In this lecture we continue the discussion of the properties of an adjacency matrix of a graph. We discuss how it can be viewed as a transition matrix of a random walk and the derandomization of randomness extraction from bit-fixing sources by performing a random walk on expander graphs. We then discuss non-oblivious bit-fixing sources and their relation to collective coin flipping and leader election. We see that it is much harder to extract randomness from such sources and we give some bounds on the size of the set of “bad guys” even to extract 1 bit. We discuss several algorithms such as majority, iterative majority, and baton-passing.

1 Adjacency matrices over graphs

We begin with some review from last time. Given a graph \( G = (V, E) \) with an adjacency matrix \( A = \begin{cases} 1 & \text{if } (i, j) \in E, \\ 0 & \text{otherwise} \end{cases} \), notice that \( A \) is a symmetric matrix.

**Fact 1**: Any real symmetric \( N \times N \) matrix has \( N \) orthonormal eigenvectors \( v_1 \ldots v_n \) corresponding to eigenvalues \( \lambda_1 \ldots \lambda_n \).

Recall: The inner product \( \langle u, v \rangle \) is defined as \( \sum_{i=1}^{N} u_i v_i \), \( \| u \| = \sqrt{\langle u, u \rangle} \), \( \forall i, \| v_i \| = 1 \), so the eigenvectors are normalized, \( \forall i \neq j \) \( \langle v_i, v_j \rangle = 0 \), meaning \( v_i \perp v_j \) and so all eigenvectors are orthonormal, and \( Av_i = \lambda_i v_i \) by the definition of an eigenvector.

**Example**: \( G \) is a hypercube of \( \{0, 1\}^n \), so \( N = 2^n \). Then for vector \( z \in \{0, 1\}^n \), \( v_z(x) = (-1)^{xz} \sqrt{\frac{1}{N}} \) and \( \lambda_z = n - 2 \times \text{weight}(z) \).

**Theorem 2**  Courant-Fischer Theorem: Assume \( \lambda_1 \geq \lambda_2 \ldots \geq \lambda_n \) and \( u \perp v_1 \ldots v_n \). Then, \( u^\top A u = \lambda_{i+1} \| u \|^2 \) and \( v_{i+1}^\top A v_{i+1} = \lambda_{i+1} \| v_{i+1} \|^2 \).

**Proof**: \( u = \sum_{j>i} c_j v_j \) where \( c_j = \langle u, v_j \rangle \) \( \Rightarrow u^\top A u = (\sum c_j v_j) A (\sum c_j v_j) = \sum_{k, j>i} c_k v_j \lambda_j \langle v_k, v_j \rangle = \sum \lambda_j c_j^2 \leq \lambda_{i+1} \sum c_j^2 = \lambda_{i+1} \| u \|^2 \). \( \square \)

**Lemma 3**: Assume \( |\lambda_1| \geq |\lambda_2| \ldots \geq |\lambda_n| \) and \( u \perp v_1 \ldots v_n \). Then, \( \| A u \| \leq |\lambda_{i+1}| \cdot \| u \| \).

**Proof**: \( \| A u \|^2 = \langle A \sum_{j>i} c_j v_j, A \sum_{j>i} c_j v_j \rangle = \langle \sum c_j \lambda_j v_j, \sum c_j \lambda_j v_j \rangle \) by applying \( A \) to its eigenvector, \( = \sum_{j>i} c_j^2 \lambda_j^2 \leq \lambda_{i+1}^2 \sum c_j^2 \) since the lambda’s were ordered by absolute value in decreasing order = \( \lambda_{i+1}^2 \| u \|^2 \). \( \square \)
2 d-regular graphs

Definition 4 $G$ is a d-regular graph if every row of $A$ has $d$ ones and everything else is 0.

Then we have that $A \cdot 1^N = d \cdot 1^N$ which implies that $1^N$ is an eigenvector corresponding to the eigenvalue $\lambda = d$.

Fact 5 1. $G$ is disconnected $\iff d$ has multiplicity greater than 1.

2. $G$-connected, $G$-bipartide $\iff -d$ is eigenvalue.

3. $G$-connected, $G$-nonbipartide $\Rightarrow |\lambda_2| \leq \frac{1}{N \cdot \text{diameter}(G)}$, in particular, $|\lambda_1|/|\lambda_1| = d$.

Usually, people normalize $A$ by dividing all values by $d$. This has several advantages:

- The eigenvectors stay the same, but every eigenvalue $\lambda_i$ is divided by $d$ which means that $\lambda_1 = 1$.
- You can now view $A$ as a transition matrix of a random walk on $G$ where $A = \left\{ \frac{1}{d} \right\}_{(i,j) \in E}$, 0 otherwise, and it corresponds to the probabilities of moving from $i$ to $j$ and the sum of all the rows is 1.

Let us now begin an examination of $A$ as a transition matrix for a random walk. Assume $p$ is a distribution, and $\sum p_i = 1$. Then we see that $(Ap)_i = \sum_{j=1}^{N} a_{ji}p_j$. So $Ap$ is a probability distribution after taking one step starting from $p$. Set $u = (\frac{1}{N}, \ldots, \frac{1}{N})$. $U$ can be viewed both as the uniform distribution and also as the eigenvector $\frac{1}{\sqrt{N}}$ corresponding to the eigenvalue $\lambda_1 = 1$. Now, take any probability distribution $p$ and write it as $p = \sum_{i=1}^{N} c_i v_i$.

Notice firstly that $c_i = \frac{1}{\sqrt{N}}$ because $\sum p_i = 1$, therefore $\sum p_i - \frac{1}{N} = 0 \Rightarrow (p - u) \perp u$. So $p = u + \sum_{i \geq 1} c_i v_i$. So we see that any distribution can be written as the uniform distribution plus a linear combination of eigenvectors. Also notice that $Au = u$.

Now let's see what happens to the distribution after taking one step using the transition matrix. $\|Ap - u\| = \|Ap - Au\| = \|A(p - u)\|$. Since we saw that $(p - u) \perp u$, we can apply the lemma to get $\|A(p - u)\| \leq |\lambda_2| \cdot \|p - u\|$. Thus, if $\lambda_2 < 1$ then taking one step got us closer to the uniform distribution. This implies that $\|A^lp - u\| \leq \|\lambda_2|^l \cdot \|p - u\|$.

Now let us examine the statistical distance. We know that for any distributions $x, y$, $SD(x, y) = 1/2 \sum_{i=1}^{N} x_i - y_i$, which by the Cauchy-Schwartz Theorem is $\leq \sqrt{\sum_{i=1}^{N} x_i y_i} = 1/2 \cdot \sqrt{N} \cdot \sqrt{\sum_{i=1}^{N} x_i y_i} = 1/2 \cdot \sqrt{N} \cdot \|x - y\|$. So $SD(A^lp, u) \leq \frac{1}{2} \cdot \sqrt{N} \cdot |\lambda_2|^l \cdot \|p - u\| \leq \sqrt{N} \cdot |\lambda_2|^l$, since we know that $\|p - u\| \leq 2$.

To make the statistical distance less than $\epsilon$, we need $\sqrt{N} \cdot |\lambda_2|^l < \epsilon$ so $l < \frac{\log N}{2 \log \frac{1}{|\lambda_2|}}$ in order to be $\epsilon$-close to uniform.
However, \( \forall G, |\lambda_2| \geq \frac{2\sqrt{d-1}}{d} \sim \frac{2}{\sqrt{d}} \) tight, and is achieved by random graphs and by Ramanujan graphs. If you use Ramanujan graphs, then you need \( l \approx \frac{\log_2 N}{\log_2 d} \) steps to get to an almost uniform distribution.

3 Application to bit-fixing sources

In order to use this for a bit-fixing source, assume you have a graph with \( N = 2^k \). The source outputs \( x_1 \ldots x_n \), with each \( x_i \in \{1 \ldots d\} \) and use the \( x_i \)'s for a random walk and then output \( k \) bits at the end of the walk from the last vertex. Only \( l \) of the \( x_i \)'s are random, and the rest are fixed. It can be shown that the random steps get you closer to uniform which the fixed steps don’t get you any further.

Assume you get an \((N, \lambda_2)\) expander graph. It is \( \epsilon \)-close to uniform if \( 2 \log \frac{1}{|\lambda_2|} \cdot l \geq \log N + 2 \log \frac{1}{\epsilon} \Rightarrow k \approx 2 \log \frac{1}{|\lambda_2|} \cdot l - 2 \log \frac{1}{\epsilon} \). If you use Ramanujan, then \( \lambda_2 \approx \frac{2}{\sqrt{d}} \Rightarrow k \approx l \log d - 2 \log \frac{1}{\epsilon} \). This is a fantastic result, since the original entropy was \( l \log d \) and you’re getting entropy minus a small constant term of bits output.

This is only true for alphabets of size \( > 2 \). For the Binary case [GRS05]: \( k = l \cdot (1 - o(1)) \).

If \( l \geq \sqrt{n} \), then one can get \( \frac{1}{\epsilon} \sim 2^{-n\Omega(1)} \). If \( l \leq \sqrt{n} \), then you can only get \( \epsilon = l^{-\Omega(1)} \).

4 Non-oblivious bit-fixing sources

- \( A \) chooses \( |S| = l \)
- bits in \( S \leftarrow \$ \) given to \( A \)
- \( A \) sets remaining \((n - l)\) bits

For the case where we extract one bit, this is analogous to collective coin flipping and leader election where \( b \) out of \( n \) players are bad and \( b = n - l \). We can use leader elections to extract a random bit by letting the elected leader flip a coin.

The simplest case is when \( l = n - 1, b = 1, \) and \( f : \{0, 1\}^n \rightarrow \{0, 1\} \). (Note that parity is completely fixed, even for \( b = 1 \).)

**Definition 6** A function \( f \) is \((b, \epsilon)\)-resilient if \( \forall \) coalitions of “bad guys” of size \( b \):

\[
Pr(f(x) = 0), Pr(f(x) = 1) \geq \epsilon
\]

(Note that this is very weak compared to our usual requirement of \( 1/2 + \delta, \frac{1}{2} - \delta \))

Historically, the goal has been \( \epsilon = \Omega(1) \). We want \( \epsilon = \frac{1}{2} - o(1) \). We can hope for \( o(1) = n^{-\omega(1)} \) as before with oblivious sources. But we will see that this result is impossible for non-oblivious sources.

It is easy to see that we can bound \( \epsilon \leq \frac{1}{2} - \Omega(\frac{k}{n}) \). i.e. even for \( b = 1, \epsilon \leq \frac{1}{2} - \frac{1}{n} \).
**Definition 7** The influence of a set $S$ over the function $f$. $I_f(S) = \Pr(f \text{ is not fixed when the vars outside } S \leftarrow \emptyset) = \Pr(S \text{ matters}).$

If $f$ has $|S| \leq b$ such that $I_f(S) \geq \delta$ then $b$ “bad guys” can force the outcome with probability $\frac{1}{2} + \Omega(\delta)$.

It is easy to see that

$$\sum_{i=1}^{n} I_f(x_i) \geq 1 \rightarrow \exists i \ I_f(x_i) \geq \frac{1}{n}$$

We can prove this lower bound $\forall$ balanced $f$ using the hypercube of dimension $n$. There are $2^n$ vertices whose coordinates correspond to the possible inputs, $x$, to $f$. Each vertex is assigned the value $f(x)$.

Let $T$ be a subset of vertices $x$ such that $T = x|f(x) = 0$, $|T| = 2^{n-1}$. Using an isoperimetric inequality on the hypercube we see that $e(T, \{0,1\}^n \setminus T) \geq 2^{n-1}$.

This is the number of 0/1 edges $(x, y\ f(x) = 0, f(y) = 1)$ which corresponds to the values of $x$ for which 1 player can change the value of $f(x)$.

The total number of edges in the hypercube is: $n 2^{n-1}$. So clearly $\frac{1}{n}$ edges are 0/1 edges. Therefore, on average, each $x_i$ has influence $\geq \frac{1}{n}$.

**Theorem 8 (KKL):** $\forall f \exists i$ such that $I_f(x_i) = \Omega\left(\frac{\log(n)}{n}\right)$.

**Corollary 9** $\forall \exists |S| = \Omega\left(\frac{\log(n)}{n}\right)$ such that $I_f(S) = \Omega(1)$.

This implies that we cannot have an $(\frac{n}{\log(n)})$-resilient function. Even if $k = 1, \epsilon = \Omega(1)$ then we have $l \geq n - O\left(\frac{n}{\log(n)}\right)$.

**Practical Algorithms**

**Majority:** standard deviation is $\sqrt{n}$. Therefore, $\exists c_1$ such that if at most $c_1 \sqrt{n}$ bad guys $\rightarrow I_f($badguys$) = O(1)$.

Therefore, majority is $(\sqrt{n}, \frac{1}{2} - O(1))$-resilient.

Can we do better?

**Iterated Majority of 3:** build a tree and keep taking the majority of 3 until we get to the root. If $n = 3^k$ then $2^k$ guys can fix the input (if placed appropriately). $2^k = n^{\log_3(2)} \approx n^{0.63}$.

**Theorem 10 (BL87)** $b$ bad guys have influence $\leq \frac{b}{n^{0.63}}$.

So we get $(n^{0.63-\epsilon}, \frac{1}{2} - O(1))$-resilient function.

**Theorem 11 (AL93)** $\exists \Omega\left(\frac{n}{\log(n)}\right)$-resilient functions.

With multi-round games we can approach this bound.

**Baton Passing** [Saks89, AL93]: pass to random person and last person left is the leader. Can tolerate $\frac{n}{(2+\epsilon)\log(n)}$ bad guys $\forall \epsilon > 0$. 

L5-4