Expander Code

Scribe: Raymond Liu

October 21, 2025

1 Expander Code

We begin by introducing the concept of low density parity check code.

Definition 1.1 (Low Density Parity Check Codes)**.** *Let $H \in \mathbb{F}_2^{m \times n}$ be a parity check matrix of $C \subseteq \mathbb{F}_2^n$ such that $C = \{x \in \mathbb{F}_2^n \mid Hx = 0\}$. If each row of H has constant number of 1's, we say it is a low density parity check code (LDPC code).*

We can alternatively think of H as the adjacency matrix of a bipartite graph $G = ([n], [m], E)$, where there is an edge between $i \in [n], j \in [m]$ iff $H(j, i) = 1$. Any word $w \in \mathbb{F}_2^n$ now corresponds to a set $S \subseteq [n]$, and any parity check now corresponds to a vertex in $[m]$. In particular, for any set $S \subseteq [n]$, it passes all parity checks if and only if S has even number of edges mapping to each of its neighbour.

With this view, we can thus define LDPC code through a bipartite graph

Definition 1.2 (LDPC Code from Bipartite Graph)**.** *Let $G = ([n], [m], E)$ be a bipartite graph where $m < n$. The induced LDPC code on this graph is defined as*

$$C(G) = \{c \in \mathbb{F}_2^n \mid \forall u \in [m], \sum_{i \in N(u)} c_i = 0\}$$

where $N(u) = \{v \in [n] \mid uv \in E\}$ is the set of neighbours of u.

We cover some basic graph definition below. As we shall see, good expansion property of the graph implies good parameters for the induced LDPC code.

Basic Graph Definition

- *Regular graph*: we say a graph is D-regular if every vertex has degree d. For a bipartite graph $G = (L, R, E)$, we say it is D-left regular if every vertex in L has degree D.
- *Neighbourhood:* for any $S \subseteq V$, denote $N(S) = \{u \in V \mid \exists v \in S, uv \in E\}$ as the neighbourhood of S.

- *Unique neighbour*: for any $S \subseteq V$, we say $v \in V$ is a unique neighbour of S if $|N(v) \cap S| = 1$, i.e. there is an unique edge from S mapping to v. We denote $U(S)$ as the set of all unique neighbours of S.
- *1-sided small set vertex expansion:* for any bipartite graph $G = (L, R, E)$, we say G has (γ, α) 1isded small set vertex expansion if for all $S \subseteq L$, $|S| \leq \gamma|L| \Rightarrow |N(S)| \geq \alpha|S|$.

Bipartite graphs with 1-sided small set vertex expansion are referred to as bipartite expander graph here.

Definition 1.3 ($(n, m, D, \gamma, \alpha)$-Bipartite Expander Graph)**.** *For any bipartite graph G, we say G is a $(n, m, D, \gamma, \alpha)$-bipartite expander if $|L(G)| = n$, $|R(G)| = m$, G is D-left regular, and it has (γ, α) 1-sided small set vertex expansion.*

LDPC code defined on bipartite expander graph are called *expander code*.

For any D-left regular graph, α is trivially upper bounded by D. As it turns out, there exists graphs where all its small sets expands near perfectly, i.e. $\alpha \approx D$.[1]

Theorem 1.4 (Existence of Good $(n, m, D, \gamma, \alpha)$-Bipartite Expander Graph)**.** *$\forall \epsilon > 0$, $\forall m, n \in \mathbb{N}$, $m < n$, there exists $(n, m, D, \gamma, D(1-\epsilon))$-bipartite expander with $D = \Theta\left(\frac{\log \frac{n}{m}}{\epsilon}\right)$, $\gamma = \Theta\left(\frac{\epsilon m}{Dn}\right)$.*

We can think of m, n as being a constant factor away from each other. The set size bound $\gamma = \Theta\left(\frac{\epsilon m}{Dn}\right)$ is near optimal as any set of size $\approx \gamma n$ has neighbourhood of size $D(1-\epsilon) \cdot \gamma n \approx m$.

Remark 1.5 (History)**.** *Expander code was created by Sisper and Spielman in 1996. At the time, no explicit construction of good bipartite expander graph (1-sided lossless expander) was known. There were known construction of Ramanujan graphs, but Ramanujan graphs were shown to have small set vertex expansion $\approx \frac{D}{2}$ as shown by Kahale, which is not strong enough for good expander code. Capalbo, Reingold, Vadhan, Wigderson 2002 gave the first explicit construction of 1-sided lossless expander, i.e. $(n, m, D, \gamma, \alpha)$-bipartite expander for arbitrarily small $\epsilon > 0$. New constructions have been given for lossless expander in recent years. See Golowich 2024 for a new simple explicit construction of $(n, m, D, \gamma, \alpha)$-bipartite expander, and Hsieh, Lubotzky, Mohanty, Reiner, Zhang 2025 on an explicit construction of two sided lossless expanders that uses the cubical complex.*

For our construction of expander code, we want the small set expansion α to be very close D. One important reason is that this allows to argue that the small sets have many unique neighbours, thus giving a lower bound for the distance of the code.

Lemma 1.6 (Large Neighbour Expansion implies Many Unique Neighbours)**.** *If G is $(n, m, D, \gamma, D(1-\epsilon))$-bipartite expander where $\epsilon < \frac{1}{2}$, then $\forall S \subseteq L(G)$ where $|S| \leq \gamma n$, $|U(S)| \geq D(1-2\epsilon) \cdot |S|$.*

Proof. Since $|N(S)| \geq D(1-\epsilon) \cdot |S|$ and S has $D|S|$ edges going out, there are at most $\epsilon \cdot D|S|$ vertices in $N(S)$ that has ≥ 2 edges mapping to it, which implies $|U(S)| \geq D(1-2\epsilon) \cdot |S|$. □

[1] Such graphs are more commonly known as lossless expanders in the literature.

We present two ways of lower bounding the distance of expander code by unique neighbour expansion.

Corollary 1.7 (Lower Bounding Distance by One Unique Neighbour)**.** *Let G be a $(n, m, D, \gamma, D(1-\epsilon))$-bipartite expander where $\epsilon < \frac{1}{2}$. Then, the minimum distance in the code $C(G)$ has $\Delta(C(G)) > \gamma n$.*

Proof. For any $S \subseteq L(G) = [n]$ with size $|S| \leq \gamma n$, $|U(S)| \geq D(1-2\epsilon) \cdot |S| > 0$. Let $v \in U(S)$ be one such unique neighbour, then, the word corresponding to S in $\mathbb{F}_2^n$ does not satisfy the parity check on v. Thus, no word of length $\leq \gamma n$ is in $C(G)$, this implies $\Delta(C(G)) > \gamma n$. □

Note that the above proof only needs the existence of one unique neighbour. The below proof fully utilizes $|U(Q)| \geq D(1-2\epsilon)\gamma n$ for small sets Q, thus obtaining a better distance bound.

Theorem 1.8 (Better Distance by Unique Neighbour Expansion)**.** *Let G be a $(n, m, D, \gamma, D(1-\epsilon))$-bipartite expander where $\epsilon < \frac{1}{2}$. Then, $\Delta(C(G)) > 2\gamma(1-\epsilon)n$.*

Proof. It suffices to show that for $S \subseteq L(G) = [n]$ with size $\gamma n < |S| < 2\gamma(1-\epsilon)n$, it has a unique neighbour.

Let $Q \subseteq S$ has size $|Q| = \gamma n$. Then, $|U(Q)| \geq D(1-2\epsilon)\gamma n$. Since $|S \setminus Q| < \gamma n(1-2\epsilon)$ and the graph is D-left regular, we have $|N(S \setminus Q)| < D \cdot \gamma n(1-2\epsilon)$. This means that there exists a vertex $v \in U(Q)$ that is not in $N(S \setminus Q)$. Thus, S has a unique neighbour, this concludes the proof. □

We exhibit our main result here.

Theorem 1.9 (Expander Codes are Asymptotically Good Code)**.** *For any $(n, m, D, \gamma, D(1-\epsilon))$-bipartite expander, the code $C(G)$ has rate $1 - \frac{m}{n}$ and distance $2\gamma(1-\epsilon)n$.*

Proof. $C(G)$ has rate $1 - \frac{m}{n}$ since it has blocklength n and m linear constraints. The distance follows by Theorem 1.8 and the existence of $(n, m, D, \gamma, D(1-\epsilon))$-bipartite expander follows by Theorem 1.4. □

Expander code also has a fast and natural decoding algorithm. Given a word $w \in \mathbb{F}_2^n$, the idea is to simply keep flipping bits in w as long it decreases the number of errors.

Algorithm 1 Flip Algorithm for Expander Codes

Require: A parity check matrix $H \in \mathbb{F}_2^{m \times n}$ and a word $w \in \mathbb{F}_2^n$.
  while $|Hx| \neq 0$ **do**
    Find a bit $i \in [n]$ such that flipping it decreases $|Hx|$
  end while

We show that this decoding algorithm can decode from $\approx \frac{1}{2} \cdot \Delta(C(G))$ errors for $(n, m, D, \gamma, D(1-\epsilon))$-bipartite expander where $\epsilon < \frac{1}{4}$.

Theorem 1.10. *Let G be a $(n, m, D, \gamma, D(1-\epsilon))$-bipartite expander where $\epsilon < \frac{1}{4}$, and let $w \in \mathbb{F}_2^n$ be a word such that its distance to a codeword $y \in C$ is at most $(1-2\epsilon) \cdot \gamma n$, then the decoding algorithm above returns y in at most m iteration.*

Proof. Let $w^{(t)}$ denote w in iteration t. We first show that as long as $\Delta(w^{(t)}, y) < \gamma n$, there exists some bit i such that flipping it decreases the number of unsatisfied constraints.

Let $S := \{i \in [n] \mid w_i^{(t)} \neq y_i\}$ be the set of errors at iteration t. Since $|S| < \gamma n$, then $|U(S)| \geq (1-2\epsilon)D|S| > \frac{D}{2}|S|$, where the second inequality follows by $\epsilon < \frac{1}{4}$. This implies there must exist a $v \in S$ with $> \frac{D}{2}$ unsatisfied constraints. Flipping v thus decreases the total bumber of unsatisfied constraints.

We now show that if we start with some $w \in \mathbb{F}_2^n$ with $(1-2\epsilon) \cdot \gamma n$ errors, then during the algorithm, $\Delta(w^{(t)}, y) < \gamma n$.

Let $S := \{i \in [n] \mid w_i^{(t)} \neq y_i\}$ be the set of errors at iteration t. Suppose $|S| = \gamma n$ for contradiction (in general, if $|S| \geq \gamma n$, then since S changes by 1 each iteration, there must exist an iteration where $|S| = \gamma n$), then $|U(S)| > (1-2\epsilon)D\gamma n$, which means there are more than $(1-2\epsilon)D\gamma n$ unsatisfied constraints for $w^{(t)}$ (note that $Hw = Hw + Hy = H\chi_S$). However, we start initially with at most γn errors, which means there were at most $(1-2\epsilon)D \cdot \gamma n$ unsatisfied constraints. Since the number of unsatisfied constraints is strictly decreasing, this is a contradiction.

Therefore, if we start with some $w \in \mathbb{F}_2^n$ where $\Delta(w, C(G)) \leq (1-2\epsilon) \cdot \gamma n$, the number of unsatisfied constraints decrease at least by 1 each iteration. Since there are at most m unsatisfied constraints, the decoding algorithm terminates in m steps. Since the distance of the code is $2\gamma(1-\epsilon)n > \gamma n$, the codeword we obtained through this decoding procedure must be y. □