Introduction to Quantum Error Correction

Jonathan Webb

August 2022

The classical repetition code

A bit being transmitted has some probability p of flipping. To increase the reliability of this channel, we perform the following encoding: $0 \to 000$, $1 \to 111$, and we instruct the receiver to interpret the message to be the state they see most frequently.

The classical repetition code

A bit being transmitted has some probability p of flipping. To increase the reliability of this channel, we perform the following encoding: $0 \to 000$, $1 \to 111$, and we instruct the receiver to interpret the message to be the state they see most frequently.

For an encoded state to be interpreted incorrectly, it must have suffered at least two bit flips. It turns out that if $p < 1/2$, the channel is more reliable if this encoding is used as opposed to sending the bit with no encoding.

By introducing redundancy, we have made transmission more reliable - this will be a common theme.

Hamming space

A Hamming space is the set of all words of fixed length made from letters of some alphabet. A code is a subset of this space. For instance, the repetition code is the set $\{000, 111\}$, which is a subset - in fact, a subspace - of $(\mathbb{Z}/2)^3$.

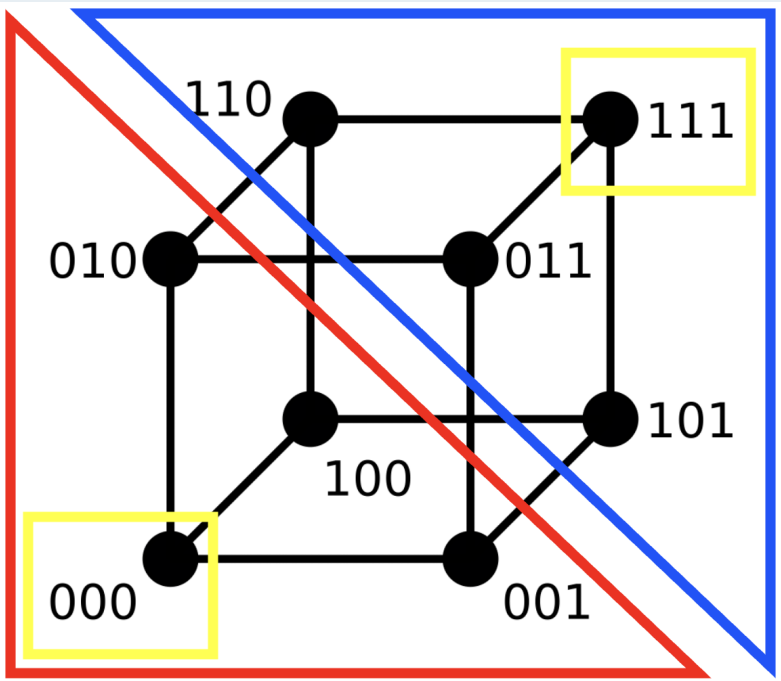
Hamming space

A Hamming space is the set of all words of fixed length made from letters of some alphabet. A code is a subset of this space. For instance, the repetition code is the set $\{000, 111\}$, which is a subset - in fact, a subspace - of $(\mathbb{Z}/2)^3$.

The Hamming distance is defined to be the number of places at which two words differ. The distance of a code is defined as the smallest distance between any two distinct codewords,

$$d = \inf_{x,y \in C} d(x, y) \text{ with } x \neq y.$$

A code with distance d can correct at most $\left\lfloor \frac{d-1}{2} \right\rfloor$ errors.



This is the Hamming space $(\mathbb{Z}/2)^3$, with the code $\{000, 111\}$ outlined in yellow. The red and blue triangles are the balls of radius 1, with respect to the Hamming metric, with these codewords at their centers. Note two things: we must move three times to get from one codeword to the other (the distance of the code is 3), and we cannot correct errors which take us outside of the error ball (the code can correct up to 1 error). This image suggests an analogy between error correction and sphere packing.

Qubit states

Regarding Dirac notation, $|v\rangle$ denotes a vector, just like v or $\mathbf{v}$; $\langle v| := (|v\rangle)^\dagger$.

A qubit is the fundamental unit of quantum information; it is the quantum analogue of a bit. We represent the state of a qubit as a vector $|\psi\rangle = a|0\rangle + b|1\rangle$ in a two-dimensional complex Hilbert space, and require that $|a|^2 + |b|^2 = 1$.

We can identify the basis elements $|0\rangle$ and $|1\rangle$ with $\begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}0\\1\end{bmatrix}$, respectively; we can then think of $|\psi\rangle$ as $\begin{bmatrix}a\\b\end{bmatrix}$.

Errors as operators

If qubit states are two-dimensional vectors with complex entries, then qubit states are changed, or acted on, by operators in the space of 2×2 complex matrices. The Pauli matrices are a set of Hermitian, unitary, and involutory 2×2 complex matrices which form a basis for $M_{2,2}(\mathbb{C})$.

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Errors as operators

If qubit states are two-dimensional vectors with complex entries, then qubit states are changed, or acted on, by operators in the space of 2×2 complex matrices. The Pauli matrices are a set of Hermitian, unitary, and involutory 2×2 complex matrices which form a basis for $M_{2,2}(\mathbb{C})$.

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

The action of each Pauli corresponds to an error on one qubit. For some $|\psi\rangle = a|0\rangle + b|1\rangle$,

$$\begin{aligned} I|\psi\rangle &= |\psi\rangle \\ X|\psi\rangle &= a|1\rangle + b|0\rangle \\ Y|\psi\rangle &= i(a|1\rangle - b|0\rangle) \\ Z|\psi\rangle &= a|0\rangle - b|1\rangle. \end{aligned}$$

Measurement

There exists a measurement operation which collapses the qubit state $|\psi\rangle = a|0\rangle + b|1\rangle$ into the state $|0\rangle$ with probability $|a|^2$, or into the state $|1\rangle$ with probability $|b|^2$.

Measurement

There exists a measurement operation which collapses the qubit state $|\psi\rangle = a|0\rangle + b|1\rangle$ into the state $|0\rangle$ with probability $|a|^2$, or into the state $|1\rangle$ with probability $|b|^2$.

Let W be a subspace of a Hilbert space spanned by an orthonormal basis $|1\rangle, \ldots, |n\rangle$. The projector P into W is the operator given by

$$P := \sum_{i=1}^{n} |i\rangle\langle i|.$$

Measurement

There exists a measurement operation which collapses the qubit state $|\psi\rangle = a|0\rangle + b|1\rangle$ into the state $|0\rangle$ with probability $|a|^2$, or into the state $|1\rangle$ with probability $|b|^2$.

Let W be a subspace of a Hilbert space spanned by an orthonormal basis $|1\rangle, \ldots, |n\rangle$. The projector P into W is the operator given by

$$P := \sum_{i=1}^{n} |i\rangle\langle i|.$$

Observables are Hermitian operators whose eigenvalues correspond to measurement outcomes. An observable M has a spectral decomposition,

$$M = \sum_{i} \lambda_i P_{\lambda_i}.$$

Measurements defined by observables are called projective measurements.

Multi-qubit systems

Given n qubits with individual states $|\psi_1\rangle, \ldots, |\psi_n\rangle$, the state of the total system of these n qubits is given by $|\psi_1\rangle \otimes \cdots \otimes |\psi_n\rangle$. However, not all quantum systems can be written in this way (as simple tensors). Entangled states, such as one of the Bell states

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}},$$

are those which cannot be written as a tensor product of subsystems. Entanglement plays a key role in quantum error correction. (Notation: $|00\rangle = |0\rangle \otimes |0\rangle$.)

Multi-qubit systems

Given n qubits with individual states $|\psi_1\rangle, \ldots, |\psi_n\rangle$, the state of the total system of these n qubits is given by $|\psi_1\rangle \otimes \cdots \otimes |\psi_n\rangle$. However, not all quantum systems can be written in this way (as simple tensors). Entangled states, such as one of the Bell states

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}},$$

are those which cannot be written as a tensor product of subsystems. Entanglement plays a key role in quantum error correction. (Notation: $|00\rangle = |0\rangle \otimes |0\rangle$.)

In a similar way, operators acting on multiple qubits are equal to the tensor product of the operators acting on individual qubits. For example, the bit flip on the second of three qubits is equal to $I \otimes X \otimes I$, which may be abbreviated as X_2.

Barriers to QEC

1. Measuring a quantum state changes it.

2. We cannot replicate quantum information. The no-cloning theorem states that there is no unitary transformation which will take us from $|\psi\rangle \otimes |e\rangle$ to $|\psi\rangle \otimes |\psi\rangle$.

3. Besides an error analogous to a classical bit flip, qubits can also suffer phase flips or continuous errors such as rotation by an arbitrary degree.

The $\{|0\rangle, |1\rangle\}$ repetition code

Suppose a qubit $|\psi\rangle = a|0\rangle + b|1\rangle$ is sent across a channel whose effect is to apply the Pauli X, or bit flip, matrix with some probability. To mitigate the effects of this noise, encode $|\psi\rangle$ as $a|000\rangle + b|111\rangle$. Then the code is the subspace spanned by $\{|000\rangle, |111\rangle\}$, subject to the constraint on the norms of a and b.

The error detection strategy is to compare the first and second qubits, then compare the second and third qubits, and use the four distinct pairs of outcomes to determine which error occurred.

The $\{|0\rangle, |1\rangle\}$ repetition code

Suppose a qubit $|\psi\rangle = a|0\rangle + b|1\rangle$ is sent across a channel whose effect is to apply the Pauli X, or bit flip, matrix with some probability. To mitigate the effects of this noise, encode $|\psi\rangle$ as $a|000\rangle + b|111\rangle$. Then the code is the subspace spanned by $\{|000\rangle, |111\rangle\}$, subject to the constraint on the norms of a and b.

The error detection strategy is to compare the first and second qubits, then compare the second and third qubits, and use the four distinct pairs of outcomes to determine which error occurred.

We do this by measuring the observables Z_1Z_2 and Z_2Z_3, for a combination of two reasons:

1. Z has eigenvalue 1 for $|0\rangle$ and eigenvalue -1 for $|1\rangle$.

2. If $Ax = \lambda x$ and $By = \mu y$, then $(A \otimes B)(x \otimes y) = \lambda\mu(x \otimes y)$.

The $\{|0\rangle, |1\rangle\}$ repetition code

Therefore $|00\rangle$ and $|11\rangle$ have eigenvalue 1 for ZZ, while $|01\rangle$ and $|10\rangle$ have eigenvalue -1 for these operators. To see an example of how these measurements help detect error, suppose the encoded state suffers a bit flip on the second qubit. Let $|\psi'\rangle = a|010\rangle + b|101\rangle$. Then,

$$\begin{aligned} Z_1 Z_2 |\psi'\rangle &= a Z_1 Z_2 |010\rangle + b Z_1 Z_2 |101\rangle \\ &= -a|010\rangle - b|101\rangle \\ &= -|\psi'\rangle. \end{aligned}$$

The $\{|0\rangle, |1\rangle\}$ repetition code

Therefore $|00\rangle$ and $|11\rangle$ have eigenvalue 1 for ZZ, while $|01\rangle$ and $|10\rangle$ have eigenvalue -1 for these operators. To see an example of how these measurements help detect error, suppose the encoded state suffers a bit flip on the second qubit. Let $|\psi'\rangle = a|010\rangle + b|101\rangle$. Then,

$$\begin{aligned} Z_1 Z_2 |\psi'\rangle &= a Z_1 Z_2 |010\rangle + b Z_1 Z_2 |101\rangle \\ &= -a|010\rangle - b|101\rangle \\ &= -|\psi'\rangle. \end{aligned}$$

Therefore $|\psi'\rangle$ is a -1 eigenvector of $Z_1 Z_2$. It turns out that this is the case for $Z_2 Z_3$ as well. Then we measure the pair $(-1, -1)$, indicating that the first and second qubits are different, and the second and third qubits are different. Assuming at most one error, the error in $|\psi'\rangle$ must have been to the second qubit.

The $\{|+\rangle, |-\rangle\}$ repetition code

Suppose a qubit $|\psi\rangle = a|0\rangle + b|1\rangle$ is sent across a channel whose effect is to apply the Pauli Z, or phase flip, matrix with some probability. To mitigate the effects of this noise, encode $|\psi\rangle$ as $a|+\,+\,+\rangle + b|-\,-\,-\rangle$, where the $\{|+\rangle, |-\rangle\}$ basis vectors are defined as

$$|\pm\rangle = \frac{|0\rangle \pm |1\rangle}{\sqrt{2}}.$$

The $\{|+\rangle, |-\rangle\}$ repetition code

Suppose a qubit $|\psi\rangle = a|0\rangle + b|1\rangle$ is sent across a channel whose effect is to apply the Pauli Z, or phase flip, matrix with some probability. To mitigate the effects of this noise, encode $|\psi\rangle$ as $a|+\,+\,+\rangle + b|-\,-\,-\rangle$, where the $\{|+\rangle, |-\rangle\}$ basis vectors are defined as

$$|\pm\rangle = \frac{|0\rangle \pm |1\rangle}{\sqrt{2}}.$$

We can move between the $\{|0\rangle, |1\rangle\}$ and $\{|+\rangle, |-\rangle\}$ bases easily via the Hadamard gate H, which has the following effects:

$$H(a|0\rangle + b|1\rangle) = a|+\rangle + b|-\rangle$$
$$X(a|+\rangle + b|-\rangle) = a|+\rangle - b|-\rangle$$
$$Z(a|+\rangle + b|-\rangle) = a|-\rangle + b|+\rangle.$$

H exchanges the roles of X and Z.

The $\{|+\rangle, |-\rangle\}$ repetition code

Error detection for this code is done nearly exactly as it was done for the $\{|0\rangle, |1\rangle\}$ code, except we instead measure the observables X_1X_2 and X_2X_3. We have chosen our code to lie in the $+1$ eigenspace of two observables whose measurement does not affect the information we are transmitting, with the rationale that measuring a state to be outside of this eigenspace corresponds to an error.

The $\{|+\rangle, |-\rangle\}$ repetition code

Error detection for this code is done nearly exactly as it was done for the $\{|0\rangle, |1\rangle\}$ code, except we instead measure the observables X_1X_2 and X_2X_3. We have chosen our code to lie in the $+1$ eigenspace of two observables whose measurement does not affect the information we are transmitting, with the rationale that measuring a state to be outside of this eigenspace corresponds to an error.

As with the prior code, each pair of measurements we observe corresponds to a distinct error. For instance, a phase flip on the first qubit results in the state $a|- + +\rangle + b|+ - -\rangle$, which is a -1 eigenvector of X_1X_2 and a $+1$ eigenvector of X_2X_3.

The Shor code

The Shor code encodes a qubit state $|\psi\rangle = a|0\rangle + b|1\rangle$ into the nine-qubit state given by the codewords below:

$$|0\rangle \to \frac{(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)}{2\sqrt{2}}$$
$$|1\rangle \to \frac{(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)}{2\sqrt{2}}.$$

This is done by combining the above repetition codes. We perform two encodings:

1. $|0\rangle \to |+++\rangle, |1\rangle \to |---\rangle$.

2. $|+\rangle \to (|000\rangle + |111\rangle)\sqrt{2}, |-\rangle \to (|000\rangle - |111\rangle)\sqrt{2}$.

The Shor code

	$\|q_1\rangle$	$\|q_2\rangle$	$\|q_3\rangle$	$\|q_4\rangle$	$\|q_5\rangle$	$\|q_6\rangle$	$\|q_7\rangle$	$\|q_8\rangle$	$\|q_9\rangle$
M_1	Z	Z							
M_2		Z	Z						
M_3				Z	Z				
M_4					Z	Z			
M_5							Z	Z	
M_6								Z	Z
M_7	X	X	X	X	X	X			
M_8				X	X	X	X	X	X

The eight measurements above, with Is and tensor products omitted for readability, are performed to detect error and indicate which recovery operation must be performed.

Discretizing error

Theorem - If a code corrects errors E and F, then it corrects $aE + bF$.

Recall that the Pauli matrices span $M_{2,2}(\mathbb{C})$, and that an error on a qubit can be modeled as a 2×2 complex matrix. We have shown that X, Y, and Z errors on one qubit can each be corrected by the Shor code. By the above theorem, the Shor code can correct an arbitrary error on one qubit.

Discretizing error

Theorem - If a code corrects errors E and F, then it corrects $aE + bF$.

Recall that the Pauli matrices span $M_{2,2}(\mathbb{C})$, and that an error on a qubit can be modeled as a 2×2 complex matrix. We have shown that X, Y, and Z errors on one qubit can each be corrected by the Shor code. By the above theorem, the Shor code can correct an arbitrary error on one qubit.

To provide a bit more detail, suppose an error E can be written as the linear combination $aI + bX + cY + dZ$. Then, the (unnormalized) qubit state $E|\psi\rangle$ can be written as the superposition $a|\psi\rangle + bX|\psi\rangle + cY|\psi\rangle + dZ|\psi\rangle$. Measuring which error occurred collapses this superposition into one of the those four states, with the result being a Pauli error that we know how to correct.

Barriers to QEC (revisited)

1. Rather than measuring corrupted states directly, we instead measure error indirectly by comparing the corrupted state's qubits through projective measurements.

2. Since we cannot replicate quantum states, we entangle the qubit we are transmitting with other qubits to make a higher-dimensional state that is more resilient to noise.

3. We can correct phase flips very similarly to how we correct bit flips, and given that a code can correct both, it can correct arbitrary errors.

The Pauli group

Define the Pauli group on n qubits, P_n, to be the set of n-fold tensor products of Pauli matrices, along with the scalars ± 1, $\pm i$. For instance, $I \otimes X \otimes Z \in P_3$.

Any two operators in the Pauli group either commute or anticommute. In the one-qubit case, commutation is a bit trivial; for multi-qubit systems it can be more interesting.

The weight of some $M \in P_n$ is defined to be the number of qubits at which M acts as a non-identity operator.

The stabilizer group

The stabilizer of a code is the group consisting of all Pauli operators M with the property that $M|\psi\rangle = |\psi\rangle \ \forall |\psi\rangle \in C$. The stabilizer is abelian. This relates to the fact that commuting operators have simultaneous eigenvectors; we want operators which share the codewords as eigenvectors, therefore they should commute.

The stabilizer group

The stabilizer of a code is the group consisting of all Pauli operators M with the property that $M|\psi\rangle = |\psi\rangle \ \forall |\psi\rangle \in C$. The stabilizer is abelian. This relates to the fact that commuting operators have simultaneous eigenvectors; we want operators which share the codewords as eigenvectors, therefore they should commute.

Given any abelian group S of Pauli operators, we can define a code to be the intersection of the $+1$ eigenspaces of each $M \in S$:

$$C = \{|\psi\rangle \text{ such that } M|\psi\rangle = |\psi\rangle \ \forall M \in S\}.$$

For a stabilizer code with r generators and n-qubit codewords, the code must have dimension 2^{n-r}.

Detecting errors

Suppose E commutes with some $M \in S$, so that $ME|\psi\rangle = EM|\psi\rangle = E|\psi\rangle$. $E|\psi\rangle$ is a $+1$ eigenvector of M, so E is not detected as an error. Conversely, suppose E and M anticommute so that $E|\psi\rangle$ is a -1 eigenvector of M. Then E is detected as an error.

Let $N(S)$ be the set of operators N which commute with all $M \in S$. A stabilizer code detects errors outside of $N(S) \setminus S$ because errors outside of $N(S)$ anticommute with stabilizers and are detected by measuring eigenvalues, while errors in S act like the identity by definition - they are trivial errors.

Detecting errors

Suppose E commutes with some $M \in S$, so that $ME|\psi\rangle = EM|\psi\rangle = E|\psi\rangle$. $E|\psi\rangle$ is a $+1$ eigenvector of M, so E is not detected as an error. Conversely, suppose E and M anticommute so that $E|\psi\rangle$ is a -1 eigenvector of M. Then E is detected as an error.

Let $N(S)$ be the set of operators N which commute with all $M \in S$. A stabilizer code detects errors outside of $N(S) \setminus S$ because errors outside of $N(S)$ anticommute with stabilizers and are detected by measuring eigenvalues, while errors in S act like the identity by definition - they are trivial errors.

Errors can be classified as detectable (not in $N(S) \setminus S$), undetectable (in $N(S)$), or trivial (in S). Define the weight of a code to be the infimum of the set of weights of the elements of $N(S) \setminus S$. The weight of a code is the weight of the smallest error it cannot detect.

Correcting errors

Correcting error requires us to know exactly which operator had its eigenvalues change, so that we can apply the proper recovery procedure. We do this by measuring the eigenvalues of every operator in the stabilizer and noting that they can change differently due to different errors. For instance, an error that commutes with the first generator but not the second will give eigenvalues of $+1$ and -1; these will be flipped for the opposite commutation.

Correcting errors

Correcting error requires us to know exactly which operator had its eigenvalues change, so that we can apply the proper recovery procedure. We do this by measuring the eigenvalues of every operator in the stabilizer and noting that they can change differently due to different errors. For instance, an error that commutes with the first generator but not the second will give eigenvalues of $+1$ and -1; these will be flipped for the opposite commutation.

It can be shown that E and F have the same error syndrome (eigenvalue measurements) iff $E^\dagger F \in N(S)$. Then if $E^\dagger F \notin N(S)$, E and F may be distinguished by their different error syndromes. Meanwhile, if $E^\dagger F \in S$, then $E|\psi\rangle = F|\psi\rangle$, and we cannot distinguish between the two errors but don't need to because they do the same thing to the codewords. Then a stabilizer code corrects errors for which $E^\dagger F \notin N(S) \setminus S$ for all possible pairs of errors (E, S).