Electronic Colloquium on Computational Complexity, Report No. 168 (2010)

Pseudorandom Generators for Combinatorial Checkerboards

Thomas Watson*

November 9, 2010

Abstract

We define a combinatorial checkerboard to be a function $f : \{1, \ldots, m\}^d \to \{1, -1\}$ of the form $f(u_1, \ldots, u_d) = \prod_{i=1}^{d} f_i(u_i)$ for some functions $f_i : \{1, \ldots, m\} \to \{1, -1\}$. This is a variant of combinatorial rectangles, which can be defined in the same way but using $\{0, 1\}$ instead of $\{1, -1\}$. We consider the problem of constructing explicit pseudorandom generators for combinatorial checkerboards. This is a generalization of small-bias generators, which correspond to the case $m = 2$.

We construct a pseudorandom generator that ϵ-fools all combinatorial checkerboards with seed length $O\big(\log m + \log d \cdot \log\log d + \log^{3/2} \frac{1}{\epsilon}\big)$. Previous work by Impagliazzo, Nisan, and Wigderson implies a pseudorandom generator with seed length $O\big(\log m + \log^2 d + \log d \cdot \log \frac{1}{\epsilon}\big)$. Our seed length is better except when $\frac{1}{\epsilon} \geq d^{\omega(\log d)}$.

1 Introduction

A central question in the theory of computation is whether randomized algorithms are more powerful than deterministic algorithms. Some computational problems, such as testing whether a succinctly described polynomial is the zero polynomial, have efficient randomized algorithms but are not known to have efficient deterministic algorithms. On the other hand, a line of research in complexity theory [59, 11, 36, 70, 35, 68, 71] has shown that under widely believed conjectures (namely the existence of nonuniformly hard functions in certain uniform complexity classes), every polynomial-time randomized algorithm solving a decision problem can be *derandomized* to yield a polynomial-time deterministic algorithm solving the same decision problem. These proofs proceed by using the hypothesized hard function to construct an efficient *pseudorandom generator*, which is an algorithm that stretches a short truly random string (the seed) to a long "pseudorandom" string that is indistinguishable from a long truly random string by any efficient algorithm. Provided the seed is short enough, one can then cycle over all the seeds in polynomial time, running the randomized algorithm using the output of the pseudorandom generator for the randomness, to get a polynomial-time deterministic algorithm for the same decision problem.

Unfortunately, there are no known results that shed light on how to unconditionally construct pseudorandom generators that fool arbitrary polynomial-time randomized algorithms. Furthermore, there is formal evidence suggesting that unconditionally derandomizing arbitrary polynomial-time algorithms is far beyond the reach of current techniques, even if we do not insist on using a pseudorandom generator [33, 32, 1, 28].

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

ISSN 1433-8092

In light of these barriers, a natural goal is to unconditionally construct pseudorandom generators with good seed lengths for restricted classes of functions. One such class of functions is those computed by *small-width read-once branching programs*, which model space-bounded computations. The theory of pseudorandomness for space-bounded computations has a long and rich history [2, 12, 56, 55, 57, 58, 60, 34, 67, 10, 8, 61, 21, 63, 66, 26, 64, 49, 42, 16, 19, 20, 69, 38], including very general results as well as improved results for special cases. One such special case is *linear functions over* $\mathbb{Z}_2$ [53, 6, 52]. Pseudorandom generators for this class of functions are called small-bias generators. It is known how to construct small-bias generators whose seed lengths are optimal up to constant factors [53, 6]. Another special case that has been considered is *combinatorial rectangles* [25, 39, 9, 43].

We consider the problem of constructing an explicit pseudorandom generator for a new class of functions, which we dub *combinatorial checkerboards*. These functions can be viewed as

- a *special case* of small-width read-once branching programs,
- a *generalization* of linear functions over $\mathbb{Z}_2$, and
- a *variant* of combinatorial rectangles.

Other classes of functions for which constructions of good pseudorandom generators are known include juntas [53, 6], constant-depth circuits [3, 54, 46, 45, 72, 13, 62, 18, 22, 41], low-degree polynomials [46, 72, 15, 17, 40, 73], and polynomial threshold functions [23, 50, 30, 31, 24].

1.1 Combinatorial Checkerboards

We give four equivalent ways of defining combinatorial checkerboards, which are parameterized by two positive integers m and d. Recall that $[m]$ denotes the set of integers $\{1, 2, \ldots, m\}$. For us, it is not important that the elements are integers; we only use $[m]$ as an arbitrary set of size m.

(1) A combinatorial checkerboard can be defined as a subset of $[m]^d$ of the following form. There are sets $S_1, \ldots, S_d \subseteq [m]$ such that a point $(u_1, \ldots, u_d) \in [m]^d$ is in the checkerboard if and only if the number of coordinates i such that $u_i \in S_i$ is odd. (See Figure 1, and note that unlike the example in the figure, the set S_i need not be a contiguous interval.) In contrast, a combinatorial rectangle can be defined similarly but where a point is in the rectangle if and only if $u_i \in S_i$ holds for all coordinates i.

(2) A combinatorial checkerboard can be defined as a function from $[m]^d$ to $\{0, 1\}$ computed by a width-2 length-d degree-m layered branching program of the following form. At layer $i \in \{1, \ldots, d\}$, the branching program reads the ith symbol of the input and transitions to layer $i + 1$, and the set of symbols that cause it to cross from top to bottom is the same as the set of symbols that cause it to cross from bottom to top (call this set S_i). The start state is the bottom node in layer 1, and the accept state is the top node in layer $d+1$. (See Figure 2.) In contrast, a combinatorial rectangle can be defined similarly but where the start state and the accept state are both on top, and at layer $i \in \{1, \ldots, d\}$, the bottom node transitions to the bottom node in layer $i + 1$ no matter what the ith symbol is (while the behavior at the top node is arbitrary).

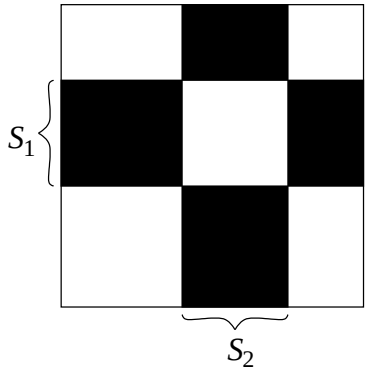

Figure 1

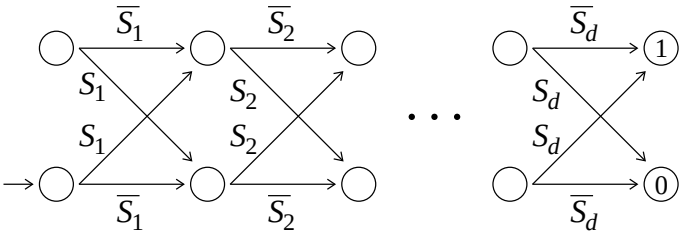

Figure 2

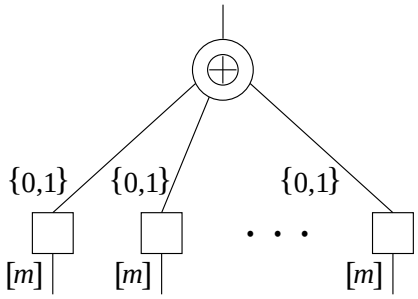

Figure 3

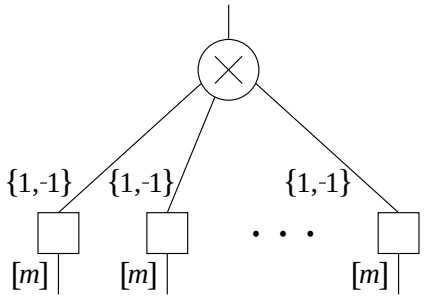

Figure 4

(3) A combinatorial checkerboard can be defined as a function from $[m]^d$ to $\{0,1\}$ computed by a circuit of the following form. There are d input wires, each carrying a symbol in $[m]$. Each input wire feeds into a "gate" that computes an arbitrary function from $[m]$ to $\{0,1\}$, and the resulting d bits are fed into an XOR gate. (See Figure 3.) In contrast, a combinatorial rectangle can be defined similarly but where the XOR gate is replaced with an AND gate.

(4) A combinatorial checkerboard can be defined as a function from $[m]^d$ to $\{1,-1\}$ computed by a circuit of the following form. There are d input wires, each carrying a symbol in $[m]$. Each input wire feeds into a "gate" that computes an arbitrary function from $[m]$ to $\{1,-1\}$, and the resulting d numbers are fed into a multiplication gate. (See Figure 4.) In contrast, a combinatorial rectangle can be defined similarly but where $\{1,-1\}$ is replaced with $\{0,1\}$.

For the rest of this paper, we adopt the fourth view.

Definition 1 (Combinatorial Checkerboards). *We say $f : [m]^d \to \{1,-1\}$ is an (m,d)-*checkerboard *if it is of the form $f(u_1,\ldots,u_d) = \prod_{i\in[d]} f_i(u_i)$ for some functions $f_i : [m] \to \{1,-1\}$. We denote this as $f = \bigotimes_{i\in[d]} f_i$.*

Definition 2 (Pseudorandom Generators). *Let $\mathcal{C}$ be a class of functions from some finite universe U to $\{1,-1\}$. We say $G : \{0,1\}^s \to U$ is an ϵ-*pseudorandom generator for $\mathcal{C}$ *if for all $f \in \mathcal{C}$, $\left|\mathrm{E}_{r\in\{0,1\}^s}\left[f(G(r))\right] - \mathrm{E}_{u\in U}[f(u)]\right| \leq \epsilon$ where r and u are both chosen uniformly at random. We say s is the* seed length *of G.*

1.2 Our Result

The rest of this paper is devoted to proving the following theorem.

Theorem 1 (Main Theorem). *There exists an explicit ϵ-pseudorandom generator for the class of (m,d)-checkerboards with seed length $O\big(\log m + \log d \cdot \log\log d + \log^{3/2}\frac{1}{\epsilon}\big)$.*

Informally, when we say *explicit* we mean that an efficient algorithm with the desired behavior is exhibited. We do not attempt to quantify the time or space efficiency parameters throughout this paper. In the case of Theorem 1, the precise meaning is that there exists a uniform deterministic algorithm $\mathcal{A}$ that takes as input the parameters m, d, ϵ and a string in $\{0,1\}^s$ (where s is the seed length), outputs an element of $[m]^d$, runs in time $\mathrm{poly}\big(\log m + d + \log\frac{1}{\epsilon}\big)$, and is such that for all m, d, ϵ the function $\mathcal{A}(m,d,\epsilon,\cdot)$ is an ϵ-pseudorandom generator for the class of (m,d)-checkerboards. A simple probabilistic argument shows that $O\big(\log m + \log d + \log\frac{1}{\epsilon}\big)$ seed length can be achieved if we allow *nonexplicit* pseudorandom generators.

Impagliazzo, Nisan, and Wigderson [34] proved a result for small-width read-once branching programs which in particular gives an explicit ϵ-pseudorandom generator for (m,d)-checkerboards with seed length $O\big(\log m + \log^2 d + \log d \cdot \log\frac{1}{\epsilon}\big)$. Our seed length is better except when $\frac{1}{\epsilon} \geq d^{\omega(\log d)}$. If m is a power of 2, then an (m,d)-checkerboard can be viewed as a polynomial over $\mathbb{Z}_2$ of degree at most $\log_2 m$ with $d \cdot \log_2 m$ variables (since each f_i can be viewed as an arbitrary function from $\mathbb{Z}_2^{\log_2 m}$ to $\mathbb{Z}_2$). Viola [73] constructed an ϵ-pseudorandom generator for n-variable, degree-k polynomials over $\mathbb{Z}_2$ with seed length $O\big(k \cdot \log n + 2^k \cdot k \cdot \log\frac{1}{\epsilon}\big)$, which yields an ϵ-pseudorandom generator for (m,d)-checkerboards with seed length $O\big(\log m \cdot \log d + m \cdot \log m \cdot \log\frac{1}{\epsilon}\big)$, assuming m is a power of 2.[1] The latter seed length is optimal when m is constant but has very poor dependence on m. If $m = 2$, then the degree of the polynomial becomes 1 (that is, a $(2,d)$-checkerboard is equivalent to a d-variable affine function over $\mathbb{Z}_2$) and the result of [73] degenerates to known constructions of small-bias generators, which have seed length $O\big(\log d + \log\frac{1}{\epsilon}\big)$.

For comparison, we mention what is known for combinatorial rectangles. The two best generators (which have incomparable seed lengths) are due to Impagliazzo, Nisan, and Wigderson [34], who achieved seed length $O\big(\log m + \log^2 d + \log d \cdot \log\frac{1}{\epsilon}\big)$, and Lu [43], who achieved seed length $O\big(\log m + \log d + \log^{3/2}\frac{1}{\epsilon}\big)$. The latter result is better than the former except when $\frac{1}{\epsilon} \geq d^{\omega(\log d)}$.

1.3 Overview of the Proof

We partition the set of (m,d)-checkerboards into a "high-weight case" and a "low-weight case". We construct a generator that fools high-weight checkerboards and a different generator that fools low-weight checkerboards, and we combine the two generators to get a single generator that fools all checkerboards. (This technique has been used before, for example in [52, 42].) We now give our definition of the weight of a checkerboard.

Definition 3 (Bias and Unbias). *The* bias *of $f : U \to \{1,-1\}$ is $\beta(f) = \big|\mathrm{E}_{u\in U}[f(u)]\big|$ where u is chosen uniformly at random, and the* unbias *is $\alpha(f) = 1 - \beta(f)$.*

Definition 4 (Weight). *The* weight *of an (m,d)-checkerboard $f = \bigotimes_{i\in[d]} f_i$ is $\sum_{i\in[d]} \alpha(f_i)$.*

[1] In the proof of Theorem 1, we note that we can assume without loss of generality that m is a power of 2. However, this is not without loss of generality when we apply the result of [73], because the reduction to the power-of-2 case blows up m to at least $4md/\epsilon$.

Observation 1. *If $f_1, \ldots, f_d, f'_1, \ldots, f'_d : [m] \to \{1, -1\}$ are such that $\bigotimes_{i\in[d]} f_i = \bigotimes_{i\in[d]} f'_i$, then for each $i \in [d]$ we have $f_i = \pm f'_i$ and thus $\alpha(f_i) = \alpha(f'_i)$. In particular, the weight of an (m, d)-checkerboard is independent of the representation as a tensor product.*

The real difficulty in proving Theorem 1 stems from the fact that the biases $\beta(f_i)$ are arbitrary numbers in $[0, 1]$. If we knew that each bias $\beta(f_i)$ were either 0 or 1, then the techniques of [42, 38] would translate straightforwardly to our setting: The techniques of [42] would immediately yield a pseudorandom generator with seed length $O\big(\log m \cdot \log \frac{1}{\epsilon} + \log d + \log \frac{1}{\epsilon} \cdot \log\log \frac{1}{\epsilon}\big)$, and the techniques of [38] would immediately yield a pseudorandom generator with seed length $O\big(\log m + \log d \cdot \log \frac{1}{\epsilon}\big)$.

For the known results on combinatorial rectangles [25, 39, 9, 43], there is an analogous (but different) notion of "bias", and these results give techniques for handling arbitrary biases in $[0, 1]$. We adapt these techniques to fool low-weight checkerboards. However, in the case of combinatorial rectangles, there basically is no "high-weight case" — to fool high-weight rectangles it suffices to fool low-weight rectangles. In our setting we are not so fortunate, and we must do something genuinely different to fool high-weight checkerboards. To accomplish the latter, we build on the techniques of [42].

The threshold we use to distinguish "high-weight" from "low-weight" is $\Theta\big(\log \frac{1}{\epsilon}\big)$.

Lemma 1 (High-Weight Case). *There exists a universal constant C such that the following holds. There exists an explicit ϵ-pseudorandom generator for the class of (m, d)-checkerboards of weight at least $C \cdot \log_2 \frac{1}{\epsilon}$ with seed length $O\big(\log m + \log d \cdot \log\log d + \log \frac{1}{\epsilon} \cdot \log\log \frac{1}{\epsilon}\big)$, provided m and d are powers of 2.*

Lemma 2 (Low-Weight Case). *There exists an explicit ϵ-pseudorandom generator for the class of (m, d)-checkerboards of weight less than $C \cdot \log_2 \frac{1}{\epsilon}$ with seed length $O\big(\log m + \log d + \log^{3/2} \frac{1}{\epsilon}\big)$, provided m and d are powers of 2, where C is the constant from Lemma 1.*

We derive Theorem 1 from Lemma 1 and Lemma 2 in Section 2. This just amounts to showing that (i) we can assume without loss of generality that m and d are powers of 2, and (ii) the two generators can be combined to fool all checkerboards. Both are simple and standard; we include the arguments for completeness.

We prove Lemma 1 in Section 3. Here is the outline of the proof. Suppose we can construct a generator with seed length $O\big(\log m + \log d \cdot \log\log d\big)$ that fools, within a constant, checkerboards of at least constant weight. Then we can use the following technique of [42] to get the final generator. First use a hash function to randomly partition the coordinates into a small number of buckets such that most buckets have at least constant weight. Then apply the hypothesized generator to each bucket, but instead of using independent seeds for the different instantiations of the hypothesized generator, sample the seeds from an appropriate pseudorandom distribution. This technique of [42] only contributes an additive $O\big(\log d + \log \frac{1}{\epsilon} \cdot \log\log \frac{1}{\epsilon}\big)$ to the seed length. Thus we just need to be able to fool, within a constant, checkerboards of at least constant weight. The heart of our proof of Lemma 1 is a new analysis of the generator of Impagliazzo, Nisan, and Wigderson [34] showing that for this special case, it suffices to use expander graphs of degree polylog d. In [42], the analysis of the corresponding part of the argument is considerably simpler because the authors exploit the fact that in their setting, the bias of each coordinate is either 0 or 1.

We prove Lemma 2 in Section 4. We take as a starting point the techniques of [9, 43]. Numerous small modifications to these techniques are needed. One bigger modification is the following. Lu's

proof [43] critically makes use of the Bonferroni inequalities, which state that the probability of a union of events is alternately upper and lower bounded by the successive truncations of the inclusion-exclusion formula. In our proof we use an alternative analogous principle which is a bit tougher to prove than the Bonferroni inequalities, but which follows from elementary combinatorial techniques, and which may be folklore.

1.4 Preliminaries

Before diving into the proofs, we mention some conventions that we use for convenience throughout the proofs. We identify $\{0,1\}^s$ with $[2^s]$, and we always use the latter notation. Thus for example, a pseudorandom generator with seed length s is a function with domain $[2^s]$. We may also identify $[2^s]$ with $[2^{s_1}] \times [2^{s_2}]$ if $s = s_1 + s_2$. We also freely flatten trees of Cartesian products of sets; for example, we identify $\big((U_1 \times U_2) \times (U_3 \times U_4)\big)$ with $U_1 \times U_2 \times U_3 \times U_4$.

2 Deriving Theorem 1 from Lemma 1 and Lemma 2

In this section we prove Theorem 1.

Definition 5. *We say $\pi : [m] \times [m] \to [m]$ is a* quasigroup operation *if for every $v \in [m]$, the mappings $u \mapsto \pi(u,v)$ and $u \mapsto \pi(v,u)$ are both permutations.*

Definition 6. *We say a class $\mathcal{C}$ of (m,d)-checkerboards is* closed under permutations *if the following holds. For all functions $f_1, \ldots, f_d : [m] \to \{1,-1\}$ and all permutations $p_1, \ldots, p_d : [m] \to [m]$, if $\bigotimes_{i\in[d]} f_i \in \mathcal{C}$ then $\bigotimes_{i\in[d]}(f_i \circ p_i) \in \mathcal{C}$.*

Definition 7. *Given $\pi : [m] \times [m] \to [m]$ and $G_1 : [2^{s_1}] \to [m]^d$ and $G_2 : [2^{s_2}] \to [m]^d$, we define $\big(G_1 +_\pi G_2\big) : [2^{s_1}] \times [2^{s_2}] \to [m]^d$ by*

$$\big(G_1 +_\pi G_2\big)(r_1, r_2)_i \;=\; \pi\big(G_1(r_1)_i, G_2(r_2)_i\big)$$

for $i \in [d]$.

Proposition 1. *Suppose $\pi : [m] \times [m] \to [m]$ is a quasigroup operation, and suppose $\mathcal{C}_1$ and $\mathcal{C}_2$ are two classes of (m,d)-checkerboards both closed under permutations. If $G_1 : [2^{s_1}] \to [m]^d$ is an ϵ-pseudorandom generator for $\mathcal{C}_1$ and $G_2 : [2^{s_2}] \to [m]^d$ is an ϵ-pseudorandom generator for $\mathcal{C}_2$, then $G = G_1 +_\pi G_2$ is an ϵ-pseudorandom generator for $\mathcal{C}_1 \cup \mathcal{C}_2$.*

Proof. Consider an arbitrary $f \in \mathcal{C}_1 \cup \mathcal{C}_2$. Assume $f \in \mathcal{C}_1$; the other case is symmetric. To show that

$$\Big|\, \mathrm{E}_{r_1\in[2^{s_1}], r_2\in[2^{s_2}]}\big[(f\circ G)(r_1,r_2)\big] - \mathrm{E}_{u\in[m]^d}[f(u)] \Big| \;\le\; \epsilon$$

it suffices to show that for each $r_2 \in [2^{s_2}]$,

$$\Big|\, \mathrm{E}_{r_1\in[2^{s_1}]}\big[(f\circ G)(r_1,r_2)\big] - \mathrm{E}_{u\in[m]^d}[f(u)] \Big| \;\le\; \epsilon. \tag{1}$$

Fix an arbitrary $r_2 \in [2^{s_2}]$ and define $(v_1, \ldots, v_d) = G_2(r_2)$. Define an (m,d)-checkerboard $f' = \bigotimes_{i\in[d]} f'_i$ where $f'_i(u) = f_i\big(\pi(u, v_i)\big)$. Observe that $(f \circ G)(r_1, r_2) = (f' \circ G_1)(r_1)$ holds for each $r_1 \in [2^{s_1}]$, and thus

$$\mathrm{E}_{r_1\in[2^{s_1}]}\big[(f\circ G)(r_1,r_2)\big] \;=\; \mathrm{E}_{r_1\in[2^{s_1}]}\big[(f'\circ G_1)(r_1)\big] \tag{2}$$

(this holds even if π is not a quasigroup operation). Observe that $\mathrm{E}_{u\in[m]}[f_i(u)] = \mathrm{E}_{u\in[m]}[f'_i(u)]$ holds for each $i \in [d]$ since π is a quasigroup operation, and thus

$$\mathrm{E}_{u\in[m]^d}[f(u)] \;=\; \mathrm{E}_{u\in[m]^d}[f'(u)]. \tag{3}$$

Since $\mathcal{C}_1$ is closed under permutations and $f \in \mathcal{C}_1$, we have $f' \in \mathcal{C}_1$. Since G_1 is an ϵ-pseudorandom generator for $\mathcal{C}_1$, we have

$$\left|\mathrm{E}_{r_1\in[2^{s_1}]}\left[(f'\circ G_1)(r_1)\right] - \mathrm{E}_{u\in[m]^d}[f'(u)]\right| \;\leq\; \epsilon. \tag{4}$$

Now Inequality (1) follows from Equality (2), Equality (3), and Inequality (4). □

Proposition 2. *Suppose that for some W there exists an explicit ϵ-pseudorandom generator G_1 for the class $\mathcal{C}_1$ of (m,d)-checkerboards of weight at least W with seed length s_1, and there exists an explicit ϵ-pseudorandom generator G_2 for the class $\mathcal{C}_2$ of (m,d)-checkerboards of weight less than W with seed length s_2. Then there exists an explicit ϵ-pseudorandom generator for the class of all (m,d)-checkerboards with seed length $s_1 + s_2$.*

Proof. Let $\pi : [m] \times [m] \to [m]$ be any explicit quasigroup operation. For example, we can identify $[m]$ with $\{0, 1, \ldots, m-1\}$ and let π be addition modulo m. Observe that both $\mathcal{C}_1$ and $\mathcal{C}_2$ are closed under permutations. Then Proposition 1 guarantees that $G_1 +_\pi G_2$ is an explicit ϵ-pseudorandom generator for $\mathcal{C}_1 \cup \mathcal{C}_2$, which is the class of all (m,d)-checkerboards. Furthermore, $G_1 +_\pi G_2$ has seed length $s_1 + s_2$. □

Proposition 2 is also used in the proof of Lemma 2. We are now ready to prove Theorem 1.

Proof of Theorem 1. Let m' be the smallest power of 2 that is at least $4md/\epsilon$, let d' be the smallest power of 2 that is at least d, and let $\epsilon' = \epsilon/2$. For the parameters m', d', ϵ', combining Lemma 1 with Lemma 2 using Proposition 2 (with $W = C \cdot \log_2 \frac{1}{\epsilon'}$ where C is the constant from Lemma 1) we find that there exists an explicit ϵ'-pseudorandom generator G' for the class of (m', d')-checkerboards with seed length $s = O\big(\log m' + \log d' \cdot \log\log d' + \log^{3/2} \frac{1}{\epsilon'}\big) = O\big(\log m + \log d \cdot \log\log d + \log^{3/2} \frac{1}{\epsilon}\big)$.

Now let $h : [m'] \to [m]$ be any explicit function such that every element of $[m]$ has at least $\lfloor \frac{m'}{m} \rfloor$ preimages and at most $\lceil \frac{m'}{m} \rceil$ preimages. Define $H : [m']^{d'} \to [m]^d$ by $H(u_1, \ldots, u_{d'}) = \big(h(u_1), \ldots, h(u_d)\big)$. Then we claim that the function $G = H \circ G'$, which also has seed length s, is an ϵ-pseudorandom generator for the class of (m,d)-checkerboards. Consider an arbitrary (m,d)-checkerboard $f = \bigotimes_{i\in[d]} f_i$, and define $f' = f \circ H$. Notice that $f' = \bigotimes_{i\in[d']} f'_i$ where

$$f'_i \;=\; \begin{cases} f_i \circ h & \text{if } i \in [d] \\ 1 & \text{otherwise} \end{cases}$$

where 1 denotes the constant 1 function on $[m']$. Since f' is an (m', d')-checkerboard, we have

$$\left|\mathrm{E}_{r\in[2^s]}\left[(f'\circ G')(r)\right] - \mathrm{E}_{u\in[m']^{d'}}[f'(u)]\right| \;\leq\; \epsilon/2. \tag{5}$$

Since $f \circ G = f' \circ G'$, we have

$$\mathrm{E}_{r\in[2^s]}\left[(f\circ G)(r)\right] \;=\; \mathrm{E}_{r\in[2^s]}\left[(f'\circ G')(r)\right]. \tag{6}$$

A simple calculation shows that for each $i \in [d]$ we have

$$\left| \mathrm{E}_{u \in [m']} \left[(f_i \circ h)(u) \right] - \mathrm{E}_{u \in [m]}[f_i(u)] \right| \;\leq\; 2m/m' \;\leq\; \epsilon/2d.$$

Thus we have

$$\begin{aligned} \left| \mathrm{E}_{u \in [m']^{d'}}[f'(u)] - \mathrm{E}_{u \in [m]^d}[f(u)] \right| &= \left| \textstyle\prod_{i \in [d]} \mathrm{E}_{u \in [m']} \left[(f_i \circ h)(u) \right] - \prod_{i \in [d]} \mathrm{E}_{u \in [m]}[f_i(u)] \right| \\ &\leq \textstyle\sum_{i \in [d]} \left| \mathrm{E}_{u \in [m']} \left[(f_i \circ h)(u) \right] - \mathrm{E}_{u \in [m]}[f_i(u)] \right| \\ &\leq \epsilon/2 \end{aligned}$$

where the second line follows by the simple fact that for all $x_1, \ldots, x_d, y_1, \ldots, y_d \in [-1, 1]$ we have $\left| \prod_{i \in [d]} x_i - \prod_{i \in [d]} y_i \right| \leq \sum_{i \in [d]} |x_i - y_i|$. Combining this with Inequality (5) and Equality (6) yields

$$\left| \mathrm{E}_{r \in [2^s]} \left[(f \circ G)(r) \right] - \mathrm{E}_{u \in [m]^d}[f(u)] \right| \;\leq\; \epsilon. \qquad \square$$

3 The High-Weight Case

This section is devoted to the proof of Lemma 1. The main component in the proof of Lemma 1 is Lemma 3 below, and the main component in the proof of Lemma 3 is Lemma 8 below. We prove these three lemmas in Section 3.1, Section 3.2, and Section 3.3 respectively.

3.1 Proof of Lemma 1

We first discuss notation. We use G to denote the generator we construct to witness Lemma 1. The parameters m, d, ϵ are fixed, with m and d powers of 2. We can also assume without loss of generality that $\log_2 \frac{1}{\epsilon}$ is a power of 2, since otherwise we could decrease ϵ to make this so, while only affecting the seed length by a constant factor. For the rest of this section, we define $b = 16 \cdot \log_2 \frac{1}{\epsilon}$, which represents the number of "buckets" of a certain hash function. We use $i \in [d]$ to index coordinates of the original checkerboard and $j \in [b]$ to index buckets. The construction has three steps, and we use s_1, s_2, s_3 to denote the contributions of the three steps to the final seed length s.

Lemma 3 (Step 1). *There exists an explicit function $G_1 : [2^{s_1}] \to [m]^d$ with $s_1 = O\big(\log m + \log d \cdot \log\log d\big)$ such that if f is an (m, d)-checkerboard of weight at least 1, then $\beta(f \circ G_1) \leq 3/4$.*

Lemma 4 (Step 2). *There exists an explicit function $G_2 : [2^{s_1}] \times [2^{s_2}] \to [2^{s_1}]^b$ with $s_2 = O\big(\log \frac{1}{\epsilon}\big)$ such that the following holds. Suppose $g = \bigotimes_{j \in [b]} g_j$ is a $(2^{s_1}, b)$-checkerboard such that*

$$\mathrm{Pr}_{j \in [b]} \left[\beta(g_j) > 3/4 \right] \;\leq\; 1/16$$

where j is chosen uniformly at random. Then $\beta(g \circ G_2) \leq \epsilon/4$.

Lemma 5 (Step 3). *There exists a universal constant $C \geq 1$ and an explicit function $G_3 : [2^{s_3}] \times [d] \to [b]$ with $s_3 = O\big(\log d + \log \frac{1}{\epsilon} \cdot \log\log \frac{1}{\epsilon}\big)$ such that the following holds. Suppose $\alpha_1, \ldots, \alpha_d \in [0, 1]$ are such that $\sum_{i \in [d]} \alpha_i \geq C \cdot \log_2 \frac{1}{\epsilon}$. Then*

$$\mathrm{Pr}_{r_3 \in [2^{s_3}]} \left[\mathrm{Pr}_{j \in [b]} \left[\textstyle\sum_{i \,:\, G_3(r_3, i) = j} \alpha_i < 1 \right] > 1/16 \right] \;\leq\; \epsilon/4$$

where r_3 and j are chosen uniformly at random.

We prove Lemma 3 in Section 3.2. The proof involves a new analysis of the generator of Impagliazzo, Nisan, and Wigderson [34] for the setting of combinatorial checkerboards. The heart of the analysis, which we call the Tree Labeling Lemma, is proven in Section 3.3.

Lovett et al. [42] implicitly proved Lemma 4, although they did not phrase it in terms of combinatorial checkerboards. Their proof (which we do not reproduce here) uses an instantiation of the generator of Impagliazzo, Nisan, and Wigderson [34].

In Lemma 5, G_3 is viewed as a family of hash functions parameterized by the first argument. Lovett et al. [42] proved Lemma 5 assuming each number α_i is 0 or 1, but their proof goes through for arbitrary $\alpha_i \in [0,1]$. Their proof (which we do not reproduce here) makes use of a concentration result for sums of k-wise independent random variables, due to Bellare and Rompel [14].

We now show how Lemma 1 follows from Lemma 3, Lemma 4, and Lemma 5.

Proof of Lemma 1. We construct a generator $G : [2^s] \to [m]^d$ with $s = s_1 + s_2 + s_3 = O\big(\log m + \log d \cdot \log\log d + \log \frac{1}{\epsilon} \cdot \log\log \frac{1}{\epsilon}\big)$ that witnesses Lemma 1. Identifying $[2^s]$ with $[2^{s_1}] \times [2^{s_2}] \times [2^{s_3}]$, we let

$$G(r_1, r_2, r_3)_i \;=\; G_1\Big(G_2(r_1, r_2)_{G_3(r_3,i)}\Big)_i$$

for $i \in [d]$. That is, the generator first runs $G_2(r_1, r_2)$ to obtain b seeds for G_1 and then it runs G_1 on each of these seeds. From the execution of G_1 on the jth seed, the generator obtains d values in $[m]$ but it only keeps those corresponding to indices in the jth bucket of the hash function $G_3(r_3, \cdot)$.

We claim that G witnesses Lemma 1. Consider an arbitrary (m,d)-checkerboard $f = \bigotimes_{i\in[d]} f_i$ of weight at least $C \cdot \log_2 \frac{1}{\epsilon}$ where C is the constant from Lemma 5. Note that

$$\beta(f) \;=\; \textstyle\prod_{i\in[d]} \beta(f_i) \;\le\; e^{-\sum_{i\in[d]} \alpha(f_i)} \;\le\; \epsilon/2.$$

We just need to argue that $\beta(f \circ G) \le \epsilon/2$, because then it follows that

$$\Big|\, \mathrm{E}_{r\in[2^s]}\big[(f\circ G)(r)\big] - \mathrm{E}_{u\in[m]^d}[f(u)] \Big| \;\le\; \beta(f\circ G) + \beta(f) \;\le\; \epsilon/2 + \epsilon/2 \;=\; \epsilon.$$

Define

$$\mathrm{Bad} \;=\; \left\{ r_3 \in [2^{s_3}] \;:\; \mathrm{Pr}_{j\in[b]}\Big[\textstyle\sum_{i\,:\,G_3(r_3,i)=j} \alpha(f_i) < 1\Big] > 1/16 \right\}$$

and let $\mathrm{Good} = [2^{s_3}] \backslash \mathrm{Bad}$. Applying Lemma 5 with $\alpha_i = \alpha(f_i)$ for each $i \in [d]$, we find that $\mathrm{Pr}_{r_3\in[2^{s_3}]}\big[r_3 \in \mathrm{Bad}\big] \le \epsilon/4$. We claim that for each $r_3 \in \mathrm{Good}$,

$$\Big|\, \mathrm{E}_{r_1\in[2^{s_1}], r_2\in[2^{s_2}]}\big[(f\circ G)(r_1, r_2, r_3)\big] \Big| \;\le\; \epsilon/4. \tag{7}$$

This will finish the argument since then

$$\begin{aligned}
\beta(f\circ G) \;&\le\; \Big|\, \mathrm{E}_{r_1\in[2^{s_1}], r_2\in[2^{s_2}], r_3\in[2^{s_3}]}\big[(f\circ G)(r_1, r_2, r_3) \;\big|\; r_3 \in \mathrm{Good}\big] \Big| + \mathrm{Pr}_{r_3\in[2^{s_3}]}\big[r_3 \in \mathrm{Bad}\big] \\
&\le\; \epsilon/4 + \epsilon/4 \\
&=\; \epsilon/2.
\end{aligned}$$

To prove the claim, fix an arbitrary $r_3 \in \mathrm{Good}$. For each $j \in [b]$ define an (m,d)-checkerboard $f^{(j)} = \bigotimes_{i\in[d]} f_i^{(j)}$ by

$$f_i^{(j)} \;=\; \begin{cases} f_i & \text{if } G_3(r_3, i) = j \\ 1 & \text{otherwise} \end{cases}$$

where 1 denotes the constant 1 function on $[m]$. Define a $(2^{s_1}, b)$-checkerboard $g = \bigotimes_{j\in[b]} g_j$ by $g_j = f^{(j)} \circ G_1$. Note that for each $r_1 \in [2^{s_1}], r_2 \in [2^{s_2}]$, we have

$$\begin{aligned}
(g \circ G_2)(r_1, r_2) &= \textstyle\prod_{j\in[b]} g_j\big(G_2(r_1, r_2)_j\big) \\
&= \textstyle\prod_{j\in[b]} \prod_{i\in[d]} f_i^{(j)}\Big(G_1\big(G_2(r_1, r_2)_j\big)_i\Big) \\
&= \textstyle\prod_{j\in[b]} \prod_{i\,:\,G_3(r_3,i)=j} f_i\Big(G_1\big(G_2(r_1, r_2)_j\big)_i\Big) \\
&= \textstyle\prod_{i\in[d]} f_i\Big(G_1\Big(G_2(r_1, r_2)_{G_3(r_3,i)}\Big)_i\Big) \\
&= (f \circ G)(r_1, r_2, r_3)
\end{aligned}$$

by commutativity of multiplication. It follows that

$$\mathrm{E}_{r_1\in[2^{s_1}],r_2\in[2^{s_2}]}\left[(f \circ G)(r_1, r_2, r_3)\right] = \mathrm{E}_{r_1\in[2^{s_1}],r_2\in[2^{s_2}]}\left[(g \circ G_2)(r_1, r_2)\right]$$

and hence to prove Inequality (7) it suffices to show that $\beta(g\circ G_2) \le \epsilon/4$. If $\sum_{i\,:\,G_3(r_3,i)=j} \alpha(f_i) \ge 1$ then the weight of $f^{(j)}$ is at least 1 and thus by Lemma 3 we have $\beta(g_j) \le 3/4$. Since $r_3 \in \mathrm{Good}$, we have $\Pr_{j\in[b]}\left[\beta(g_j) > 3/4\right] \le 1/16$. Thus Lemma 4 implies that $\beta(g \circ G_2) \le \epsilon/4$, as desired. □

3.2 Proof of Lemma 3

In this section we prove Lemma 3. The proof uses explicit constructions of expander graphs. One can view a $(2^n, 2^k, \lambda)$-expander as a symmetric $2^n \times 2^n$ matrix M of nonnegative integers such that each row and each column sums to 2^k, and such that every eigenvalue of $M/2^k$, except the first, is at most λ in absolute value. An equivalent way of viewing an expander is as a regular symmetric directed multigraph on 2^n vertices with degree 2^k whose adjacency matrix is M. A third view, which we use, is a function $E : [2^{n+k}] \to [2^n] \times [2^n]$ that maps the edges (of which there are 2^{n+k}, and which are identified with the elements of $[2^{n+k}]$ in an arbitrary way) to their (head, tail) pairs.

Definition 8. *A $(2^n, 2^k, \lambda)$-expander is a function $E : [2^{n+k}] \to [2^n] \times [2^n]$ such that the $2^n \times 2^n$ matrix M defined by $M_{\nu_1,\nu_2} = \left|E^{-1}(\nu_1, \nu_2)\right|$ satisfies the following: M is symmetric, each row and each column sums to 2^k, and every eigenvalue of $M/2^k$, except the first, is at most λ in absolute value.*

Many explicit constructions of good expanders are known [47, 27, 37, 7, 5, 44, 48, 51, 65]. The Gabber-Galil construction [27] in particular yields the following.

Lemma 6. *For every $\lambda > 0$ there exists an integer $k = O\big(\log \frac{1}{\lambda}\big)$ such that for all integers $n \ge 1$ there exists an explicit $(2^n, 2^k, \lambda)$-expander.*

We use the classic Expander Mixing Lemma, which is generally attributed to [4].

Lemma 7 (Expander Mixing Lemma). *For every $(2^n, 2^k, \lambda)$-expander E, every $S \subseteq [2^n]$, and every $T \subseteq [2^n]$, we have*

$$\begin{aligned}
&\left|\Pr_{\mu\in[2^{n+k}]}\left[E(\mu) \in S \times T\right] - \Pr_{\nu\in[2^n]}[\nu \in S] \cdot \Pr_{\nu\in[2^n]}[\nu \in T]\right| \\
\le\ &\lambda\sqrt{\Pr_{\nu\in[2^n]}[\nu \in S] \cdot \Pr_{\nu\in[2^n]}[\nu \in T]}
\end{aligned}$$

where μ and ν are both chosen uniformly at random.

Definition 9 (Cartesian Product with Respect to E). *Given a $(2^n, 2^k, \lambda)$-expander E and two functions $h_1, h_2 : [2^n] \to U$ for some finite U, we define $h_1 \times_E h_2 : [2^{n+k}] \to U \times U$ by*

$$\big(h_1 \times_E h_2\big)(\mu) \;=\; \Big(h_1\big(E(\mu)_1\big), h_2\big(E(\mu)_2\big)\Big).$$

In other words, $h_1 \times_E h_2 = \big(h_1 \times h_2\big) \circ E$.

Definition 10 (Tensor Product with Respect to E). *Given a $(2^n, 2^k, \lambda)$-expander E and two functions $h_1, h_2 : [2^n] \to \{1, -1\}$, we define $h_1 \otimes_E h_2 : [2^{n+k}] \to \{1, -1\}$ by*

$$\big(h_1 \otimes_E h_2\big)(\mu) \;=\; h_1\big(E(\mu)_1\big) \cdot h_2\big(E(\mu)_2\big).$$

In other words, $h_1 \otimes_E h_2 = \big(h_1 \otimes h_2\big) \circ E$.

Observation 2. *For all $(2^n, 2^k, \lambda)$-expanders E and all functions $g_1, g_2 : [2^n] \to U$ and $h_1, h_2 : U \to \{1, -1\}$ for some finite U, we have*

$$\big(h_1 \otimes h_2\big) \circ \big(g_1 \times_E g_2\big) \;=\; \big(h_1 \circ g_1\big) \otimes_E \big(h_2 \circ g_2\big).$$

Proposition 3. *For all $(2^n, 2^k, \lambda)$-expanders E and all functions $h_1, h_2 : [2^n] \to \{1, -1\}$, we have*

$$\beta\big(h_1 \otimes_E h_2\big) \;\le\; \beta(h_1)\beta(h_2) + \lambda \cdot \big(\alpha(h_1) + \alpha(h_2)\big).$$

Proof. We may assume without loss of generality that $\mathrm{E}_{\nu \in [2^n]}[h_1(\nu)] \ge 0$ and $\mathrm{E}_{\nu \in [2^n]}[h_2(\nu)] \ge 0$ because replacing h_1 by $-h_1$ and/or h_2 by $-h_2$ changes none of the quantities in the inequality we wish to prove. Now we have

$$\begin{aligned}
&\beta\big(h_1 \otimes_E h_2\big) - \beta(h_1)\beta(h_2) \\
\le\;& \Big|\mathrm{E}_{\mu \in [2^{n+k}]}\big[\big(h_1 \otimes_E h_2\big)(\mu)\big] - \mathrm{E}_{\nu \in [2^n]}[h_1(\nu)] \cdot \mathrm{E}_{\nu \in [2^n]}[h_2(\nu)]\Big| \\
=\;& \Big|\mathrm{E}_{\mu \in [2^{n+k}]}\big[\big(h_1 \otimes_E h_2\big)(\mu)\big] - \mathrm{E}_{\nu_1 \in [2^n], \nu_2 \in [2^n]}\big[\big(h_1 \otimes h_2\big)(\nu_1, \nu_2)\big]\Big| \\
=\;& 2 \cdot \Big|\Pr_{\mu \in [2^{n+k}]}\big[\big(h_1 \otimes_E h_2\big)(\mu) = -1\big] - \Pr_{\nu_1 \in [2^n], \nu_2 \in [2^n]}\big[\big(h_1 \otimes h_2\big)(\nu_1, \nu_2) = -1\big]\Big| \\
\le\;& 2 \cdot \Big|\Pr_{\mu \in [2^{n+k}]}\big[\big(h_1 \times_E h_2\big)(\mu) = (-1, 1)\big] - \Pr_{\nu_1 \in [2^n], \nu_2 \in [2^n]}\big[\big(h_1 \times h_2\big)(\nu_1, \nu_2) = (-1, 1)\big]\Big| + \\
& 2 \cdot \Big|\Pr_{\mu \in [2^{n+k}]}\big[\big(h_1 \times_E h_2\big)(\mu) = (1, -1)\big] - \Pr_{\nu_1 \in [2^n], \nu_2 \in [2^n]}\big[\big(h_1 \times h_2\big)(\nu_1, \nu_2) = (1, -1)\big]\Big|
\end{aligned}$$

by simple manipulations. Since every row of the matrix corresponding to E has the same sum, we find that

$$\Pr_{\mu \in [2^{n+k}]}\big[\big(h_1 \times_E h_2\big)(\mu)_1 = -1\big] \;=\; \Pr_{\nu_1 \in [2^n]}\big[h_1(\nu_1) = -1\big]$$

and thus

$$\begin{aligned}
&\Big|\Pr_{\mu \in [2^{n+k}]}\big[\big(h_1 \times_E h_2\big)(\mu) = (-1, 1)\big] - \Pr_{\nu_1 \in [2^n], \nu_2 \in [2^n]}\big[\big(h_1 \times h_2\big)(\nu_1, \nu_2) = (-1, 1)\big]\Big| \\
=\;& \Big|\Pr_{\mu \in [2^{n+k}]}\big[\big(h_1 \times_E h_2\big)(\mu) = (-1, -1)\big] - \Pr_{\nu_1 \in [2^n], \nu_2 \in [2^n]}\big[\big(h_1 \times h_2\big)(\nu_1, \nu_2) = (-1, -1)\big]\Big| \\
\le\;& \lambda\sqrt{\Pr_{\nu \in [2^n]}\big[h_1(\nu) = -1\big] \cdot \Pr_{\nu \in [2^n]}\big[h_2(\nu) = -1\big]}
\end{aligned}$$

$= \frac{1}{2}\lambda\sqrt{\alpha(h_1)\alpha(h_2)}$

by applying the Expander Mixing Lemma with $S = h_1^{-1}(-1)$ and $T = h_2^{-1}(-1)$ and using the fact that $\alpha(h_1) = 2\cdot\Pr_{\nu\in[2^n]}\big[h_1(\nu) = -1\big]$ and $\alpha(h_2) = 2\cdot\Pr_{\nu\in[2^n]}\big[h_2(\nu) = -1\big]$. A symmetric argument gives

$$\Big|\Pr_{\mu\in[2^{n+k}]}\big[\big(h_1 \times_E h_2\big)(\mu) = (1,-1)\big] - \Pr_{\nu_1\in[2^n],\nu_2\in[2^n]}\big[\big(h_1\times h_2\big)(\nu_1,\nu_2) = (1,-1)\big]\Big| \le \frac{1}{2}\lambda\sqrt{\alpha(h_1)\alpha(h_2)}.$$

Putting the pieces together, we have

$$\beta\big(h_1 \otimes_E h_2\big) - \beta(h_1)\beta(h_2) \le 2\lambda\sqrt{\alpha(h_1)\alpha(h_2)} \le \lambda\cdot\big(\alpha(h_1)+\alpha(h_2)\big)$$

by the arithmetic mean – geometric mean inequality. □

In the proof of Proposition 3 we showed a stronger bound than the one given in the statement, and we weakened it via the arithmetic mean – geometric mean inequality. We did this because our arguments exploit the additive structure of the weaker bound. A result similar to Proposition 3 was proven in [42], though the proof in that paper is direct (not going through the Expander Mixing Lemma) and achieves a slightly different bound.

We are now ready to prove Lemma 3.

Proof of Lemma 3. The generator G_1 we construct has the same form as the generator of Impagliazzo, Nisan, and Wigderson [34] but with a different setting of parameters.

For $\ell = 0,\dots,\log_2 d$ we define a function $G_1^{(\ell)} : \big[2^{\log_2 m + k\ell}\big] \to [m]^{2^\ell}$ as follows, where $k = O(\log\log d)$ is the value corresponding to $\lambda = \frac{1}{8\log_2 d}$ according to Lemma 6. We let $G_1^{(0)}$ be the identity function, and for $\ell > 0$ we let $G_1^{(\ell)} = G_1^{(\ell-1)} \times_{E_\ell} G_1^{(\ell-1)}$ where E_ℓ is an explicit $\big(2^{\log_2 m + k(\ell-1)}, 2^k, \lambda\big)$-expander. Then we take $G_1 = G_1^{(\log_2 d)}$. Note that the seed length of G_1 is $s_1 = \log_2 m + k\cdot\log_2 d = O\big(\log m + \log d\cdot\log\log d\big)$ as required.

We claim that G_1 witnesses Lemma 3. Let $f = \bigotimes_{i\in[d]} f_i$ be an arbitrary (m,d)-checkerboard of weight at least 1. Consider a full binary tree with exactly d leaves which correspond to the coordinates $i = 1,\dots,d$ from left to right. Let ρ denote the root. We say the leaves are at level 0, their parents are at level 1, and so on, with ρ at level $\log_2 d$. For each node v at level ℓ we define a function $f^{(v)} : [m]^{2^\ell} \to \{1,-1\}$ as follows. If v is a leaf, say the ith one, then we let $f^{(v)} = f_i$. If v is an internal node with children v_1 and v_2 then we let $f^{(v)} = f^{(v_1)} \otimes f^{(v_2)}$. In other words, $f^{(v)} = \bigotimes_{i\in[2^\ell]} f_i^{(v)}$ where $f_i^{(v)} = f_{i_v+i-1}$ where i_v is the index of the leftmost leaf in v's subtree. Observe that $f = f^{(\rho)}$. For each node v at level ℓ we define $F^{(v)} = f^{(v)} \circ G_1^{(\ell)}$.

Thus our goal is to show that $\beta\big(F^{(\rho)}\big) \le 3/4$. For each node v at level $\ell \ge 1$ with children v_1 and v_2, applying Observation 2 with $h_1 = f^{(v_1)}$ and $h_2 = f^{(v_2)}$ and $g_1 = g_2 = G_1^{(\ell-1)}$ we find that $F^{(v)} = F^{(v_1)} \otimes_{E_\ell} F^{(v_2)}$. Now we have the following two things.

(i) For each internal node v with children v_1 and v_2, applying Proposition 3 with $h_1 = F^{(v_1)}$ and $h_2 = F^{(v_2)}$ we find that

$$\beta\big(F^{(v)}\big) \le \beta\big(F^{(v_1)}\big)\beta\big(F^{(v_2)}\big) + \lambda\cdot\Big(\alpha\big(F^{(v_1)}\big) + \alpha\big(F^{(v_2)}\big)\Big).$$

(ii) For each leaf v, say the ith one, we have $f^{(v)} \circ G_1^{(0)} = f_i$, and hence

$$\sum_{\text{leaves } v} \alpha\big(F^{(v)}\big) \;=\; \sum_{i \in [d]} \alpha(f_i) \;\geq\; 1.$$

Using the notation $\beta_v = \beta\big(F^{(v)}\big)$ and $\alpha_v = \alpha\big(F^{(v)}\big)$ for each node v, Lemma 3 now follows immediately from the following lemma, which is proved in Section 3.3. □

Lemma 8 (Tree Labeling Lemma). *Suppose a full binary tree with d leaves has each node v labeled with numbers $\alpha_v, \beta_v \in [0,1]$ with $\alpha_v + \beta_v = 1$, such that*

(i) for each internal node v with children v_1 and v_2 we have $\beta_v \leq \beta_{v_1}\beta_{v_2} + \lambda \cdot \big(\alpha_{v_1} + \alpha_{v_2}\big)$ where $\lambda = \frac{1}{8 \log_2 d}$, and

(ii) $\sum_{\text{leaves } v} \alpha_v \geq 1$.

Then the root node ρ satisfies $\beta_\rho \leq 3/4$.

3.3 Proof of the Tree Labeling Lemma

We now prove Lemma 8. We give the intuition in Section 3.3.1 and then the formal argument in Section 3.3.2.

3.3.1 Intuition

Very roughly, the intuition is as follows. For each node v, β_v represents an approximation to the product of β_w over all the leaves w in v's subtree. Thus for the root ρ, β_ρ represents an approximation to $\prod_{\text{leaves } v} \beta_v \leq e^{-\sum_{\text{leaves } v} \alpha_v} \leq 1/e$. However, each internal node v introduces an "error" of $\lambda \cdot \big(\alpha_{v_1} + \alpha_{v_2}\big)$ in addition to the errors already accumulated at the children v_1 and v_2. If these errors are small on average throughout the tree, then β_ρ will be small. On the other hand, if the errors are large on average, then this means the labels α_v are large on average and hence the labels β_v are small on average, so we expect β_ρ to be small in this case as well. However, this is just intuition for why the lemma is true, and the formal argument does not follow the dichotomy suggested by this intuition.

In the formal argument we attempt to reduce to a "worst-case scenario", by which we mean that all the inequalities in both (i) and (ii) in the statement of Lemma 8 hold as equalities. Provided the tree obeys a certain "monotonicity" property, we can decrease the α labels and increase the β labels to reach such a worst-case scenario. For a worst-case scenario, we can argue that the "errors" (as in the previous paragraph) must be small on average, and thus the new value of β_ρ must be small. Since we only increased all the β labels, the original value of β_ρ must also be small.

It turns out that the aforementioned monotonicity property is obeyed provided the β labels of all nodes are not too small. What if β_v is too small for some node v? Then we would like to conclude that β_ρ is small. Unfortunately, in general it might be the case that $\beta_\rho > \beta_v$, for example if v is a child of ρ and the other child of ρ has a β label very close to 1.[2] However, a calculation shows that β_ρ cannot be too much larger than β_v, so we are still safe.

[2]This issue would arise even if we tried to take advantage of the stronger version of Proposition 3 that results by not applying the arithmetic mean – geometric mean inequality.

3.3.2 Formal Argument

First, suppose there exists a node v with $\beta_v \le 1/2$. Then we can prove $\beta_\rho \le 3/4$ as follows. Let $v_0, v_1, v_2, \ldots, v_\ell$ denote the path from v to ρ, with $v = v_0$ and $\rho = v_\ell$. Then for each $i \in \{1, \ldots, \ell\}$, we have $\beta_{v_i} \le \beta_{v_{i-1}} + 2\lambda$ by condition (i) in the statement of Lemma 8. Since $\ell \le \log_2 d$, we conclude that

$$\beta_\rho \;\le\; \beta_v + 2\lambda \cdot \log_2 d \;=\; \beta_v + 1/4 \;\le\; 3/4.$$

Thus we are done, assuming there exists a node v with $\beta_v \le 1/2$. To prove the latter, suppose for contradiction that $\beta_v > 1/2$ holds for all v. We show that this implies $\beta_\rho \le 1/2$, which is a contradiction.

Let us assign new labels $\alpha'_v, \beta'_v \in [0,1]$ with $\alpha'_v + \beta'_v = 1$ to each node v as follows. For the leaf nodes, let α'_v equal α_v except arbitrarily decrease some of them so as to achieve $\sum_{\text{leaves } v} \alpha'_v = 1$. Then working our way up the tree, for each internal node v with children v_1 and v_2 let $\beta'_v = \beta'_{v_1}\beta'_{v_2} + \lambda \cdot \left(\alpha'_{v_1} + \alpha'_{v_2}\right)$. For each internal node v with children v_1 and v_2, define the error label $\delta'_v = \lambda \cdot \left(\alpha'_{v_1} + \alpha'_{v_2}\right)$, and for a leaf v define $\delta'_v = 0$. We say the leaves are at level 0, their parents are at level 1, and so on, with ρ at level $\log_2 d$. We have the following three claims.

Claim 1. *For each node v, we have $\beta'_v \le \prod_{\text{leaves } w \text{ in } v\text{'s subtree}} \beta'_w + \sum_{\text{nodes } w \text{ in } v\text{'s subtree}} \delta'_w$.*

Claim 2. *For each node v at level 2 or higher with children v_1 and v_2, we have $\delta'_v \le \delta'_{v_1} + \delta'_{v_2}$.*

Claim 3. *For each node v, we have $\beta'_v \ge \beta_v$.*

We now stitch the three claims together to get the desired contradiction. By Claim 1 we have

$$\beta'_\rho \;\le\; \textstyle\prod_{\text{leaves } v} \beta'_v + \sum_{\text{nodes } v} \delta'_v \;\le\; e^{-\sum_{\text{leaves } v} \alpha'_v} + \sum_{\text{nodes } v} \delta'_v \;=\; 1/e + \sum_{\text{nodes } v} \delta'_v.$$

Claim 2 implies that for each $\ell \in \{2, \ldots, \log_2 d\}$ we have

$$\textstyle\sum_{\text{nodes } v \text{ at level } \ell} \delta'_v \;\le\; \sum_{\text{nodes } v \text{ at level } \ell - 1} \delta'_v.$$

Since $\sum_{\text{leaves } v} \delta'_v = 0$ and

$$\textstyle\sum_{\text{nodes } v \text{ at level } 1} \delta'_v \;=\; \lambda \cdot \sum_{\text{leaves } v} \alpha'_v \;=\; \lambda$$

we find that $\sum_{\text{nodes } v} \delta'_v \le \lambda \cdot \log_2 d = 1/8$. Using Claim 3 we conclude that

$$\beta_\rho \;\le\; \beta'_\rho \;\le\; 1/e + 1/8 \;\le\; 1/2$$

which is the desired contradiction.

We now prove the three claims. Claim 3 is the only part where we need the assumption that $\beta_v > 1/2$ holds for all nodes v.

Proof of Claim 1. We prove this by structural induction on the tree. If v is a leaf then this holds trivially with equality. Suppose v is an internal node with children v_1 and v_2 and the claim holds for v_1 and v_2. Then we have

$$\begin{aligned} \beta'_v &= \beta'_{v_1}\beta'_{v_2} + \delta'_v \\ &\le \beta'_{v_1} \textstyle\prod_{\text{leaves } w \text{ in } v_2\text{'s subtree}} \beta'_w + \delta'_v + \sum_{\text{nodes } w \text{ in } v_2\text{'s subtree}} \delta'_w \end{aligned}$$

$$\begin{aligned} &\leq \prod_{\text{leaves } w \text{ in } v_1\text{'s subtree}} \beta'_w \cdot \prod_{\text{leaves } w \text{ in } v_2\text{'s subtree}} \beta'_w + \\ &\quad \delta'_v + \sum_{\text{nodes } w \text{ in } v_1\text{'s subtree}} \delta'_w + \sum_{\text{nodes } w \text{ in } v_2\text{'s subtree}} \delta'_w \\ &= \prod_{\text{leaves } w \text{ in } v\text{'s subtree}} \beta'_w + \sum_{\text{nodes } w \text{ in } v\text{'s subtree}} \delta'_w \end{aligned}$$

where the first inequality follows by the induction hypothesis for v_2 and by $\beta'_{v_1} \leq 1$, and the second inequality follows by the induction hypothesis for v_1 and by $\prod_{\text{leaves } w \text{ in } v_2\text{'s subtree}} \beta'_w \leq 1$. □

Proof of Claim 2. Consider a node v at level 2 or higher with children v_1 and v_2. Let $v_{1,1}$ and $v_{1,2}$ be v_1's children, and let $v_{2,1}$ and $v_{2,2}$ be v_2's children. Note that

$$\beta'_{v_1} \;\geq\; \beta'_{v_{1,1}}\beta'_{v_{1,2}} \;\geq\; 1 - \alpha'_{v_{1,1}} - \alpha'_{v_{1,2}}.$$

Thus $\alpha'_{v_1} \leq \alpha'_{v_{1,1}} + \alpha'_{v_{1,2}}$ and similarly $\alpha'_{v_2} \leq \alpha'_{v_{2,1}} + \alpha'_{v_{2,2}}$. It follows that

$$\begin{aligned} \delta'_v &= \lambda \cdot \left(\alpha'_{v_1} + \alpha'_{v_2}\right) \\ &\leq \lambda \cdot \left(\alpha'_{v_{1,1}} + \alpha'_{v_{1,2}} + \alpha'_{v_{2,1}} + \alpha'_{v_{2,2}}\right) \\ &= \delta'_{v_1} + \delta'_{v_2}. \end{aligned}$$

□

Proof of Claim 3. We prove this by structural induction on the tree. For a leaf v, $\beta'_v \geq \beta_v$ holds by definition. For an internal node v with children v_1 and v_2, assume that $\beta'_{v_1} \geq \beta_{v_1}$ and $\beta'_{v_2} \geq \beta_{v_2}$. Then

$$\begin{aligned} \beta'_v - \beta_v &\geq \Big(\beta'_{v_1}\beta'_{v_2} + \lambda \cdot \left(\alpha'_{v_1} + \alpha'_{v_2}\right)\Big) - \Big(\beta_{v_1}\beta_{v_2} + \lambda \cdot \left(\alpha_{v_1} + \alpha_{v_2}\right)\Big) \\ &= \left(\beta'_{v_1} - \beta_{v_1}\right)\left(\beta_{v_2} - \lambda\right) + \left(\beta'_{v_2} - \beta_{v_2}\right)\left(\beta'_{v_1} - \lambda\right) \\ &\geq 0 \end{aligned}$$

since $\beta_{v_2} > 1/2 \geq \lambda$ and $\beta'_{v_1} \geq \beta_{v_1} > 1/2 \geq \lambda$. □

This finishes the proof of Lemma 8.

4 The Low-Weight Case

This section is devoted to the proof of Lemma 2. The main component in the proof of Lemma 2 is Lemma 14 below, and one of the key tools in the proof of Lemma 14 is Lemma 16 below. We prove these three lemmas in Section 4.1, Section 4.2, and Section 4.3 respectively.

4.1 Proof of Lemma 2

We first discuss notation. The parameters m, d, ϵ are fixed, with m and d powers of 2. Let C be the constant from Lemma 1. Given a function $h : U_1 \times U_2 \to U_3$, we use the notation $h(u_1, \cdot)$ to represent the function from U_2 to U_3 that maps u_2 to $h(u_1, u_2)$. The construction has five steps, and we use s_1, s_2, s_3, s_4, s_5 to denote the contributions of the five steps to the final seed length s.

Lemma 9 (Step 1). *There exists an explicit $\epsilon/4$-pseudorandom generator $G_1 : [2^{s_1}] \times [m_1]^{d_1} \to [m]^d$ for the class of (m, d)-checkerboards of weight less than $C \cdot \log_2 \frac{1}{\epsilon}$ with $s_1 = O\big(\log d + \log \frac{1}{\epsilon}\big)$ and $m_1 = (m + d)^{O(1)}$ and $d_1 = \big(\frac{1}{\epsilon}\big)^{O(1)}$ such that for all (m, d)-checkerboards f and all $r_1 \in [2^{s_1}]$, $f \circ G_1(r_1, \cdot)$ is an (m_1, d_1)-checkerboard. Moreover, m_1 and d_1 are powers of 2.*

Lemma 10 (Step 2). *There exists an explicit $\epsilon/4$-pseudorandom generator $G_2 : [2^{s_2}] \times [m_2]^{d_2} \to [m_1]^{d_1}$ for the class of (m_1, d_1)-checkerboards with $s_2 = O\big(\log m + \log d + \log \frac{1}{\epsilon}\big)$ and $m_2 = \big(\frac{1}{\epsilon}\big)^{O(1)}$ and $d_2 = d_1$ such that for all (m_1, d_1)-checkerboards f and all $r_2 \in [2^{s_2}]$, $f \circ G_2(r_2, \cdot)$ is an (m_2, d_2)-checkerboard. Moreover, m_2 is a power of 2.*

Lemma 11 (Step 3). *There exists an explicit $\epsilon/4$-pseudorandom generator $G_3 : [2^{s_3}] \times [m_3]^{d_3} \to [m_2]^{d_2}$ for the class of (m_2, d_2)-checkerboards of weight less than $C \cdot \log_2 \frac{1}{\epsilon/2}$ with $s_3 = O\big(\log^{3/2} \frac{1}{\epsilon}\big)$ and $m_3 = 2^{O(\log^{3/2} \frac{1}{\epsilon})}$ and $d_3 = 2^{O(\log^{1/2} \frac{1}{\epsilon})}$ such that for all (m_2, d_2)-checkerboards f and all $r_3 \in [2^{s_3}]$, $f \circ G_3(r_3, \cdot)$ is an (m_3, d_3)-checkerboard. Moreover, m_3 and d_3 are powers of 2.*

Lemma 12 (Step 4). *There exists an explicit $\epsilon/4$-pseudorandom generator $G_4 : [2^{s_4}] \to [m_3]^{d_3}$ for the class of (m_3, d_3)-checkerboards with $s_4 = O\big(\log^{3/2} \frac{1}{\epsilon}\big)$.*

Lemma 13 (Step 5). *There exists an explicit $\epsilon/2$-pseudorandom generator $G_5 : [2^{s_5}] \to [m_2]^{d_2}$ for the class of (m_2, d_2)-checkerboards of weight at least $C \cdot \log_2 \frac{1}{\epsilon/2}$ with $s_5 = O\big(\log \frac{1}{\epsilon} \cdot \log\log \frac{1}{\epsilon}\big)$.*

The parameters in the first four steps are essentially the same as those used by Lu [43]. At a very high level, the motivation for these steps is as follows. Applying the generator of Impagliazzo, Nisan, and Wigderson [34] directly would give a seed length with poor dependence on the dimension d, so the plan is to first reduce the dimension and then (Step 4) apply the generator of [34]. Step 3 reduces the dimension quite nicely, balancing a tradeoff between how much the dimension is reduced and the cost in seed length to accomplish this dimension reduction. However, achieving the strong parameters of Step 3 requires that the parameter m has been reduced to $\big(\frac{1}{\epsilon}\big)^{O(1)}$. Step 2 accomplishes this, but it requires the dimension to have already been reduced to $\big(\frac{1}{\epsilon}\big)^{O(1)}$. Fortunately, the latter can be accomplished (Step 1) without any further requirements. We refer to Lu's paper [43] for more intuition about the parameters. Step 5 is needed because the dimension reduction steps only work for low-weight checkerboards, but Step 1 and Step 2 do not always preserve the low-weight property.

Lemma 9 and Lemma 11 are special cases of a more general result (Lemma 14 below) which is stated and proven in Section 4.2. The proof is an adaptation of an argument due to Lu [43] (which itself generalizes an argument due to Armoni et al. [9]).

Lemma 10 follows from a result of Nisan and Zuckerman [60], which uses extractors. Lu [43] used a similar lemma for combinatorial rectangles, which he obtained by plugging in an extractor due to Goldreich and Wigderson [29] and giving a somewhat simplified version of Nisan and Zuckerman's argument for his setting. Lemma 10 can be obtained using the same extractor with essentially the same parameters as in [43]. Although Lu's simplified argument does not directly work for combinatorial checkerboards, Nisan and Zuckerman's original argument still applies, yielding Lemma 10. We do not reproduce the proof of Lemma 10 here.

Lemma 12 is an instantiation of the generator of Impagliazzo, Nisan, and Wigderson [34], and Lemma 13 is just an instantiation of Lemma 1. We now prove a simple proposition showing how the above pseudorandom generators can be composed with each other (a similar proposition was used in [9, 43] though with different terminology).

Proposition 4. *Suppose $G'' : [2^{s''}] \times [m'']^{d''} \to [m']^{d'}$ is an ϵ'-pseudorandom generator for some class $\mathcal{C}'$ of (m', d')-checkerboards such that for all (m', d')-checkerboards f and all $r'' \in [2^{s''}]$, $f \circ G''(r'', \cdot)$ is an (m'', d'')-checkerboard. Further suppose $G''' : [2^{s'''}] \to [m'']^{d''}$ is an ϵ''-pseudorandom*

generator for the class of all (m'', d'')-checkerboards. Then the function $G' : [2^{s''}] \times [2^{s'''}] \to [m']^{d'}$ defined by $G'(r'', r''') = G''\big(r'', G'''(r''')\big)$ is an $(\epsilon' + \epsilon'')$-pseudorandom generator for $\mathcal{C}'$.

Proof. Consider any $f \in \mathcal{C}'$. By the pseudorandom property of G'' we get

$$\left| \mathrm{E}_{r'' \in [2^{s''}], u \in [m'']^{d''}} \left[(f \circ G'')(r'', u) \right] - \mathrm{E}_{u \in [m']^{d'}} [f(u)] \right| \;\leq\; \epsilon'. \tag{8}$$

For each $r'' \in [2^{s''}]$, since $f \circ G''(r'', \cdot)$ is an (m'', d'')-checkerboard, by the pseudorandom property of G''' we get

$$\left| \mathrm{E}_{r''' \in [2^{s'''}]} \left[\Big(\big(f \circ G''(r'', \cdot) \big) \circ G''' \Big)(r''') \right] - \mathrm{E}_{u \in [m'']^{d''}} \left[\big(f \circ G''(r'', \cdot) \big)(u) \right] \right| \;\leq\; \epsilon''.$$

Noticing that

$$\Big(\big(f \circ G''(r'', \cdot) \big) \circ G''' \Big)(r''') \;=\; (f \circ G')(r'', r''')$$

and

$$\big(f \circ G''(r'', \cdot) \big)(u) \;=\; (f \circ G'')(r'', u)$$

we find that

$$\begin{aligned} & \left| \mathrm{E}_{r'' \in [2^{s''}], r''' \in [2^{s'''}]} \left[(f \circ G')(r'', r''') \right] - \mathrm{E}_{r'' \in [2^{s''}], u \in [m'']^{d''}} \left[(f \circ G'')(r'', u) \right] \right| \\ \leq\; & \mathrm{E}_{r'' \in [2^{s''}]} \left[\left| \mathrm{E}_{r''' \in [2^{s'''}]} \left[(f \circ G')(r'', r''') \right] - \mathrm{E}_{u \in [m'']^{d''}} \left[(f \circ G'')(r'', u) \right] \right| \right] \\ \leq\; & \epsilon''. \end{aligned}$$

Combining this with Inequality (8) yields

$$\left| \mathrm{E}_{r'' \in [2^{s''}], r''' \in [2^{s'''}]} \left[(f \circ G')(r'', r''') \right] - \mathrm{E}_{u \in [m']^{d'}} [f(u)] \right| \;\leq\; \epsilon' + \epsilon''. \qquad \square$$

Lemma 2 now follows straightforwardly.

Proof of Lemma 2. Combining G_3 with G_4 using Proposition 4 yields an explicit $\epsilon/2$-pseudorandom generator for the class of (m_2, d_2)-checkerboards of weight less than $C \cdot \log_2 \frac{1}{\epsilon/2}$ with seed length $s_3 + s_4$. Combining this generator with G_5 using Proposition 2 we obtain an explicit $\epsilon/2$-pseudorandom generator for the class of all (m_2, d_2)-checkerboards with seed length $s_3 + s_4 + s_5$. Combining this generator with G_2 using Proposition 4 yields an explicit $3\epsilon/4$-pseudorandom generator for the class of (m_1, d_1)-checkerboards with seed length $s_2 + s_3 + s_4 + s_5$. Finally, combining this generator with G_1 using Proposition 4 we obtain an explicit ϵ-pseudorandom generator for the class of (m, d)-checkerboards of weight less than $C \cdot \log_2 \frac{1}{\epsilon}$ with seed length $s_1 + s_2 + s_3 + s_4 + s_5 = O\big(\log m + \log d + \log^{3/2} \frac{1}{\epsilon}\big)$. $\square$

4.2 Dimension Reduction

In this section, m, d, ϵ are free parameters (not necessarily the same as the fixed values that were assumed throughout Section 4.1). Again, C is the constant from Lemma 1.

Lemma 14 (Dimension Reduction). *Let m and d be powers of 2, and let $2 \leq k \leq d$ be an integer parameter. There exists an explicit ϵ-pseudorandom generator $G : [2^s] \times [m']^{d'} \to [m]^d$ for the class of (m,d)-checkerboards of weight less than $C \cdot \log_2 \frac{1}{\epsilon}$ with $s = k \cdot \max(\log_2 d, \log_2 d')$ and $m' = \max(d,m)^k$ and $d' = 2^{O(\frac{1}{k} \cdot \log \frac{1}{\epsilon})}$ such that for all (m,d)-checkerboards f and all $r \in [2^s]$, $f \circ G(r,\cdot)$ is an (m',d')-checkerboard. Moreover, m' and d' are powers of 2.*

To obtain Lemma 9, just use $k = 2$ and plug in $\epsilon/4$ for ϵ. To obtain Lemma 11, just use $k = \Theta\big(\log^{1/2} \frac{1}{\epsilon}\big)$ and plug in m_2 for m, d_2 for d, and $\epsilon/4$ for ϵ. In both instantiations, the generator given by Lemma 14 actually fools a slightly larger class of checkerboards than necessary.

In the proof of Lemma 14 we employ the standard k-wise independent generator. It turns out that using *almost* k-wise independence does not improve the final seed length in Lemma 2, for the same reason it does not help in [43]. We refer to the latter paper for a discussion of this issue.

Lemma 15. *Let n_1, n_2, k be positive integers. There exists an explicit function $h : [2^s] \times [2^{n_1}] \to [2^{n_2}]$ with $s = k \cdot \max(n_1, n_2)$ such that for every $S \subseteq [2^{n_1}]$ with $|S| \leq k$, the random variables $h(r, \nu_1)$ for $\nu_1 \in S$ are fully independent and uniformly distributed, where $r \in [2^s]$ is chosen uniformly at random.*

We also need the following tool.

Lemma 16 (Parity Version of Bonferroni Inequalities). *Let $E_1, \ldots, E_\ell$ be events in any finite probability space. Let p be the probability that an odd number of E_i's hold. For $k = 1, \ldots, \ell$ let*

$$t_k \;=\; \textstyle\sum_{\kappa=1}^{k} (-2)^{\kappa-1} \sum_{S \subseteq [\ell] \,:\, |S| = \kappa} \Pr\big[\bigcap_{i \in S} E_i\big].$$

Then (i) $p \leq t_k$ if k is odd, (ii) $p \geq t_k$ if k is even, and (iii) $p = t_\ell$.

We prove Lemma 16 in Section 4.3 below. We are now ready to prove Lemma 14.

Proof of Lemma 14. Let $d' = 2^{\lceil \frac{2C+1}{k-1} \log_2 \frac{1}{\epsilon} \rceil}$. By convention, we use the notation $i \in [d]$, $j \in [d']$, $u \in [m]$, $w \in [m']$, and $r \in [2^s]$. Let $h_1 : [2^s] \times [d] \to [d']$ and $h_2 : [m'] \times [d] \to [m]$ be k-wise independent generators given by Lemma 15. For $r \in [2^s]$ and $j \in [d']$ we define $I_{r,j} = \big\{ i \in [d] \;:\; h_1(r,i) = j \big\}$.

For $i \in [d]$ we define

$$G(r, w_1, \ldots, w_{d'})_i \;=\; h_2\big(w_{h_1(r,i)}, i\big).$$

That is, we use h_1 to partition the d coordinates into d' buckets, and for each bucket we use an independent seed for h_2 to generate the symbols for the coordinates in that bucket. We claim that G witnesses Lemma 14. For an arbitrary (m,d)-checkerboard $f = \bigotimes_{i \in [d]} f_i$ and arbitrary $r \in [2^s]$, define the (m',d')-checkerboard $f' = \bigotimes_{j \in [d']} f'_j$ where

$$f'_j(w) \;=\; \textstyle\prod_{i \in I_{r,j}} f_i\big(h_2(w,i)\big).$$

Observe that $f'(w_1, \ldots, w_{d'}) = \prod_{i \in [d]} f_i\big(G(r, w_1, \ldots, w_{d'})_i\big)$. Thus $f' = f \circ G(r,\cdot)$ and hence $f \circ G(r,\cdot)$ is an (m',d')-checkerboard. It remains to prove that G is an ϵ-pseudorandom generator for the class of (m,d)-checkerboards of weight less than $C \cdot \log_2 \frac{1}{\epsilon}$. Fix an arbitrary (m,d)-checkerboard $f = \bigotimes_{i \in [d]} f_i$ of weight less than $C \cdot \log_2 \frac{1}{\epsilon}$.

Claim 4. *For every $r \in [2^s]$ and $j \in [d']$ we have*

$$\left| \mathrm{E}_{w\in[m']}\left[\prod_{i\in I_{r,j}} f_i\big(h_2(w,i)\big)\right] - \mathrm{E}_{(u_1,\ldots,u_d)\in[m]^d}\left[\prod_{i\in I_{r,j}} f_i(u_i)\right]\right| \;\le\; \sum_{S\subseteq I_{r,j}\,:\,|S|=k}\prod_{i\in S}\alpha(f_i).$$

Proof of Claim 4. Fix arbitrary $r \in [2^s]$ and $j \in [d']$. For $i \in [d]$ let $\mu_i = -\operatorname{sgn}\big(\mathrm{E}_{u\in[m]}[f_i(u)]\big)$ (and if $\mathrm{E}_{u\in[m]}[f_i(u)] = 0$ then let $\mu_i = \pm 1$ arbitrarily). Note that

$$\Pr_{u\in[m]}\big[f_i(u) = \mu_i\big] \;=\; \alpha(f_i)/2.$$

Define $b_{r,j} = (-1)^{|I_{r,j}|+1}\prod_{i\in I_{r,j}}\mu_i$. For every $(u_1,\ldots,u_d) \in [m]^d$, we have $\prod_{i\in I_{r,j}} f_i(u_i) = b_{r,j}$ if and only if the number of $i \in I_{r,j}$ such that $f_i(u_i) = \mu_i$ is odd. Relative to our fixed r and j, for any distribution $\mathcal{D}$ on $[m]^d$ and any integer $k' \ge 1$ we define

$$t_{k'}^{(\mathcal{D})} \;=\; \sum_{\kappa=1}^{k'}(-2)^{\kappa-1}\sum_{S\subseteq I_{r,j}\,:\,|S|=\kappa}\Pr_{(u_1,\ldots,u_d)\sim\mathcal{D}}\big[\forall i\in S\;:\;f_i(u_i)=\mu_i\big].$$

Applying Lemma 16 identifying $I_{r,j}$ with $[\ell]$ and letting $E_{i'}$ be the event that $f_i(u_i) = \mu_i$ where i is the i'th element of $I_{r,j}$, we find that $\Pr_{(u_1,\ldots,u_d)\sim\mathcal{D}}\big[\prod_{i\in I_{r,j}} f_i(u_i) = b_{r,j}\big]$ lies between $t_{k-1}^{(\mathcal{D})}$ and $t_k^{(\mathcal{D})}$ inclusive. (This follows from part (i) and part (ii) when $k \le |I_{r,j}|$ and from part (iii) when $k > |I_{r,j}|$.) Now let $\mathcal{U}$ denote the uniform distribution over $[m]^d$, and let $\mathcal{D}$ be the distribution $\big(h_2(w,1),\ldots,h_2(w,d)\big)$ where $w \in [m']$ is chosen uniformly at random. By the k-wise independence of h_2, we have $t_k^{(\mathcal{D})} = t_k^{(\mathcal{U})}$ and $t_{k-1}^{(\mathcal{D})} = t_{k-1}^{(\mathcal{U})}$. Now since $\Pr_{w\in[m']}\big[\prod_{i\in I_{r,j}} f_i\big(h_2(w,i)\big) = b_{r,j}\big]$ and $\Pr_{(u_1,\ldots,u_d)\in[m]^d}\big[\prod_{i\in I_{r,j}} f_i(u_i) = b_{r,j}\big]$ both lie between $t_{k-1}^{(\mathcal{U})}$ and $t_k^{(\mathcal{U})}$ inclusive, we have

$$\begin{aligned}
&\left|\Pr_{w\in[m']}\left[\prod_{i\in I_{r,j}} f_i\big(h_2(w,i)\big) = b_{r,j}\right] - \Pr_{(u_1,\ldots,u_d)\in[m]^d}\left[\prod_{i\in I_{r,j}} f_i(u_i) = b_{r,j}\right]\right| \\
&\le\; \big|t_k^{(\mathcal{U})} - t_{k-1}^{(\mathcal{U})}\big| \\
&=\; 2^{k-1}\sum_{S\subseteq I_{r,j}\,:\,|S|=k}\Pr_{(u_1,\ldots,u_d)\in[m]^d}\big[\forall i\in S\;:\;f_i(u_i)=\mu_i\big] \\
&=\; 2^{k-1}\sum_{S\subseteq I_{r,j}\,:\,|S|=k}\prod_{i\in S}\alpha(f_i)/2 \\
&=\; \tfrac{1}{2}\sum_{S\subseteq I_{r,j}\,:\,|S|=k}\prod_{i\in S}\alpha(f_i).
\end{aligned}$$

It follows that

$$\begin{aligned}
&\left|\mathrm{E}_{w\in[m']}\left[\prod_{i\in I_{r,j}} f_i\big(h_2(w,i)\big)\right] - \mathrm{E}_{(u_1,\ldots,u_d)\in[m]^d}\left[\prod_{i\in I_{r,j}} f_i(u_i)\right]\right| \\
&=\; 2\cdot\left|\Pr_{w\in[m']}\left[\prod_{i\in I_{r,j}} f_i\big(h_2(w,i)\big) = b_{r,j}\right] - \Pr_{(u_1,\ldots,u_d)\in[m]^d}\left[\prod_{i\in I_{r,j}} f_i(u_i) = b_{r,j}\right]\right| \\
&\le\; \sum_{S\subseteq I_{r,j}\,:\,|S|=k}\prod_{i\in S}\alpha(f_i).
\end{aligned}$$

This finishes the proof of the claim. □

We now continue with the proof of Lemma 14. We have

$$\left|\mathrm{E}_{r\in[2^s],(w_1,\ldots,w_{d'})\in[m']^{d'}}\big[(f\circ G)(r,w_1,\ldots,w_{d'})\big] - \mathrm{E}_{(u_1,\ldots,u_d)\in[m]^d}\big[f(u_1,\ldots,u_d)\big]\right|$$

$$
\begin{aligned}
&\leq \mathrm{E}_{r\in[2^s]}\left[\left|\mathrm{E}_{(w_1,\ldots,w_{d'})\in[m']^{d'}}\left[(f\circ G)(r,w_1,\ldots,w_{d'})\right]-\mathrm{E}_{(u_1,\ldots,u_d)\in[m]^d}\left[f(u_1,\ldots,u_d)\right]\right|\right]\\
&= \mathrm{E}_{r\in[2^s]}\left[\left|\prod_{j\in[d']}\mathrm{E}_{w\in[m']}\left[\prod_{i\in I_{r,j}}f_i\big(h_2(w,i)\big)\right]-\prod_{j\in[d']}\mathrm{E}_{(u_1,\ldots,u_d)\in[m]^d}\left[\prod_{i\in I_{r,j}}f_i(u_i)\right]\right|\right]\\
&\leq \mathrm{E}_{r\in[2^s]}\left[\sum_{j\in[d']}\left|\mathrm{E}_{w\in[m']}\left[\prod_{i\in I_{r,j}}f_i\big(h_2(w,i)\big)\right]-\mathrm{E}_{(u_1,\ldots,u_d)\in[m]^d}\left[\prod_{i\in I_{r,j}}f_i(u_i)\right]\right|\right]\\
&\leq \mathrm{E}_{r\in[2^s]}\left[\sum_{j\in[d']}\sum_{S\subseteq I_{r,j}\,:\,|S|=k}\prod_{i\in S}\alpha(f_i)\right]\\
&= \sum_{j\in[d']}\sum_{S\subseteq[d]\,:\,|S|=k}\Pr_{r\in[2^s]}[S\subseteq I_{r,j}]\cdot\prod_{i\in S}\alpha(f_i)\\
&= \sum_{j\in[d']}\sum_{S\subseteq[d]\,:\,|S|=k}\tfrac{1}{(d')^k}\cdot\prod_{i\in S}\alpha(f_i)\\
&= \tfrac{1}{(d')^{k-1}}\sum_{S\subseteq[d]\,:\,|S|=k}\prod_{i\in S}\alpha(f_i)\\
&\leq \tfrac{1}{(d')^{k-1}}\cdot\tfrac{1}{k!}\sum_{(i_1,\ldots,i_k)\in[d]^k}\prod_{\kappa\in[k]}\alpha(f_{i_\kappa})\\
&= \tfrac{1}{(d')^{k-1}}\cdot\tfrac{1}{k!}\Big(\sum_{i\in[d]}\alpha(f_i)\Big)^k\\
&< \tfrac{1}{(d')^{k-1}}\cdot\tfrac{1}{k!}\big(C\cdot\log_2\tfrac{1}{\epsilon}\big)^k
\end{aligned}
$$

where the fourth line follows by the simple fact that for all $x_1,\ldots,x_{d'},y_1,\ldots,y_{d'}\in[-1,1]$ we have $\big|\prod_{j\in[d']}x_j-\prod_{j\in[d']}y_j\big|\leq\sum_{j\in[d']}|x_j-y_j|$, the fifth line follows by Claim 4, the seventh line follows by the k-wise independence of h_1, and the other lines follow by simple manipulations. All that remains is to show that $\frac{1}{(d')^{k-1}}\cdot\frac{1}{k!}\big(C\cdot\log_2\frac{1}{\epsilon}\big)^k\leq\epsilon$. We have $(d')^{k-1}\geq\big(\frac{1}{\epsilon}\big)^{2C+1}$. Using the standard bound $k!\geq\big(\frac{k}{e}\big)^k$ we have

$$
\tfrac{1}{k!}\big(C\cdot\log_2\tfrac{1}{\epsilon}\big)^k \;\leq\; \left(\tfrac{e\cdot C\cdot\log_2\frac{1}{\epsilon}}{k}\right)^k \;\leq\; e^{C\cdot\log_2\frac{1}{\epsilon}} \;\leq\; \big(\tfrac{1}{\epsilon}\big)^{2C}
$$

where the middle inequality follows by a little calculus (and holds no matter what k is). We conclude that $\frac{1}{(d')^{k-1}}\cdot\frac{1}{k!}\big(C\cdot\log_2\frac{1}{\epsilon}\big)^k\leq\big(\frac{1}{\epsilon}\big)^{2C}/\big(\frac{1}{\epsilon}\big)^{2C+1}=\epsilon$. This finishes the proof of Lemma 14.□

4.3 Proof of Lemma 16

Let Ω denote the sample space. Let $P:\Omega\to\{0,1\}$ be the indicator for the event that an odd number of E_i's hold. For $k=1,\ldots,\ell$ let $T_k:\Omega\to\mathbb{Z}$ be defined by

$$
T_k(x) \;=\; \sum_{\kappa=1}^{k}(-2)^{\kappa-1}\sum_{S\subseteq[\ell]\,:\,|S|=\kappa}\chi_S(x)
$$

where $\chi_S:\Omega\to\{0,1\}$ is the indicator for the event $\bigcap_{i\in S}E_i$. We have $p=\mathrm{E}[P]$ and $t_k=\mathrm{E}[T_k]$. To prove the lemma, it suffices to show that for all $x\in\Omega$, (i) $P(x)\leq T_k(x)$ if k is odd, (ii) $P(x)\geq T_k(x)$ if k is even, and (iii) $P(x)=T_\ell(x)$.

Fix an arbitrary $x\in\Omega$ and let $\ell_x=\big|\{i\;:\;x\in E_i\}\big|$. Note that

$$
T_k(x) \;=\; \sum_{\kappa=1}^{\min(k,\ell_x)}(-2)^{\kappa-1}\binom{\ell_x}{\kappa} \;=\; \tfrac{1}{2}-\tfrac{1}{2}T'_k(x)
$$

where

$$
T'_k(x) \;=\; \sum_{\kappa=0}^{\min(k,\ell_x)}(-2)^{\kappa}\binom{\ell_x}{\kappa}.
$$

Now if $k \geq \ell_x$ then by the binomial theorem we have

$$T'_k(x) \;=\; \textstyle\sum_{\kappa=0}^{\ell_x} 1^{\ell_x-\kappa}(-2)^\kappa \binom{\ell_x}{\kappa} \;=\; (-1)^{\ell_x}$$

and thus $P(x) = T_k(x)$. This establishes (iii) since $\ell \geq \ell_x$, and it establishes (i) and (ii) assuming $k \geq \ell_x$. Now assume $k \leq \ell_x$. We claim that $T'_k(x) \leq -1$ if k is odd, and $T'_k(x) \geq 1$ if k is even. Assuming the claim, (i) follows by $P(x) \leq 1$, and (ii) follows by $P(x) \geq 0$. We already established the claim for $k = \ell_x$, and the claim trivially holds for $k = 0$. For $\kappa = 0, \ldots, \ell_x$ define $\tau_\kappa(x) = 2^\kappa \binom{\ell_x}{\kappa}$ so that $T'_k(x) = \sum_{\kappa=0}^{k} (-1)^\kappa \tau_\kappa(x)$. Note that the sequence $\tau_0(x), \tau_1(x), \ldots, \tau_{\ell_x}(x)$ is unimodal, since for $\kappa \geq 1$ we have

$$\tau_\kappa(x)/\tau_{\kappa-1}(x) \;=\; 2(\ell_x - \kappa + 1)/\kappa$$

which is at least 1 when $\kappa \leq \frac{2}{3}(\ell_x + 1)$ and at most 1 when $\kappa \geq \frac{2}{3}(\ell_x + 1)$. We now show by induction on $k = 0, 1, \ldots, \lfloor \frac{2}{3}(\ell_x + 1) \rfloor$ that the claim holds for these values of k (and a symmetric argument shows by induction on $k = \ell_x, \ell_x - 1, \ldots, \lceil \frac{2}{3}(\ell_x + 1) \rceil$ that the claim holds for these values of k). For the base cases, we know the claim holds for $k = 0$, and for $k = 1$ we have $T'_k(x) = 1 - 2\ell_x \leq -1$ since $\ell_x \geq k = 1$. Now assuming the claim holds for $k - 2$, we prove it for k. Assume k is even (a symmetric argument handles the case k is odd). We have $T'_{k-2}(x) \geq 1$ and $T'_k(x) = T'_{k-2}(x) - \tau_{k-1}(x) + \tau_k(x) \geq 1$ since $\tau_k(x) \geq \tau_{k-1}(x)$. This finishes the proof of Lemma 16.

Acknowledgments

I thank Siu Man Chan, Siu On Chan, Anindya De, and Luca Trevisan for helpful discussions.

References

[1] S. Aaronson and D. van Melkebeek. A note on circuit lower bounds from derandomization. Technical Report TR10-105, Electronic Colloquium on Computational Complexity, 2010.

[2] M. Ajtai, J. Komlós, and E. Szemerédi. Deterministic simulation in LOGSPACE. In *Proceedings of the 19th ACM Symposium on Theory of Computing*, pages 132–140, 1987.

[3] M. Ajtai and A. Wigderson. Deterministic simulation of probabilistic constant-depth circuits. *Advances in Computing Research – Randomness and Computation*, 5:199–223, 1989.

[4] N. Alon and F. Chung. Explicit construction of linear sized tolerant networks. *Discrete Mathematics*, 72(1-3):15–19, 1988.

[5] N. Alon, Z. Galil, and V. Milman. Better expanders and superconcentrators. *Journal of Algorithms*, 8(3):337–347, 1987.

[6] N. Alon, O. Goldreich, J. Håstad, and R. Peralta. Simple constructions of almost k-wise independent random variables. *Random Structures and Algorithms*, 3(3):289–304, 1992.

[7] N. Alon and V. Milman. λ_1, isoperimetric inequalities for graphs, and superconcentrators. *Journal of Combinatorial Theory, Series B*, 38(1):73–88, 1985.

[8] R. Armoni. On the derandomization of space-bounded computations. In *Proceedings of the 2nd International Workshop on Randomization and Computation*, pages 47–59, 1998.

[9] R. Armoni, M. Saks, A. Wigderson, and S. Zhou. Discrepancy sets and pseudorandom generators for combinatorial rectangles. In *Proceedings of the 37th IEEE Symposium on Foundations of Computer Science*, pages 412–421, 1996.

[10] R. Armoni, A. Ta-Shma, A. Wigderson, and S. Zhou. An $O(\log^{4/3}(n))$ space algorithm for (s,t) connectivity in undirected graphs. *Journal of the ACM*, 47(2):294–311, 2000.

[11] L. Babai, L. Fortnow, N. Nisan, and A. Wigderson. BPP has subexponential time simulations unless EXPTIME has publishable proofs. *Computational Complexity*, 3:307–318, 1993.

[12] L. Babai, N. Nisan, and M. Szegedy. Multiparty protocols, pseudorandom generators for logspace, and time-space trade-offs. *Journal of Computer and System Sciences*, 45(2):204–232, 1992.

[13] L. Bazzi. Polylogarithmic independence can fool DNF formulas. *SIAM Journal on Computing*, 38(6):2220–2272, 2009.

[14] M. Bellare and J. Rompel. Randomness-efficient oblivious sampling. In *Proceedings of the 35th IEEE Symposium on Foundations of Computer Science*, pages 276–287, 1994.

[15] A. Bogdanov. Pseudorandom generators for low degree polynomials. In *Proceedings of the 37th ACM Symposium on Theory of Computing*, pages 21–30, 2005.

[16] A. Bogdanov, Z. Dvir, E. Verbin, and A. Yehudayoff. Pseudorandomness for width 2 branching programs. Technical Report TR09-070, Electronic Colloquium on Computational Complexity, 2009.

[17] A. Bogdanov and E. Viola. Pseudorandom bits for polynomials. *SIAM Journal on Computing*, 39(6):2464–2486, 2010.

[18] M. Braverman. Polylogarithmic independence fools AC^0 circuits. *Journal of the ACM*, 57(5), 2010.

[19] M. Braverman, A. Rao, R. Raz, and A. Yehudayoff. Pseudorandom generators for regular branching programs. In *Proceedings of the 51st IEEE Symposium on Foundations of Computer Science (to appear)*, 2010.

[20] J. Brody and E. Verbin. The coin problem, and pseudorandomness for branching programs. In *Proceedings of the 51st IEEE Symposium on Foundations of Computer Science (to appear)*, 2010.

[21] J.-Y. Cai, V. Chakaravarthy, and D. van Melkebeek. Time-space tradeoff in derandomizing probabilistic logspace. *Theory of Computing Systems*, 39(1):189–208, 2006.

[22] A. De, O. Etesami, L. Trevisan, and M. Tulsiani. Improved pseudorandom generators for depth 2 circuits. In *Proceedings of the 14th International Workshop on Randomization and Computation*, pages 504–517, 2010.

[23] I. Diakonikolas, P. Gopalan, R. Jaiswal, R. Servedio, and E. Viola. Bounded independence fools halfspaces. *SIAM Journal on Computing*, 39(8):3441–3462, 2010.

[24] I. Diakonikolas, D. Kane, and J. Nelson. Bounded independence fools degree-2 threshold functions. In *Proceedings of the 51st IEEE Symposium on Foundations of Computer Science (to appear)*, 2010.

[25] G. Even, O. Goldreich, M. Luby, N. Nisan, and B. Velickovic. Efficient approximation of product distributions. *Random Structures and Algorithms*, 13(1):1–16, 1998.

[26] L. Fortnow and A. Klivans. Linear advice for randomized logarithmic space. In *Proceedings of the 23rd International Symposium on Theoretical Aspects of Computer Science*, pages 469–476, 2006.

[27] O. Gabber and Z. Galil. Explicit constructions of linear-sized superconcentrators. *Journal of Computer and System Sciences*, 22(3):407–420, 1981.

[28] O. Goldreich. In a world of P = BPP. Technical Report TR10-135, Electronic Colloquium on Computational Complexity, 2010.

[29] O. Goldreich and A. Wigderson. Tiny families of functions with random properties: A quality-size trade-off for hashing. *Random Structures and Algorithms*, 11(4):315–343, 1997.

[30] P. Gopalan, R. O'Donnell, Y. Wu, and D. Zuckerman. Fooling functions of halfspaces under product distributions. In *Proceedings of the 25th IEEE Conference on Computational Complexity*, pages 223–234, 2010.

[31] P. Harsha, A. Klivans, and R. Meka. An invariance principle for polytopes. In *Proceedings of the 42nd ACM Symposium on Theory of Computing*, pages 543–552, 2010.

[32] R. Impagliazzo and V. Kabanets. Derandomizing polynomial identity tests means proving circuit lower bounds. *Computational Complexity*, 13(1-2):1–46, 2004.

[33] R. Impagliazzo, V. Kabanets, and A. Wigderson. In search of an easy witness: Exponential time vs. probabilistic polynomial time. *Journal of Computer and System Sciences*, 65(4):672–694, 2002.

[34] R. Impagliazzo, N. Nisan, and A. Wigderson. Pseudorandomness for network algorithms. In *Proceedings of the 26th ACM Symposium on Theory of Computing*, pages 356–364, 1994.

[35] R. Impagliazzo, R. Shaltiel, and A. Wigderson. Reducing the seed length in the Nisan-Wigderson generator. *Combinatorica*, 26(6):647–681, 2006.

[36] R. Impagliazzo and A. Wigderson. P = BPP if E requires exponential circuits: Derandomizing the XOR lemma. In *Proceedings of the 29th ACM Symposium on Theory of Computing*, pages 220–229, 1997.

[37] S. Jimbo and A. Maruoka. Expanders obtained from affine transformations. *Combinatorica*, 7(4):343–355, 1987.

[38] M. Koucký, P. Nimbhorkar, and P. Pudlak. Pseudorandom generators for group products. Technical Report TR10-113, Electronic Colloquium on Computational Complexity, 2010.

[39] N. Linial, M. Luby, M. Saks, and D. Zuckerman. Efficient construction of a small hitting set for combinatorial rectangles in high dimension. *Combinatorica*, 17(2):215–234, 1997.

[40] S. Lovett. Unconditional pseudorandom generators for low degree polynomials. *Theory of Computing*, 5(1):69–82, 2009.

[41] S. Lovett, P. Mukhopadhyay, and A. Shpilka. Pseudorandom generators for $\mathrm{CC}_0[p]$ and the Fourier spectrum of low-degree polynomials over finite fields. In *Proceedings of the 51st IEEE Symposium on Foundations of Computer Science (to appear)*, 2010.

[42] S. Lovett, O. Reingold, L. Trevisan, and S. Vadhan. Pseudorandom bit generators that fool modular sums. In *Proceedings of the 13th International Workshop on Randomization and Computation*, pages 615–630, 2009.

[43] C.-J. Lu. Improved pseudorandom generators for combinatorial rectangles. *Combinatorica*, 22(3):417–434, 2002.

[44] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

[45] M. Luby and B. Velickovic. On deterministic approximation of DNF. *Algorithmica*, 16(4-5):415–433, 1996.

[46] M. Luby, B. Velickovic, and A. Wigderson. Deterministic approximate counting of depth-2 circuits. In *Proceedings of the 2nd Israel Symposium on Theory of Computing Systems*, pages 18–24, 1993.

[47] G. Margulis. Explicit constructions of expanders. *Problemy Peredaci Informacii*, 9(4):71–80, 1973.

[48] G. Margulis. Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators. *Problems of Information Transmission*, 24(1):39–46, 1988.

[49] R. Meka and D. Zuckerman. Small-bias spaces for group products. In *Proceedings of the 13th International Workshop on Randomization and Computation*, pages 658–672, 2009.

[50] R. Meka and D. Zuckerman. Pseudorandom generators for polynomial threshold functions. In *Proceedings of the 42nd ACM Symposium on Theory of Computing*, pages 427–436, 2010.

[51] M. Morgenstern. Existence and explicit constructions of $q+1$ regular Ramanujan graphs for every prime power q. *Journal of Combinatorial Theory, Series B*, 62(1):44–62, 1994.

[52] E. Mossel, A. Shpilka, and L. Trevisan. On epsilon-biased generators in NC^0. *Random Structures and Algorithms*, 29(1):56–81, 2006.

[53] J. Naor and M. Naor. Small-bias probability spaces: Efficient constructions and applications. *SIAM Journal on Computing*, 22(4):838–856, 1993.

[54] N. Nisan. Pseudorandom bits for constant depth circuits. *Combinatorica*, 11(1):63–70, 1991.

[55] N. Nisan. Pseudorandom generators for space-bounded computation. *Combinatorica*, 12(4):449–461, 1992.

[56] N. Nisan. On read-once vs. multiple access to randomness in logspace. *Theoretical Computer Science*, 107(1):135–144, 1993.

[57] N. Nisan. RL $\subseteq$ SC. *Computational Complexity*, 4:1–11, 1994.

[58] N. Nisan, E. Szemerédi, and A. Wigderson. Undirected connectivity in $O(\log^{1.5}(n))$ space. In *Proceedings of the 33rd IEEE Symposium on Foundations of Computer Science*, pages 24–29, 1992.

[59] N. Nisan and A. Wigderson. Hardness vs. randomness. *Journal of Computer and System Sciences*, 49(2):149–167, 1994.

[60] N. Nisan and D. Zuckerman. Randomness is linear in space. *Journal of Computer and System Sciences*, 52(1):43–52, 1996.

[61] R. Raz and O. Reingold. On recycling the randomness of states in space bounded computation. In *Proceedings of the 31st ACM Symposium on Theory of Computing*, pages 159–168, 1999.

[62] A. Razborov. A simple proof of Bazzi's Theorem. *ACM Transactions on Computation Theory*, 1(1), 2009.

[63] O. Reingold. Undirected connectivity in log-space. *Journal of the ACM*, 55(4), 2008.

[64] O. Reingold, L. Trevisan, and S. Vadhan. Pseudorandom walks on regular digraphs and the RL vs. L problem. In *Proceedings of the 38th ACM Symposium on Theory of Computing*, pages 457–466, 2006.

[65] O. Reingold, S. Vadhan, and A. Wigderson. Entropy waves, the zig-zag graph product, and new constant-degree expanders. *Annals of Mathematics*, 155(1):157–187, 2002.

[66] E. Rozenman and S. Vadhan. Derandomized squaring of graphs. In *Proceedings of the 9th International Workshop on Randomization and Computation*, pages 436–447, 2005.

[67] M. Saks and S. Zhou. $\mathrm{BP_HSPACE}(S) \subseteq \mathrm{DSPACE}(S^{3/2})$. *Journal of Computer and System Sciences*, 58(2):376–403, 1999.

[68] R. Shaltiel and C. Umans. Simple extractors for all min-entropies and a new pseudorandom generator. *Journal of the ACM*, 52(2):172–216, 2005.

[69] J. Sima and S. Zak. A polynomial time construction of a hitting set for read-once branching programs of width 3. Technical Report TR10-088, Electronic Colloquium on Computational Complexity, 2010.

[70] M. Sudan, L. Trevisan, and S. Vadhan. Pseudorandom generators without the XOR lemma. *Journal of Computer and System Sciences*, 62(2):236–266, 2001.

[71] C. Umans. Pseudo-random generators for all hardnesses. *Journal of Computer and System Sciences*, 67(2):419–440, 2003.

[72] E. Viola. Pseudorandom bits for constant-depth circuits with few arbitrary symmetric gates. *SIAM Journal on Computing*, 36(5):1387–1403, 2007.

[73] E. Viola. The sum of d small-bias generators fools polynomials of degree d. *Computational Complexity*, 18(2):209–217, 2009.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de