Quantum Coding Theory **(UC Berkeley CS294, Spring 2024)**

Lecture 6: Linear Codes and Stabilizer Codes

February 9, 2024

Lecturer: John Wright *Scribe: Pranav Trivedi*

1 Linear Codes

Last time we saw the general theory of error correction, but we did not discuss how to design error correcting codes. In particular we did not provide a systematic way of constructing quantum error correcting codes. To that end, we will finish discussing classical linear codes and see how they will give rise to two families of quantum error correcting codes: stabilizer codes and CSS codes.

In the previous lecture, we saw a classical linear code $C \subseteq \{0,1\}^n$ that is a linear subspace. One way to characterize C is as the span of linearly independent basis of vectors $g_1, \dots, g_k$. Equivalently, if we let

$$G = \begin{bmatrix} — & g_1 & — \\ — & g_2 & — \\ & \vdots & \\ — & g_k & — \end{bmatrix}$$

then $C = \{x \cdot G \mid x \in \{0,1\}^k\}$ and G is called the *generator matrix* for G.

Definition 1.1. A *parity check* is a $h \in \{0,1\}^n$ such that $h \cdot c = 0 \pmod 2$ for all $c \in C$.

The set of parity checks is denoted

$$C^\perp := \{h \mid h \cdot c = 0 \forall c \in C\}.$$

Note that if $h_1, h_2 \in C^\perp$ then $h_1 + h_2 \in C^\perp$ so $C^\perp$ is a linear subspace. Using the parity checks, we can construct another matrix known as the parity check matrix:

$$H = \begin{bmatrix} — & h_1 & — \\ — & h_2 & — \\ & \vdots & \\ — & h_{n-k} & — \end{bmatrix}$$

where $h_1, h_2, \dots, h_{n-k}$ is a linearly independent basis of parity checks.

Claim 1.2. *A linear code C is generated by a basis of k vectors if and only if $C^\perp$ is generated by a basis of $n-k$ vectors.*

Proof. Fix any nonzero parity check $h \in C^{\perp}$. Let $X_i = \{u \in \{0,1\}^n : h \cdot u = i \pmod 2\}$ for $i = 0, 1$. For any nonzero $x \in \{0,1\}^n$ such that $h \cdot x = 1 \pmod 2$, if $v \in X_0$ then $x + v \in X_1$. and if $v \in X_1$ then $x + v \in X_0$. Hence, we can partition $\{0,1\}^n$ into pairs $(v, x+v)$ where the first member is in X_0 and the second is in X_1. Therefore, h divides in $\{0,1\}^n$ in half. If we have $n-k$ independent parity checks then the resulting space will have dimension k since we half the number of vectors in the space with every additional independent parity check. □

For any $c \in C$ we know that $h \cdot c = 0 \pmod 2$ by definition of h. This implies that $C \subseteq \{c \mid h \cdot c = 0, \forall h \in C^{\perp}\}$, but even more is true.

Fact 1.3. $C = \{c \mid h \cdot c = 0, \forall h \in C^{\perp}\}$

Proof. The RHS can be rewritten as the set

$$C' := \{c \mid H \cdot c = 0\} = \ker H$$

where H is the parity check matrix. This is because any parity check can be written as a linear combination of the $n-k$ linearly independent basis of parity checks. Since the rank of H is $n-k$, the dimension of $\ker H = k$. So $C \subseteq C'$, but they have the same finite dimension so $C = C'$. □

This means we can define the code as the kernel of the parity check matrix or as the row space of the generator matrix. Additionally,

$$C = \{c \mid h \cdot c = 0, \forall h \in C^{\perp}\} = (C^{\perp})^{\perp}$$

so we have that $(C^{\perp})^{\perp} = C$.

Since we defined a linear code to be any linear subspace of $\{0,1\}^n$ then by above, $C^{\perp}$ is also a linear code and is known as the *dual code* of C. The generator matrix for $C^{\perp}$ is H and its parity check matrix is G.

Now for all $c = (c_1, \ldots, c_n) \in C$ we have that

$$H \cdot c^T = \begin{bmatrix} — & h_1 & — \\ — & h_2 & — \\ & \vdots & \\ — & h_{n-k} & — \end{bmatrix} \cdot \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} h_1 \cdot c \\ h_2 \cdot c \\ \vdots \\ h_{n-k} \cdot c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

and more generally

$$H \cdot G^T = \begin{bmatrix} — & h_1 & — \\ — & h_2 & — \\ & \vdots & \\ — & h_{n-k} & — \end{bmatrix} \cdot \begin{bmatrix} | & | & \cdots & | \\ g_1 & g_2 & \cdots & g_k \\ | & | & \cdots & | \end{bmatrix} = 0.$$

This leads to the discussion of syndromes in the classical case. Notice that if we choose any $c \in C$ and add an error $e \in E$, our set of errors. Then for any parity check h, we have

$$h \cdot (c+e) = h \cdot c + h \cdot e = 0 + h \cdot e = h \cdot e.$$

More generally,

$$H(c+e)^T = Hc^T + He^T = He^T$$

so the syndrome depends only on the error and not the codeword. This is exactly what we saw in the the quantum setting.

In the classical setting, for E to be correctable, we need $He_1^T \neq He_2^T$ for all $e_1, e_2 \in E$. In words, we see that distinct errors lead to distinct corruption. This is in contrast with the quantum setting where we saw distinct errors lead to the same corruption. The weight of an error that can be corrected depends on the (Hamming) distance between codewords.

Definition 1.4. The *distance* of a code C is defined as $d(C) = \min_{x \neq y \in C} \Delta(x, y)$ where Δ is the Hamming distance. If C is linear, we have that

$$\min_{x \neq y \in C} \Delta(x,y) = \min_{x \neq y \in C} \Delta(x-y, 0) = \min_{x \neq 0 \in C} \Delta(x, 0) = \min_{x \neq 0 \in C} \mathrm{wt}(x).$$

So the distance of a linear code is the minimum weight of a nonzero codeword.

Now we characterize the errors that can be corrected given the distance of a code.

Fact 1.5. *Let C be a code with distance $2t+1$. Then we can correct errors of weight $\leq t$.*

Proof. Suppose the transmitted codeword is c_1 and y is the received codeword. Then $\Delta(c_1, y) \leq t$. Assume for the sake of contradiction, we incorrectly decode y as a codeword $c_2 \neq c_1 \in C$. Then $\Delta(c_2, y) \leq \Delta(c_1, y)$. However, this leads to

$$\Delta(c_1, c_2) \leq \Delta(c_1, y) + \Delta(c_2, y) \leq 2\Delta(c_1, y) \leq 2t < 2t+1.$$

This contradicts is minimum distance of C. □

Finally, note that if $e_1 \neq e_2$ are errors of weight $\leq t$ then $0 < \mathrm{wt}(e_1 + e_2) \leq 2t$ so $H(e_1+e_2)^T \neq 0$ since nonzero codewords have weight at least $2t+1$. Therefore, $He_1^T \neq He_2^T$.

2 Back to the Quantum Setting

We will now use this background on classical linear codes to develop a construction for quantum error correcting codes.

Definition 2.1. Let $\Pi = \{\Pi_0, \Pi_1\}$ be a two outcome projective measurement. The corresponding *binary observable* is $O = \Pi_0 - \Pi_1$.

For example $Z = |0\rangle\langle 0| - |1\rangle\langle 1|$ and $X = |+\rangle\langle +| - |-\rangle\langle -|$ so the "Z basis" is $\{|0\rangle, |1\rangle\}$ and the "X basis" is $\{|+\rangle, |-\rangle\}$.

Fact 2.2. *If O is Hermitian matrix then O is a binary observable if and only if $O^2 = I$.*

A Hermitian matrix has real eigenvalues so if $O^2 = I$ then the eigenvalues of O are ± 1. We can write a projector Π_0 for the $+1$ eigenvalue and Π_1 for the -1 eigenvalue and obtain our observable. For example $Y^2 = I$ so Y is an observable.

Fact 2.3. *Measure $\{\Pi_0, \Pi_1\}$ on $|\psi\rangle$ then we always observe 0 if and only if $O|\psi\rangle = |\psi\rangle$. Notice that*

$$O|\psi\rangle = \Pi_0|\psi\rangle - \Pi_1|\psi\rangle = \Pi_0|\psi\rangle$$

since we always observe 0. This implies $\Pi_0|\psi\rangle = |\psi\rangle$. In other words, $|\psi\rangle$ is a $+1$ eigenvector. Similarly, we always observe 1 if and only if $\Pi_1|\psi\rangle = -|\psi\rangle$ so $|\psi\rangle$ is a -1 eigenvector.

It is easy to see that $|0\rangle$ is a $+1$ eigenvector of Z and $|1\rangle$ is a -1 eigenvector of Z. And similarly $|+\rangle$ is a $+1$ eigenvector of X and $|-\rangle$ is a -1 eigenvector of X.

Additionally, suppose we want to measure the pairs of parities of qubits but we do not care to measure each of the measurements individually. Notice that if $\Pi_A = \{A_0, A_1\}$ and $\Pi_B = \{B_0, B_1\}$ and they correspond to observables O_A and O_B, respectively. Let $\Pi_0 = A_0 \otimes B_0 + A_1 \otimes B_1$ and $\Pi_1 = A_0 \otimes B_1 + A_1 \otimes B_0$. Then

$$O = A_0 \otimes B_0 + A_1 \otimes B_1 - A_0 \otimes B_1 + A_1 \otimes B_0 = (A_0 - A_1) \otimes (B_0 - B_1) = O_A \otimes O_B.$$

Consider the 3 qubit bit flip code. If we apply ZZI to the state $|\psi\rangle = a|000\rangle + b|111\rangle$ then we obtain $|\psi\rangle$ so it is a $+1$ eigenvector. On the other hand, it is easy to check that $|\phi\rangle = a|010\rangle + b|101\rangle$ is a -1 eigenvector. So the parity checks for this code are ZZI and IZZ.

Extending this to the 9 qubit Shor code, we have the parity checks $Z_1Z_2, Z_2Z_3, Z_4Z_5, Z_5Z_6, Z_7Z_8, Z_8Z_9$ that will check for any bit flips. To check for phase flips we need to check that any two blocks of 3 qubits have the same phase so we need the checks $X_1X_2X_3X_4X_5X_6$ and $X_3X_4X_5X_6X_7X_8X_9$.

For example,

$$\begin{aligned}
&X_1X_2X_3X_4X_5X_6(|000\rangle + (-1)^a|111\rangle)(|000\rangle + (-1)^b|111\rangle) \\
&= (|111\rangle + (-1)^a|000\rangle)(|111\rangle + (-1)^b|000\rangle) \\
&= (-1)^{(a+b)}(|000\rangle + (-1)^a|111\rangle)(|000\rangle + (-1)^b|111\rangle) \\
&= \begin{cases} +1 \text{ eigenvalue if} & a = b \\ -1 \text{ eigenvalues if} & a \neq b. \end{cases}
\end{aligned}$$

Now we will generalize this idea to construct quantum codes. Given a set of parity checks $\{P_i\}$ we let $C = \{|\psi\rangle : P_i|\psi\rangle = |\psi\rangle \, \forall i\}$.

Fact 2.4. *If P_i and P_j are parity checks for C then so is P_iP_j.*

Proof. For all $|\psi\rangle \in C$ we have

$$P_iP_j|\psi\rangle = P_i|\psi\rangle = |\psi\rangle.$$

□

This fact implies that the set of parity checks form a subgroup of the Paulis.

We will provide a brief recap on some properties of Pauli matrices.The Pauli matrices are observables because $X^2 = Y^2 = Z^2 = I$. We also have the relations $XY = iZ, YZ = iX$, and $ZX = iY$. Lastly all the Pauli matrices anti-commute because $XY = -YX, XZ = -ZX, YZ = -ZY$.

However, products of Paulis will commute if they have an even number of locations with differences. For example,

$$(X \otimes X) \cdot (Z \otimes Z) = (XZ) \otimes (XZ) = (-ZX) \otimes (-ZX) = (ZX) \otimes (ZX) = (Z \otimes Z) \cdot (X \otimes X).$$

In general, the product of n Paulis $P_1 \otimes \cdots \otimes P_n$ and $Q_1 \otimes \cdots \otimes Q_n$ will always commute if the number of locations with $P_i \neq Q_i$ is even and will anticommute otherwise.

Definition 2.5. The n-qubit Pauli group Pauli$_n$ is the set of matrices $\pm P_1 \otimes \cdots \otimes P_n$ which are the observables and the matrices $\pm i P_1 \otimes \cdots \otimes P_n$ are the non observables. Here $P_i \in \{I, X, Y, Z\}$.

This leads us to the following characterization of a stabilizer quantum code.

Definition 2.6. Let S be a subgroup of Pauli$_n$. The *stabilizer code* is $C(S) = \{|\psi\rangle : P|\psi\rangle = |\psi\rangle, \forall P \in S\}$. We call S the stabilizer group and each P a stabilizer.

We will end with a couple facts about stabilizer codes that complete their characterization.

Fact 2.7. *For $C(S)$ to be nonempty we need*

- *For all $P, Q \in S$, we need $PQ = QP$. Notice*

  $$PQ|\psi\rangle = P|\psi\rangle = |\psi\rangle = Q|\psi\rangle = QP|\psi\rangle,$$

  but P and Q either commute or anticommute so by the above, it must be the case that $PQ = QP$.

- *S contains only $\pm P_1 \otimes \cdots \otimes P_n$ and not $\pm i P_1 \otimes \cdots \otimes P_n$. The former has ± 1 eigenvalues and the latter is not even Hermitian.*

- *S does not contain $-I$, as no state is stabilized by $-I$. For a further discussion of this point, see Lecture 8.*