Lecture 6: Bounding ϵ-Biased Sets

Topics in Pseudorandomness and Complexity Theory (Spring 2018)
Rutgers University
Swastik Kopparty
Scribes: Heman Gandhi, Steve Hsu

1 Review: ϵ-Biased Sets

Definition 1. *A collection* $(x_1, \ldots, x_n) \in \{0,1\}^n$ *of random variables is called* ϵ*-biased if* $\forall y \in \{0,1\}^n, \Pr[\langle y, x\rangle = 0] \in [\frac{1-\epsilon}{2}, \frac{1+\epsilon}{2}]$.

Fact 2. *A 0-biased distribution is uniform.*

This fact will be proven later on.

Theorem 3. *There are* ϵ*-biased distributions on* $\{0,1\}^n$ *which are uniform over a multiset of size* $O(\frac{n}{\epsilon^2})$.

2 Constructing ϵ-Biased Sets

The construction of ϵ-Biased sets has been consistently improving. A simple $O(\frac{n^2}{\epsilon^2})$ will be shown here. However, the following are known:

1. Before 2013: $O(\frac{n^2}{\epsilon^2})$ or $O(\frac{n}{\epsilon^3})$.
2. 2013, by Ben-Aroya and Ta-Shma: $O(\frac{n}{\epsilon^2}^{1.25})$.
3. 2017, by Ta-Shma: $O(\frac{n}{\epsilon^{2+o(1)}})$.

We do now know if there is a construction of an ϵ-biased distribution of size $O(\frac{n}{\epsilon^2})$ though we know they exist.

We now present the construction of an ϵ-biased distribution of size $O(\frac{n^2}{\epsilon^2})$ by Alon-Goldreich-Halsted-Peralta.

2.1 Aside: $\mathbb{F}_2^n$ and $\mathbb{F}_{2^n}$

For a prime p, we can consider $\mathbb{Z}/p\mathbb{Z}$. This is the construction of $\mathbb{F}_p$: with operations modulo p.

Similarly, we can take $\mathbb{F}_2[x]$, the polynomials with coefficients in $\mathbb{F}_2$. Here, let q be an irreducible polynomial (one that does not factor) and have degree n. $q(x)\mathbb{F}_2[x] \subset \mathbb{F}_2[x]$ is an ideal and we can define $\mathbb{F}_{2^n} = \mathbb{F}_2[x]/q(x)\mathbb{F}_2[x]$. This looks like $\mathbb{F}_{2^n} = \{a(x) \in \mathbb{F}_2[x] : \deg(a) < n\}$.

As an additive group, $\mathbb{F}_{2^n} \cong \mathbb{F}_2^n$ as polynomial addition can be viewed as component-wise addition. In fact, we define $\phi : \mathbb{F}_{2^n} \to \mathbb{F}_2^n$ to be this linear map in $\mathbb{F}_2$.

Multiplication is not so nice, though a formula depending on q exists. This is an invertible operation thanks to the GCD algorithm and the division theorem and their applicability in $\mathbb{F}_2[x]$.

2.2 The Construction

Let $S = \{(\langle\phi(\alpha), z\rangle, \langle\phi(\alpha^2), z\rangle, \ldots, \langle\phi(\alpha^n), z\rangle) \forall \alpha \in \mathbb{F}_{2^m}, z \in F_2^m\}$

$|S| = 2^{2m}$.

We now prove that S is ϵ-biased.

Proof. What we want is $\Pr_{x\in S}[\langle x, y\rangle = 0]$ being close to 1/2 for $y \neq 0$. The following equivalences give us all we need:

$$\Pr_{x\in S}[\langle x, y\rangle = 0] = \Pr_{\alpha\in\mathbb{F}_{2^m}, z\in\mathbb{F}_2^m}[\sum_{i=1}^{n} \langle\phi(\alpha^i), z\rangle y_i = 0] = \Pr_{\alpha,z}[\langle\phi(\sum_{i=1}^{n} y_i\alpha^i), z\rangle = 0]$$

We define $p_y(t) = \sum_{i=1}^{n} y_i t^i$, a polynomial that we evaluate at α. In addition, we condition on whether $\phi(p_y(\alpha)) = 0$. This is relevant since $y \neq 0$ so $\phi(p_y(\alpha))$ is not zero everywhere.

$$\Pr_{\alpha,z}[\langle\phi(\sum_{i=1}^{n} y_i\alpha^i), z\rangle = 0] = \Pr_{\alpha,z}[\langle\phi(p_y(\alpha)), z\rangle = 0] = \Pr_{\alpha}[\phi(p_y(\alpha)) = 0] * 1 + \Pr_{\alpha}[\phi(p_y(\alpha)) = 0] * \frac{1}{2}$$

We can bound $p = \Pr_\alpha[\phi(p_y(\alpha)) = 0] \leq \frac{n}{2^n}$ since a p_y has at most n zeros out of the 2^n possible values. Then, $\Pr_{\alpha,z}[\langle\phi(p_y(\alpha)), z\rangle = 0] = p + \frac{1-p}{2} = \frac{1+p}{2}$. $\frac{1+p}{2} \in [\frac{1}{2}, 1 + \frac{n}{2^m}]$.

We can choose $m = \log(\frac{n}{\epsilon})$ so that S is ϵ-biased of size $O(\frac{n^2}{\epsilon^2})$.

□

We note that S can contain duplicates so it better thought of as a multi-set. For $z = 0$, regardless of α, we get the 0 vector in $\mathbb{F}_2^n$ as our entry in S.

We also note that there are deterministic ways to get $q \in \mathbb{F}_2[x]$ in poly-n time.

3 ϵ-Biased Sets and Expander Graphs

Let $S \subset \mathbb{F}_2^n$ be an ϵ-biased set. We consider the Cayley graph.

Definition 4. *The Cayley graph of a group G and some $T \subset G$, is called $Cay(G; T)$ and defined to be a graph with vertices in G and edge set $\{(a, a + x) : a \in G, x \in S\}$.*

Fact 5. *We have the following clear facts about Cayley graphs:*

- *The graph is $|S|$-regular. This is possibly around $O(\frac{n}{\epsilon^2})$.*
- *Over $\mathbb{F}_2^n$, this is undirected since $x = x^{-1}$.*
- *2^n vertices when formed over $\mathbb{F}_2^n$*

We now put $G = \text{Cay}(\mathbb{F}_2^n; S)$ and consider its eigenvalues. In doing so, we define the additive character.

Definition 6. *Let $y \in \mathbb{F}_{2^n}$. Let $\psi_y : \mathbb{F}_{2^n} \to \mathbb{R}$ with $\psi_y(a) = (-1)^{\langle y,a \rangle}$. ψ is called an additive character and we know that $\psi_y(a+b) = \psi_y(a)\psi_y(b)$ and that $\psi_{x+y}(a) = \psi_x(a)\psi_y(a)$.*

Claim 7. *For all $y \in \mathbb{F}_{2^n}$, ψ_y is an eigenvector of G.*

Proof. Let M be the normalized adjacency matrix of G. $(M\psi_y)(a) = \mathbb{E}_{b \text{ adjacent to } a}[\psi_y(b)]$, where $M\psi_y(a)$ is formed by evaluating ψ_y at every point in $\mathbb{F}_{2^n}$ and taking the matrix-vector product. Then, using lots of linearity (LOL), we can establish:

$$(M\psi_y)(a) = \mathop{\mathbb{E}}_{x \in S}[\psi_y(a+x)] = \mathop{\mathbb{E}}_{x \in S}[\psi_y(a)\psi_y(x)] = \psi_y(a) \mathop{\mathbb{E}}_{x \in S}[\psi_y(x)]$$

so ψ_y is an eigenvector of eigenvalue $\mathbb{E}_{x \in S}[\psi_y(x)]$. This eigenvalue is bounded by: $|\mathbb{E}_{x \in S}[\psi_y(x)]| = |\mathbb{E}_{x \in S}[(-1)^{\langle y,x \rangle}]| < \epsilon$ if $y \neq 0$. $\mathbb{E}_{x \in S}[(-1)^{\langle y,x \rangle}] = 1$ if $y = 0$.

Hence, $\text{Cay}(\mathbb{F}_{2^n}; S)$ is a $|S|$-regular ϵ-expander. □

4 Size of ϵ-biased sets

Claim 8. *Any ϵ-biased set $S \subseteq \{0,1\}^n$ has size at least $O(n/\epsilon^2)$.*

Proof. Let S be an ϵ-biased subset of $\{0,1\}^n$. Consider the distribution $S + S + \cdots + S$ (S added to itself t times). This the is random variable $\mu_t = x_1 + x_2 + \cdots + x_t$, where each x_i is uniform over S. We also define the function $\mu_t(z)$ to be the probability mass function of μ_t. For any $y \in \{0,1\}^n$, $\mathbb{E}_{z \in \mu_t}[(-1)^{\langle z,y \rangle}] = \mathbb{E}_{x_1,\ldots,x_t}[(-1)^{\langle x_1+\cdots+x_t,y \rangle}] = \mathbb{E}_{x_1,\ldots,x_t}[(-1)^{\sum_{i=1}^t \langle x_i,y \rangle}] = \mathbb{E}_{x_1,\ldots,x_t}[\prod_{i=1}^t (-1)^{\langle x_i,y \rangle}] = \prod_{i=1}^t \mathbb{E}_{x_i}[(-1)^{\langle x_i,y \rangle}]$, which is at most ϵ^t in absolute value by relation to the tth power of $\text{Cay}(\mathbb{F}_2^n; S)$. Note that $\mu_t(z) = \frac{1}{|S|^t}\text{Card}\{(x_1,\ldots,x_t) : \sum_{i=1}^t x_i = z\}$.

Definition 9. *Let f and g be functions from $\mathbb{F}_2^n$ to $\mathbb{R}$. We define the convolution of f and g as $(f \star g)(a) = \mathbb{E}_{b_1,b_2:b_1+b_2=a}[f(b_1)g(b_2)]$.*

Notice that $\widehat{(f \star g)}(y) = \mathbb{E}_a[(f \star g)(a)\psi_y(a)] = \mathbb{E}_{a,b_1,b_2:b_1+b_2=a}[f(b_1)g(b_2)\psi_y(a)] =$
$\mathbb{E}_{b_1,b_2}[f(b_1)\psi_y(b_1)g(b_2)\psi_y(b_2)] = \mathbb{E}_{b_1}[f(b_1)\psi_y(b_1)]\,\mathbb{E}_{b_2}[g(b_2)\psi_y(b_2)] = \hat{f}(y)\hat{g}(y)$.

For example, let $f = \frac{2^n}{|S|}1_S$. Then $(f \star f)(a) = \mathbb{E}_{b_1,b_2:b_1+b_2=a}[f(b_1)f(b_2)] =$
$(\frac{2^n}{|s|})^2 \frac{1}{2^n} \sum_{b_1,b_2:b_1+b_2=a} 1_S(b_1)1_S(b_2) = (\frac{2^n}{|s|})^2 \frac{1}{2^n}\text{Card}\{(b_1,b_2) \in S \times S : b_1 + b_2 = a\} =$
$\frac{2^n}{|S|^2}\text{Card}\{(b_1,b_2) \in S \times S : b_1 + b_2 = a\} = 2^n\mu_2(a)$.

More generally, f convoluted with itself t times is $f^{\star t}(a) = (f \star \cdots \star f)(a) = \mathbb{E}_{b_1,\ldots,b_t: \sum_{i=1}^t b_i = a}[\prod_{i=1}^t f(b_i)] = \frac{1}{2^{n(t-1)}} \sum_{b_1,\ldots,b_t: \sum_{i=1}^t b_i = a} (\frac{2^n}{|S|})^t \prod_{i=1}^t 1_S(b_i) = \frac{2^n}{|S|^t} \mathrm{Card}\{(b_1,\ldots,b_t) \in S^t : \sum_{i=1}^t b_i = a\} = 2^n \mu_t(a)$. Notice that $f^{\star t}/2^n = \mu_t$ is ϵ^t-biased.

Definition 10. *The support of μ_t, written supp(μ_t), is the number of points z where $\mu_t(z) \neq 0$.*

Since t is much smaller than $|S|$, $\mathrm{supp}(\mu_t) \leq \sum_{k=0}^t \binom{|S|}{k} \leq t\binom{|S|}{t}$.

Notice that $\hat{\mu}_t(y) = \mathbb{E}_a[\mu_t(a)\psi_y(a)] = \frac{1}{2^n} \sum_a \mu_t(a)\psi_y(a) = \frac{1}{2^n} \mathbb{E}_{a \in \mu_t} \mu_t(a)\psi_y(a) \leq \frac{\epsilon^t}{2^n}$.

Then $\mathrm{supp}(\mu_t) \geq \left(\frac{(\sum_a \mu_t(a))^2}{\sum_a \mu_t^2(a)}\right) = \frac{1}{\sum_a \mu_t^2(a)}$.

Since $\mathbb{E}_a[\mu_t^2(a)] = \sum_y \hat{\mu}_t^2(a) = \frac{1}{2^n} + \frac{2^n-1}{2^n}\epsilon^{2t}$, $\mathrm{supp}(\mu_t) \geq \frac{2^n}{1+(2^n-1)\epsilon^{2t}}$.

The inequalities $t\binom{|S|}{t} \geq \frac{2^n}{1+(2^n-1)\epsilon^{2t}}$ and $(\frac{|S|}{t})^t \leq \binom{|S|}{t} \leq (\frac{e|S|}{t})^t$ imply that $t(\frac{e|S|}{t})^t \geq \frac{2^n}{1+(2^n-1)\epsilon^{2t}}$.

Take t such that $\epsilon^{2t} \in \Theta(\frac{1}{2^n})$, that is, $t = \frac{n}{2\log(1/\epsilon)}$.

Since $\sqrt[t]{2^n} = 1/\epsilon^2$, $|S| \geq (\frac{2^n}{O(1)t})^t \frac{t}{e} \in O(\frac{n}{2e\log(1/\epsilon)\epsilon^2}) = O(\frac{n}{\epsilon^2 \log(1/\epsilon)})$, as desired. □

Claim 11. *If μ is 0-biased, then μ is uniform.*

Proof. Let μ be a 0-biased distribution. Let y be nonzero. Since μ is 0-biased, by definition, we have that $\mathbb{E}_x[\mu(x)\psi_y(x)] = \frac{1}{2^n}\sum_x \mu(x)\psi_y(x) = \frac{1}{2^n}\mathbb{E}_{x\in\mu}[\psi_y(x)] \leq 0$, so $\langle \mu, \psi_y \rangle = 0$ for all nonzero y. This implies that μ is parallel to ψ_0, so μ is uniform, as desired. □

5 Relation to k-wise independence

Claim 12. *If μ is ϵ-biased, then $\|\mu - U\|_1 \leq \epsilon 2^{n/2}$.*

Proof. By definition, $\|\mu - U\|_1 = \sum_x |\mu(x) - (1/2^n)| \leq \sqrt{\sum_x (\mu(x) - \frac{1}{2^n})^2}\sqrt{2^n}$. Therefore, $\mathbb{E}[(\mu - U)^2] = \langle \mu - U, \mu - U \rangle = \sum_y ((\widehat{\mu - U})(y))^2 = \sum_y (\hat{\mu}(y) - \hat{U}(y))^2 \leq (2^n - 1)(\frac{\epsilon}{2^n})^2 \leq (\epsilon^2/2^n)$. Hence, if $y \neq 0$, then $|\hat{\mu}(y)| \leq \epsilon/2^n$ and $\hat{U}(y) = 0$ and if $y = 0$, then $|\hat{\mu}(y)| = |\hat{U}(y)| = 1/2^n$. $\sum_x (\mu(x) - 1/2^n)^2 \leq \epsilon^2$, so $\|\mu - U\|_1 \leq \epsilon 2^{n/2}$, as desired. □

Let μ be an ϵ-biased distribution on $\mathbb{F}_2^n$. Let $x_1, \ldots, x_n$ be randomly sampled from μ. Then for $I \subseteq [n]$ such that $|I| \leq k$, the distribution $(x_{i_1}, x_{i_2}, \ldots, x_{i_\ell})$ is $\epsilon 2^{k/2}$-close to uniform over $\mathbb{F}_2^{|I|}$ by the claim above. Taking $\epsilon = \delta 2^{k/2}$, we can obtain a δ-almost k-wise independent distribution which we can sample with $2\log\frac{n}{\epsilon} = 2\log n + k + 2\log(1/\delta)$ random bits.