Telecom Paris Telecom 202b, Coding Theory

ASSIGNMENT 5 WITH SOLUTIONS

Exercise 1 (Random graphs are good expanders)**.** In this exercise, we will show that a randomly chosen bipartite graph is a good expander with high probability. Recall that a bipartite graph with n left vertices, m right vertices, and left degree D is a (γ, α) expander if for all subsets S of left vertices with $|S| \leq \gamma n$, we have $|N(S)| > \alpha|S|$. Here, $N(S)$ denotes the set of neighbours of S.

Let us pick a random graph in the following manner: For each left vertex, pick D neighbours uniformly at random from the set of all $\binom{m}{D}$ subsets of right vertices. This is done independently for each vertex. Call the resulting random graph $\mathcal{G}$. We want to show that for all sufficiently large n, and $m > 3n/4$, $D > 32$, $\gamma = 1/(10D)$, $\alpha = 5D/8$

$$\Pr[\mathcal{G} \text{ is not a } (\gamma, \alpha) \text{ expander}] \leq 1/10.$$

1. Choose any set of left vertices S and set of right vertices T, with $|S| = s \leq \gamma n$ and $|T| \leq \alpha s$. Compute the probability that $N(S) \subset T$.

2. Argue that

$$\Pr[\mathcal{G} \text{ is not a } (\gamma, \alpha) \text{ expander}] = \Pr[\exists S \subset \mathcal{L}, T \subset \mathcal{R} : |S| \leq \gamma n, |T| \leq \alpha|S|,\ N(S) \subset T]$$

   where $\mathcal{L}, \mathcal{R}$ denote the set of left and right vertices respectively.

3. Use the first two parts to get an upper bound on the probability that $\mathcal{G}$ is not an expander.

4. Using the bound $\binom{a}{b} \leq \left(\frac{ea}{b}\right)^b$, prove that as long as $m > 3n/4$, $D > 32$, $\gamma = 1/10$, $\alpha = 5D/8$, the probability that $\mathcal{G}$ is not an expander is $o(1)$.

Solution. 1. Each left vertex chooses D neighbours uniformly at random, and independently of other vertices.

$$\Pr[N(S) \subset T] \leq \left(\frac{\binom{t}{D}}{\binom{m}{D}}\right)^s \leq \left(\frac{t}{m}\right)^{Ds}$$

2. This is straightforward. The random graph $\mathcal{G}$ is not an expander if and only if there exists some subset S of left vertices with $|S| \leq \gamma n$ whose neighbourhood is of size less than or equal to $\alpha|S|$. Moreover, we can restrict ourseleves to the case where $|T| = \alpha|S|$.

3. Using the union bound,

$$\Pr[\mathcal{G} \text{ is not an expander}] \leq \sum_{s=2}^{\gamma n} \sum_{t=D}^{\alpha s} \binom{n}{s} \binom{m}{t} \left(\frac{\alpha|S|}{m}\right)^{Ds}$$

$$\Pr[\mathcal{G} \text{ is not an expander}] \leq \sum_{s=1}^{\gamma n} \binom{n}{s} \binom{m}{\alpha s} \left(\frac{\alpha s}{m}\right)^{Ds}$$

4. Considering $\alpha = 5D/8$ and $\gamma = 1/(10D)$ we have

$$\begin{aligned}
\Pr\left[\mathcal{G} \text{ is not an expander}\right] &\leq \sum_{s=1}^{n/10D} \binom{n}{s} \binom{m}{5Ds/8} \left(\frac{5Ds}{8m}\right)^{sD} \\
&\leq \sum_{s=1}^{n/10D} \left(\frac{ne}{s}\right)^{s} \left(\frac{8me}{5Ds}\right)^{5Ds/8} \left(\frac{5Ds}{8m}\right)^{sD} \\
&\leq \sum_{s=1}^{n/10D} 20^{-s} \\
&\leq 1/10
\end{aligned}$$

The term inside the summation is maximized when $s = \gamma n$ and $t = \alpha\gamma n$. Using the bound on the binomial coefficient,

$$\begin{aligned}
\Pr[\mathcal{G} \text{ is not an expander}] &\leq (\gamma n)(\alpha\gamma n) \binom{n}{\gamma n} \binom{m}{\alpha\gamma n} \left(\frac{\binom{\alpha\gamma n}{D}}{\binom{m}{D}}\right)^{\gamma n} \\
&\leq (\gamma n)(\alpha\gamma n) \left(\frac{e}{\gamma}\right)^{\gamma n} \left(\frac{me}{\alpha\gamma n}\right)^{\alpha\gamma n} \left(\frac{e\alpha\gamma n}{m}\right)^{\gamma D n} \\
&= \alpha\gamma^2 n^2 e^{n\gamma(+\alpha+D)} \left(\frac{1}{\gamma}\right)^{\gamma n} \left(\frac{\alpha\gamma n}{m}\right)^{\gamma n(D-\alpha)} \\
&\leq \alpha\gamma^2 n^2 e^{n(1+13D/8)/10} 10^{n/10} \times \left(\frac{D}{12}\right)^{3nD/80} \\
&= o(1).
\end{aligned}$$

Reference: "Expander graphs and their applications" by S. Hoory, N. Linial, A. Wigderson, `https://www.cs.huji.ac.il/~nati/PAPERS/expander_survey.pdf`

Exercise 2 (Minimum distance)**.** Let $\mathcal{G}$ be an $(n, m, D, \gamma, D(1-\epsilon))$ be an expander graph for some $0 < \epsilon < 1/2$. Given any set of left vertices S, a right vertex v is said to be a unique neighbour of S if it is adjacent to exactly one vertex in S. Let $U(S)$ denote the set of unique neighbours of S.

1. Fix any set of left vertices S such that $|S| \leq \gamma n$. How many edges leave S? Using this, compute an upper bound on the number of vertices in $N(S)$ that have more than one incident edge from S.

2. Use the above to argue that $|U(S)| \geq D(1-2\epsilon)|S|$.

3. Use the second part to argue that the minimum distance of the corresponding expander code is at least γn.

   Hint: Choose any nonzero codeword and label the left vertices by the codeword bits. Let S be the support set of vertices labelled 1. What can you say about $U(S)$?

Solution. 1. The number of edges leaving S is $D|S|$. Since the graph is an expander, S has at least $(1-\epsilon)D|S|$ neighbours. Since there are $D|S|$ edges, by the pigeonhole principle, at most $\epsilon D|S|$ neighbours of S can have more than one incident edge from S.

2. S has at least $(1-\epsilon)D|S|$ neighbours, of which at most $\epsilon D|S|$ neighbours of S can have more than one incident edge from S. Therefore, $|U(S)| \geq (1-2\epsilon)D|S|$.

3. Let S be the support of the nonzero codeword (vertices labelled 1) of least Hamming weight. This is a valid codeword if and only if $U(S)$ is empty. This is because for any parity check in $U(S)$, exactly one neighbour comes from S and the rest are outside S. Hence, this equation cannot be satisfied. Using part 2, we know that for every S of size less than or equal to γn, we have $|U(S)| > 0$. Hence, the minimum distance is at least γn.

Exercise 3 (Minimum distance, alternative method)**.** Let $\mathcal{G}$ be an $(n, m, D, \gamma, D(1-\epsilon))$ expander. We will now show that the minimum distance of the corresponding expander code is $\geq 2\gamma(1-\epsilon)n$. We will prove this by contradiction. For the sake of contradiction, suppose that c^n is a nonzero codeword of Hamming weight less than $2\gamma(1-\epsilon)n$.

- Label the set of left vertices of $\mathcal{G}$ using c^n, and let S be the set of vertices labelled 1. Note that $|S|$ is equal to the Hamming weight of c^n. What is the size of $U(S)$?
- Pick $Q \subset S$ such that $|Q| = \gamma n$. Compute the size of $U(Q)$ and $N(S \backslash Q)$ and argue that $|U(S)| > 0$.

Solution. 1. Using the same argument as in the last part of the previous question, $|U(S)|$ must be 0 for a valid codeword.

2. Since $|Q| = \gamma n$ and the graph is an expander, we have $|U(Q)| \geq (1-2\epsilon)\gamma Dn$. Now, $|S \backslash Q| < (1-2\epsilon)\gamma n$. The number of edges leaving this is strictly less than $(1-2\epsilon)\gamma Dn$. Therefore, the number of vertices in $U(Q)$ that might have a neighbour in $S \backslash Q$ is strictly less than $(1-2\epsilon)\gamma Dn$, which implies that $U(S) > 0$.

Using the above, we see that the minimum distance is at least $2\gamma(1-\epsilon)n$.

Exercise 4 (Encoding/decoding complexity of expander codes)**.** Expander codes have low encoding and decoding complexity.

- What is the encoding complexity of an expander code?
- What is the computational complexity in each iteration of decoding an expander code? Use this to find the worst-case computational complexity assuming that the number of errors is less than $\gamma(1-2\epsilon)n$.

Solution. 1. An expander code is linear. Hence, the encoding complexity is $O(n^2)$.

2. In each iteration, we need to find a left vertex which has more unsatisfied neighbours than satisfied ones, and also update the parities at the right vertices. The complexity is $O(n)$.

   At each iteration the bit-flipping algorithm takes $O(n)$ time to find a variable with a number of violated clauses larger than the number of correct clauses. At each step the total number of violated clauses decreases by at least one. Hence, the total number of steps is $O(n)$ and hence the overall decoding complexity is $O(n^2)$.