\neq\perp \text{ and } (x_1,\ldots,x_k)\in W_S\big]\\
\leq\ &\frac{1}{\gamma t}\sum_{S\subseteq[k]}|S|\cdot\Pr_{x_1,\ldots,x_k}\big[A'(x_1,\ldots,x_k)\neq\perp \text{ and } (x_1,\ldots,x_k)\in W_S\big]\\
=\ &\frac{1}{\gamma t}\sum_{i\in[k]}\sum_{S\ni i}\Pr_{x_1,\ldots,x_k}\big[A'(x_1,\ldots,x_k)\neq\perp \text{ and } (x_1,\ldots,x_k)\in W_S\big]\\
=\ &\frac{1}{\gamma t}\sum_{i\in[k]}\Pr_{x_1,\ldots,x_k}\Big[A'(x_1,\ldots,x_k)\neq\perp \text{ and } x_i\in B \text{ and } i\in\sigma\big(C,f^k(x_1,\ldots,x_k),b\big)\Big]\\
=\ &\frac{1}{\gamma t}\sum_{i\in[k]}\sum_{x\in B}\Pr_{x_1,\ldots,x_k}\Big[A'(x_1,\ldots,x_k)\neq\perp \text{ and } i\in\sigma\big(C,f^k(x_1,\ldots,x_k),b\big) \ \Big|\ x_i=x\Big]\cdot\Pr_{x_1,\ldots,x_k}[x_i=x]\\
=\ &\frac{1}{\gamma t2^n}\sum_{x\in B}\sum_{i\in[k]}\Pr_{x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_k}\big[A'(x_1,\ldots,x_{i-1},x,x_{i+1},\ldots,x_k)\neq\perp \text{ and } i\in\sigma(C,y,b)\big]\\
=\ &\frac{1}{\gamma t2^n}\sum_{x\in B}k\cdot\Pr_{i,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_k}\big[A'(x_1,\ldots,x_{i-1},x,x_{i+1},\ldots,x_k)\neq\perp \text{ and } i\in\sigma(C,y,b)\big]\\
<\ &\frac{1}{\gamma t2^n}\sum_{x\in B}k\cdot\epsilon t/4k
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\epsilon/4}{\gamma 2^n} \cdot |B| \\
&\leq \epsilon/2
\end{aligned}
$$

where the second and fifth lines follow by the disjointness of the W_S's, the last line follows by $\left|f^{-1}(b)\right| \leq 2^n$, and the remaining lines follow by simple rearrangements. For the seventh line, we used the fact that $x \in B$ implies $f(x) = b$ and thus $y = f^k(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_k)$. □

We now finish the proof of Theorem 2. First we note that for all $x \in \{0,1\}^n$ and all choices of randomness, Algorithm 2 does indeed output either $f(x)$ or $\perp$. This follows trivially from the facts that $A'(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_k) \neq \perp$ implies

$$
A'(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_k) \;=\; f'(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_k)
$$

by property (i) of A', and $i \in \sigma(C, y, b)$ implies $f'(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_k) = f(x)$, where y is as in line 3 of Algorithm 2. Next we observe that for each $x \in G$, we have

$$
\Pr_{\text{randomness}}\left[\text{Algorithm 2 outputs } \perp\right] \;\leq\; \left(1 - \epsilon t/4k\right)^{\frac{k}{t} \cdot \frac{4}{\epsilon} \ln \frac{2}{\delta}} \;\leq\; \delta/2.
$$

Therefore, picking $x \in f^{-1}(b)$ uniformly at random, we have

$$
\begin{aligned}
\Pr_{x,\text{ randomness}}\left[\text{Algorithm 2 outputs } \perp\right] \;&\leq\; \Pr_x\left[x \in B\right] + \\
&\qquad \mathop{\mathrm{E}}_x\left[\Pr_{\text{randomness}}\left[\text{Algorithm 2 outputs } \perp\right] \;\middle|\; x \in G\right] \\
&\leq\; \delta/2 + \mathop{\mathrm{E}}_x\left[\delta/2 \;\middle|\; x \in G\right] \\
&=\; \delta
\end{aligned}
$$

where the second inequality follows by Claim 2 and by the above observation. It follows that there exists a setting of the randomness such that

(i) Algorithm 2 outputs $f(x)$ or $\perp$ for all $x \in \{0,1\}^n$, and

(ii) $\Pr_x\left[\text{Algorithm 2 outputs } \perp\right] \leq \delta$ where $x \in f^{-1}(b)$ is chosen uniformly at random.

To get a circuit A that witnesses that f is not (s, δ, b)-hard, just hard-wire the randomness and the correct responses to the $\sigma(C, \cdot, b)$ queries (which only depend on the randomness and not on x), and plug in the hypothesized circuit A'. In fact, the iterations for which $i \notin \sigma(C, y, b)$ for this particular choice of randomness can simply be eliminated. Since A' has size $\leq s'$ and Algorithm 2 makes $\leq \frac{k}{t} \cdot \frac{4}{\epsilon} \ln \frac{2}{\delta}$ queries to A', A has size $\leq s' \cdot \frac{k}{t} \cdot \frac{4}{\epsilon} \ln \frac{2}{\delta} = s$. Note that Algorithm 2 can trivially be implemented with *nonadaptive* access to A'.

4 Proof of Theorem 4

Let $(\mathrm{Enc}, \mathrm{Dec})$ be an $(n, n', \delta, \epsilon, \alpha)$-black-box non-errorless to errorless hardness amplification with $\delta < 1/8$ and $\epsilon \geq 1/1.01^{1.01^{2^n}}$. We use the notation $\Delta(f_1, f_2)$ for the relative Hamming distance between bit strings f_1 and f_2. We begin with a completely standard claim that asserts the existence of a good error-correcting code. We include the proof for completeness.

Claim 3. *There exists an $F \subseteq \{0,1\}^{2^n}$ such that*

(1) $\forall f_1, f_2 \in F$: if $f_1 \neq f_2$ then $\Delta(f_1, f_2) > 2\delta$, and

(2) $|F| = 4\log_3 \frac{1}{\epsilon}$.

Proof. Pick $F \subseteq \{0,1\}^{2^n}$ randomly by choosing $f_1, \ldots, f_{4\log_3 \frac{1}{\epsilon}} \in \{0,1\}^{2^n}$ independently uniformly at random and setting $F = \left\{f_1, \ldots, f_{4\log_3 \frac{1}{\epsilon}}\right\}$. To prove (1) and (2), it suffices to show that

$$\Pr\left[\exists i_1, i_2 \in \{1, \ldots, 4\log_3 \tfrac{1}{\epsilon}\} : i_1 \neq i_2 \text{ and } \Delta(f_{i_1}, f_{i_2}) \leq 2\delta\right] \;<\; 1.$$

For each pair $i_1 \neq i_2$, since $\delta < 1/8$ we have

$$\Pr\left[\Delta(f_{i_1}, f_{i_2}) \leq 2\delta\right] \;\leq\; \Pr\left[\Delta(f_{i_1}, f_{i_2}) < 1/4\right] \;<\; e^{-2^n/16}$$

by applying Theorem 5 with X_j as the indicator variable for $f_{i_1}(x) \neq f_{i_2}(x)$ where x is the jth string in $\{0,1\}^n$, and with $\tau = 2^n$ and $\pi = 1/2$. By a union bound, the probability in question is at most $\left(4\log_3 \frac{1}{\epsilon}\right)^2 \cdot e^{-2^n/16} < 1$ since $\epsilon \geq 1/1.01^{1.01^{2^n}}$. □

For the rest of the proof of Theorem 4 we fix a set F as in Claim 3.

Claim 4. *There exists an $E \subseteq F$ and an $A' \in \{0,1,\perp\}^{2^{n'}}$ such that*

(i) $\forall f \in E$: $A'(x') \in \{f'(x'), \perp\}$ for all $x' \in \{0,1\}^{n'}$, where $f' = Enc(f)$, and

(ii) $\Pr_{x'}[A'(x') = \perp] \leq 1 - \epsilon$, and

(iii) $|E| = \log_3 \frac{1}{\epsilon}$.

Proof. For each $x' \in \{0,1\}^{n'}$ let $m_{x'} = \text{majority}_{f \in F}\text{Enc}(f)(x')$, breaking a tie arbitrarily, and define

$$M_{x'} \;=\; \left\{f \in F \;:\; \text{Enc}(f)(x') = m_{x'}\right\}.$$

For any set $E \subseteq F$, define

$$M_E^{-1} \;=\; \left\{x' \in \{0,1\}^{n'} \;:\; E \subseteq M_{x'}\right\}.$$

We construct a sequence of sets $E_0 \subseteq E_1 \subseteq E_2 \subseteq \cdots \subseteq E_{\log_3 \frac{1}{\epsilon}} \subseteq F$ such that for all i, $|E_i| = i$ and $\left|M_{E_i}^{-1}\right| \geq 2^{n'}/3^i$. Then we can take $E = E_{\log_3 \frac{1}{\epsilon}}$ and

$$A'(x') \;=\; \begin{cases} m_{x'} & \text{if } x' \in M_E^{-1} \\ \perp & \text{otherwise} \end{cases}$$

and (i), (ii), and (iii) all follow immediately. We do the construction inductively. The base case $i = 0$ is trivial. Now assume $i \in \left\{0, 1, \ldots, \log_3 \frac{1}{\epsilon} - 1\right\}$ and we have a set $E_i \subseteq F$ with $|E_i| = i$ and $\left|M_{E_i}^{-1}\right| \geq 2^{n'}/3^i$. For each $x' \in M_{E_i}^{-1}$, since $i \leq |F|/4$ we have

$$\Pr_{f \in F \setminus E_i}[f \in M_{x'}] \;=\; \frac{|M_{x'} \setminus E_i|}{|F \setminus E_i|} \;\geq\; \frac{\frac{1}{2}|F| - i}{|F| - i} \;\geq\; 1/3$$

where f is chosen uniformly at random. Thus for some $f \in F \backslash E_i$ we have $\Pr_{x' \in M_{E_i}^{-1}}[f \in M_{x'}] \geq 1/3$ where x' is chosen uniformly at random. For this fixed f, setting $E_{i+1} = E_i \cup \{f\}$ we have $|E_{i+1}| = |E_i| + 1 = i + 1$ and

$$\left|M_{E_{i+1}}^{-1}\right| \;=\; \left|\left\{x' \in M_{E_i}^{-1} \;:\; f \in M_{x'}\right\}\right| \;=\; \left|M_{E_i}^{-1}\right| \cdot \Pr_{x' \in M_{E_i}^{-1}}[f \in M_{x'}] \;\geq\; \left|M_{E_i}^{-1}\right|/3 \;\geq\; 2^{n'}/3^{i+1}.$$

This finishes the induction step. □

Now to prove Theorem 4, suppose for contradiction that $2^\alpha < \log_3 \frac{1}{\epsilon}$. Then by the pigeonhole principle there must exist $f_1, f_2 \in E$ such that $f_1 \neq f_2$ and the advice string corresponding to f_1 and A' equals the advice string corresponding to f_2 and A'. Call this advice string a, and let $A = \mathrm{Dec}(A', a)$. By Definition 5, we must have $\Delta(A, f_1) \leq \delta$ and $\Delta(A, f_2) \leq \delta$. But this is impossible because $\Delta(f_1, f_2) > 2\delta$ by property (1) in Claim 3.

Acknowledgments

I thank anonymous reviewers for helpful comments.

References

[1] A. Bogdanov and M. Safra. Hardness Amplification for Errorless Heuristics. In *Proceedings of the 48th IEEE Symposium on Foundations of Computer Science*, pages 418-426, 2007.

[2] J. Buresh-Oppenheim, V. Kabanets, and R. Santhanam. Uniform Hardness Amplification in NP via Monotone Codes. Electronic Colloquium on Computational Complexity, Technical Report TR06-154, 2006.

[3] G. Cohen, S. Litsyn, and G. Zémor. Upper Bounds on Generalized Distances. *IEEE Transactions on Information Theory*, 40: 2090-2092, 1994.

[4] P. Gopalan and V. Guruswami. Hardness Amplification Within NP Against Deterministic Algorithms. In *Proceedings of the 23rd IEEE Conference on Computational Complexity*, pages 19-30, 2008.

[5] O. Goldreich, N. Nisan, and A. Wigderson. On Yao's XOR Lemma. Electronic Colloquium on Computational Complexity, Technical Report TR95-050, 1995.

[6] V. Guruswami. List Decoding from Erasures. Chapter 10 of List Decoding of Error-Correcting Codes, *Lecture Notes in Computer Science* 3282, 2005.

[7] V. Guruswami and S. Vadhan. A Lower Bound on List Size for List Decoding. In *Proceedings of the 8th International Workshop on Randomization and Computation*, pages 318-329, 2005.

[8] A. Healy, S. Vadhan, and E. Viola. Using Nondeterminism to Amplify Hardness. *SIAM Journal on Computing*, 35(4): 903-931, 2006.

[9] R. Impagliazzo. Hard-Core Distributions for Somewhat Hard Problems. In *Proceedings of the 36th IEEE Symposium on Foundations of Computer Science*, pages 538-545, 1995.

[10] R. Impagliazzo, R. Jaiswal, V. Kabanets, and A. Wigderson. Uniform Direct Product Theorems: Simplified, Optimized, and Derandomized. *SIAM Journal on Computing*, 39(4): 1637-1665, 2010.

[11] R. Impagliazzo and A. Wigderson. P=BPP Unless E has Subexponential Circuits: Derandomizing the XOR Lemma. In *Proceedings of the 29th ACM Symposium on Theory of Computing*, pages 220-229, 1997.

[12] A. Klivans and R. Servedio. Boosting and Hard-Core Sets. *Machine Learning*, 53(3): 217-238, 2003.

[13] L. Levin. Average Case Complete Problems. *SIAM Journal on Computing*, 15(1): 285-286, 1986.

[14] L. Levin. One-Way Functions and Pseudorandom Generators. *Combinatorica*, 7(4): 357-363, 1987.

[15] C.-J. Lu, S.-C. Tsai, and H.-L. Wu. Improved Hardness Amplification in NP. *Theoretical Computer Science*, 370(1-3): 293-298, 2007.

[16] C.-J. Lu, S.-C. Tsai, and H.-L. Wu. On the Complexity of Hardness Amplification. *IEEE Transactions on Information Theory*, 54(10): 4575-4586, 2008.

[17] R. O'Donnell. Hardness Amplification Within NP. *Journal of Computer and System Sciences*, 69(1): 68-94, 2004.

[18] R. Shaltiel and E. Viola. Hardness Amplification Proofs Require Majority. *SIAM Journal on Computing*, 39(7): 3122-3154, 2010.

[19] L. Trevisan. List Decoding Using the XOR Lemma. In *Proceedings of the 44th IEEE Symposium on Foundations of Computer Science*, pages 126-135, 2003.

[20] L. Trevisan. On Uniform Amplification of Hardness in NP. In *Proceedings of the 37th ACM Symposium on Theory of Computing*, pages 31-38, 2005.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de