Bounded independence vs. moduli

Ravi Boppana * Johan Håstad † Chin Ho Lee ‡ Emanuele Viola ‡

July 5, 2016

Abstract

Let $k = k(n)$ be the largest integer such that there exists a k-wise uniform distribution over $\{0,1\}^n$ that is supported on the set $S_m := \{x \in \{0,1\}^n : \sum_i x_i \equiv 0 \bmod m\}$, where m is any integer. We show that $\Omega(n/m^2 \log m) \le k \le 2n/m+2$. For $k = O(n/m)$ we also show that any k-wise uniform distribution puts probability mass at most $1/m+1/100$ over S_m. Finally, for any fixed odd m we show that there is $k = (1-\Omega(1))n$ such that any k-wise uniform distribution lands in S_m with probability exponentially close to $|S_m|/2^n$; and this result is false for any even m.

*Research Affiliate, Department of Mathematics, Massachusetts Institute of Technology

†KTH-Royal Institute of Technology, work done in part while visiting the Simons Institute. Supported by the Swedish Research Council.

‡College of Computer and Information Science, Northeastern University. Supported by NSF grant CCF-1319206. Work done in part while visiting Harvard University, with support from Salil Vadhan's Simons Investigator grant, and in part while at the Simons Institute for the Theory of Computing.

1 Introduction and our results

A distribution on $\{0,1\}^n$ is k-wise uniform if any k bits are uniform in $\{0,1\}^k$. Researchers have analyzed various classes of tests that cannot distinguish distributions with k-wise uniformity from uniform. Such tests include (combinatorial) rectangles [EGL+98] (cf. [CRS00]), bounded-depth circuits [Baz09, Raz09, Bra10, Tal14], and halfspaces [DGJ+10, GOWZ10, DKN10], to name a few. We say that such tests are *fooled* by distributions with bounded independence.

In this work we consider the mod m tests, defined next.

Definition 1. *For an input length n, and an integer m, we define the set $S_m := \{x \in \{0,1\}^n : \sum_i x_i \equiv 0 \bmod m\}$.*

These tests have been intensely studied at least since circuit complexity theory hit the wall of gates computing mod m for composite m in the 80's. However, the effect of bounded independence on mod m tests does not seem to have been known before this paper.

Our first main result is that there exist distributions with linear uniformity that are supported on S_m.

Theorem 2. *There exists a $c > 0$ such that the following holds.*

For every integer $m \geq 2$, there exists a $k \geq cn/m^2 \log m$ and a k-wise uniform distribution over $\{0,1\}^n$ that is supported on S_m.

This proves a conjecture in [LV15] where this question is also raised. Their motivation was a study of the "mod 3" dimension of k-wise uniform distributions, started in [MZ09], which is the dimension of the space spanned by the support of the distribution over GF(3). [LV15] shows that $k = 100 \log n$-wise uniformity with dimension $\leq n^{0.49}$ would have applications to pseudorandomness. It also exhibits a distribution with dimension $n^{0.72}$ and uniformity $k = 2$. Theorem 2 yields a distribution with dimension $n-1$ and $\Omega(n)$-wise uniformity.

We then prove three results, summarized in the next theorem, that show that k-wise uniformity does fool mod m when k is large. (1) shows that the largest possible value of k in Theorem 2 is $k \leq 2(n+1)/m + 2 \leq (1 - \Omega(1))n$. (2) shows that when k is larger than $(1-\gamma)n$ for a constant γ depending only on m then k-wise uniformity fools S_m with exponentially small error when m is odd. The proof of (2) however does not carry to the setting of $k < n/2$, for any m. So we establish (3) which gives a worse error bound but allows for k to become smaller for larger m, specifically $k = O(n/m)$ for constant error. The error bound in (3) and the density of S_m are such that (3) only provides a meaningful upper bound on the probability that the k-wise uniform distribution lands in S_m, but not a lower bound. In fact, we conjecture that no lower bound is possible in the sense that there is $c > 0$ such that for every m there is a cn-wise uniform distribution supported on the complement of S_m.

The combination of (2) and (3) implies that for $k = \min\{O(n/m), (1 - \Omega(1))n\}$ any k-wise uniform distribution puts probability mass at most $1/m + 1/100$ over S_m for odd m.

Theorem 3. *Let m be an integer.*

(1) For $k \geq 2n/m + 2$, a k-wise uniform distribution over $\{0,1\}^n$ cannot be supported on S_m.

(2) Suppose m is odd, then there is a $\gamma > 0$ depending only on m such that for any $(1-\gamma)n$-wise uniform distribution D over $\{0,1\}^n$, $|\Pr[D \in S_m] - |S_m|/2^n| \le 2^{-\gamma n}$.

(3) There exists a universal constant c such that for every $\varepsilon > 0$, $n \ge cm^2 \log(m/\varepsilon)$, and any $c(n/m)(1/\varepsilon)^2$-wise uniform distribution D over $\{0,1\}^n$, $\Pr[D \in S_m] \le |S_m|/2^n + \varepsilon$.

In our results the sum s of n bits $x_i \in \{0,1\}$ is constrained to be divisible by m. This setting was chosen for convenience, but our techniques apply in greater generality. For example we obtain the same results if we instead constrain s to be c mod m for any fixed c.

We also note that (2) is false for any even m because the uniform distribution on S_2 has uniformity $k = n-1$ but puts about $2/m$ mass on S_m, a set which as we shall see later (cf. Remark 1) has density about $1/m$.

Organization. Theorem 3 is a little easier to prove than Theorem 2, but uses overlapping lemmas. So we start by proving Theorem 3 in Section 2. Then in Section 3 we prove Theorem 2.

2 Proof of Theorem 3

In this section we prove Theorem 3. We start with the following theorem which will give (1) in Theorem 3 as a corollary.

Theorem 4. *Let $I \subseteq \{0,1,\ldots,n\}$ be a subset of size $|I| \le n/2$. There does not exist a $2|I|$-wise uniform distribution on $\{0,1\}^n$ that is supported on $S := \{x \in \{0,1\}^n : \sum_i x_i \in I\}$.*

Proof. Suppose there exists such a distribution D. Consider the n-variate nonzero real polynomial p defined by

$$p(x) := \prod_{i \in I}(-i + \sum_{j=1}^{n} x_j).$$

Note that $p(x) = 0$ when $x \in S$. And so $\mathrm{E}[p^2(D)] = 0$ in particular. However, since p^2 has degree at most $2|I|$, we have $\mathrm{E}[p^2(D)] = \mathrm{E}[p^2(U)] > 0$, where U is the uniform distribution over $\{0,1\}^n$, a contradiction. $\square$

Proof of (1) in Theorem 3. When I corresponds to the mod m test S_m, $|I| \le n/m + 1$. $\square$

We now move to (2) in Theorem 3. First we prove a lemma that estimates the sum $\sum_{x \in S_m} (-1)^{\sum_{i=1}^k x_i}$. Similar bounds have been established elsewhere, cf. e.g. Theorem 2.9 in [VW08], but we do not know of a reference with an explicit dependence on m, which will be used in the next section. (2) follows from bounding above the tail of the Fourier coefficients of the indicator function of S_m.

Lemma 5. *For any $1 \le k \le n-1$, $|\sum_{x \in S_m} (-1)^{\sum_{i=1}^k x_i}| \le 2^n \left(\cos \frac{\pi}{2m}\right)^n$, while for $k = 0$ $|\sum_{x \in S_m} (-1)^{\sum_{i=1}^k x_i} - 2^n/m| \le 2^n \left(\cos \frac{\pi}{2m}\right)^n$. For odd m the first bound also holds for $k = n$.*

Proof. Consider an expansion of

$$p(y) = (1-y)^k(1+y)^{n-k}$$

into 2^n terms indexed by $x \in \{0,1\}^n$ where $x_i = 0$ indicates that we take the term 1 from the i'th factor. It is easy to see that the coefficient of y^d is $\sum_{|x|=d}(-1)^{\sum_{i=1}^k x_i}$. Denote $\zeta := e^{2\pi i/m}$ as the m-th root of unity. Recall the identity

$$\frac{1}{m}\sum_{j=0}^{m-1}\zeta^{jd} = \begin{cases} 1 & \text{if } d \equiv 0 \bmod m \\ 0 & \text{otherwise.} \end{cases}$$

Thus the sum we want to bound is equal to

$$\frac{1}{m}\sum_{j=0}^{m-1} p(\zeta^j).$$

Note that $p(\zeta^0) = p(1) = 0$ for $k \neq 0$ while for $k = 0$, $p(\zeta^0) = 2^n$. For the other terms we have the following bound.

Claim 6. *For $1 \leq j \leq m-1$, $|p(\zeta^j)| \leq 2^n \left(\cos\frac{\pi}{2m}\right)^k \left(\cos\frac{\pi}{m}\right)^{n-k}$.*

Proof. As $|1+e^{i\theta}| = 2|\cos(\theta/2)|$ and $|1-e^{i\theta}| = 2|\sin(\theta/2)|$ we have

$$\begin{aligned} |p(\zeta^j)| &= |1-\zeta^j|^k|1+\zeta^j|^{n-k} \\ &= 2^n\left(\sin\frac{j\pi}{m}\right)^k\left(\cos\frac{j\pi}{m}\right)^{n-k} \\ &\leq 2^n\left(\cos\frac{\pi}{2m}\right)^k\left(\cos\frac{\pi}{m}\right)^{n-k}, \end{aligned}$$

where the last inequality holds for odd m because (1) $\sin\frac{j\pi}{m}$ is largest when $j = \frac{m-1}{2}$ or $j = \frac{m+1}{2}$, (2) $\sin(\frac{\pi}{2} - x) = \cos x$, and (3) $\cos\frac{j\pi}{m}$ is largest when $j = 1$ or $j = m-1$. For even m the term with $j = m/2$ is 0, as in this case we are assuming that $k < n$, and the bounds for odd m are valid for the other terms. □

Therefore, for $k \neq 0$ we have

$$\left|\sum_{x\in S_m}(-1)^{\sum_{i=1}^k x_i}\right| = \frac{m-1}{m}\cdot 2^n\left(\cos\frac{\pi}{2m}\right)^k\left(\cos\frac{\pi}{m}\right)^{n-k} \leq 2^n\left(\cos\frac{\pi}{2m}\right)^k\left(\cos\frac{\pi}{m}\right)^{n-k},$$

and we complete the proof using the fact that $\cos(\pi/m) \leq \cos(\pi/2m)$. For $k = 0$ we also need to include the term $p(1) = 2^n$ which divided by m gives the term $2^n/m$. □

Remark 1. *Clearly the lemma for $k = 0$ simply is the well known fact that the cardinality of S_m is very close to $2^n/m$. Equivalently, if x is uniform in $\{0,1\}^n$ then the probability that $\sum_i x_i \in S_m$ is very close to $1/m$. The same holds for the probability that $\sum_i x_i \equiv c \bmod m$ for any fixed c. This can be seen by using the polynomial $y^{-c}p(y)$ in the above proof.*

Proof of (2) in Theorem 3. Let $f : \{0,1\}^n \to \{0,1\}$ be the characteristic function of S_m. We first bound above the nonzero Fourier coefficients of f. Let $S = S_m$. By Lemma 5, we have for any β with $|\beta| = k > 0$,

$$|\hat{f}_\beta| = 2^{-n} \sum_{x \in S} (-1)^{\sum_{i=1}^k x_i} \le \left(\cos \frac{\pi}{2m}\right)^n \le 2^{-\alpha n},$$

where $\alpha = -\ln \cos(\pi/2m)$ depends only on m. Thus, if D is k-wise uniform,

$$|\mathrm{E}[f(D)] - \mathrm{E}[f(U)]| \le \sum_{|\beta|>k} |\hat{f}_\beta| \cdot |\mathrm{E}_{x\sim D}[(-1)^{\sum x_i \beta_i}]| \le \sum_{|\beta|>k} |\hat{f}_\beta| \le 2^{-\alpha n} \sum_{t=k+1}^{n} \binom{n}{t} = 2^{-\alpha n} \sum_{t=0}^{n-k-1} \binom{n}{t}.$$

For $k \ge (1-\delta)n$, we have an upper bound of $2^{n(H(\delta)-\alpha)}$. Pick δ small enough so that $H(\delta) \le \alpha/2$. The result follows by setting $\gamma := \min\{\alpha/2, \delta\}$. $\square$

Note that the above proof fails when m is even as we cannot handle the term with $|\beta| = n$. Finally, we prove (3) in Theorem 3. We use approximation theory.

Proof of (3) in Theorem 3. Let $f\colon \{0,1\}^n \to \{0,1\}$ be the characteristic function of S_m. The proof amounts to exhibiting a real polynomial p in n variables of degree $d = c(n/m)(1/\varepsilon)^2$ such that $f(x) \le p(x)$ for every $x \in \{0,1\}^n$, and $\mathrm{E}[p(U)] \le \varepsilon$ for U uniform over $\{0,1\}^n$. To see that this suffices, note that $\mathrm{E}[p(U)] = \mathrm{E}[p(D)]$ for any distribution D that is d-wise uniform. Using this and the fact that f is non-negative, we can write

$$0 \le \mathrm{E}[f(U)] \le \mathrm{E}[p(U)] \le \varepsilon \quad \text{and} \quad 0 \le \mathrm{E}[f(D)] \le \mathrm{E}[p(D)] \le \varepsilon.$$

Hence, $|\mathrm{E}[f(U)] - \mathrm{E}[f(D)]| \le \varepsilon$. This is the method of sandwiching polynomials from [Baz09].

Let us write $f = g(\sum_i x_i/n)$, for $g\colon \{0, 1/n, \dots, 1\} \to \{0,1\}$. We exhibit a univariate polynomial q of degree d such that $g(x) \le q(x)$ for every x, and the expectation of q under the binomial distribution is at most ε. The polynomial p is then $q(\sum_i x_i/n)$.

Consider the continuous, piecewise linear function $s : [-1,1] \to [0,1]$ defined as follows. The function is always 0, except at intervals of radius a/n around the inputs x where g equals 1, i.e., inputs x such that nx is divisible by m. In those intervals it goes up and down like a 'Λ', reaching the value of 1 at x. We set $a = \varepsilon m/10$.

By Jackson's theorem, see e.g. [Car, Theorem 7.4] or [Che66], for a degree $d = O(n\varepsilon^{-1}a^{-1}) = O(n\varepsilon^{-2}m^{-1})$, there exists a univariate polynomial q' of degree d that approximates s with pointwise error $\varepsilon/10$. Our polynomial q is defined as $q := q' + \varepsilon/10$.

It is clear that $g(x) \le q(x)$ for every $x \in \{0, 1/n, \dots, 1\}$. It remains to estimate $\mathrm{E}[q(U)]$.

As q' is a good approximation of s we have $\mathrm{E}[q(U)] \le 2\varepsilon/10 + \mathrm{E}[s(U)]$. We noted in Remark 1 that the remainder modulo m of $\sum x_i$ is δ-close to uniform for $\delta = \cos(\pi/2m)^n = e^{-O(n/m^2)}$. Now the function s, as a function of $\sum x_i$, is a periodic function with period m and if we feed the uniform distribution over $\{0, 1/n, \dots, m/n\}$ into s we have $\mathrm{E}[s] \le \varepsilon/10$. It follows that if n is at least a large constant times $m^2(\log(1/\varepsilon) + \log m)$, we have $\mathrm{E}[s(U)] \le 2\varepsilon/10$ and we conclude that $\mathrm{E}[q(U)] \le 4\varepsilon/10$. $\square$

3 Proof of Theorem 2

In this section we prove Theorem 2. Let I be a subset of $\{0, 1, \ldots, n-1, n\}$ and $S \subseteq \{0, 1\}^n$ be the subset of strings whose sum $\sum_i x_i$ belongs to I. Let U_S be the uniform distribution over S. We are going to construct a k-wise uniform distribution starting from U_S and changing the weights of $k+1$ slices of the Hamming cube. In particular, our distribution will be symmetric. We note that since S is symmetric, if there is a k-wise uniform distribution supported on it then by a simple symmetrization argument there must also be a symmetric one.

Let ε_t be the bias of a parity of size t under U_S, i.e., $\varepsilon_t := E_{x \in U_S}[(-1)^{\sum_{i=1}^t x_i}]$. Note that because we are working with symmetric distributions, all parities of the same size have the same bias. Now let $\varepsilon(t, \ell)$ be the bias of a parity of size t over the uniform distribution on strings that sum to ℓ. Note that $\varepsilon(t, \ell)$ is a scaled version of the Kravchuk polynomial of degree t in the variable ℓ.

We note that $\varepsilon_t = \sum_{\ell \in I} \Pr_{x \sim U_S}[\sum_j x_j = \ell] \cdot \varepsilon(t, \ell)$.

Now let $a_0 < a_1 < \cdots < a_k$ be $k+1$ points in I that are closest to $n/2$ and let i^* be an index that maximizes $|a_i - \frac{n}{2}|$. Finally let p_i be the probability over x drawn from U_S that x sums to a_i.

We are going to change the p_i to $p_i - \Delta_i$ with the goal of making ε_t zero for every $1 \le t \le k$. The effect of the substitution on ε_t is to decrease it by $\sum_{0 \le i \le k} \Delta_i \varepsilon(t, a_i)$.

Thus our goal is to find Δ_i's so that

$$\sum_{i=0}^{k} \Delta_i \varepsilon(t, a_i) = \varepsilon_t, \quad \forall t \in \{1, 2, \ldots, k\}$$

$$\sum_{i=0}^{k} \Delta_i = 0,$$

$$0 \le p_i - \Delta_i \le 1, \quad \forall i \in \{0, \ldots, k\}.$$

Let M be the $(k+1) \times (k+1)$ matrix $M_{t,i} := \varepsilon(t, a_i)$ where $t, i \in \{0, \ldots, k\}$. Let $\Delta := (\Delta_0, \ldots, \Delta_k)^T$ and $b := (0, \varepsilon_1, \ldots, \varepsilon_k)^T$. Then the first two conditions form the linear system

$$M\Delta = b.$$

We will show that there is a unique solution Δ to this system.

To satisfy the third condition, note that p_{i^*} is the smallest among all the p_i's. It will also be the case that $p_{i^*} \le 1/2$. Thus if $\|\Delta\|_\infty \le p_{i^*}$ we will also satisfy the third condition and have a k-wise uniform distribution supported on S.

Consider the expression $n^{-t}(\sum_{j=1}^{n}(-1)^{x_j})^t$. If we expand this, cancel factors that appear twice, and collect terms, we can rewrite it as

$$n^{-t}(\sum_{j=1}^{n}(-1)^{x_j})^t = \sum_{r=0}^{t} \gamma_{t,r} \binom{n}{r}^{-1} \sum_{|\beta|=r} (-1)^{\sum x_i \beta_i},$$

for some choice of non-negative values $\gamma_{t,r}$, which by plugging in $x_1 = x_2 = \ldots = x_n = 0$ can be seen to satisfy $\sum_{r=0}^{t} \gamma_{t,r} = 1$.

Let $\alpha_i := (n-2a_i)/n$. Taking expectation in the above equation over all the x's with sum equal to a_i we have for every $i \in \{0,1,\dots,k\}$,

$$\alpha_i^t = ((n-2a_i)/n)^t = \sum_{r=0}^{t} \gamma_{t,r} \binom{n}{r}^{-1} \sum_{|\beta|=r} \mathrm{E}[(-1)^{\sum x_i \beta_i}] = \sum_{r=0}^{t} \gamma_{t,r} \varepsilon(r, a_i). \qquad \text{(A)}$$

Let M_r be the r-th row of M. We construct a new matrix V from M by applying the following row operations R to M: For every t, set $V_t = \sum_{r=0}^{t} \gamma_{t,r} M_r$. It follows from equation (A) that $V_{t,i} = \alpha_i^t$, and so $V = RM$ is a Vandermonde matrix, which is invertible. Hence,

$$\Delta = V^{-1} R b$$

is a unique solution.

Therefore it suffices to show that $\|\Delta\|_\infty \le p_{i^*}$. Note that $\|\Delta\|_\infty \le \|V^{-1}\|_\infty \|Rb\|_\infty$, where the ∞ norm of a matrix is the maximum sum of the absolute values along any one row.

Moreover, since $(Rb)_t = \sum_{r=0}^{t} \gamma_{t,r} b_r$ and $\sum_{r=0}^{t} \gamma_{t,r} = 1$, we have $\|Rb\|_\infty \le \|b\|_\infty$. Hence, it suffices to bound above $\|V^{-1}\|_\infty$ and $\|b\|_\infty$.

Roadmap for the following claims. To get an idea of the following claims, consider the case $m = 3$ and $k = o(n)$. We first show in Claim 7 that $\|V^{-1}\|_\infty \le 2^{o(n)}$. Then we find it convenient to bound $\|b\|_\infty$ and p_{i^*} multiplied by $|S|$. We show that $|S| p_{i^*} \ge 2^{n(1-o(1))}$ in Claim 8. We note that Claims 7, 8 and 9 hold for any symmetric subset S. Finally, in Claim 10 we use the definition of S to obtain bounds on a_{i^*} and b, and show that $|S| \|b\|_\infty \le (2-\Omega(1))^n$. Altogether,

$$\|V^{-1}\|_\infty |S| \|b\|_\infty \le 2^{o(n)} (2-\Omega(1))^n \le 2^{n(1-\Omega(1))} \le |S| p_{i^*},$$

as desired.

Claim 7. $\|V^{-1}\|_\infty \le (k+1)(\frac{4en}{k})^k$.

Proof. Since V is a Vandermonde matrix, we can specify the entries of its inverse explicitly. As shown in e.g. [Tur66] we have

$$V^{-1}_{i,k-j} = (-1)^{k-j} \left(\sum_{\substack{|\beta|=j \\ i \notin \beta}} \alpha^\beta \right) \cdot \left(\prod_{s \neq i} (\alpha_s - \alpha_i)^{-1} \right).$$

We now give an upper bound on each of the factors on the R.H.S.

Bounding $\sum_{|\beta|=j, i\notin\beta} \alpha^\beta$: Since $|\alpha_i| \le 1$, this is bounded by the number of terms, $\binom{k}{j}$, and hence by 2^k.

Bounding $\prod_{s\neq i}(\alpha_s-\alpha_i)^{-1}$: Since the difference between every pair of distinct a_i, a_j is at least 1, we have

$$\prod_{s\neq i}(a_s-a_i)\geq (k/2)!^2$$

when k is even and is at least $(\frac{k+1}{2})(\frac{k-1}{2})!^2$ when k is odd. By a crude form of Stirling's formula, $n!\geq (n/e)^n$, and so we get the lower bound $(k/2e)^k$ in either case. Hence,

$$\prod_{s\neq i}(\alpha_s-\alpha_i)^{-1}\leq n^k\prod_{s\neq i}(a_s-a_i)^{-1}\leq (\frac{2en}{k})^k.$$

Putting the bounds together, we have

$$\|V^{-1}\|_\infty\leq (k+1)\max_{i,j}|V^{-1}_{i,j}|\leq (k+1)(\frac{4en}{k})^k.$$

□

Now we give a lower bound on p_{i^*}.

Claim 8. $p_{i^*}|S|\geq \frac{2^{n\left(1-\alpha_{i^*}^2\right)}}{n+1}$.

Proof. Using the inequalities $\binom{n}{i}\geq\frac{2^{nH(i/n)}}{n+1}$ and $H(\frac{1-\varepsilon}{2})\geq 1-\varepsilon^2$, we have

$$p_{i^*}|S|=\binom{n}{a_{i^*}}\geq\frac{2^{nH\left(\frac{1-\alpha_{i^*}}{2}\right)}}{n+1}\geq\frac{2^{n\left(1-\alpha_{i^*}^2\right)}}{n+1}.$$

□

Therefore,

$$\frac{p_{i^*}|S|}{\|V^{-1}\|_\infty}\geq\frac{2^{n\left(1-\alpha_{i^*}^2\right)}}{(n+1)(k+1)(\frac{4en}{k})^k}\geq e^{nf(k,n,a_{i^*})},$$

where

$$f(k,n,a_{i^*}):=\ln 2\cdot\left(1-\alpha_{i^*}^2\right)-\frac{k}{n}\left(\ln\frac{4en}{k}\right)-o(1).$$

We conclude with the following claim.

Claim 9. *If $e^{nf(k,n,a_{i^*})}\geq\max_{1\leq t\leq k}\sum_{x\in S}(-1)^{\sum_{i=1}^t x_i}$, then there exists a k-wise uniform distribution supported on S.*

Proof. We just showed

$$\frac{p_{i^*}|S|}{\|V^{-1}\|_\infty}\geq e^{nf(k,n,a_{i^*})}\geq\max_{1\leq t\leq k}\sum_{x\in S}(-1)^{\sum_{i=1}^t x_i}=\|b\|_\infty|S|.$$

Hence, $\|\Delta\|_\infty\leq\|V^{-1}\|_\infty\|b\|_\infty\leq p_{i^*}$.

□

3.1 Zero modulo m

We have that S_m consists of all strings with $\sum x_i \equiv 0 \bmod m$. If follows that $|\alpha_{i^*}| \leq (k+1)m/2n$. We now give an upper bound on $\|b\|_\infty |S|$.

Claim 10. *$\|b\|_\infty |S| \leq e^{ng(n,m)}$, where $g(n,m) := \ln 2 - \frac{1}{2}\left(\frac{\pi}{2m}\right)^2$.*

Proof. Note that $\|b\|_\infty |S| = \sum_{x \in S} (-1)^{\sum_{i=1}^k x_i}$. By Lemma 5,

$$\sum_{x \in S} (-1)^{\sum_{i=1}^k x_i} \leq 2^n \left(\cos \frac{\pi}{2m}\right)^n \leq e^{ng(n,m)},$$

where in the last two inequalities we used the fact that $\ln \cos(x) \leq -\frac{x^2}{2}$ for $x \in [0, \pi/2)$. □

We are now ready to prove Theorem 2.

Proof of Theorem 2. Recall that $|\alpha_{i^*}| \leq (k+1)m/2n$. By Claim 10 and Claim 9, it suffices to show that $f(k,n,a_{i^*}) - g(n,m)$ is positive, where recall

$$\begin{aligned} f(k,n,a_i^*) &= \ln 2 \cdot \left(1 - \alpha_{i^*}^2\right) - \frac{k}{n}\left(\ln \frac{4en}{k}\right) - o(1) \\ &\geq \ln 2 \cdot \left(1 - (\frac{(k+1)m}{2n})^2\right) - \frac{k}{n}\left(\ln \frac{4en}{k}\right) - o(1) \end{aligned}$$

and

$$g(n,m) := \ln 2 - \frac{1}{2}\left(\frac{\pi}{2m}\right)^2.$$

Indeed, we have

$$f(k,n,a_{i^*}) - g(n,m) \geq \frac{1}{2}\left(\frac{\pi}{2m}\right)^2 - \frac{k}{n}\left(\ln \frac{4en}{k}\right) - \ln 2 \cdot \left(\frac{(k+1)m}{2n}\right)^2 - o(1),$$

and choosing $k = \frac{\varepsilon n}{m^2 \ln m}$ for a sufficiently small ε makes this quantity positive. □

References

[Baz09] Louay M. J. Bazzi. Polylogarithmic independence can fool DNF formulas. *SIAM J. Comput.*, 38(6):2220–2272, 2009.

[Bra10] Mark Braverman. Polylogarithmic independence fools AC^0 circuits. *J. of the ACM*, 57(5), 2010.

[Car] Neal Carothers. A short course on approximation theory. Available at http://personal.bgsu.edu/~carother/Approx.html.

[Che66] E. Cheney. *Introduction to approximation theory.* McGraw-Hill, New York, New York, 1966.

[CRS00] Suresh Chari, Pankaj Rohatgi, and Aravind Srinivasan. Improved algorithms via approximations of probability distributions. *J. Comput. System Sci.*, 61(1):81–107, 2000.

[DGJ+10] Ilias Diakonikolas, Parikshit Gopalan, Ragesh Jaiswal, Rocco A. Servedio, and Emanuele Viola. Bounded independence fools halfspaces. *SIAM J. on Computing*, 39(8):3441–3462, 2010.

[DKN10] Ilias Diakonikolas, Daniel Kane, and Jelani Nelson. Bounded independence fools degree-2 threshold functions. In *51th IEEE Symp. on Foundations of Computer Science (FOCS)*. IEEE, 2010.

[EGL+98] Guy Even, Oded Goldreich, Michael Luby, Noam Nisan, and Boban Velickovic. Efficient approximation of product distributions. *Random Struct. Algorithms*, 13(1):1–16, 1998.

[GOWZ10] Parikshit Gopalan, Ryan O'Donnell, Yi Wu, and David Zuckerman. Fooling functions of halfspaces under product distributions. In *25th IEEE Conf. on Computational Complexity (CCC)*, pages 223–234. IEEE, 2010.

[LV15] Chin Ho Lee and Emanuele Viola. Some limitations of the sum of small-bias distributions. Available at http://www.ccs.neu.edu/home/viola/, 2015.

[MZ09] Raghu Meka and David Zuckerman. Small-bias spaces for group products. In *13th Workshop on Randomization and Computation (RANDOM)*, volume 5687 of *Lecture Notes in Computer Science*, pages 658–672. Springer, 2009.

[Raz09] Alexander A. Razborov. A simple proof of Bazzi's theorem. *ACM Transactions on Computation Theory (TOCT)*, 1(1), 2009.

[Tal14] Avishay Tal. Tight bounds on The Fourier Spectrum of AC^0. *Electronic Colloquium on Computational Complexity*, Technical Report TR14-174, 2014. www.eccc.uni-trier.de/.

[Tur66] L. Richard Turner. Inverse of the Vandermonde matrix with applications. 1966. NASA technical note D-3547 available at http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19660023042.pdf.

[VW08] Emanuele Viola and Avi Wigderson. Norms, XOR lemmas, and lower bounds for GF(2) polynomials and multiparty protocols. *Theory of Computing*, 4:137–168, 2008.