

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.

Part I

Classical Error Correction Code.

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 [Hamming], 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-off, 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 $\rho \rightarrow 1$. 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: $[n, 1, n]$, $[n, n-1, 2]$, $[n, n, 1]$ [eczoo'mds]. In particular, there are no good binary codes that obtain equality (And no binary code which get close to the equality exists). Let's review the polynomial code family [Reed1960PolynomialCO], 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, [Tanner]:

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 □

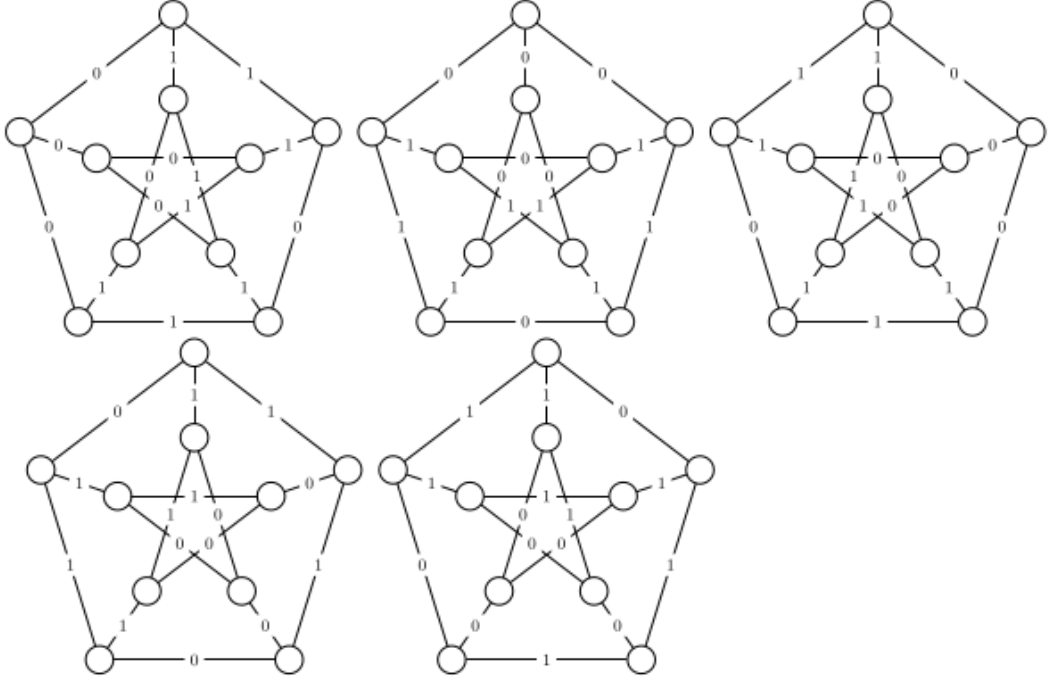


Figure 1: Peterson Graph.

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:

$$C_{i+1} = \mathcal{T}(K_{n(C_i)+1}, C_i)$$

$$C_0 = \text{Some simple } \Delta[1, \rho_0, \delta_0] \text{ code.}$$

Where $n(C_i)$ represents the code length of the i th 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 $\frac{1}{2} \sum |V_i| \cdot |i| = \frac{1}{2} 2 \cdot n = n$.

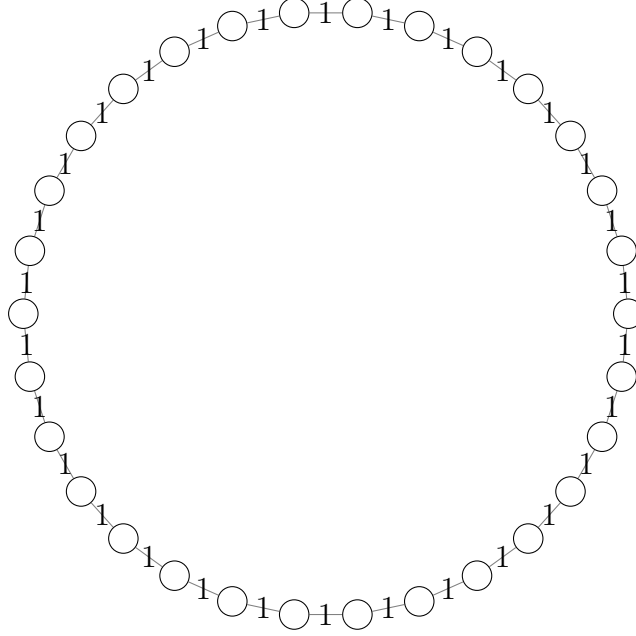


Figure 2: The 1^n assignment on the cycle graph. Any vertex compute parity $1+1 = 0$, therefore all the restrictions are satisfied and $1^n \in \mathcal{T}(\text{cycle}, \text{parity})$.

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.

Part II

Quantum Error Correction Codes.

5 Introduction.

It is widely believed that quantum machines have a significant advantage over classical machines in a wide range of computational tasks [grover1996fast], [ahuja1999quantum]. Simple algorithms which can be interpreted as the quantum version of scanning all the options, reducing the running time by the square root of the classical magnitude. Nevertheless, Shor has demonstrated a polynomial-depth quantum circuit that solves the hidden abelian subgroup problem [Shor'1997], which is considered a breakthrough, as it made the computer science community believe that a quantum computer could offer an exponential advantage.

Despite a consensus on the superiority of the ideal quantum computing model, it is still uncertain whether it is possible to implement such a machine in a noisy environment. Still, simply pointing out the existence of noise is not sufficient to negate the feasibility of computation. Evidence of this is that classical computers also experience a certain rate of errors. Therefore, to fully comprehend the difficulty, let us compare two main factors that made it a challenging task.

First is the magnitude of the error rate; classical computers also have errors, and sometimes we witness system failures (e.g. the blue screen). The error rate of modern computers is so low that the probability of errors propagating stays negligible, even if the length of the computation is polynomial in the scale of what is considered a reasonable input size. It's worth mentioning that in exascale computing, when supercomputers perform around 10^{18} operations per second, it is difficult to overlook faults. In quantum computing, we become aware of their existence much earlier.

The second difference, which is a subtle point, is that quantum states are susceptible to additional types of errors. In addition to the possibility of bit-flip errors, a quantum state may also experience a change in phase. For example, consider the initial state $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, and suppose that due to noise the state is transformed into $\frac{1}{\sqrt{4}}(\sqrt{3}|0\rangle + |1\rangle)$. Classical circuits are oblivious to such faults, meaning that their operation would remain unchanged as no error has occurred. Quantum circuits, however, are usually affected and may fail. Furthermore, when designing a decoder for quantum error correction codes, one must ensure that the decoding process does not introduce bit-flip errors if a classical code is used to protect against phase flips.

6 Quantum Noise.

Definition 6.1 (Bit and phase flip.). *Consider a quantum state $|\psi\rangle$ encoded in the computation base. We will say that a bit-flip occurs in a scenario the operator Pauli X is applied on one of our state's qubits. The bit-flip event could be considered as exactly as the standard bit-flip error in the classical regime. Similarly, phase-flip*

occurs when the Pauli Z is applied on one of the qubits.

However, even though quantum noise is so violent, it has been proven that any ideal circuit of polynomial depth can be transformed into a robust circuit at poly-logarithmic cost [aharonov1999faulttolerant]. In other words, there is a threshold: if physicists provide qubits and a finite gate set with a noise rate below that threshold, then BQP, the class of polynomial-time ideal quantum computation, is feasible and can be computed on a realistic machine.

The basic idea in [aharonov1999faulttolerant] was to show the existence of quantum error correction codes, which would enable logic operations to be performed in a way that prevents errors from propagating. This process involves separating the operation into two stages: the operation itself and an error correction stage. This comes with an additional cost in terms of space and time, but it can reduce the probability of the final state being faulty. The trade-off between the resources needed and the rate of decrease defines the threshold. If the balance is positive, then the process can be repeated in a recursive manner, and after log-log iterations, the failure probability will decay to zero. At the same time, the circuit will scale to a maximum of poly-logarithmic width and depth.

Let's return to the repetition code presented in Chapter 2. We would like to have an analog; a first and natural attempt might consider duplicating copies of the state. Unfortunately, copying a general state is not a linear operation and therefore cannot be done in the circuit model (or any other believed to be feasible). In particular, there is no circuit U which can simultaneously duplicate the states $|0\rangle, |1\rangle, |+\rangle, |-\rangle$.

To overcome the issue, Shor came up with the nine-qubit code [Ninequ], which at first glance might seem a naive straightforward implementation of “duplication”, but instead uses a clever insight about quantumness in general. Any operation can be seen as a linear (and even unitary) operation over a subspace embedded in a large enough dimension. The encoding is given as follows:

$$\begin{aligned} |\bar{0}\rangle &= \frac{1}{2\sqrt{2}} (|000\rangle + |111\rangle)^{\otimes 3} \\ |\bar{1}\rangle &= \frac{1}{2\sqrt{2}} (|000\rangle - |111\rangle)^{\otimes 3} . \end{aligned}$$

For convenience, let us use the notation $|\mathbf{GHZ}^\pm\rangle = |0^m\rangle \pm |1^m\rangle$. We can also consider the Shor code over m^2 qubits, which is defined as above, such that any logical state contains m products over m qubits. Therefore, the state $|\bar{0}\rangle$ over m^2 qubits can be written as $|\mathbf{GHZ}^+\rangle^m$. We are now ready to prove a statement regarding the robustness.

Lemma 6.1. *The Shor code over 9 qubits enable to correct a single either bit or phase flip.*

It is evident that a single bit-flip error can be handled in the same way as in the conventional case. The decoder will check if any of the triples have the same value, and if not, it will correct it by majority. To create a decoder that can also correct a phase-flip error, we need the following statement. In this chapter, we denote the Hadamard gate over m qubits as H^m .

Claim 6.1. $H^m |\mathbf{GHZ}^\pm\rangle = \sum_{x \cdot \mathbf{1} = 2^\pm} |x\rangle$

Proof.

$$\begin{aligned} H^m |\mathbf{GHZ}^\pm\rangle &= H^m |0^m\rangle \pm H^m |1^m\rangle = \sum_{x \in \mathbb{F}_2^m} |x\rangle \pm \sum_{x \in \mathbb{F}_2^m} (-1)^{x \cdot \mathbf{1}} |x\rangle \\ &= \sum_{x \in \mathbb{F}_2^m} (1 \pm (-1)^{x \cdot \mathbf{1}}) |x\rangle = \sum_{x \cdot \mathbf{1} = 2^\pm} |x\rangle \end{aligned}$$

□

Now it is clear how to correct a phase flip. One can apply the Hadamard transform and compute the parity of each triple. By the assumption that only a single phase flip may occur, either all the triples have the same parity or the faulted one has an opposite parity and needs to be corrected. Thus, we obtain an $[[9, 1, 3]]$ quantum error correction code. Asymptotically, this is an $[[m^2, 1, m]]$ code.

7 CSS Codes.

The Shor code is a specific case of the more general CSS (Calderbank-Shor-Steane) code [Calderbank'1996]. A family composed by two binary codes C_X, C_Z such that $C_Z^\perp \subset C_X$.

8 qLDPC Codes.

As exactly as in the classic case, qLDPC codes are codes in which any check act non trivially on at most a constant number of qubits, It was proved that using a good Quantum LDPC code one can achieve a fault tolerance threshold theorem at the cost of only constant overhead¹ [gottesman2014faulttolerant]. We are now about to embark on a detailed review of the first quantum LDPC code [Dennis'2002].

Recall that one way to present a code is by define the parity check matrix, Consider the $l \times l$ Tours, namely the Cayley graph of the group product $\mathbb{Z}_l \times \mathbb{Z}_l$. Associate any coordinate (bit/qubit) with an edge on the Tours. And consider the following two restrictions:

1. Each vertex requires form its local view, the bits lay on its supported edges, To has an even party. We will refer to this type of check as *cross check*.
2. Similarly, each face requires the same from its supported edges, but computes the parity in a different (specific) base. That it, the face first rotates the qubits by applying the Hadamard transform on them, and then computes their XOR. Finally, the qubits are rotated back to the computation base. We shall refer to this type of check as *face check*.

¹under the assumption of having an efficient decoder.

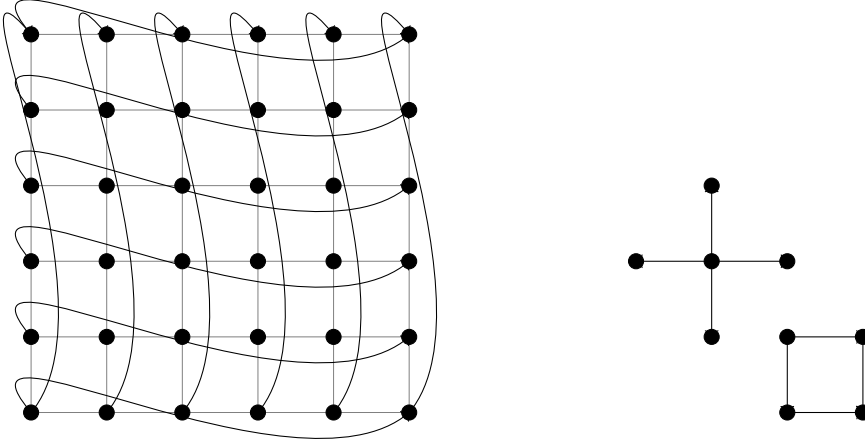


Figure 3: On the left is the Toric Graph. On the right are cross and face checks.

Claim 8.1. *The $l \times l$ Toric code is a CSS code, with dimension 2 and distance $\Theta(l)$.*

Proof. Consider a pair of cross and a face checks. If they are not intersect (share edges) then, obviously, they commute. So suppose they share an edge. Now there are a finite number of cases (4) in which a cross check can intersect a face, and for all of them we have they must to intersect in an exactly two edges. Therefore the crosses checks commute with all the faces checks.

Now, denote by C_X and C_Z the linear codes defined by the crosses and faces checks. Observe that C_X contain all the subgraph of the torus which have only even degree, namely all the loops. The following claim will be used to show that all the low weight loops (squeezable loops) are in C_Z^\perp .

Claim 8.2. *Let c be an assignment of ones on a unite square, namely a closed loop at length 4, then $c \in C_Z^\perp$.*

Proof. Denote by c' a codeword of C_Z , therefore the parity sum induced by c' on any square equals zero, Particularly the induced parity on the square supporting c is zero. But that parity is exactly $c \cdot c'$. As it true for any $c' \in C_Z$ we obtain that $c \in C_Z^\perp$. \square

Claim 8.3 (Veblen's theorem). *The set of even subgraph is a linear space spanned by simple cycles (vertices degree equals exactly 2).*

Proof. Let $P \subset E$ be an even subgraph then it must to have a simple cycle denote it by P' , now notice that P/P' is also an even subgraph. To see that consider a vertex v , As P' is simple cycle, substituting P' is either not effect the degree of v in P/P' or it decrees v degree by exactly 2 so $d_{P/P'}(v) = d_P(v) - 2$. In both cases $d_{P/P'}(v)$ is even. By repeating recursively until $P/P' = \{\emptyset\}$ we get a decomposition of P into a sum of simple cycles. \square

Claim 8.4. Associate with any vertex of the Torus a coordinate in $\mathbb{Z}_l \times \mathbb{Z}_l$. Consider a simple cycle P subset of the Torus. Denote P by the vertices composing it $v_0 v_1 \dots v_k$ arranged in order. Consider a vertex $v_i \in P$ and denote $v = (x, y) \in \mathbb{Z}_l \times \mathbb{Z}_l$. Then there exists a vertex $u \in P$, $u \neq v_{i-1} v_{i+1}$ that shares one of the coordinates of v . Put differently, there exists z such that either $u = (z, y)$ or $u = (x, z)$.

Proof. Assume without loss of generality that $v_{i-1} = (x, y - 1)$. Now denote by f the projection on the second coordinate, $f(a, b) = b$. and observe that for any j the distance $f(v_{j+1}) - f(v_j)$ satisfies:

$$f(v_{j+1}) - f(v_j) = \begin{cases} \pm(l-1) & v_j = (\cdot, l-1), v_{j+1} = (\cdot, 0) \\ \in \{\pm 1, 0\} & \text{else} \end{cases}$$

So if there is no $u \neq v_i$ such that $f(u) = f(v_i) = f(v_{i-1}) + 1$ then there must to be an edge $\{v_j, v_{j+1}\}$, $v_j = (\cdot, l-1)$, $v_{j+1} = (\cdot, 0)$ such P pass through it. So

$$\begin{aligned} f(v_{i-1}) &\leq f(v_{i+1}) + |f(v_{i+2}) - f(v_{i+1})| + \dots + -(l-1) \\ &\leq f(v_{i+1}) - |P| - (l-1) \end{aligned}$$

But $|P| \leq l$. □

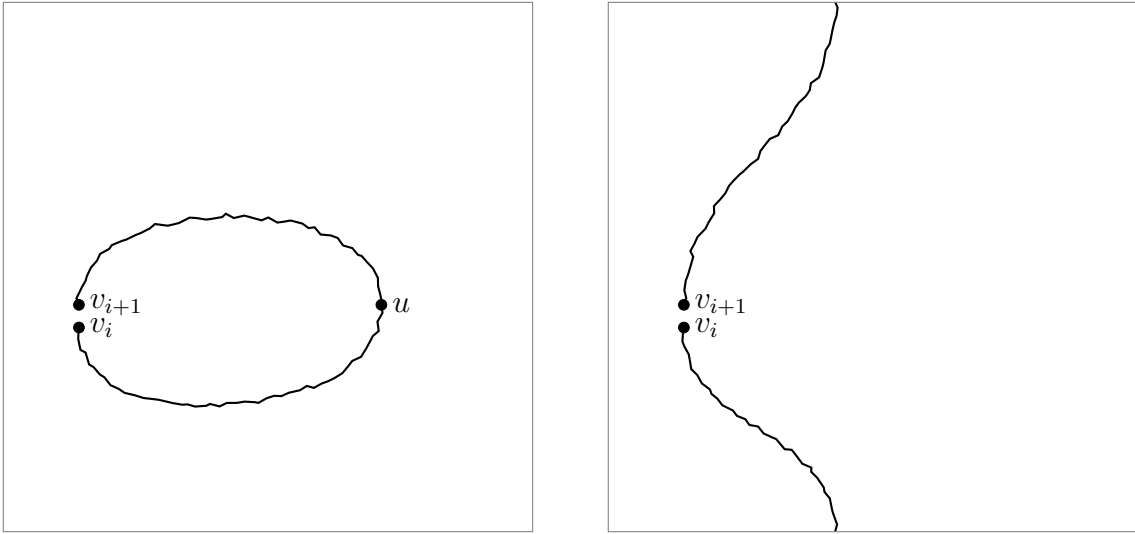


Figure 4: A simple cycle P that passes through a vertex u , with the same y -coordinate as v_i , on the left versus a cycle path for which there is no such u .

Definition 8.1. *Horizontal and Vertical diameters.* Let P be defined again as exactly as in ???. We will say that the horizontal diameter of P is:

$$\max_{v, u \in P} \min \{|v_x - u_x|, |v_x + l - u_x|\}$$

Similarly, we define the vertical diameter to be:

$$\max_{v, u \in P} \min \{|v_y - u_y|, |v_y + l - u_y|\}$$

Claim 8.5. *If P is a non empty simple cycle with vertical and horizontal diameters d_1, d_2 , then it can either be a square or can be decomposed into two simple cycles P_1, P_2 with either the vertical diameter of both of them being strictly less than d_1 and their horizontal diameters being at most d_2 , or the horizontal diameter of both of them being strictly less than d_2 and their vertical diameters being at most d_1 .*

Proof. If P is not a square, it must have either a vertical or horizontal diameter greater than 1. We will assume, without loss of generality, that the vertical diameter is greater than 1. Let $v_0 v_1 \dots v_k$ be the vertices that compose P , and let v_i, v_j be the vertices at maximal distance. We will pick the first vertex $v_{i'}$ from the left of v_i such that $v_{i',y} \neq v_{i,y}$. By ??, there must be at least one vertex v_l in P such that $v_{i,y} = v_{l,y}$ and is the closest in horizontal distance to v_i . We will denote by P^\uparrow the path $v_i, v_{i+1} \dots v_{l-1}, v_l$ and by P^\downarrow the path $v_l, v_{l+1} \dots v_{i-1}, v_i$. We will also denote by L the straight line $v_i, (v_{i,x} + 1, v_{i,y}), (v_{i,x} + 2, v_{i,y}), \dots, v_l$.

Observes that the vertical diameter will be lower by exactly 1 than the vertical diameter of P , and the horizontal distance between any point on L to other point on P will be lower than the horizontal distance of between the end of L (namely $\{v_i, v_l\}$), to the same point. So if $P^\uparrow \cup L, P^\downarrow \cup L$ are simple cycles then we get the desire. Suppose that $P^\uparrow \cup L$ is not a simple cycle. This means that the path $v_i, v_{i+1} \dots v_{l-1}, v_l \dots (v_{i,x} + 2, v_{i,y}), (v_{i,x} + 1, v_{i,y}) v_i$ must contain an inner loop. This loop cannot be supported on P^\uparrow alone, because $P^\uparrow \subset P$ so the inner loop would have to be contained also in P , which is a contradiction. Therefore, the inner loop is also supported on L , so it follows that there is a vertex in L with degree > 2 in $P^\uparrow \cup L$, namely a vertex $u \neq v_i, v_l$ such that $u_y = v_{i,y}$, and also $u \in P$. But by the construction of L , $|u_x - v_{i,x}| < |v_{l,x} - v_{i,x}|$, which is a contradiction to the fact that v_l was chosen to minimize the horizontal distance. Thus, we have our claim. \square

We have now established that $c \in C_X$ with $|c| < l$ can be decomposed into simple cycles, each of which has a weight less than l . Applying ?? recursively, we can further decompose each of these cycles into a sum of unit squares. However, ?? states that unit squares are in C_Z^\perp , so $c \in C^\perp$. \square

Note that the subtraction of any two l -length vertical cycles is a code word of C'_Z (one can be obtained by adding unit rectangles to the other). Therefore, we have only two non-trivial cosets in C_X/C'_Z . Hence, the dimension of the code is two.

9 Quantum Expander Codes.

As similar to the classical case, the next natural question to ask is whether there are codes with positive rates. The quantum expanders were the first quantum LDPC codes to achieve a square-root distance and positive rate [Tillich'2014; Leverrier'2015]. The leading insight was the idea that the Toric code could be represented as a variant product of the repetition code. For example, consider the cross restriction in Figure ??; that restriction can be obtained by gluing two vertices of two different cycle graphs.

Definition 9.1. For any two matrices A, B , with the same number of rows, denote by $[A, B]$ the matrix obtained by attach B next to A from right. Let $H_1, H_2 \in \mathbb{F}_2^{n \times r}$ be the parity check matrices. Define the bit and the phase parity checks matrices to be:

$$\begin{aligned} H_X &= [H_1 \otimes I_r \quad I_n \otimes H_2^\top] \\ H_Z &= [I_r \otimes H_2 \quad H_1^\top \otimes I_n] \end{aligned}$$

The matrices are orthogonal to each other as $H_X H_Z^\top = H_1 \otimes H_2^\top + H_1 \otimes H_2^\top = 0$ and therefore the pair define a valid CSS code. We will call to that code the Hyperproduct and denote it by $Q(H_1 \times H_2)$.

Obliviously, if $H_1, H_2, H_1^\top, H_2^\top$ are parity checks matrices of an LDPC codes, so are H_X, H_Z as their maximal row weight is at most two times larger.

```

latex.matrix_delimiters('['']')M = matrix(GF(2), [for_inrange(3)])repetition =
codes.LinearCode(M)H = repetition.parity_check_matrix().stack(matrix(GF(2), [1, 1, 0]))H1 =
identity_matrix(3).tensor_product(H).augment(H.transpose().tensor_product(identity_matrix(3)))Hstr =
latex(H)H1str = latex(H1)

```

Example 9.1. The Toric code could be thought as the Hyperproduct of the repetition code with himself. The parity check matrices of the codes are given follow. The left 3×3 matrix corresponds to the repetition code while the right 18×9 corresponds to the vertices check of the Toric code.

$$HstrH1str$$

Claim 9.1. Let $A, B \in \mathbb{M}_{n \times r_1}, \mathbb{M}_{n \times r_2}$ then $\dim \ker[A, B]$ is $\dim \ker A + n$.

Proof. The proof is omitted. □

Claim 9.2. Let k_1, k_2 be the demission of the codes with the full rank parity check matrices H_1, H_2 . Then the dimension of the Hyperproduct code is $\geq k_1 k_2$.

Proof. We will find the dimensions of each of the classical codes defined by $\ker H_X$ and $\ker H_Z$. Notice that the length of the $H \otimes I_n$ equals $n \times n = n^2$, And assuming the fullness of the ranks, the length of $I_{r_1} \otimes H_2^\top$ is $r_1 \cdot r_2$. Thus the length of $\ker H_X$ is $n^2 + r_1 r_2$. Now, recall that for any matrix A it holds that $\dim \ker(A \otimes I_l) = l \cdot \dim \ker A$. Therefore using ?? we obtain that the dimension of $\ker H_X$ is $k_1 n + r_1 r_2$.

By the same arguments we have that $\dim C_Z = k_2 n + r_1 r_2$. Thus the dimension of the quantum code is:

$$\begin{aligned} \dim Q(C_X, C_Z) &= \dim C_X + \dim C_Z - (n^2 + (n - k_1)(n - k_2)) \\ &= (k_1 + k_2)n + 2(n - k_1)(n - k_2) - (n^2 + (n - k_1) \cdot (n - k_2)) \\ &= k_1 k_2 \end{aligned}$$

□

Remark 9.1. Let H'_1 be a parity check matrix obtained by puncturing columns from H_1 , denote by k'_1 the dimension of that code. Then the Hyperproduct $Q(H'_1 \times H_2)$ is a CSS code with dimension $k'_1 k_2$. Moreover if the number of columns left after the puncturing is less the distance of $\ker H_1$ then it must hold that $k'_1 = 0 \Rightarrow \dim Q(H'_1 \times H_2) = 0$. Otherwise, one can take a non trivial codeword of $\ker H'_1$ and extending it to a valid codeword of $\ker H_1$ by set any punctured coordinate of it to zero. The yielded codeword has weight less than d which is contradiction.

Claim 9.3. Denote by d the minimal distance of $\ker H_1$. Any codeword x of $C_X = \ker H_X$ with edge at most d belongs to C_Z^\perp .

Proof. Define by H'_1 the matrix which obtained by puncturing from H_1 the columns associated with the coordinates e_i such that subspace corresponding to $e_i \otimes I_n$ doesn't support x . Denote the by S the set of the reaming coordinates. For example if $H_1 \otimes I_n = I_2 \otimes I_2$ and $x = [1, 0, 0, 0]$ then H'_1 is the unit matrix $[1, 0]^\top$ obtained by puncturing the second column of $H_1 = I_2$, and $S = \{e_1\}$.

As $|x| < d$ we have that $|S| < d$, Namely H'_1 supported on less than d coordinates and therefore $\dim Q(H'_1 \times H_2) = k'_1 k_2 = 0$. Thus, by the fact that for any CSS code $\dim C = \dim C_X - \dim C_Z^\perp$ it follows that $\dim \ker H'_X = \dim \ker H_Z^\perp \Rightarrow \ker H'_X = \ker H_Z^\perp$. Denote by x' the restriction of x to the columns of H'_X and clearly, by the definition of the construction, x' is a codeword of $\ker H'_X$. Thus x' is also codeword of $\ker H_Z^\perp$ and by the same argument, x is also a codeword of $\ker H_Z^\perp$. \square

Immediately from ?? we obtain that existences of quantum LDPC codes with positive rate and $\Theta(\sqrt{n})$ distance by taking the Hyperproduct of two classical expander codes.

Theorem 9.1. There exists an infinity family of QLDPC codes with positive rate and $\Theta(\sqrt{n})$ distance.

Part III

Appendix.

10 Expander Codes

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

Definition 10.1. 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 Mixing Lemma. We refer the reader to [hoory2006expander] for more detailed survey.

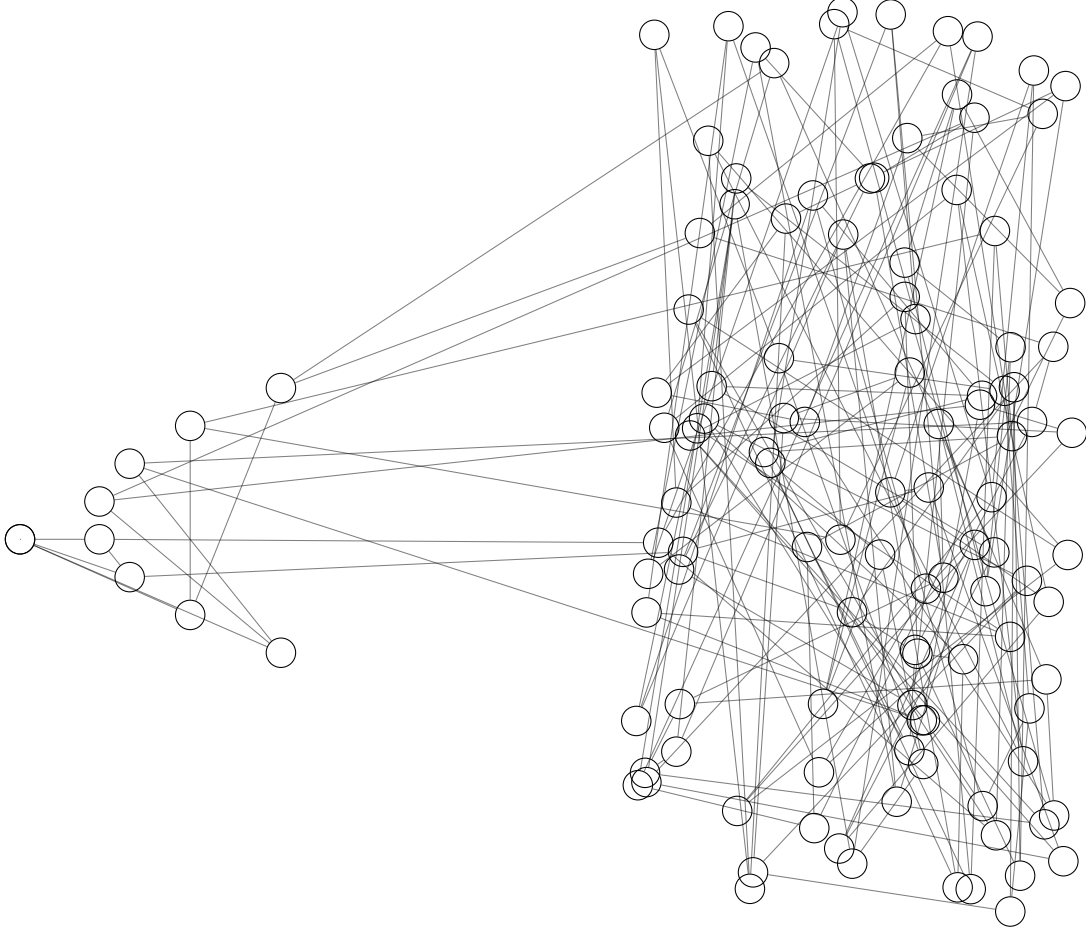


Figure 5: Expander Graph. Any small set of vertices, like the isolated set on the left, has more edges leaving it than intermediate edges. The exact amount is controlled by the expansion factor.

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 Tanner 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 Mixing Lemma we have that:

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

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 \square

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.

11 Randomized Constructions.

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

Claim 11.1. *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 11.2. *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 $j \neq k$.

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

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

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

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

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:

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

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

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

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 . \square