Quantum Information Theory - 67749
Recitation 3, June 3, 2025

1 Recitation Overview - Quantum Error Correction Codes.

In the last lectures, we saw that quantum states can be considered as resources, presenting the quantum AEP concept and the compression theorem. In this recitation, we continue that line. We'll start by completing the compression theorem proof. Then we will present (and represent) simple applications of quantum states, including going over superdense decoding and quantum teleportation again, and presenting gate teleportation and quantum money. Then we will strengthen our understanding of mixed states by presenting and proving Uhlmann's theorem. Finally, we will discuss EPR distillation (entanglement purification), a process by which one tries to extract from a noisy EPR mixture much better EPR pairs.

2 Codes in General. Notations and Definitions.

Here we focus only on linear binary codes, which one could think about as linear subspaces of $\mathbb{F}_2^n$. A common way to measure resilience is to ask how many bits an evil entity needs to flip such that the corrupted vector will be closer to another vector in that space than the original one. Those ideas were formulated by Hamming [Ham50], who presented the following definitions.

Definition 2.1. *Let $n \in \mathbb{N}$ and $\rho, \delta \in (0,1)$. We say that C is a **binary linear code** with parameters $[n, \rho n, \delta n]$. If C is a subspace of $\mathbb{F}_2^n$, and the dimension of C is at least ρn. In addition, we call the vectors belong to C codewords and define the distance of C to be the minimal number of different bits between any codewords pair of C.*

From now on, we will use the term code to refer to linear binary codes, as we don't deal with any other types of codes. Also, even though it is customary to use the above parameters to analyze codes, we will use their percent forms called the relative distance and the rate of code, matching δ and ρ correspondingly.

Definition 2.2. *A **family of codes** is an infinite series of codes. Additionally, suppose the rates and relative distances converge into constant values ρ, δ. In that case, we abuse the notation and call that family of codes a code with $[n, \rho n, \delta n]$ for fixed $\rho, \delta \in [0, 1)$, and infinite integers $n \in \mathbb{N}$.*

Notice that the above definition contains codes with parameters attending to zero. From a practical view, it means that either we send too many bits, more than a constant amount, on each bit in the original message. Or that for big enough n, adversarial, limited to changing only a constant fraction of the bits, could disrupt the transmission. That distinction raises the definition of good codes.

Definition 2.3. *We will say that a family of codes is a **good code** if its parameters converge into positive values.*

3 Singleton Bound

To get a feeling of the behavior of the distance-rate trade-of, Let us consider the following two codes; each demonstrates a different extreme case. First, define the repetition code $C_r \subset \mathbb{F}_2^{n \cdot r}$, In which, for a fixed integer r, any bit of the original string is duplicated r times. Second, consider the parity check code $C_p \subset \mathbb{F}_2^{n+1}$, in which its codewords are only the vectors with even parity. Let us analyze the repetition code. Clearly, any two n-bits different messages must have at least a single different bit. Therefore their corresponding encoded codewords have to differ in at least r bits. Hence, by scaling r, one could achieve a higher distance as he wishes. Sadly the rate of the code decays as $n/nr = 1/r$. In contrast, the parity check code adds only a single extra bit for the original message. Therefore scaling n gives a family which has a rate attends to $\rho \to 1$. However, flipping any two different bits of a valid codeword is conversing the parity and, as a result, leads to another valid codeword.

To summarize the above, we have that, using a simple construction, one could construct the codes $[r, 1, r]$, $[r, r-1, 2]$. Each has a single perfect parameter, while the other decays to the worst.

Besides being the first bound, Singleton bound demonstrates how one could get results by using relatively simple elementary arguments. It is also engaging to ask why the proof yields a bound that, empirically, seems far from being tight.

Theorem (Singleton Bound.)**.** *For any linear code with parameter $[n, k, d]$, the following inequality holds:*

$$k + d \leq n + 1$$

Proof. Since any two codewords of C differ by at least d coordinates, we know that by ignoring the first $d-1$ coordinate of any vector, we obtain a new code with one-to-one corresponding to the original code. In other words, we have found a new code with the same dimension embedded in $\mathbb{F}_2^{n-d+1}$. Combine the fact that dimension is, at most, the dimension of the container space, we get that:

$$\dim C = 2^k \leq 2^{n-d+1} \Rightarrow k + d \leq n + 1$$

□

It is also well known that the only binary codes that reach the bound are: $[n, 1, n]$, $[n, n-1, 2]$,$[n, n, 1]$ [AF22]. In particular, there are no good binary codes that obtain equality(And no binary code which get close to the equality exits). Let's review the polynomial code family [RS60], which is a code over none binary field that achieve the Singleton Bound.

Next, we will review Tanner's construction, that in addition to being a critical element to our proof, also serves as an example of how one can construct a code with arbitrary length and positive rate.

4 Tanner Code

The constructions require two main ingredients: a graph Γ, and for simplicity, we will restrict ourselves to a Δ regular graph, Yet notice that the following could be generalize straightforwardly for graphs with degree at most Δ. The second ingredients is a ;small' code C_0 at length equals the graph's regularity, namely $C_0 = [\Delta, \rho\Delta, \delta\Delta]$. We can think about any bit string at length Δ as an assignment over the edges of the graph. Furthermore, for every vertex $v \in \Gamma$, we will call the bit string, which is set on its edges, the local view of v. Then we can define, [Tan81]:

Definition 4.1. *Let $C = \mathcal{T}(\Gamma, C_0)$ be all the codewords which, for any vertex $v \in \Gamma$, the local view of v is a codeword of C_0. We say that C is a **Tanner code** of Γ, C_0. Notice that if C_0 is a binary linear code, So C is.*

Example 4.1. *Consider the Petersen graph Γ, which is a regular graph with degree 3. Let C_0 be the set of all words with even parity. It follows that C_0 contains all even-length binary strings of length 3: 000, 110, 101, and 011. However, the size of $\mathcal{T}(\Gamma, C_0)$ is significantly larger, as shown in Figure fig. 1. Specifically, any rotation of the inner and outer cycles simultaneously gives rise to another valid codeword, so any assignments that are not invariant under these rotations would produce five additional valid codewords.*

Lemma 4.1. *Tanner codes have a rate of at least $2\rho - 1$.*

Proof. The dimension of the subspace is bounded by the dimension of the container minus the number of restrictions. So assuming non-degeneration of the small code restrictions, we have that any vertex count exactly $(1-\rho)\Delta$ restrictions. Hence,

$$\dim C \geq \frac{1}{2}n\Delta - (1-\rho)\Delta n = \frac{1}{2}n\Delta(2\rho - 1)$$

Clearly, any small code with rate $> \frac{1}{2}$ will yield a code with an asymptotically positive rate □

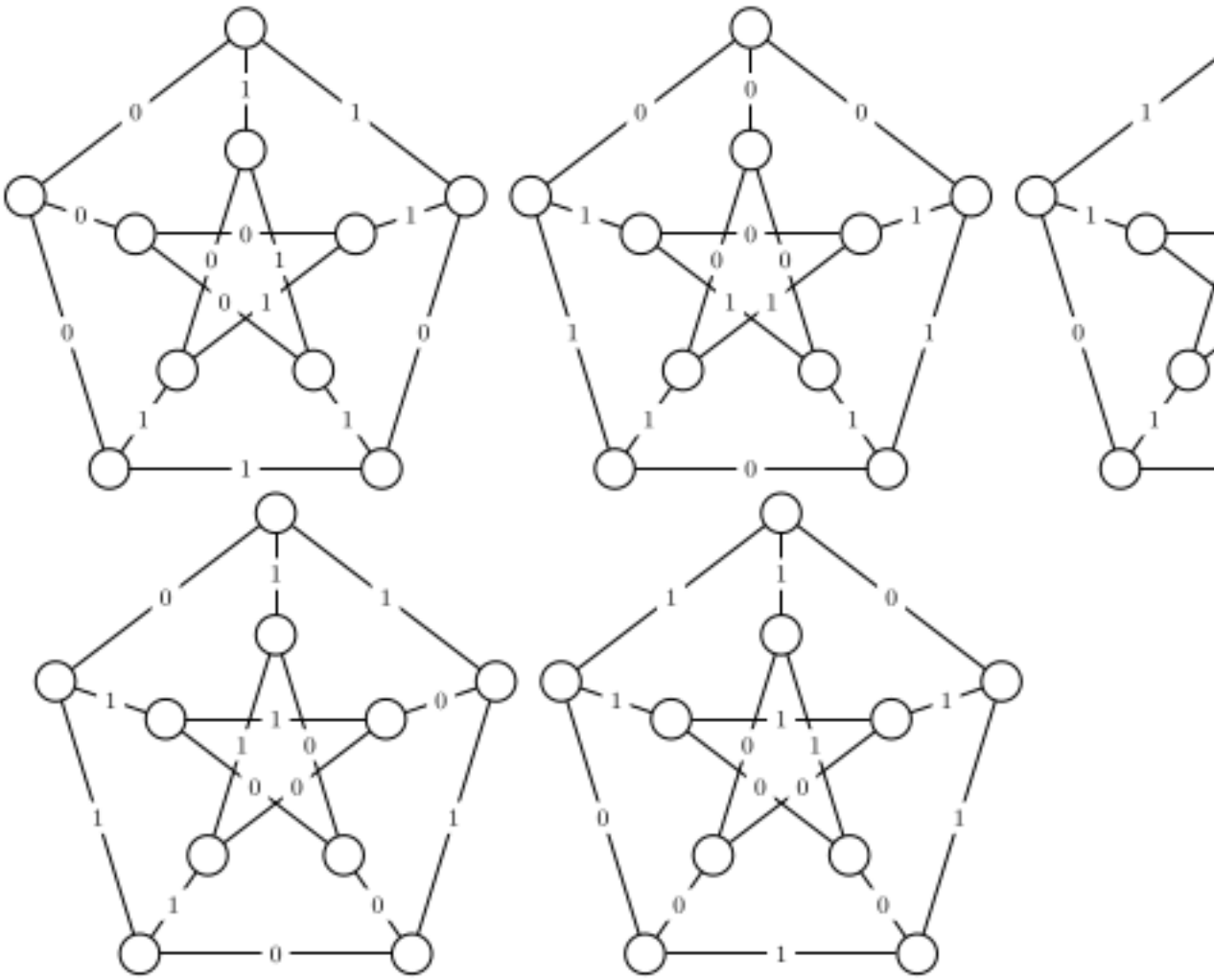

Figure 1: Peterson Graph.

Based on Lemma 4.1, we can obtain a recipe for constructing codes with a almost non-vanishing rate for arbitrarily large lengths and dimensions. This recipe involves concatenating a series of Tanner codes over complete graphs. To be more precise, we can define a family of codes as follows:

$$\begin{aligned} C_{i+1} &= \mathcal{T}\left(K_{n(C_i)+1}, C_i\right) \\ C_0 &= \text{ Some simple } \Delta[1, \rho_0, \delta_0] \text{ code.} \end{aligned}$$

Where $n(C_i)$ represents the code length of the ith code. Repeating the process described above $\log^*_\Delta(n)$ times allows us to extend the initial code $\Delta[1, \rho_0]$ to $n[1, \sim 2\rho^{\log^*_\Delta(n)}]$. Interestingly, any family of finite groups generated by a constant-size generator set can define a family of codes by utilizing their Cayley graphs as a basis for Tanner codes.

Once we have seen that Tanner codes enable us to achieve rates, the next natural question to ask is about the distance of the codes. Achieving a linear distance requires

a little bit more from the graphs, but to understand this idea better, let us return to the repetition code. For instance, the repetition code can be presented as a Tanner code over the cycle graph.

Example 4.2. *In this representation, each vertex checks if the bits on its edges are equal. A valid codeword is an assignment in which all the bits are equal, since otherwise, there would be an edge with no supporting vertex. An illustration of a legal assignment is provided in section 4.*

Recall that the distance of a linear code is the minimal weight of the non-zero codewords. Consider a codeword $c \in C$ and group the vertices by four sets V_i such that V_i is the set of vertices that see $i \in \{0,1\}^2$. Since $c \in C$, we have that $|V_{10}| = |V_{01}| = 0$. Additionally, any vertex in V_{00} is not connected to V_{11}, which gives us two possible cases: either all the vertices in V_{11} are isolated, or the graph is not connected. Hence, the distance of the code is equal to $\frac{1}{2}\sum |V_i| \cdot |i| = \frac{1}{2} 2 \cdot n = n$.

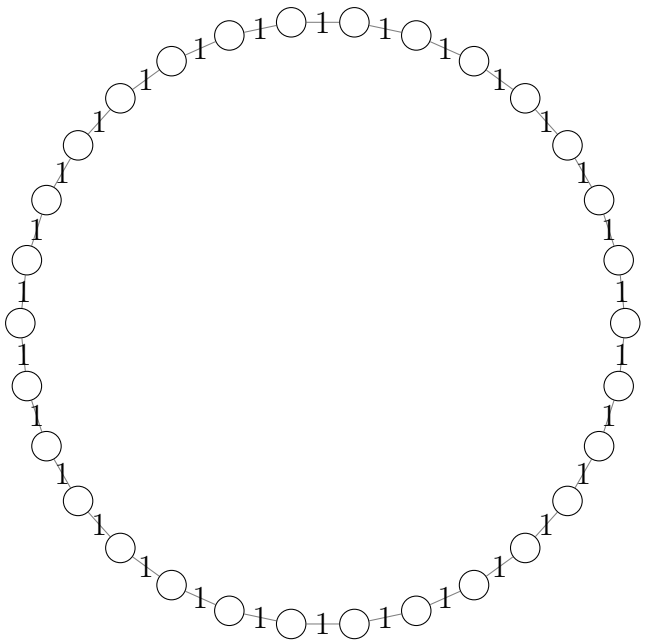

Figure 2: The 1^n assignment on the cycle graph. Any vertex compute parity $1+1 = 0$, therefore all the restrictions are satisfied and $1^n \in \mathcal{T}$ (cycle , parity).

It is worth mentioning that, in the literature, the repetition code is not usually given as an example of a Tanner code. However, this example will come up again later in the chapter on quantum codes, when we discuss the Toric code, its relation to the hyperproduct code, and how it can be seen as a hyperproduct of two cycle codes.

Furthermore, analyzing the repetition code gives a clue as to how, in certain cases, one might prove a lower bound on the code distance. We would like to say that, if the weight of the code word is below the distance, then it must be that there is at least one vertex that has a non-trivial local view which is not a codeword in

C_0. Put differently, we cannot spread a small weight codeword over $\{V_i\}$, defined above, without expanding into subsets corresponding to low $|i|$. Next, we are going to present the Expander codes, which are Tanner codes constructed from graphs with good algebraic expansion.

5 Expander Codes

We saw how a graph could give us arbitrarily long codes with a positive rate. We will show, Sipser's result [SS96] that if the graph is also an expander, we can guarantee a positive relative distance.

Definition 5.1. *Denote by λ the second eigenvalue of the adjacency matrix of the Δ-regular graph. For our uses, it will be satisfied to define expander as a graph $G = (V, E)$ such that for any two subsets of vertices $T, S \subset V$, the number of edges between S and T is at most:*

$$| E(S,T) - \frac{\Delta}{n}|S||T| | \leq \lambda\sqrt{|S||T|}$$

This bound is known as the Expander Mixining Lemma. We refer the reader to [HLW06] for more deatilied survery.

Theorem. *Theorem, let C be the Tanner Code defined by the small code $C_0 = [\Delta, \delta\Delta, \rho\Delta]$ such that $\rho \geq \frac{1}{2}$ and the expander graph G such that $\delta\Delta \geq \lambda$. C is a good LDPC code.*

Proof. We have already shown that the graph has a positive rate due to the Tunner construction. So it's left to show also the code has a linear distance. Fix a codeword $x \in C$ and denote By S the support of x over the edges. Namely, a vertex $v \in V$ belongs to S if it connects to nonzero edges regarding the assignment by x, Assume towards contradiction that $|x| = o(n)$. And notice that $|S|$ is at most $2|x|$, Then by The Expander Mixining Lemma we have that:

$$\begin{aligned}\frac{E(S,S)}{|S|} &\leq \frac{\Delta}{n}|S| + \lambda \\ &\leq_{n\to\infty} o(1) + \lambda\end{aligned}$$

Namely, for any such sublinear weight string, x, the average of nontrivial edges for the vertex is less than λ. So there must be at least one vertex $v \in S$ that, on his local view, sets a string at a weight less than λ. By the definition of S, this string cannot be trivial. Combining the fact that any nontrivial codeword of the C_0 is at weight at least $\delta\Delta$, we get a contradiction to the assumption that v is satisfied, videlicet, x can't be a codeword □

Now, we have an explicit construction for good classical LDPC codes. Next, we present a randomized construction which gives us non-LDPC, but good, codes of arbitrary length. We have to present this construction rather than stopping here, as for proving good quantum codes, we have to rely on assumptions about the small code which hold only for codes of sufficient length.

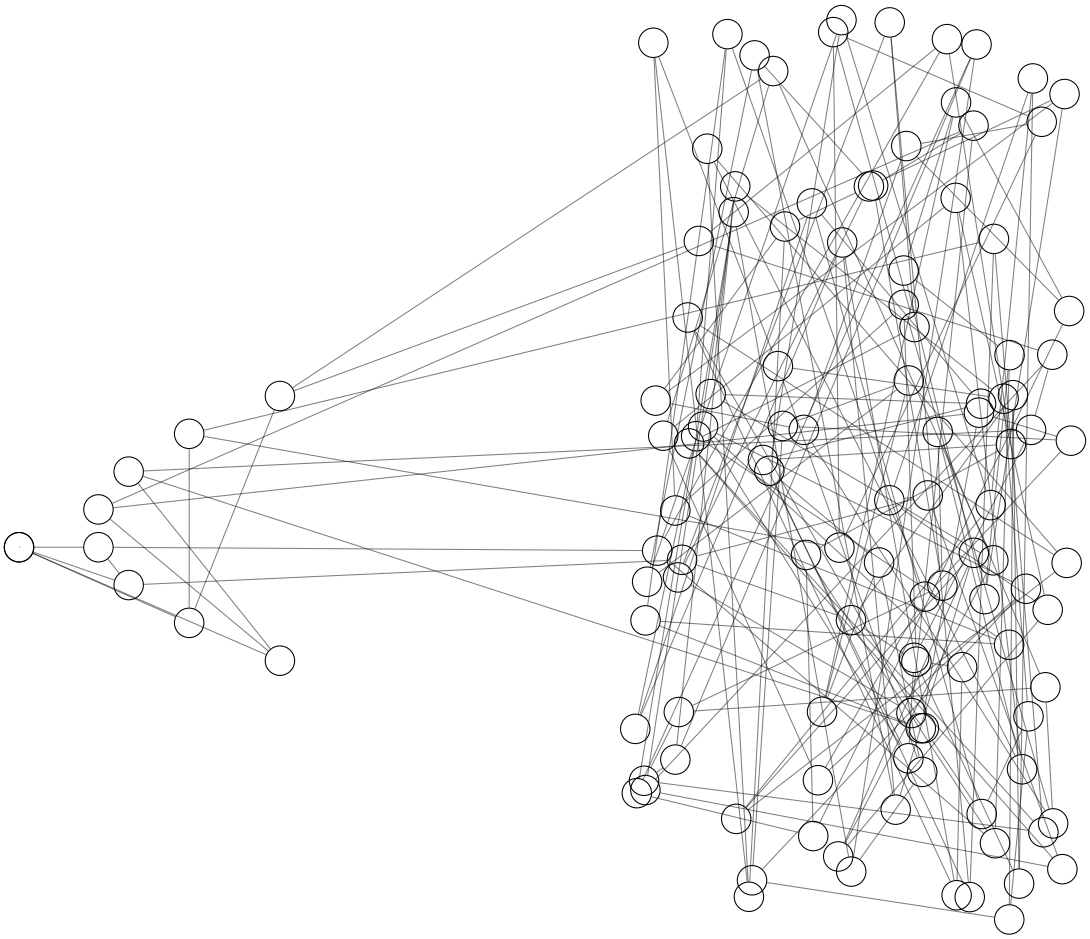

Figure 3: Expander Graph. Any small set of vertices, like the isolated set on the left, has more edges leaving it than intermediate edges. The exact amount is controlled by the expansion factor.

6 Randomized Constructions.

Randomized constraction reffers to taking as eithr generator or pairty check matrix a matrix that sampled uniformly random. That is, we are going show that if the generator matrix A is a random matrix in $M(\mathbb{F}_2^{k\times n})$ then $\operatorname{Im} A$ is a good code.

Claim 6.1. *Let $n, k \in \mathbb{N}$ such $k = \Theta(n)$ and let A be a random matrix in $M(\mathbb{F}_2^{k\times n})$ then the code generated by A is good code with high probability.*

Let's start with the following claim:

Claim 6.2. *Let A be a random matrix in $M(\mathbb{F}_2^{k\times n})$. For any non-zero $x \in \mathbb{F}$, we have that Ax is distributed uniformly.*

Proof. By the fact that $x \neq 0$, there exists at least one coordinate $i \in [k]$ such that

$x_i \neq 0$. Thus, we have

$$\begin{aligned}(Ax)_j &= \sum_k A_{jk}x_k = \sum_{i\neq k} A_{jk}x_k + A_{ji}x_i \\ &= \sum_{i\neq k} A_{jk}x_k + A_{ji}\end{aligned}$$

Notice that due to the fact that $\mathbb{F}_2$ is a field, there is exactly one assignment that satisfies the equation conditioned on all the values A_{jk} where $j \neq k$.

$$\begin{aligned}\mathbf{Pr}\left[(Ax)_j = 1\right] &= \sum_{A_{jk};k\neq i} \mathbf{Pr}\left[(Ax)_j = 1 \mid A_{jk}; k \neq i\right] \mathbf{Pr}\left[A_{jk}; k \neq i\right] \\ &= \frac{1}{2}\end{aligned}$$

Therefore, any coordinate of Ax is distributed uniformly $\Rightarrow$ Ax is distributed uniformly. □

proof of 6.1. By the uniformity of Ax, we obtain that the expected Hamming weight of Ax is:

$$\mathbf{E}\left[|Ax|\right] = \mathbf{E}\left[\sum_i^n (Ax)_i\right] = \frac{1}{2}n$$

As the coordinates of A_x are independent (each row of A is sampled separately), we can use the Hoff's bound to conclude that:

$$\mathbf{Pr}\left[\,||Ax| - \mathbf{E}\left[|Ax|\right]| \geq \left(\frac{1}{2} - \delta\right) n\right] \leq e^{-n\left(\frac{1}{2}-\delta\right)^2}$$

Now, we will use the union bound to show that any $x \in \mathbb{F}_2^k$, Ax is of weight at least δ.

$$\mathbf{Pr}\left[|Ax| \geq \delta : \forall x \in \mathbb{F}_2^k\right] \geq 1 - |\mathbb{F}_2^k| \cdot e^{-n\left(\frac{1}{2}-\delta\right)^2}$$

Let $k = \rho n$ and note that the statement holds when $\rho \geq \left(\frac{1}{2} - \delta\right)^2$. In other words, for any $k = \Theta(n)$, there exists a positive δ such that with high probability, the distance of Im is at least δn. □

References

[Ham50] R. W. Hamming. "Error detecting and error correcting codes". In: *The Bell System Technical Journal* 29.2 (1950), pp. 147–160. DOI: 10.1002/j.1538-7305.1950.tb00463.x.

[RS60] Irving S. Reed and Gustave Solomon. “Polynomial Codes Over Certain Finite Fields”. In: *Journal of The Society for Industrial and Applied Mathematics* 8 (1960), pp. 300–304.

[Tan81] R. Tanner. “A recursive approach to low complexity codes”. In: *IEEE Transactions on Information Theory* 27.5 (1981), pp. 533–547. DOI: 10.1109/TIT.1981.1056404.

[SS96] M. Sipser and D.A. Spielman. “Expander codes”. In: *IEEE Transactions on Information Theory* 42.6 (1996), pp. 1710–1722. DOI: 10.1109/18.556667.

[HLW06] Shlomo Hoory, Nathan Linial, and Avi Wigderson. “Expander graphs and their applications”. In: *Bulletin of the American Mathematical Society* 43.4 (2006), pp. 439–561.

[AF22] “Maximum distance separable (MDS) code”. In: *The Error Correction Zoo*. Ed. by Victor V. Albert and Philippe Faist. 2022. URL: https://errorcorrectionzoo.org/c/mds.