Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 191 (2017)

Prediction from Partial Information and Hindsight, an Alternative Proof

Alexander V. Smal* Navid Talebanfard†

Abstract

Let X be a random variable distributed over n-bit strings with $H(X) \geq n - k$, where $k \ll n$. Using subadditivity we know that the average coordinate has high entropy. Meir and Wigderson [1] showed that a random coordinate looks random to an adversary who is allowed to query around n/k other coordinates non-deterministically. They used this result to obtain top-down arguments in depth-3 circuit lower bounds. In this note we give an alternative proof of their main result which tightens their parameters. Our proof is inspired by a paper of Paturi, Pudlák and Zane [3] who gave a non-trivial k-SAT algorithm and tight depth-3 circuit lower bounds for parity.

1 Introduction

Motivated by developing top-down arguments in circuit complexity, Meir and Wigderson [1] recently studied the following question: let $X = (X_1, \ldots, X_n) \in \{0,1\}^n$ be a random variable with entropy at least $n - k$. An adversary who knows the distribution of X and a uniformly chosen coordinate $i \in [n]$ needs to predict the value of X_i. He is allowed to query q coordinates of X other than i. How large should q be so that the adversary has non-negligible advantage? The answer is $\omega(n/k)$ and it holds even if the adversary can make his queries non-deterministically. This is formally captured as follows.

Definition 1. A *witness* for a coordinate $i \in [n]$ is a pair (Q, a) where $Q \subseteq [n] \backslash \{i\}$ and $a \in \{0,1\}^{|Q|}$. The witness *appears* in a string $x \in \{0,1\}^n$ if $x|_Q = a$. The *length* of the witness is $|Q|$.

Definition 2. A *q-family of witnesses* $\mathcal{F}$ for a coordinate $i \in [n]$ is a set of witnesses for i of length at most q. We say that a string $X \in \{0,1\}^n$ *satisfies* $\mathcal{F}$, denoted by $X \models \mathcal{F}$, if at least one of the witnesses in $\mathcal{F}$ appears in X. For a random string $X \in \{0,1\}^n$, a bit $b \in \{0,1\}$ and $0 \leq \epsilon \leq 1$, we say that $\mathcal{F}$ *ϵ-predicts* $X_i = b$ if

$$\Pr[X_i = b \mid X \text{ satisfies } \mathcal{F}] \geq \frac{1}{2} + \frac{\epsilon}{2}.$$

Theorem 1 ([1])**.** *Let X be a random variable over $\{0,1\}^n$ such that $H(X) \geq n - k$, and $q \leq n$. For $\epsilon \in (0,1]$, every $i \in [n]$ and $b \in \{0,1\}$ let $\mathcal{F}_i^b$ be a q-family of witnesses that ϵ-predicts $X_i = b$ and let σ_i be the probability that either $X \models \mathcal{F}_i^0$ or $X \models \mathcal{F}_i^1$. Then $\bar{\sigma} := \mathbb{E}_i[\sigma_i]$ is at most $\frac{300 \cdot k \cdot q}{\epsilon^3 \cdot n}$.*

*St. Petersburg Department of Steklov Mathematical Institute of Russian Academy of Sciences, avsmal@gmail.com.

†Institute of Mathematics of the Czech Academy of Sciences, talebanfard@math.cas.cz. Supported by ERC Advanced Grant 339691 (FEALORA).

ISSN 1433-8092

The application of this theorem in circuit lower bounds uses $\epsilon = 1$ (see [1]). The focus of this note is an improvement of Theorem 1 for this case and the background which gave rise to the proof, interestingly coming from similar questions in circuit lower bounds.

Definition 3. Given a random string $X \in \{0,1\}^n$, a coordinate $i \in [n]$ and a bit $b \in \{0,1\}$, a *b-certificate* for i is a witness (Q, a) such that

$$\Pr[X_i = b \mid X|_Q = a] = 1.$$

Theorem 2. *Let X be a random variable over $\{0,1\}^n$ such that $H(X) \geq n - k$, and $q \leq n$. For every $i \in [n]$ and $b \in \{0,1\}$ let σ_i be the probability that X contains any certificate for X_i. Then $\bar{\sigma} := \mathbb{E}_i[\sigma_i]$ is at most $\frac{k \cdot (q+1)}{n}$.*

It is easy to compare this theorem with Theorem 1 for case $\epsilon = 1$. Note that the bound in Theorem 2 is tight. For any $\sigma \in [0,1]$ we can consider a random variable $X = (X_1, X_2, \ldots, X_n)$ defined by the following process: group the first σn coordinates in blocks of size $q+1$. Then for every block assign all coordinates except the last one uniformly at random and set the last one to be the xor of the previous ones. Finally assign all remaining coordinates $X_{\sigma n+1}, X_{\sigma n+2}, \ldots, X_n$ uniformly at random. It is easy to see that

$$H(X) = \frac{\sigma n}{q+1} \cdot q + (1-\sigma) \cdot n = n - \frac{\sigma n}{q+1}.$$

Applying Theorem 2 with $k = \sigma n/(q+1)$ we get $\bar{\sigma} \leq \sigma$. On the other hand each of the first σn coordinates has certificate (with probability 1) of size q that predicts it, hence $\bar{\sigma} \geq \sigma$.

2 Background for the proof

Let ψ be a k-CNF formula in n variables. We say that a satisfying assignment α is *isolated* if flipping any single bit of α falsifies some clause. Paturi, Pudlák and Zane [3] showed that the number of isolated solutions is upper bounded by $2^{(1-1/k)n}$. This immediately implies a $2^{n/k-1}$ lower bound for parity against Σ_k^3 circuits (depth-3 OR-AND-OR circuits with bottom fan-in bounded by k). Furthermore since they proved this via an efficient encoding, they also managed to obtain an algorithm for k-SAT. Following a similar line of reasoning Paturi, Pudlák, Saks and Zane [2] showed that if every pair of solutions of a k-CNF disagree on a super-constant number of bits, the number of satisfying assignments is upper bounded by $2^{(1-\pi^2/6k+o(1))n}$. This subsequently gave an improved lower bound for error-correcting codes with proper parameters and an improved k-SAT algorithm which still remains the best known. The best Σ_k^3 lower bounds are $2^{\Omega(n/k)}$ and the best upper bound bound for k-SAT is $2^{(1-\Omega(1/k))n}$. Improving the dependence on k remains a major challenge in complexity theory. One way to make progress in this direction would be to make the arguments of [3] and [2] more flexible through replacing the efficient encoding by direct entropy arguments. This brings us to the result of Meir and Wigderson.

Let X be a uniformly chosen isolated solution of ψ. We can generate a $(k-1)$-family of certificates $\mathcal{F}$ for X as follows. For every isolated solution α and every $i \in [n]$, flip the ith bit. This falsifies some clause $C_{\alpha,i}$ which means that under α, $C_{\alpha,i}$ is satisfied only by x_i. We then add the set of variables other than x_i in $C_{\alpha,i}$ and their values under α to $\mathcal{F}_i$. By the construction of $\mathcal{F}$ for every $i \in [n]$ we have that $\Pr[X \models \mathcal{F}_i] = 1$ and thus $\mathbb{E}_i[\sigma_i] = 1$. Theorem 1 then implies

that $H(X) \leq (1 - 1/300k)n$, since otherwise we would have $\mathbb{E}_i[\sigma_i] < 1$. However [3] gives the tight bound $H(X) \leq (1 - 1/k)n$. This suggests that it might be possible to improve Theorem 1 and this is what we achieved.

Our proof is inspired by the argument of [3] and it bounds the entropy by considering random permutations of the bits, whereas the proof of [1] considers random splittings of X in two parts. Unfortunately our improvement does not imply any new result. However we would like to stress our original motivation which led to the current proof and pose the question of reproving or improving the circuit lower bound of [2] as follows.

Question 1. *Let ψ be a k-CNF whose satisfying assignments are pairwise of super-constant Hamming distance. Let X be uniformly distributed over the set of satisfying assignments. Can we upper bound $H(X)$ directly without adhering to the explicit encoding of [2]?*

3 Proof of Theorem 2

Let $X = (X_1, \ldots, X_n)$ be a random variable over $\{0,1\}^n$. For $i \in [n]$ we will use $X_{<i}$ as a shorthand for $(X_1, \ldots, X_{i-1})$ and let $\mathcal{F}_i$ be the set of all certificates for X_i.

Using the chain rule one can derive the following equality

$$H(X) = H(X_1) + H(X_2 \mid X_1) + \cdots + H(X_n \mid X_{<n}).$$

The same equality holds if we permute the coordinates of X:

$$H(X) = H(X^\pi) = H(X_1^\pi) + H(X_2^\pi \mid X_1^\pi) + \cdots + H(X_n^\pi \mid X_{<n}^\pi).$$

where $\pi \in S_n$ is a permutation of $[n]$ and X^π is a random variable equal to X permuted with π, i.e. $X^\pi = (X_1^\pi, X_2^\pi, \ldots, X_n^\pi) = (X_{\pi(1)}, X_{\pi(2)}, \ldots, X_{\pi(n)})$.

Lemma 1. *For any $i \in [n]$ and $\alpha \in \{0,1\}^n$ such that α contains a certificate (Q, a) for X_i of size at most q,*

$$\mathop{\mathbb{E}}_{\pi}[H(X_i \mid X^\pi_{<\pi^{-1}(i)} = \alpha^\pi_{<\pi^{-1}(i)})] \leq 1 - \frac{1}{q+1}.$$

Proof. If for a fixed permutation π all indices in Q precede $\pi^{-1}(i)$, i.e., $\forall j \in Q, \pi^{-1}(j) < \pi^{-1}(i)$, then $H(X_i \mid X^\pi_{<\pi^{-1}(i)} = \alpha^\pi_{<\pi^{-1}(i)}) = H(X_i \mid X|_Q = a) = 0$ by the definition of certificate. We have

$$\begin{aligned}
\mathop{\mathbb{E}}_{\pi}[H(X_i \mid X^\pi_{<\pi^{-1}(i)} = \alpha^\pi_{<\pi^{-1}(i)})] \leq{}& \Pr_\pi[\forall j \in Q, \pi^{-1}(j) < \pi^{-1}(i)] \cdot H(X_i \mid X|_Q = a) \\
&+ \Pr_\pi[\exists j \in Q, \pi^{-1}(j) > \pi^{-1}(i)] \cdot H(X_i) \\
={}& \left(1 - \frac{1}{|Q|+1}\right) \cdot H(X_i) \\
\leq{}& 1 - \frac{1}{|Q|+1} \leq 1 - \frac{1}{q+1}.
\end{aligned}$$

Hence we get the statement of the lemma. □

Let us bound the expectation of $H(X)$ over all permutations of $[n]$.

$$\begin{aligned}
\mathop{\mathbb{E}}_{\pi}[H(X^\pi)] &= \mathop{\mathbb{E}}_{\pi}\left[\sum_{i=1}^{n} H(X_i^\pi \mid X_{<i}^\pi)\right] && \text{(chain rule)}\\
&= \mathop{\mathbb{E}}_{\pi}\left[\sum_{i=1}^{n} \sum_{\beta\in\{0,1\}^{i-1}} \Pr[X_{<i}^\pi = \beta]\cdot H(X_i^\pi \mid X_{<i}^\pi = \beta)\right] && \text{(expansion)}\\
&= \mathop{\mathbb{E}}_{\pi}\left[\sum_{i=1}^{n} \sum_{\beta\in\{0,1\}^{i-1}} \sum_{\substack{\alpha\in\{0,1\}^n\\ \alpha_{<i}^\pi=\beta}} \Pr[X=\alpha]\cdot H(X_i^\pi \mid X_{<i}^\pi = \beta)\right] && \text{(expansion)}\\
&= \mathop{\mathbb{E}}_{\pi}\left[\sum_{i=1}^{n} \sum_{\alpha\in\{0,1\}^n} \Pr[X=\alpha]\cdot H(X_i^\pi \mid X_{<i}^\pi = \alpha_{<i}^\pi)\right] && (\alpha \text{ extends } \beta)\\
&= \mathop{\mathbb{E}}_{\pi}\left[\sum_{j=1}^{n} \sum_{\alpha\in\{0,1\}^n} \Pr[X=\alpha]\cdot H(X_j \mid X_{<\pi^{-1}(j)}^\pi = \alpha_{<\pi^{-1}(j)}^\pi)\right] && (j=\pi(i))\\
&= \mathop{\mathbb{E}}_{\pi}\Bigg[\sum_{j=1}^{n} \sum_{\substack{\alpha\in\{0,1\}^n\\ \alpha\models\mathcal{F}_j}} \Pr[X=\alpha]\cdot H(X_j \mid X_{<\pi^{-1}(j)}^\pi = \alpha_{<\pi^{-1}(j)}^\pi)\\
&\qquad + \sum_{j=1}^{n} \sum_{\substack{\alpha\in\{0,1\}^n\\ \alpha\not\models\mathcal{F}_j}} \Pr[X=\alpha]\cdot H(X_j \mid X_{<\pi^{-1}(j)}^\pi = \alpha_{<\pi^{-1}(j)}^\pi)\Bigg]
\end{aligned}$$

By the linearity of expectation, we have

$$\begin{aligned}
\mathop{\mathbb{E}}_{\pi}[H(X^\pi)] &= \sum_{j=1}^{n} \sum_{\substack{\alpha\in\{0,1\}^n\\ \alpha\models\mathcal{F}_j}} \Pr[X=\alpha]\cdot \mathop{\mathbb{E}}_{\pi}[H(X_j \mid X_{<\pi^{-1}(j)}^\pi = \alpha_{<\pi^{-1}(j)}^\pi)]\\
&\quad + \sum_{j=1}^{n} \sum_{\substack{\alpha\in\{0,1\}^n\\ \alpha\not\models\mathcal{F}_j}} \Pr[X=\alpha]\cdot \mathop{\mathbb{E}}_{\pi}[H(X_j \mid X_{<\pi^{-1}(j)}^\pi = \alpha_{<\pi^{-1}(j)}^\pi)]
\end{aligned}$$

Applying Lemma 1 we get

$$\begin{aligned}\mathop{\mathbb{E}}_{\pi}[H(X^{\pi})] &\leq \sum_{j=1}^{n} \sum_{\substack{\alpha\in\{0,1\}^n \\ \alpha\models\mathcal{F}_j}} \Pr[X=\alpha]\cdot\Big(1-\frac{1}{q+1}\Big) + \sum_{j=1}^{n} \sum_{\substack{\alpha\in\{0,1\}^n \\ \alpha\not\models\mathcal{F}_j}} \Pr[X=\alpha]\cdot 1 \\ &= \sum_{j=1}^{n}\left(\sigma_j\cdot\Big(1-\frac{1}{q+1}\Big)+(1-\sigma_j)\cdot 1\right) = \sum_{j=1}^{n}\Big(1-\frac{\sigma_i}{q+1}\Big) \\ &= n-\frac{n\cdot\bar{\sigma}}{q+1}.\end{aligned}$$

Having in mind that $\mathbb{E}_{\pi}[H(X^{\pi})] = \mathbb{E}_{\pi}[H(X)] = H(X) \geq n-k$ we get $\bar{\sigma} \leq \frac{k\cdot(q+1)}{n}$.

References

[1] O. Meir and A. Wigderson. Prediction from partial information and hindsight, with application to circuit lower bounds. *Electronic Colloquium on Computational Complexity (ECCC)*, 24:149, 2017.

[2] R. Paturi, P. Pudlák, M. E. Saks, and F. Zane. An improved exponential-time algorithm for k-SAT. *J. ACM*, 52(3):337–364, 2005.

[3] R. Paturi, P. Pudlák, and F. Zane. Satisfiability coding lemma. *Chicago J. Theor. Comput. Sci.*, 1999, 1999.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il