Introduction to Quantum Computation

These lecture notes are based on the *Introduction to Quantum Computation* course given at EPFL for the academic year 2024/2025 by Prof. Olivier Lévêque and Prof. Rüdiger Urbanke. The content of the course was originally created by Prof. Nicolas Macris and is based on his lecture notes.

Department of Computer Science - École Polytechnique Fédérale de Lausanne

Author: Arthur Aimone

2024/2025

Chapter 8

Quantum Error Correction

This chapter explains how to do quantum information processing reliably in the presence of noise. We begin by developing the basic theory of quantum error-correcting codes with the bit-flip and the phase-flip error models. We will then introduce the Shor code to deal with both bit-flips and phase-flips. Finally, we will briefly study the Steane code which is more efficient than Shor's code.

One could think that an equivalent argument of performing repetition on classical code is to consider the entangled system $|\phi\rangle \otimes |\phi\rangle \otimes |\phi\rangle$ where $|\phi\rangle = \alpha|0\rangle + \beta|1\rangle$. However this cannot happen by the No-cloning Theorem (cf. Remark **??**)! Instead, our repetition code will consist of taking $|0\rangle \mapsto |000\rangle$ and $|1\rangle \mapsto |111\rangle$. This isn't cloning since only the computational basis may be copied. Thus, for a general state $|\phi\rangle$, we will consider its repetition equivalent to be our "codewords" given by,

$$|\phi\rangle = \alpha|0\rangle + \beta|1\rangle \mapsto |\psi\rangle = \alpha|000\rangle + \beta|111\rangle. \tag{8.1}$$

In the first part of this chapter we consider two error models, one involving bit-flips (via an X gate) and the other involving phase flips (via a Z gate). Recall from Definition (**??**) that a Z gate is defined as,

$$Z = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{8.2}$$

8.1 Projective Measurements and Observables

In the previous chapters, we mostly discussed measurements in the computational basis. For one qubit, this means that the possible outcomes are 0 and 1, and a state

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \qquad |\alpha|^2 + |\beta|^2 = 1, \tag{8.3}$$

is measured as 0 with probability $|\alpha|^2$ and as 1 with probability $|\beta|^2$. In this chapter we will need a slightly more general language. For instance, in quantum error correction we will often measure whether two qubits have the same value or different values, without learning the individual values of the qubits. This is exactly what the formalism of **projective measurements** and **observables** allows us to describe.

Projective measurements

Let $\mathcal{H}$ be a finite-dimensional Hilbert space. Recall that a linear operator P on $\mathcal{H}$ is called a **projector** if

$$P^2 = P. \tag{8.4}$$

In quantum mechanics we also require projectors to be Hermitian, namely $P^\dagger = P$, where $P^\dagger$ denotes the conjugate transpose of P. Geometrically, such an operator projects a vector onto a subspace of $\mathcal{H}$.

Definition 8.1.1 (Projective measurement)**.** *A* ***projective measurement*** *is a finite family of projectors*

$$\{P_m\}_{m\in\mathcal{M}} \tag{8.5}$$

labelled by possible measurement outcomes $m \in \mathcal{M}$, satisfying

$$P_m P_{m'} = 0 \quad \text{for } m \neq m', \qquad \sum_{m\in\mathcal{M}} P_m = \mathbb{I}. \tag{8.6}$$

If the system is in a state $|\psi\rangle$, then the probability of outcome m is

$$\mathbb{P}(m) = \| P_m |\psi\rangle \|^2 = \langle\psi| P_m |\psi\rangle . \tag{8.7}$$

If outcome m occurs and $\mathbb{P}(m) > 0$, then the post-measurement state is

$$|\psi_m\rangle = \frac{P_m |\psi\rangle}{\sqrt{\mathbb{P}(m)}}. \tag{8.8}$$

The conditions in the definition have a simple interpretation. The equation $P_m P_{m'} = 0$ says that the different outcomes correspond to mutually orthogonal alternatives. The equation $\sum_m P_m = \mathbb{I}$ says that the list of possible outcomes is complete: some outcome must occur.

Example 8.1.2 (Computational basis measurement)**.** *For a single qubit, the usual computational basis measurement is the projective measurement*

$$P_0 = |0\rangle\langle 0|, \qquad P_1 = |1\rangle\langle 1|. \tag{8.9}$$

Indeed, $P_0 + P_1 = \mathbb{I}$ and $P_0 P_1 = 0$. If $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, then

$$\mathbb{P}(0) = \langle\psi| P_0 |\psi\rangle = |\alpha|^2, \qquad \mathbb{P}(1) = \langle\psi| P_1 |\psi\rangle = |\beta|^2. \tag{8.10}$$

After outcome 0, the state becomes $|0\rangle$; after outcome 1, the state becomes $|1\rangle$.

Projective measurements can also be less fine-grained than measuring every qubit. For example, they may reveal only one piece of information about a multi-qubit state.

Example 8.1.3 (Measuring only the first qubit)**.** *Consider a two-qubit system. Measuring only the first qubit in the computational basis is described by the projectors*

$$P_0 = |0\rangle\langle 0| \otimes \mathbb{I}, \qquad P_1 = |1\rangle\langle 1| \otimes \mathbb{I}. \tag{8.11}$$

The second qubit is not measured. For instance, take

$$|\psi\rangle = \frac{1}{2}\big(|00\rangle + |01\rangle + |10\rangle + |11\rangle\big). \tag{8.12}$$

Then

$$\mathbb{P}(0) = \|P_0|\psi\rangle\|^2 = \frac{1}{2}, \qquad \mathbb{P}(1) = \|P_1|\psi\rangle\|^2 = \frac{1}{2}. \tag{8.13}$$

If the outcome is 0*, then the post-measurement state is*

$$\frac{P_0|\psi\rangle}{\sqrt{1/2}} = \frac{|00\rangle + |01\rangle}{\sqrt{2}}. \tag{8.14}$$

Notice that the measurement did not decide whether the second qubit was 0 *or* 1*.*

A useful special case is when the state is already inside one of the subspaces of the measurement. Suppose that $P_m|\psi\rangle = |\psi\rangle$ for some outcome m. Then $\mathbb{P}(m) = 1$ and the state is not changed by the measurement. This observation will be important for error syndrome measurements in error-correcting codes.

Observables

Instead of listing projectors directly, it is often convenient to package a projective measurement into a single matrix.

Definition 8.1.4 (Observable)**.** *An* ***observable*** *is a Hermitian operator $A = A^\dagger$ on the state space. Since A is Hermitian, its eigenvalues are real and it has a spectral decomposition*

$$A = \sum_\lambda \lambda P_\lambda, \tag{8.15}$$

where the sum runs over the distinct eigenvalues of A, and P_λ is the projector onto the eigenspace of eigenvalue λ.
Measuring the observable A means performing the projective measurement $\{P_\lambda\}_\lambda$. The possible measurement outcomes are the eigenvalues λ, and the probability of outcome λ is

$$\mathbb{P}(\lambda) = \langle\psi| P_\lambda |\psi\rangle. \tag{8.16}$$

Thus an observable does two things at once. First, it specifies the subspaces onto which the measurement projects. Second, it assigns a numerical value, namely an eigenvalue, to each possible outcome. The same projective measurement may be described with different numerical labels, but in quantum error correction it is especially convenient to use the labels $+1$ and -1.

Example 8.1.5 (The observable Z)**.** *Recall that*

$$Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{8.17}$$

The eigenvectors of Z are $|0\rangle$ and $|1\rangle$, with

$$Z|0\rangle = +|0\rangle, \qquad Z|1\rangle = -|1\rangle. \tag{8.18}$$

Therefore

$$Z = (+1)|0\rangle\langle 0| + (-1)|1\rangle\langle 1|. \tag{8.19}$$

Measuring Z is the same physical measurement as measuring in the computational basis, except that we label the outcomes by $+1$ and -1 instead of by 0 and 1. If $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, then

$$\mathbb{P}(+1) = |\alpha|^2, \qquad \mathbb{P}(-1) = |\beta|^2. \tag{8.20}$$

Example 8.1.6 (The observable X)**.** *The Pauli operator*

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{8.21}$$

has eigenvectors

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \qquad |-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}, \tag{8.22}$$

with eigenvalues $+1$ and -1 respectively. Hence

$$X = (+1)|+\rangle\langle +| + (-1)|-\rangle\langle -|. \tag{8.23}$$

Measuring X means measuring in the $\{|+\rangle, |-\rangle\}$ basis. For $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$,

$$\mathbb{P}(+1) = |\langle +|\psi\rangle|^2 = \left|\frac{\alpha+\beta}{\sqrt{2}}\right|^2, \qquad \mathbb{P}(-1) = |\langle -|\psi\rangle|^2 = \left|\frac{\alpha-\beta}{\sqrt{2}}\right|^2. \tag{8.24}$$

In a circuit, an X-basis measurement can be implemented by first applying a Hadamard gate H and then measuring in the computational basis.

Expectation values

Since the outcome of measuring an observable A is a real number, we can ask for its average value over many repetitions of the same experiment.

Definition 8.1.7 (Expectation value)**.** *Let*

$$A = \sum_{\lambda} \lambda P_{\lambda} \tag{8.25}$$

be an observable, and let the system be in state $|\psi\rangle$. The ***expectation value*** *of A in state $|\psi\rangle$ is*

$$\mathbb{E}[A] = \sum_{\lambda} \lambda \mathbb{P}(\lambda). \tag{8.26}$$

Equivalently,

$$\mathbb{E}[A] = \langle\psi| A |\psi\rangle . \tag{8.27}$$

Indeed,

$$\sum_{\lambda} \lambda \mathbb{P}(\lambda) = \sum_{\lambda} \lambda \langle\psi| P_{\lambda} |\psi\rangle = \langle\psi| \left(\sum_{\lambda} \lambda P_{\lambda} \right) |\psi\rangle = \langle\psi| A |\psi\rangle . \tag{8.28}$$

This formula is often the fastest way to compute averages.

Example 8.1.8 (Expectation values of Z and X)**.** *Let $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. For Z we get*

$$\mathbb{E}[Z] = \langle\psi| Z |\psi\rangle = |\alpha|^2 - |\beta|^2. \tag{8.29}$$

This is also clear from the probabilities: the outcome is $+1$ with probability $|\alpha|^2$ and -1 with probability $|\beta|^2$.
For X we get

$$\mathbb{E}[X] = \langle\psi| X |\psi\rangle = \overline{\alpha}\beta + \overline{\beta}\alpha = 2\,\mathrm{Re}(\overline{\alpha}\beta). \tag{8.30}$$

For example, if $|\psi\rangle = |+\rangle$, then $\mathbb{E}[X] = 1$ because the measurement of X always returns $+1$.

Multi-qubit observables and parity checks

For an n-qubit system, we write Z_i for the operator that applies Z to qubit i and the identity to all other qubits. For example, on three qubits,

$$Z_1 = Z \otimes \mathbb{I} \otimes \mathbb{I}, \qquad Z_2 = \mathbb{I} \otimes Z \otimes \mathbb{I}, \qquad Z_3 = \mathbb{I} \otimes \mathbb{I} \otimes Z. \tag{8.31}$$

Similarly,

$$Z_1 Z_2 = Z \otimes Z \otimes \mathbb{I}. \tag{8.32}$$

This is an observable because it is Hermitian. Moreover, since $Z^2 = \mathbb{I}$, we have

$$(Z_1 Z_2)^2 = \mathbb{I}, \tag{8.33}$$

so its only possible eigenvalues are $+1$ and -1.

On a computational basis state $|x_1 x_2 x_3\rangle$, where each $x_i \in \{0, 1\}$,

$$Z_1 Z_2 |x_1 x_2 x_3\rangle = (-1)^{x_1 + x_2} |x_1 x_2 x_3\rangle . \tag{8.34}$$

Therefore the outcome of measuring $Z_1 Z_2$ tells us the parity of the first two bits:

$$\begin{aligned} +1 &\longleftrightarrow x_1 = x_2, \\ -1 &\longleftrightarrow x_1 \neq x_2. \end{aligned} \tag{8.35}$$

Importantly, this measurement does not tell us the values of x_1 and x_2 separately. It only tells us whether they agree or disagree.

Proposition 8.1.9 (Projectors for a ± 1 observable)**.** *Let S be an observable such that $S^2 = \mathbb{I}$. Then the only possible outcomes of measuring S are $+1$ and -1, and the corresponding projectors are*

$$P_+ = \frac{\mathbb{I} + S}{2}, \qquad P_- = \frac{\mathbb{I} - S}{2}. \tag{8.36}$$

Proof. First observe that

$$P_+^2 = \left(\frac{\mathbb{I} + S}{2}\right)^2 = \frac{\mathbb{I} + 2S + S^2}{4} = \frac{\mathbb{I} + S}{2} = P_+, \tag{8.37}$$

and similarly $P_-^2 = P_-$. Also,

$$P_+ P_- = \frac{(\mathbb{I} + S)(\mathbb{I} - S)}{4} = \frac{\mathbb{I} - S^2}{4} = 0, \qquad P_+ + P_- = \mathbb{I}. \tag{8.38}$$

Thus P_+ and P_- form a projective measurement. If $S|\psi\rangle = |\psi\rangle$, then $P_+|\psi\rangle = |\psi\rangle$ and $P_-|\psi\rangle = 0$, so the outcome is certainly $+1$. If $S|\psi\rangle = -|\psi\rangle$, then the outcome is certainly -1. □

For example, measuring $Z_1 Z_2$ is described by

$$P_+ = \frac{\mathbb{I} + Z_1 Z_2}{2}, \qquad P_- = \frac{\mathbb{I} - Z_1 Z_2}{2}. \tag{8.39}$$

On two qubits, these projectors can also be written as

$$P_+ = |00\rangle\langle 00| + |11\rangle\langle 11|, \qquad P_- = |01\rangle\langle 01| + |10\rangle\langle 10|. \tag{8.40}$$

Thus P_+ projects onto the subspace where the two bits are equal, while P_- projects onto the subspace where they are different.

Example 8.1.10 (Why parity measurements are useful)**.** *Consider the three-qubit state*

$$|\psi\rangle = \alpha\,|000\rangle + \beta\,|111\rangle\,. \tag{8.41}$$

Then

$$Z_1Z_2\,|\psi\rangle = |\psi\rangle\,, \qquad Z_2Z_3\,|\psi\rangle = |\psi\rangle\,. \tag{8.42}$$

Hence measuring either Z_1Z_2 or Z_2Z_3 gives the outcome $+1$ with probability 1, and the state is not changed.
Now suppose that a bit-flip occurred on the first qubit:

$$|\psi'\rangle = X_1\,|\psi\rangle = \alpha\,|100\rangle + \beta\,|011\rangle\,. \tag{8.43}$$

Then

$$Z_1Z_2\,|\psi'\rangle = -\,|\psi'\rangle\,, \qquad Z_2Z_3\,|\psi'\rangle = |\psi'\rangle\,. \tag{8.44}$$

So the two measurement outcomes are $(-1,+1)$. This tells us that the first qubit was flipped. Notice that the measurement reveals the error syndrome, but it does not reveal the encoded amplitudes α and β.

Commuting observables

In quantum error correction we often measure several observables, such as Z_1Z_2 and Z_2Z_3. It is important that the observables used for syndrome measurements commute.

Remark 8.1.11 (Why commutation matters)**.** *If two observables A and B commute, meaning $AB = BA$, then they can be simultaneously diagonalized: there is a basis made of vectors that are eigenvectors of both A and B. In this case, it makes sense to speak about a joint measurement outcome, such as*

$$(Z_1Z_2, Z_2Z_3) = (+1, -1). \tag{8.45}$$

If two observables do not commute, such as X and Z on the same qubit, then measuring one of them can disturb the outcome statistics of the other.

Summary

The main points to remember are the following.

- A projective measurement is specified by orthogonal projectors $\{P_m\}_m$ which sum to the identity.
- An observable is a Hermitian operator $A = \sum_\lambda \lambda P_\lambda$. Measuring A means projecting onto its eigenspaces, and the possible outcomes are its eigenvalues.
- The expectation value of A in state $|\psi\rangle$ is

$$\mathbb{E}[A] = \langle\psi|\,A\,|\psi\rangle\,. \tag{8.46}$$

- Pauli products such as Z_1Z_2 and $X_1X_2X_3X_4X_5X_6$ are observables with outcomes ± 1. These are the observables used as stabilizer measurements in quantum error-correcting codes.

8.2 The Bit-Flip Error Model

This error model is very similar in its construction as its classical counterpart. We first start with error detection.

Example 8.2.1 (Bit-flip in position 1)**.** *Consider a bit flip in position 1, then for an input state* $|\psi\rangle = \alpha|000\rangle + \beta|111\rangle$ *the error will yield,*

$$|\psi\rangle \mapsto |\psi'\rangle = X_1|\psi\rangle = (X \otimes \mathbb{I} \otimes \mathbb{I})|\psi\rangle = \alpha|100\rangle + \beta|011\rangle\,. \tag{8.47}$$

Now, the quantum equivalent of syndromes are given by measurements using the observables $Z_1Z_2 := Z \otimes Z \otimes \mathbb{I}$ and $Z_2Z_3 := \mathbb{I} \otimes Z \otimes Z$ with possible eigenvalues $\lambda_\pm = \pm 1$. As an aside, remember the parity-check matrix of the classical repetition code,

$$H = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \tag{8.48}$$

where now the first row $\begin{pmatrix} 1 & 1 & 0 \end{pmatrix}$ corresponds to the observable Z_1Z_2 and the second row $\begin{pmatrix} 0 & 1 & 1 \end{pmatrix}$ corresponds to the observable Z_2Z_3.

• Assume no error has happened such that $|\psi'\rangle = \alpha|000\rangle + \beta|111\rangle$. Then, the actions of the observables Z_1Z_2 and Z_2Z_3 on $|\psi'\rangle$ are,

$$Z_1Z_2|\psi'\rangle = \alpha|000\rangle + (-1)(-1)\beta|111\rangle = \alpha|000\rangle + \beta|111\rangle = (+1)|\psi'\rangle\,. \tag{8.49}$$

$$Z_2Z_3|\psi'\rangle = \alpha|000\rangle + (-1)(-1)\beta|111\rangle = \alpha|000\rangle + \beta|111\rangle = (+1)|\psi'\rangle\,. \tag{8.50}$$

• Assume now that one bit was flipped (without loss of generality assume the first one) such that $|\psi'\rangle = \alpha|100\rangle + \beta|011\rangle$. Then, the actions of the observables Z_1Z_2 and Z_2Z_3 on $|\psi'\rangle$ are,

$$Z_1Z_2|\psi'\rangle = -\alpha|100\rangle - \beta|011\rangle = (-1)|\psi'\rangle\,. \tag{8.51}$$

$$Z_2Z_3|\psi'\rangle = \alpha|100\rangle + \beta|011\rangle = (+1)|\psi'\rangle\,. \tag{8.52}$$

In summary, with the measurements of the **stabilizers** Z_1Z_2 and Z_2Z_3 we obtain:

$$(Z_1Z_2, Z_2Z_3) \equiv (+1,+1) \longleftrightarrow \text{no bit flip}$$

$$(Z_1Z_2, Z_2Z_3) \equiv (-1,+1) \longleftrightarrow \text{bit flip in position 1}$$

$$(Z_1Z_2, Z_2Z_3) \equiv (-1,-1) \longleftrightarrow \text{bit flip in position 2}$$

$$(Z_1Z_2, Z_2Z_3) \equiv (+1,-1) \longleftrightarrow \text{bit flip in position 3}$$

The brackets (Z_1Z_2, Z_2Z_3) are considered as our new syndromes. Finally observe that $|\psi'\rangle$ is an eigenvector of Z_1, Z_2 and Z_3, as a result we have no state perturbation!
We now move on with error correction. In this case it is relatively straightforward.

$$(Z_1Z_2, Z_2Z_3) \equiv (+1,+1) \longrightarrow \text{do nothing}$$

$$(Z_1Z_2, Z_2Z_3) \equiv (-1,+1) \longrightarrow \text{apply } X_1$$

$$(Z_1Z_2, Z_2Z_3) \equiv (-1,-1) \longrightarrow \text{apply } X_2$$

$$(Z_1Z_2, Z_2Z_3) \equiv (+1,-1) \longrightarrow \text{apply } X_3$$

After applying X_1, X_2 or X_3 we will get back the state $|\psi\rangle = \alpha|000\rangle + \beta|111\rangle$. Note that X_1, X_2, X_3 need to be applied **after** the Zs to $|\psi'\rangle$, which is **not** an eigenvector of X_1, X_2, X_3.

In summary, for a bit-flip, $|\psi\rangle = \alpha|000\rangle + \beta|111\rangle \mapsto |\psi'\rangle = X_1|\psi\rangle = \alpha|100\rangle + \beta|011\rangle$. For error detection we find $Z_1Z_2|\psi'\rangle = (-1)|\psi'\rangle$ and $Z_2Z_3|\psi'\rangle = (+1)|\psi'\rangle$. Finally, for error correction set $X_1|\psi'\rangle = X_1X_1|\psi\rangle = X_1^2|\psi\rangle = |\psi\rangle$.

8.3 The Phase-Flip Error Model

Recall that the action of Z on the computational basis states is $Z|0\rangle = |0\rangle$ and $Z|1\rangle = -|1\rangle$. For a phase-flip, set a new code in the form $|0\rangle \mapsto |+++\rangle$ and $|1\rangle \mapsto |---\rangle$ such that $|\psi\rangle = \alpha|+++\rangle + \beta|---\rangle$. Now, the action of Z on $|+\rangle$ and $|-\rangle$ is $Z|+\rangle = |-\rangle$ and $Z|-\rangle = |+\rangle$ (we are in the same scenario as before).
Now, suppose that we have a phase-flip in the first qubit. Then,

$$Z_1|\psi\rangle = \alpha|-++\rangle + \beta|+--\rangle. \tag{8.53}$$

For error detection, we will use the observables X_1X_2 and X_2X_3. For error correction, if we have $(-1,+1)$ we apply Z_1 to recover the original state.

8.4 Shor's Code

In the two previous sections we have seen how to deal with bit-flip errors and phase-flip errors on their own. But how do we handle the situation if we have bit-flip and phase-flip errors together? For this, we will use Shor's code which corresponds to a concatenation of the two previous codes. The idea can be seen in two steps.

- The following transformations are useful to deal with phase-flips:

$$|0\rangle \mapsto |+++\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle+|1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle+|1\rangle}{\sqrt{2}} \tag{8.54}$$

$$|1\rangle \mapsto |---\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle-|1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle-|1\rangle}{\sqrt{2}} \tag{8.55}$$

- The following transformations are useful to deal with bit-flips:

$$|0\rangle \mapsto \frac{|000\rangle+|111\rangle}{\sqrt{2}} \otimes \frac{|000\rangle+|111\rangle}{\sqrt{2}} \otimes \frac{|000\rangle+|111\rangle}{\sqrt{2}} =: |0\rangle_{\text{Shor}} \tag{8.56}$$

$$|1\rangle \mapsto \frac{|000\rangle-|111\rangle}{\sqrt{2}} \otimes \frac{|000\rangle-|111\rangle}{\sqrt{2}} \otimes \frac{|000\rangle-|111\rangle}{\sqrt{2}} =: |1\rangle_{\text{Shor}} \tag{8.57}$$

Here, we note that $k = 1$ and $n = 9$ and codewords are of the form $|\psi\rangle = \alpha|0\rangle_{\text{Shor}} + \beta|1\rangle_{\text{Shor}}$.

Proposition 8.4.1. *Shor's code protects against a bit-flip and/or a phase-flip.*

Proof. Consider an initial state $|\psi\rangle = \alpha|0\rangle_{\text{Shor}} + \beta|1\rangle_{\text{Shor}}$ and suppose that the output state $|\psi'\rangle$ suffered from a bit-flip and/or a phase-flip. For the bit-flip, we may consider the stabilizers $Z_1Z_2, Z_2Z_3, Z_4Z_5, Z_5Z_6, Z_7Z_8$ and Z_8Z_9. These measurements will not perturb the state and they will indeed provide some information on the bit-flip. Next we can consider the stabilizers $X_1X_2X_3X_4X_5X_6$ and $X_4X_5X_6X_7X_8X_9$ such that they will provide information on the phase-flip. Again, these stabilizers will not perturb the state.

□

Remark that these operators commute and $|\psi'\rangle$ is an eigenvector of all of them.

For error correction, the game is just to apply the correct Z gate or X gate. To see how this goes, let us consider some examples.

Example 8.4.2 (Bit-flip error on bit 3)**.** *In this case, only Z_2Z_3 has eigenvalue -1. Thus one applies X_3 to correct the error.*

Example 8.4.3 (Phase-flip error on bit 5)**.** *In this case, both $X_1X_2X_3X_4X_5X_6$ and $X_4X_5X_6X_7X_8X_9$ have eigenvalue -1. Thus one doesn't know where the phase-flip occurred among bits 4,5 or 6, however one can still correct the error by applying $Z_4Z_5Z_6$.*

Example 8.4.4 (Bit-flip and phase-flip error on bit 4)**.** *In this case, Z_4Z_5, $X_1X_2X_3X_4X_5X_6$ and $X_4X_5X_6X_7X_8X_9$ have eigenvalue -1. This bit-phase flip can be corrected by applying both X_4 and Z_4 [note that depending on the order of application, this might generate a global (-1) phase since $X_4Z_4 = -Z_4X_4$. However, the error will be corrected anyway.]*

8.5 Steane's Code

Recall the Hamming code Hamming$(7,4,3)$ with its parity-check matrix,

$$H = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}, \tag{8.58}$$

along with the generator matrix,

$$G = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}. \tag{8.59}$$

We recall the sets $\mathcal{C}_H = \{c = u \cdot G, u \in \mathbb{F}_2^4\} = \{c \in \mathbb{F}_2^7 : Hc^T = 0\}$ and $\mathcal{C}_H^\perp = \{c = u \cdot H, u \in \mathbb{F}_2^3\}$. Steane's code is more efficient than Shor's code. The insight is to take the classical Hamming code (which

corrects single-bit flips) and turn it into a quantum code via the CSS (Calderbank–Shor–Steane) construction. The Hamming code (length 7, dimension 4) has a (3×7) parity-check matrix H. Its row-space (dimension 3) is a classical code of distance 4. In CSS language, one chooses two nested classical codes $C_2 \subset C_1$ here with $C_1 = \text{Hamming}(7,4,3)$ and $C_2 = C_1^{\perp}$. This nesting gives rise to a Steane$(7,1,3)$ quantum code. The transformations $|0\rangle_S$ and $|1\rangle_S$ are thus superpositions of the 8 classical length-7 codewords in C_2. We may set the following transformations for $|0\rangle_S$ and $|1\rangle_S$, where we are looking for the row-space of H:

$$
\begin{aligned}
|0\rangle \mapsto |0\rangle_S &= \frac{1}{\sqrt{8}} \sum_{c \in \mathcal{C}_H^{\perp}} |c\rangle \\
&= \frac{1}{\sqrt{8}} \big(|0000000\rangle + |1010101\rangle + |0110011\rangle + |0001111\rangle + |0111100\rangle + |1011010\rangle + |1101001\rangle + |1100110\rangle \big)
\end{aligned}
$$

$$
\begin{aligned}
|1\rangle \mapsto |1\rangle_S = X|0\rangle_S &= \frac{1}{\sqrt{8}} \sum_{c \in \mathcal{C}_H^{\perp}} |c \oplus 1111111\rangle \\
&= \frac{1}{\sqrt{8}} \big(|1111111\rangle + |0101010\rangle + |1001100\rangle + |1110000\rangle + |1000011\rangle + |0100101\rangle + |0010110\rangle + |0011001\rangle \big)
\end{aligned}
$$

The corresponding stabilizers (operators that leave the code invariant) are given by $g_1 = X_4X_5X_6X_7$, $g_2 = X_2X_3X_6X_7$, $g_3 = X_1X_3X_5X_7$ and $g_4 = Z_4Z_5Z_6Z_7$, $g_5 = Z_2Z_3Z_6Z_7$, $g_6 = Z_1Z_3Z_5Z_7$ (these correspond to the 1's in the matrix H). Note that these stabilizers form a **group S**, which is a group generated by these stabilizers.

Proposition 8.5.1. *The stabilizers commute, i.e. $g_ig_j = g_jg_i$ for all i, j.*

Proof. We will show this for a specific example, where the rest may be shown in a similar way. Recall that $X_iZ_i = -Z_iX_i$,

$$g_2g_6 = X_2X_3X_6X_7Z_1Z_3Z_5Z_7 = Z_1X_2X_3Z_3Z_5X_6X_7Z_7 = Z_1X_2Z_3X_3Z_5X_6Z_7X_7 = g_6g_2. \quad (8.60)$$

In the second equality we have just rearranged the gates in order of their action. For the general case, remark that there is always an even number of common X's and Z's between two stabilizers. □

Proposition 8.5.2. *Codewords are invariant to stabilizers, i.e. $g_i|0\rangle_S = |0\rangle_S$ and $g_i|1\rangle_S = |1\rangle_S$.*

Proof. We again prove this for a specific example, where the rest may be shown in a very similar way.

$$g_1|0\rangle_S = g_1 \frac{1}{\sqrt{8}} \sum_{c \in \mathcal{C}_H^{\perp}} |c\rangle = \frac{1}{\sqrt{8}} \sum_{c \in \mathcal{C}_H^{\perp}} |c \oplus 0001111\rangle = \frac{1}{\sqrt{8}} \sum_{c \in \mathcal{C}_H^{\perp}} |c\rangle,$$

where we have used the fact that $|c \oplus 0001111\rangle$ is an element of the group $\mathcal{C}_H^\perp$. For $|1\rangle_S$ one finds,

$$\begin{aligned} g_1 |1\rangle_S = X_4X_5X_6X_7 \frac{1}{\sqrt{8}} \sum_{c\in\mathcal{C}_H^\perp} |c \oplus 1111111\rangle &= \frac{1}{\sqrt{8}} \sum_{c\in\mathcal{C}_H^\perp} |c \oplus 1111111 \oplus 0001111\rangle \\ &= \frac{1}{\sqrt{8}} \sum_{c\in\mathcal{C}_H^\perp} |c \oplus 0001111 \oplus 1111111\rangle \\ &= \frac{1}{\sqrt{8}} \sum_{c\in\mathcal{C}_H^\perp} |c \oplus 1111111\rangle = |1\rangle_S , \end{aligned}$$

where we have used the fact that $c \oplus 0001111$ is an element of the group $\mathcal{C}_H^\perp$. Thus we have invariance for g_1, and the rest follows. We also have an identical reasoning for Z-stabilizers.

□

Remark 8.5.3. *Note that the invariance property demonstrates that $|\psi\rangle$ is a codeword of the Steane code if and only if for all $g \in S$, $g(\alpha |0\rangle_S + \beta |1\rangle_S) = g|\psi\rangle = |\psi\rangle$.*

Proposition 8.5.4. *The Steane code can correct any single error, i.e. either $\{X_i, Y_i, Z_i\}_{i=1}^7$.*

Proof. We again prove this for a specific example, where the rest may be shown in a very similar way. Recall that the X and Z gates commute between each other. For $g_1 = X_4X_5X_6X_7$, consider a bit-flip in position 7, $|\psi'\rangle = X_7|\psi\rangle$. Then,

$$g_1 |\psi'\rangle = g_1X_7|\psi\rangle = X_7g_1|\psi\rangle = X_7|\psi\rangle = |\psi'\rangle , \tag{8.61}$$

where we have used the invariance of codewords to stabilizers [Proposition (8.5.2)].

□

Example 8.5.5 (Syndrome for a bit-flip in position 7)**.** *Let $|\psi'\rangle = X_7(\alpha|0\rangle_S + \beta|1\rangle_S)$. Then: $g_1|\psi'\rangle = (+1)|\psi'\rangle$, $g_2|\psi'\rangle = (+1)|\psi'\rangle$, $g_3|\psi'\rangle = (+1)|\psi'\rangle$, $g_4|\psi'\rangle = (-1)|\psi'\rangle$, $g_5|\psi'\rangle = (-1)|\psi'\rangle$ and $g_6|\psi'\rangle = (-1)|\psi'\rangle$. Thus the syndrome is $(1,1,1,-1,-1,-1)$. From the three last positions in the syndrome $(-1,-1,-1)$, we get the error 111 corresponding to the binary decomposition of 7, i.e. X error in position 7.*

Example 8.5.6 (Other examples of syndromes)**.** *Consider the following:*

- *X_3 error $\longrightarrow$ syndrome $= (1,1,1,\underbrace{1,-1,-1}) \longrightarrow 011 \longrightarrow X$ error in position 3.*
- *Z_4 error $\longrightarrow$ syndrome $= (\underbrace{-1,1,1},1,1,1) \longrightarrow 100 \longrightarrow Z$ error in position 4.*
- *$Y_6 = iX_6Z_6$ error $\longrightarrow$ syndrome $= (\underbrace{-1,-1,1}_{110},\underbrace{-1,-1,1}_{110}) \longrightarrow$ error in position 6.*

8.6 Building Reliable Quantum Gates using the Steane Code

First recall the seven gates we will consider in this section,

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = iXZ = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

$$S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix}, \quad CNOT = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

With the Steane code we have the following encoding, yielding a 7 qubit state:

$$\alpha |0\rangle + \beta |1\rangle \xrightarrow{\text{encoding}} \alpha |0\rangle_S + \beta |1\rangle_S . \tag{8.62}$$

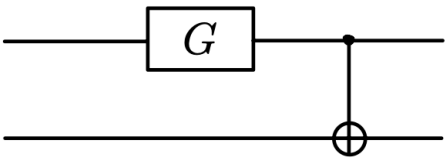
- The **idea** in this section is the following: On the one hand, for some gate G,

$$G(\alpha |0\rangle + \beta |1\rangle) \xrightarrow{\text{encoding}} |\psi\rangle_S , \tag{8.63}$$

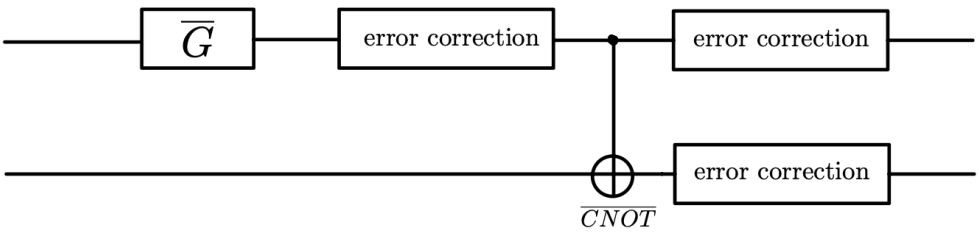
and on the other hand, for some gate $\overline{G}$,

$$\alpha |0\rangle + \beta |1\rangle \xrightarrow{\text{encoding}} \alpha |0\rangle_S + \beta |1\rangle_S \xrightarrow{\overline{G}} \overline{G}(\alpha |0\rangle_S + \beta |1\rangle_S). \tag{8.64}$$

- The **aim** is to find seven gates $\overline{G} = \{\overline{X}, \overline{Y}, \overline{Z}, \overline{H}, \overline{S}, \overline{T}, \overline{CNOT}\}$ that act on all the bits of the Steane code. We want the two outputs in (8.63) and (8.64) to match, such that $|\psi\rangle_S = \overline{G}(\alpha |0\rangle_S + \beta |1\rangle_S)$. We now understand that we want an alternative to first applying a gate G on a state $|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$ and then encoding it, by first encoding the state $|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$ and then applying some gate $\overline{G}$. From the point of view of a circuit, we are looking (for example) to go from a circuit of the form,



to one where we consider a gate $\overline{H}$ as a 7-qubit gate, and $\overline{CNOT}$ as a 14-qubit gate, [error correction gate = decoder gate]



Proposition 8.6.1. *For $U = \{X, Y, Z\}$, $\overline{U}$ is given by $\overline{U} = U_1 U_2 ... U_7 = U^{\otimes 7}$.*

Proof. We check this gate by gate.

- For gate X on the one hand, with $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$,

$$X(\alpha|0\rangle + \beta|1\rangle) = \alpha|1\rangle + \beta|0\rangle \xrightarrow{\text{Steane}} \alpha|1\rangle_S + \beta|0\rangle_S . \tag{8.65}$$

On the other hand,

$$\overline{X}(\alpha|0\rangle_S + \beta|1\rangle_S) = \alpha X^{\otimes 7}|0\rangle_S + \beta X^{\otimes 7}|1\rangle_S = \alpha|1\rangle_S + \beta|0\rangle_S . \tag{8.66}$$

- For gate Z on the one hand, with $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$,

$$Z(\alpha|0\rangle + \beta|1\rangle) = \alpha|0\rangle - \beta|1\rangle \xrightarrow{\text{Steane}} \alpha|0\rangle_S - \beta|1\rangle_S . \tag{8.67}$$

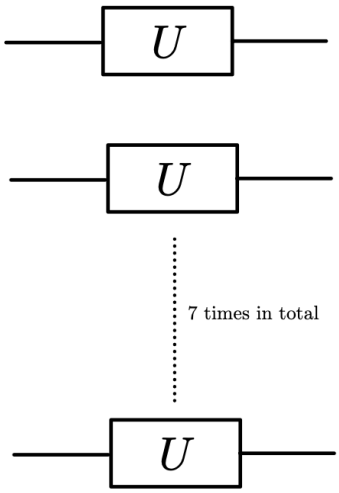
On the other hand,

$$\overline{Z}(\alpha|0\rangle_S + \beta|1\rangle_S) = \alpha Z^{\otimes 7}|0\rangle_S + \beta Z^{\otimes 7}|1\rangle_S = \alpha|0\rangle_S - \beta|1\rangle_S , \tag{8.68}$$

where we have used the facts that $|0\rangle_S$ has an even number of 1's and $|1\rangle_S$ has an odd number of 1's.

- $Y = iXZ$ is also fine since X and Z work independently.

□



Proposition 8.6.2. *For the gate S, $\overline{S}$ is given by $\overline{S} = Z_1 S_1 ... Z_7 S_7 = (ZS)^{\otimes 7}$.*

Proof. Since $S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$, we have for a state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$,

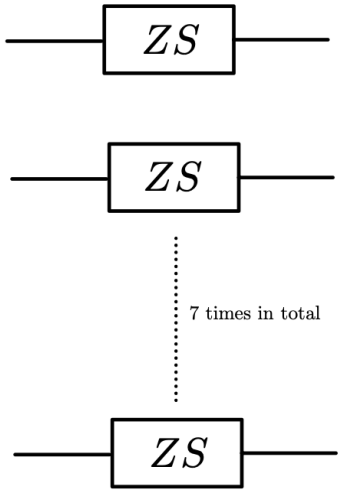
$$S(\alpha|0\rangle + \beta|1\rangle) = \alpha|0\rangle + i\beta|1\rangle \xrightarrow{\text{encoding}} \alpha|0\rangle_S + i\beta|1\rangle_S . \tag{8.69}$$

On the other hand, for $\alpha|0\rangle + \beta|1\rangle \xrightarrow{\text{encoding}} \alpha|0\rangle_S + \beta|1\rangle_S$, we need to find a gate $\overline{S}$ such that,

$$\overline{S}(\alpha|0\rangle_S + \beta|1\rangle_S) = \alpha|0\rangle_S + i\beta|1\rangle_S . \tag{8.70}$$

However, remark that $\overline{S} = S_1 ... S_7 = S^{\otimes 7}$ would be fine for the term $\alpha |0\rangle_S$ (since we get either a factor of $i^0 = 1$ or $i^4 = 1$) but not for $i\beta |1\rangle_S$ (since we would get $-i$ instead of i). Thus, by setting $\overline{S} = Z_1 S_1 ... Z_7 S_7 = (ZS)^{\otimes 7}$ will yield an extra minus sign.

□



Proposition 8.6.3. *For the gate H, $\overline{H}$ is given by $\overline{H} = H_1 ... H_7 = (H)^{\otimes 7}$.*

Proof. For gate H on the one hand, with $|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$,

$$H(\alpha |0\rangle + \beta |1\rangle) = \alpha \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}} \right) + \beta \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}} \right) \xrightarrow{\text{Steane}} \alpha \left(\frac{|0\rangle_S + |1\rangle_S}{\sqrt{2}} \right) + \beta \left(\frac{|0\rangle_S - |1\rangle_S}{\sqrt{2}} \right). \quad (8.71)$$

On the other hand, for an encoding $\alpha |0\rangle + \beta |1\rangle \xrightarrow{\text{Steane}} \alpha |0\rangle_S + \beta |1\rangle_S$ we would like,

$$\overline{H}(\alpha |0\rangle_S + \beta |1\rangle_S) = \alpha \left(\frac{|0\rangle_S + |1\rangle_S}{\sqrt{2}} \right) + \beta \left(\frac{|0\rangle_S - |1\rangle_S}{\sqrt{2}} \right). \quad (8.72)$$

The tricky part here is to make sure that we would get constant flips in the signs (e.g. when it hits a qubit $|1\rangle$). For this, we have to verify two key points. In the following, we will use the conjugation relations $HX = ZH \iff HXH = Z$ and $HZ = XH \iff HZH = X$ (recall $H^2 = \mathbb{I}$).

(i) We need to make sure that $\overline{H}$ keeps codewords in the code (i.e. maps codewords into codewords). For $|\psi\rangle_S = \alpha |0\rangle_S + \beta |1\rangle_S$ we require, for all $g_i \in S$,

$$g_i \overline{H} |\psi\rangle_S = \overline{H} |\psi\rangle_S . \quad (8.73)$$

In other words, this is equivalent to verifying that $\overline{H} g_i \overline{H} \in S$. This is true, since for example $\overline{H} g_1 \overline{H} = (H_1 ... H_7) X_4 X_5 X_6 X_7 (H_1 ... H_7) = (H_4 X_4 H_4) ... (H_7 X_7 H_7) = Z_4 Z_5 Z_6 Z_7 \in S$. It is not difficult to check that the remaining cases also hold.

(ii) The conjugation relations also hold for $\overline{H}$. Indeed,

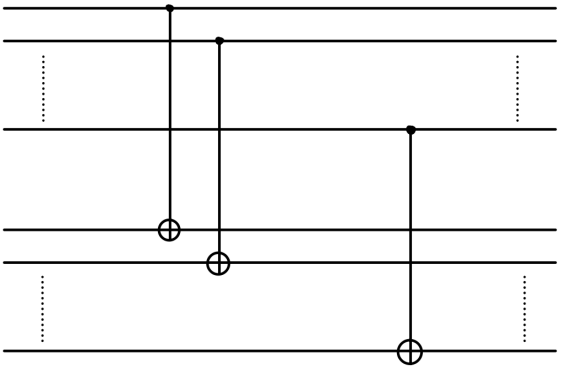
$$\overline{H}\ \overline{X}\ \overline{H} = (H_1 ... H_7)(X_1 ... X_7)(H_1 ... H_7) = (H_1 X_1 H_1) ... (H_7 X_7 H_7) = Z_1 ... Z_7 = \overline{Z}. \quad (8.74)$$

$$\overline{H}\ \overline{Z}\ \overline{H} = (H_1...H_7)(Z_1...Z_7)(H_1...H_7) = (H_1Z_1H_1)...(H_7Z_7H_7) = X_1...X_7 = \overline{X}. \tag{8.75}$$

Therefore, from points (i) and (ii) we conclude that relation (8.72) holds.

□

We will not explicitly study $CNOT \mapsto \overline{CNOT}$, however we note that an error in $\overline{CNOT}$ might propagate but it will do so into two separate 7-qubit blocks (which is fine). The circuit for $\overline{CNOT}$ is given by,



Finally, we will not study either the gate $T \mapsto \overline{T}$ as it needs more attention than the other gates.