1 Part 1: Explicit, almost optimal $\epsilon$-biased sets. Lecturer: Matthew Dippel; Scribe: Willy Quach

In this lecture we discuss explicit construction of $\epsilon$-biased sets with almost optimal support size.

**Definition 1.** $\epsilon$-biased sets. A set $S \subseteq \{0, 1\}^n$ is $\epsilon$-biased if for all linear test $a$:

$$\left| \Pr_{x \in S}[\langle a, x \rangle = 1] - \Pr_{x \in S}[\langle a, x \rangle = 0] \right| \leq \epsilon.$$ 

In this lecture, we will focus on proving the following theorem:

**Theorem 2.** There is an explicit construction of an $\epsilon$-biased set $S \subseteq \{0, 1\}^n$ such that $|S| = O\left(\frac{n}{\epsilon^2 d} \right)$ where $d = o(1)$.

Note that we saw in class a construction an $\epsilon$-biased set $S$ with $|S| = O\left(n \right)$ and the size of any $\epsilon$-biased set is lower bounded by $\Omega\left(\frac{n}{\epsilon^2 \log 1/\epsilon}\right)$.

The basic idea is the following fact:

**Claim 3.** If $S$ is $\epsilon$-biased, then for all $k \geq 1$, the sum of $k$ i.i.d samples from $S$ is $\epsilon^k$-biased.

This is not enough to give an explicit construction by itself, as the support size grows roughly exponentially with $k$.

The idea is to leverage this fact, but using pseudorandomness for the samples. More precisely, we will start with some $\epsilon_0$-biased set $S$ for some constant $\epsilon_0$, and map the elements of $S$ onto the nodes of an expander graph (recall that taking a somewhat short random walk over an expander graph leads to a distribution close to uniform over the vertices of the graph). Then, we hope that the sum of the elements seen while randomly walking through the graph is a good $\epsilon$-biased set.

More precisely, (long) random walks on expanders are good parity samplers: for any linear test $a$, define $B_a = \{v \in V \mid \langle a, v \rangle = 1\}$. Then:

$$|\Pr[\text{The walk hits } B_a \text{ an odd number of times }] - 1/2| \leq \epsilon.$$
Note that this is better than drawing $t$-wise independent samples (where $t$ is the length of the random walk). Indeed, setting such a $v$ such that $\sum_{i=0}^{t} v_i = 0$ implies that $\sum_{i=0}^{t} \langle a, v_i \rangle = 0$ for all linear test $a$ (and therefore fails the parity test).

The main idea is the following: take $G$ to be an expander graph, such that we map the elements of $S$ onto its vertices. If its degree is too large, then sampling a random walk on $G$ costs too much. Instead, we consider another expander $H$ whose vertices correspond to edges connected to a fixed vertex in $G$; in other words, the number of vertices in $H$ is the degree of $G$, and each vertex of $H$ corresponds to a next edge to take for the walk in $G$. Therefore, a random walk on $H$ induces a random walk on $G$ as well (again, where the vertex reached on $H$ defines the next step to take on $G$). If the degree of $H$ is much less than the degree of $G$, this allows to have a much smaller support; and one can hope that if $H$ is an expander, then the random walk on $G$ actually achieves the desired properties.

More formally, we consider an expander $G = (N_1, D_1, \lambda_1)$ (where $N_1$ is the number of vertices, $D_1$ the degree, and $\lambda_1 = \max\{\lambda_2(G), \lambda_n(G)\}$), and $H = (N_2, D_2, \lambda_2)$ where $N_2 = D_1^s$ for some parameter $s$ (think of $s$ as a large constant). In particular, each vertex of $H$ can be viewed as a list of $s$ elements in $[D_1]$. Then, any random walk on $H$ induces a deterministic walk on $G$ in the following way: to take the $\ell$th step, take a step on $G$ where the edge is determined by the $\ell \mod s$-th element of the current edge in $H$ (again, this element is an element in $[D_1]$), so it defines an edge going out the current vertex in $G$), and then take a step on $H$. Intuitively, this corresponds to apply the procedure described above, with $s$ parallel copies of $H$.

Such a construction allows to get the following parameters (on input $n$, $\epsilon$): Take $d = \Theta\left(\left(\frac{\log \log 1/\epsilon}{\log 1/\epsilon}\right)^{1/3}\right)$, and $H$ such that $s = 1/d$, $D_2 = s^{4s}$ (for instance $H$ can be taken to be the Cayley Graph over $\mathbb{Z}_2^{\log |D_2|}$, the initial distribution is an $\epsilon_0$-biased distribution with support size $\tilde{O}(n/\epsilon_0^2)$, with $\epsilon_0 = 1/D_2$. Take $G$ to be a Ramanujan expander with $D_1 = O(1/\lambda_1^3)$, $N_1 \approx |S|$. Then it suffices to consider a random walk of length $t$, where $t$ is the smallest integer such that $\lambda_2^{(1-4d)/(1-d)t} \leq \epsilon$ (and in particular $t \geq 1/d^2$).

Let us show how such a random walk allows to reduce the bias, even in the case when we do not use an outer graph $H$. The main idea is to express
the bias of the resulting distribution using linear algebra.

We start with a $\epsilon_0$-biased distribution over $G$ (say, that $\epsilon_0$ is a constant, for simplicity). Suppose $N_1$ is such that $N_1 \in [(1-\beta)n, n]$ or $N_1 \in [(1-\beta)2n, 2n]$ for some small constant $\beta$. We sample a random walk of length $t$. Let $\alpha$ be a best linear distinguisher for the resulting distribution, and define:

$$S_b = \{v \in N_1 \mid \langle \alpha, v \rangle = b\}, \text{ and } \Pi_b \text{ to be the projection on } S_b, \text{ where } b \in \{0, 1\}. \text{ Let } \Pi = \Pi_0 - \Pi_1. \text{ Let } \Upsilon \text{ be the resulting distribution of the random walk. Let } p_{\text{even}}(S_1) \text{ (respectively } p_{\text{odd}}(S_1)) \text{ be the probability that the random walk visits } S_1 \text{ an even (respectively odd) number of times. Let } 1 \text{ be the unit vector colinear with } (1, \ldots, 1).

**Theorem 4.**

We have:

1. $\text{Bias}(\Upsilon) = |p_{\text{even}}(S_1) - p_{\text{odd}}(S_1)|$;
2. $p_{\text{even}}(S_1) - p_{\text{odd}}(S_1) = \sum_{b_0 \ldots b_t \in \{0, 1\}} (-1)^{b_t} \Pi_{b_t} G \cdots \Pi_{b_1} G \Pi_{b_0} 1$;
3. $p_{\text{even}}(S_1) - p_{\text{odd}}(S_1) = 1^T (\Pi G)^t \Pi 1$;
4. $\| (\Pi G)^2 \| \leq \epsilon_0 + 2\beta + 2\lambda$;
5. $\text{Bias}(\Upsilon) \leq (\epsilon_0 + 2\beta + 2\lambda)^{\lfloor t/2 \rfloor}$.

We prove item 4: if $v$ is of norm 1, we can write $v = v^\parallel + v^\perp$ along $\text{Span}(1)$ and its orthogonal, such that $Gv^\parallel = v^\parallel = \|v^\parallel\|1$. Then:

$$\| (\Pi G)^2 \| \leq \| (\Pi G)^2 v^\parallel \| \leq \| (\Pi G)^2 v^\parallel\| + \| (\Pi G)^2 v^\perp\|,$$

$$\leq \| v^\parallel \| \| \Pi G \Pi 1 \| + \| \Pi G \| \| v^\perp \|,$$

$$\leq \| \Pi G (\Pi 1)^\parallel\| + \| \Pi G (\Pi 1)^\perp\| + \| v^\perp \|,$$

$$\leq \| \Pi 1\| + 2\lambda.$$

Then, note that $\| \Pi 1\| = \| \langle \Pi 1, 1 \rangle \| = \left| \frac{|S_0| - |S_1|}{N_1} \right|$. As the initial distribution $\Upsilon_0$ is $\epsilon_0$ biased and we removed at most $\beta n$ elements we have:

$$\|S_0| - |S_1| \leq \frac{1 + \epsilon_0}{2} n - \left( \frac{1 - \epsilon_0}{2} n - \beta n \right) \leq (\epsilon_0 + 2\beta)N_1.$$
2 Part 2: Quadratic Time Hardness of the Longest Common Subsequence Problem. Lecturer: Tanay Mehta

Let us focus on Fine-Grained Complexity, which mainly establishes lower bounds on the hardness of problems in P (assuming the hardness of a few problems).

The main conjectured hard problems in fine-grained complexity are the following:

- **3SUM**: given a set in $S \subset [-n^3, n^3]$ of size $n$, find three elements $a, b, c$ such that $a + b = c$. Its conjectured hardness is $n^{2-o(1)}$ time.

- **APSP (All Pairs Shortest Paths)**: given a weighted graph $G$, compute the (weighted) distance between all pairs of vertices. Its conjectured hardness is $n^{3-o(1)}$ time.

- **OV (Orthogonal Vectors)**: given two sets $U, V$ of vectors in $\{0, 1\}^d$, decide if there exists $u \in U, v \in V$ such that $\langle u, v \rangle = 0$. Its conjectured hardness is $n^{2-o(1)}$ time for $d = \omega(\log n)$ (and is in general $\approx n^2d$).

Interestingly, the hardness of OV is implied by the Strong Exponential Time Hypothesis (SETH).

**Definition 1.** The Strong Exponential Time Hypothesis states that:

$$\forall \epsilon > 0, \exists k, k\text{-SAT requires } 2^{(1-\epsilon)n} \text{ time.}$$

**Claim 2.** Assuming SETH, OV requires $\Omega(n^2d)$ time to solve.

The reduction from $k$-SAT to OV is surprisingly simple: given a SAT instance $\phi$ on $n$ variables and $m$ clauses, split the variables into two disjoint sets $A, B$ of size $n/2$, and define:

$U = \{\vec{u} \in \{0, 1\}^m, \vec{u}_i = 0 \text{ if and only if the } i\text{th clause is satisfied by some partial assignment } a \in A\},$

$V = \{\vec{v} \in \{0, 1\}^m, \vec{v}_i = 0 \text{ if and only if the } i\text{th clause is satisfied by some partial assignment } b \in B\}.$
Then $\phi$ is satisfiable if and only if there is a pair of orthogonal vectors across $U,V$ (were each contains $2^{n/2}$ vectors, one for each possible partial assignment in $A$ and $B$, respectively).

In the following, we will be more interested in an extension of the OV problem:

**Definition 3.** The Most-OV problem consists in, given an integer $r$, and two sets $U,V \in (\{0,1\}^d)^n$ of $n$ vectors of dimension $d$, decide if there exists $u \in U,v \in V$ such that $\langle u,v \rangle \leq r$.

Recall that a subsequence of some string $z = z_1 \ldots z_n$ is a string $z_{i_1} \ldots z_{i_k}$ where $\{i_j\}_j$ is an increasing sequence of integers. In particular, a subsequence does not necessarily consist in consecutive letters in the original string.

**Definition 4.** The Longest Common Subsequence (LCS) problem consists in, given two strings $P_1, P_2$ of length $n$ over some alphabet $\Sigma$, compute the length of their Longest Common Subsequence.

We will prove the following theorem:

**Theorem 5.** If there exists some $\epsilon > 0$ such that LCS over an alphabet of size 7 can be solved in $O(n^{2-\epsilon})$ time, then Most-OV can be solved in $O(n^{2-\epsilon}d)$ time.

We will next sketch the proof of the theorem.

Define Weighted LCS (WLCS) to be the LCS problem with weights on the elements of the alphabet; the goal is then to maximize the weight of a common subsequence. Note that WeightedLCS reduces to LCS: if $\alpha \in \Sigma$ has weight $w$, simply define a morphism that maps $\alpha$ to $\alpha^w$.

Therefore, it suffices to reduce Most-OV to Weighted LCS.

Let $\{\alpha\}_{[n]}, \{\beta\}_{[n]}$ be a Most-OV instance, and let $\Sigma = \{0,\ldots,6\}$.

Define the following Coordinate Gadgets:

$$CG_1(\alpha, i) = \begin{cases} 5465 & \text{if } \alpha_i = 0 \\ 545 & \text{otherwise} \end{cases}$$

$$CG_2(\beta, i) = \begin{cases} 5645 & \text{if } \beta_i = 0 \\ 565 & \text{otherwise} \end{cases}$$

and define weights $w(5) = X := 100d$, $w(4) = w(6) = 1$.

Note that: $WLCS(CG_1(\alpha, i), CG_2(\beta, i)) = \begin{cases} 2X + 1 & \text{if } \alpha_i \beta_i = 0 \\ 2X & \text{otherwise} \end{cases}$.
Define now the following Vector Gadgets:

\[ VG_1(\alpha) = 1 \circ o_{i=1}^{d} CG_1(\alpha, i), \]
\[ VG_2(\beta) = o_{i=1}^{d} CG_2(\beta, i) \circ 1, \]
with weight \( w(1) = A := (r + 1)2X + (d - (r + 1))(2X + 1). \)

Claim 6. If \( \langle \alpha, \beta \rangle \leq r \), then:

\[ WLCS(VG_1(\alpha), VG_2(\beta)) \geq A + 1 = r \cdot 2X + (d - r)(2X + 1). \]

The claim above follows directly from the construction.

Claim 7. If \( \langle \alpha, \beta \rangle > r \), then:

\[ WLCS(VG_1(\alpha), VG_2(\beta)) = A. \]

To see this, note that 1 is a common subsequence, so that the WLCS is at least \( A \).

Furthermore, if 1 is not taken in the subsequence we can assume without loss of generality that the 5's map to each other as letters in the subsequences, and at least \( r + 1 \) letters in between that match, with weight 1 each. The inequality follows.

We can now build the sequences for the WLCS problem. Define:

\[ P_2 = 3o(o_{i=1}^{n-1}(0 \circ VG_2(1^d) \circ 2 \circ 3))o(o_{i=1}^{n-1}(0 \circ VG_2(\beta^i) \circ 2 \circ 3))o(o_{i=1}^{n}(0 \circ VG_2(1^d) \circ 2 \circ 3)); \]
\[ P_1 := 3^{|P_2|} o(o_{i=1}^{n}(0 \circ VG_1(\alpha^i) \circ 2)) o 3^{|P_2|}, \]

with weights \( w(3) = A^2 \) and \( w(0) = w(2) = A^4 \).

With some additional work, one can show that \( P_1 \) and \( P_2 \) have their WLCS greater than \( n \cdot (2A^4 + 1) + 2nA^2 \) if and only if there are no vectors in \( \{\alpha\}_{[n]} \), \( \{\beta\}_{[n]} \) with inner product less than \( r \).