Electronic Colloquium on Computational Complexity, Report No. 33 (2022)

A better-than-3 $\log n$ depth lower bound for De Morgan formulas with restrictions on top gates

Ivan Mihajlin[1] and Anastasia Sofronova[1, 2]

[1]Leonhard Euler International Mathematical Institute in Saint Petersburg
[2]St. Petersburg Department of Steklov Mathematical Institute of Russian Academy of Sciences
ivmihajlin@gmail.com, ana.a.sofronova@gmail.com

March 6, 2022

Abstract

We prove that a modification of Andreev's function is not computable by $(3+\alpha-\varepsilon)\log n$ depth De Morgan formula with $(2\alpha-\varepsilon)\log n$ layers of AND gates at the top for any $0 < \alpha < \frac{1}{5}$ and any constant $\varepsilon > 0$. In order to do this, we prove a weak variant of Karchmer-Raz-Wigderson conjecture. To be more precise, we prove the existence of two functions $f\colon \{0,1\}^n \to \{0,1\}$ and $g\colon \{0,1\}^n \to \{0,1\}^n$ such that $f(g(x) \oplus y)$ is not computable by depth $(1+\alpha-\varepsilon)n$ formulas with $(2\alpha-\varepsilon)n$ layers of AND gates at the top. We do this by a top-down approach, which was only used before for depth-3 model.

Our technical contribution includes combinatorial insights into structure of composition with random boolean function, which led us to introducing a notion of well-mixed sets. A set of functions is well-mixed if, when composed with a random function, it does not have subsets that agree on large fractions of inputs. We use probabilistic method to prove the existence of well-mixed sets.

1 Introduction

Proving lower bounds on Boolean formulas remains one of the fundamental problems in complexity theory. Specifically, one of the major open question here is separating classes **P** and $\mathbf{NC}^1$ by proving a super-logarithmic depth lower bound for a function from **P**. The long line of prior work includes [Sub61], [Khr71], [And87], [PZ91], [IN93] up to the currently best depth lower bound $(3-o(1))\log n$ from the celebrated paper by Håstad [Hås98] which stands unbeaten for two decades up to lower order terms [Tal14].

Karchmer, Raz and Wigderson [KRW95] proposed an approach for attacking this problem, introducing *a block composition* of two Boolean functions:

Definition 1. *The block composition* $f \diamond g\colon (\{0,1\}^m)^n \to \{0,1\}$ *of two Boolean functions* $f\colon \{0,1\}^n \to \{0,1\}$ *and* $g\colon \{0,1\}^m \to \{0,1\}$ *is defined as follows:*

$$(f \diamond g)(x_1, \ldots, x_n) = f(g(x_1), \ldots, g(x_n))$$

where $x_i \in \{0,1\}^m$.

Let $D(f)$ be the minimal depth of a formula computing f. It is easy to see that $f \diamond g$ can be computed by a formula of depth $D(f) + D(g)$. Karchmer, Raz and Wigderson conjectured that this bound is roughly optimal.

ISSN 1433-8092

Conjecture 1 (KRW conjecture)**.** *For any non-constant functions* $f : \{0,1\}^n \to \{0,1\}$ *and* $g : \{0,1\}^m \to \{0,1\}$*:*

$$D(f \diamond g) \approx D(f) + D(g)$$

The symbol "approximately equal" could be interpreted in a number of ways, but pretty much all reasonable interpretations, should the conjecture be proven, imply $\mathbf{P} \not\subseteq \mathbf{NC}^1$. In fact, while in the original conjecture there is $\forall$ quantifier for both f and g, the existence of such g for every f would be quite enough.

As an example, let us formulate a weaker version of conjecture, from which $\mathbf{P} \not\subseteq \mathbf{NC}^1$ would still follow.

Conjecture 2 (KRW conjecture, weaker version)**.** *There exists a constant* ε *such that for any* n *and* m *and any non-constant* $f : \{0,1\}^n \to \{0,1\}$ *there exists* $g : \{0,1\}^m \to \{0,1\}$ *such that*

$$D(f \diamond g) \geq D(f) + \varepsilon m$$

KRW conjecture was extensively studied in a series of work [EIRS01], [HW90], [GMWW17], [DM16], [KM18], [dRMN$^+$20], mostly from communication complexity point of view. To present a thorough overview, we include the necessary definitions as well.

For a function $f : \{0,1\}^n \to \{0,1\}$, let KW_f (Karchmer-Wigderson game for a function f) be a communication problem, where Alice gets $x \in f^{-1}(0)$, Bob gets $y \in f^{-1}(1)$, and they need to find i such that $x_i \neq y_i$. In [KW88] it was observed that $D(f) = CC(\mathrm{KW}_f)$, where $CC(R)$ denotes the minimal depth of a communication protocol solving relation R.

KRW conjecture can be reformulated in those terms as $CC(\mathrm{KW}_{f \diamond g}) \approx CC(\mathrm{KW}_f) + CC(\mathrm{KW}_g)$.

Karchmer-Wigderson games have been successfully applied to a monotone setting, separating monotone $\mathbf{NC}^1$ and $\mathbf{NC}^2$ [KW88]. There have been attempts to tackle monotone KRW conjecture [dRMN$^+$20], where the authors introduced also a semi-monotone setup. [EIRS01, HW90] proved a lower bound for a block-composition of two *universal relations*.

Definition 2. $U_n = \{(x, y, i) \mid x, y \in \{0,1\}^n, x_i \neq y_i\} \cup \{(x, x, \bot) \mid x \in \{0,1\}^n\}$

In a sense, universal relation generalizes KW_f for any f, since a protocol for U_n can be used to solve KW_f as well. The difference is that in U_n, inputs for players do not come from two disjoint sets. The same way as universal relation generalizes KW games for functions, their composition generalizes KW games for composition of functions. There does not seem to be any formula lower bounds that follow from communication lower bounds involving universal relation, but it can be considered more of a stepping stone to hone our techniques before dealing with actual function.

Currently there exist lower bounds in the following setups:

- a lower bound on $U_n \diamond U_n$ [EIRS01, HW90];
- a lower bound on $\mathrm{KW}_f \diamond U_n$ for any f [GMWW17, KM18];
- a lower bound on $U_n \diamond \mathrm{KW}_g$ and $U_n \boxplus_2 \mathrm{KW}_g$ for some g [MS21].

Here the operation $\boxplus_m$ is defined in the following way:

Definition 3 ([MS21])**.** *XOR-composition* $f \boxplus_m g \colon (\{0,1\}^n)^n \to \{0,1\}$ *of functions* $f \colon \{0,1\}^n \to \{0,1\}$ *and* $g \colon \{0,1\}^n \to \{0,1\}^n$ *is defined as follows:*

$$(f \boxplus_m g)(x_1, \ldots, x_m) = f(g(x_1) \oplus \ldots \oplus g(x_m))$$

.

where $x_i \in \{0,1\}^n$.

In [MS21] authors also define another version of XOR-composition, which we use throughout the paper:

Definition 4 ([MS21])**.** *XOR-composition* $f \boxplus g \colon \{0,1\}^{2n} \to \{0,1\}$ *of two function* $f \colon \{0,1\}^n \to \{0,1\}$ *and* $g \colon \{0,1\}^n \to \{0,1\}^n$ *is defined as follows:*

$$(f \boxplus g)(x, y) \coloneqq f(g(x) \oplus y)$$

This differs from the definition of $\boxplus_2$ in the sense that g is not applied to a second argument.

In the same paper, the authors stated a variant of KRW conjecture using XOR-composition instead of block-composition. This variant of the conjecture also implies $\mathbf{P} \not\subseteq \mathbf{NC}^1$. Moreover, [MS21] also introduced a variant of XOR-KRW regarding size of the formulas. If proved, this would imply a supercubic lower bound on a modified Andreev's function.

To outline a general idea of how different variants of KRW work, for $\mathbf{P} \not\subseteq \mathbf{NC}^1$ we need to prove a variant of conjecture of the form "$\forall f \exists g$ such that the depth of a formula for $f \circ g$ (for a reasonable definition of $\circ$) noticeably increases in comparison to a depth of a formula for f". For beating cubic size lower bound for Andreev's function, we only need to prove that "$\exists f, g$ such that the formula size for $f \circ g$ is big enough".

The next step following the lower bounds mentioned above would be getting rid of universal relation as both inner and outer parts of the composition, since for formula lower bounds of any form we need proper functions there. We are able to do that with restrictions on top gates of the formula:

Theorem 1 (Main theorem)**.** *With probability* $1 - o(1)$ *for a random function* $f \colon \{0,1\}^n \to \{0,1\}$*, there exists a function* $g \colon \{0,1\}^n \to \{0,1\}^n$*, such that* $f \boxplus g$ *is not computable by an AND of* $2^{(2\alpha-\varepsilon)n}$ *formulas of size at most* $2^{(1-\alpha-\varepsilon)n}$ *for any* $0 < \alpha \leq \frac{1}{5} - 0.01$ *and any constant* $0 < \varepsilon < \alpha$.

AND in the statement could be replaced by OR, and 0.01 serves as an arbitrary small constant separating α from $\frac{1}{5}$. We also obtain a new lower bound on modified Andreev's function, again with the restriction on top gates of the formula.

Our approach equips a technique from [EIRS01] of tracking a suitable subadditive measure with new combinatorial insights.

The plan of the proof of Theorem 1 could be briefly summarized as follows:

- we sample a set of functions such that any big enough subset of its compositions with f have very few zeroes in common;
- we track a certain subadditive measure while we walk down the trees of formulas for the compositions;
- we consider different subformulas obtained in this way, and argue that they cannot represent too many functions g at once, or the measure would be too small;
- then a counting argument gives a lower bound on maximal size of such subformulas.

The paper is organized in the following way. First we give the necessary definitions and some warm-up lemmas. Then we prove Theorem 1 and derive a lower bound on modified Andreev's function from it, assuming the existence of a set of functions with required combinatorial properties. Then we prove the existence of such set in Section 4.

While in our proof we rely on the fact that the top gate of a formula has a big fan-in (which, in a setting with fan-in 2, corresponds to having a top subtree of gates of the same type and big enough depth), this restriction seems somehow artificial. We believe that there is a possibility that this method could be adapted for the general case.

2 Preliminaries

2.1 Notation

Let us recall the definition of the XOR-composition.

Definition 4 ([MS21])**.** *XOR-composition* $f \boxplus g \colon \{0,1\}^{2n} \to \{0,1\}$ *of two function* $f \colon \{0,1\}^n \to \{0,1\}$ *and* $g \colon \{0,1\}^n \to \{0,1\}^n$ *is defined as follows:*

$$(f \boxplus g)(x,y) \coloneqq f(g(x) \oplus y)$$

Let us list some notation in regard to formula complexity.

Definition 5. *Let* $L(f)$ *be the minimum number of leaves in a formula* F *over basis* $\{\wedge, \vee, \neg\}$ *such that it computes* f*.*

Let $h(x,y)$ be a function of two variables. We denote as h^x a function: $h^x(y) \coloneqq h(x,y)$.

We also introduce the following shortcut notation for dealing with matrices.

Let $M = X \times Y$ be a matrix. We denote a submatrix $A \times Y$ for $A \subseteq X$ as M^A. An element of a matrix, located in row indexed by x, $x \in X$ and column indexed by y, $y \in Y$, is denoted as $M[x,y]$. Analogously, we denote a row indexed by x as $M[x]$.

For the rest of the paper, we consider only boolean matrices whose rows and columns are indexed by $X \coloneqq \{0,1\}^n$ and $Y \coloneqq \{0,1\}^n$.

Definition 6. *For a function* $h \colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$*, we define a matrix* M_h*:*

$$M_h[x,y] \coloneqq h(x,y)$$

Definition 7. *For a pair of functions* $f \colon \{0,1\}^n \to \{0,1\}$ *and* $g \colon \{0,1\}^n \to \{0,1\}^n$ *we define a matrix* $M_{f,g}$*:*

$$M_{f,g}[x,y] \coloneqq f(g(x) \oplus y)$$

Definition 8. *For a function* $f \colon \{0,1\}^n \to \{0,1\}$ *and a set of functions* Z *from* $\{0,1\}^n \to \{0,1\}^n$ *we define a matrix* $M_{f,Z}$*:*

$$M_{f,Z}[x,y] \coloneqq \bigvee_{g \in Z} f(g(x) \oplus y)$$

For the rest of the paper, $N \coloneqq 2^n$.

As we use f and g solely to denote first and second arguments of XOR-composition, we always imply the following domains and ranges for them: $f : \{0,1\}^n \to \{0,1\}$, $g : \{0,1\}^n \to \{0,1\}^n$. Analogously, x and y are implied to be vectors from $\{0,1\}^n$.

2.2 Warm-up lemmas

Lemma 1. *For a random f and arbitrarily fixed g and x we have: $L((f \boxplus g)^x) \geq N^{1-o(1)}$ with probability $1 - o(1)$.*

Proof. As we fix g and x, $f(g(x) \oplus y)$ depends only on y, so let $h(y) \coloneqq f(g(x) \oplus y)$.

Any formula for h can be transformed into a formula for f by adding negations to the variables y_i for all i where $g(x)_i = 1$. But, as a random function, f does not have a formula of size less than $N^{1-o(1)}$ with high probability [RS42]. □

The next lemma gives us connection between formula complexity of a function f and the size of $f^{-1}(0)$:

Lemma 2. *For $f\colon \{0,1\}^n \to \{0,1\}$, it holds that $L(f) \leq |f^{-1}(0)| \cdot n$.*

Proof. Let us compose a CNF formula for f. For any $x \in f^{-1}(0)$, we write a clause of n variables which becomes violated iff we substitute x to those variables. It is easy to check that a CNF formula composed exactly of $|f^{-1}(0)|$ such clauses represents function f. This formula has $|f^{-1}(0)| \cdot n$ leaves, which proves the inequality. □

Lemma 3. *Let $h(x, y)$ be a function of two variables. Then $L(h) \geq L(h^x)$ for any x.*

Proof. Taking a formula, computing h, we get a formula, computing h^x, by hardwiring the value of x into it. □

3 Proof of the main theorem

In this section, we prove Theorem 1 and, as a corollary, a lower bound on modified Andreev's function.

3.1 Modified Andreev's function

Definition 9. *Modified Andreev's function $Andr'$ takes $(3\log n + 1)n$ bits and outputs 1. We will treat it's inputs as:*

- *first $2n \log n$ bits are $2 \log n$ strings of size n;*
- *next $n \log n$ bits represent a description of a function from $\{0,1\}^{\log n}$ to $\{0,1\}^{\log n}$;*
- *last n bits represent a description of a function from $\{0,1\}^{\log n}$ to $\{0,1\}$.*

$$Andr'(x_1, \ldots, x_{2\log n}, g, f) = (f \boxplus g)\left(\bigoplus x_1, \ldots, \bigoplus x_{2\log n}\right)$$

where $\bigoplus z$ is parity of the vector z.

Theorem 2. *Modified Andreev's function is not computable by an AND of $n^{2\alpha-\varepsilon}$ formulas of size at most $n^{3-\alpha-\varepsilon}$ for any $0 < \alpha \leq \frac{1}{5} - 0.01$ and any $0 < \varepsilon < \alpha$.*

Again, AND in the statement could be replaced by OR.

As size of the formula is at most exponential in its depth, it immediately follows that:

Theorem 3. *Modified Andreev's function is not computable by an a $(3 + \alpha - 2\varepsilon)\log n$-depth formula with $(2\alpha - \varepsilon)\log n$ layers of AND gates at the top for any $0 < \alpha \leq \frac{1}{5} - 0.01$ and any constant $0 < \varepsilon < \alpha$.*

Note that lower bounds for different values of α are incomparable, since for increasing depth we have to pay with increasing number of same-type layers.

This beats current lower bounds on the formula depth of an explicit function, but it is conditioned on the form of the formula. To the best of our knowledge this is the first lower bound with restriction on the top of the formula.

Let us first show how a lower bound for a modified Andreev's function follows from Theorem 1. We will follow the classical proof of hardness for Andreev's function. We will take a look at how modified version behaves under random restriction R_p, where $p \coloneqq \frac{2 \ln \log n}{n}$ and show that it both shrinks well and remains hard with high probability. The only technical difficulty we need to overcome is that we need shrinkage to occur for many subformulas simultaneously. To do this we use concentration inequality on shrinkage proved in [IMZ12].

Let R_p be a distribution on partial assignments, such that for any variable x we independently assign:

- $x \coloneqq *$ with probability p;
- $x \coloneqq 1$ with probability $(1-p)/2$;
- $x \coloneqq 0$ with probability $(1-p)/2$;

Lemma 4 ([IMZ12])**.** *For any $p \geq \frac{1}{\sqrt{(L(f))}}$:*

$$\Pr\left[L(f \mid R_p)) \geq p^2 L(f)^{1+o(1)}\right] \leq \frac{1}{L(f)^{11}}$$

As in [IMZ12] this lemma is proved somewhat implicitly, we refer the reader to the Appendix A.1 for a more detailed explanation for which families of random restrictions their result works.

Proof of Theorem 2 from Theorem 1.

We can take any pair of functions f, g and hardwire it into $Andr'$. We pick those for which $f \boxplus g$ is not computable by AND of $n^{2\alpha-\delta}$ formulas of size at most $n^{1-\alpha-\delta}$. Such functions exist due to Theorem 1.

Let us take a look at how $Andr'$ with hardwired f and g behaves under random restriction R_p, where $p \coloneqq \frac{2 \ln \log n}{n}$:

$$\Pr\left[\text{all variables in a block are fixed by a } R_p\right] = (1-p)^n =$$

$$= \left(1 - \frac{2 \ln \log n}{n}\right)^n \leq e^{-2 \ln \log n} = (\log n)^{-2}$$

As there are $2 \log n$ input blocks, with probability $1 - o(1)$ we have at least one variable in each block that is not fixed. We pick exactly one such variable per block and fix other variables to arbitrary values. Now as there is exactly one variable in each block which is not fixed we end up with a function that is equal to $f \boxplus g$ up to possible negation of some variables. This means that with high probability $Andr'$ is not computable by AND of $n^{2\alpha-\delta}$ formulas of size at most $n^{1-\alpha-\delta}$ under random restriction R_p.

Now suppose that modified Andreev's function equals to $\bigwedge_{i=1}^{n^{2\alpha-\varepsilon}} a_i$, where each a_i is computable by $n^{3-\alpha-\varepsilon}$ size formula for some $\varepsilon > \delta > 0$.

We prove that under restriction R_p, all a_i shrink to a size less than $n^{1-\alpha-\delta}$.

For any i we have 3 different cases depending on $L(a_i)$:

- $L(a_i) < n^{1-\alpha-\varepsilon}$. In this case we are already done, as the formula size cannot increase under random restriction.

- $p \geq \frac{1}{\sqrt{L(a_i)}}$, $L(a_i) \geq n^{1-\alpha-\varepsilon}$. In this case, we apply Lemma 4. Since $L(a_i) \leq n^{3-\alpha-\varepsilon}$, $\Pr\left[L(a_i \mid R_p) \geq n^{1-\alpha-\varepsilon+o(1)}\right] \leq L(a_i)^{-11} \leq \frac{1}{n^5}$
- $p < \frac{1}{\sqrt{L(a_i)}}$, $L(a_i) \geq n^{1-\alpha-\varepsilon}$. In this case, we invoke monotonicity of $Pr[L(a_i \mid R_p) \geq k]$ on p with fixed k. Let $q \coloneqq \frac{1}{\sqrt{L(a_i)}}$, then $\Pr\left[L(a_i \mid R_p) \geq q^{2+o(1)}L(a_i)\right] \leq \Pr\left[L(a_i \mid R_q) \geq L(a_i)^{o(1)}\right] \leq L(a_i)^{-11} \leq \frac{1}{n^5}$.

Hence, with probability $1 - o(1)$ there is no i such that $L(a_i \mid R_p) \geq n^{1-\alpha-\delta} \geq n^{1-\alpha-\varepsilon+o(1)}$ for $\delta < \varepsilon$. Then with probability very close to 1 function $Andr'$ under random restriction R_p is computable by AND of $n^{2\alpha-\delta}$ formulas of size at most $n^{1-\alpha-\delta}$, which is a contradiction.

3.2 Proof of Theorem 1

Now we prove Theorem 1.

Theorem 1 (Main theorem)**.** *With probability* $1 - o(1)$ *for a random function* $f\colon \{0,1\}^n \to \{0,1\}$*, there exists a function* $g\colon \{0,1\}^n \to \{0,1\}^n$*, such that* $f \boxplus g$ *is not computable by an AND of* $2^{(2\alpha-\varepsilon)n}$ *formulas of size at most* $2^{(1-\alpha-\varepsilon)n}$ *for any* $0 < \alpha \leq \frac{1}{5} - 0.01$ *and any constant* $0 < \varepsilon < \alpha$*.*

To achieve that, we need to define a notion of *well-mixed set of functions.*

Definition 10 (Well-mixed set of functions)**.** *A set of functions* G *from* $\{0,1\}^n \to \{0,1\}^n$ *is* (Q, D, P)*-well-mixed for* f *if* $\forall Z \subset G, |Z| = Q$*, there exist a set* $K \subset \{0,1\}^n$*,* $|K| \leq P$*, such that* $M_{f,Z}^{X\setminus K}$ *has no more than* D *zeroes in total.*

We call the set K unlucky rows*, and we call the set* $X \setminus K$ lucky rows*.*

Informally, we say that a set is well-mixed if all of its subsets behave close to how a random subset of functions would behave.

The key assumption for proving Theorem 1 is that there exists large enough well-mixed set of functions G with suitable parameters. We leave proving its existence until next section, and for now let us prove Theorem 1 under this assumption.

Let us recall that N stands for 2^n, as this notation is used heavily below.

Proof. Let us randomly pick a function f. Then we take a set of functions G such that $|G| = N^{\frac{1}{4}N^{1-\alpha}}$ and G is $(N^\alpha, 2N^{2-2\alpha}, 2N^{1-\alpha})$-well-mixed. We aim to show that $f \boxplus g$ is hard for some $g \in G$.

We assume the contrary. Suppose for all $g \in G$ the XOR-composition $f \boxplus g$ can be represented as AND of small enough formulas. Formally, for any $g \in G$:

$$f \boxplus g = \bigwedge_{i=1}^{N^{2\alpha-\varepsilon}} h_{g,i}$$

where all $h_{g,i}$ are computable by formulas of size $N^{1-\alpha-\varepsilon}$. For any g we fix the smallest formula for $f \boxplus g$ of such form, and we denote it F_g.

We define a measure $C(h)$ of any function h of two arguments:

$$C(h) = \sum_{x \in X} L(h^x)$$

Let us note that $C(f \boxplus g) \geq N \cdot L(f) \geq N^{2-o(1)}$.

We prove that this measure is subadditive in the following sense:

Lemma 5. *Let $h(x,y) = \circ(g_1, g_2, \ldots, g_k)(x,y)$, where $\circ$ is either $\wedge$ or $\vee$. Then $C(h) \leq C(g_1) + \ldots + C(g_k)$.*

Proof. If we prove that the inequality holds for any specific x, namely, $L(h^x) \leq L(g_1^x) + \ldots + L(g_k^x)$, the lemma immediately follows.

Since $h(x,y) = \circ(g_1, g_2, \ldots, g_k)(x,y)$, we can construct a formula computing h^x, applying operation $\circ$ to formulas computing each g_i^x. The inequality follows. □

It means that for any $g \in G$ and $F_g = \bigwedge_{i=1}^{N^{2\alpha-\varepsilon}} h_{g,i}$ we can fix $1 \leq i_g \leq N^{2\alpha-\varepsilon}$ such that for h_{g,i_g} the measure is big enough, namely:

$$C(h_{g,i_g}) \geq N^{2-2\alpha+\varepsilon-o(1)}$$

We are going to use two different upper bounds on $L(h_{g,i_g}^x)$ for any x.

If for some h_{g,i_g} it is true that $L(h_{g,i_g}) \geq N^{1-\alpha}$, then we are done, as the corresponding subformula should be big enough. Otherwise, for any $x \in X$:

$$L(h_{g,i_g}^x) \leq L(h_{g,i_g}) \leq N^{1-\alpha}$$

.

On the other hand,

$$L(h_{g,i_g}^x) \leq |(h_{g,i_g}^x)^{-1}(0)| \cdot n$$

by Lemma 2.

Now let Z_h be a subset of G such that $\forall g \in Z_h : h_{g,i_g} = h$ for some fixed h.

Suppose $|Z_h| \geq N^\alpha$. We are going to come to a contradiction with this assumption. First, without losing generality, assume $|Z_h| = N^\alpha$, as we can take a subset of Z_h of exactly this size. Since G is $(N^\alpha, 2N^{2-2\alpha}, 2N^{1-\alpha})$-well-mixed, we can consider a matrix M_{f,Z_h} and distinguish set of unlucky rows K and set of lucky rows $X \setminus K$ in it.

From the properties of well-mixed set, we know that:

- $|K| \leq 2N^{1-\alpha}$;
- there is an upper bound on number of zeroes in $M_{f,Z_h}^{X\setminus K}$, which are exactly common zeroes of all $g \in Z_h$ on $X \setminus K$:
$$\sum_{x \notin K} \Big| \bigcap_{g \in Z_h} (g^x)^{-1}(0) \Big| \leq 2N^{2-2\alpha}$$

Let us also note that $M_h \geq M_{f,Z_h}$, since $(f \boxplus g)(x,y) = 1$ for any $g \in Z_h$ implies $h(x,y) = 1$.

$$\begin{aligned} C(h) = \sum_x L(h_x) = \sum_{x \in K} L(h_x) + \sum_{x \notin K} L(h_x) \leq \\ \leq \sum_{x \in K} L(h_x) + \sum_{x \notin K} |(h^x)^{-1}(0)| \cdot n \leq \\ \leq 2N^{1-\alpha} \cdot N^{1-\alpha} + \sum_{x \notin K} \Big| \bigcap_{g \in Z_h} (g^x)^{-1}(0) \Big| \cdot n \leq \\ \leq 2N^{2-2\alpha} + 2N^{2-2\alpha} \cdot n = N^{2-2\alpha+o(1)} \end{aligned}$$

since we have no more than $2N^{2-2\alpha}$ zeroes in all lucky rows in M_{f,Z_h}, therefore, the number of zeroes in rows $x \notin K$ in M_h does not exceed this as well.

We know that $C(h) \geq N^{2-2\alpha+\varepsilon-o(1)}$ though, so this is a contradiction.

So $|Z_h| < N^\alpha$. Then number of different h's is at least $\frac{|G|}{N^\alpha} \geq N^{\frac{1}{4}N^{1-\alpha}(1-o(1))}$.

Now let $s \coloneqq \max_h L(h)$. The number of functions with formulas of size at most s is at most $s^{s(1+o(1))}$, so $s^{s(1+o(1))} \geq N^{\frac{1}{4}N^{1-\alpha}(1-o(1))}$ and $s \log s(1+o(1)) \geq \frac{1}{4}N^{1-\alpha}\log N(1-o(1))$. Then $s \geq N^{1-\alpha-\varepsilon}$, and this completes the proof.

□

4 The existence of a well-mixed set of functions

Let us restate the definition of a well-mixed set:

Definition 10 (Well-mixed set of functions)**.** *A set of functions G from $\{0,1\}^n \to \{0,1\}^n$ is (Q, D, P)-well-mixed for f if $\forall Z \subset G, |Z| = Q$, there exist a set $K \subset \{0,1\}^n$, $|K| \leq P$, such that $M_{f,Z}^{X\setminus K}$ has no more than D zeroes in total.*

We call the set K unlucky rows*, and we call the set $X \setminus K$* lucky rows*.*

The goal of this section is to prove the following theorem:

Theorem 4. *For $0 < \alpha \leq \frac{1}{5} - 0.01$, let G be a family of functions $\{0,1\}^n \to \{0,1\}^n$ of size $N^{\frac{1}{4}N^{1-\alpha}}$, each sampled uniformly at random. Then G is $(N^\alpha, 2N^{2-2\alpha}, 2N^{1-\alpha})$-well-mixed set with probability at least $1 - \frac{1}{\exp(\exp(n))}$.*

Note that we sample G as a family of functions, with possible repetitions, which we use heavily in the proof. But with probability at least $\left(1 - \frac{1}{N^N - N^{\frac{1}{4}N^{1-\alpha}}}\right)^{N^{\frac{1}{4}N^{1-\alpha}}} \geq e^{-\frac{2}{N^{3N/4}}} \geq 1 - \frac{2}{N^{3N/4}}$ all functions turn out to be different from each other.

The plan of the proof is as follows:

- We identify which rows are supposed to be unlucky, and there will be two kinds of those: good and bad.
- For each kind of unlucky rows, we do three steps:
  - we prove that for a fixed row the probability to be of this kind for a random family of functions is very low;
  - we amplify this probability further, calculating the probability that some number of rows fail us simultaneously;
  - we sum the error over subsets of the random family.

The first kind of unlucky rows would be *a bad row.*

Definition 11. *Let Z be a family of functions $\{0,1\}^n \to \{0,1\}^n$, each function sampled uniformly at random. We call x a bad row for Z, if $\dim\left(span\left(\bigcup_{g\in Z}\{g(x)\}\right)\right) \leq 2\alpha n$, and a good row otherwise. Here the vector space is defined over $\mathcal{F}_2$ in a natural way.*

We bound the probability that a given x is bad for a random Z:

Lemma 6. *Let Z be a random family of functions $\{0,1\}^n \to \{0,1\}^n$ of size Q. Then for a fixed x:*

$$\Pr_Z\left[\dim\left(span\left(\bigcup_{g\in Z}\{g(x)\}\right)\right) \leq 2\alpha n\right] \leq N^{2\alpha n}\left(\frac{1}{N^{1-2\alpha}}\right)^Q$$

Proof. The proof is a simple counting argument. To have a vector subspace of dimension $2\alpha n$, we need to pick $2\alpha n$ generating vectors. There are $2^{2\alpha n} = N^{2\alpha}$ vectors in such space, and for each of Q functions from Z we pick its value in point x from those possibilities, versus N^Q possibilities in general case. So we have:

$$\begin{aligned}
&\Pr_Z\left[\dim\left(\operatorname{span}\left(\bigcup_{g\in Z}\{g(x)\}\right)\right) \le 2\alpha n\right] \le \\
&\le \binom{N}{2\alpha n}\left(N^{2\alpha}\right)^Q N^{-Q} \le \\
&\le N^{2\alpha n}\left(\frac{1}{N^{1-2\alpha}}\right)^Q
\end{aligned}$$

□

Now let us prove that with high probability, we have no more than $N^{1-\alpha}$ bad rows for a random Z.

Lemma 7. *Let Z be as in Lemma 6. Then*

$$\Pr_Z\left[\exists V \subset X, |V| = N^{1-\alpha} : \forall x \in V : \dim\left(span\left(\bigcup_{g\in Z}\{g(x)\}\right)\right) \le 2\alpha n\right] \le N^{-(1-2\alpha)QN^{1-\alpha}(1-o(1))}$$

if $Q = \omega(n)$.

Proof. Since each function in Z is sampled uniformly at random, for each x the event of "being bad" is independent of others. So for random Z the probability that a fixed set of $N^{1-\alpha}$ x's is bad can be bounded by $\left(N^{2\alpha n}\left(\frac{1}{N^{1-2\alpha}}\right)^Q\right)^{N^{1-\alpha}}$. After that we apply a union bound over all sets of x's.

We get:

$$\begin{aligned}
N^{N^{1-\alpha}} N^{2\alpha n N^{1-\alpha}}\left(\frac{1}{N^{1-2\alpha}}\right)^{QN^{1-\alpha}} &= \left(\frac{N^{2\alpha n+1}}{N^{(1-2\alpha)Q}}\right)^{N^{1-\alpha}} = \\
&= N^{-(1-2\alpha)Q(1-\frac{2\alpha n+1}{(1-2\alpha)Q})N^{1-\alpha}} \le \\
&\le N^{-(1-2\alpha)QN^{1-\alpha}(1-o(1))}
\end{aligned}$$

for $Q = \omega(n)$.

□

At last, we apply union bound over all choices of Z from G.

Lemma 8. *Let G be as in Theorem 4. Then:*

$$\Pr_G\left[\exists Z \subset G, |Z| = Q, \exists V \subset X, |V| = N^{1-\alpha} \colon \forall x \in V \colon x \text{ is bad for } Z\right] \le N^{-\frac{Q}{4}N^{1-\alpha}(1-o(1))}$$

Proof. For any specific $Z \subset G$ we can apply Lemma 7, as each function in G is sampled uniformly at random. We sum the error over all possible choices of Z and get:

$$|G|^Q N^{-(1-2\alpha)QN^{1-\alpha}(1-o(1))} = \left(\frac{N^{\frac{1}{4}N^{1-\alpha}}}{N^{(1-2\alpha)N^{1-\alpha}(1-o(1))}}\right)^Q \le N^{-\frac{Q}{4}N^{1-\alpha}(1-o(1))}$$

□

Hence, for every $Z \subset G$ there are no more than $N^{1-\alpha}$ bad rows with high probability.

We consider bad rows unlucky. After we take those out of consideration, we prove that almost all good rows are lucky.

To do that, first we consider a technical Lemma:

Lemma 9. *Let f be a function $\{0,1\}^n \to \{0,1\}$, chosen uniformly randomly, S be any set of functions from $\{0,1\}^n \to \{0,1\}^n$ such that for a fixed x values $g(x)$ for all $g \in S$ are linearly independent, $2 \le |S| < \frac{n}{2}$. Then:*

$$\Pr_f \left[M_{f,S}[x] \text{ has } 2N2^{-|S|} \text{ zeroes} \right] \le 2^{(n+1)|S| - 2^{n-2|S|-2}}$$

This is a corollary from the following statement by Kaave Hosseini (from personal communication):

Lemma 10. *Let $B \subset \{0,1\}^n$ be a random set, where each point lies with probability $\frac{1}{2}$, $L = x_1, \ldots, x_k$ be linearly independent vectors, $k \ge 2$. Then:*

$$\Pr_B \left[|(B + x_1) \cap \ldots \cap (B + x_k)| \ge 2^{n-k+1} \right] \le 2^{(n+1)k - 2^{n-2k-2}}$$

First we show how Lemma 9 follows from Lemma 10.

Proof of 9 from 10: The value of $M_{f,S}[x, y]$ is zero iff $\forall g \in S : f(g(x) \oplus y) = 0$. Plugging $B \coloneqq f^{-1}(0)$, $k \coloneqq |S|$, $L \coloneqq \{g(x)\}_{g \in S}$ in 10, we get 9, since zeroes of $f(g(x) \oplus y)$ are shifted by vector $g(x)$ in comparison with zeroes of f.

Now we prove Lemma 10.

Proof. To upper bound $|(B + x_1) \cap \ldots \cap (B + x_k)|$, let us consider the following sum:

$$A \coloneqq \sum_{y \in \{0,1\}^n} \chi(y + x_1)\chi(y + x_2) \ldots \chi(y + x_k)$$

where χ is a characteristic function of a set B. This sum equals $|(B + x_1) \cap \ldots \cap (B + x_k)|$ exactly.

Let us now extend set $L = \{x_1, \ldots, x_k\}$ to a basis and split $\{0,1\}^n$ onto 2^k parts, depending on whether a vector contains x_i in its decomposition onto basis vectors or not. Each part contains 2^{n-k} different points. Let us consider one of those parts, without losing generality we pick:

$$P = \{y \mid y \text{ does not contain any of } x_i \text{ in decomposition}\}$$

In the sum $A_P \coloneqq \sum\limits_{y \in P} \chi(y + x_1) \ldots \chi(y + x_k)$ every point in $\{0,1\}^n$ occurs as an argument of χ no more than once, so all summands and all multipliers in them are independent.

If we interpret every summand as a Bernoulli variable which equals 1 with probability $\frac{1}{2^k}$, we can apply Chernoff bounds.

The exact form that we use here is the following:

Theorem 5 (Chernoff bounds, multiplicative form)**.** *For independent random $X_1, \ldots, X_n$ from $\{0,1\}$ and $0 \le \delta \le 1$:*

$$\Pr\left[\sum_i X_i \ge (1 + \delta)\mu \right] \le e^{-\frac{\delta^2 \mu}{3}}$$

where μ is the expected value of $\sum_i X_i$.

Here the expected value of A_P equals 2^{n-2k}.

$$\Pr_B\left[A_P > (1+\delta)2^{n-2k}\right] < e^{(-\delta^2)2^{n-2k-2}}$$

This holds regardless of the choice of a part P. We apply union bound and sum the error over all possible parts:

$$\Pr_B\left[A > (1+\delta)2^{n-k}\right] \le \Pr_B\left[\exists P\colon A_P > (1+\delta)2^{n-2k}\right] < 2^k \cdot e^{(-\delta^2)2^{n-2k-2}}$$

At last, we apply union bound over all possible choices of L:

$$\Pr_B\left[\exists L\colon A > (1+\delta)2^{n-k}\right] < 2^{nk} \cdot 2^k \cdot e^{(-\delta^2)2^{n-2k-2}}$$

We pick $\delta \coloneqq 1$ and get:

$$\Pr_B\left[A > 2^{n-k+1}\right] < 2^{(n+1)k-2^{n-2k-2}}$$

This finishes the proof of the Lemma.

□

We use the Lemma to bound the number of zeroes in good rows.

Lemma 11. *Let x be a good row for a family of functions Z. Then:*

$$\Pr_f\left[M_{f,Z}[x] \text{ has } 2N^{1-2\alpha} \text{ zeroes}\right] \le 2^{-N^{\alpha+0.05}(1-o(1))}$$

Proof. Since x is good for Z, we can find a subset $S \subset Z$ such that $|S| = 2\alpha n$ and $\{g(x)\}_{g\in S}$ are linearly independent. By Lemma 9 and the fact that $n - 2 \cdot 2\alpha n \ge (\alpha + 0.05)n$, we immediately have an upper bound on number of zeroes in a row x:

$$\begin{aligned}
&\Pr_f\left[M_{f,Z}[x] \text{ has } 2N^{1-2\alpha} \text{ zeroes}\right] \le \\
&\le \Pr_f\left[M_{f,S}[x] \text{ has } 2N^{1-2\alpha} \text{ zeroes}\right] \le \\
&\le 2^{(n+1)2\alpha n - 2^{(\alpha+0.05)n-2}} = \\
&= 2^{-N^{\alpha+0.05}(1-o(1))}
\end{aligned}$$

□

Lemma 12. *Let Z be a random family of functions from $\{0,1\}^n \to \{0,1\}^n$. Then:*

$$\Pr_{f,Z}\Big[\exists V,V \text{ is a set of good rows for } Z, |V| = N^{1-\alpha}\colon \forall x \in V\colon M_{f,Z}[x] \text{ has } 2N^{1-2\alpha} \text{ zeroes}\Big] \le$$
$$\le 2^{-N^{1.05}(1-o(1))}$$

Proof. Again, as Z is picked uniformly at random, the events regarding every x are independent for a fixed V. Now let us sum the error over all choices of V:

$$\Pr_{f,Z}\left[\exists V, V \text{ is a set of good rows for } Z, |V| = N^{1-\alpha} \colon \forall x \in V \colon M_{f,Z}[x] \text{ has } 2N^{1-2\alpha} \text{ zeroes}\right] \leq$$
$$\leq \binom{N}{N^{1-\alpha}} \left(2^{-N^{\alpha+0.05}(1-o(1))}\right)^{N^{1-\alpha}} \leq$$
$$\leq \binom{N}{N^{1-\alpha}} 2^{-N^{1.05}(1-o(1))} \leq$$
$$\leq 2^{N^{1-\alpha}\cdot \log N} \cdot 2^{-(1-o(1))N^{1.05}} \leq$$
$$\leq 2^{-(1-o(1))N^{1.05}}$$

□

Now, we are ready to sum over all possible choices of Z in a random G.

Lemma 13. *Let G be as in Theorem 4. Then:*

$$\Pr_{f}\Big[\exists Z \subset G, |Z| = Q, \exists V, V \text{ is a set of good rows for } Z, |V| = N^{1-\alpha} \colon$$
$$\forall x \in V \colon M_{f,Z}[x] \text{ has } 2N^{1-2\alpha} \text{ zeroes}\Big] \leq N^{Q\frac{1}{4}N^{1-\alpha}} 2^{-(1-o(1))N^{1.05}}$$

Proof. This is a union bound over all choices of a subset Z of size Q from set G of size $N^{\frac{1}{4}N^{1-\alpha}}$. □

Now we are ready to prove Theorem 4.

Let us take a family G of functions from $\{0,1\}^n \to \{0,1\}^n$ of size $N^{\frac{1}{4}N^{1-\alpha}}$, where each function is sampled uniformly at random. With probability $1 - \frac{1}{\exp(\exp(n))}$ all functions turn out to be different from each other.

Let us take any $Z \subset G$ of size N^α. We plug $Q \coloneqq N^\alpha$ into Lemma 8 and get that with probability at least $1 - N^{-N^{1-\alpha}}$ we have no more than $N^{1-\alpha}$ bad rows for our Z.

Plugging this parameter in Lemma 13, we get that with probability $1 - N^{\frac{1}{4}N} 2^{-N^{1.05}(1-o(1))} = 1 - 2^{-N^{1.05}(1-o(1))}$ no more than $N^{1-\alpha}$ good rows have $2N^{1-2\alpha}$ zeroes in them.

So, with very high probability, all properties hold. We say that both bad rows and good rows with at least $2N^{1-2\alpha}$ zeroes are our unlucky rows, and this is a set K from Theorem 4.

Now, as every lucky row has at most $2N^{1-2\alpha}$ zeroes, and there are no more than N lucky rows, there are no more than $2N^{2-2\alpha}$ zeroes in all lucky rows, which gives us the statement of Theorem 4.

5 Notes and open questions

Note that our result works for α that is separated from $\frac{1}{5}$ by some constant. With more accurate analysis in Section 4 it might be possible to push α up. So the first natural question is: can we prove variant of the main theorem for $\alpha < \frac{1}{2}$?

The second question concerns the balance of the parameters. In our current result, one can say that we are trading one arbitrary layer for two AND layers. Can this trade-off be made more favorable?

But the main question that arises from our work is whether this method could be adapted to work without restrictions on top gates of the formula, or with weaker restrictions. The first natural extension would be to to prove a lower bound for AC_0 formula on top of arbitrary De Morgan formulas.

Conjecture 3. *For any positive integer d, there exists $\alpha > 0$ and $c > 1$ such that modified Andreev's function is not computable by an AC_0 formula of depth d and size $n^{c\alpha}$ on top of $(3-\alpha)\log n$-depth De Morgan formulas.*

6 Acknowledgements

The authors would like to thank Kaave Hosseini for his invaluable insights on Lemma 10, Alexander Kulikov, Alexander Smal and Artur Riazanov for fruitful discussions and comments on the draft.

7 Another important statement

The contents of this section had to be modified for the reasons of safety of the authors, compared to the original version, as the recently adopted Russian laws effectively establish censorship.

Nevertheless, we are deeply upset about the events currently happening in Ukraine and wish for peace more than anything.

References

[And87] Alexander E. Andreev. On a method for obtaining more than quadratic effective lower bounds for the complexity of π-schemes. *Moscow University Mathematics Bulletin*, 42(1):24–29, 1987.

[DM16] Irit Dinur and Or Meir. Toward the KRW composition conjecture: Cubic formula lower bounds via communication complexity. In Ran Raz, editor, *31st Conference on Computational Complexity, CCC 2016, May 29 to June 1, 2016, Tokyo, Japan*, volume 50 of *LIPIcs*, pages 3:1–3:51. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016.

[dRMN+20] Susanna F. de Rezende, Or Meir, Jakob Nordström, Toniann Pitassi, and Robert Robere. KRW composition theorems via lifting. In Sandy Irani, editor, *61st IEEE Annual Symposium on Foundations of Computer Science, FOCS 2020, Durham, NC, USA, November 16-19, 2020*, pages 43–49. IEEE, 2020.

[EIRS01] Jeff Edmonds, Russell Impagliazzo, Steven Rudich, and Jirí Sgall. Communication complexity towards lower bounds on circuit depth. *Comput. Complex.*, 10(3):210–246, 2001.

[GMWW17] Dmitry Gavinsky, Or Meir, Omri Weinstein, and Avi Wigderson. Toward better formula lower bounds: The composition of a function and a universal relation. *SIAM J. Comput.*, 46(1):114–131, 2017.

[Hås98] Johan Håstad. The shrinkage exponent of de morgan formulas is 2. *SIAM J. Comput.*, 27(1):48–64, 1998.

[HW90] Johan Håstad and Avi Wigderson. Composition of the universal relation. In *Advances In Computational Complexity Theory, Proceedings of a DIMACS Workshop, New Jersey, USA, December 3-7, 1990*, pages 119–134, 1990.

[IMZ12] Russell Impagliazzo, Raghu Meka, and David Zuckerman. Pseudorandomness from shrinkage. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, pages 111–119. IEEE, 2012.

[IN93] Russell Impagliazzo and Noam Nisan. The effect of random restrictions on formula size. *Random Struct. Algorithms*, 4:121–134, 1993.

[Khr71] Valeriy Mihailovich Khrapchenko. Complexity of the realization of a linear function in the class of II-circuits. *Mathematical Notes of the Academy of Sciences of the USSR*, 9(1):21–23, 1971.

[KM18] Sajin Koroth and Or Meir. Improved composition theorems for functions and relations. In Eric Blais, Klaus Jansen, José D. P. Rolim, and David Steurer, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2018, August 20-22, 2018 - Princeton, NJ, USA*, volume 116 of *LIPIcs*, pages 48:1–48:18. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018.

[KRW95] Mauricio Karchmer, Ran Raz, and Avi Wigderson. Super-logarithmic depth lower bounds via the direct sum in communication complexity. *Computational Complexity*, 5(3/4):191–204, 1995.

[KW88] Mauricio Karchmer and Avi Wigderson. Monotone circuits for connectivity require super-logarithmic depth. In *Proceedings of the 20th Annual ACM Symposium on Theory of Computing, May 2-4, 1988, Chicago, Illinois, USA*, pages 539–550, 1988.

[MS21] Ivan Mihajlin and Alexander Smal. Toward better depth lower bounds: The XOR-KRW conjecture. In Valentine Kabanets, editor, *36th Computational Complexity Conference, CCC 2021, July 20-23, 2021, Toronto, Ontario, Canada (Virtual Conference)*, volume 200 of *LIPIcs*, pages 38:1–38:24. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021.

[PZ91] Michael S. Paterson and Uri Zwick. Shrinkage of de morgan formulae under restriction. In *Proceedings of the 32nd Annual Symposium on Foundations of Computer Science*, SFCS '91, page 324–333, USA, 1991. IEEE Computer Society.

[RS42] J. Riordan and C. Shannon. The number of two-terminal series-parallel networks. *J Math Phys*, 21, 01 1942.

[Sub61] Bella Abramovna Subbotovskaya. Realization of linear functions by formulas using $\wedge$, $\vee$, $\neg$. In *Doklady Akademii Nauk*, volume 136-3, pages 553–555. Russian Academy of Sciences, 1961.

[Tal14] Avishay Tal. Shrinkage of de morgan formulae by spectral techniques. *Proceedings - Annual IEEE Symposium on Foundations of Computer Science, FOCS*, pages 551–560, 12 2014.

A Appendix

A.1 Notes on Shrinkage Lemma

As the formulation of Lemma 4 is proven in [IMZ12] somewhat implicitly, we give some notes on that. The following notation is consistent with [IMZ12].

Let *a p-regular distribution* on partial assignments to variables $x_1, \ldots, x_n$ be the one for which for any variable x_i: $Pr[x_i = *] = p$.

For random restriction ρ, let $\mathrm{supp}(\rho) \coloneqq \{x_i \mid \rho(x_i) \neq *\}$.

[IMZ12] use a notion of an independent sequence of restrictions. While it is defined naturally on families of restrictions to which shrinkage is applicable, defining this in general can be quite technical. Below we explain the usage of this notion in [IMZ12].

The authors construct a random restriction ρ as a sequence of r k-wise independent restrictions $\rho_1, \ldots, \rho_r$. First they sample ρ_1, and then ρ_2 is sampled independently on those variables x for which $\rho_1(x) = *$ and so on. Note that supports of such restrictions are not independent of each other, so formalizing this in general would require some accuracy with probability space. Nevertheless, this construction is truly independent for a sequence of uniform distributions with the same parameter.

Let us now present a general statement, which follows [IMZ12] almost to a letter.

Lemma 14 (Lemma 4.8 in [IMZ12])**.** *Let $\Gamma \coloneqq 2$. For a constant $c \geq 11$, $p \geq m^{-1/\Gamma}$, $r \geq 11$, a formula f on $\leq m$ variables with $L(f) = m$ and any p-regular random restriction ρ which is a sequence of r independent $(q = p^{\frac{1}{r}})$-regular k-wise independent restrictions:*

$$\Pr\left[L(f \mid \rho) \geq 2^{3c \log^{2/3} n} p^{\Gamma} m\right] \leq m^{-c}$$

Note that in [IMZ12] authors argue the existence of such distribution and take great care of minimizing the number of random bits needed to generate it. Nevertheless, their proof works for any distribution with mentioned properties and we use it for uniform distribution.

The uniform distribution R_p can be broken up to a sequence of r distributions R_q, where $q = p^{\frac{1}{r}}$. All of them are k-wise independent for any k, so, plugging in the parameters along with $c = 11$, we get Lemma 4.

In terms of uniform distribution this statement was also mentioned in [Tal14].

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il