1 Review of notation

A $d$-regular undirected graph $G$ on $n$ vertices, has spectral gap $\gamma = 1 - \lambda$. It has an associated random walk matrix $M = \frac{1}{d}A_G$, where $A_G$ is the adjacency matrix for $G$.

2 Graph reduction

Given an algorithm $A$ that is correct with probability $3/4$, uses $m$ random bits, and runs in time $T$, how can we leverage this algorithm to reduce the error to $2^{-k}$?

Naively, we can run $A$ many times and take the majority, but this takes $O(Tk)$ time and $O(mk)$ random bits. With pairwise independence, we can use only $O(m + k)$ random bits, but need $O(T^{2k})$ time. Expander graphs will allow us to do better, achieving a runtime of $O(Tk)$ and $O(mk)$ random bits.

We start with an expander graph with nodes from the set $V = \{0, 1\}^m$. We randomly choose a starting point $v_1$ (this takes $m$ random bits), then do a random walk for $t - 1$ steps, arriving at vertices $v_2, \ldots, v_t$. Note that this requires an additional $\log d$ random bits for each step.

First, we will prove a result for algorithms with 1-sided error ($\text{RP}$).

**Theorem 2.1 (Hitting property of random walks).** For all $B \subseteq V$, let $\mu_B = \frac{|B|}{n}$ be the density of $B$. Then for a random walk $v_1, \ldots, v_t$,

$$\Pr \left[ \bigvee_{i=1}^{t} v_i \in B \right] \leq (\mu_B + \lambda(1 - \mu_B))^t.$$

3 Proof of hitting property

Let $P$ be the $n \times n$ diagonal matrix with $P_{ii} = 1$ if $i \in B$, and $P_{ii} = 0$ otherwise.

**Claim 3.1.**

$$\Pr \left[ \bigvee_{i=1}^{t} v_i \in B \right] = |uP(MP)^{t-1}|$$

where $u$ is the uniform vector with $u_i = 1/n$.

**Proof.** We prove a similar statement: the probability that the first $t$ steps of the random walk are all in $B$ and the $t$th vertex is $i$ is given by $(uP(MP)^{t-1})_i$. Note that this directly implies our claim. We will prove this by induction on $t$.

When $t = 1$, if $i \in B$, then $(uP)_i = 1/n$ which is the probability we desire. Similarly, if $i \notin B$, $(uP)_i = 0$.

Now if we assume the statement holds for $t - 1$, then $(uP(MP)^{t-2} \cdot M)_i$ is the probability we are in vertex $i$ on the $t$th step and all $t - 1$ vertices were in $B$. We multiply by $P$ to ensure we only have positive probability if $i$ is in $B$. So we find that $(uP(MP)^{t-1})_i$ the probability that all $t$ vertices are in $B$ and the last vertex is $i$. □
3.1 Matrix decomposition

We will now look at some related ideas that will help us finish proving the theorem.

**Definition 3.2.** We say the spectral norm of a matrix $A$ is

$$
\|A\| = \max_{x \in \mathbb{R}^n} \frac{\|xA\|_2}{\|x\|_2}.
$$

It is easy to confirm the following properties of the spectral norm:

- $\|cA\| = c\|A\|$
- $\|AB\| \leq \|A\|\|B\|$
- $\|A + B\| \leq \|A\| + \|B\|$
- $\|xA\|_2 \leq \|x\|_2\|A\|$

**Lemma 3.3 (Matrix decomposition).** For random walk matrix $M$ on graph $G$ with spectral gap $\gamma = 1 - \lambda$,

$$M = \gamma J + \lambda E,$$

where $J$ is the matrix with all entries equal to $1/n$, and $\|E\| \leq 1$.

**Proof.** Let $E = \frac{1}{\lambda}(M - \gamma J)$. For any vector $v$, we can decompose it as $v = v_1 + v_2$, where $v_1 = \langle v_1, u \rangle u$ and $v_2 = v - v_1$. (Note that $v_2 \perp u$).

Then

$$v_1 E = \langle v_1, u \rangle u E
= \frac{\langle v_1, u \rangle}{\lambda} (u M - \gamma u J)
= \frac{\langle v_1, u \rangle}{\lambda} (u - \gamma u)
= \frac{\langle v_1, u \rangle}{\lambda} (u \lambda)
= \langle v_1, u \rangle u = v_1.$$

And $v_2 E = \frac{1}{\lambda}(v_2 M - \gamma v_2 J)$. First, we see that $v_2 J = 0$, since $v_2 \perp u$. Then

$$\langle v_2 E, u \rangle = \frac{1}{\lambda}(v_2 M u^T)
= \frac{1}{\lambda}(v_2 u^T) = 0,$$

so $v_2 E \perp u$ (and also $v_2 E \perp v_1$). We also have that

$$\|v_2 E\|_2 = \frac{1}{\lambda}\|v_2 M\|_2
\leq \frac{1}{\lambda}\|v_2\|_2 = \|v_2\|_2$$

because $v_2 \perp u$ which is the eigenvector corresponding to the largest eigenvalue. So $v_2$ is scaled by at most the second largest eigenvalue, $\lambda$. 
Combining the above results, we get that
\[
\|vE\|_2^2 = \|v_1E + v_2E\|_2^2 = \|v_1E\|_2^2 + \|v_2E\|_2^2 \leq \|v_1\|_2^2 + \|v_2\|_2^2 = \|v\|_2^2
\]
which by the definition of the spectral norm, \(\|E\| \leq 1\).

Now we can use this matrix decomposition to make some progress.

**Claim 3.4.**
\[
\|PMP\| \leq \mu_B + \lambda(1 - \mu_B).
\]

**Proof.** We just use the decomposition on \(M\) and do algebra.
\[
\|PMP\| = \|P(\gamma J + \lambda E)\|
\]
\[
\leq \gamma \|PJP\| + \lambda \|PEP\|
\]
\[
\leq \gamma \|PJP\| + \lambda \|E\| \|P\| = \gamma \|PJP\| + \lambda
\]
since both \(P\) and \(E\) have norms bounded by 1.

Now consider any vector \(x\). Let \(y = xP\). Then since \(yJ = (\sum_i y_i) u\)
\[
\|xPJP\|_2 = \|yJP\|_2
\]
\[
= \left\| \left( \sum_i y_i \right) uP \right\|_2
\]
\[
\leq \left\| \sum_i y_i \right\|_2 \cdot \|uP\|_2
\]
\[
= |\langle 1_B, x \rangle| \cdot \|u\|_2 \quad \text{where } 1_B \text{ is the indicator vector for } B
\]
\[
\leq \sqrt{\mu_B n} \|x\|_2 \cdot \sqrt{\mu_B/n} \quad \text{(Cauchy-Schwarz)}
\]
\[
= \mu_B \|x\|_2
\]
where we use the fact that \(P\) (and \(y\)) have at most \(\mu_B n\) non-zero entries. Since this is true for all \(x\), \(\|PJP\| \leq \mu_B\). And so,
\[
\|PMP\| \leq \gamma \mu_B + \lambda = \mu_B + \lambda(1 - \mu_B).
\]

Now, we can finally finish the proof of the hitting property. Here, we make use of the fact that \(P(MP)^t = P(PMP)^t\) because \(P = P^2\).
\[
|uP(MP)^{t-1}| \leq \sqrt{\mu_B n} \cdot \|uP(PMP)^{t-1}\|_2 \quad \text{(Cauchy-Schwarz)}
\]
\[
\leq \sqrt{\mu_B n} \cdot \|uP\|_2 \|PMP\|^{t-1}
\]
\[
\leq \sqrt{\mu_B n} \sqrt{\mu B/n} (\mu_B + \lambda(1 - \mu_B))^{t-1}
\]
\[
= \mu (\mu_B + \lambda(1 - \mu_B))^{t-1}
\]
\[
\leq (\mu_B + \lambda(1 - \mu_B))^{t}.
\]
4 Chernoff bound for expanders

To extend our result for 2-sided error (BPP), we need the following theorem.

**Theorem 4.1.** Given a graph $G$ on $n$ vertices, let $f : [n] \rightarrow [0,1]$ be any function. For a random walk $v_1, \ldots, v_t$, we have

$$\Pr \left[ \left| \frac{1}{t} \sum_i f(v_i) - \mathbb{E} f \right| \geq \lambda + \epsilon \right] \leq 2e^{-\Omega(\epsilon^2 t)}.$$

Due to time, we did not cover the proof, but it is theorem 4.22 in the Pseudorandomness book.

5 Final remarks

If we want to reduce $\lambda$, (for example to reduce the bias from the above theorem and use expanders to sample $f$,) we can take an expander $G$ raised to some power $k$, and our $\lambda$ becomes $\lambda^k$. But this also increases our degree bound from $d$ to $d^k$, which means we need more random bits for each step in our random walk.

A final distinction on expanders and how explicit they must be. One notion is a mildly explicit expander, which means we can construct the expander in $\text{poly}(n)$ time. But this is bad in our application, because $n = 2^m$, so this actually requires exponential time and space. Since we only care about the neighbors, we can instead use fully explicit expanders, which allow you to find the $i$th adjacent vertex in $O(\log n)$ time.