comput. complex. **20** (2011), 145 – 171

1016-3328/11/010145-27
published online April 19, 2011
DOI 10.1007/s00037-011-0003-7

computational complexity

COMPLEXITY OF HARD-CORE SET PROOFS

Chi-Jen Lu, Shi-Chun Tsai, and Hsin-Lung Wu

Abstract. We study a fundamental result of Impagliazzo (*FOCS'95*) known as the hard-core set lemma. Consider any function $f : \{0,1\}^n \to \{0,1\}$ which is "mildly hard", in the sense that any circuit of size s must disagree with f on at least a δ fraction of inputs. Then, the hard-core set lemma says that f must have a hard-core set H of density δ on which it is "extremely hard", in the sense that any circuit of size $s' = O(s/(\frac{1}{\varepsilon^2}\log(\frac{1}{\varepsilon\delta})))$ must disagree with f on at least $(1-\varepsilon)/2$ fraction of inputs from H.

There are three issues of the lemma which we would like to address: the loss of circuit size, the need of non-uniformity, and its inapplicability to a low-level complexity class. We introduce two models of hard-core set proofs, a strongly black-box one and a weakly black-box one, and show that those issues are unavoidable in such models.

First, we show that using any strongly black-box proof, one can only prove the hardness of a hard-core set for smaller circuits of size at most $s' = O(s/(\frac{1}{\varepsilon^2}\log\frac{1}{\delta}))$. Next, we show that any weakly black-box proof must be inherently non-uniform—to have a hard-core set for a class G of functions, we need to start from the assumption that f is hard against a non-uniform complexity class with $\Omega(\frac{1}{\varepsilon}\log|G|)$ bits of advice. Finally, we show that weakly black-box proofs in general cannot be realized in a low-level complexity class such as $\mathsf{AC}^0[p]$—the assumption that f is hard for $\mathsf{AC}^0[p]$ is not sufficient to guarantee the existence of a hard-core set.

Keywords. Hard-core set, hardness amplification, black-box proofs.

Subject classification. 68Q05, 68Q10, 68Q17.

1. Introduction

The hardness of a function is a fundamental notion in complexity theory. Informally speaking, a function f is hard if any circuit of small size must fail to compute it correctly on some inputs. More precisely, we can characterize the hardness by parameters δ and s and say that f is δ-hard (or has hardness δ) for size s if any circuit of size at most s must fail to compute f correctly on at least δ fraction of inputs. An important question is: given a δ-hard function for size s, can we transform it into a harder function with hardness $\delta' > \delta$ for size about s? This is known as the task of hardness amplification, and it plays a crucial role in the study of derandomization, in which one would like to obtain an extremely hard function, with hardness $(1-\varepsilon)/2$, so that it looks like a random function and can be used to construct pseudo-random generators, see Babai *et al.* (1993), Impagliazzo (1995), Impagliazzo & Wigderson (1997), Sudan *et al.* (2001), Yao (1982).

One may wonder whether the hardness of a function mostly comes from a large subset of inputs that are extremely hard to compute. So a natural question is: given any δ-hard function for size s, is there always a subset of inputs of density about δ on which f is extremely hard for circuits of size about s? A seminal result of Impagliazzo (1995) answers this affirmatively. He showed that any δ-hard function for size s indeed has a subset H of the inputs with density δ on which f has hardness $(1-\varepsilon)/2$ for circuits of size $s' = O(s/(\frac{1}{\varepsilon^2}\log\frac{1}{\delta\varepsilon}))$. Such a set H is called an ε-hard-core set for size s'.

In addition to answering a very basic question in complexity theory, the hard-core set lemma has found applications in learning theory, see Barak *et al.* (2009), Klivans & Servedio (2003) and cryptography, see Holenstein (2005), and has become an important tool in the study of pseudo-randomness. It can be used to provide an alternative proof of Yao's celebrated XOR Lemma, see Impagliazzo (1995), or to construct a pseudo-random generator directly from a mildly hard function, bypassing the XOR lemma, see Sudan *et al.* (2001). Recently, it has become a key ingredient in the study of hardness amplification for functions in NP, see Healy *et al.* (2006), O'Donnell (2004), Trevisan (2003, 2005). In spite of

its importance, there are some issues of the hard-core set lemma which are still not well understood and have become the bottlenecks in some applications. This calls for a more thorough study of the lemma.

The first issue is the loss of circuit size. Note that in Impagliazzo's result, the hardness on the hard-core set, although increased, is actually measured against circuits of a smaller size s', as opposed to the initial size s. This loss of circuit size was later reduced by Klivans & Servedio (2003) who showed the existence an ε-hard-core set of density $\delta/2$ for size $s' = O(s/(\frac{1}{\varepsilon^2}\log\frac{1}{\delta}))$. Then, a natural question is: can the size s' be further improved to, say, $\Omega(s)$? The second issue is non-uniformity. In all previous works, see Barak *et al.* (2009), Impagliazzo (1995), Klivans & Servedio (2003), even when one only wants to have a hard-core set that is hard for uniform algorithms, one still needs to start from a function that is hard for non-uniform circuits, or algorithms supplied with advice strings. In fact, this becomes the bottleneck in Trevisan's work on uniform hardness amplification for functions in NP, see Trevisan (2003, 2005), in which he showed how to amplify hardness from $\delta = 1 - 1/\mathrm{poly}(n)$ to $\delta' = (1-\varepsilon)/2$ against BPP algorithms, with $\varepsilon = 1/\log^c n$ for a small constant $c < 1$. Note that ideally one would like to have the hardness amplified to $(1-\varepsilon)/2$ with $\varepsilon \leq 1/\mathrm{poly}(n)$, and what prevents Trevisan from reaching this goal (or even $\varepsilon = 1/\log n$) is the long advice strings needed by existing proofs of the hard-core set lemma, see Barak *et al.* (2009), Impagliazzo (1995), Klivans & Servedio (2003). This is because all these proofs need to start from a function that is hard against algorithms with advice strings of length $\Omega(1/\varepsilon^2)$, but existing techniques are only able to remove advice strings of length $O(\log n)$. On the other hand, it is known that hardness amplification for functions in a higher complexity class, such as EXP, only requires advice strings of length $O(\log(1/\varepsilon))$, see Sudan *et al.* (2001). So a natural question is: can the length of the advice strings needed in the hard-core set lemma be reduced? The third issue is that the lemma currently does not apply to a low-level complexity class such as $\mathsf{AC}^0[p]$. The reason is that existing proofs of the lemma, see Barak *et al.* (2009), Impagliazzo (1995), Klivans & Servedio

(2003), all need to start from the assumption that f is hard for a complexity class which is high enough to include the majority function with input length at least $\Omega(1/\varepsilon)$, in order to show that f has a hard-core set. This prevents ones from applying hardness amplification to obtain an average-case hard function, and consequently constructing a good pseudo-random generator for such a complexity class still remains an open problem. Thus, an interesting question is: for any function f which is δ-hard for $\mathsf{AC}^0[p]$, does it always have an ε-hard-core set for $\mathsf{AC}^0[p]$ with $\varepsilon = 1/\mathrm{poly}(n)$?

All these three issues seem inherent in existing proofs of the hard-core set lemma, and they look difficult to resolve. One may wonder whether they are indeed impossible to avoid. However, proving such negative results appears to require proving circuit lower bounds, which seems to be far beyond our reach. Therefore, we would like to identify general and reasonable models for the hard-core set lemma in which such negative results can actually be proven.

Black-Box Hard-Core Set Proofs. The hard-core set lemma, when stated in the contrapositive way, basically says that given any function f with no hard-core set for small circuits (on any such subset H, there is a small circuit C_H with a good correlation with f), one can find a small circuit C that is close to f. A closer look at Impagliazzo's proof shows that the circuit C is simply the weighted majority of a small subset of those circuits $\{C_H\}$. In fact, one can replace the class of small circuits $\{C_H\}$ by any class G of functions, and Impagliazzo's proof shows that given any f with no hard-core set for functions in G, one can construct a function C that is close to f by taking a weighted majority of a small subset of functions in G. Note that C only uses those functions in G as an oracle.

This motivates us to define our first model of hard-core set proofs as follows. We say that a (non-uniform) oracle algorithm $\textsc{Dec}^{(\cdot)}$ with a decision function $D : \{0,1\}^q \to \{0,1\}$ realizes a *strongly black-box* (δ, ε, k)-proof (of a hard-core set) if the following holds. First, Dec will be given a family $G = \{g_1, \ldots, g_k\}$ of functions as oracle together with a multi-set $I = \{i_1, \ldots, i_q\}$

as its advice, and for any input x, it will query the functions $g_{i_1}, \ldots, g_{i_q}$, all at x, and then output $D(g_{i_1}(x), \ldots, g_{i_q}(x))$. Moreover, it satisfies the property that for any G and for any f which has no ε-hard-core set of density $\Omega(\delta)$ for G, there exists a multiset I of size q such that the function $\textsc{Dec}^{G,I}$ is δ-close to f (i.e., $\textsc{Dec}^{G,I}(x) \neq f(x)$ for at most δ fraction of x). We call q the query complexity of Dec and observe that it relates to the loss of circuit size in the hard-core set lemma, with $s' \leq s/q$. In order to have $D(g_{i_1}(x), \ldots, g_{i_q}(x))$ computed by a circuit of size s, one needs each g_{i_j} to be computed by a circuit of size $s' \leq s/q$. Thus, from the assumption that f is hard for size s, one can only show that f has a hard-core set for size s'. That is, one has a loss of the circuit size by a factor of q. Note that all the known hard-core set proofs, see Barak *et al.* (2009), Impagliazzo (1995), Klivans & Servedio (2003), are in fact done in such a strongly black-box way.

Our second model of hard-core set proofs generalizes the first one by removing the constraint on how the algorithm Dec works; the algorithm Dec and its advice now are allowed to be of arbitrary form. We say that a (non-uniform) oracle algorithm $\textsc{Dec}^{(\cdot)}$ realizes a *weakly black-box* (δ, ε, k)-proof (of a hard-core set) if the following holds. For any family G of k functions and for any function f that has no ε-hard-core set of density $\Omega(\delta)$ for G, there exists an advice string α such that $\textsc{Dec}^{G,\alpha}$ is δ-close to f.

Our Results. We have three results, which give negative answers to the three questions we raised before, with respect to our models of black-box proofs. Note that our lower bounds for our second model (weakly black-box one) also hold for our first model as the first model is a special case of the second one.

Our first result shows that any *strongly* black-box (δ, ε, k)-proof must require a query complexity of $q = \Omega(\frac{1}{\varepsilon^2} \log \frac{1}{\delta})$. Our lower bound explains why it is very difficult to have a smaller loss of circuit size in the hard-core set lemma; in fact, any strongly black-box proof must suffer a loss of such a large factor q. Note that our lower bound is tight as it is matched (up to a constant factor) by the upper bounds from Klivans & Servedio (2003) and Barak *et al.* (2009).

Our second result shows that any *weakly* black-box (δ, ε, k)-proof must require an advice string of length $\Omega(\frac{1}{\varepsilon} \log k)$. This explains why it is difficult to have a uniform version of the hard-core set lemma; in fact, any weakly black-box proof is inherently non-uniform. Moreover, one cannot hope to improve Trevisan's uniform hardness amplification results, see Trevisan (2003, 2005), by reducing the length of the advice string needed in the hard-core set proof, unless one can come up with a non-black-box approach. Note that from the query upper bound of Klivans & Servedio (2003), one can immediately have an upper bound of $O(\frac{1}{\varepsilon^2}(\log \frac{1}{\delta}) \log k)$ on the length of the advice string, which has a gap from our lower bound. It is not clear which bound can be further improved, but our feeling is that this upper bound may likely be improved.

Our third result shows that no *weakly* black-box (δ, ε, k)-proof can be implemented in a low-level complexity class such as $\mathsf{AC}^0[p]$ for a prime p, when $\delta < 1/20$ and $\varepsilon \leq 1/n$. More precisely, we show that the function DEC realizing such a black-box proof can be used to approximate the majority function, but on the other hand, the majority function cannot be approximated by an $\mathsf{AC}^0[p]$ circuit. Therefore, one cannot have a hard-core set lemma for $\mathsf{AC}^0[p]$, unless one can prove it in a non-black-box way.

Bounds from Hardness Amplification. There is no previous result directly on the lower bounds of hard-core set proofs. However, one can obtain such bounds from lower bounds for the task of hardness amplification, see Lu *et al.* (2008), Viola (2006). This is because the hard-core set lemma can be used for hardness amplification, as shown in Impagliazzo (1995), and a closer inspection shows that a black-box hard-core set proof in fact yields a hardness amplification in a similar black-box model.

In particular, one can conclude the following. First, using a result of Viola (2006) for hardness amplification, we can derive a lower bound of $\min(\frac{1}{10\varepsilon}, \frac{n}{5 \log n})$ on the query complexity of any strongly black-box (δ, ε, k)-proof. Note that this bound is always smaller than our bound. Second, we can use the result in Lu *et al.* (2008) to derive a lower bound of $\Omega(\log \frac{(1-2\delta)^2}{\varepsilon})$ on the length of the advice string for any weakly black-box (δ, ε, k)-proof. Note that

this bound is exponentially worse than ours. Finally, we can use another result of Viola (2006) to show that for any weakly black-box (δ, ε, k)-proof, if the function DEC satisfies the additional condition that it only needs an advice string of logarithmic (in the circuit size of DEC) length, then it cannot be implemented by an $\mathsf{AC}^0[p]$ circuit. Note that this additional condition is not required in our result and our proof is much simpler. On the other hand, under this additional condition, Viola achieved something stronger: such DEC can be used as oracle gates by an AC^0 circuit to compute the majority function exactly, instead of approximately.

Subsequent Results. Shortly after the conference version of our paper appeared in Lu *et al.* (2007), Shaltiel & Viola (2008) improved the results in Viola (2006) for hardness amplification. More precisely, they showed that any black-box proof that amplifies hardness from δ to $(1-\varepsilon)/2$ and uses only an advice string of polynomial length must require $\Omega(\frac{1}{\varepsilon^2}\log\frac{1}{\delta})$ oracle queries and can be used to compute the majority function on $1/\varepsilon$ bits. Consequently, their results imply our first and third results, but nevertheless, our proofs are considerably simpler than theirs.

Our second result is seemly related to the recent result of Barak *et al.* (2009) on what they called a "uniform" version of the hard-core set lemma. More precisely, they provided a uniform algorithm that can output a circuit that is δ-close to f when given access to a very powerful oracle. The oracle, when in turn given oracle access to any measure of density δ, can return a circuit as a weak hypothesis such that $C(x) = f(x)$ with probability $(1+\varepsilon)/2$ when x is sampled according to the measure. Note that not only is their oracle different from ours, but their oracle also needs the help from the algorithm for providing the measure it needs in order to generate the weak hypothesis. In fact, their result is more naturally seen as providing a uniform boosting algorithm. Since their model is different from ours, their positive result does not contradict our negative one.

Our Techniques. Recall that a black-box (δ, ε, k)-proof requires a universal algorithm DEC which works for any f and G such that f has no ε-hard-core set for G. To establish a negative result, it

suffices to show the existence of f and G that violate the guarantee of DEC.

To obtain our query lower bound, suppose we have a strongly black-box proof with the function DEC making only a small number of queries. We show the existence of f and $G = \{g_1, \ldots, g_k\}$ for which DEC fails, using a probabilistic argument. Choose f randomly and then choose $g_1, \ldots, g_k$ independently as k noisy versions of f, with each $g_i(x)$ being $f(x)$ altered with independent noise of rate $(1-2\varepsilon)/2$. We can show that f is unlikely to have an ε-hard-core set for G, because it is unlikely to have a subset on which every g_i has a large deviation from f, when k is large enough. On the other hand, we can show that if the function DEC does not make enough queries to functions in G, there is a good chance that it is not close to f. This implies the existence of G and f for which DEC fails to work. Thus, we conclude that the query complexity must be high.

To obtain our advice lower bound, we show the existence of a family $G = \{g_1, \ldots, g_k\}$ of functions such that one can find a large collection Γ of functions with the property that every function in Γ has no hard-core set for G, but no two functions in Γ are close. As each function in Γ is a legitimate candidate for f, DEC^G must use an advice string of length $\log_2 |\Gamma|$ to specify the correct candidate. Again, we show the existence using a probabilistic argument, with G chosen randomly. The candidates for Γ are functions G_I, with $I = \{i_1, \ldots, i_t\}$, defined as $G_I(x) = \text{MAJ}(g_{i_1}(x), \ldots, g_{i_t}(x))$, where MAJ denotes the majority function. We will let $t = \lfloor 1/\varepsilon \rfloor$, so that every G_I has a good correlation with some g_i for $i \in I$, which implies that G_I has no ε-hard-core set for G. On the other hand, for any G_I and G_J with small $I \cap J$, they are likely to be far away because for any input x, $\sum_{i \in I \cap J} g_i(x)$ is likely to be small, so there is a good chance that the values of $G_I(x)$ and $G_J(x)$ are dominated by $\sum_{i \in I \setminus J} g_i(x)$ and $\sum_{j \in J \setminus I} g_j(x)$, respectively, and hence there is a good chance that $G_I(x) \neq G_J(x)$. This implies that with high probability, each G_I is far away from many other G_J's, and by Turán's well-known theorem, there must be many G_I's that are far away from each other, and they form the set Γ.

To prove that no weakly black-box (δ, ε, k)-proof can be realized in $\mathsf{AC}^0[p]$ for a prime p, we show that given such a black-box proof with the function DEC, one can find some G such that DEC^G can be used to approximate the majority function. Again, we use the observation that for any G, the function G_I, with $|I| = t \leq 1/\varepsilon$, has no ε-hard-core set for G. We choose the family $G = \{g_1, \ldots, g_k\}$ with g_i defined as $g_i(x) = x_i$ (the i-th bit of x) for $i \leq n$ and $g_i(x) = 0$ otherwise, for $x \in \{0,1\}^n$ (assuming $t \leq n \leq k$), and let $f = G_I$ with $I = \{1, \ldots, t\}$. As f has no ε-hard-core set for G, there must exist an advice string α such that $\text{DEC}^{G,\alpha}(x) = G_I(x) = \text{MAJ}(x_1, \ldots, x_t)$ for at least a δ fraction of x, and by an averaging argument there must exist some fixed $\bar{x}_{t+1}, \ldots, \bar{x}_n$ such that $\text{DEC}^{G,\alpha}(x_1, \ldots, x_t, \bar{x}_{t+1}, \ldots, \bar{x}_n) = \text{MAJ}(x_1, \ldots, x_t)$ for at least a δ fraction of $(x_1, \ldots, x_t)$. By hard-wiring α and $\bar{x}_{t+1}, \ldots, \bar{x}_n$ into the circuit for DEC, we get a circuit that is δ-close to the majority function on t bits.

Organization of this paper. First, in Section 2, we give some preliminaries and define our two models for black-box proofs of hard-core set. In Section 3, we prove a query lower bound for strongly black-box proofs. Then, we show a lower bound on the advice length needed in weakly black-box proofs in Section 4. Finally, in Section 5, we show that no weakly black-box proofs can be realized in $\mathsf{AC}^0[p]$.

2. Preliminaries

For any $n \in \mathbb{N}$, let $[n]$ denote the set $\{1, \ldots, n\}$ and let $\mathcal{U}_n$ denote the uniform distribution over $\{0,1\}^n$. For a finite set X, we will also use X to denote the uniform distribution over it when there is no confusion. For a string $x \in \{0,1\}^n$, we let x_i denote the i-th bit of x. Let F_n denote the set of all Boolean functions $f : \{0,1\}^n \to \{0,1\}$. Let $\mathsf{SIZE}(s)$ be the class of Boolean functions computed by (non-uniform) circuits of size s. Let $\mathsf{AC}^0[p](s)$ denote the class of Boolean functions computed by constant-depth circuits of size s equipped with $\mod_p$ gates (which output 0 exactly when the input bits sum to 0 modulo p), and let $\mathsf{AC}^0[p] = \mathsf{AC}^0[p](\text{poly}(n))$. Given a multi-set (or simply a

set) S, we let $|S|$ denote the number of elements in it, counting multiplicity. Given a set $G = \{g_1, \ldots, g_k\} \subseteq F_n$, together with a multi-set $I = \{i_1, \ldots, i_q\} \subseteq [k]$ of indices, let g_I denote the function such that $g_I(x) = (g_{i_1}(x), \cdots, g_{i_q}(x))$ for $x \in \{0,1\}^n$. We say that two functions f and g in F_n are δ-close if $\Pr_{x \in \mathcal{U}_n}[f(x) \neq g(x)] \leq \delta$. All the logarithms in this paper will have base two.

We will need the following simple lower bound on the tail probability of binomial distribution.

FACT 2.1. *Let $Z_1, \ldots, Z_t$ be i.i.d. binary random variables, with $\mathrm{E}[Z_i] = \mu$ for every $i \in [t]$. Suppose $\Omega(\frac{1}{\sqrt{t}}) = \varepsilon \leq \frac{1}{3}$. Then, we have the following:* (1) *if $\mu \leq \frac{1}{2} + \varepsilon$, then $\Pr[\sum_{i \in [t]} Z_i \leq \frac{1-\varepsilon}{2} t] \geq 2^{-O(\varepsilon^2 t)}$, and* (2) *if $\mu \geq \frac{1}{2} - \varepsilon$, then $\Pr[\sum_{i \in [t]} Z_i \geq \frac{1+\varepsilon}{2} t] \geq 2^{-O(\varepsilon^2 t)}$.*

PROOF. First, consider the case that $\mu \leq \frac{1}{2} + \varepsilon$. Note that if we fix ε and vary μ, then the probability gets smaller when μ gets larger, because $\sum_{i=1}^{t} Z_i$ becomes more unlikely to have a small value when the probability of $Z_i = 1$ becomes higher. Thus, it suffices to show the lower bound for the case of $\mu = \frac{1}{2} + \varepsilon$. Then,

$$\begin{aligned}
\Pr\left[\sum_{i=1}^{t} Z_i \leq \frac{1-\varepsilon}{2} t\right] &= \sum_{0 \leq j \leq \frac{1-\varepsilon}{2} t} \binom{t}{j} \cdot \left(\frac{1+2\varepsilon}{2}\right)^{j} \left(\frac{1-2\varepsilon}{2}\right)^{t-j} \\
&\geq \sum_{\frac{1-2\varepsilon}{2} t \leq j \leq \frac{1-\varepsilon}{2} t} \binom{t}{j} \cdot \left(\frac{1+2\varepsilon}{2}\right)^{j} \left(\frac{1-2\varepsilon}{2}\right)^{t-j} \\
&\geq \frac{\varepsilon t}{2} \cdot \binom{t}{\frac{1-2\varepsilon}{2} t} \cdot \left(\frac{1+2\varepsilon}{2}\right)^{\frac{1-2\varepsilon}{2} t} \left(\frac{1-2\varepsilon}{2}\right)^{\frac{1+2\varepsilon}{2} t}.
\end{aligned}$$

Using the inequality that $\binom{t}{\alpha t} \geq \frac{1}{O(\sqrt{t})} (\frac{1}{\alpha})^{\alpha t} (\frac{1}{1-\alpha})^{(1-\alpha)t}$ from Stirling's formula, the above becomes

$$\frac{\varepsilon t}{O(\sqrt{t})} \left(\frac{2}{1-2\varepsilon}\right)^{\frac{1-2\varepsilon}{2} t} \left(\frac{2}{1+2\varepsilon}\right)^{\frac{1+2\varepsilon}{2} t} \left(\frac{1+2\varepsilon}{2}\right)^{\frac{1-2\varepsilon}{2} t} \left(\frac{1-2\varepsilon}{2}\right)^{\frac{1+2\varepsilon}{2} t},$$

which is at least

$$\begin{aligned}
\Omega(\varepsilon\sqrt{t}) \left(\frac{1-2\varepsilon}{1+2\varepsilon}\right)^{2\varepsilon t} &= \Omega(\varepsilon\sqrt{t}) \left(1 - \frac{4\varepsilon}{1+2\varepsilon}\right)^{2\varepsilon t} \\
&\geq \Omega(\varepsilon\sqrt{t}) 2^{-O(\varepsilon^2 t)} \geq 2^{-O(\varepsilon^2 t)},
\end{aligned}$$

where the first inequality uses the fact that $1 - x \geq 2^{-cx}$ for some constant c when $x(= \frac{4\varepsilon}{1+2\varepsilon}) \leq \frac{4}{5}$, and the last inequality follows from the condition that $t = \Omega(\frac{1}{\varepsilon^2})$.

The second case with $\mu \geq \frac{1-2\varepsilon}{2}$ follows immediately from the first case by symmetry. More precisely, define new random variables $Y_1, \ldots, Y_t$, with $Y_i = 1 - Z_i$ for $i \in [t]$, and then we can get the desired bound by applying the bound of the first case to these new variables. □

We will also need the following result, known as Turán's Theorem, which can be found in standard textbooks (see e.g. Alon & Spencer (2000)).

Fact 2.2 (Turán 1941). *Given a graph $G = (V, E)$, let d_v denote the degree of a vertex v. Then, the size of its maximum independent set is at least $\sum_{v \in V} \frac{1}{d_v+1}$.*

2.1. Hardness and Hard-Core Set Lemma. We say that a function is hard if no small circuit computes a function that is close to it. Formally, we define the hardness of a function as follows.

Definition 2.3. *We say that a function $f \in F_n$ is δ-hard (or has hardness δ) for size s, if for any $C \in \mathsf{SIZE}(s), Pr_{x \in \mathcal{U}_n}[C(x) \neq f(x)] > \delta$.*

Impagliazzo introduced the following notion of a hard-core set of a hard function.

Definition 2.4 (Impagliazzo 1995). *We say that a function $f \in F_n$ has an ε-hard-core set $H \subseteq \{0,1\}^n$ for size s, if for any $C \in \mathsf{SIZE}(s)$, $\Pr_{x \in H}[C(x) \neq f(x)] > \frac{1-\varepsilon}{2}$.*

Now we can state Impagliazzo's hard-core set lemma, see Impagliazzo (1995), which is the focus of our paper.

Lemma 2.5 (Impagliazzo 1995). *Any function $f \in F_n$ which is δ-hard for size s must have an ε-hard-core set H for size s', with $|H| \geq \delta 2^n$ and $s' = O(s/(\frac{1}{\varepsilon^2} \log \frac{1}{\delta\varepsilon}))$.*

Note that in this lemma, the hardness on the set H is measured against a smaller circuit size s', as compared to the original circuit

size s. This was later improved by Klivans & Servedio (2003) to $s' = O(s/(\frac{1}{\varepsilon^2}\log\frac{1}{\delta}))$ but at the expense of having a slightly smaller hard-core set of size $\delta 2^{n-1}$. A closer look at their proofs shows that they work for the more general setting with hardness measured against any class of functions instead of just circuits. For this, let us first formalize the notion that a function has no hard-core set for a class $G \subseteq F_n$.

DEFINITION 2.6. *Given a set $G = \{g_1, \ldots, g_k\} \subseteq F_n$, we say that a function $f \in F_n$ is (δ, ε, G)-easy if for any $H \subseteq \{0,1\}^n$ of size $\delta 2^n$, there is a function $g \in G$ such that $\Pr_{x\in H}[g(x) \neq f(x)] \leq \frac{1-\varepsilon}{2}$.*

Then, from Impagliazzo (1995) and its improvement in Klivans & Servedio (2003), one actually has the following.

LEMMA 2.7. *For some $q = O(\frac{1}{\varepsilon^2}\log\frac{1}{\delta})$, there exists a function $D : \{0,1\}^q \to \{0,1\} \in \mathsf{SIZE}(\mathrm{poly}(q))$ such that for some constant, c the following holds. For any $G = \{g_1, \ldots, g_k\} \subseteq F_n$, if a function $f \in F_n$ is $(c\delta, \varepsilon, G)$-easy, then there is a multi-set I with $|I| = q$ such that $\Pr_x[D(g_I(x)) \neq f(x)] \leq \delta$.*

In Impagliazzo (1995), $c = 1$ and D is the majority function (and $q = O(\frac{1}{\varepsilon^2}\log\frac{1}{\delta\varepsilon})$), while in Klivans & Servedio (2003), $c = 1/2$ and D is a majority of majority functions.

2.2. Black-Box Proofs of Hard-Core Sets. Now we introduce our two models for black-box proofs of hard-core sets. The first one is stronger than the second.

DEFINITION 2.8. *We say that a (non-uniform) oracle algorithm $\mathrm{DEC}^{(\cdot)}$ realizes a strongly black-box (δ, ε, k)-proof (of a hard-core set) if for some $q \in \mathbb{N}$ it has a decision function $D : \{0,1\}^q \to \{0,1\}$ such that for some constant c the following holds. For any $G = \{g_1, \ldots, g_k\} \subseteq F_n$, if a function $f \in F_n$ is $(c\delta, \varepsilon, G)$-easy, then there is a multi-set I with $|I| = q$ such that $\mathrm{DEC}^{G,I}(x) = D(g_I(x))$ and $\Pr_x\left[\mathrm{DEC}^{G,I}(x) \neq f(x)\right] \leq \delta$. We call q the query complexity of* DEC.

In this model, I can be seen as the advice, which is of the form of a multi-set $I = \{i_1, \ldots, i_q\}$, and the algorithm DEC is restricted to be of the following form: on input x, it queries the functions

$g_{i_1}, \ldots, g_{i_q}$ all on the input x, applies the function D on the q answer bits, and outputs $D(g_{i_1}(x), \ldots, g_{i_q}(x))$. Note that all the known proofs of the hard-core set lemma are in fact done in our first model, see Barak *et al.* (2009), Impagliazzo (1995), Klivans & Servedio (2003). Our second model generalizes the first one by removing the format constraint on the algorithm Dec and its advice. That is, the algorithm Dec and its advice now are allowed to be of arbitrary form.

Definition 2.9. *We say that a (non-uniform) oracle algorithm* $\text{Dec}^{(\cdot)}$ *realizes a weakly black-box* (δ, ε, k)*-proof (of a hard-core set) if for some constant* c *the following holds. For any* $G = \{g_1, \ldots, g_k\} \subseteq F_n$*, if a function* $f \in F_n$ *is* $(c\delta, \varepsilon, G)$*-easy, then there is an advice string* α *such that* $\Pr_x[\text{Dec}^{G,\alpha}(x) \neq f(x)] \leq \delta$.

Note that in the two definitions above, we do not place any constraint on the computational complexity of Dec. Our first two results show that even having an unbounded computational power, Dec still needs to make a sufficient number of queries and use a sufficiently long advice string in these two models, respectively. On the other hand, our third result targets the computational complexity of Dec and shows that it cannot be implemented in a low-level complexity class such as $\mathsf{AC}^0[p]$. Here, we say that the oracle algorithm Dec^G can be *implemented* in a circuit class if the function Dec^G can be computed by a circuit in the class equipped with functions from G as oracle gates.

3. Query Complexity in Strongly Black-Box Proofs

In this section, we give a lower bound on the query complexity of any strongly black-box hard-core set proof. Formally, we have the following.

Theorem 3.1. *Suppose* $2^{-c_1 n} \leq \varepsilon, \delta < c_2$*, and* $\omega(\frac{1}{\varepsilon^2} \log \frac{1}{\delta}) \leq k \leq 2^{2^{c_3 n}}$*, for small enough constants* $c_1, c_2, c_3 > 0$*. Then, any strongly black-box* (δ, ε, k)*-proof must have a query complexity of* $q = \Omega(\frac{1}{\varepsilon^2} \log \frac{1}{\delta})$.

Our lower bound is optimal since it is matched (up to a constant factor) by the upper bound of Klivans and Servedio, see Lemma 2.7. Note that the assumption $k \geq \omega(\frac{1}{\varepsilon^2}\log\frac{1}{\delta})$ is reasonable, since in standard settings of the hard-core set lemma, G typically consists of circuits of polynomial (or larger) size, which gives $k = |G| \geq 2^{\mathrm{poly}(n)}$.

The roadmap for the proof is the following. Consider any DEC that realizes such a strongly black-box proof. We would like to show the existence of a function f and a family $G = \{g_1, \ldots, g_k\}$ of functions such that f is $(c\delta, \varepsilon, G)$-easy (with c being the constant associated with DEC), but the algorithm DEC without making enough queries cannot approximate f well. We will prove their existence by a probabilistic argument.

Now we proceed to the proof of the theorem. Suppose the parameters ε, δ, k satisfy the condition stated in the theorem. Suppose we have such a black-box proof realized by an oracle algorithm DEC with the decision function D. Consider k *independent* random functions $b_1, \ldots, b_k$ from F_n, which will serve as noise vectors, such that for any i and x, the value of $b_i(x)$ is chosen independently with

$$\Pr[b_i(x) = 0] = \frac{1+2\varepsilon}{2}.$$

Let f be a perfectly random function from F_n, so that $\Pr[f(x) = 1] = \frac{1}{2}$ independently for any x, and let $g_1, \ldots, g_k$ be k independent noisy versions of f defined as $g_i(x) = f(x) \oplus b_i(x)$, for any i and x. Let $B = \{b_1, \ldots, b_k\}$ and $G = \{g_1, \ldots, g_k\}$. First, we show that f is likely to be $(c\delta, \varepsilon, G)$-easy, where f and B are chosen randomly and G is determined by f and B as described above.

LEMMA 3.2. *If* $k = \omega(\frac{1}{\varepsilon^2}\log\frac{1}{\delta})$, *then* $\Pr_{f,B}[f \text{ is not } (c\delta, \varepsilon, G)\text{-easy}] = o(1)$.

PROOF. Consider any $H \subseteq \{0,1\}^n$ of size $c\delta 2^n$. We call f hard on H if $\Pr_{x\in H}[g_i(x) \neq f(x)] > \frac{1-\varepsilon}{2}$ for every i. Note that for any i, the random variables $b_i(x)$'s, for $x \in H$, are i.i.d. with $\mathrm{E}[b_i(x)] = \frac{1-2\varepsilon}{2}$, and $b_i(x) = 1$ exactly when $g_i(x) \neq f(x)$. Thus,

the probability that f is hard on H equals

$$\Pr_B\left[\forall i \in [k] : \sum_{x\in H} b_i(x) > \frac{1-\varepsilon}{2}|H|\right] = \prod_{i\in[k]} \Pr_{b_i}\left[\sum_{x\in H} b_i(x) > \frac{1-\varepsilon}{2}|H|\right] \leq \left(2^{-\Omega(\varepsilon^2\delta 2^n)}\right)^k,$$

where the equality is due to the fact that each b_i is independent from others and the inequality uses the Chernoff bound.

Recall that f is not $(c\delta, \varepsilon, G)$-easy exactly when f is hard on some H of size $c\delta 2^n$. Then, a union bound, over the choice of H, shows that it happens with probability at most

$$\binom{2^n}{c\delta 2^n} \cdot \left(2^{-\Omega(\varepsilon^2\delta 2^n)}\right)^k \leq \left(\frac{e}{c\delta}\right)^{c\delta 2^n} \cdot 2^{-\Omega(\varepsilon^2\delta 2^n k)} \leq 2^{O(\delta 2^n \log\frac{1}{\delta})} \cdot 2^{-\Omega(\varepsilon^2\delta 2^n k)},$$

which is $o(1)$ when $k = \omega(\frac{1}{\varepsilon^2}\log\frac{1}{\delta})$ and $\delta < c_2$ for some small enough constant c_2. □

Next, we show that with a small q, DEC is unlikely to approximate f well. Recall that for a multi-set $I = \{i_1, \ldots, i_q\} \subseteq [k]$, $g_I(x)$ denotes $(g_{i_1}(x), \ldots, g_{i_q}(x))$. We say that DEC can δ-approximate f if there is a multi-set $I \subseteq [k]$ with $|I| = q$ such that $D \circ g_I$ is δ-close to f (i.e., $\Pr_x[D(g_I(x)) \neq f(x)] \leq \delta$).

LEMMA 3.3. *If* $q = o(\frac{1}{\varepsilon^2}\log\frac{1}{\delta})$, *then* $\Pr_{f,B}[$DEC *can* δ*-approximate* $f] = o(1)$.

PROOF. Consider any multi-set $I \subseteq [k]$ with $|I| = q$. First, we show the following. □

CLAIM 3.4. *For any* $x \in \{0,1\}^n$, $\Pr_{f,B}[D(g_I(x)) \neq f(x)] \geq 2\delta$.

PROOF. Let $\tilde{I}$ denote the set of elements from I, removing multiplicity, and $\tilde{D}$ the function such that $\tilde{D}(g_{\tilde{I}}(x)) = D(g_I(x))$. For example, for $I = \{1,1,2\}$, we have $\tilde{I} = \{1,2\}$ and $\tilde{D}(g_1(x), g_2(x)) = D(g_1(x), g_1(x), g_2(x))$. Then

$$\Pr_{f,B}[D(g_I(x)) \neq f(x)] = \Pr_{f,B}[\tilde{D}(g_{\tilde{I}}(x)) \neq f(x)] = \frac{1}{2}p(0) + \frac{1}{2}p(1),$$

where $p(v) = \Pr_{f,B}[\tilde{D}(g_{\tilde{I}}(x)) = v \mid f(x) = 1 - v]$ for $v \in \{0,1\}$, so it suffices to give a lower bound for either $p(0)$ or $p(1)$. Let $\tilde{I} = \{i_1, \ldots, i_{\tilde{q}}\}$, where $\tilde{q}$ is clearly at most q. Assume without loss of generality that $|\tilde{D}^{-1}(1)| \geq 2^{\tilde{q}-1}$, and we will give a lower bound for $p(1)$ (otherwise, we bound $p(0)$ in a similar way). Let $Z = (Z_1, \ldots, Z_{\tilde{q}})$ denote the sequence of random variables $(b_{i_1}(x), \ldots, b_{i_{\tilde{q}}}(x))$, which are i.i.d. with $\mathrm{E}[Z_i] = \frac{1-2\varepsilon}{2}$. Note that $g_i(x) = b_i(x)$ when $f(x) = 0$, so

$$p(1) = \Pr_B\left[\tilde{D}(b_{i_1}(x), \ldots, b_{i_{\tilde{q}}}(x)) = 1\right] = \sum_{y \in \tilde{D}^{-1}(1)} \Pr[Z = y].$$

The above is the sum of $|\tilde{D}^{-1}(1)| \geq 2^{\tilde{q}-1}$ values from the $2^{\tilde{q}}$ values: $\Pr[Z = y]$ for $y \in \{0,1\}^{\tilde{q}}$, so it is clearly no less than the sum of the $2^{\tilde{q}-1}$ smallest values from them. Observe that $\Pr[Z = y] = (\frac{1-2\varepsilon}{2})^{\#_1(y)}(\frac{1+2\varepsilon}{2})^{\tilde{q}-\#_1(y)}$, where $\#_1(y)$ denotes the number of 1's in the string y, so $\Pr[Z = y] \leq \Pr[Z = y']$ whenever $\#_1(y) \geq \#_1(y')$. As a result, $p(1)$ is at least

$$\sum_{y:\#_1(y) > \frac{1}{2}\tilde{q}} \Pr[Z = y] = \Pr\left[\sum_{i \in [\tilde{q}]} Z_i > \frac{1}{2}\tilde{q}\right]$$

$$\geq \Pr\left[\sum_{i \in [\tilde{q}]} Z_i > \frac{1+\varepsilon}{2}\tilde{q}\right] \geq 2^{-O(\varepsilon^2 \tilde{q})},$$

by Fact 2.1 (2). So when $\tilde{q} \leq q = o(\frac{1}{\varepsilon^2}\log\frac{1}{\delta})$, we have $\Pr_{f,B}[D(g_I(x)) \neq f(x)] \geq \frac{1}{2}p(1) \geq 2\delta$. □

Now for any fixed multi-set I with $|I| = q$, let Y_x, for $x \in \{0,1\}^n$, denote the binary random variable such that $Y_x = 1$ if and only if $D(g_I(x)) \neq f(x)$. Clearly, they are i.i.d. random variables, and we know from above that $\mathrm{E}[Y_x] \geq 2\delta$ for any x. So by Chernoff bound,

$$\Pr_{f,G}[D \circ g_I \text{ is } \delta\text{-close to } f] = \Pr\left[\sum_x Y_x \leq \delta 2^n\right] \leq 2^{-\Omega(\delta^2 2^n)}.$$

Then, a union bound over the choices of $I \subseteq [k]$ with $|I| = q$ gives

$$\Pr_{f,B}[\exists I \text{ with } |I| = q : D \circ g_I \text{ is } \delta\text{-close to } f] \leq k^q \cdot 2^{-\Omega(\delta^2 2^n)} \leq o(1),$$

since $\varepsilon, \delta \geq 2^{-c_1 n}$ and $k \leq 2^{2^{c_3 n}}$, for small enough constants c_1, $c_3 > 0$. □

From Lemma 3.2 and Lemma 3.3, we conclude that for ε, δ, k as in Theorem 3.1, there exist $f \in F_n$ and $G = \{g_1, \ldots, g_k\} \subseteq F_n$ which satisfy the following:

- f is $(c\delta, \varepsilon, G)$-easy, and
- for every multi-set $I \subseteq [k]$ of size $q = o(\frac{1}{\varepsilon^2} \log \frac{1}{\delta})$, $\Pr_x [D(g_I(x)) \neq f(x)] > \delta$.

Therefore, any Dec realizing a strongly black-box (δ, ε, k)-proof must have $q = \Omega(\frac{1}{\varepsilon^2} \log \frac{1}{\delta})$. This proves Theorem 3.1.

Remark 3.5. *As pointed out by an anonymous reviewer, our proof of* Theorem 3.1 *is similar in spirit to that used by* Freund (1995) *to show a lower bound on the number of calls a boosting algorithm must make to a weak learner.*

4. Advice Complexity in Weakly Black-Box Proofs

In this section, we show a length lower bound on the advice string needed in any weakly black-box hard-core set proof. This explains why a uniform version of the hard-core set lemma is hard to come by and any black-box proof is inherently non-uniform. Formally, we have the following.

Theorem 4.1. *Suppose* $2^{-c_1 n} \leq \varepsilon, \delta < c_2$, *and* $\frac{1}{\varepsilon^3} \leq k \leq 2^{2^{c_3 n}}$, *for small enough constants* $c_1, c_2, c_3 > 0$. *Then, any weakly black-box* (δ, ε, k)*-proof must need an advice string of length* $\Omega(\frac{1}{\varepsilon} \log k)$.

As a comparison, the proof of Klivans and Servedio, see Lemma 2.7, provides an upper bound of $O(\frac{1}{\varepsilon^2} \log \frac{1}{\delta} \log k)$ on the advice length, so there is a gap of a factor $O(\frac{1}{\varepsilon} \log \frac{1}{\delta})$ between our lower bound and their upper bound. As in Theorem 3.1, one can also argue that the range assumption on the parameters is reasonable.

The roadmap for the proof is the following. Consider any DEC realizing such a weakly black-box proof. We will show the existence of a family $G = \{g_1, \ldots, g_k\} \subseteq F_n$ with respect to which we can find a large collection Γ of functions satisfying the following two properties: (1) any function in Γ is $(c\delta, \varepsilon, G)$-easy (with c the constant associated with DEC) and (2) any two functions in Γ are not 2δ-close. This then implies a lower bound of $\log|\Gamma|$ on the advice length. Again, we will show the existence by a probabilistic argument.

Now we proceed to the proof of the theorem. First, we independently sample k perfectly random functions $g_1, \ldots, g_k \in F_n$ (independently for any i and x, $g_i(x) = 1$ with probability exactly $\frac{1}{2}$), and let $G = \{g_1, \ldots, g_k\}$. Now for any set $I = \{i_1, \ldots, i_t\} \subseteq [k]$, let G_I be the function such that $G_I(x) = \textsc{Maj}(g_{i_1}(x), \ldots, g_{i_t}(x))$, where MAJ denotes the majority function. Then, we have the following lemma, which follows from the known discriminator lemma stating that any majority gate has a good correlation with one of its input bits, see Goldmann *et al.* (1992), Pisier (1981), Hajnal *et al.* (1993). For completeness, we also give its proof.

LEMMA 4.2. *Let $G = \{g_1, \ldots, g_k\}$ be any set of functions from F_n. Then, for any $I \subseteq [k]$, the function G_I is $(c\delta, \frac{1}{|I|}, G)$-easy.*

PROOF. Let I be a multi-set of size t. For any $H \subseteq \{0,1\}^n$, consider the $|H| \times t$ matrix M such that for $x \in H$ and $j \in I$, $M_{x,j} = 1$ if $g_j(x) = G_I(x)$ and 0 otherwise. Clearly, each row of M has more 1's than 0's, so the fraction of 1's must be at least $\frac{1}{2}(1 + \frac{1}{t})$ (otherwise, the number of 1's minus the number of 0's is less than $t \cdot \frac{1}{t} = 1$, a contradiction). Then, by an averaging argument, some column must have at least this fraction of 1's. That is, for any $H \subseteq \{0,1\}^n$ (including those of size $c\delta 2^n$), there exists a function $g_j \in G$ such that

$$\Pr_{x \in H}[g_j(x) = G_I(x)] \geq \frac{1}{2}\left(1 + \frac{1}{t}\right).$$

Therefore, G_I is a $(c\delta, \frac{1}{t}, G)$-easy function. □

Let $t = \left\lfloor \frac{1}{\varepsilon} \right\rfloor$, let $V^t = \{I \subseteq [k] : |I| = t\}$, and consider the class $\{G_I : I \in V^t\}$ of functions. Lemma 4.2 tells us that each

function in the class is $(c\delta, \varepsilon, G)$-easy. Our next step is to find a large collection of functions from this class such that any two of them are not close. Note that whether or not two functions G_I, G_J are close really depends on the choice of G. We will show that if I and J have a small intersection, then G_I and G_J are unlikely to be close for a random G.

LEMMA 4.3. *For any $I, J \in V^t$ with $|I \cap J| < \frac{t}{2}$, $Pr_G[G_I$ is 2δ-close to $G_J] \leq 2^{-\Omega(2^n)}$.*

PROOF. Consider any such I and J. First, we prove the following. □

CLAIM 4.4. *For any $x \in \{0,1\}^n$, $\Pr_G[G_I(x) \neq G_J(x)] = \Omega(1)$.*

PROOF. Note that for any x, $g_1(x), \ldots, g_k(x)$ can be seen as a sequence of i.i.d. binary random variables $Z_1, \ldots, Z_k$, with $\mathrm{E}[Z_i] = \frac{1}{2}$ for each i. Let Z_I denote the subsequence of random variables Z_i for $i \in I$, and similarly for Z_J. Thus, our goal is to show that $\Pr[\textsc{Maj}(Z_I) \neq \textsc{Maj}(Z_J)] = \Omega(1)$.

Let $K = I \cap J$, $I_1 = I \setminus K$, and $J_1 = J \setminus K$, and note that $|K| < |I_1|, |J_1|$. Consider the following three events.

- A_1: $\left|\sum_{i \in K} Z_i - \frac{|K|}{2}\right| \leq \frac{1}{2}\sqrt{|K|}$. By Chernoff bound, $\Pr[\neg A_1] < \gamma$ for a constant $\gamma < 1$, so $\Pr[A_1] = \Omega(1)$.
- A_2: $\sum_{i \in I_1} Z_i \leq \frac{1}{2}(|I_1| - \sqrt{|I_1|})$. By Fact 2.1 (1) with $\mu = \frac{1}{2}$, $\Pr[A_2] = \Omega(1)$.
- A_3: $\sum_{i \in J_1} Z_i \geq \frac{1}{2}(|J_1| + \sqrt{|J_1|})$. By Fact 2.1 (2) with $\mu = \frac{1}{2}$, $\Pr[A_3] = \Omega(1)$.

Now observe that if $A_1 \wedge A_2$, then

$$\sum_{i \in I} Z_i \leq \frac{1}{2}(|K| + |I_1| + \sqrt{|K|} - \sqrt{|I_1|}) < \frac{|K| + |I_1|}{2} = \frac{|I|}{2},$$

which implies that $\textsc{Maj}(Z_I) = 0$. Similarly, if $A_1 \wedge A_3$, then

$$\sum_{i \in J} Z_i \geq \frac{1}{2}(|K| + |J_1| - \sqrt{|K|} + \sqrt{|J_1|}) > \frac{|K| + |J_1|}{2} = \frac{|J|}{2},$$

which implies that $\textsc{Maj}(Z_J) = 1$. That is, if $A_1 \wedge A_2 \wedge A_3$, then $\textsc{Maj}(Z_I) = 0 \wedge \textsc{Maj}(Z_J) = 1$, so $\textsc{Maj}(Z_I) \neq \textsc{Maj}(Z_J)$. Since the events A_1, A_2, A_3 are independent from each other (as each depends on a separate set of random variables), we have

$$\begin{aligned}\Pr[\textsc{Maj}(Z_I) \neq \textsc{Maj}(Z_J)] &\geq \Pr[A_1 \wedge A_2 \wedge A_3] \\ &= \Pr[A_1] \cdot \Pr[A_2] \cdot \Pr[A_3] \geq \Omega(1).\end{aligned}$$

□

We also give an alternative proof suggested by an anonymous reviewer as follows.

Proof. (Alternative proof of Claim 4.4) For convenience, let us use ± 1 for binary values and consider the sequence of i.i.d. binary random variables $Y_1, \ldots, Y_k$, with $Y_i \in \{-1, 1\}$ and $\mathrm{E}[Y_i] = 0$ for each i. For $I \subseteq [k]$, let $\tilde{Y}_I$ denote the sum $\sum_{i \in I} Y_i$. For an integer v, let $\mathrm{sign}(v) = 1$ if $v \geq 0$ and $\mathrm{sign}(v) = 0$ if $v < 0$. Then, we have

$$\Pr_G [G_I(x) \neq G_J(x)] = \Pr\left[\mathrm{sign}(\tilde{Y}_I) \neq \mathrm{sign}(\tilde{Y}_J)\right].$$

Let $K = I \cap J$, $I_1 = I \setminus K$, and $J_1 = J \setminus K$, and note that $|K| < |I_1|, |J_1|$. Then

$$\begin{aligned}&\Pr\left[\mathrm{sign}(\tilde{Y}_I) \neq \mathrm{sign}(\tilde{Y}_J)\right] \\ &\quad \geq \sum_v \Pr\left[\tilde{Y}_K = v\right] \\ &\qquad \cdot \Pr\left[\mathrm{sign}(\tilde{Y}_I) = \mathrm{sign}(v) \wedge \mathrm{sign}(\tilde{Y}_J) \neq \mathrm{sign}(v) \mid \tilde{Y}_K = v\right] \\ &\quad = \sum_v \Pr\left[\tilde{Y}_K = v\right] \cdot \Pr\left[\mathrm{sign}(\tilde{Y}_I) = \mathrm{sign}(v) \mid \tilde{Y}_K = v\right] \\ &\qquad \cdot \Pr\left[\mathrm{sign}(\tilde{Y}_J) \neq \mathrm{sign}(v) \mid \tilde{Y}_K = v\right].\end{aligned}$$

Note that conditioned on $\tilde{Y}_K = v$, the probability of $\mathrm{sign}(\tilde{Y}_I) = \mathrm{sign}(v)$ is at least $\Pr[\tilde{Y}_{I_1} \geq 0]$ when $v \geq 0$ and at least $\Pr[\tilde{Y}_{I_1} \leq 0]$ when $v < 0$, and we know by symmetry that $\Pr[\tilde{Y}_{I_1} \geq 0] = \Pr[\tilde{Y}_{I_1} \leq 0] \geq \frac{1}{2}$. Thus, for any v,

$$\Pr\left[\mathrm{sign}(\tilde{Y}_I) = \mathrm{sign}(v) \mid \tilde{Y}_K = v\right] \geq \frac{1}{2},$$

and consequently,

$$\begin{aligned}\Pr\left[\operatorname{sign}(\tilde{Y}_I) \neq \operatorname{sign}(\tilde{Y}_J)\right] &\geq \sum_v \Pr\left[\tilde{Y}_K = v\right] \cdot \frac{1}{2} \\ &\quad \cdot \Pr\left[\operatorname{sign}(\tilde{Y}_J) \neq \operatorname{sign}(v) \mid \tilde{Y}_K = v\right] \\ &= \frac{1}{2} \cdot \Pr\left[\operatorname{sign}(\tilde{Y}_J) \neq \operatorname{sign}(\tilde{Y}_K)\right].\end{aligned}$$

Now consider any $J_2 \subseteq J_1$ of the same size as K, and note that $\operatorname{sign}(\tilde{Y}_J) \neq \operatorname{sign}(\tilde{Y}_K)$ whenever all the following three events hold.

- B_1: The absolute value of $\tilde{Y}_{J_2}$ is at least that of $\tilde{Y}_K$.
- B_2: $\operatorname{sign}(\tilde{Y}_{J_2}) \neq \operatorname{sign}(\tilde{Y}_K)$ or $\tilde{Y}_{J_2} = 0$.
- B_3: $\operatorname{sign}(\tilde{Y}_{J_1 \setminus J_2}) \neq \operatorname{sign}(\tilde{Y}_K)$.

Since J_2 has the same size as K, $\tilde{Y}_{J_2}$ and $\tilde{Y}_K$ each are equally likely to have an absolute value at least that of the other by symmetry, which implies that $\Pr[B_1] \geq \frac{1}{2}$. Next, $\Pr[B_2 \mid B_1]$ is at least $\Pr[\tilde{Y}_{J_2} \geq 0 \mid B_1]$ when $\operatorname{sign}(\tilde{Y}_K) = 0$ and at least $\Pr[\tilde{Y}_{J_2} \leq 0 \mid B_1]$ when $\operatorname{sign}(\tilde{Y}_K) = 1$. Observe that both conditional probabilities are the same by symmetry, so we have $\Pr[B_2 \mid B_1] \geq \frac{1}{2}$. Finally, since $\tilde{Y}_{J_1 \setminus J_2}$ is independent of $\tilde{Y}_{J_2}$ and $\tilde{Y}_K$, we have $\Pr[B_3 \mid B_1 \wedge B_2] = \Pr[B_3]$, which is at least $\Pr[\tilde{Y}_{J_1 \setminus J_2} > 0]$ when $\operatorname{sign}(\tilde{Y}_K) = 0$ and at least $\Pr[\tilde{Y}_{J_1 \setminus J_2} < 0]$ when $\operatorname{sign}(\tilde{Y}_K) = 1$. Note that both probabilities are the same by symmetry, and they are equal to $\frac{1}{2}(1 - \Pr[\tilde{Y}_{J_1 \setminus J_2} = 0]) \geq \frac{1}{4}$ because $|J_1 \setminus J_2| \geq 1$ and $\Pr[\tilde{Y}_{J_1 \setminus J_2} = 0] \leq \frac{1}{2}$, which implies that $\Pr[B_3] \geq \frac{1}{4}$. As a result, we have

$$\begin{aligned}\Pr\left[\operatorname{sign}(\tilde{Y}_J) \neq \operatorname{sign}(\tilde{Y}_K)\right] &\geq \Pr[B_1 \wedge B_2 \wedge B_3] \\ &= \Pr[B_1] \cdot \Pr[B_2 \mid B_1] \cdot \Pr[B_3 \mid B_1 \wedge B_2] \\ &\geq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{4},\end{aligned}$$

which then implies that

$$\Pr_G[G_I(x) \neq G_J(x)] = \Pr[\operatorname{sign}(\tilde{Y}_I) \neq \operatorname{sign}(\tilde{Y}_J)] \geq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{32}.$$

□

From this claim, we next show that G_I and G_J are unlikely to agree on many x. For any $x \in \{0,1\}^n$, consider the binary random variable Y_x such that $Y_x = 1$ if and only if $G_I(x) \neq G_J(x)$. From the above claim, we know that $\mathrm{E}_G[Y_x] \geq c_0$ for some constant c_0. So by Chernoff bound, we have

$$\Pr_G[G_I \text{ is } 2\delta\text{-close to } G_J] = \Pr\left[\sum_x Y_x \leq 2\delta 2^n\right] \leq 2^{-\Omega((c_0-2\delta)^2 2^n)} \leq 2^{-\Omega(2^n)},$$

as we assume that $\delta < c_2$ for a small enough constant c_2. This completes the proof of Lemma 4.3. □

Call G nice if for any $I, J \in V^t$ with $|I \cap J| < \frac{t}{2}$, G_I is not 2δ-close to G_J. By a union bound,

$$\Pr_G[G \text{ is not nice}] \leq \binom{k}{t}^2 \cdot 2^{-\Omega(2^n)} \leq 2^{2t\log k} \cdot 2^{-\Omega(2^n)} < 1,$$

as we assume that $t \leq \frac{1}{\varepsilon} \leq 2^{c_1 n}$ and $k \leq 2^{2^{c_3 n}}$, for small enough constants $c_1, c_3 > 0$. This guarantees the existence of a nice G. From now on, we will fix on one such G.

Consider the undirected graph $\mathcal{G} = (V, E)$ where $V = \{G_I : I \in V^t\}$ and E consists of those pairs of G_I, G_J which are 2δ-close to each other. Then, we have the following.

Lemma 4.5. *$\mathcal{G}$ has an independent set of size at least $k^{\Omega(t)}$.*

Proof. Since G is nice, there cannot be an edge between vertices G_I and G_J if $|I \cap J| < \frac{t}{2}$. Thus, the degree of any vertex G_I is at most the number of G_J with $|I \cap J| \geq \frac{t}{2}$, which is at most

$$\sum_{\frac{t}{2} \leq i < t} \binom{t}{i}\binom{k-t}{t-i} \leq \sum_{\frac{t}{2} \leq i < t} \binom{t}{i}\binom{k}{\frac{t}{2}} \leq 2^t \binom{k}{\frac{t}{2}} \leq \left(\frac{8ek}{t}\right)^{t/2} \leq k^{t/2},$$

where the first and last inequalities, respectively, hold under the conditions that $k \geq \frac{1}{\varepsilon^3} \geq t^3$ and $t = \left\lfloor \frac{1}{\varepsilon} \right\rfloor$ is at least a large enough constant, while the third inequality uses the fact that $\binom{n}{m} \leq (\frac{en}{m})^m$.

Then by Turán's theorem (Fact 2.2), $\mathcal{G}$ has an independent set of size

$$\binom{k}{t}\frac{1}{k^{t/2}+1} \geq \left(\frac{k}{t}\right)^t \frac{1}{k^{t/2}+1} \geq k^{2t/3}\frac{1}{k^{t/2}+1} \geq k^{\Omega(t)},$$

where the second inequality follows from the assumption that $k \geq t^3$. □

Now we are ready to finish the proof of the theorem. From Lemma 4.5, we know that $\mathcal{G}$ has an independent set Γ of size $k^{\Omega(t)}$. Note that any two $G_I, G_J \in \Gamma$ are not 2δ-close. Furthermore, we know from Lemma 4.2 that every $G_I \in \Gamma$ is $(c\delta, \varepsilon, G)$-easy, since $|I| = t \leq \frac{1}{\varepsilon}$. Therefore, an advice string of length $\log|\Gamma| = \Omega(t \log k) = \Omega(\frac{1}{\varepsilon}\log k)$ is required, because for each advice string α, $\textsc{Dec}^{G,\alpha}$ can only be δ-close to at most one $G_I \in \Gamma$. This proves Theorem 4.1.

5. No Weakly Black-Box Proof in $\mathsf{AC}^0[p]$

In this section, we show that no weakly black-box hard-core set proof can be implemented in $\mathsf{AC}^0[p]$. More precisely, we have the following.

Theorem 5.1. *Suppose* $1 < \delta < 1/20$, $0 < \varepsilon < 1$, $k \geq n$, *and* p *a prime. Let* $t = \min(\lfloor 1/\varepsilon \rfloor, n)$. *Then, no weakly black-box* (δ, ε, k)*-proof can be implemented in* $\mathsf{AC}^0[p](2^{\mathrm{poly}(\log t)})$.

The idea is the following. Suppose we have a function $\textsc{Dec}$ realizing such a black-box proof. Let $I = [t]$ and note that $1/t \geq \varepsilon$. From the previous section, we know that for any G, the function G_I is $(c\delta, \varepsilon, G)$-easy (with c the constant associated with $\textsc{Dec}$), so there must exist some advice α such that $\textsc{Dec}^{G,\alpha}$ is δ-close to the function G_I, which is the majority function over $g_1, \ldots, g_t$. As will be shown later, by defining G properly, we can use $\textsc{Dec}^{G,\alpha}$ to approximate the majority function on t input variables. Then, we need the following lower bound on the majority function.

Lemma 5.2. *For any* $C : \{0,1\}^t \to \{0,1\}$ *in* $\mathsf{AC}^0[p](2^{\mathrm{poly}(\log t)})$ *and for a large enough* t, *we have* $\Pr_x[C(x) \neq \textsc{Maj}(x)] \geq 1/20$.

PROOF. From Smolensky (1987), we know that for any function $C : \{0,1\}^t \to \{0,1\}$ in $\mathsf{AC}^0[p](2^{\mathrm{poly}(\log t)})$, there is a polynomial Q over $GF(p)$ of degree $\mathrm{poly}(\log t)$ such that $\Pr_x[C(x) \neq Q(x)] \leq 2^{-\mathrm{poly}(\log t)}$. From Szegedy (1989), Tarui (1991), we know that for any such polynomial Q, for a large enough t, $\Pr_x[Q(x) \neq \textsc{Maj}(x)] \geq 1/10$, so we have $\Pr_x[C(x) \neq \textsc{Maj}(x)] \geq 1/10 - 2^{-\mathrm{poly}(\log t)} \geq 1/20$. □

We define the function g_i as $g_i(x) = x_i$ for $i \in [n]$ and $g_i(x) = 0$ otherwise, for $x \in \{0,1\}^n$. Let $G = \{g_1, \ldots, g_k\}$. Then $G_I(x) = \textsc{Maj}(x_1, \ldots, x_t)$ for any $x \in \{0,1\}^n$, so there must be some advice α such that $\Pr_x[\textsc{Dec}^{G,\alpha}(x) \neq \textsc{Maj}(x_1, \ldots, x_t)] \leq \delta$, and by an averaging argument there must be some fixed $\bar{x}_{t+1}, \ldots, \bar{x}_n$ such that

$$\Pr_{x_1,\ldots,x_t}[\textsc{Dec}^{G,\alpha}(x_1, \ldots, x_t, \bar{x}_{t+1}, \ldots, \bar{x}_n) \neq \textsc{Maj}(x_1, \ldots, x_t)] \leq \delta.$$

Such α and $\bar{x}_{t+1}, \ldots, \bar{x}_n$ can be hard-wired into the circuit for DEC and observe that all the oracle gates can be removed as every oracle query can be answered by some input bit x_i or a fixed constant. So if $\textsc{Dec}^G$ can be computed by an $\mathsf{AC}^0[p](2^{\mathrm{poly}(\log t)})$ circuit equipped with oracle gates from G, we can obtain from it an $\mathsf{AC}^0[p](2^{\mathrm{poly}(\log t)})$ circuit (without oracle gates) which is δ-close to the majority function on t bits and contradicts Lemma 5.2. This proves Theorem 5.1.

Acknowledgements

The authors would like to thank anonymous referees for their useful comments. The research of Chi-Jen Lu was supported in part by the National Science Council under the Grant NSC97-2221-E-001-012-MY3. The research of Shi-Chun Tsai was supported in part by the National Science Council of Taiwan under contracts NSC-97-2221-E-009-064-MY3 and NSC-98-2221-E-009-078-MY3. The research of Hsin-Lung Wu was supported in part by the National Science Council of Taiwan under contracts NSC-97-2218-E-305-001-MY2.

References

Noga Alon & Joel H. Spencer (2000). *The Probabilistic Method*. Wiley, New York, 2nd edition.

László Babai, Lance Fortnow, Noam Nisan & Avi Wigderson (1993). BPP has subexponential time simulations unless EXPTIME has publishable proofs. *Computational Complexity* **3**(4), 307–318.

Boaz Barak, Moritz Hardt & Satyen Kale (2009). The uniform hardcore lemma via approximate Bregman projections. In *Proceedings of the 20th Annual ACM-SIAM Symposium on Discrete Algorithms*, 1193–1200.

Yoav Freund (1995). Boosting a Weak Learning Algorithm by Majority. *Information and Computation* **121**(2), 256–285.

Mikael Goldmann, Johan Hastad & Alexander A. Razborov (1992). Majority Gates VS. General Weighted Threshold Gates. *Computational Complexity* **2**, 277–300.

Andras Hajnal, Wolfgang Maass, Pavel Pudlak, Mario Szegedy & Gyorgy Turan (1993). Threshold Circuits of Bounded Depth. *Journal of Computer System Sciences* **46**(2), 129–154.

Alexander Healy, Salil P. Vadhan & Emanuele Viola (2006). Using Nondeterminism to Amplify Hardness. *SIAM Journal of Computing* **35**(4), 903–931.

Thomas Holenstein (2005). Key agreement from weak bit agreement. In *Proceedings of the 37th ACM Symposium on Theory of Computing*, 664–673.

Russell Impagliazzo (1995). Hard-Core Distributions for Somewhat Hard Problems. In *Proceedings of the 36th Annual IEEE Symposium on Foundations of Computer Science*, 538–545.

Russell Impagliazzo & Avi Wigderson (1997). P = BPP if E Requires Exponential Circuits: Derandomizing the XOR Lemma. In *Proceedings of the 29th ACM Symposium on Theory of Computing*, 220–229.

Adam R. Klivans & Rocco A. Servedio (2003). Boosting and Hard-Core Set Construction. *Machine Learning* **51**(3), 217–238.

Chi-Jen Lu, Shi-Chun Tsai & Hsin-Lung Wu (2007). On the Complexity of Hard-Core Set Constructions. In *Proceedings of the 34th International Colloquium on Automata, Languages and Programming*, 183–194.

Chi-Jen Lu, Shi-Chun Tsai & Hsin-Lung Wu (2008). On the Complexity of Hardness Amplification. *IEEE Transactions on Information Theory* **54**(10), 4575–4586.

Ryan O'Donnell (2004). Hardness amplification within NP. *Journal of Computer System Sciences* **69**(1), 68–94.

Gilles Pisier (1981). Remarques sur un resultat non publié de B. Maurey. *Seminaire Analyse fonctionnelle* **1**(12), 1980–1981.

Ronen Shaltiel & Emanuele Viola (2008). Hardness amplification proofs require majority. In *Proceedings of the 40th annual ACM symposium on Theory of computing*, 589–598.

Roman Smolensky (1987). Algebraic methods in the theory of lower bounds for Boolean circuit complexity. In *Proceedings of the 19th annual ACM symposium on Theory of computing*, 77–82.

Madhu Sudan, Luca Trevisan & Salil Vadhan (2001). Pseudorandom generators without the XOR lemma. *Journal of Computer System Sciences* **62**(2), 236–266.

Márió Szegedy (1989). *Algebraic methods in lower bounds for computational models with limited communication.* Ph.D. thesis.

Jun Tarui (1991). Degree complexity of Boolean functions and its applications to relativized separations. In *Proceedings of the 6th Annual IEEE Conference on Structure in Complexity Theory*, 382–390.

Luca Trevisan (2003). List-Decoding Using The XOR Lemma. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, 126–135.

Luca Trevisan (2005). On uniform amplification of hardness in NP. In *Proceedings of the 37th ACM Symposium on Theory of Computing*, 31–38.

Pál Turán (1941). On an extremal problem in graph theory. *Matematikai és Fizikai Lapok* **48**, 436–452.

Emanuele Viola (2006). *The complexity of hardness amplification and derandomization*. Ph.D. thesis.

Andrew C. Yao (1982). Theory and application of trapdoor functions. In *Proceedings of the 23rd Annual Symposium on Foundations of Computer Science*, 80–91.

Manuscript received 10 November 2007

Chi-Jen Lu
Institute of Information Science,
Academia Sinica,
Taipei, Taiwan.
cjlu@iis.sinica.edu.tw

Shi-Chun Tsai
Department of Computer Science,
National Chiao Tung University,
Hsinchu, Taiwan.
sctsai@csie.nctu.edu.tw

Hsin-Lung Wu
Department of Computer Science
and Information Engineering,
National Taipei University,
Taipei, Taiwan.
hsinlung@mail.ntpu.edu.tw