ε-discrepancy sets and their applications for interpolation of sparse polynomials

Noga Alon[*] Yishay Mansour[†]

February 22, 2002

Abstract

We present a simple explicit construction of a probability distribution supported on $(p-1)^2$ vectors in Z_p^n, where $p \geq n/\varepsilon$ is a prime, for which the absolute value of each nontrivial Fourier coefficients is bounded by ε. This construction is used to derandomize the algorithm of [Man92] that interpolates a sparse polynomial in polynomial time in the bit complexity model.

1 Introduction

Given a set $A \subset Z_p^n$, for each $\alpha \in Z_p^n$ define

$$DISC_A(\alpha) = \frac{1}{|A|} \left| \sum_{z \in A} \omega^{<\alpha,z>} \right|,$$

where ω is the pth root of unity over the complex numbers, i.e. $\omega = e^{2\pi i/p}$.

Definition 1 *A set $A \subset Z_p^n$ is an ε discrepancy set if for any $\alpha \neq \vec{0}$, $DISC_A(\alpha) \leq \varepsilon$.*

In this note we present a simple explicit construction as follows.

Theorem 1.1 *For any prime p and any $n > 1$ there exists an explicit set $A_p^n \subset Z_p^n$, such that $|A_p^n| = (p-1)^2$ and A_p^n is an $\frac{n-1}{p-1}$ discrepancy set.*

The construction is a *mod p* variant of one of the binary constructions presented in [AGHP90]. Another construction with related properties appears in [AMN90]. The main advantage of the present construction is its simplicity and the elementary proof of its properties.

[*]Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel-Aviv University, Tel-Aviv, Israel. Research supported in part by a USA-Israeli BSF grant and by the Fund for Basic Research administered by the Israel Academy of Sciences.

[†]Department of Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel-Aviv University, Tel-Aviv, Israel. Research supported in part by a grant of the Israeli Ministry of Science and Technology and by The Israel Science Foundation administered by The Israel Academy of Science and Humanities.

Our main application for ε discrepancy sets is the derandomization of the interpolation algorithm of [Man92]. Using an ε discrepancy set we can test whether a sparse multivariate polynomial is identically zero, which is a major task in any multivariate interpolation algorithm. Other possible applications are mentioned as well.

Other previous works on sparse multi-variate polynomial interpolation include the work of Zippel [Zip79], which gives a probabilistic algorithm, that of Grigoriev and Karpinski [GK87], for interpolation of a sparse permanent, and the work of Ben-Or and Tiwari [BOT88].

Our construction of ε-discrepancy sets can be viewed as a real-value analog of the "ε-bias" distribution [NN90, AGHP90], which is defined over boolean variables and guarantees that the absolute value of each of its nontrivial Fourier coefficients is bounded by ε.

2 Construction of an ε discrepancy set

Let p be a prime and let Z_p^* denote the multiplicative group of the finite field Z_p. For $x, y \in Z_p^*$ put $v_{x,y} = (y, yx, yx^2, \ldots, yx^{n-1})$. Define $A_p^n = \{v_{x,y} \mid x, y \in Z_p^*\}$. Note that the size of the set A_p^n is $(p-1)^2$.

For $a \in Z_p$ and $\alpha \in Z_p^n$ define $n_{a,\alpha}$ by

$$n_{a,\alpha} = |\{v_{x,y} \mid x, y \in Z_p^*, < v_{x,y}, \alpha >= a\}|$$

Claim 2.1 *Let $\alpha \neq \vec{0}$. If $a, b \neq 0$ then $n_{a,\alpha} = n_{b,\alpha}$. In addition, $n_{0,\alpha} \leq (n-1)(p-1)$*

Proof: Consider the inner product,

$$< v_{x,y}, \alpha >= \sum_{i=0}^{n-1} yx^i \alpha_i = yP_\alpha(x)$$

where $P_\alpha(x)$ is the polynomial with $\vec{\alpha}$ as the vector of its coefficients, i.e. $P_\alpha(x) = \sum_i \alpha_i x^i$. We are interested in the number of solutions x, y of the equation

$$yP_\alpha(x) = a.$$

Fix $x \in Z_p^*$. Clearly, if $P_\alpha(x) \neq 0$ then for each $y \in Z_p^*$ there is a different nonzero value to $yP_\alpha(x)$. Hence, each value in Z_p^* is generated by a unique y (in this case). On the other hand if $P_\alpha(x) = 0$, then $< v_{x,y}, \alpha >= 0$ for every $y \in Z_p^*$. Since $P_\alpha(x) = 0$ for at most $n-1$ different $x \in Z_p^*$, we conclude that $n_{0,\alpha} \leq (n-1)(p-1)$. □

Theorem 2.2 *A_p^n is an $\frac{n-1}{p-1}$ discrepancy set.*

Proof: By the construction of A_p^n,

$$DISC_{A_p^n}(\alpha) = \frac{1}{(p-1)^2} \left| \sum_{x,y \in Z_p} \omega^{<\alpha, v_{x,y}>} \right|.$$

We can rewrite this as,

$$DISC_{A_p^n}(\alpha) = \frac{1}{(p-1)^2}\left|\sum_{a\in Z_p} n_{a,\alpha}\omega^a\right|.$$

Recall that $\sum_{i=0}^{p-1}\omega^i = 0$. Since for $a \neq 0$, $n_{a,\alpha} = k$ is the same, the only non-zero contribution is from $a = 0$, showing that $\sum_{a\in Z_p} n_{a,\alpha}\omega^a = n_{0,\alpha} - k$. Since $n_{0,\alpha} \leq (n-1)(p-1)$ and $0 \leq k \leq p-1$, we have that,

$$DISC_{A_p^n}(\alpha) \leq \frac{n-1}{p-1},$$

completing the proof of the theorem. □

3 Applications

3.1 Interpolation of Multivariate Polynomials

Let $P(x_1,\ldots,x_n) = \sum_{i=0}^{t} c_i x_1^{e_{i,1}}\cdots x_n^{e_{i,n}}$ be a multivariate polynomial with t integer coefficients. Let $L_1(P)$ denote the sum of the absolute values of the coefficients of P, i.e. $L_1(P) = \sum_{i=0}^{t}|c_i|$.

Lemma 3.1 *Let A be an ε discrepancy set, and $P(x_1,\ldots,x_n) = c_0 + \sum_{i=1}^{t} c_i x_1^{e_{i,1}}\cdots x_n^{e_{i,n}}$. Then,*

$$|E_{(z_1,\ldots,z_n)\in A}[P(\omega^{z_1},\ldots,\omega^{z_n})] - c_0| \leq \varepsilon L_1(P)$$

where E is the expectation over the uniform distribution of vectors from A.

Proof: Let $\vec{e_i} = (e_{i,1},\ldots,e_{i,n})$. By the linearity of expectation,

$$E_{\vec{z}=(z_1,\ldots,z_n)\in A}[P(\omega^{z_1},\ldots,\omega^{z_n})] = c_0 + \sum_{i=1}^{t} c_i E[\omega^{<\vec{z},\vec{e_i}>}].$$

Since A is an ε discrepancy set and $\vec{e_i} \neq \vec{0}$,

$$|E[\omega^{<\vec{z},\vec{e_i}>}]| \leq \varepsilon,$$

and the assertion of the lemma follows. □

By the same argument one can show the following.

Claim 3.2 *Let A be an ε discrepancy set, and $P(x_1,\ldots,x_n) = \sum_{i=0}^{t} c_i x_1^{e_{i,1}}\cdots x_n^{e_{i,n}}$. Then,*

$$\left|E_{(z_1,\ldots,z_n)\in A}\left[\|P(\omega^{z_1},\ldots,\omega^{z_n})\|^2\right] - \sum_{i=0}^{t} c_i^2\right| \leq \varepsilon L_1^2(P)$$

where E is the expectation over the uniform distribution of vectors from A.

The above claim gives an immediate tool to test if a sparse multivariate polynomial is zero (assuming that its coefficients are integers and bounded). Since the coefficients are integers, then either $\sum_{i=0}^{t} c_i^2$ is at least one or it is zero. By choosing $p > 2nL_1^2(P)$ we guarantee that the error is less than 1/2, and therefore, by the above claim, we can distinguish between the two cases.

We next demonstrate the derandomization on the algorithm of [Man92]. The idea, as in [Zip79], is to interpolate the variables one by one. Since we have an upper bound, say t, on the number of non-zero coefficients, there would be at most t terms to consider. The assumption here is that we have a black box that outputs the value of $P(x_1, \dots, x_n)$ for any desired $(x_1, \dots, x_n)$, and our objective is to determine the coefficients of P. From the analysis it follows that this is possible even if our black box only outputs (sufficiently accurate) approximations of the values of P.

Initially, we can rewrite P as,

$$P(x_1, \dots, x_n) = \sum_{j=0}^{d} x_1^j P_j(x_2, \dots, x_n).$$

We are interested in determining which of the P_j's are not the zero polynomial. To perform this we note that, for a prime $p > d$,

$$P_j(x_2, \dots, x_n) = \frac{1}{p} \sum_{k=0}^{p-1} P(\omega^k, x_2, \dots, x_n) \omega^{-kj}.$$

For each $\vec{z} = (z_2, \dots, z_n) \in A_p^{n-1}$ we can compute $P_j(z_2, \dots, z_n)$ by using the above identity, and then compute $E_z[\|P_j(z)\|^2]$.

In general we define $P_{e_1,\dots,e_k}$ as follows,

$$P(x_1, \dots, x_n) = \sum_{e_1=0}^{d} \cdots \sum_{e_k=0}^{d} x_1^{e_1} \cdots x_k^{e_k} P_{e_1,\dots,e_k}(x_{k+1}, \dots, x_n),$$

i.e. $P_{e_1,\dots,e_k}(x_{k+1}, \dots, x_n)$ has all the terms that include $x_1^{e_1} \cdots x_k^{e_k}$. By the properties of the discrete Fourier transform we have that,

$$P_{e_1,\dots,e_k}(x_{k+1}, \dots, x_n) = \frac{1}{p^k} \sum_{j_1=0}^{p-1} \cdots \sum_{j_k=0}^{p-1} P(\omega^{j_1}, \dots, \omega^{j_k}, x_{k+1}, \dots, x_n) \omega^{-e_1 j_1} \cdots \omega^{-e_k j_k}.$$

In order to test whether $P_{e_1,\dots e_k} \not\equiv 0$, we estimate its norm by computing,

$$E_{(z_{k+1},\dots,z_n) \in A_p^{n-k}} \left[\left\| E_{(z_1,\dots,z_k) \in A_p^k} [P(\omega^{z_1}, \dots, \omega^{z_n}) \omega^{-\sum_{j=1}^{k} e_j z_j}] \right\|^2 \right].$$

The interpolation works in phases. At the kth phase we determine all the vectors $(e_1, \dots e_k)$, such that $P_{e_1,\dots e_k} \not\equiv 0$, given all the the vectors $(e_1, \dots e_{k-1})$, such that $P_{e_1,\dots e_{k-1}} \not\equiv 0$. Since the polynomial P has only t non-zero coefficients, at any phase we need to maintain at most t vectors. At the end we have all the terms, i.e. $\vec{e}_i$, and need only to determine the coefficients.

3.2 Univariate polynomials

In [Kat89, AIK$^+$90] and, more explicitly, in [RSW93] it is shown how to construct explicitly a set $B \subset Z_p$, such that $|B| = O((\log p/\varepsilon)^c)$, for some constant c, and such that for any $\alpha \neq 0, \alpha \in Z_p$,

$$\frac{1}{|B|}\left|\sum_{z\in B}\omega^{\alpha z}\right| \leq \varepsilon.$$

We can use the set B to interpolate any sparse univariate polynomial of (high) degree d ($\leq p$). Recall that if $P(x) = \sum_{i=0}^{p} a_i x^i$ then $a_k = (1/p)\sum_{j=0}^{p} P(\omega^j)\omega^{-kj}$. Hence, averaging over B would add an additive error of at most $\varepsilon L_1(P)$ to any coefficient.

Using such constructions we can reduce the size of A_p^n when $\varepsilon >> 1/p^{1/c}$ to $O(\frac{n}{\varepsilon}(\log p/\varepsilon)^c)$. To do so, simply modify the construction above by letting x vary over an arbitrary subset of cardinality n/ε of Z_p and by letting y vary over a subset $B \subset Z_p$, that has the above properties. It is easy to check that the discussion in Section 2 implies that the modified set is a 2ε-discrepancy set in Z_p^n. By Proposition 7' in [AR94], for $\epsilon > p^{-n/2}$ the size of any ϵ discrepancy set for Z_p^n is at least $\Omega(\frac{n \log p}{\epsilon^2 \log(n \log p/\epsilon^2)})$ showing that the last construction is not far from the optimum.

3.3 Axis Parallel boxes

The sets A_p^n can be used to approximate the expectation of any function P with a small value of $L_1(P)$. As an illustration, consider the function $f_a(x) = 1$ if $x < a$ and $f_a(x) = 0$ otherwise, where $x \in Z_p$. For the intersection of k such functions, i.e. $F(\vec{x}) = \prod_{j=1}^{k} f_{a_j,i_j}(\vec{x})$, one can show that $L_1(F) = O(\log^k p)$, and hence the set A_p^n can be used to approximate the expectation of F (which is the fraction of the volume of the corresponding box in Z_p^n) within an additive error of $O(\frac{n \log^k p}{p})$. In fact, by replacing F by a smooth function that approximates it this error term can be improved to $O(\frac{n \log^k (p/n)}{p})$. Since for this example this is a weaker estimate than those obtained by the constructions in [EGL$^+$92] and [CRS94] we omit the details.

4 Conclusion and Open questions

We showed that the set A_p^n is an $\frac{n-1}{p-1}$ discrepancy set, and $|A_p^n| = O(p^2)$. The modified construction described in Subsection 3.2 provides an ϵ-discrepancy set of size polynomial in $\log p/\varepsilon$ and linear in n.

For the interpolation problem we need that p is larger than the degree of the polynomial in each variable. Therefore, the modified set can be useful here. It is not difficult to check (see Proposition 6' in [AR94]) that a random set of $\Theta(\frac{n}{\varepsilon^2}\log p)$ vectors from Z_p^n would almost surely be an ε discrepancy set, and as mentioned above this is nearly best possible. However the problem of finding an explicit construction of such a small ε-discrepancy set remains open.

Acknowledgment We would like to thank Avi Wigderson for helpful comments and suggestions.

References

[AGHP90] Noga Alon, Oded Goldreich, Johan Hastad, and Rene Peralta. Simple constructions of almost k-wise independent random variables. In 31^{th} *Annual Symposium on Foundations* *of Computer Science, St. Louis, Missouri*, pages 544–553, October 1990. Also: Random Structures and Algorithms 3 (1992), 289-304.

[AIK+90] M. Ajtai, H. Iwaniec, J. Komlós, J. Pintz, and E. Szemerédi. Construction of a thin set with small fourier coefficients. *Bull. London Math. Soc.*, 22:583–590, 1990.

[AMN90] Y. Azar, R. Motwani, and J. Naor. Approximating arbitrary probability distributions using small sample spaces. Manuscript, 1990.

[AR94] N. Alon and Y. Roichman. Random cayley graphs and expanders. *Random Structures and Algorithms*, 5:271–284, 1994.

[BOT88] M. Ben-Or and P. Tiwari. A deterministic algorithm for sparse multivariate polynomial interpolation. In *Proceedings of the* 20^{th} *Annual ACM Symposium on Theory of Computing*, pages 301–309, May 1988.

[CRS94] Suresh Chari, Pankaj Rohatgi, and Aravind Srinivasan. Improved algorithms via approximations of probability distributions. In *Proceedings of the* 26^{th} *Annual ACM Symposium on Theory of Computing*, pages 584-592, May 1994.

[EGL+92] G. Even, O. Goldreich, M. Luby, N. Nisan, and B. Velickovic. Approximations of general independent distributions. In *Proceedings of the* 24^{th} *Annual ACM Symposium on Theory of Computing*, pages 10–16, May 1992.

[GK87] D. Y. Grigoriev and M. Karpinski. The matching problem for bipartite graphs with polynomially bounded permanents is in NC. In 28^{th} *Annual Symposium on Foundations of Computer Science, Los Angeles, California*, pages 166–172, October 1987.

[Kat89] N. M. Katz. An estimate for character sums. *J. AMS* 2:197–200, 1989.

[Man92] Yishay Mansour. Randomized approxmation and interpolation of sparse polynomials. In *ICALP*, pages 261-272, July 1992.

[NN90] Joseph Naor and Moni Naor. Small bias probability spaces: efficient construction and applications. In *Proceedings of the* 22^{nd} *Annual ACM Symposium on Theory of Computing, Baltimore, Maryland*, pages 213–223, May 1990.

[RSW93] A. Razborov, E. Szemerédi, and A. Wigderson. Constructing small sets that are uniform in arithmetic progressions. *Combinatorics, Probability and Computing* 2:513–518, 1993.

[Zip79] R. E. Zippel. Probabilistic algorithms for sparse polynomials. In *EUROSAM*, pages 216–226. Springer Lecture notes in computer science, vol. 72, 1979.