Small sample spaces cannot fool low degree polynomials

Noga Alon * Ido Ben-Eliezer † Michael Krivelevich ‡

Abstract

A distribution D on a set $S \subset \mathbb{Z}_p^N$ ϵ-fools polynomials of degree at most d in N variables over $\mathbb{Z}_p$ if for any such polynomial P, the distribution of $P(x)$ when x is chosen according to D differs from the distribution when x is chosen uniformly by at most ϵ in the ℓ_1 norm. Distributions of this type generalize the notion of ϵ-biased spaces and have been studied in several recent papers. We establish tight bounds on the minimum possible size of the support S of such a distribution, showing that any such S satisfies

$$|S| \geq c_1 \cdot \left(\frac{(\frac{N}{2d})^d \cdot \log p}{\epsilon^2 \log(\frac{1}{\epsilon})} + p \right).$$

This is nearly optimal as there is such an S of size at most

$$c_2 \cdot \frac{(\frac{3N}{d})^d \cdot \log p + p}{\epsilon^2}.$$

1 Introduction

Let P be a polynomial in N variables over $\mathbb{Z}_p$ of degree at most d. Let D be a distribution over a set S of vectors from $\mathbb{Z}_p^N$, and denote by U_N the uniform distribution on $\mathbb{Z}_p^N$. The distribution

*School of Mathematics, Institute for Advanced Study, Princeton, NJ 08540, USA, and Schools of Mathematics and Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel. Email: nogaa@tau.ac.il. Research supported in part by the Israel Science Foundation, by a USA-Israeli BSF grant and by the Hermann Minkowski Minerva Center for Geometry at Tel Aviv University.

†School of Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel. Email: idobene@post.tau.ac.il. This work forms part of the author's Ph.D. thesis.

‡School of Mathematical Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel. E-mail: krivelev@post.tau.ac.il. Research supported in part by USA-Israel BSF Grant 2006322, and by grant 526/05 from the Israel Science Foundation.

D is an ϵ-*approximation* of U_N with respect to P if

$$\sum_{a \in \mathbb{Z}_p} \left| \Pr_{x \sim D} [P(x) = a] - \Pr_{x \sim U_N} [P(x) = a] \right| \leq \epsilon.$$

We say that S (with the distribution D) is an (ϵ, N, d)-*biased space* if it is an ϵ-approximation with respect to any polynomial on N variables of degree at most d. Note that D is not necessarily a uniform distribution over its support S.

The case $d = 1$ is known as ϵ-*biased spaces*. Many works deal with such spaces, including efficient constructions, lower bounds and applications (see, for example, [3, 4, 5, 7, 16, 17, 19] and their references).

Luby et al. [15] gave an explicit construction for the general case, but the size of their sample space S is $2^{2^{O(\sqrt{\log (N/\epsilon)})}}$ even for the case $d = 2$. They used it to construct a deterministic approximation algorithm to the probability that a given depth-2 circuit outputs a certain value on a random input.

Bogdanov [8] gave better constructions that work for fields of size at least $poly(d, \log N, \frac{1}{\epsilon})$. Bogdanov and Viola [10] suggested a construction for general fields. The construction is the sum of d copies of ϵ-biased spaces, and the sample size is $N^d \cdot f(\epsilon, d)$ for some function f. However, the analysis of their construction relies on the so called "Inverse Gowers Conjecture" which was recently shown to be false [14]. Lovett [13] proved unconditionally that the sum of 2^d copies of ϵ-biased spaces fools polynomials of degree d, thus giving an explicit construction of size $(\frac{N}{\epsilon})^{2^{O(d)}}$. Later, Viola [20] proved that the sum of d copies is sufficient. This yields an explicit construction of size $\frac{N^d}{\epsilon^{O(2^d)}}$ using the best known constructions of ϵ-biased spaces. Recently, Bogdanov et. al. [9] showed how to fool width-2 branching programs using such distributions.

Here we study the minimum possible size of (ϵ, N, d)-biased spaces. Bogdanov and Viola [10] observed that for $p = 2$ and $\epsilon < 2^{-d}$ every such space is of size at least $\binom{N}{d}$. Their argument is very simple: The set of polynomials of degree at most d forms a linear space of dimension $\sum_{i=0}^{d} \binom{N}{i} > \binom{N}{d}$. If S is of size less than $\binom{N}{d}$ then there is a non-zero polynomial P such that $P(x) = 0$ for every $x \in S$, and since every non zero polynomial is not zero with probability at least 2^{-d} (as follows, for example, by considering the minimal distance of the Reed-Muller code of order d) we get the desired bound. However, their bound doesn't depend on ϵ and, for small values of ϵ, is far from optimal and also from the known bound for ϵ-biased space, which is nearly optimal for $d = 1$. Our main contribution is a nearly tight lower bound on the size of such spaces as a a function of all four parameters ϵ, N,d and p. Note that as spaces of this type can be useful in derandomization, where the running time of the resulting algorithms is

proportional to the size of the space, it is interesting to get a tight bound for their smallest possible size.

Theorem 1. *There is an absolute constant $c_1 > 0$ so that for every $\epsilon \geq d \cdot 2^{-\frac{N}{2d}}, d \leq \frac{N}{10}$, every (ϵ, N, d)-biased space over $\mathbb{Z}_p$ has size at least*

$$\max\left\{c_1 \cdot \frac{(\frac{N}{2d})^d \log p}{\epsilon^2 \log(\frac{1}{\epsilon})}, p(1-\epsilon)\right\}.$$

We also observe that this bound is nearly tight by proving the following simple statement:

Proposition 2. *There is an absolute constant $c_2 > 0$ so that for every $d \leq \frac{N}{10}$ there is an (ϵ, N, d)-biased space over $\mathbb{Z}_p$ of size at most $c_2 \cdot \frac{(\frac{3N}{d})^d \log p + p}{\epsilon^2}$.*

The proofs are described in the next section; for completeness, we include some of the details in the appendix. The final section contains some concluding remarks. Throughout the proofs we omit all floor and ceiling signs whenever these are not crucial.

2 Proofs

In this section we present the proofs of our results. The proof of our main result, Theorem 1, lower bounding the size of an (ϵ, N, d)-biased set, is given in Section 2.1. The proof of the upper bound (Proposition 2) is in Section 2.2.

2.1 Lower bound

First we observe that a bound of $p(1-\epsilon)$ follows easily as otherwise the distribution doesn't fool every polynomial P for which $P(x)$ is the uniform distribution (for example, all the linear polynomials). Let N be the number of variables and let d be the degree of the polynomial. Assume for simplicity that $N = nd$, where n is an integer. For every $i \geq 1$ define the set of variables $S_i = \{x_{i,1}, ..., x_{i,n}\}$. A monomial over $\mathbb{Z}_p$ is called d-partite if it has the form $\prod_{1 \leq i \leq d} x_{i,j_i}$, and a polynomial over $\mathbb{Z}_p$ is called d-partite if it is a sum of d-partite monomials. Note that d-partite polynomials are homogeneous polynomials of degree d.

Let $P_{n,d}$ be the uniform distribution on the set of d-partite polynomials. A random element in $P_{n,d}$ is a sum of d-partite monomials, where every one of the possible n^d monomials has a random coefficient selected uniformly and independently from $\mathbb{Z}_p$.

An assignment to the variables $\{x_i\}$ is non-trivial if there is an i such that $x_i \neq 0$. Similarly, if $v_1, v_2, ..., v_n \in V$ for some vector space V, a linear combination $\sum_i \alpha_i v_i$ is non-trivial if there is i such that $\alpha_i \neq 0$. For a prime p, a polynomial P over $\mathbb{Z}_p$ is *δ-balanced* if

$$\sum_{a \in \mathbb{Z}_p} \left| \frac{|\{x : P(x) = a\}|}{p^N} - \frac{1}{p} \right| \leq \delta.$$

A polynomial is *balanced* if it is 0-balanced. We have the following key lemma:

Lemma 1. *The probability $\phi(n, d)$ that a random element from $P_{n,d}$ is $d \cdot p^{-\frac{n}{2}}$-balanced is at least*

$$1 - p^{-(\frac{n}{2})^d + 2(\frac{n}{2})^{d-1} + \sum_{i=0}^{d-3} (\frac{n}{2})^i (\frac{n^2}{4} + n)} \geq 1 - p^{-(\frac{n}{2})^d + 4(\frac{n}{2})^{d-1}}.$$

Proof. We apply induction on d. For $d = 1$, as every non-trivial linear polynomial is balanced, we have $\phi(n, 1) = 1 - p^{-n}$, and the statement holds. Assuming that the statement is valid for d, we prove it for $d+1$. A random $(d+1)$-partite polynomial P can be represented as $\sum_{i=1}^{n} x_{1,i} P_i$, where for every i, P_i is a random polynomial (distributed uniformly and independently over $P_{n,d}$) over the sets of variables $S_2, S_3, ..., S_{d+1}$. Denote the set $\{P_i\}$ of polynomials by $\mathcal{P}$. We use the following claim:

Claim 1. *With probability at least $1 - p^{-(\frac{n}{2})^d + 2(\frac{n}{2})^{d-1} + \sum_{i=0}^{d-3} (\frac{n}{2})^i (\frac{n^2}{4} + n)}$ over the choice of polynomials in $\mathcal{P}$, there is a subset $B \subseteq \mathcal{P}$ of size at least $\frac{n}{2}$ such that for any non-trivial choice of $\{\alpha_i\}$, the polynomial $\sum_{P_i \in B} \alpha_i P_i$ is $d \cdot p^{-\frac{n}{2}}$-balanced.*

Proof. Let $B_0 := \emptyset$. In the i'th step, we consider the polynomial P_i. If P_i as well as all its combinations with elements from B_{i-1} are $d \cdot p^{-\frac{n}{2}}$-balanced, we set $B_i := B_{i-1} \bigcup \{P_i\}$, otherwise we call the step *bad* and let $B_i := B_{i-1}$. After the last polynomial, set $B := B_n$. We want to bound the probability that there are more than $\frac{n}{2}$ bad steps. Consider a certain step i and assume that $|B_{i-1}| < \frac{n}{2}$. Since P_i is a random polynomial, the sum of P_i with every fixed polynomial is uniformly distributed over the set $P_{n,d}$. By the induction hypothesis, it is $d \cdot p^{-\frac{n}{2}}$-balanced with probability at least $1 - p^{-(\frac{n}{2})^d + 2(\frac{n}{2})^{d-1} + \sum_{i=0}^{d-3} (\frac{n}{2})^i (\frac{n^2}{4} + n)}$. By the union bound, the probability that the step is bad is at most $2^{n/2} \cdot p^{-(\frac{n}{2})^d + 2(\frac{n}{2})^{d-1} + \sum_{i=0}^{d-3} (\frac{n}{2})^i (\frac{n^2}{4} + n)}$. We bound the probability that there are more than $\frac{n}{2}$ bad steps. For $d = 2$ the probability is at most

$$\binom{n}{\frac{n}{2}} \left(2^{n/2} \cdot p^{-n}\right)^{n/2} \leq p^{-(\frac{n}{2})^2 + n}.$$

For $d \geq 3$, we have:

$$\binom{n}{\frac{n}{2}} \left(2^{n/2} \cdot p^{-(\frac{n}{2})^d + 2((\frac{n}{2})^{d-1}) + \sum_{i=0}^{d-3}(\frac{n}{2})^i(\frac{n^2}{4}+n)}\right)^{n/2} \leq p^{-(\frac{n}{2})^{d+1} + 2((\frac{n}{2})^d) + \sum_{i=0}^{d-2}(\frac{n}{2})^i(\frac{n^2}{4}+n)}.$$

The claim follows. □

Assume that the condition of the claim holds, and without loss of generality assume that $\{P_1, P_2, ..., P_{n/2}\} \subseteq B$. Let $P' = \sum_{i=1}^{n/2} x_{1,i}P_i$. By Claim 1, for every non-trivial assignment of the variables $\{x_{1,i}\}$, the obtained polynomial is $d \cdot p^{-\frac{n}{2}}$-balanced. The probability that the assignment of the variables $\{x_{1,i}\}$ is trivial is $p^{-\frac{n}{2}}$. Therefore, P' is δ-balanced, where

$$\delta \leq p^{-\frac{n}{2}} + d \cdot p^{-\frac{n}{2}} = (d+1) \cdot p^{-\frac{n}{2}}. \tag{1}$$

We use this fact to prove that the polynomial P is $(d+1)\cdot p^{-\frac{n}{2}}$-balanced. For every assignment of the variables from $\bigcup_{2\leq i\leq d+1} S_i$, P reduces to a linear polynomial, which depends only on the variables from S_1. Denote by $\mu(P)$ (respectively, $\mu(P')$) the probability over the assignments of $\bigcup_{2\leq i\leq d+1} S_i$ that P (respectively, P') reduces to a trivial linear polynomial. Clearly $\mu \leq \mu'$ and μ is an upper bound on the imbalance of P. Therefore, it is sufficient to prove that μ' is bounded by $(d+1) \cdot p^{-\frac{n}{2}}$. To this end, note that whenever P' is reduced to a constant polynomial it is actually reduced to the zero polynomial. Therefore, as the bias of P' is bounded by $(d+1)\cdot p^{-\frac{n}{2}}$, the lemma follows.

□

We construct a set of polynomials Q as follows. Let $r = \log(\frac{1}{1-\phi(n,d)}) - 1 \geq (\frac{N}{2d})^d - 4(\frac{N}{2d})^{d-1} - 1$. For every $1 \leq i \leq r$ let q_i be a polynomial distributed uniformly and independently over $P_{n,d}$. Denote by Q the set of all non-trivial combinations of $\{q_1, ..., q_r\}$.

By the union bound and by Lemma 1, with positive probability all the elements of Q are $d \cdot q^{-\frac{n}{2}}$-balanced. Fix Q to be such a set. It follows also that the vectors $q_1, q_2, ..., q_r$ are linearly independent (otherwise Q contains the zero vector, which is not $d \cdot q^{-\frac{n}{2}}$-balanced). Therefore, $|Q| \geq q^{(\frac{N}{2d})^d - 4(\frac{N}{2d})^{d-1} - 1} - 1$.

The following lemma is due to Alon [2]:

Lemma 2 ([2])**.** *There exists an absolute positive constant c so that the following holds. Let B be an n by n real matrix with $b_{i,i} \geq \frac{1}{2}$ for all i and $|b_{i,j}| \leq \epsilon$ for all $i \neq j$ where $\frac{1}{2\sqrt{n}} \leq \epsilon \leq \frac{1}{4}$.*

Then the rank of B satisfies

$$rank(B) \geq \frac{c \log n}{\epsilon^2 \log(\frac{1}{\epsilon})}.$$

Here we need the following complex variant of the lemma:

Lemma 3. *There exists an absolute positive constant c so that the following holds. Let C be an n by n complex matrix with $|c_{i,i}| \geq \frac{1}{2}$ for all i and $|c_{i,j}| \leq \epsilon$ for all $i \neq j$ where $\frac{1}{2\sqrt{n}} \leq \epsilon \leq \frac{1}{4}$. Then the rank of C satisfies*

$$rank(C) \geq \frac{c \log n}{\epsilon^2 \log(\frac{1}{\epsilon})}.$$

We give the proof of this lemma in the appendix. For completeness we also reproduce there the proof of Lemma 2.

We are now ready to prove Theorem 1:

Proof of Theorem 1. Suppose that W is an (ϵ, N, d)-biased space, and that $W = \{w_1, w_2, ..., w_m\}$, $\Pr[w_i] = t_i$. Define a $|Q|$-by-m complex matrix U whose rows are indexed by the elements of Q and whose columns are indexed by the elements of W. Set $U_{q,w_i} = (\xi_p)^{q(w_i)}\sqrt{t_i}$, where ξ_p is a primitive root of unity of order p and the value of $q(w_i)$ is computed over $\mathbb{Z}_p$. Note that by our choice of Q and the definition of an (ϵ, N, d)-biased space, for every $q \in Q$:

$$|\sum_{i=1}^{m} (\xi_p)^{q(w_i)} \cdot t_i| \leq 2\epsilon + d \cdot q^{-\frac{n}{2}} \leq 3\epsilon.$$

Note that the first term in the last inequallity follows from the fact that W is ϵ-close to the uniform distribution and that for every two roots of unity z_1, z_2, by the triangle inequallity $|z_1 - z_2| \leq 2$. Also, obviously:

$$\sum_{i=1}^{m} t_i = 1.$$

For every two distinct polynomials $q_1, q_2 \in Q$, the polynomial $q_1 - q_2$ is also in Q, and for every w_i we have

$$(\xi_p)^{(q_1-q_2)(w_i)} = (\xi_p)^{q_1(w_i)} \cdot (\xi_p)^{-q_2(w_i)}.$$

Set $A = UU^*$. For every distinct $q_1, q_2 \in Q$ we have:

$$|A_{q_1,q_2}| = |\sum_{i=1}^{m} (\xi_p)^{(q_1-q_2)(w_i)} \cdot t_i| \leq 3\epsilon.$$

All the diagonal entries in A are 1. Since the rank of U is at most m the rank of A is also at most m. By Lemma 3:

$$m \geq rank(A) \geq c \cdot \frac{\log |Q|}{\epsilon^2 \log (\frac{1}{\epsilon})} \geq c_1 \cdot \frac{(\frac{N}{2d})^d \cdot \log p}{\epsilon^2 \log (\frac{1}{\epsilon})}.$$

The desired result follows.

□

2.2 Upper bound

Here we prove the simple upper bound:

Proof of Propostion 2. Consider a random set $R \subseteq \mathbb{Z}_p^N$ of size $m = \frac{\log (p)(\frac{3N}{d})^d + p}{\epsilon^2}$. We bound the probability that for a given polynomial P, the uniform distribution on R is not an ϵ-approximation with respect to P.

Let $L \subset \mathbb{Z}_p$, and let $\mu_L = m \sum_{a \in L} \Pr_{x \in U_n} [p(x) = a]$ be the expected number of vectors from R such that P evaluates to elements from L. By the Chernoff bounds (see, e.g., [6], Appendix A), we have:

$$\Pr_R [\Pr_{x \in U_n} [P(x) \in L] - \Pr_{x \in R} [P(x) \in L] > \epsilon] \leq e^{-\mu_L \cdot (\frac{\epsilon m}{\mu_L})^2/2} \leq e^{-m\epsilon^2/2}.$$

By the union bound over all 2^p possible sets L, the probability that the uniform distribution on R is not an ϵ-approximation is at most $e^{-m\epsilon^2/2+p}$.

The number of normalized monomials of degree at most d is exactly the number of ways to put d identical balls in $N+1$ distinct bins, and is bounded by

$$\binom{d+N}{d} \leq \left(\frac{e(N+d)}{d}\right)^d \leq \left(\frac{3N}{d}\right)^d.$$

Therefore the total number of polynomials of degree at most d is at most

$$p^{(\frac{3N}{d})^d} = 2^{(\frac{3N}{d})^d \cdot \log p}.$$

By applying the union bound, with high probability the uniform distribution on R is an ϵ-approximation with respect to any polynomial on N variables with degree at most d, and the theorem follows.

□

3 Concluding Remarks

For $p \ll \binom{n}{d}$, the ratio between the lower and upper bounds is $c \cdot (2e)^d \log\left(\frac{1}{\epsilon}\right)$ for some constant c. In particular, for fixed d the ratio is $\Theta(\log\left(\frac{1}{\epsilon}\right))$. This matches the ratio between the best known upper and lower bounds in the case $d = 1$ that corresponds to ϵ-biased spaces.

Our bound is valid only for $\epsilon \geq d \cdot p^{-\frac{N}{2d}}$. As noted in [2], for $\epsilon \leq p^{-\frac{N}{2}}$ every ϵ-biased space must be essentially the whole space (even for $d = 1$). It may be interesting to close the gap between $p^{-\frac{N}{2}}$ and $d \cdot p^{-\frac{N}{2d}}$. We note that as we can see from the proof of Lemma 1, for every d-partite polynomial P, $\Pr_x[P(x) = 0] \geq \Pr_x[P(x) = a]$ for every $0 \neq a \in Z_p$. By considering a random assignment of the variables from S_1 we get that the imbalance of the polynomial is at least $p^{-\frac{N}{d}}$. Hence we need a different approach in order to deal with smaller values of ϵ.

Recently, Schechtman and Shraibman [18] proved a strengthening of Lemma 2. They showed that under the conditions of Lemma 2, if A is also positive-semidefinite then we need only an upper bound on the values of non-diagonal entries, instead of an upper bound on their absolute values. In our case, for $p = 2$ the matrix A is positive semidefinite, and we can thus relax the conditions and establish a similar lower bound for the size of the support of any distribution in which no polynomial attains the value zero with probability bigger by $\epsilon/2$ than the probability it attains it in the uniform distribution. That is, for $p = 2$ the lower bound for the size of the distribution holds, even if there is no lower bound on the probability that each polynomial attains the value zero.

Lemma 1 can also be formulated in the language of error correcting codes. For given N and d, it states that every Reed-Muller code with parameters N and d contains a dense linear subcode in which every nontrivial codeword is balanced.

Recently, Dvir and Shpilka [12] gave an efficient encoding and decoding procedures for the construction of sum of d copies of ϵ-biased spaces.

Acknowledgements. We thank Avi Wigderson and Shachar Lovett for helpful comments.

References

[1] *N. Alon*, Problems and results in extremal combinatorics, I, Discrete Math. 273 (2003), 31-53.

[2] *N. Alon*, Perturbed identity matrices have high rank: proof and applications, to appear in Combinatorics, Probability and Computing.

[3] *N. Alon, A. Andoni, T. Kaufman, K. Matulef, R. Rubinfeld and N. Xie*, Testing k-wise and almost k-wise independence, Proceedings of the 39th Annual ACM Symposium, STOC 2007, 496-505.

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

[5] *N. Alon and Y. Roichman*, Random Cayley graphs and expanders, Random Structures and Algorithms 5 (1994), 271-284.

[6] *N. Alon and J. Spencer*, The Probabilistic Method, Second Edition, Wiley, New York, 2000.

[7] *E. Ben-Sasson, M. Sudan, S. Vadhan and A. Wigderson*, Randomness-efficient low degree tests and short PCPs via epsilon-biased sets, Proceedings of the 35th Annual ACM Symposium, STOC 2003, pp. 612-621.

[8] *A. Bogdanov*, Pseudorandom generators for low degree polynomials. Proceedings of the 37th Annual ACM Symposium, STOC 2005, 21-30.

[9] *A. Bogdanov, Z. Dvir, E. Verbin and A. Yehudayoff*, Pseudorandomness for width 2 branching programs. Manuscript, 2008.

[10] *A. Bogdanov and E. Viola*, Pseudorandom bits for polynomials, Proceedings of the 38th Annual Symposium on Foundations of Computer Science (FOCS), 2007.

[11] *B. Codenotti, P. Pudlák and G. Resta*, Some structural properties of low-rank matrices related to computational complexity, Theoret. Comput. Sci. 235 (2000), 89–107.

[12] *Z. Dvir and A. Shpilka*, Noisy interpolating sets for low degree polynomials, to appear in Proceedings of the 23th IEEE Conference on Computational Complexity (CCC), 2008.

[13] *S. Lovett*, Unconditional pseudorandom generators for low degree polynomials, to appear in Proceedings of the 40th Annual ACM Symposium, STOC 2008.

[14] *S. Lovett, R. Meshulam and A. Samorodnitsky*, Inverse conjecture for the Gowers norm is false, to appear in Proceedings of the 40th Annual ACM Symposium, STOC 2008.

[15] *M. Luby, B. Velickovic and A. Wigderson*, Deterministic approximate counting of depth-2 circuits, Proceedings of the 2nd ISTCS (Israeli Symposium on Theoretical Computer Science), pp. 18-24, 1993.

[16] *R. Motwani, J. Naor and M. Naor*, The probabilistic method yields deterministic parallel algorithms, JCSS 49(3) (1994), 478-516.

[17] *J. Naor and M. Naor*, Small bias probability spaces: efficient constructions and applications, Proceedings of the 22th Annual ACM Symposium, STOC 1990, pp. 213-223.

[18] *G. Schechtman and A. Shraibman*, Lower bounds for local versions of dimension reductions, Manuscript, 2007.

[19] *A. Shpilka*, Constructions of low-degree and error-correcting ϵ-biased generators, Proceedings of the 21st Annual IEEE Conference on Computational Complexity (CCC), Prague, Czech Republic, 2006, pp. 33-45.

[20] *E. Viola*, The sum of d small-bias generators fools polynomials of degree d, to appear in Proceedings of the 23th IEEE Conference on Computational Complexity (CCC), 2008.

A A complex variant of Lemma 2

In this section we reproduce the proof of Lemma 2 (omitting the final detailed computation) as given in [2], and also prove Lemma 3.

We start with the following lemma from which Lemma 2 will follow:

Lemma 4. *There exists an absolute positive constant c so that the following holds. Let B be an n by n real matrix with $b_{i,i} = 1$ for all i and $|b_{i,j}| \leq \epsilon$ for all $i \neq j$. If $\frac{1}{\sqrt{n}} \leq \epsilon < 1/2$, then*

$$rank(B) \geq \frac{c}{\epsilon^2 \log(1/\epsilon)} \log n.$$

We need the following well known lemma proved, among other places, in [11], [1].

Lemma 5. *Let $A = (a_{i,j})$ be an n by n real, symmetric matrix with $a_{i,i} = 1$ for all i and $|a_{i,j}| \leq \epsilon$ for all $i \neq j$. If the rank of A is d, then*

$$d \geq \frac{n}{1 + (n-1)\epsilon^2}.$$

In particular, if $\epsilon \leq \frac{1}{\sqrt{n}}$ then $d > n/2$.

Proof. Let $\lambda_1, \ldots, \lambda_n$ denote the eigenvalues of A, then their sum is the trace of A, which is n, and at most d of them are nonzero. Thus, by Cauchy-Schwartz, $\sum_{i=1}^{n} \lambda_i^2 \geq d(n/d)^2 = n^2/d$. On the other hand, this sum is the trace of $A^t A$, which is precisely $\sum_{i,j} a_{i,j}^2 \leq n + n(n-1)\epsilon^2$. Hence $n + n(n-1)\epsilon^2 \geq n^2/d$, implying the desired result. □

Lemma 6. *Let $B = (b_{i,j})$ be an n by n matrix of rank d, and let $P(x)$ be an arbitrary polynomial of degree k. Then the rank of the n by n matrix $(P(b_{i,j}))$ is at most $\binom{k+d}{k}$. Moreover, if $P(x) = x^k$ then the rank of $(P(b_{i,j}))$ is at most $\binom{k+d-1}{k}$.*

Proof. Let $\mathbf{v_1} = (v_{1,j})_{j=1}^{n}, \mathbf{v}_2 = (v_{2,j})_{j=1}^{n}, \ldots, \mathbf{v_d} = (v_{d,j})_{j=1}^{n}$ be a basis of the row-space of B. Then the vectors $(v_{1,j}^{k_1} \cdot v_{2,j}^{k_2} \cdots v_{d,j}^{k_d})_{j=1}^{n}$, where $k_1, k_2, \ldots, k_d$ range over all non-negative integers whose sum is at most k, span the rows of the matrix $(P(b_{i,j}))$. In case $P(x) = x^k$ it suffices to take all these vectors corresponding to $k_1, k_2, \ldots, k_d$ whose sum is precisely k. □

Proof of Lemma 4. We may and will assume that B is symmetric, since otherwise we simply apply the result to $(B + B^t)/2$ whose rank is at most twice the rank of B. Put $d = rank(B)$. If $\epsilon \leq 1/n^\delta$ for some fixed $\delta > 0$, the result follows by applying Lemma 5 to a $\lfloor \frac{1}{\epsilon^2} \rfloor$ by $\lfloor \frac{1}{\epsilon^2} \rfloor$ principal submatrix of B. Thus we may assume that $\epsilon \geq 1/n^\delta$ for some fixed, small $\delta > 0$. Put $k = \lfloor \frac{\log n}{2\log(1/\epsilon)} \rfloor$, $n' = \lfloor \frac{1}{\epsilon^{2k}} \rfloor$ and note that $n' \leq n$ and that $\epsilon^k \leq \frac{1}{\sqrt{n'}}$. By Lemma 6 the rank of the n' by n' matrix $(b_{i,j}^k)_{i,j \leq n'}$ is at most $\binom{d+k}{k} \leq (\frac{e(k+d)}{k})^k$. On the other hand, by Lemma 5, the rank of this matrix is at least $n'/2$. Therefore

$$\left(\frac{e(k+d)}{k}\right)^k \geq \frac{n'}{2} = \frac{1}{2}\lfloor \frac{1}{\epsilon^{2k}} \rfloor,$$

and the desired result follows by some simple (though somewhat tedious) manipulation, which we omit. □

Proof of Lemma 2. Let $C = (c_{i,j})$ be the n by n diagonal matrix defined by $c_{i,i} = 1/b_{i,i}$ for all i. Then every diagonal entry of CB is 1 and every off-diagonal entry is of absolute value at most 2ϵ. The result thus follows from Lemma 4. □

Proof of Lemma 3. Let P be an n by n diagonal matrix defined by $p_{i,i} = 1/c_{i,i}$ and set $D = CP$. Then every diagonal entry of D is 1 and every off-diagonal entry is of absolute value at most 2ϵ. Set $D' = (D + D^*)/2$. Then D' is a real matrix and $rank(D') \leq 2 \cdot rank(D)$. The desired result follows by applying Lemma 4 to D'. $\square$