Paper review for Steane code (CSS code)

Zheyuan Wu

December 9, 2025

Abstract

A quantum error-correcting code is defined to be a unitary mapping (encoding) of k qubits (two-state quantum systems) into a subspace of the quantum state space of n qubits such that if any t of the qubits undergo arbitrary decoherence, not necessarily independently, the resulting n qubit state can be used to faithfully reconstruct the original quantum state of the k encoded qubits.

For this project, I will build a self-contained report for Steane's code [3] that is readable for undergraduates who have just taken some coding and information theory classes, assuming no knowledge of quantum computing and quantum information theory.

1 Problem setting and motivation

We will use the notation defined in class and $[n] = \{1, \ldots, n-1, n\}$, (yes, we use 1-indexed in computer science), each in natural numbers. And $\mathbb{F}_q$ is the finite field with q elements.

This notation system is annoying since in mathematics, A^* is the transpose of A, but since we are using literature in physics, we keep the notation of A^* to denote the transpose of an element. (A could be a complex number, or a complex matrix. In mathematics we use $\overline{A}$ to denote the transpose) In this report, I will try to make the notation as consistent as possible and follow the **physics** convention in this report. So every vector you see will be in $|\psi\rangle$ form. And we will avoid using the $\langle v, w\rangle$ notation for inner product, as it is used in math, we will use $\langle v|w\rangle$ or $\langle v, w\rangle$ to denote the inner product.

Asymptotic rate $k/n = 1 - 2H_2(2t/n)$, where H_2 is the binary entropy function

$$H_2 = -p\log_2(p) - (1-p)\log_2(1-p)$$

1.1 Linear algebra 102

First, we will introduce some notations in linear algebra [1] that we will use in quantum information theory and quantum computing.

The main vector space we are interested in is $\mathbb{C}^n$; therefore, all the linear operators we defined are from $\mathbb{C}^n$ to $\mathbb{C}^n$.

We denote a vector in vector space as $|\psi\rangle = (z_1, \ldots, z_n)$ (might also be infinite dimensional, and $z_i \in \mathbb{C}$).

A natural inner product space defined on $\mathbb{C}^n$ is given by the Hermitian inner product:

$$\langle\psi|\varphi\rangle = \sum_{i=1}^{n} z_i z_i^*$$

This satisfies the following properties:

1. $\langle\psi|\sum_i \lambda_i |\varphi\rangle = \sum_i \lambda_i \langle\psi|\varphi\rangle$ (linear on the second argument. Note that in physics [4] we use linear on the second argument and conjugate linear on the first argument. But in

math, we use linear on the first argument and conjugate linear on the second argument [1]. As promised in the beginning, we will use the physics convention in this report.)

2. $\langle\varphi|\psi\rangle = (\langle\psi|\varphi\rangle)^*$
3. $\langle\psi|\psi\rangle \geq 0$ with equality if and only if $|\psi\rangle = 0$

Here ψ is just a label for the vector, and you don't need to worry about it too much. This is also called the ket, where the counterpart:

- $\langle\psi\rangle$ is called the bra, used to denote the vector dual to ψ; such an element is a linear functional if you really want to know what that is.
- $\langle\psi|\varphi\rangle$ is the inner product between two vectors, and $\langle\psi| A |\varphi\rangle$ is the inner product between $A|\varphi\rangle$ and $\langle\psi|$, or equivalently $A^\dagger \langle\psi|$ and $|\varphi\rangle$.
- Given a complex matrix $A = \mathbb{C}^{n\times n}$,
  1. A^* is the complex conjugate of A. i.e.,
  $$A = \begin{bmatrix} 1+i & 2+i & 3+i \\ 4+i & 5+i & 6+i \\ 7+i & 8+i & 9+i \end{bmatrix}, A^* = \begin{bmatrix} 1-i & 2-i & 3-i \\ 4-i & 5-i & 6-i \\ 7-i & 8-i & 9-i \end{bmatrix}$$
  2. $A^\top$ denotes the transpose of A. i.e.,
  $$A = \begin{bmatrix} 1+i & 2+i & 3+i \\ 4+i & 5+i & 6+i \\ 7+i & 8+i & 9+i \end{bmatrix}, A^\top = \begin{bmatrix} 1+i & 4+i & 7+i \\ 2+i & 5+i & 8+i \\ 3+i & 6+i & 9+i \end{bmatrix}$$
  3. $A^\dagger = (A^*)^\top$ denotes the complex conjugate transpose, referred to as the adjoint, or Hermitian conjugate of A. i.e.,
  $$A = \begin{bmatrix} 1+i & 2+i & 3+i \\ 4+i & 5+i & 6+i \\ 7+i & 8+i & 9+i \end{bmatrix}, A^\dagger = \begin{bmatrix} 1-i & 4-i & 7-i \\ 2-i & 5-i & 8-i \\ 3-i & 6-i & 9-i \end{bmatrix}$$
  4. A is unitary if $A^\dagger A = AA^\dagger = I$.
  5. A is hermitian (self-adjoint in mathematics literature) if $A^\dagger = A$.

1.1.1 Motivation of Tensor product

Recall from the traditional notation of product space of two vector spaces V and W, that is, $V \times W$, is the set of all ordered pairs $(|v\rangle, |w\rangle)$ where $|v\rangle \in V$ and $|w\rangle \in W$.

The space has dimension $\dim V + \dim W$.

We want to define a vector space with the notation of multiplication of two vectors from different vector spaces.

That is

$$(|v_1\rangle + |v_2\rangle) \otimes |w\rangle = (|v_1\rangle \otimes |w\rangle) + (|v_2\rangle \otimes |w\rangle)$$
$$|v\rangle \otimes (|w_1\rangle + |w_2\rangle) = (|v\rangle \otimes |w_1\rangle) + (|v\rangle \otimes |w_2\rangle)$$

and enables scalar multiplication by

$$\lambda(|v\rangle \otimes |w\rangle) = (\lambda|v\rangle) \otimes |w\rangle = |v\rangle \otimes (\lambda|w\rangle)$$

And we wish to build a way to associate the basis of V and W with the basis of $V \otimes W$. That makes the tensor product a vector space with dimension $\dim V \times \dim W$.

Definition 1. *Definition of linear functional*
A linear functional is a linear map from V to $\mathbb{F}$.

Note the difference between a linear functional and a linear map.
A generalized linear map is a function $f : V \to W$ satisfying the condition.

- $f(|u\rangle + |v\rangle) = f(|u\rangle) + f(|v\rangle)$
- $f(\lambda\,|v\rangle) = \lambda f(|v\rangle)$

Definition 2. *A bilinear functional is a bilinear function $\beta : V \times W \to \mathbb{F}$ satisfying the condition that $|v\rangle \to \beta(|v\rangle, |w\rangle)$ is a linear functional for all $|w\rangle \in W$ and $|w\rangle \to \beta(|v\rangle, |w\rangle)$ is a linear functional for all $|v\rangle \in V$.*

The vector space of all bilinear functionals is denoted by $\mathcal{B}(V, W)$.

Definition 3. *Let V, W be two vector spaces.*
Let V' and W' be the dual spaces of V and W, respectively, that is $V' = \{\psi : V \to \mathbb{F}\}$ and $W' = \{\phi : W \to \mathbb{F}\}$, ψ, ϕ are linear functionals.
The tensor product of vectors $v \in V$ and $w \in W$ is the bilinear functional defined by $\forall(\psi, \phi) \in V' \times W'$ given by the notation

$$(v \otimes w)(\psi, \phi) \coloneqq \psi(v)\phi(w)$$

The tensor product of two vector spaces V and W is the vector space $\mathcal{B}(V', W')$
Notice that the basis of such vector space is the linear combination of the basis of V' and W', that is, if $\{e_i\}$ is the basis of V' and $\{f_j\}$ is the basis of W', then $\{e_i \otimes f_j\}$ is the basis of $\mathcal{B}(V', W')$.
That is, every element of $\mathcal{B}(V', W')$ can be written as a linear combination of the basis.
Since $\{e_i\}$ and $\{f_j\}$ are bases of V' and W', respectively, then we can always find a set of linear functionals $\{\phi_i\}$ and $\{\psi_j\}$ such that $\phi_i(e_j) = \delta_{ij}$ and $\psi_j(f_i) = \delta_{ij}$.
Here $\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$ is the Kronecker delta.

$$V \otimes W = \left\{ \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} \phi_i(v) \psi_j(w) : \phi_i \in V', \psi_j \in W' \right\}$$

Note that $\sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} \phi_i(v) \psi_j(w)$ is a bilinear functional that maps $V' \times W'$ to $\mathbb{F}$.
This enables basis-free construction of vector spaces with proper multiplication and scalar multiplication.
This vector space is equipped with the unique inner product $\langle v \otimes w, u \otimes x \rangle_{V \otimes W}$ defined by

$$\langle v \otimes w, u \otimes x \rangle = \langle v, u \rangle_V \langle w, x \rangle_W$$

In practice, we ignore the subscript of the vector space and just write $\langle v \otimes w, u \otimes x \rangle = \langle v, u \rangle \langle w, x \rangle$.
In this report, we will use the following definition for a quantum system. There are many variations for the definition of Hilbert space (In mathematics, the Hilbert space is the complete inner product space, but here in physics, for most of the time, we ignore the complete requirement and just use the inner product space).
We use $\mathcal{H}$ to denote the Hilbert space. (In mathematics we use $\mathscr{H}$)

Definition 4. *A two-state quantum system is the Hilbert space $\mathcal{H}_2^n$ over n qubits (finite dimensional) generated by the complex vectors $|b_0\rangle, |b_1\rangle, \ldots, |b_{2^n-1}\rangle$ where b_i is the representation of the number i in binary. It is equivalent to $\mathcal{H}_2^{\otimes n}$ (The tensor product of n two-state quantum systems $\mathcal{H}_2$).*

Each $\mathcal{H}_2$ is a representation of a qubit quantum system.[3]

One might ask, what is the fundamental difference between a quantum system and a classical system, and why can we not directly apply those theorems in classical computers to a quantum computer? It turns out that quantum error-correcting codes are hard due to the following definitions and features for quantum computing.

Definition 5. *All quantum operations can be constructed by composing four kinds of transformations: (adapted from Chapter 10 of [2])*

1. *Unitary operations. $U(\cdot)$ for any quantum state. It is possible to apply a non-unitary operation for an open quantum system, but that is usually not the focus for quantum computing and usually leads to non-recoverable loss of information that we wish to obtain.*
2. *Extend the system. Given a quantum state $\rho \in \mathcal{H}^N$, we can extend it to a larger quantum system by "entangle" (For this report, you don't need to worry for how quantum entanglement works) it with some new states $\sigma \in \mathcal{H}^K$ (The space where the new state dwells is usually called ancilla system) and get $\rho' = \rho \otimes \sigma \in \mathcal{H}^N \otimes \mathcal{K}$.*
3. *Partial trace. Given a quantum state $\rho \in \mathcal{H}^N$ and some reference state $\sigma \in \mathcal{H}^K$, we can trace out some subsystems and get a new state $\rho' \in \mathcal{H}^{N-K}$.*
4. *Selective measurement. Given a quantum state, we measure it and get a classical bit; unlike the classical case, the measurement is a probabilistic operation. (More specifically, this is some projection to a reference state corresponding to a classical bit output. For this report, you don't need to worry about how such a result is obtained and how the reference state is constructed.)*

During quantum computing in practice, it is hard to isolate the quantum from the environment, and the one actually doing computations. This results in decoherence process $\rho \to \rho'$ where $\mathcal{H}_K$ is the error from the environment, which leads our system to extend with environment and "lose information", that is, if you make measurement based on initial states but assume no noise is introduced, the distribution of result will be different than the one expected and it is impossible to recover if we don't know what noise is (which holds true for most of the cases). One intuitive explanation for that is similar to Jenson's inequality, where some concentration on getting a certain value is "dispersed" to extra outcome states where measurement is not expected.

Generally, the following few operations are mostly used for creating quantum circuits, which can be found in section 4.2 of [4]

Definition 6. *Let $|psi\rangle = a|0\rangle + b|1\rangle$ be a single qubit vector where a, b are complex numbers satisfying $|a|^2 + |b|^2 = 1$. Operations on a qubit must preserve this norm, and thus are described by 2×2 unitary matrices. The most commonly used are Pauli matrices, listed as follows:*

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

And a few other gates that will be used in this report:
Hadamard gate (H)

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

And additional two bit operation $|c\rangle |t\rangle \to |c\rangle |t \oplus c\rangle$ *given as*

$$\text{CNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Measurement will destroy the quantum state, and recovery is impossible.

Theorem 7. *No-cloning theorem: There is no way to clone a general quantum state via a unitary operation.*

The proof is adapted from the book [4]

Proof. Suppose we have a quantum system with two slots A and B, the data slot starts out in an unknown but pure quantum state $|\psi\rangle$. This is the state which is to be copied into slot Bm the target slot. We assume that the target slot starts in some standard pure state $|s\rangle$. Thus, the initial state of the copying machine is $|\psi\rangle \otimes |s\rangle$.

Assume there exists some unitary operator U such that $U(|\psi\rangle \otimes |s\rangle) = |\psi\rangle \otimes |\psi\rangle$.

Consider two pure states $|\psi\rangle$ and $|\varphi\rangle$, such that $U(|\psi\rangle \otimes |s\rangle) = |\psi\rangle \otimes |\psi\rangle$ and $U(|\varphi\rangle \otimes |s\rangle) = |\varphi\rangle \otimes |\varphi\rangle$. The inner product of the two equation yields:

$$\langle\psi|\varphi\rangle = (\langle\psi|\varphi\rangle)^2$$

This equation has only two solutions, either $\langle\psi|\varphi\rangle = 0$ or $\langle\psi|\varphi\rangle = 1$.

If $\langle\psi|\varphi\rangle = 0$, then $|\psi\rangle = |\varphi\rangle$, no cloning for trivial case.

If $\langle\psi|\varphi\rangle = 1$, then $|\psi\rangle$ and $|\varphi\rangle$ are orthogonal, which does not hold for general quantum states.

□

1.2 Theoretical upper bound for quantum error-correcting code

From the quantum information capacity of a quantum channel, we can deduce the upper bound for a quantum error-correcting code. Due to time constraints, we use this as a lemma, and the reader can check the relevant literature for proofs.

$$\min\{1 - H_2(2t/3n), H_2(\frac{1}{2} + \sqrt{(1 - t/n)t/n})\}$$

1.3 Quantum error-correcting code from binary linear error-correcting code

Recall from classical linear error-correcting codes, we have a code (n, k, ℓ) with n qubits, k data qubits, and ℓ error-correcting qubits.

All the operations will be done in $\mathbb{F}_2 = \{0, 1\}$.

Consider two binary vectors $v = [v_1, \ldots, v_n], v_i \in \{0, 1\}$ and $w = [w_1, \ldots, w_n], w_i \in \{0, 1\}$ with size n.

Recall from our lecture that.

d denotes the Hamming weight of a vector.

$d_H(v, w) = \sum_{i=1}^{n} \begin{cases} 0 & \text{if } v_i = w_i \\ 1 & \text{if } v_i \neq w_i \end{cases}$ denotes the Hamming distance between v and w.

$\text{supp}(v) = \{i \in [n] : v_i \neq 0\}$ denotes the support of v.

$v|_S$ denotes the projection of v onto the subspace S, we usually denote the S by a set of coordinates, that is $S \subseteq [n]$.

When projecting a vector v onto a another vector w, we usually write $v|_E \coloneqq v|_{\text{supp}\, w}$.

When we have two vectors, we may use $v \leqslant w$ (Note that this is different from the $\leq$ sign) to mean $\text{supp}(v) \subseteq \text{supp}(w)$.

Example 2. *Let* $v = [1,0,0,1,1,1,1]$ *and* $w = [1,0,0,1,0,0,1]$*, then* $\text{supp}(v) = \{1,4,5,6,7\}$*,* $\text{supp}(w) = \{1,4,7\}$*. Therefore* $w \leqslant v$*.*

$v|_w = [v_1, v_4, v_7] = [1,1,0]$

$\mathcal{C}$ denotes the code, a set of arbitrary binary vectors with length n.

$d(\mathcal{C}) = \{d(v,w)|v,w \in \mathcal{C}\}$ denotes the minimum distance of the code.

If $\mathcal{C}$ is linear, then the minimum distance is the minimum Hamming weight of a non-zero codeword.

A $[n,k,d]$ linear code is a linear code of n bits codeword with k message bits that can correct d errors.

$R := \frac{\dim \mathcal{C}}{n}$ is the rate of code $\mathcal{C}$.

$\mathcal{C}^\perp := \{v \in \mathbb{F}_2^n : v \cdot w = 0 \text{ for all } w \in \mathcal{C}\}$ is the dual code of a code $\mathcal{C}$. From linear algebra, we know that $\dim \mathcal{C}^\perp + \dim \mathcal{C} = n$.

Example 3. *Consider the* $[7,4,3]$ *Hamming code with generator matrix* G*.*

$$G = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{bmatrix}$$

All the codewords for the $[7,4,3]$ *Hamming code are*

$$\mathcal{C} = \begin{matrix} 0000000 & 0001011 & 0010110 & 0011101 \\ 0100111 & 0101100 & 0110001 & 0111010 \\ 1000101 & 1001110 & 1010011 & 1011000 \\ 1100010 & 1101001 & 1110100 & 1111111 \end{matrix}$$

Common error defined in the quantum information theory involves bit flips and phase flips, introduced in Section 8.3.3 of [4].

Definition 8. *The bit-flip channel flips the state of a qubit from* $|0\rangle$ *to* $|1\rangle$ *with probability* $1-p$*, the operation element is defined as follows. Let* $|0\rangle, |1\rangle$ *be the basis of matrices for* $\mathbb{F}^{2\times 2}$*.*

$$E_0 = \sqrt{p}I = \sqrt{p}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad E_1 = \sqrt{1-p}X = \sqrt{1-p}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Definition 9. *The phase-flip channel flips the state of a qubit with probability* $1-p$*; the operation element is defined as follows. Let* $|0\rangle, |1\rangle$ *be the basis of matrices for* $\mathbb{F}^{2\times 2}$*.*

$$E_0 = \sqrt{p}I = \sqrt{p}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad E_1 = \sqrt{1-p}Z = \sqrt{1-p}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Generally, the error channel might be more general (Amplitude damping, Depolarizing channel) than that, but to simplify our discussion, we only consider bit-flip and phase-flip channels as our main source of noise.

4 Tools and related topics

Next, we will introduce some tools for quantum error correction.

4.1 Shor code

First, we will introduce a coding scheme analogous to repetition code in our class, but in quantum computing. The Shor code [5]

Proposition 10. *Assume that the decoherence process only affects one qubit of our superposition, while the other qubits remain unchanged.*

There exists a quantum coding scheme that encodes 8 to 9 bits that corrects 1 error. $[9, 1, 3]$ *code exists.*

Recover 1 qubit from a 9-qubit quantum system.

First we define the decoherence process introduced in [5]. That is, the basis bit $|e_0\rangle|1\rangle \mapsto |a_0\rangle|0\rangle|a_1\rangle|1\rangle$ and $|e_0\rangle|0\rangle \mapsto |a_2\rangle|0\rangle|a_3\rangle|1\rangle$. Where $|a_0\rangle, |a_1\rangle, |a_2\rangle, |a_3\rangle$ are states of the environment. We use $E(|\phi\rangle)$ to represent the decoherence process.

The encoding process for the Shor code goes as follows:

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

The decoding process is slightly different from the classical cases. Recall that in the classical cases, we can directly use the measurement to determine the error position. In quantum computing, we can make a measurement on the extra 8 bits to determine if $|\psi\rangle$ has an error or not, but it may not help us restore the correct bit $|\psi\rangle$ since measurement is not recoverable.

We need to restore the target state $|\psi\rangle$ using the information from the measurement.

Suppose, without loss of generality, the decoherence process affects encoded $|0\rangle$, then we have the following superposition to decode:

$$E(|0_L\rangle) = \frac{1}{\sqrt{2}}\left[(|a_0\rangle|0\rangle + |a_1\rangle|1\rangle)|00\rangle + (|a_2\rangle|0\rangle + |a_3\rangle|1\rangle)|11\rangle\right]$$

In terms of Bell basis $|\phi^+\rangle = \frac{1}{2\sqrt{2}}(|000\rangle + |111\rangle)$ and $|\phi^-\rangle = \frac{1}{2\sqrt{2}}(|000\rangle - |111\rangle)$, we have the following:

$$\begin{aligned} E(|0_L\rangle) = &\frac{1}{2\sqrt{2}}(|a_0\rangle + |a_3\rangle)(|000\rangle + |111\rangle) \\ &+ \frac{1}{2\sqrt{2}}(|a_0\rangle - |a_3\rangle)(|000\rangle - |111\rangle) \\ &+ \frac{1}{2\sqrt{2}}(|a_1\rangle + |a_2\rangle)(|100\rangle + |011\rangle) \\ &+ \frac{1}{2\sqrt{2}}(|a_1\rangle - |a_2\rangle)(|100\rangle - |011\rangle) \end{aligned}$$

and $|1\rangle$ will have

$$\begin{aligned} E(|1_L\rangle) = &\frac{1}{2\sqrt{2}}(|a_0\rangle + |a_3\rangle)(|000\rangle - |111\rangle) \\ &+ \frac{1}{2\sqrt{2}}(|a_0\rangle - |a_3\rangle)(|000\rangle + |111\rangle) \\ &+ \frac{1}{2\sqrt{2}}(|a_1\rangle + |a_2\rangle)(|100\rangle - |011\rangle) \\ &+ \frac{1}{2\sqrt{2}}(|a_1\rangle - |a_2\rangle)(|100\rangle + |011\rangle) \end{aligned}$$

Suppose $|\psi\rangle = \alpha\,|0_L\rangle + \beta\,|1_L\rangle$, then the error with the encoding can be written as a superposition of four terms:

$$|\psi\rangle\,, X_1\,|\psi\rangle\,, Z_1\,|\psi\rangle\,, X_1Z_1\,|\psi\rangle$$

Measuring the error syndrome collapse the superposition into a single term, (This has to do with the fact that these states are orthogonal therefore perfectly distinguishable, the detailed proof worth another section but the intuition is that projection for the orthogonal states, or called measurement, is either 0 or 1, corresponding to the probability of measuring the outcome with 1 or 0 if otherwise) which tells us the error operation E and how to recover them by inverse operation for I, X_1, Z_1, X_1Z_1.

The whole process can be done using the quantum circuits described below:

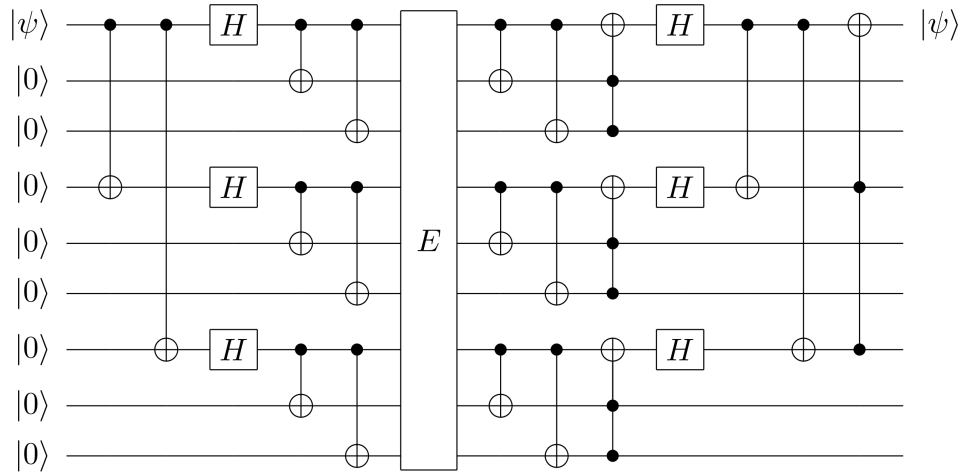

Figure 1: Encoding and decoding process for the Shor code using controlled-NOT gates and Hadamard gates.

4.2 CSS code (Steane code)

This is a quantum error-correction code actually introduced in our selection [3]. A special instance of that is Steane's code. It's basically the quantum version of the classical Hamming code.

Proposition 11. *Let $\mathcal{C}_1, \mathcal{C}_2$ are $[n, k_1]$, $[n, k_2]$ classic linear code such that $\mathcal{C}_2 \subset \mathcal{C}_1$ and $\mathcal{C}_1$ and $\mathcal{C}_2^{\perp}$ both correct t errors. The CSS code $[n, k_1 - k_2]$ quantum code correcting t qubits, namely the CSS code of $\mathcal{C}_1$ over $\mathcal{C}_2$ is a quantum error-correcting code by the following construction.*

Let $x \in \mathcal{C}_1$ be any codeword of $\mathcal{C}_1$. The encoding for $|x\rangle \to |x + \mathcal{C}_2\rangle$ is the following:

$$|x + \mathcal{C}_2\rangle = \frac{1}{\sqrt{|\mathcal{C}_2|}} \sum_{y \in \mathcal{C}_2} |x + y\rangle$$

The quantum code $CSS(\mathcal{C}_1, \mathcal{C}_2)$ is the span of the states $|x + \mathcal{C}_2\rangle$ for all $x \in \mathcal{C}_1$, which is a $[n, k_1 - k_2]$ quantum code.

Suppose x' is an element of $\mathcal{C}_1$ such that $x - x' \in \mathcal{C}_2$, by linearity of linear code, $|x + \mathcal{C}_2\rangle = |x' + \mathcal{C}_2\rangle$, thus the state $|x + \mathcal{C}_2\rangle$ depends only upon the coset of $\mathcal{C}_1/\mathcal{C}_2$.

If x, x' belongs to different coset of $\mathcal{C}_2$, then for no $y, y' \in \mathcal{C}_2$ such that $x + y = x' + y'$.

This shows that $|x + \mathcal{C}_2\rangle$ and $|x' + \mathcal{C}_2\rangle$ are orthogonal states.

We want to show that the quantum code $CSS(\mathcal{C}_1, \mathcal{C}_2)$ is a quantum error-correcting code that is capable of correcting t bit and phase flip errors. Therefore, decoding of the CSS code is valid and possible. The proof is adapted from [4] Section 10.4.2

Proof. Suppose the bit flip errors are described by an n bit vector e_1 with 1s where bit flips occurred, and 0s represent the bit remains the same. e_2 be the n bit vector where 1s where phase flips occurred and 0s represents the phase remains the same.

Let $|x + \mathcal{C}_2\rangle$ be the original state, then the corrupted state is:

$$\frac{1}{\sqrt{|\mathcal{C}_2|}} \sum_{y \in \mathcal{C}_2} (-1)^{(x+y)\cdot e_2} |x + y + e_1\rangle$$

The term $(-1)^{(x+y)\cdot e_2}$ generalized the phase error and $|x + y + e_1\rangle$ gives the bit flip errors.

First, we detect and remove bit flip errors.

We initialized the ancilla system by taking tensor products of $|x + y + e_1\rangle$ with the zero states, then we have $|x + y + e_1\rangle |0\rangle$. By applying reversible parity check matrix H_1 of code $\mathcal{C}_1$, $|x + y + e_1\rangle |0\rangle$ becomes $|x + y + e_1\rangle |H_1(x + y + e_1)\rangle = |x + y + e_1\rangle |He_1\rangle$.

Since $(x + y) \in \mathcal{C}_1$, the operation will map the state

$$\frac{1}{\sqrt{|\mathcal{C}_2|}} \sum_{y \in \mathcal{C}_2} (-1)^{(x+y)\cdot e_2} |x + y + e_1\rangle \to \frac{1}{\sqrt{|\mathcal{C}_2|}} \sum_{y \in \mathcal{C}_2} (-1)^{(x+y)\cdot e_2} |x + y + e_1\rangle |He_1\rangle$$

Then we take the measurement operator on the ancilla and decode the linear code $\mathcal{C}_1$ to obtain the e_1 since $\mathcal{C}_1$ can correct up to t errors.

$$\frac{1}{\sqrt{|\mathcal{C}_2|}} \sum_{y \in \mathcal{C}_2} (-1)^{(x+y)\cdot e_2} |x + y + e_1\rangle |He_1\rangle \to \frac{1}{\sqrt{|\mathcal{C}_2|}} \sum_{y \in \mathcal{C}_2} (-1)^{(x+y)\cdot e_2} |x + y + e_1\rangle$$

By applying the NOT gate on the error bits detected, we have

$$\frac{1}{\sqrt{|\mathcal{C}_2|}} \sum_{y \in \mathcal{C}_2} (-1)^{(x+y)\cdot e_2} |x + y + e_1\rangle \to \frac{1}{\sqrt{|\mathcal{C}_2|}} \sum_{y \in \mathcal{C}_2} (-1)^{(x+y)\cdot e_2} |x + y\rangle$$

Second, we detect and remove phase flip errors.

We first detect phase flip errors by applying Hadamard gates to each qubit.

$$\frac{1}{\sqrt{|\mathcal{C}_2|}} \sum_{y \in \mathcal{C}_2} (-1)^{(x+y)\cdot e_2} |x + y\rangle \to \frac{1}{\sqrt{|\mathcal{C}_2|2^n}} \sum_{z \in \mathbb{F}_2^n} \sum_{y \in \mathcal{C}_2} (-1)^{(x+y)\cdot(e_2+z)} |z\rangle$$

Let $z' = z + e_2$, then the state can be rewritten as

$$\frac{1}{\sqrt{|\mathcal{C}_2|2^n}} \sum_{z \in \mathbb{F}_2^n} \sum_{y \in \mathcal{C}_2} (-1)^{(x+y)\cdot(e_2+z)} |z\rangle \to \frac{1}{\sqrt{|\mathcal{C}_2|2^n}} \sum_{z' \in \mathbb{F}_2^n} \sum_{y \in \mathcal{C}_2} (-1)^{(x+y)\cdot z'} |z' + e_2\rangle$$

Suppose $z' \in \mathcal{C}_2^\perp$, then $\sum_{y \in \mathcal{C}_2} (-1)^{y \cdot z'} = |\mathcal{C}_2|$, and if $z' \notin \mathcal{C}_2^\perp$, then $\sum_{y \in \mathcal{C}_2} (-1)^{y \cdot z'} = 0$

Therefore, we can rewrite the state as

$$\frac{1}{\sqrt{|\mathcal{C}_2|2^n}} \sum_{z \in \mathbb{F}_2^n} \sum_{y \in \mathcal{C}_2} (-1)^{(x+y)\cdot z} |z' + e_2\rangle \to \frac{1}{\sqrt{|\mathcal{C}_2|2^n}} \sum_{z' \in \mathcal{C}_2^\perp} (-1)^{x \cdot z'} |z' + e_2\rangle$$

Note that the form is exactly the same as the bit flip cases. Therefore, we can correct it using the same procedure and obtain the following with error e_2.

$$\frac{1}{\sqrt{|\mathcal{C}_2|2^n}} \sum_{z' \in \mathcal{C}_2^\perp} (-1)^{x \cdot z'} |z' + e_2\rangle \to \frac{1}{\sqrt{|\mathcal{C}_2|2^n}} \sum_{z' \in \mathcal{C}_2^\perp} (-1)^{x \cdot z'} |z'\rangle$$

Apply the Hadamard gate to the bits to the states where $e_2 = 0$, since $\forall |\phi\rangle, H^2 |\phi\rangle = |\phi\rangle$.

$$\frac{1}{\sqrt{|\mathcal{C}_2|2^n}} \sum_{z' \in \mathcal{C}_2^\perp} (-1)^{x \cdot z'} |z'\rangle \to \frac{1}{\sqrt{|\mathcal{C}_2|}} \sum_{y \in \in \mathcal{C}_2} |x + y\rangle$$

Which restores the desired state.

□

Example 5. *Let $\mathcal{C}_1$ be the $[7, 4, 3]$ Hamming code and $\mathcal{C}_2$ be $\mathcal{C}_1^\perp$, then the Steane code is the CSS code of $\mathcal{C}_1$ over $\mathcal{C}_2$.*

This code contains $2^{n-k_1-k_2}$ orthogonal states, in this case, $k_1 = 4$ and $k_2 = 3$. Therefore, our quantum error correcting code will map $4 - 3 = 1$ qubits into 7 qubits.

$$\begin{aligned}
|0\rangle \to \frac{1}{2\sqrt{2}} \sum_{v \in H_\mathcal{C}} |v\rangle = \frac{1}{2\sqrt{2}} \big[& |0000000\rangle + |1010101\rangle + |0110011\rangle + |1100110\rangle \\
& + |0001111\rangle + |1011010\rangle + |0111100\rangle + |1101001\rangle \big] \\
|1\rangle \to \frac{1}{2\sqrt{2}} \sum_{v \in H_\mathcal{C}} |v + e\rangle = \frac{1}{2\sqrt{2}} \big[& |1111111\rangle + |0101010\rangle + |1001100\rangle + |0011001\rangle \\
& + |1110000\rangle + |0100101\rangle + |1000011\rangle + |0010110\rangle \big]
\end{aligned}$$

Where e is the all-1 vector.

6 Evaluation of paper

This paper provides a natural mapping of traditional linear codes to the quantum error correction code, which inspires future research and applications on more complicated quantum error correction algorithms like stabilizer codes and surface code.

7 Limitation and suggestions

These codes generally deal with the case where the decoherence process only affects one qubit of our superposition, while the other qubits remain unchanged.

However, generally, the error in quantum computing is continuous, and it is hard to distinguish the error via a discrete measurement.

An interesting topic is to evaluate the bounds of the correctable error beyond bit and phase flip for CSS code and show the limit of the method in terms of Bloch sphere transformation for the multi-qubit system.

8 Further direction and research

By the distinct properties of quantum computation and information theory derived from the setting, there are many interesting applications and ways to correct errors by using entanglement and measurement operations.

Some interesting topics are the Toric code and the surface code.

This method gives a $[2nm + n + m + 1, 1, min(n, m)]$ error correcting code that only needs local stabilizer checks and really interests me. I want to know more about this code and how it works, if time permits, and these questions will stay in my mind when I have a good chance to work on it.

References

[1] Sheldon Jay Axler. *Linear algebra done right*. Springer, 2024.

[2] Ingemar Bengtsson and Karol Zyczkowski. *Geometry of Quantum States: An Introduction to Quantum Entanglement*. Cambridge University Press, 2006.

[3] A. R. Calderbank and Peter W. Shor. "Good quantum error-correcting codes exist". In: *Physical Review A* 54.2 (Aug. 1996), pp. 1098–1105. ISSN: 1094-1622. DOI: 10.1103/physreva.54.1098. URL: http://dx.doi.org/10.1103/PhysRevA.54.1098.

[4] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum information Michael A. Nielsen & Isaak L. Chuang*. Cambridge Univ. Press, 2010.

[5] Peter W. Shor. "Scheme for reducing decoherence in quantum computer memory". In: *Phys. Rev. A* 52 (4 Oct. 1995), R2493–R2496. DOI: 10.1103/PhysRevA.52.R2493. URL: https://link.aps.org/doi/10.1103/PhysRevA.52.R2493.