THEORY OF COMPUTING, Volume 12 (11), 2016, pp. 1–17
www.theoryofcomputing.org

# Anti-concentration for Polynomials of Independent Random Variables

Raghu Meka* Oanh Nguyen Van Vu†

Received August 11, 2015; Revised June 16, 2016; Published August 25, 2016

**Abstract:** We prove anti-concentration results for polynomials of independent random variables with arbitrary degree. Our results extend the classical Littlewood-Offord result for linear polynomials, and improve several earlier estimates.

We discuss applications in two different areas. In complexity theory, we prove near-optimal lower bounds for computing the PARITY function, addressing a challenge in complexity theory posed by Razborov and Viola, and also address a problem concerning the OR function. In random graph theory, we derive a general anti-concentration result on the number of copies of a fixed graph in a random graph.

**ACM Classification:** F.1.1, G.2.2

**AMS Classification:** 68Q05, 68R10

**Key words and phrases:** complexity theory, random polynomials, anti-concentration, parity, random graphs

## 1 Introduction

Let $\xi$ be a Rademacher random variable (taking value $\pm 1$ with probability $1/2$) and $A = \{a_1, \dots, a_n\}$ be a multi-set in $\mathbb{R}$ (here $n \to \infty$). Consider the random sum

$$S := a_1\xi_1 + \cdots + a_n\xi_n$$

where $\xi_i$ are i. i. d. copies of $\xi$.

*Supported by NSF Career Award CCF-1553605.
†Supported by NSF grant DMS-1307797 and AFORS grant FA9550-12-1-0083.

© 2016 Raghu Meka, Oanh Nguyen, and Van Vu
Licensed under a Creative Commons Attribution License (CC-BY) DOI: 10.4086/toc.2016.v012a011

In 1943, Littlewood and Offord, in connection with their studies of random polynomials [16], raised the problem of estimating $\mathbf{P}(S \in I)$ for *arbitrary* coefficients $a_i$. They proved the following remarkable theorem:

**Theorem 1.1.** *There is a constant $B$ such that the following holds for all $n$. If all coefficients $a_i$ have absolute value at least 1, then for any open interval $I$ of length 1,*

$$\mathbf{P}(S \in I) \leq Bn^{-1/2} \log n\,.$$

Shortly after the Littlewood-Offord result, Erdős [10] removed the $\log n$ term to obtain the optimal bound using an elegant combinatorial proof. Littlewood-Offord type results are commonly referred to as anti-concentration (or small-ball) inequalities. Anti-concentration results have been developed by many researchers through decades, and have recently found important applications in the theories of random matrices and random polynomials; see, for instance, [19] for a survey.

The goal of this paper is to extend Theorem 1.1 to higher degree polynomials. Consider

$$P(x_1, \ldots, x_n) := \sum_{S \subset \{1,\ldots,n\}; |S| \leq d} a_S \prod_{j \in S} x_j\,. \tag{1.1}$$

The first result in this direction, due to Costello, Tao, and the third author, [8], is

**Theorem 1.2.** *There is a constant $B$ such that the following holds for all $d, n$. If there are $mn^{d-1}$ coefficients $a_S$ of absolute value at least 1, then for any open interval $I$ of length 1,*

$$\mathbf{P}\big(P(\xi_1, \ldots, \xi_n) \in I\big) \leq Bm^{-2^{-(d^2+d)/2}}\,.$$

The exponent $2^{-(d^2+d)/2}$ tends very fast to zero with $d$, and it is desirable to improve this bound. For the case $d = 2$, Costello [7] obtained the optimal bound $n^{-1/2+o(1)}$. In a more recent paper [21], Razborov and Viola proved

**Theorem 1.3.** *There is a constant $B$ such that the following holds for all $d, n$. If there are pairwise disjoint subsets $S_1, \ldots, S_r$ each of size $d$ such that the $a_{S_i}$ have absolute value at least 1 for all $i$, then for any open interval $I$ of length 1,*

$$\mathbf{P}\big(P(\xi_1, \ldots, \xi_n) \in I\big) \leq Br^{-g(d)}\,, \qquad \textit{where} \qquad g(d) = \frac{1}{d2^{d+1}}\,.$$

This theorem improves the bound in Theorem 1.2 to $m^{-g(d)}$ via a simple counting argument.

Researchers in analysis have also considered anti-concentration of polynomials, for entirely different reasons. Carbery and Wright [6] consider polynomials with $\xi_i$ being i. i. d. Gaussian and prove the following bound.

**Theorem 1.4** ([6, Theorem 8])**.** *There is a constant $B$ such that*

$$\mathbf{P}\big(|P(\xi_1, \ldots, \xi_n)| \leq \varepsilon\, \mathbf{Var}(P(\xi_1, \ldots, \xi_n))^{1/2}\big) \leq Bd\varepsilon^{1/d}\,.$$

See [18] for related results. The above result has been extended by Mossel, O'Donnell and Oleszkiewicz [17] to general variables, at a cost of an extra term on the right hand side depending on how spread out the coefficients of $P$ are.

The goal of this paper is to further improve these anti-concentration bounds, with several applications in complexity theory. Our new results will be nearly optimal in a wide range of parameters. Let $[n] = \{1,2,\ldots,n\}$. Following [21], we define the rank of polynomials as follows.

**Definition 1.5.** For a degree-$d$ multilinear polynomial[1] of the form (1.1), the *rank* of $P$, denoted by $\operatorname{rank}(P)$, is the largest integer $r$ such that there exist disjoint sets $S_1,\ldots,S_r \subseteq [n]$ of size $d$ with $|a_{S_j}| \geq 1$, for $j \in [r]$.

Our first main result concerns the Rademacher case. Let $\xi_i, i = 1,\ldots,n$ be i. i. d. Rademacher random variables.

**Theorem 1.6.** *There is an absolute constant $B$ such that the following holds for all $d,n$. For any polynomial $P$ of the form* (1.1) *with rank $r \geq 2$ and for any interval $I$ of length 1, we have*

$$\mathbf{P}\big(P(\xi_1,\ldots,\xi_n) \in I\big) \leq \min\left(\frac{Bd^{4/3}\sqrt{\log r}}{r^{\frac{1}{4d+1}}}\,,\ \frac{\exp(Bd\log d\log\log r + Bd^2\log d)}{\sqrt{r}}\right).$$

For the case when $d$ is fixed, it has been conjectured [19] that $\mathbf{P}(P(\xi_1,\ldots,\xi_n) \in I) = O(r^{-1/2})$. This conjectural bound is a natural generalization of the Erdős-Littlewood-Offord result and is optimal, as shown by taking $P = (\xi_1 + \cdots + \xi_n)^d$, with $n$ even. For this $P$, the rank $r = \Theta(n)$ and $\mathbf{P}(|P| \leq 1/2) = \mathbf{P}(P = 0) = \Theta(n^{-1/2})$. Our result confirms this conjecture up to some polylog term $(\log r)^{Bd\log d}$.

In applications it is important that we can allow the degree $d$ to tend to infinity with $n$. Our bounds in Theorem 1.6 are non-trivial for degrees up to $c\log r/\log\log r$, for some positive constant $c$. Up to the $\log\log$ term, this is as good as it gets, as we cannot hope to get any non-trivial bound for polynomials of degree $\log_2 r$. For example, the degree-$d$ polynomial on $2^d \cdot d$ variables defined by

$$P(\xi) = \sum_{i=1}^{2^d}\prod_{j=1}^{d}(\xi_{ij}+1)\,,$$

where $\xi_{ij}$ are i. i. d. Rademacher random variables, has $r = 2^d$ and $\mathbf{P}(P(\xi) = 0) = \Omega(1)$.

Next, we generalize our result to non-Rademacher distributions. As a first step, we consider the $p$-biased distribution on the hypercube. For $p \in (0,1)$, let $\mu_p$ denote the Bernoulli variable with $p$-biased distribution:

$$\mathop{\mathbf{P}}_{\xi\sim\mu_p}(\xi = 0) = 1-p\,,$$
$$\mathop{\mathbf{P}}_{\xi\sim\mu_p}(\xi = 1) = p$$

and let $\mu_p^n$ be the product distribution on $\{0,1\}^n$.

[1] A polynomial is multilinear if the degree of each individual variable is at most 1 in the polynomial.

**Theorem 1.7.** *There is an absolute constant $B$ such that the following holds. Let $P$ be a polynomial of the form* (1.1) *whose rank is $r \geq 2$. Let $p$ be such that $\tilde{r} := 2^d \alpha^d r \geq 3$ where $\alpha := \min\{p, 1-p\}$. Then for any interval $I$ of length 1,*

$$\mathbf{P}_{\xi \sim \mu_p^n}(P(\xi) \in I) \leq \min\left(\frac{Bd^{4/3}(\log \tilde{r})^{1/2}}{(\tilde{r})^{1/(4d+1)}}, \frac{\exp(Bd\log d \log\log \tilde{r} + Bd^2 \log d)}{\sqrt{\tilde{r}}}\right).$$

The distribution $\mu_p^n$ plays an essential role in probabilistic combinatorics. For example, it is the ground distribution for the random graphs $G(N,p)$ (with $n := \binom{N}{2}$). We discuss an application in the theory of random graphs in the next section.

Finally, we present a result that applies to virtually all sets of independent random variables, with a weak requirement that these variables do not concentrate on a short interval.

**Theorem 1.8.** *There is an absolute constant $B$ such that the following holds. Let $\xi_1, \dots, \xi_n$ be independent (but not necessarily i. i. d.) random variables. Let $P$ be a polynomial of the form* (1.1) *whose rank is $r \geq 2$. Assume that there are positive numbers $p$ and $\varepsilon$ such that for each $1 \leq i \leq n$, there is a number $y_i$ such that $\min\{\mathbf{P}(\xi_i \leq y_i), \mathbf{P}(\xi_i > y_i)\} = p$ and $\mathbf{P}(|\xi_i - y_i| \geq 1) \geq \varepsilon$. Assume furthermore that $\tilde{r} := (p\varepsilon)^d r \geq 3$. Then for any interval $I$ of length 1*

$$\mathbf{P}\big(P(\xi_1, \dots, \xi_n) \in I\big) \leq \min\left(\frac{Bd^{4/3}(\log \tilde{r})^{1/2}}{(\tilde{r})^{1/(4d+1)}}, \frac{\exp(Bd\log d \log\log \tilde{r} + Bd^2 \log d)}{\sqrt{\tilde{r}}}\right).$$

Notice that even in the Gaussian case, Theorem 1.8 is incomparable to Theorem 1.4. If we use Theorem 1.4 to bound $\mathbf{P}(P \in I)$ for an interval $I$ of length 1, then we need to set $\varepsilon = \mathbf{Var}(P)^{-1/2}$, and the resulting bound becomes $B/(\mathbf{Var}\, P)^{1/2d}$. For sparse polynomials, it is typical that $r$ is much larger than $(\mathbf{Var}\, P)^{1/d}$ and in this case our bound is superior. To illustrate this point, let us fix a constant $d > c > 0$ and consider

$$P := \sum_{S \subset \{1,\dots,n\}, |S|=d} a_S \prod_{i \in S} x_i$$

where $a_S$ are i. i. d. random Bernoulli variables with $\mathbf{P}(a_S = 1) = n^{-c}$. It is easy to show that the following hold with probability $1 - o(1)$.

- For any set $X \subset \{1, \dots, n\}$ of size at least $n/2$, there is a subset $S \subset X, |S| = d$, such that $a_S = 1$.
- The number of nonzero coefficients is at most $n^{d-c}$.

In other words, these two conditions are typical for a sparse polynomial with roughly $n^{d-c}$ nonzero coefficients. On the other hand, if the above two conditions holds, then we have $\mathbf{Var}(P) \leq n^{d-c}$ and $r \geq n/2d$ (by a trivial greedy algorithm). Our bound implies that

$$\mathbf{P}(P \in I) \leq C(d) n^{-1/2+o(1)}$$

while Carbery-Wright bound only gives

$$\mathbf{P}(P \in I) \leq C(d) n^{-1/2+c/2d}.$$

The rest of the paper is organized as follows. In Section 2, we discuss applications in complexity theory and graph theory, with one long proof delayed to Section 5. The proof of Theorem 1.6 is contained in Section 3. The generalizations are discussed in Section 4.

All asymptotic notations are used under the assumption that $n$ tends to infinity. All the constants are absolute, unless otherwise noted. Throughout the paper, log denotes the natural logarithm.

## 2 Applications

### 2.1 Applications in complexity theory

We use our anti-concentration results to prove lower bounds for approximating Boolean functions by polynomials in the *Hamming metric*. The notion of approximation we consider is as follows.

**Definition 2.1.** Let $\varepsilon > 0$ and $\mu$ be a distribution on $\{0,1\}^n$. For a Boolean function $f : \{0,1\}^n \to \{0,1\}$ and a polynomial $P : \mathbb{R}^n \to \mathbb{R}$, we say $P$ $\varepsilon$-approximates $f$ with respect to[2] $\mu$ if

$$\mathop{\mathbf{P}}_{x\sim\mu}\big(P(x) = f(x)\big) > 1-\varepsilon.$$

We define $d_{\mu,\varepsilon}(f)$ to be the least $d$ such that there is a degree-$d$ polynomial which $\varepsilon$-approximates $f$ with respect to $\mu$.

An alternate (dual) way to view the above notion is in terms of distributions over low-degree polynomials—"randomized polynomials"—which approximate the function in the worst-case. In particular, by Yao's min-max principle, $d_{\mu,\varepsilon}(f) \le d$ for every distribution $\mu$ if and only if there exists a distribution $\mathcal{D}$ over polynomials of degree at most $d$ which approximates $f$ in the worst case: for all $x$, $\mathbf{P}_{P\sim\mathcal{D}}[P(x) = f(x)] > 1-\varepsilon$.

Approximating Boolean functions by polynomials in the Hamming metric was first considered in the works of Razborov [20] and Smolensky [22] over fields of finite characteristic as a technique for proving lower bounds for small-depth circuits. This was also studied in a similar context over real numbers by Beigel, Reingold, and Spielman [4] and Aspnes, Beigel, Furst, and Rudich [3]. The latter paper uses them to prove lower bounds for $\mathsf{AC}^0$. More recently, in a remarkable result, Williams [24] (see also [25, 1]) used polynomial approximations in Hamming metric to obtain the best known algorithms for all-pairs shortest path and other related algorithmic questions. Here, we study lower bounds for the existence of such approximations.

**Approximating Parity.** Let $\mathrm{Par}_n : \{0,1\}^n \to \{0,1\}$ denote the parity function:

$$\mathrm{Par}_n(x) = x_1 \oplus x_2 \oplus \cdots \oplus x_n$$

(where arithmetic is mod 2).

[2]We drop $\mu$ in the description when it is clear from context or if it is the uniform distribution.

In [21], Razborov and Viola introduced another way to look at the problem of approximating parity by low-degree polynomials in the Hamming metric. For two functions $f,g:\{0,1\}^n \to \mathbb{R}$, define their "correlation" to be the quantity

$$\mathrm{Cor}_n(f,g) = \mathop{\mathbf{P}}_x\big(f(x)=g(x)\big) - 1/2\,,$$

where $x$ is uniformly distributed over $\{0,1\}^n$. They highlighted the following challenge.

**Challenge 2.2.** Exhibit an explicit Boolean function $f:\{0,1\}^n \to \{0,1\}$ such that for any real polynomial $P$ of degree at most $\log_2 n$, we have

$$\mathrm{Cor}_n(f,P) \le o(1/\sqrt{n})\,.$$

This challenge is motivated by studies in complexity theory and has connections to many other problems, such as the famous rigidity problem; see [21] for more discussion.

The Parity function seems to be a natural candidate in problems like this. Razborov and Viola, using Theorem 1.3, proved

**Theorem 2.3** ([21])**.** *For all sufficiently large $n$, $\mathrm{Cor}_n(\mathrm{Par}_n,P) \le 0$ for any real polynomial $P$ of degree at most $(1/2)\log_2\log_2 n$.*

With Theorem 1.6, we obtain the following improvement, which gets us to within a $\log\log n$ factor of Challenge 2.2.

**Theorem 2.4.** *For all sufficiently large $n$, $\mathrm{Cor}_n(\mathrm{Par}_n,P) \le 0$ for any real polynomial $P$ of degree at most*

$$\frac{\log n}{15\log\log n}\,.$$

*Proof.* Let $d$ be the degree of $P$. Following the arguments in the proof of [21, Theorem 1.1], we can assume that $P$ contains at least $\sqrt{n}$ pairwise disjoint subsets $S_i$ each of size $d$ with non-zero coefficients. It suffices to show that the probability that $P$ outputs a Boolean value is at most $1/2$. By replacing $P$ by $q(x_1,\dots,x_n) := P((x_1+1)/2,\dots,(x_n+1)/2)$, we can convert $P$ into a polynomial of the same degree defined on $\{\pm 1\}^n$, in other words, on Rademacher variables. Then by Theorem 1.6, this probability is bounded by

$$2B\frac{d^{4/3}\log^{1/2} n}{n^{1/(8d+2)}}\,.$$

This is less than $1/2$ for every

$$d \le \frac{\log n}{15\log\log n}$$

when $n$ is sufficiently large. □

**Approximating** AND/OR. One of the main building blocks in obtaining polynomial approximations in the Hamming metric is the following result for approximating the OR function.[3]

**Claim 2.5.** *For all $\varepsilon \in (0,1)$ and all distributions $\mu$ over $\{0,1\}^n$, there exists a polynomial $P : \mathbb{R}^n \to \mathbb{R}$ of degree $O((\log n)(\log 1/\varepsilon))$ such that $\mathbf{P}_{x\sim\mu}(P(x) = \mathrm{OR}(x)) > 1-\varepsilon$.*

By iteratively applying the above claim, Aspnes, Beigel, Furst, and Rudich [3] showed that $\mathsf{AC}^0$ circuits of depth $d$ have $\varepsilon$-approximating polynomials of degree at most

$$O\big(((\log s)(\log(1/\varepsilon)))^d \cdot (\log(s/\varepsilon))^{d-1}\big).$$

We prove that the following lower bound for such approximations:

**Theorem 2.6.** *There is a constant $c > 0$ and a distribution $\mu$ on $\{0,1\}^n$ such that for any polynomial $P : \{0,1\}^n \to \mathbb{R}$ of degree $d < c(\log\log n)/(\log\log\log n)$,*

$$\mathop{\mathbf{P}}_{x\sim\mu}\big(P(x) = \mathrm{OR}(x)\big) < 2/3\,.$$

To the best of our knowledge no $\omega(1)$ lower bound was known for approximating the OR function. We give an explicit distribution (directly motivated by the upper bound construction in [3]) under which OR has no $1/3$-error polynomial approximation. We define a distribution on $\{0,1\}^n$ by defining a corresponding random variable $x$ taking values in $\{0,1\}^n$ by the following procedure.

1. With probability $1/2$ output $x = (0,\dots,0)$.
2. With probability $1/2$ pick an index $i \in [D]$ uniformly at random and output $x \sim \mu^n_{2^{-a^i}}$ for some suitably chosen parameters $a, D$.

The analysis then proceeds at a high level as in the lower bound for parity. However, we need some extra care with the inductive argument as unlike for parity, we cannot consider arbitrary assignments of zeros and ones to some subset of the coordinates of the OR function. We get around this hurdle by instead only considering fixing parts of the input to 0 and decreasing the bias $p$ to make sure that these coordinates are indeed set to 0 with high probability. The details are deferred to Section 5.

## 2.2 The number of small subgraphs in a random graph

Consider the Erdős-Rényi random graph $G(N,p)$. Let $H$ be a small fixed graph (a triangle or $C_4$, say). The problem of counting the number of copies of $H$ in $G(N,p)$ is a fundamental topic in the theory of random graphs (see, for instance, the textbooks [5, 13]). In fact, we can talk about a more general problem of counting the number of copies of $H$ in a random subgraph of any deterministic graph $G$ on $N$ vertices formed by choosing each edge of $G$ with probability $p$. Here, $G'$ is said to be a copy of $H$ in a graph $G$ if $G'$ is isomorphic to $H$ and the vertex set and edge set of $G'$ are subsets of those of $G$. We denote this random variable by $F(H,G,p)$. In this setting we understand that $H$ has constant size, and the size of $G$ tends to infinity.

[3] $\mathrm{OR}(x_1,\dots,x_n)$ is 1 if any of the bits $x_i$ is non-zero.

It has been noticed that $F$ can be written as a polynomial in terms of the edge-indicator random variables. For example, the number of $C_4$ copies (cycle of length 4) is

$$\sum_{i,j,k,l} \xi_{ij}\xi_{jk}\xi_{kl}\xi_{li}$$

where the summation is over all quadruples $ijkl$ which form a $C_4$ in $G$ and the Bernoulli random variable $\xi_{ij}$ represents the edge $ij$. Clearly, any polynomial of this type has $n = e(G)$ Bernoulli random variables $\xi_{ij}$ with mean $p$, and its degree equals the number of edges of $H$. The rank $r$ of $F$ is exactly the size of the largest collection of edge-disjoint copies of $H$ in $G$.

The polynomial representation has been useful in proving *concentration* (i. e., *large deviation*) results for $F$ (see [15, 23], for instance). Interestingly, it has turned out that we can also use this to derive anti-concentration result, in particular bounds on the probability that the random graph has exactly $m$ copies of $H$.

By Theorem 1.7, we have

**Corollary 2.7.** *Assume that $p$ is a constant in $(0,1)$. Then for fixed $H$ and any integer $m$ which may depend on $G$*

$$\mathbf{P}(F(H,G,p) = m) \leq r^{-1/2+o(1)},$$

*where $r$ is the size of the largest collection of edge-disjoint copies of $H$ in $G$. In particular, if $G = K_n$, then*

$$\mathbf{P}(F(H,K_n,p) = m) \leq n^{-1/2+o(1)}.$$

A copy $G'$ of $H$ in a graph $G$ is said to be an *induced* copy if $G'$ contains all the edges in $G$ whose endpoints are both in the vertex set of $G'$. A similar argument can be used to deal with the number of induced copies of $H$, which can also be written as a polynomial with degree at most $\binom{v}{2}$, with $v$ being the number of vertices of $H$. Details are left out as an exercise.

Finally, let us mention that in a recent paper [11], Gilmer and Kopparty obtained a precise estimate for $\mathbf{P}(F(H,K_n,p) = m)$ in the case when $H$ is a triangle.[4] Their approach relies on a careful treatment of the characteristic function. It remains to be seen if this method applies to our more general setting.

## 3 Proof of Theorem 1.6

There are two proofs to this theorem. In the initial version of the paper, we had a longer, but more elementary proof based on the following ideas. To prove the first bound in Theorem 1.6, we first consider a simple case when the polynomial $P$ is sufficiently "nice," which basically means a fair contribution from each variable to $P$. These polynomials are said to be *regular* and their regularity allows one to efficiently relate the anti-concentration property of polynomials of Rademacher random variables to that of Gaussian ones and then use Theorem 1.4 for Gaussian case.

To complete the argument, we use a regularity lemma which shows that any polynomial can be written as a small-depth decision tree where most leaves are labeled by polynomials which are either (1) regular or (2) polynomials which are fixed in sign with high probability over a uniformly random input. In the

[4] We would like to thank J. Kahn for pointing out this reference.

first case, you get a regular polynomial of high rank (as the tree is shallow) and we apply the previous argument. In the second case, we argue directly that the anti-concentration probability is small.

To prove the second bound of Theorem 1.6, we follow the same conceptual approach but adopt a more careful analysis following the work of Kane [14]. We again use the regularity lemma to reduce to regular polynomials. And for a sufficiently regular polynomial, consider partitioning the variables into blocks of equal size. One can show that with significant probability, there exists a block on which the restricted polynomial has small standard deviation compared to its mean, which guarantees good anti-concentration bound.[5]

Here we present a shorter proof suggested by Daniel Kane. It shows that the desired anti-concentration probability can be bounded in terms of *average sensitivity* of *polynomial threshold functions* (see Lemma 3.4) and then uses the bounds on average sensitivity of [9] and [14]. The two proofs are quite close in spirit. The current one is shorter and perhaps more elegant, benefiting from Lemma 3.4, which is of independent interest.

To start the proof, by covering the interval $I$ by smaller intervals, we can assume that $I$ has length $2/3$. Let $S_1,\dots,S_r$ be the disjoint sets in the definition of $\mathrm{rank}(P)$. By conditioning on the random variables outside $\bigcup S_i$, we can assume that $n = dr$.

By subtracting the center of $I$ from $P$, we can then assume that $I = [-1/3, 1/3]$.

Let $f = \mathrm{sign}(P+2/5)$ and $g = \mathrm{sign}(P-1/3)$ where $\mathrm{sign}(a) := \mathbf{1}_{a>0}$. Note that the term $2/5$ can be replaced by anything that is slightly greater than $1/3$. For each $k\in[n]$ for each $x = (x_1,\dots,x_n) \in \{\pm 1\}^n$, let $x^k := (x_1,\dots,-x_k,\dots,x_n)$ be the point obtained by negating the $k$-th coordinate. Define the influence of the $k$-th variable to $f$ (and similarly for $g$) to be

$$\mathrm{Inf}_k(f) := \mathbf{P}\big(f(\xi_1,\dots,\xi_n) \neq f(\xi_1,\dots,,-\xi_k,\dots,\xi_n)\big)\,.$$

For each $S_i$, we claim that

$$\mathbf{P}\left(|P(\xi_1,\dots,\xi_n)| \le 1/3\right) \le 2^d \sum_{k\in S_i} \left(\mathrm{Inf}_k(f)+\mathrm{Inf}_k(g)\right)\,. \tag{3.1}$$

Assuming (3.1), taking the sum over $i\in[r]$, we get

$$r\,\mathbf{P}\left(|P(\xi_1,\dots,\xi_n)| \le 1/3\right) \le 2^d\left(\mathrm{AS}(f)+\mathrm{AS}(g)\right) \tag{3.2}$$

where

$$\mathrm{AS}(f) := \sum_{k=1}^{n} \mathrm{Inf}_k(f)$$

is called *average sensitivity* of $f$. Then the bounds in Theorem 1.6 follow from the following corresponding bounds on average sensitivity from Diakonikolas et al. [9] and Kane [14], respectively.

**Theorem 3.1** ([9, Theorem 1.1])**.** *There exists an absolute constant $C$ such that for any polynomial $Q$ of degree $d$ on $n$ random Rademacher variables, we have*

$$\mathrm{AS}(h) \le C^d(\log n)n^{1-1/(4d-2)}$$

*where $h = \mathrm{sign}(Q)$.*

[5]An interested reader can find this proof at http://arxiv.org/abs/1507.00829 [version 4].

Notice that by applying this bound, we only get a slightly weaker bound than the first bound in Theorem 1.6. Our claimed bound is proved by the same techniques as in [9] but with more careful analysis. We again refer the interested reader to our old proof.

**Theorem 3.2** ([14, Theorem 1.2]). *Under the assumptions of Theorem 3.1, we have*

$$\mathrm{AS}(h) \le \sqrt{n}(\log n)^{Cd\log d}C^{d^2\log d}.$$

To prove (3.1), let $\mathcal{G}$ be the set of all $x \in \{\pm 1\}^n$ such that $|P(x)| \le 1/3$ and let $\mathcal{G}'$ be the set of all $x \in \{\pm 1\}^n$ such that there exists $k \in S_i$ for which $f(x) \ne f(x^k)$ or $g(x) \ne g(x^k)$. It suffices to show that

$$|\mathcal{G}| \le 2^d|\mathcal{G}'|. \tag{3.3}$$

Without loss of generality, assume that $S_i = [d]$. Let $x \in \mathcal{G}$. Since

$$\mathbf{Var}_{\xi_1,\dots,\xi_d} P(\xi_1,\dots,\xi_d,x_{d+1},\dots,x_n) \ge 1,$$

$P(\xi_1,\dots,\xi_d,x_{d+1},\dots,x_n)$ cannot always stay in an interval of length less than 1. Thus, there exist $x'_1,\dots,x'_d$ such that

$$|P(x'_1,\dots,x'_d,x_{d+1},\dots,x_n)| > 2/5.$$

And so, either $f(x) \ne f(x'_1,\dots,x'_d,x_{d+1},\dots,x_n)$ or $g(x) \ne g(x'_1,\dots,x'_d,x_{r+1},\dots,x_n)$. Going bit by bit, there exists a $k \in S_i$ and an $x' \in \{\pm 1\}^n$ that differs from $x$ only in the first $d$ elements such that $f(x') \ne f(x'^k)$ or $g(x') \ne g(x'^k)$. In other words, $x' \in \mathcal{G}'$.

Since $x'$ and $x$ differ only in the first $d$ elements, the map from $\mathcal{G}$ to $\mathcal{G}'$ that maps $x$ to $x'$ is at most $2^d$-to-1, proving (3.3). This completes the proof of (3.1) and thereby completing the proof of Theorem 1.6. □

**Remark 3.3.** Inequality (3.2) is interesting in its own right. It readily implies the following bound.

**Lemma 3.4.** *Let $P$ be a degree-$d$ polynomial of the form* (1.1) *with rank $r \ge 2$. Then*

$$\mathbf{P}(|P(\xi_1,\dots,\xi_n)| \le 1/3) \le \frac{2^{d+1}}{r}\max_Q \mathrm{AS}(\mathrm{sign}(Q))$$

*where the maximum runs over all degree-$d$ polynomials $Q$ on $dr$ variables.*

It is worth mentioning the following conjecture.

**Conjecture 3.5** (Gotsman-Linial Conjecture [12]). *Let $Q$ be a degree-$d$ polynomial on $n$ variables. Then*

$$\mathrm{AS}(\mathrm{sign}(Q)) \le 2^{-n+1}\sum_{k=0}^{d-1}\binom{n}{\lfloor (n-k)/2\rfloor}(n-\lfloor (n-k)/2\rfloor). \tag{3.4}$$

For a fixed $d$, the right-hand side of (3.4) is $O(n^{-1/2})$. Thus, the Gotsman-Linial Conjecture implies the Conjecture in [19] that $\mathbf{P}(P(\xi_1,\dots,\xi_n) \in I) = O(r^{-1/2})$ for $P$ and $I$ as in the assumptions of Theorem 1.6.

# 4 General distributions

## 4.1 Proof of Theorem 1.7

We reduce the $p$-biased case to the uniform distribution at the expense of a loss in the rank of the polynomial and then apply Theorem 1.6.

First notice that if $x \sim \mu_p$, then $1-x \sim \mu_{1-p}$. And so, by replacing the polynomial $P$ by $Q(x_1,\dots,x_n) = P(1-x_1,\dots,1-x_n)$, we can exchange the roles of $p$ and $1-p$. Therefore, without loss of generality, we assume that $\alpha = p \le 1/2$.

Our assumption $2^d p^d r \ge 3$ guarantees that $\log\log(2^d p^d r) = \Omega(1)$ and hence by choosing the implicit constants on the right-hand side of Theorem 1.7 to be sufficiently large, we can assume that $2^d p^d r$ is greater than 100 (say).

Let $\eta_1,\dots,\eta_n$ and $\xi'_1,\dots,\xi'_n$ be independent Bernoulli random variables with $\mathbf{P}(\eta_i = 0) = 1/2$ and $\mathbf{P}(\xi'_i = 0) = 1-2p$. Let $\xi_i = \eta_i \xi'_i$ then $\xi_1,\dots,\xi_n$ are i. i. d. Bernoulli variables with $\mathbf{P}(\xi_i = 0) = 1-p$. Therefore, we need to bound $\mathbf{P}(P(\xi_1,\dots,\xi_n) \in I)$.

From the definition of $\mathrm{rank}(P)$, there exist disjoint sets $S_1,\dots,S_r$ such that $|a_{S_j}| \ge 1$ for all $j = 1,\dots,r$. We have

$$P(\xi_1,\dots,\xi_n) = \sum_{S\subset[n],|S|\le d} \left( a_S \prod_{i\in S} \xi'_i \right) \prod_{i\in S} \eta_i\,.$$

Conditioning on the variables $\xi'_i$, $P$ becomes a polynomial of degree $d$ in terms of the $\eta_i$, whose coefficients associated with $S_j$ are

$$b_{S_j} := a_{S_j} \prod_{i\in S_j} \xi'_i\,.$$

For each such $j$, we have

$$\mathop{\mathbf{P}}_{\xi'_1,\dots,\xi'_n}(|b_{S_j}| \ge 1) = \mathbf{P}(\xi'_i = 1, \forall i \in S_j) = (2p)^d\,.$$

Now, since the sets $S_j$ are disjoint, the events $|b_{S_j}| \ge 1$ are independent. Define

$$X = \sum_{j=1,\dots,r} \mathbf{1}_{|b_{S_j}|\ge 1}\,.$$

By Chernoff's bound (see, for example, [2, Corollary A.1.14]), we have, for $0 < \gamma < 1$,

$$\mathbf{P}(|X - \mathbf{E}X| \ge \gamma \mathbf{E}X) \le 2e^{-\gamma^2 \mathbf{E}X/3}\,.$$

Setting $\gamma = 1/2$, we conclude that with probability at least $1-\exp(-2^{d-1}p^d r/6)$, there are at least $2^{d-1}p^d r$ indices $j$ with $|b_j| \ge 1$. Conditioning on this event, we obtain a polynomial of degree $d$ in terms of $\eta_1,\dots,\eta_n$ which has rank at least $2^{d-1}p^d r$. The theorem now follows from applying Theorem 1.6 to this polynomial and noting that the additional error of $\exp(-2^{d-1}p^d r/6)$ is smaller than both terms from Theorem 1.6. $\square$

### 4.2 Proof of Theorem 1.8

By replacing $P(x_1,\dots,x_n)$ by $Q(x_1,\dots,x_n) = P(x_1+y_1,\dots,x_n+y_n)$ and $\xi_i$ by $\xi_i - y_i$, we can also assume without loss of generality that $y_i = 0$ for all $i$. Furthermore, we can assume that $\mathbf{P}(\xi_i \le 0) = p$ for all $i$. Indeed, if for some $i$, $\mathbf{P}(\xi_i > 0) = p$, we replace $\xi_i$ by $-\xi_i$ and modify the polynomial $P$ accordingly to reduce to the case $\mathbf{P}(\xi_i < 0) = p$. And then the proof runs along the same lines as in the case $\mathbf{P}(\xi_i \le 0) = 0$.

For each $i = 1,\dots,n$, let $\xi_i^+$ and $\xi_i^-$ be independent random variables satisfying

$$\mathbf{P}(\xi_i^+ \in A) = \mathbf{P}(\xi_i \in A \mid \xi_i > 0) \qquad \text{and} \qquad \mathbf{P}(\xi_i^- \in A) = \mathbf{P}(\xi_i \in A \mid \xi_i \le 0)$$

for all measurable subsets $A \subset \mathbb{R}$. Let $\eta_1,\dots,\eta_n$ be i. i. d. random Bernoulli variables (independent of all previous random variables) such that $\mathbf{P}(\eta_i = 0) = p$. Let $\xi_i' = \eta_i \xi_i^+ + (1-\eta_i)\xi_i^-$, then $\xi_i'$ and $\xi_i$ have the same distribution. Therefore, it suffices to bound the probability that $P(\xi_1',\dots,\xi_n')$ belongs to $I$. We have

$$\begin{aligned} P(\xi_1',\dots,\xi_n') &= P\big(\eta_1(\xi_1^+ - \xi_1^-) + \xi_1^-, \dots, \eta_n(\xi_n^+ - \xi_n^-) + \xi_n^-\big) \\ &= \sum_{S\subset[n],|S|=d} \left( a_S \prod_{i\in S} (\xi_i^+ - \xi_i^-) \right) \prod_{i\in S} \eta_i + Q, \end{aligned}$$

where $Q$ is some polynomial that has degree $< d$ in terms of the $\eta_i$ when all the $\xi_i^\pm$ are fixed. From the definition of rank$(P)$, let $S_1,\dots,S_r$ be disjoint subsets of $[n]$ with $|a_{S_j}| \ge 1$ for all $1 \le j \le r$. Conditioning on the variables $\xi_i^\pm$, the polynomial $P$ becomes a polynomial of degree $d$ in terms of the $\eta_i$, whose coefficients associated with $S_j$ are

$$b_{S_j} := a_{S_j} \prod_{i\in S_j} (\xi_i^+ - \xi_i^-)$$

accordingly. For each such $j$, we have

$$\mathop{\mathbf{P}}_{\xi_1^\pm,\dots,\xi_n^\pm} (|b_{S_j}| \ge 1) \ge \mathbf{P}(\xi_i^+ - \xi_i^- \ge 1, \forall i \in S_j).$$

Since $\xi_i^+ \ge 0 \ge \xi_i^-$ a. e., we have

$$2\,\mathbf{P}(\xi_i^+ - \xi_i^- \ge 1) \ge \mathbf{P}(\xi_i^+ \ge 1) + \mathbf{P}(\xi_i^- \le -1) = \mathbf{P}(|\xi_i| \ge 1) \ge \varepsilon\,.$$

Hence,

$$\mathop{\mathbf{P}}_{\xi_1^\pm,\dots,\xi_n^\pm} (|b_{S_j}| \ge 1) \ge 2^{-d}\varepsilon^d\,.$$

Now, since the sets $S_j$ are disjoint, the events $|b_{S_j}| \ge 1$ are independent. Therefore, using a Chernoff-type bound as in the proof of Theorem 1.7, we can conclude that with probability at least $1 - \exp(-2^{-d}\varepsilon^d r/12)$, there are at least $r2^{-d}\varepsilon^d/2$ indices $j$ with $|b_j| \ge 1$. Conditioning on this event, we obtain a polynomial of degree $d$ in terms of $\eta_1,\dots,\eta_n$ which has rank at least $r2^{-d}\varepsilon^d/2$. Using Theorem 1.7, we obtain the desired bound. $\square$

## 5 Proof of Theorem 2.6

Let $a$ be an integer to be chosen later. Let $D = \lfloor \log_a(\log_2 n - 1) \rfloor$ be the largest integer such that $2^{-a^D} \geq 2/n$. Let $\mu$ be the distribution obtained by the following procedure:

1. With probability $1/2$ output $x = \overline{0}$ (the all 0's vector).
2. With probability $1/2$ pick an index $i \in \{1, \ldots, D\}$ uniformly at random and output $x \sim \mu^n_{2^{-a^i}}$.

We next show that for some constant $c > 0$, there exists no polynomial $P$ of degree

$$d < c \frac{\log \log n}{\log \log \log n}$$

such that $\mathbf{P}_{x \sim \mu}(P(x) = \mathrm{OR}(x)) \geq 2/3$. Let $P$ be such a polynomial. Then, necessarily, $P(\overline{0}) = 0$; as

$$\mathop{\mathbf{P}}_{x \sim \mu}(P(x) = 0) \leq \frac{1}{2} + \frac{1}{2}\left(1 - 2^{-a^D}\right)^n \leq \frac{1}{2} + \frac{1}{2}\left(1 - \frac{2}{n}\right)^n < \frac{2}{3},$$

there must exist a set of indices $I \subseteq [D]$ with $|I| \geq \Omega(D)$ such that for all $i \in I$,

$$\mathop{\mathbf{P}}_{x \sim \mu_{2^{-a^i}}}(P(x) = 1) = \Omega(1).$$

Let $I = \{i_1 < i_2 < \cdots < i_k\}$ and for $\ell \in [k]$, let $p_\ell = 2^{-a^{i_\ell}}$. Now, by Theorem 1.7 applied to the polynomial $P - 1$ and $x \sim \mu^n_{p_1}$, we get that either $\mathrm{rank}(P) \leq (3/2p_1)^d$ or

$$\Omega(1) = \mathbf{P}(P(x) = 1) \leq O(d^{4/3}) \frac{\log(\mathrm{rank}(P)(2p_1)^d)^{1/2}}{(\mathrm{rank}(P)(2p_1)^d)^{1/(4d+1)}}.$$

Hence, in any case, $\mathrm{rank}(P) \leq r_1 = d^{O(d)}/p_1^d$. This in turn implies that there exists a set $S_1 \subseteq [n]$ of $r_1 \cdot d$ indices such that the polynomial $P_1 = P_{S_1}$ obtained by assigning the value 0 to the variables in $S_1$ is of degree at most $d - 1$. For a set $S \subseteq [n]$ and a distribution $\mu$ on $\{0, 1\}$, let $\mu^S$ denote the distribution where the coordinates in $S$ are chosen independently according to $\mu$. Then, for $x \sim \mu^{[n]}_{p_2}$,

$$\begin{aligned}
\Omega(1) &= \mathop{\mathbf{P}}_x(P(x) = 1) \\
&= \mathbf{P}(x_{S_1} = 0) \cdot \mathop{\mathbf{P}}_x(P(x) = 1 \mid x_{S_1} = 0) + \mathbf{P}(x_{S_1} \neq 0) \cdot \mathop{\mathbf{P}}_x(P(x) = 1 \mid x_{S_1} \neq 0) \\
&\leq \mathop{\mathbf{P}}_{x \sim \mu_{p_2}^{[n] \setminus [S_1]}}(P_1(x) = 1) + \mathbf{P}(x_{S_1} \neq 0) \\
&\leq \mathop{\mathbf{P}}_{x \sim \mu_{p_2}^{[n] \setminus [S_1]}}(P_1(x) = 1) + |S_1| \cdot p_2.
\end{aligned}$$

Thus,

$$\mathop{\mathbf{P}}_{x \sim \mu_{p_2}^{[n] \setminus [S_1]}}(P_1(x) = 1) \geq \Omega(1) - d^{O(d)+1} \cdot (p_2/p_1^d) = \Omega(1) - d^{O(d)+1} 2^{-a^{i_2} + da^{i_1}} \geq \Omega(1) - d^{O(d)} 2^{-a^{i_1}},$$

for $a \geq 2d$. Further, note that $P_1(\overline{0}) = 0$.

Iterating the argument with $P_1$ and so forth, we get a sequence of polynomials $P_1, P_2, \ldots, P_{k-1}$ such that for $1 \leq j \leq \min(d, k-1)$, $P_j$ is of degree at most $d - j$, $P_j(\overline{0}) = 0$ and for $x \sim \mu_{p_{j+1}}^{[n]\setminus(S_1 \cup \cdots \cup S_j)}$,

$$\mathbf{P}_x(P_j(x) = 1) = \Omega(1) - d^{O(d)+j} 2^{-a}.$$

This clearly leads to a contradiction if $k > d$ and $a \geq Cd \log d$ for a large enough constant $C$ (so that the right-hand side of the above equation is non-zero for $j = d$).

Therefore, setting $a = Cd \log d$, for a sufficiently large constant $C$, we must have $k = \Omega(D) \leq d$. That is, $\log_2(n-1) = a^{O(d)} = d^{O(d)}$. Thus, we must have $d = \Omega(1)(\log\log n)/(\log\log\log n)$, completing the proof of Theorem 2.6.

## 6 Acknowledgements

The authors would like to thank Daniel Kane for simplifying the proof of Theorem 1.6 and the anonymous referees for their very helpful remarks.

## References

[1] AMIR ABBOUD, RICHARD RYAN WILLIAMS, AND HUACHENG YU: More applications of the polynomial method to algorithm design. In *Proc. 26th Ann. ACM-SIAM Symp. on Discrete Algorithms (SODA'15)*, pp. 218–230. ACM Press, 2015. [doi:10.1137/1.9781611973730.17] 5

[2] NOGA ALON AND JOEL H. SPENCER: *The Probabilistic Method*. John Wiley & Sons, 2004. 11

[3] JAMES ASPNES, RICHARD BEIGEL, MERRICK L. FURST, AND STEVEN RUDICH: The expressive power of voting polynomials. *Combinatorica*, 14(2):135–148, 1994. Preliminary version in STOC'91. [doi:10.1007/BF01215346] 5, 7

[4] RICHARD BEIGEL, NICK REINGOLD, AND DANIEL A. SPIELMAN: The perceptron strikes back. In *Proc. 6th IEEE Conf. on Structure in Complexity Theory (SCTC'91)*, pp. 286–291. IEEE Comp. Soc. Press, 1991. [doi:10.1109/SCT.1991.160270] 5

[5] BÉLA BOLLOBÁS: *Random Graphs*. Volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge Univ. Press, 2001. [doi:10.1017/CBO9780511814068] 7

[6] ANTHONY CARBERY AND JAMES WRIGHT: Distributional and $L^q$ norm inequalities for polynomials over convex bodies in $\mathbb{R}^n$. *Math. Res. Lett.*, 8(3):233–248, 2001. [doi:10.4310/MRL.2001.v8.n3.a1] 2

[7] KEVIN P. COSTELLO: Bilinear and quadratic variants on the Littlewood-Offord problem. *Israel J. Math.*, 194(1):359–394, 2013. [doi:10.1007/s11856-012-0082-4, arXiv:0902.1538] 2

[8] KEVIN P. COSTELLO, TERENCE TAO, AND VAN H. VU: Random symmetric matrices are almost surely nonsingular. *Duke Math. J.*, 135(2):395–413, 2006. [doi:10.1215/S0012-7094-06-13527-5, arXiv:math/0505156] 2

[9] ILIAS DIAKONIKOLAS, PRASAD RAGHAVENDRA, ROCCO A. SERVEDIO, AND LI-YANG TAN: Average sensitivity and noise sensitivity of polynomial threshold functions. *SIAM J. Comput.*, 43(1):231–253, 2014. Preliminary version in STOC'10. [doi:10.1137/110855223, arXiv:0909.5011] 9, 10

[10] PAUL ERDŐS: On a lemma of Littlewood and Offord. *Bull. Amer. Math. Soc.*, 51(12):898–902, 1945. Available at Project Euclid. 2

[11] JUSTIN GILMER AND SWASTIK KOPPARTY: A local central limit theorem for triangles in a random graph. *Random Structures Algorithms*, 48(4):732–750, 2016. [doi:10.1002/rsa.20604, arXiv:1412.0257] 8

[12] CRAIG GOTSMAN AND NATHAN LINIAL: Spectral properties of threshold functions. *Combinatorica*, 14(1):35–50, 1994. [doi:10.1007/BF01305949] 10

[13] SVANTE JANSON, TOMASZ ŁUCZAK, AND ANDRZEJ RUCIŃSKI: *Random Graphs*. Volume 45. Wiley-Interscience, 2000. [doi:10.1002/9781118032718] 7

[14] DANIEL M. KANE: The correct exponent for the Gotsman-Linial conjecture. *Comput. Complexity*, 23(2):151–175, 2014. Preliminary version in CCC'13. [doi:10.1007/s00037-014-0086-z, arXiv:1210.1283] 9, 10

[15] JEONG HAN KIM AND VAN H. VU: Concentration of multivariate polynomials and its applications. *Combinatorica*, 20(3):417–434, 2000. [doi:10.1007/s004930070014] 8

[16] JOHN EDENSOR LITTLEWOOD AND ALBERT CYRIL OFFORD: On the number of real roots of a random algebraic equation (III). *Rec. Math. [Mat. Sbornik] N.S.*, 12(54)(3):277–286, 1943. MathNet. 2

[17] ELCHANAN MOSSEL, RYAN O'DONNELL, AND KRZYSZTOF OLESZKIEWICZ: Noise stability of functions with low influences: Invariance and optimality. *Ann. of Math.*, 171(1):295–341, 2010. Preliminary version in FOCS'05. [doi:10.4007/annals.2010.171.295] 3

[18] FEDOR NAZAROV, MIKHAIL SODIN, AND ALEXANDER VOL'BERG: The geometric Kannan-Lovász-Simonovits lemma, dimension-free estimates for volumes of sublevel sets of polynomials, and distribution of zeros of random analytic functions. *Algebra i Analiz*, 14(2):214–234, 2002. Available at Math-Net.Ru. [arXiv:math/0108212] 3

[19] HOI H. NGUYEN AND VAN H. VU: Small ball probability, inverse theorems, and applications. In *Erdős Centennial*, volume 25 of *Bolyai Soc. Math. Studies*, pp. 409–463. Springer, 2013. [doi:10.1007/978-3-642-39286-3_16, arXiv:1301.0019] 2, 3, 10

[20] ALEXANDER A. RAZBOROV: Lower bounds on the size of bounded depth circuits over a complete basis with logical addition. *Math. Notes*, 41(4):333–338, 1987. [doi:10.1007/BF01137685] 5

[21] ALEXANDER A. RAZBOROV AND EMANUELE VIOLA: Real advantage. *ACM Trans. Comput. Theory*, 5(4):17:1–17:8, 2013. Preliminary version in ECCC. [doi:10.1145/2540089] 2, 3, 6

[22] ROMAN SMOLENSKY: Algebraic methods in the theory of lower bounds for Boolean circuit complexity. In *Proc. 19th STOC*, pp. 77–82. ACM Press, 1987. [doi:10.1145/28395.28404] 5

[23] VAN H. VU: Concentration of non-Lipschitz functions and applications. *Random Structures Algorithms*, 20(3):262–316, 2002. [doi:10.1002/rsa.10032] 8

[24] RICHARD RYAN WILLIAMS: Faster all-pairs shortest paths via circuit complexity. In *Proc. 46th STOC*, pp. 664–673. ACM Press, 2014. [doi:10.1145/2591796.2591811, arXiv:1312.6680] 5

[25] RICHARD RYAN WILLIAMS: The polynomial method in circuit complexity applied to algorithm design (invited talk). In *Proc. 34th Internat. Conf. on Foundation of Software Tech. and Theoret. Comput. Sci. (FSTTCS'14)*, volume 29 of *LIPIcs*, pp. 47–60. Springer, 2014. [doi:10.4230/LIPIcs.FSTTCS.2014.47] 5

## AUTHORS

Raghu Meka
Assistant professor
University of California, Los Angeles
raghum@cs.ucla
`http://www.raghumeka.org/`

Oanh Nguyen
Ph. D. student
Yale University
New Haven, CT
oanh.nguyen@yale.edu
`http://users.math.yale.edu/public_html/People/otn2.html`

Van Vu
Professor
Yale University
New Haven, CT
van.vu@yale.edu
`http://users.math.yale.edu/public_html/People/vuha.html`

## ABOUT THE AUTHORS

RAGHU MEKA is an assistant professor in the Computer Science department at UCLA. He is broadly interested in complexity theory, learning, and probability theory. He got his Ph. D. from UT Austin under the direction of David Zuckerman. He spent two years as a postdoctoral fellow at the Institute for Advanced Study, Princeton and Rutgers University.

OANH NGUYEN is a graduate student in mathematics at Yale University. Her undergraduate advisor, Professor Duong Minh Duc, at University of Science, Ho Chi Minh City, Vietnam, played an important role in deepening her interest in analysis. She gained much of her knowledge and interest in probability from her graduate advisor, Van Vu. Her research interests include analysis, combinatorics, and probability theory.

VAN VU completed his undergraduate studies at Eötvös University (Budapest) in 1994. He then moved to Yale and wrote his Ph. D. thesis under the direction of László Lovász. He is currently the Percey F. Smith professor of mathematics at Yale, after having spent time at the Institute for Advanced Study, Microsoft Research, UC San Diego, and Rutgers. His research interests include probability, number theory, combinatorics, and theoretical computer science.