AALBORG UNIVERSITY

DENMARK

Puncturing Entanglement-Assisted Stabilizers

Master Thesis

Mathematics-Technology

Author:
Nicolai P. B. Pedersen

Supervisors:
Jaron S. Gundersen
Jakob K. Søndergaard

May 26, 2026

Written with pdfLaTeX.

Department of Mathematical Sciences
Thomas Manns Vej 23
Aalborg University
http://www.math.aau.dk

AALBORG UNIVERSITY
STUDENT REPORT

Title:
Puncturing Entanglement-Assisted Stabilizers

Theme:
Quantum Error-Correction

Project Period:
Fall 2025 - Spring 2026

Project Group:
MT10-10

Author:
Nicolai P. B. Pedersen

Supervisors:
Jaron S. Gundersen
Jakob K. Søndergaard

Number of Pages:
106

Date of Completion:
May 26, 2026

Abstract:

Quantum computing has attracted significant attention since it was proposed in the 1980's. Several quantum algorithms are known to provide a great advantage in computational complexity compared to their classical counterparts. However, these algorithms require a large number of fault-tolerant qubits; an unfeasible constraint in today's noisy intermediate-scale quantum (NISQ) computers. Consequently, quantum error correction constitutes a fundamental field in the development of quantum computers. This thesis studies the well established stabilizers along with their generalized entanglement-assisted (ea) counterparts. The study culminates in a puncturing method for the receiver's qubits of an ea stabilizer based on a previously developed method for non-ea stabilizers. This method allows for a dynamic error-correction setup where the ea stabilizer is customized to the availability of entanglement-resources. The developed method transforms an ea stabilizer of parameters $[[N, k, d; c]]$ into an ea stabilizer of parameters $[[N, k, d'; c-1]]$. The method was tested on a $[[5, 1, 5; 4]]$ ea stabilizer, yielding a $[[5, 1, d'; 3]]$ ea stabilizer with $d' \in \{1, 2, 3, 4\}$ depending on the chosen puncturing scenario.

Department of Mathematical Sciences
Thomas Manns Vej 23
Aalborg University
http://www.math.aau.dk

AALBORG UNIVERSITY
STUDENT REPORT

Titel:
Punktering af Entanglement-Assisterede Stabilisatorer

Tema:
Kvante Fejlkorrektion

Projektperiode:
Efterår 2025 - Forår 2026

Projektgruppe:
MT10-10

Forfatter:
Nicolai P. B. Pedersen

Vejledere:
Jaron S. Gundersen
Jakob K. Søndergaard

Antal sider:
106

Afleveringsdato:
May 26, 2026

Abstract:

Kvanteberegning har tiltrukket sig betydelig opmærksomhed, siden det blev foreslået i 1980'erne. Adskillige kvantealgoritmer har vist en stor fordel i beregningsmæssig kompleksitet sammenlignet med deres klassiske modparter. Disse algoritmer kræver dog et stort antal fejltolerante qubits; en uopnåelig begrænsning i nutidens støjende mellemstore kvantecomputere. Derfor udgør kvantefejlkorrektion et fundamentalt felt i udviklingen af kvantecomputere. Denne afhandling studerer de veletablerede stabilisatorer sammen med deres generaliserede entanglement-assisterede (ea) modparter. Studiet kulminerer i en punkteringsmetode for modtagerens qubits i en ea stabilisator baseret på en tidligere udviklet metode til ikke-ea stabilisatorer. Denne metode muliggør en dynamisk fejlkorrektionsopsætning, hvor ea stabilisatoren tilpasses tilgængeligheden af entanglement-ressourcer. Den udviklede metode transformerer en ea stabilisator med parametrene $[[N, k, d; c]]$ til en ea-stabilisator med parametrene $[[N, k, d'; c-1]]$. Metoden blev testet på en $[[5, 1, 5; 4]]$ ea stabilisator, hvilket resulterede i en $[[5, 1, d'; 3]]$ ea stabilisator med $d' \in \{1, 2, 3, 4\}$ afhængigt af det valgte punkteringsscenarie.

Preface

This thesis was written by Nicolai Peder Bülow Pedersen in a project conducted during the period from September 2025 to May 2026. The thesis is a culmination of the Mathematics-Technology master's programme at Aalborg university, a two-year programme combining rigorous mathematics with engineering.

The software developed throughout the project is attached to the report and described in detail in appendix C.

The author wishes to thank the supervisors Jaron S. Grundersen and Jakob K. Søndergaard for their support and guidance during the project.

Aalborg University, May 26, 2026

Nomenclature

The study of quantum error-correcting codes involves the fields of physics, mathematics and engineering, each of which may apply different symbols and notations. The following nomenclature summarizes the conventions adopted throughout this thesis.

Symbol	**Description**
$\mathbb{R}$	The set of all real numbers
$\mathbb{C}$	The set of all complex numbers
$\mathbb{F}_2$	The Galois field of order 2
$\mathcal{A}$	Physical system
$\mathcal{A}_1 \cdots \mathcal{A}_N$	Composition of physical systems
$\mathcal{H}$	Hilbert space
$\mathcal{H}_\mathcal{A}$	State space of physical system $\mathcal{A}$
$\|\psi\rangle, \|\phi\rangle$	State vectors
$\langle\psi\|, \langle\phi\|$	Dual state vectors
$\langle\psi\|\phi\rangle$	Inner product of the state vectors $\|\psi\rangle$ and $\|\phi\rangle$
$\|\psi\rangle\langle\phi\|$	Outer product of the state vectors $\|\psi\rangle$ and $\|\phi\rangle$
$\otimes$	Tensor product
ρ, σ	Density operators
$F(\rho, \sigma)$	Fidelity between ρ and σ
X, Y, Z	Pauli matrices
I	Identity matrix
δ_{ij}	Kronecker delta
$H^\dagger$	Hermitian conjugate of H
α^*	Complex conjugate of $\alpha \in \mathbb{C}$

Symbol	**Description**
$\mathcal{E}, \mathcal{F}, \mathcal{R}$	Quantum channels
$\mathcal{P}_N$	The N-fold Pauli group.
$\langle g_1, \ldots, g_N \rangle$	Generating set for the group G.
$N_{\mathcal{P}_N}(G)$	Normalizer of subgroup G of $\mathcal{P}_N$ in $\mathcal{P}_N$
$C_{\mathcal{P}_N}(G)$	Centralizer of subgroup G of $\mathcal{P}_N$ in $\mathcal{P}_N$
$[N, k, d]$	Classical linear code encoding k bits in N with minimum distance d
$[[N, k, d]]$	Quantum stabilizer encoding k qubits in N with minimum distance d
$[[N, k, d; c]]$	Quantum entanglement-assisted stabilizer encoding k qubits in N with minimum distance d relying on c Bell pairs

Contents

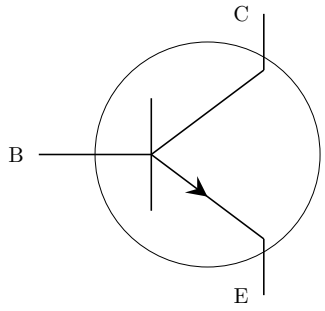
1 | Introduction

Merriam-Webster defines a computer as: “a programmable usually electronic device that can store, retrieve, and process data” [Merriam-Webster]. While this is a somewhat open definition, most people associate a computer with the laptop they use to perform their job duties or surf the internet. Although this is the most common computer encountered in our everyday life (along with our modern mobile phone, tablets, and smart screens), this is not the way people would interpret a computer in the past and perhaps this interpretation will also change in the future. In recent times, the concept of quantum computing has become more widespread; from a proposal by Richard Feynman in 1981 to working quantum computers in the present. The actual impact of these machines still depends largely on their evolution, however, to comment on their importance, it is beneficial to consider the classical computers as we know them; their origin, evolution, capabilities, and lack thereof [Preskill, 2021, pp.2-4].

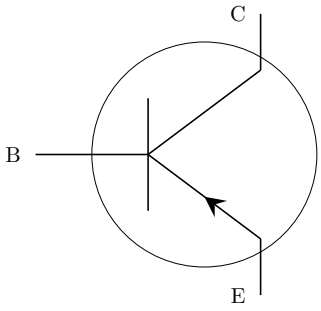
1.1 Classical Computers

The modern classical computer is a central part of everyday life. It allows people to maintain contact with their relatives and friends over vast distances in real time. It simplifies many professions, both by automating physically demanding positions and by compressing data, making information easier to store and share. The younger generation of today was born into this digital age, making it seem as if the computer is an indispensable part of the world, even though the digital computer as we know it is less than a century old.

Strictly speaking, the brain satisfies Merriam-Webster’s definition of a computer, making the first computer at least as old as humanity itself. However, restricting attention to man-made devices capable of performing operations based on human input, the picture is a bit more nuanced. Devices easing calculations of certain processes, e.g. the abacus, have been around for centuries (even millennia); whether or not these are considered computers depends on context. The digital computer has its origin in the 1940s. These computers store data in the binary domain; as strings of ones and zeroes. The first versions consisted of vacuum tubes to manipulate the data. In the 1940s the transistor was invented and about a decade later these replaced the vacuum tubes.



(a) NPN transistor.



(b) PNP transistor.

Figure 1.1: Circuit symbols for NPN and PNP transistors.

A transistor is a physical device with three legs that is used in electrical circuits. The three legs are labeled “collector”, ”emitter”, and ”base”. The current flowing through the collector-emitter circuit is controlled by the current applied to the base leg. Identifying no current flow with a 0 and current flow with a 1, this device effectively realized the bit system. Although not the only important specification of a pc, the number of transistors is vital for storing and retrieving data. Other important specifications are the number of operations the computer can do per second and how the memory (made up of transistors) is located (SSD, RAM, Cache etc.). For reference, the guidance system used during the moon landings in the 1960s and 1970s, which was only a couple decades after the birth of the digital computer, had 72 kb of memory, 4 kb of ram and performed at $14,245$ FLOPS while the computer used to write this report has 477 Gb of memory, 16 Gb of ram and performs at 0.55 TFLOPS. An average modern computer thus performs around $30,000,000$ times faster than the Apollo guidance system (AGS) and contains about $5,000,000$ times more memory. While these numbers may seem extreme, they are minuscule compared to the supercomputer “El Capitan” located in the United States. What is considered the world’s fastest supercomputer has a total of $45,613$ Tb of memory and performs at $2,792.9$ PFLOPS [Top500]. That is approximately 600 billion times more memory and $196,000$ billion times more operations per second compared to the AGS. Considering that the AGS were able to put mankind on the moon, this raises the question: Do problems exist for which even modern supercomputers lack computing power [O’Regan, 2026, pp.3-5], [Averill, 2022, p.1], [LLNL]?

Surprisingly, the answer is yes. To protect messages with confidential information, encryption can be used; this makes the message unreadable without the correct key. One such encryption scheme is the RSA. This scheme has been widely used, including in the older “NemID”. Informally, breaking the RSA scheme relies on finding two prime factors p and q of some number $N = pq$ (according to the fundamental theorem of number theory, this representation is unique, making the problem unambiguous [Hansen et al., 2013, p.60]). The problem is that this number, N, is represented using at least $1,024$ bits, making it a number of at least ≈ 300 decimals. Although the search only requires evaluation of the primes up to $\sqrt{N}$ in the “easiest” case, this potentially requires checking all prime numbers consisting of less than or equal to ≈ 150 decimals. Since there are approximately $\frac{\sqrt{N}}{\ln(\sqrt{N})}$ primes between 1 and $\sqrt{N}$, this amounts to $\approx 2.9 \cdot 10^{147}$ primes of less than or equal to

150 decimals. Even if the El Capitan supercomputer could check a prime in a single flop, it would still require 10^{129} seconds or $\approx 10^{121.5}$ years to check all primes of less than or equal to 150 decimals, making the task infeasible on a classical computer [Hansen et al., 2013, pp.64, 73-74].

Factoring large integers into their prime components may appear infeasible due to the immense size of the numbers involved. However, quantum computers offer fundamentally more efficient approaches to this problem. In particular, Shor's algorithm demonstrates that prime factorization can be performed in polynomial time on a quantum computer, compared to exponential/sub-exponential time on a classical computer [Shor, 1997]. More generally, quantum computation enables faster algorithms for certain classes of problems that are believed to be intractable on classical computers.

1.2 Quantum Computers

The quantum computer is fundamentally different from the classical computer. As mentioned, the classical computer operates in the binary domain; it stores, processes, and retrieves data as bits. The quantum computer, on the other hand, operates on quantum bits or qubits; briefly, a qubit is described by the state of a quantum system with a two-dimensional state space. The state space is some complex Hilbert space, and the state vector is a unit vector in this Hilbert space. Although a two-dimensional Hilbert space can be abstract, they are all unitarily isomorphic to $\mathbb{C}^2$, so for intuition and calculations, the latter is usually used. Defining the basis vectors of $\mathbb{C}^2$ by

$$|0\rangle := \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \text{and} \qquad |1\rangle := \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

A qubit is then described by a superposition of these two basis vectors

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle\,, \qquad |\alpha|^2 + |\beta|^2 = 1. \tag{1.1}$$

This may seem as though a qubit potentially stores infinite information; however, not all state vectors can be distinguished by measurement, greatly diminishing the potential information a qubit stores. A quantum state may equivalently be represented by a density operator; an operator on the state space. The strength and weaknesses of the two descriptions will be discussed later, for now it suffices to accept that both provides an equally valid description of qubits and how they are manipulated in a quantum computer. Associating 0 with $|0\rangle$ and 1 with $|1\rangle$, the bit arises as a special case of the qubit, making the classical computer a special case of the quantum computer. This may seem as though the classical computer is obsolete compared to its quantum counterpart. However, since the classical computer is more advanced in terms of development and practical deployment, deciding which to use is largely determined by the nature of the task at hand [Nielsen and Chuang, 2010, pp.13, 80, 98].

1.2.1 Quantum Supremacy and Quantum Algorithms

The era of quantum computers has come a long way from the proposal by Feynman in 1981. Today, multiple companies and states invest heavily in the development of quantum

hardware and software, and several quantum computers have been build; some with commercial access. A large breakthrough in the field was in 2019 when google claimed quantum supremacy: They had successfully completed a task on the Google "Sycamore" quantum processor featuring 54 qubits. The task was to sample from a quantum circuit. The classical computer would need to simulate the circuit, keeping track of the amplitudes (α and β of (1.1)) of each qubit as they pass through the circuit, whereas the quantum computer simply executes the circuit on qubits and measures the result. The quantum computer could obtain 3 million samples in about 600 seconds, something that was estimated to take 50 trillion cores hours; even with 10 million cores this amounts to more than 570 years [Arute et al., 2019, pp.506, 508-509].

Although very synthetic, the task completed by the Google quantum computer is part of a broader list of tasks made more efficient by quantum computers: The simulation of quantum systems. To understand why classical computers are inefficient, let a classical computer store α and β of (1.1) as k-bit complex numbers. For an n qubit quantum system, this would mean storing $k2^n$ bits of information, updating each as the qubits evolve in the system. A quantum computer, on the other hand, requires only the n qubits (depending on the circuit, error correction may introduce the need for extra qubits); the quantum system encodes its evolution directly in these qubits. Altogether, this means the runtime and memory evolve exponentially in n for the classical computer, while only polynomial in n for the quantum computer. Since

$$\lim_{n\to\infty} \frac{n^a}{b^n} = 0, \qquad b > 1, a \in \mathbb{R},$$

for large n, the runtime of a quantum computer is negligible compared to the classical runtime [Preskill, 2021, p.7].

Quantum computers excel at not only simulating quantum systems. Algorithms for numerous other tasks have been proposed. Although some may be demonstrative tasks, such as the Deutsch-Jozsa algorithm able to determine whether a function $f : \mathbb{F}_2^N \to \mathbb{F}_2$ is constant or balanced (preimage of 0 and 1 is of equal cardinality). On a classical computer, this could potentially require evaluating the function for $2^{N-1} + 1$ different inputs; a quantum computer can do it in a single function evaluation. Others solve a very requested task, such as Shor's algorithm which provides an efficient solution to the prime factorization problem. However, the practical implementation of Shor's algorithm is years ahead of our time. Recent studies show 20 million qubits are necessary to factor a $2,048$ bit number in 8 hours. For reference, the biggest quantum computer to date is the IBM Condor having 1,121 qubits [Gidney and Ekerå, 2021], [Bergou et al., 2021, p.117-119].

1.2.2 Physical Implementation and Quantum Gates

Classical computers use transistors to represent bits that are then manipulated through different gate setups called logic gates. Quantum computers, instead, operate on qubits and manipulate them using quantum gates. These qubits can be represented in different ways. For these implementations to be compared, the notion of decoherence and fidelity are introduced.
The quantum computer manipulates qubits, transforming their state. The state of a

qubit, given by (1.1), is prone to noise and will eventually change even without any active intervention; the time it takes is called the decoherence time. Depending on the qubit, this decoherence time can range from fractions of nanoseconds to seconds or even years. A longer decoherence time is generally favorable, as this imposes less constraints on the quantum gate time; the time it takes for a qubit to be manipulated by a quantum gate. Denoting the decoherence time by T_D and the initialization time of the qubit by T_0, the interval $[T_0, T_0 + T_D]$ is referred to as the coherence time window of the qubit [Nielsen and Chuang, 2010, pp.278-279].
The fidelity of a quantum gate is an expression of how well the output of the gate matches the intended or ideal output of the gate. If the gate introduces error, the actual output may deviate from the ideal output. More rigorously, the fidelity of two quantum states given by density operators ρ and σ is defined as

$$F(\rho, \sigma) := \mathrm{tr}\left(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right).$$

Using Uhlmann's theorem, it is evident that $0 \leq F(\rho, \sigma) \leq 1$ with $F(\rho, \sigma) = 1$ if and only if $\rho = \sigma$. Thus, the closer the gate fidelity is to 1, the better [Nielsen and Chuang, 2010, pp.409-411].

The 5 most common quantum computer implementations are summarized below along with the advantages, disadvantages, and companies involved in the realization and advancement [SpinQ].

1. **Superconducting qubits**: These are among the most widespread quantum computers. They rely on circuits operating at cryogenic temperature, making the resistance of the circuit negligible, hence the name superconducting. This type of quantum computer have high gate speed and fidelity, and they are suitable for scalability. However, they require very low temperatures and the coherence time is short compared to other quantum computers. The Google Sycamore and the largest quantum computer to date, the IBM Condor with 1,121 qubits, are of this type.

2. **Trapped ions**: These types of quantum computers manipulate ions (electrically charged atoms) in vacuum using electromagnetic fields. These types of computers have long coherence time and high fidelity gates. However, the gate speeds are lower and such computers are not easily scalable and require vacuum in the implementation. IonQ and Honeywell focus on this type of quantum computer.

3. **Quantum Dots**: A quantum dot computer uses confined electrons as qubits. The electrons are confined using semiconductors. The spin of the electrons are then manipulated using magnetics fields. Quantum dots have high gate fidelity and speed. They are scalable and relies on existing advanced semiconductor technology. However, they are sensitive to noise and requires cryogenic temperature. Intel and Microsoft are contributors to the advancement of this type of quantum computer.

4. **Photons**: This type uses particles of light as qubits. These qubits are then manipulated using beam splitters, mirrors, etc. These computers have extremely fast gates, they require no use of vacuum or cryogenic temperature, and the qubits are not prone to environmental noise. However, the qubits are generally more difficult

to make interact and the gates have lower fidelity. Furthermore, while the lower environmental noise makes them optimal for transmission over distances, the photons can be lost. They are not easily scalable, since not losing qubits and/or making them interact as the number increases remains an issue. Companies involved in this type of quantum computer includes PsiQuantum and Xanadu.

5. **Neutral Atom**: This type traps atoms and manipulates them using lasers. The qubits are then the state of the atom, e.g. the energy level of an electron. Such computers have a long coherence time, high gate fidelity, and atoms can easily be packed together, making the scalability promising. However, measurements can be difficult to perform in larger scale, and the gate speeds are not as fast as in other types of quantum computers. Aliro Quantum and researchers at Harvard are involved in this type of quantum computer.

Similarly to classical computers, quantum computers, regardless of the physical choice of implementation, manipulate qubits, and thus quantum information, through the use of quantum gates. Considering a single bit, the only two operations that can be performed on such a bit is the identity operation, leaving the bit unchanged, and the bit flip or NOT gate changing the bit from 0 to 1 or vice versa. In the quantum case, the picture is a bit more nuanced. For reasons later elaborated upon, quantum systems evolve linearly and unitarily. This means that for a single qubit, the operation $|\phi\rangle = A\,|\psi\rangle$ is a valid quantum gate if and only if $A \in \mathbb{C}^{2\times 2}$ is unitary. Although only two classical gates exist on a single bit, infinitely many quantum gates exist on a single qubit. In general, any unitary matrix $U \in \mathbb{C}^{2\times 2}$ can be decomposed as

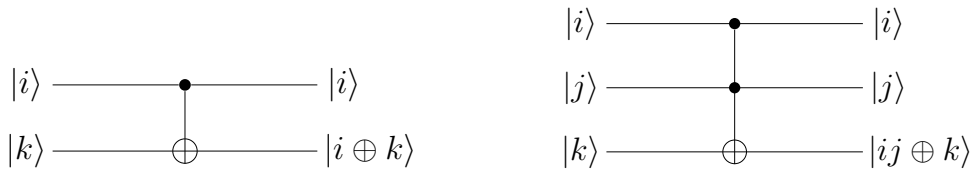
$$U = e^{i\alpha} \begin{bmatrix} e^{-i\beta} & 0 \\ 0 & e^{i\beta} \end{bmatrix} \begin{bmatrix} \cos(\delta) & -\sin(\delta) \\ \sin(\delta) & \cos(\delta) \end{bmatrix} \begin{bmatrix} e^{-i\gamma} & 0 \\ 0 & e^{i\gamma} \end{bmatrix}, \tag{1.2}$$

$$= \begin{bmatrix} \cos(\delta)e^{i(\alpha-\beta-\gamma)} & -\sin(\delta)e^{i(\alpha-\beta+\gamma)} \\ \sin(\delta)e^{i(\alpha+\beta-\gamma)} & \cos(\delta)e^{i(\alpha+\beta+\gamma)}, \end{bmatrix} \tag{1.3}$$

for some $\alpha, \beta, \delta, \gamma \in [0; 2\pi]$. Thus, any quantum gate can be described by some $x \in [0; 2\pi] \times [0; 2\pi] \times [0; 2\pi] \times [0; 2\pi]$. Some commonly encountered are the three Pauli matrices X, Y, and Z and the Hadamard gate, H, defined by

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

In the multi-qubit case, the state space is composed by tensor products. This means that any operator on the multi-qubit system is a superposition of tensor products. Since quantum systems evolve unitarily, the composed operator must be unitary. Commonly encountered multi-qubit gates are the CNOT and Toffoli (CCNOT) gates. Their circuit symbols and actions are shown in figure 1.2. Since the gates are linear, it suffices to know their action on some basis of the vector space they operate on.



(a) The CNOT gate. (b) The Toffoli gate.

Figure 1.2: Circuit symbols for the CNOT and Toffoli gate. Here, $i, j, k \in \mathbb{F}_2$.

A remarkable fact in the quantum realm, opposed to the classical one, is that any quantum gate is described by some unitary matrix U, and thus any quantum gate is invertible. This means that no information is lost when a quantum system evolves. A simple circuit composed of the Hadamard gate and the CNOT gate is shown below. This circuit takes as input $|00\rangle$ and outputs $|\Phi^+\rangle := \frac{1}{2}(|00\rangle + |11\rangle)$ known as one of the four Bell states. This bell state is an example of two entangled qubits. Entanglement is a powerful tool in quantum mechanics described more extensively later.

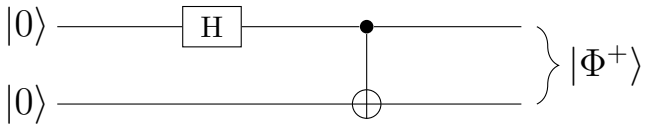


Figure 1.3: Example of a quantum circuit.

Consider a quantum system on N qubits. Such a quantum system evolves according to some unitary $U \in \mathbb{C}^{2^N \times 2^N}$, while infinite such matrices exist, they can all be realized by the composition of a finite set of unitary matrices. One such set is $\{\text{CNOT}, H, S, T\}$ where

$$T = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{bmatrix} \qquad S = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}.$$

The gates in the set $\{\text{CNOT}, H, S, T\}$ are called universal gates. Figure 1.4 illustrates the Toffoli gate realized by the universal gates.

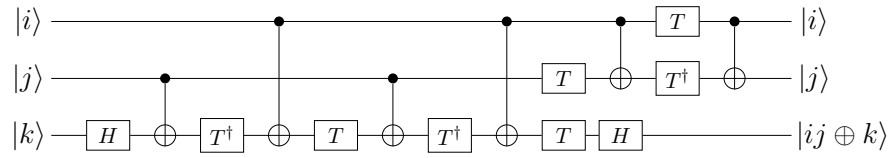


Figure 1.4: The Toffoli gate realized by the CNOT gate, the Hadamard gate and the T gate.

An inherent problem with quantum gates, almost not existent in the classical counterpart, is the presence of noise. Classical gates may achieve error rates lower than $10^{-10}\%$, less than a single error in a trillion uses. Quantum gates, on the other hand, experience error

rates higher by several magnitudes. Although the error rate depends on the choice of quantum computer, for single qubit gates, it typically ranges from 1% down to 0.01%, that is, an error in every $100 - 10,000$ uses. For two-qubit gates, the error rate is typically higher, but can be as low as a single error for every $2,000$ uses. A way of achieving a lower error rate is by correcting some of the errors that may occur. Classical computers do this using techniques from the classical error-correction field. This theory has been extensively studied since its origin with Claude Shannon in the middle of the 20th century. For obvious reasons, the quantum error-correction field has not yet been studied as extensively; however, powerful tools have been constructed [Spectrum], [Science], [Nielsen and Chuang, 2010, pp.xxx-xxxi, 17-22].

1.2.3 Quantum Error Correction

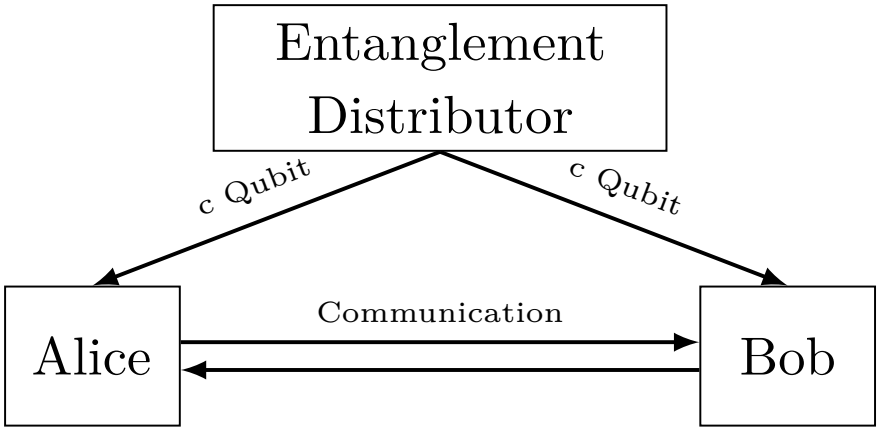
The field of quantum error correction concerns the study of correcting errors that may occur to a qubit being; transformed by a gate, transmitted between quantum computers, or even undergoing decoherence while sitting idle in memory. Such errors may be mitigated by encoding a number of qubits into a larger number of qubits, thus adding redundant information that can be used to correct errors. The errors a coding scheme can correct, and thus the significance of the code, depends on the number of redundant qubits added and the assumption on the errors.

An important class of quantum error-correcting codes is the stabilizer codes. This class of codes will be defined in more depth later; for now, it suffices to know that they can be defined in terms of a check matrix. The check matrix of a stabilizer code, which encodes k qubits into N qubits for some $N, k \in \mathbb{N}$ where $k \leq N$, is a matrix $C \in \mathbb{F}_2^{N-k\times 2N}$. For the stabilizer to have a non-trivial coding space, any two rows of the check matrix, $[\overline{a}_i \ \overline{b}_i], [\overline{a}_j \ \overline{b}_j] \in \mathbb{F}_2^{2N}$ where $\overline{a}_i, \overline{a}_j, \overline{b}_i, \overline{b}_j \in \mathbb{F}_2^N$ and $i, j \in \{1, \ldots, N-k\}$, must be symplectic orthogonal with respect to the symplectic form

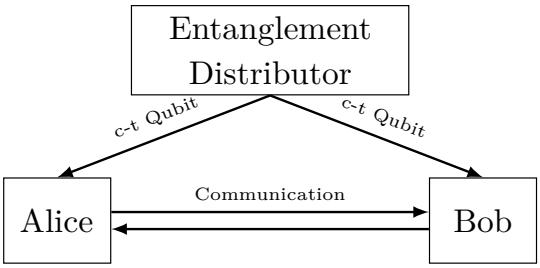
$$([\overline{a}_i \ \overline{b}_i], [\overline{a}_j \ \overline{b}_j])_S := \left\langle \overline{a}_i \middle| \overline{b}_j \right\rangle + \left\langle \overline{b}_i \middle| \overline{a}_j \right\rangle .$$

This requirement greatly limits the number of matrices $A \in \mathbb{F}_2^{N-k\times 2N}$ defining a valid stabilizer. However, this requirement can be overcome by using the most powerful resource in the quantum toolbox; entanglement. Given any matrix $A \in \mathbb{F}_2^{N-k\times 2N}$, the rows can be suitably extended to satisfy the self-orthogonality constraint. This extension is realized by distributing entangled qubits between the sender and the receiver. The stabilizer codes realized in this way are part of the class of entanglement-assisted quantum error-correcting codes (EAQECCs). Based on this idea, any classical linear code can be the foundation for a stabilizer code. A classical linear code with parameters $[N, k, d]$ is characterized by messages, $u \in \mathbb{F}_2^k$ encoded by a generator matrix $G \in \mathbb{F}_2^{k\times N}$ such that the resulting encoded message is given by $c = uG \in \mathbb{F}_2^N$. The coding space is the k dimensional subspace given by the row space of the generator matrix. This also implies that the rows are required to be linearly independent. The distance d of such codes is the minimum number of non-zero entries in any non-zero codeword. Associated with such codes is a parity check matrix defined to be the matrix $H \in \mathbb{F}_2^{N-k\times N}$ such that $GH^T = 0 \in \mathbb{F}_2^{k\times N-k}$. The rows of the check matrix are also required to be linearly independent, and its row space is exactly the subspace orthogonal to the coding space. For classical linear code with parameters

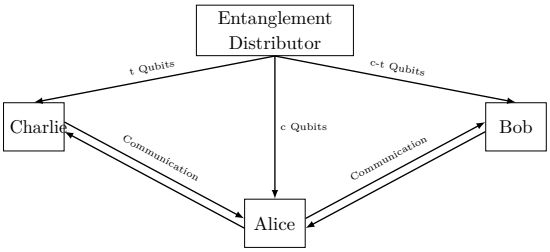
$[2N, N+k, d]$ the parity check matrix defines a valid stabilizer if the rows are symplectic orthogonal, if not, it can be paired with entangled qubit pairs, such that the unity defines a valid stabilizer. The number of entangled qubit pairs can be rigorously found depending on the parity check matrix chosen. However, introducing entanglement in the code introduces a problem; the new resource is not free and needs to be continuously supplied when using the code. As already mentioned, entangled qubits can be realized by the Hadamard gate and the CNOT gate, both of which are faulty, especially the CNOT gate. Although this problem may be overcome by a simple overproduction of entangled qubits, this may not be feasible in practice. Denote by k the number of logical qubits to be encoded, c the number of entangled qubit pairs that the code requires, and s the number of ancilla qubits needed for encoding. Then $N-k = s+c$ and the total number of qubits needed for the code is $2c+s+k$. Thus, if the code is to be used every T seconds, it is necessary to have c entangled qubit pairs ready each T seconds. Depending on T and the hardware used to create the entangled qubit pairs, creating and distributing c entangled qubit pairs each T seconds may be difficult or even impossible. If instead only $c-t$ entangled qubit pairs could be distributed, for some $1 \leq t \leq c-1$, the stabilizer code previously defined that relies on c entangled qubit pairs is useless; a new stabilizer that requires only $c-t$ entangled qubit pairs is needed. Figure 1.5 show two setups in which this problem could arise. In the first setup, less entangled qubit pairs are distributed simply due to hardware constraint; the entanglement resource cannot be distributed in the given time frame. In the second setup, a new network node appears. The entanglement resource available thus has to be split between the nodes (assuming the hardware prohibits availability of additional resources), leading to less available resource between each node-pair. This report proposes a simple solution to the limited entanglement resource described based on the notion of puncturing [Justesen and Høholdt, 2004, pp.3-5].



(a) The original setup.



(b) Less entangled qubits available due to hardware constraints.



(c) Available resource split between multiple nodes.

Figure 1.5: The original setup and the two additional setups leaving the two original nodes with less entangled qubit pairs available.

Puncturing has been a known procedure on classical codes for decades. Briefly, the punctured code of a $[N, k, d]$ classical linear code is obtained by removing a column from the generator matrix. Assuming that the rows remain linear independent after removing this column, the resulting code has parameters $[N-1, k, d']$ where $d-1 \leq d' \leq d$. This procedure increases the rate of the code since $\frac{k}{N} < \frac{k}{N-1}$. Since $d' \leq d$ it comes at the cost of being less efficient in error correction, as any code can correct the error e if and only if $t < \frac{d}{2}$ where t is the number of non-zero entries in e. This general idea can be applied with care to stabilizer codes. A stabilizer code that encodes k logical qubits into N physical qubits is called a $[[N, k, d]]$ stabilizer. The distance d requires additional theory to define, but it describes the error-correcting abilities of the code analogues to the classical distance. A recent study has found a method to puncture such stabilizers that produces a new stabilizer of parameters $[[N-1, k, d']]$. Consider an entanglement-assisted stabilizer of parameters $[[N, k, d; c]]$ where c is the number of entangled qubits needed. This report investigates whether the previously defined puncturing of a stabilizer can be used to puncture the code $[[N, k, d; c]]$ to a new code $[[N', k', d'; c-1]]$. This will solve the problem of available entanglement resources as no new stabilizer need to be defined, one simply puncture the already defined stabilizer t times, leading to a stabilizer only requiring $c-t$ entangled qubit pairs [Gundersen et al., 2025], [Justesen and Høholdt, 2004, p.15].

The problem is collected in the following problem formulation:

Can an algorithm, capable of puncturing the number of entangled qubit pairs required for an entanglement-assisted stabilizer, be constructed?

If such a method could be developed, the transformation of the parameter c under the puncturing is required to decrease by 1. However, no constraint is placed on the transformation of the other parameters, N, k, and d. To this end, it is required that the method maintain k and provide an upper bound on the decrease of d. In the case of N, the only requirement is that it does not increase to ensure that the communication rate does not decrease, and it transforms in a predictable manner.

2 | Quantum Mechanics

Unlike the classical bit belonging to $\mathbb{F}_2$, a qubit is a state vector of a quantum system, therefore belonging to $\mathbb{C}^2$ as will be postulated shortly. This allows for the potential storage of much more information; the problem is accessing it. However, before diving into quantum information theory, it is first important to understand the mathematical foundation of quantum systems; learning which actions are allowed and which are not in such systems. This chapter is primarily based on [Nielsen and Chuang, 2010].

Throughout the report, the first and second basis vector of $\mathbb{C}^2$ will be denoted as

$$|0\rangle := \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad |1\rangle := \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Thus, any vector $|\psi\rangle \in \mathbb{C}^2$ can be expressed by $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$ for some $\alpha, \beta \in \mathbb{C}$. It will be postulated that a time evolution of a quantum system is described by linear transformations, three important of such are given by the Pauli matrices defined below.

$$X := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad Z := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad Y := \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}.$$

The Pauli matrices acting on the basis vectors reveals their importance;

$$X\,|0\rangle = |1\rangle\,,\ X\,|1\rangle = |0\rangle\,, \qquad Z\,|0\rangle = |0\rangle\,,\ Z\,|1\rangle = -\,|1\rangle\,,, \qquad Y\,|0\rangle = i\,|1\rangle\,,\ Y\,|1\rangle = -i\,|0\rangle\,.$$

The X Pauli matrix corresponds to a qubit flip, changing $|0\rangle$ to $|1\rangle$ and vice versa. The Z changes the phase of the second basis vector, but leaves the first unchanged. The Y Pauli matrix corresponds to a qubit and a phase flip, thus combining both the X and the Z Pauli matrices [Nielsen and Chuang, 2010, p.65].

2.1 The Postulates of Quantum Mechanics

The foundation of quantum mechanics lies on four postulates describing what a quantum system is, how it evolves in time, how to measure it, and how composite quantum systems operate. First, the description of what a quantum system is.

Postulate 2.1
The state space of any isolated quantum system, $\mathcal{A}$, is a complex Hilbert space, $\mathcal{H}_{\mathcal{A}}$. The system is completely described by its state vector, $|\psi\rangle \in \mathcal{H}_{\mathcal{A}}$.

[Nielsen and Chuang, 2010, p.80]

The state vector mentioned in the first postulate is a unit vector in the state space, that is, $\langle\psi|\psi\rangle = 1$. The second postulate concerns time evolutions of quantum systems.

Postulate 2.2
Given a state vector, $|\psi\rangle$, of a closed quantum system, $\mathcal{A}$, its time evolution is described by the Schrödinger equation:

$$i\hbar\frac{d\,|\psi\rangle}{dt} = H\,|\psi\rangle\,.$$

[Nielsen and Chuang, 2010, p.82]

The quantity $\hbar$ of postulate 2.2 is the reduced Planck's constant given by

$$\hbar = \frac{h}{2\pi} = 1.054571817 \times 10^{34} J \cdot s,$$

where h is Planck's constant.
The matrix H of postulate 2.2 is the Hamiltonian of the quantum system. It a fixed Hermitian operator, that is, $H^\dagger = H$, where $H^\dagger := (H^*)^T$ is the hermitian conjugation corresponding to the conjugate-transpose of a matrix.
Let $\hat{H} := (i\hbar)^{-1}H$, the differential equation of postulate 2.2 simplifies to

$$\frac{d\,|\psi\rangle}{dt} = \hat{H}\,|\psi\rangle\,.$$

The solution of such a differential equation is given by

$$\begin{aligned}|\psi(t_2)\rangle &= e^{(t_2-t_1)\hat{H}}\,|\psi(t_1)\rangle\,,\\ &= U(t_1,t_2)\,|\psi(t_1)\rangle\,,\end{aligned}$$

where $U(t_1,t_2) := e^{(t_2-t_1)\hat{H}}$. The operator U is unitary, meaning $U^\dagger U = I$. To see this, first realize $\hat{H}$ is anti-hermitian since

$$\hat{H}^\dagger = ((i\hbar)^{-1})^* H^\dagger = (-i\hbar)^{-1}H = -\hat{H}.$$

Thus,

$$(U(t_1,t_2))^\dagger U(t_1,t_2) = \left(e^{(t_2-t_1)\hat{H}}\right)^\dagger e^{(t_2-t_1)\hat{H}} = e^{(t_2-t_1)\hat{H}^\dagger} e^{(t_2-t_1)\hat{H}} = I.$$

With the introduction of the unitary transformation U a new version of postulate 2.2 appears:

The time evolution of a closed quantum system, $\mathcal{A}$, is described by a unitary transformation $U(t_1, t_2)$ relating the state vector at time t_1, to the state vector at time t_2 by

$$|\psi(t_2)\rangle = U(t_1, t_2)\,|\psi(t_1)\rangle\,.$$

The Pauli matrices presented in the beginning of the chapter are easily verified to be unitary (and Hermitian), and thus describes a valid time evolution of a quantum system [Nielsen and Chuang, 2010, pp.81-83].

Postulate 2.3
Quantum measurements are described by a set of measurement operators $\{M_j\}_{j=1}^N$ on the state space, $\mathcal{H}_\mathcal{A}$, of the quantum system, $\mathcal{A}$, being measured. Let the quantum system, $\mathcal{A}$, have state vector $|\psi\rangle$ before a measurement, then the probability that the measurement yields outcome j is given by Born's rule

$$p_j = \langle\psi|\, M_j^\dagger M_j\,|\psi\rangle\,. \tag{2.1}$$

The state vector after the measurement is,

$$|\psi\rangle_j^{\text{post}} := \frac{M_j\,|\psi\rangle}{\sqrt{\langle\psi|\, M_j^\dagger M_j\,|\psi\rangle}}. \tag{2.2}$$

[Nielsen and Chuang, 2010, p.84]

By postulate 2.3, it must be the case that

$$1 = \sum_{j=1}^N p_j = \sum_{j=1}^N \langle\psi|\, M_j^\dagger M_j\,|\psi\rangle\,.$$

This implies

$$\sum_{j=1}^N M_j^\dagger M_j = I. \tag{2.3}$$

This is called a completeness equation and the measurement set $\{M_j\}_{j=1}^N$ is said to satisfy the completeness relation. The operators $E_j = M_j^\dagger M_j$ are called Positive Operator-Valued Measures (POVM). Considering (2.1), the POVMs, $\{E_j\}_{j=1}^N$, describes the probabilities of each measurement outcome.
An important case of the above is when the measurement operators are orthogonal projectors, that is, when $M_j^\dagger = M_j$ and $M_j M_i = \delta_{ji} M_j$. Denote by $\{P_j := M_j\}_{j=1}^N$ a set of such orthogonal projective measurements. Let $\{a_j\}_{j=1}^N$ denote a set of real numbers, the operator given by

$$O = \sum_{j=1}^N a_j P_j$$

is then called an observable. The possible outcomes of the projective measurement involving the observable O are its eigenvalues $\{a_j\}_{j=1}^N$ and the probability of measuring the eigenvalue a_j is given by Born's rule from postulate 2.3

$$p_j = \langle\psi| P_j |\psi\rangle .$$

If the state before the measurement was $|\psi\rangle$, the state immediately after the measurement yielding outcome a_j is given by

$$|\psi\rangle_j^{\text{post}} = \frac{P_j |\psi\rangle}{\sqrt{\langle\psi| P_j |\psi\rangle}}.$$

Measuring directly after using the same projections always yields the same outcome as the previous measurement, since

$$p_i = \frac{\langle\psi| P_j P_i P_j |\psi\rangle}{\langle\psi| P_j |\psi\rangle} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{else} \end{cases}$$

and the post-measurement state is also unaffected;

$$(|\psi\rangle_j^{\text{post}})_j^{\text{post}} = \frac{P_j |\psi\rangle_j^{\text{post}}}{\sqrt{\langle\psi|_j^{\text{post}} P_j |\psi\rangle_j^{\text{post}}}} = \frac{P_j \frac{P_j|\psi\rangle}{\sqrt{\langle\psi|P_j|\psi\rangle}}}{\sqrt{\frac{\langle\psi|P_j}{\sqrt{\langle\psi|P_j|\psi\rangle}} P_j \frac{P_j|\psi\rangle}{\sqrt{\langle\psi|P_j|\psi\rangle}}}} = \frac{P_j |\psi\rangle}{\sqrt{\langle\psi| P_j |\psi\rangle}} = |\psi\rangle_j^{\text{post}} .$$

With the introduction of measurements in postulate 2.3 comes also a constraint of quantum mechanics. The fact that two non-orthogonal states cannot be reliably distinguished. Consider first the case of N state vectors $\{|\psi_i\rangle\}_{i=1}^N$, and let the set be orthogonal. Let the quantum system be in state $k \in \{1, \ldots, N\}$ before the measurement. To determine k, define measurement operators by $M_i := |\psi_i\rangle \langle\psi_i|$ for $i \in \{1, \ldots, N\}$ and $M_0 = I - \sum_{i=1}^N |\psi_i\rangle \langle\psi_i| = I - \sum_{i=1}^N M_i$. These measurement operators satisfy the completeness relation as

$$\begin{aligned} \sum_{i=0}^{N} M_i^\dagger M_i &= M_0 M_0 + \sum_{i=1}^{N} |\psi_i\rangle \langle\psi_i|\psi_i\rangle \langle\psi_i| \\ &= I - 2\sum_{i=1}^{N} M_i + \sum_{i=1}^{N}\sum_{j=1}^{N} M_i M_j + \sum_{i=1}^{N} M_i = I, \end{aligned}$$

where it has been used that $M_i^\dagger = M_i = M_i M_i$ and $M_i M_j = 0$ for any $i, j \in \{0, 1, \ldots, N\}$ with $i \neq j$. By postulate 2.3 it then follows

$$p_i = \langle\psi_k| M_i^\dagger M_i |\psi_k\rangle = \langle\psi_k| M_i |\psi_k\rangle = \langle\psi_k|\psi_i\rangle \langle\psi_i|\psi_k\rangle = \begin{cases} 1, & \text{if } i = k, \\ 0, & \text{else,} \end{cases}$$

for $i \in \{1, \ldots, N\}$ and

$$p_0 = \langle\psi_k| M_0^\dagger M_0 |\psi_k\rangle = \langle\psi_k| M_0 |\psi_k\rangle = \langle\psi_k| \left(|\psi_k\rangle - \sum_{i=1}^{N} |\psi_i\rangle \langle\psi_i|\psi_k\rangle \right) = 0.$$

The measurement thus yields outcome k with probability 1 and the state after the measurement, according to (2.2), is

$$|\psi_k\rangle_k^{\text{post}} = \frac{|\psi_k\rangle \langle\psi_k|\psi_k\rangle}{\sqrt{\langle\psi_k| M_k^\dagger M_k |\psi_k\rangle}} = |\psi_k\rangle .$$

The state before the measurement can then be determined reliably by the measurement. The following theorem concerns the case for which the set of state vectors is not assumed to be orthogonal [Nielsen and Chuang, 2010, pp.85-88].

Theorem 2.4: Distinguishing non-orthogonal states
Let $\{|\psi_i\rangle\}_{i=1}^N$ be a set of non-orthogonal state vectors. Let a quantum system, $\mathcal{A}$, be in state $|\psi_k\rangle$ for some $k \in \{1, \ldots, N\}$. No set of measurement operators exists such the state $|\psi_k\rangle$ can be reliably determined.

[Nielsen and Chuang, 2010, p.87]

Proof
Let $\{|\psi_i\rangle\}_{i=1}^{N_1}$ be a set of state vectors and without loss of generality assume $\langle\psi_1|\psi_2\rangle \neq 0$. Let $k = 1 \vee k = 2$, meaning the quantum system is in either state $|\psi_1\rangle$ or state $|\psi_2\rangle$. Let further $\{M_j\}_{j=1}^{N_2}$ be a set of measurements and S_i be the set of measurements for which the outcome reveals the system was in state $|\psi_i\rangle$, i.e. $S_i := \{j \mid f(|\psi_i\rangle_j^{\text{post}}) = |\psi_i\rangle\}$ where f are some determination rule. To reliably determine the state of the system it would require

$$1 = \sum_{j \in S_k} p_j = \sum_{j \in S_k} \langle\psi_k| M_j^\dagger M_j |\psi_k\rangle = \langle\psi_k| E_{S_k} |\psi_k\rangle , \qquad E_{S_i} := \sum_{j \in S_i} M_j^\dagger M_j,$$

that is,

$$\langle\psi_1| E_{S_1} |\psi_1\rangle = 1 \wedge \langle\psi_2| E_{S_2} |\psi_2\rangle = 1 \tag{2.4}$$

Assume (2.4) holds. According to (2.3)

$$I = \sum_{j=1}^{N_2} M_j^\dagger M_j = \sum_{i=1}^{N_1} E_{S_i} \implies \sum_{i=1}^{N_1} \langle\psi_k| E_{S_i} |\psi_k\rangle = 1,$$

which implies $\langle\psi_1| E_{S_2} |\psi_1\rangle = 0 \wedge \langle\psi_2| E_{S_1} |\psi_2\rangle = 0$ due to (2.4). By the definition of E_{S_2} it is evident that it is hermitian and positive semidefinite, it then follows it has a unique hermitian square root, hence

$$0 = \langle\psi_1| E_{S_2} |\psi_1\rangle = \langle\psi_1| \left(\sqrt{E_{S_2}}\right)^\dagger E_{S_2} |\psi_1\rangle = \left\langle \sqrt{E_{S_2}}\psi_1 \middle| \sqrt{E_{S_2}}\psi_1 \right\rangle = \left\| \sqrt{E_{S_2}} |\psi_1\rangle \right\|^2$$

which implies $\sqrt{E_{S_2}} |\psi_1\rangle = 0$. Next, decompose $|\psi_2\rangle = \alpha |\psi_1\rangle + \beta |\phi\rangle$ where $\langle\psi_1|\phi\rangle = 0$, $\langle\phi|\phi\rangle = 1$, $|\alpha|^2 + |\beta|^2 = 1$ and $|\beta|^2 < 1$ then

$$\sqrt{E_{S_2}} |\psi_2\rangle = \alpha\sqrt{E_{S_2}} |\psi_1\rangle + \beta\sqrt{E_{S_2}} |\phi\rangle = \beta\sqrt{E_{S_2}} |\phi\rangle .$$

Taking the norm and applying the inequality

$$\langle\phi| E_{S_2} |\phi\rangle \leq \sum_{i=1}^{N_1} \langle\phi| E_{S_i} |\phi\rangle = 1$$

yields

$$\langle\psi_2| E_{S_2} |\psi_2\rangle = |\beta|^2 \langle\phi| E_{S_2} |\phi\rangle < 1.$$

This is a contradiction to (2.4). ∎

An important concept regarding measurements, is relative- and global phase. Let $|\psi_1\rangle = \sum_{i=1}^{N} \alpha_i |\phi_i\rangle$ and $|\psi_2\rangle = \sum_{i=1}^{N} \beta_i |\phi_i\rangle$, thus $|\psi_1\rangle$ and $|\psi_2\rangle$ are superposition of the same states but possibly with different amplitudes. If $\alpha_i = e^{i\theta_i}\beta_i$ for some $\theta_i \in \mathbb{R}$, α_i and β_i are said to differ by a relative phase factor of $e^{i\theta}$. The requirement $\alpha_i = e^{i\theta}\beta_i$ implies $|\alpha_i| = |\beta_i|$, meaning α_i and β_i differs only by a relative phase if they lie on the same circle periphery in the complex plane. If the relative phase factor are equal for all i, it is called a global phase. In this case $|\psi_1\rangle = e^{i\theta} |\psi_2\rangle$. The key difference between relative- and global phase, is that from an observers point of view, two states differing by a global phase have the same statistics and are thus identical. To see this let $\{M_j\}_{j=1}^{N}$ be a set of measurement operators then according to (2.1)

$$p_j = \langle\psi_1| M_j^\dagger M_j |\psi_1\rangle = \langle\psi_2| e^{-i\theta} M_j^\dagger M_j e^{i\theta} |\psi_2\rangle = \langle\psi_2| M_j^\dagger M_j |\psi_2\rangle .$$

The fact that this is not the case if the phase is only relative, can be seen by considering the below example

$$|\psi_1\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \qquad |\psi_2\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

The amplitudes of the states differ by a relative phase factor of 1 and -1, respectively, however, they are easily verified to be orthogonal and they can hence be distinguished, meaning they cannot have the same statistics. Because of this phase condition, if a quantum system is in state $|\psi\rangle$, the physical state is said to be the ray generated by $|\psi\rangle$. A ray generated by $|\psi\rangle$ are all the vectors of the form $\{\alpha |\psi\rangle \mid \alpha \in \mathbb{C}\}$. However, if we impose the constraint of unit length, then

$$\langle\psi| \alpha^* \alpha |\psi\rangle = |\alpha|^2 \langle\psi|\psi\rangle = 1 \implies |\alpha|^2 = 1.$$

The complex scalar will hence be of the form $\alpha = e^{i\theta}$ for some $\theta \in \mathbb{R}$ [Nielsen and Chuang, 2010, p.93].
The fact that two states differing by a global phase are identical allows for a convenient visual representation of state vectors. Let $|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$ with $|\alpha|^2 + |\beta|^2 = 1$, transforming to polar coordinates, the state vector can be expressed by

$$|\psi\rangle = e^{i\gamma} \left(\cos\left(\frac{\theta}{2}\right) |0\rangle + e^{i\varphi} \sin\left(\frac{\theta}{2}\right) |1\rangle \right), \qquad \theta \in [0, \pi],\ \varphi \in [0, 2\pi].$$

Ignoring the global phase $e^{i\gamma}$, the state vector is described by the two angles $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi]$;

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right) |0\rangle + e^{i\varphi} \sin\left(\frac{\theta}{2}\right) |1\rangle , \qquad \theta \in [0, \pi],\ \varphi \in [0, 2\pi].$$

This state vector can be visualized on the Bloch Sphere; a 3-dimensional unit ball shown in figure 2.1. The Bloch Sphere can also be used to describe transformations of state vectors.

Consider the X-matrix shown in the introduction of the chapter. Ignoring a global phase and using the connection between sine and cosine, the X-matrix acting on a state vector is expressed by

$$X\left|\psi\right\rangle = \cos\left(\frac{\pi-\theta}{2}\right)\left|0\right\rangle + e^{-i\varphi}\sin\left(\frac{\pi-\theta}{2}\right)\left|1\right\rangle.$$

This action is visualized in the Bloch Sphere in figure 2.1 [Nielsen and Chuang, 2010, p.15].

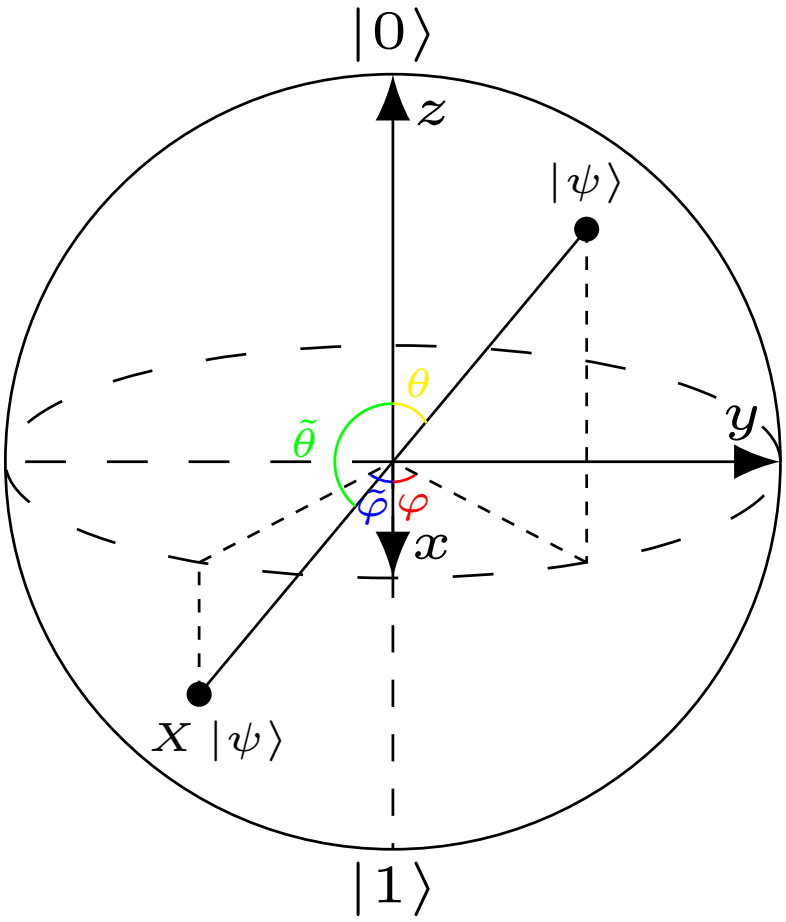


Figure 2.1: The transformation of the X matrix on the Bloch Sphere. Here $\tilde{\theta} := \pi - \theta$ and $\tilde{\varphi} := -\varphi$.

Before presenting the last postulate of quantum mechanics, the tensor product will be introduced. The tensor product space of two vector spaces $\mathcal{H}_{\mathcal{A}_1}$ and $\mathcal{H}_{\mathcal{A}_2}$ denoted $\mathcal{H}_{\mathcal{A}_1} \otimes \mathcal{H}_{\mathcal{A}_1}$ is a vector space, consisting of the set of all linear combinations of tensor product $\sum_{i=1}^{N} \alpha_i \left(\left|a_{1i}\right\rangle \otimes \left|a_{2i}\right\rangle\right)$ where $\left|a_{1i}\right\rangle \in \mathcal{H}_{\mathcal{A}_1}$ and $\left|a_{2i}\right\rangle \in \mathcal{H}_{\mathcal{A}_2}$. By definition the tensor product satisifies the following three properties: Let $\alpha_i \in \mathbb{C}$, $\left|a_1\right\rangle, \left|a_{11}\right\rangle, \left|a_{12}\right\rangle \in \mathcal{H}_{\mathcal{A}_1}$ and $\left|a_{21}\right\rangle, \left|a_{22}\right\rangle, \left|a_2\right\rangle \in \mathcal{H}_{\mathcal{A}_2}$, then

1. $$\alpha_i(\left|a_1\right\rangle \otimes \left|a_2\right\rangle) = (\alpha_i \left|a_1\right\rangle) \otimes \left|a_2\right\rangle = \left|a_1\right\rangle \otimes (\alpha_i \left|a_2\right\rangle),$$

2. $$(|a_{11}\rangle + |a_{12}\rangle) \otimes |a_2\rangle = |a_{11}\rangle \otimes |a_2\rangle + |a_{12}\rangle \otimes |a_2\rangle,$$

3. $$|a_1\rangle \otimes (|a_{21}\rangle + |a_{22}\rangle) = |a_1\rangle \otimes |a_{21}\rangle + |a_1\rangle \otimes |a_{22}\rangle.$$

The tensor product $\otimes : \mathcal{H}_{\mathcal{A}_1} \times \mathcal{H}_{\mathcal{A}_2} \to \mathcal{H}_{\mathcal{A}_1} \otimes \mathcal{H}_{\mathcal{A}_2}$ is hence a bilinear map with a scalar multiplication property (1.). An inner product can also be defined on the tensor product space using the inner product on $\mathcal{H}_{\mathcal{A}_1}$ and $\mathcal{H}_{\mathcal{A}_2}$ by

$$\left(\sum_{i=1}^{N_1} \alpha_i(|a_{1i}\rangle \otimes a_{2i}), \sum_{j=1}^{N_2} \beta_j \left(\left|a'_{1j}\right\rangle \otimes \left|a'_{2j}\right\rangle\right)\right) := \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \alpha_i^* \beta_j \left\langle a_{1i}\middle|a'_{1j}\right\rangle \left\langle a_{2i}\middle|a'_{2j}\right\rangle.$$

The tensor product between two matrices $A \in \mathbb{C}^{M\times N}$ and $B \in \mathbb{C}^{P\times Q}$ can now be defined as

$$A \otimes B = \begin{bmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{M1} & \cdots & a_{MN} \end{bmatrix} \otimes \begin{bmatrix} b_{11} & \cdots & b_{1Q} \\ \vdots & \ddots & \vdots \\ b_{P1} & \cdots & b_{PQ} \end{bmatrix} := \begin{bmatrix} a_{11}B & \cdots & a_{1N}B \\ \vdots & \ddots & \vdots \\ a_{M1}B & \cdots & a_{MN}B \end{bmatrix}.$$

The result thus yields a matrix $A \otimes B \in \mathbb{C}^{MP\times NQ}$. The vector version is a special case of the above. Consider the example below of the tensor product of the first basis vector of $\mathbb{C}^2$ with itself and notice the notation:

$$|0\rangle\,|0\rangle := |00\rangle := |0\rangle \otimes |0\rangle = \begin{bmatrix} 1 & 0 \end{bmatrix}^T \otimes \begin{bmatrix} 1 & 0 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}^T.$$

The tensor product between two vectors of $\mathbb{C}^2$ are thus a vector in $\mathbb{C}^4$. The tensor product are the central theme of the fourth postulate, concerning composite quantum systems [Nielsen and Chuang, 2010, pp.72-74].

Postulate 2.5
The state space of a composition of N quantum systems, $\mathcal{A} = \mathcal{A}_1 \cdots \mathcal{A}_N$, is the tensor product of the state space of each of the quantum systems, $\mathcal{H}_{\mathcal{A}} = \bigotimes_{i=1}^{N} \mathcal{H}_{\mathcal{A}_i}$. Furthermore, if system i has state vector $|\psi_i\rangle$, the state vector of the composite quantum system is $|\psi\rangle = \bigotimes_{i=1}^{N} |\psi_i\rangle$.

[Nielsen and Chuang, 2010, p.94]

The introduction of tensor products in postulate 2.5 may seem surprising. One could expect the composition of quantum systems to be described by product spaces rather than tensor product spaces. However, product spaces describes the composite system independently through each subsystem, in reality, this is not the case as the introduction of entanglement shows.

2.2 Entanglement and Superdensecoding

Entanglement is one of the many surprises of quantum mechanics.

Definition 2.6: Entanglement
Let $\mathcal{A} = \mathcal{A}_1 \cdots \mathcal{A}_N$ be a composite quantum system with state space $\mathcal{H}_\mathcal{A} = \bigotimes_{i=1}^N \mathcal{H}_{\mathcal{A}_i}$. A state vector $|\psi\rangle \in \mathcal{H}_\mathcal{A}$ is an entangled state vector if no state vectors, $|\psi_i\rangle \in \mathcal{H}_{\mathcal{A}_i}$, exists such $|\psi\rangle = \bigotimes_{i=1}^N \alpha_i |\psi_i\rangle$.

[Nielsen and Chuang, 2010, p.95]

The four bell states (or EPR pairs) are examples of entangled states.

Example 1: Bell States
The four Bell states are given by:

$$\begin{aligned}
\frac{|00\rangle + |11\rangle}{\sqrt{2}} &= \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}^T \\
\frac{|00\rangle - |11\rangle}{\sqrt{2}} &= \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix}^T \\
\frac{|01\rangle + |10\rangle}{\sqrt{2}} &= \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}^T \\
\frac{|01\rangle - |10\rangle}{\sqrt{2}} &= \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix}^T .
\end{aligned}$$

Writing the first bell state in terms of a tensor product between two state vectors in $\mathbb{C}^2$

$$\begin{bmatrix} a_1 & a_2 \end{bmatrix}^T \otimes \begin{bmatrix} b_1 & b_2 \end{bmatrix}^T = \begin{bmatrix} a_1 b_1 & a_1 b_2 & a_2 b_1 & a_2 b_2 \end{bmatrix}^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}^T .$$

This gives rise to the following equations,

$$a_1 b_1 = \frac{1}{\sqrt{2}} \wedge (a_1 = 0 \vee b_2 = 0) \wedge (a_2 = 0 \vee b_1 = 0) \wedge a_2 b_2 = \frac{1}{\sqrt{2}}.$$

Obviously, all of these cannot be satisfied at once. The same applies to the other bell states, and they are hence said to be entangled states.

Entanglement entwines two quantum systems such one cannot make a change in one of the systems without affecting the other. This proves very useful in quantum coding theory. An example of how i can be used is superdensecoding. This allows for the communication of two bits of information between a sender and a receiver, while transmitting only a single qubit between the two. The setup is as follows: A third party prepares an entangled state vector, this state is distributed between two parties, Alice and Bob. Alice now wishes to send two bits of information to Bob using only the single entangled qubit she was sent from the third party. The setup is shown in figure 2.2.

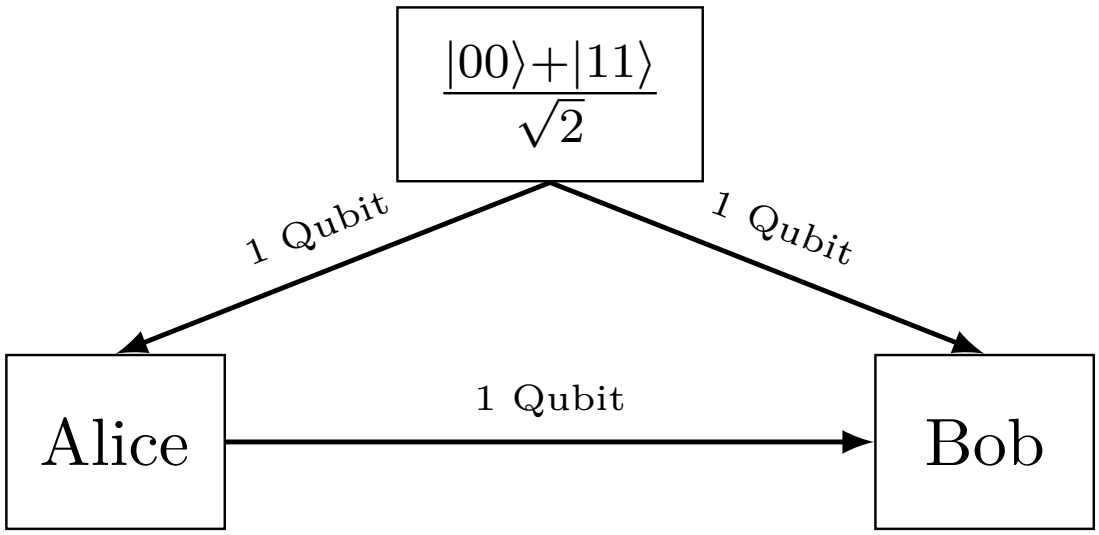


Figure 2.2: The setup of superdensecoding. A third party prepares a bell state and sends it to Alice and Bob. Alice then performs an operation on her qubit and sends it to Bob.

The way Alice is able to send two bits of information using only her single qubit, is by performing the correct operation on it. Consider the following codebook consisting of the Bell states:

$$\begin{aligned} 00 &: \frac{|00\rangle + |11\rangle}{\sqrt{2}} \\ 01 &: \frac{|00\rangle - |11\rangle}{\sqrt{2}} \\ 10 &: \frac{|01\rangle + |10\rangle}{\sqrt{2}} \\ 11 &: \frac{|01\rangle - |10\rangle}{\sqrt{2}} \end{aligned}$$

Given Alice's entangled qubit, she can perform the following operation to encode the information she wishes

$$\begin{aligned} (I \otimes I)\frac{|00\rangle + |11\rangle}{\sqrt{2}} &= \frac{|00\rangle + |11\rangle}{\sqrt{2}} \\ (Z \otimes I)\frac{|00\rangle + |11\rangle}{\sqrt{2}} &= \frac{|00\rangle - |11\rangle}{\sqrt{2}} \\ (X \otimes I)\frac{|00\rangle + |11\rangle}{\sqrt{2}} &= \frac{|01\rangle + |10\rangle}{\sqrt{2}} \\ (iY \otimes I)\frac{|00\rangle + |11\rangle}{\sqrt{2}} &= \frac{|01\rangle - |10\rangle}{\sqrt{2}}. \end{aligned}$$

After the operation has been performed, Alice sends her qubit to Bob. From example 1 it is clear the Bell states are orthogonal. Bob can then, by measurement, determine the exact state of the qubits reliably, revealing the two bits of information Alice wished to communicate [Nielsen and Chuang, 2010, pp.97-98].

2.3 The Density Operator

The previous section introduced the postulates of quantum mechanics, postulates which relied on the state vectors of the quantum systems. An equivalent foundation can be laid using the notion of density operators.

Definition 2.7: The Density Operator
Let $\{|\psi_i\rangle\}_{i=1}^N$ be a set of state vectors. Suppose a quantum system is in state $|\psi_i\rangle$ with probability p_i. The operator defined defined by

$$\rho := \sum_{i=1}^{N} p_i \, |\psi_i\rangle \langle\psi_i|$$

is called the density operator.

[Nielsen and Chuang, 2010, p.99]

The set consisting of the state vectors together with their respective probabilities, $\{p_i, |\psi_i\rangle\}_{i=1}^N$, is called an ensemble of pure states. If $N = 1$ the system is said to be in a pure state, else it is said to be in a mixed state. For a system in a pure state, the density operator satisfies

$$\operatorname{tr}(\rho\rho) = \operatorname{tr}(|\psi\rangle \langle\psi|\psi\rangle \langle\psi|) = \operatorname{tr}(|\psi\rangle \langle\psi|) = \langle\psi|\psi\rangle = 1.$$

Conversely, if the system is in a mixed state. Then

$$\begin{aligned}
\operatorname{tr}(\rho\rho) &= \operatorname{tr}\left(\sum_{i=1}^{N} p_i \, |\psi_i\rangle \langle\psi_i| \sum_{j=1}^{N} p_j \, |\psi_j\rangle \langle\psi_j|\right) \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N} p_i p_j \operatorname{tr}(|\psi_i\rangle \langle\psi_i|\psi_j\rangle \langle\psi_j|) \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N} p_i p_j \langle\psi_i|\psi_j\rangle \langle\psi_j|\psi_i\rangle \\
&= \sum_{i=1}^{N} p_i^2 + \sum_{i\neq j} p_i p_j |\langle\psi_i|\psi_j\rangle|^2 \\
&\leq \sum_{i=1}^{N} p_i^2 + \sum_{i\neq j} p_i p_j = \left(\sum_{i=1}^{N} p_i\right)^2 = 1,
\end{aligned}$$

with equality if and only if $|\psi_i\rangle = \alpha \, |\psi_j\rangle$ for $i \neq j$ and $\alpha \in \mathbb{C}$. However, from earlier it was seen that this implies $\alpha = e^{i\theta}$ for some $\theta \in \mathbb{R}$, meaning the states differ only by a global phase. In this case the system is said to be in a pure state, since they are a mixture of the same physical state [Bergou et al., 2021, p.18].
The trace condition of $\rho\rho$ can therefore be used to determine whether the system is in a pure or mixed state. Another unique property of the density operator, which proves useful when restating the postulates of the previous section in terms of the density operator, is shown in the following theorem.

Theorem 2.8: Properties of the density operator
The operator $\rho : \mathcal{H}_\mathcal{A} \to \mathcal{H}_\mathcal{A}$ is a density operator corresponding to the ensemble of pure states $\{p_i, |\psi_i\rangle\}$ if and only if $\mathrm{tr}(\rho) = 1$ and $\langle\phi| \rho |\phi\rangle \geq 0$ for all $|\phi\rangle \in \mathcal{H}_\mathcal{A}$.

[Nielsen and Chuang, 2010, p.101]

Proof
Let ρ be the density operator corresponding to the ensemble of pure states $\{p_i, |\psi_i\rangle\}_{i=1}^N$, that is,

$$\rho = \sum_{i=1}^{N} p_i |\psi_i\rangle \langle\psi_i| .$$

Then

$$\mathrm{tr}(\rho) = \mathrm{tr}\left(\sum_{i=1}^{N} p_i |\psi_i\rangle \langle\psi_i|\right) = \sum_{i=1}^{N} p_i \,\mathrm{tr}(|\psi_i\rangle \langle\psi_i|) = \sum_{i=1}^{N} p_i = 1,$$

and

$$\langle\phi| \rho |\phi\rangle = \sum_{i=1}^{N} p_i \langle\phi|\psi_i\rangle \langle\psi_i|\phi\rangle = \sum_{i=1}^{N} p_i |\langle\phi|\psi_i\rangle|^2 \geq 0.$$

Conversely, suppose the operator ρ satisfies the trace and positivity condition. Due to the positivity condition, it follows that ρ is hermitian and therefore has a spectral decomposition, hence

$$\rho = \sum_{i=1}^{N} \lambda_i |\psi_i\rangle \langle\psi_i| ,$$

where $\{\lambda_i\}_{i=1}^N$ are the non-negative eigenvalues of ρ and $\{|\psi_i\rangle\}_{i=1}^N$ are the corresponding orthonormal eigenvectors. From the trace condition it follows

$$1 = \mathrm{tr}(\rho) = \sum_{i=1}^{N} \lambda_i.$$

The operator ρ is hence a density operator corresponding to the ensemble of pure states $\{\lambda_i, |\psi_i\rangle\}_{i=1}^N$. ∎

The trace and positivity condition are the foundation of the density operator and will be the defining factor of the first postulate.

Postulate 2.9
The state space of any isolated quantum system, $\mathcal{A}$, is a complex hilbert space $\mathcal{H}_\mathcal{A}$. The quantum system is completely described by its density operator, $\rho : \mathcal{H}_\mathcal{A} \to \mathcal{H}_\mathcal{A}$, which is a positive semidefinite operator with trace equal to one. Let the quantum system be in state $\rho_i = |\psi_i\rangle \langle\psi_i|$ where $|\psi_i\rangle \in \mathcal{H}_\mathcal{A}$ with probability p_i, the density operator of the system is then $\sum_i p_i \rho_i$.

[Nielsen and Chuang, 2010, p.102]

The second postulate of quantum mechanics, describing the time evolution of a quantum system, can now be motivated with the notion of density operators rather than state vectors. Let $\{p_i, |\psi_i\rangle\}_{i=1}^N$ be an ensemble of pure states. Suppose the quantum system was in state $|\psi(t_1)\rangle$ at time t_1, then after some time evolution the system will be in state $U(t_1, t_2)\,|\psi(t_2)\rangle$ where U are some unitary transformation, according to postulate 2.2. Denote by $\rho(t_1)$ the density operator before the time evolution, then after the time evolution, the density operator is given by

$$\begin{aligned}\rho(t_2) &= \sum_{i=1}^{N} p_i U(t_1, t_2)\,|\psi_i(t_1)\rangle\,\langle\psi_i(t_1)|\,U^\dagger(t_1, t_2) \\ &= U(t_1, t_2)\left(\sum_{i=1}^{N} p_i\,|\psi_i(t_1)\rangle\,\langle\psi_i(t_1)|\right)U^\dagger(t_1, t_2) = U(t_1, t_2)\rho(t_1)U^\dagger(t_1, t_2).\end{aligned}$$

This is exactly the statement of the second postulate [Nielsen and Chuang, 2010, p.99].

Postulate 2.10
The time evolution of a closed quantum system is described by a unitary transformation, $U(t_1, t_2)$ relating the state at time t_1 to the state at time t_2 by

$$\rho(t_2) = U(t_1, t_2)\rho(t_1)U^\dagger(t_1, t_2).$$

[Nielsen and Chuang, 2010, p.102]

The density operator can also be used to describe measurements, which was the central principle of postulate 2.3. Let $\{|\psi_i\rangle\}_{i=1}^{N_1}$ be a set of state vector and $\{M_j\}_{j=1}^{N_2}$ be a set of measurement operators. Let the quantum system be in state $|\psi_i\rangle$ with probability p_i, the associated density operator of the quantum system is then given by $\rho = \sum_{i=1}^{N_1} p_i\,|\psi_i\rangle\,\langle\psi_i|$. The probability of getting measurement j, if the quantum system was in state $|\psi_i\rangle$ is

$$p_{j|i} := \langle\psi_i|\,M_j^\dagger M_j\,|\psi_i\rangle = \mathrm{tr}\Big(M_j^\dagger M_j\,|\psi_i\rangle\,\langle\psi_i|\Big) \tag{2.5}$$

according to postulate 2.3 and the fact that if $\{|\phi_1\rangle = |\psi_i\rangle, |\phi_2\rangle, \ldots, |\phi_{N_3}\rangle\}$ is an orthonormal basis for the vector space which $|\psi_i\rangle$ belongs to and A are some operator on that vector space, then

$$\mathrm{tr}(A\,|\psi_i\rangle\,\langle\psi_i|) = \sum_{j=1}^{N_3}\langle\phi_j|\,A\,|\psi_i\rangle\,\langle\psi_i|\phi_j\rangle = \langle\psi_i|\,A\,|\psi_i\rangle\,.$$

By the law of total probability it then follows

$$p_j = \sum_{i=1}^{N_1} p_{j|i}p_i = \sum_{i=1}^{N_1} p_i\,\mathrm{tr}\Big(M_j^\dagger M_j\,|\psi_i\rangle\,\langle\psi_i|\Big) = \mathrm{tr}\Big(M_j^\dagger M_j\rho\Big), \tag{2.6}$$

which describes the probability of getting outcome j of a measurement in terms of the density operator rather than the state vector. The state after the measurement can also be

described in terms of the density operator. By postulate 2.3, if the measurement resulted in outcome j

$$|\psi_i\rangle_j^{\text{post}} = \frac{M_j |\psi_i\rangle}{\sqrt{\langle\psi_i| M_j^\dagger M_j |\psi_i\rangle}}.$$

The ensemble of states after the measurement is therefore $\{p_{i|j}, |\psi_i\rangle_j^{\text{post}}\}$. The corresponding density operator after the measurement is hence

$$\begin{aligned}\rho_j^{\text{post}} &= \sum_{i=1}^{N_1} p_{i|j} |\psi_i\rangle_j^{\text{post}} \langle\psi_i|_j^{\text{post}} \\ &= \sum_{i=1}^{N_1} p_{i|j} \frac{M_j |\psi_i\rangle}{\sqrt{\langle\psi_i| M_j^\dagger M_j |\psi_i\rangle}} \frac{\langle\psi_i| M_j^\dagger}{\sqrt{\langle\psi_i| M_j^\dagger M_j |\psi_i\rangle}},\end{aligned}$$

using Bayes formula along with (2.5) and (2.6) the post-measurement density operator can be rewritten as

$$\begin{aligned}\rho_j^{\text{post}} &= \sum_{i=1}^{N_1} \frac{p_{j|i} p_i}{p_j} \frac{M_j |\psi_i\rangle \langle\psi_i| M_j^\dagger}{\langle\psi_i| M_j^\dagger M_j |\psi_i\rangle} \\ &= \frac{M_j \sum_{i=1}^{N_1} p_i |\psi_i\rangle \langle\psi_i| M_j^\dagger}{\text{tr}\Big(M_j^\dagger M_j \rho\Big)} = \frac{M_j \rho M_j^\dagger}{\text{tr}\Big(M_j^\dagger M_j \rho\Big)}\end{aligned}$$

The above expression for the post-measurement density operator together with (2.6) is exactly the foundation of the next postulate [Nielsen and Chuang, 2010, pp.99-100].

Postulate 2.11
Quantum measurements are described by a set of measurement operators $\{M_j\}_{j=1}^N$ on the state space, $\mathcal{H}_\mathcal{A}$, of the quantum system, $\mathcal{A}$, being measured. Let the quantum system be in state ρ before a measurement. The probability of that the measurement yields outcome j is given by

$$p_j = \text{tr}\Big(M_j \rho M_j^\dagger\Big).$$

The state of the quantum system after the measurement is given by

$$\rho_j^{\text{post}} = \frac{M_j \rho M_j^\dagger}{\text{tr}\Big(M_j \rho M_j^\dagger\Big)}.$$

[Nielsen and Chuang, 2010, p.102]

A surprising fact of the density operator, is that two different ensembles can correspond to the same density operator. The density operator is therefore not uniquely determined by a single ensemble. The next theorem, however, determines when two ensembles can be associated to the same density operator.

Theorem 2.12: Generating Sets for the Density Operator
Let $\mathcal{A}$ be a quantum system with state space, $\mathcal{H}_\mathcal{A}$. Let $\{p_i, |\psi_i\rangle\}_{i=1}^{N_1}$ and $\{q_j, |\phi_j\rangle\}_{j=1}^{N_2}$ be two ensembles of state vectors in $\mathcal{H}_\mathcal{A}$. The two sets generate the same density operator if and only if

$$\sqrt{p_i}\,|\psi_i\rangle = \sum_{j=1}^{N_2} u_{ij}\sqrt{q_j}\,|\phi_j\rangle$$

where $u_{ij} \in \mathbb{C}$ are the ij^{th} entry of a unitary matrix. The set with smallest cardinality is padded with vectors having probability zero such the two are of equal cardinality.

[Nielsen and Chuang, 2010, pp.103-104]

Proof
Let $\{p_i, |\psi_i\rangle\}_{i=1}^{N_1}$ and $\{q_j, |\phi_j\rangle\}_{j=1}^{N_2}$ be the sets of the theorem. Without loss of generality, assume the second set has the smallest cardinality and define a new set $\{q_j, |\phi_j\rangle\}_{j=1}^{N_1}$ where $q_j = 0$ for $j > N_2$.
First, assume $\sqrt{p_i}\,|\psi_i\rangle = \sum_{j=1}^{N_1} u_{ij}\sqrt{q_j}\,|\phi_j\rangle$. Then

$$\begin{aligned}
\sum_{i=1}^{N_1} p_i\,|\psi_i\rangle\langle\psi_i| &= \sum_{i=1}^{N_1}\left(\sum_{j=1}^{N_1} u_{ij}\sqrt{q_j}\,|\phi_j\rangle\right)\left(\sum_{k=1}^{N_1} u_{ik}^*\sqrt{q_j}\,\langle\phi_k|\right)\\
&= \sum_{i=1}^{N_1}\sum_{j=1}^{N_1}\sum_{k=1}^{N_1} u_{ij}u_{ik}^* q_j\,|\phi_j\rangle\langle\phi_k|\\
&= \sum_{j=1}^{N_1}\sum_{k=1}^{N_1}\sum_{i=1}^{N_1} (u_{ij}u_{ik}^*)\,q_j\,|\phi_j\rangle\langle\phi_k|\\
&= \sum_{j=1}^{N_1}\sum_{k=1}^{N_1} \delta_{jk} q_j\,|\phi_j\rangle\langle\phi_k|\\
&= \sum_{j=1}^{N_1} q_j\,|\phi_j\rangle\langle\phi_j|\,.
\end{aligned}$$

The two sets thus generate the same density operator.
Conversely, let

$$\rho = \sum_{i=1}^{N_1} p_i\,|\psi_i\rangle\langle\psi_i| = \sum_{j=1}^{N_1} q_j\,|\phi_j\rangle\langle\phi_j|\,.$$

ρ is easily seen to be hermitian, and by theorem 2.8 it is also positive semidefinite. By the spectral theorem, it then has a decomposition $\rho = \sum_{k=1}^{N_3}\lambda_k\,|a_k\rangle\langle a_k|$, where $\{|a_k\rangle\}_{k=1}^{N_3}$ are orthonormal and $\{\lambda_k\}_{k=1}^{N_3}$ are strictly positive (Notice $N_3 = \text{rank}(\rho) \leq \sum_{j=1}^{N_1}\text{rank}(q_j\,|\phi_j\rangle\langle\phi_j|) = N_2 \leq N_1$). Let $|b\rangle$ be an arbitrary vector orthogonal to the set $\{|a_k\rangle\}_{k=1}^{N_3}$, then

$$\langle b|\,\rho\,|b\rangle = \langle b|\sum_{k=1}^{N_3}\lambda_k\,|a_k\rangle\langle a_k|b\rangle = \sum_{k=1}^{N_3}\lambda_k\,\langle b|a_k\rangle\langle a_k|b\rangle = 0,$$

hence

$$0 = \langle b| \rho |b\rangle = \langle b| \sum_{i=1}^{N_1} p_i |\psi_i\rangle \langle\psi_i|b\rangle = \sum_{i=1}^{N_1} p_i \langle b|\psi_i\rangle \langle\psi_i|b\rangle = \sum_{i=1}^{N_1} p_i |\langle b|\psi_i\rangle|^2$$

implying all $\langle b|\psi_i\rangle = 0$ meaning $|b\rangle$ are also orthogonal to the set $\{|\psi_i\rangle\}_{i=1}^{N_1}$. This means any vector $|\psi_i\rangle$ must be in the span of $\{|a_k\rangle\}_{k=1}^{N_3}$, thus can be expressed as

$$|\psi_i\rangle = \sum_{k=1}^{N_3} c_{ik} |a_k\rangle .$$

Using this to rewrite ρ

$$\rho = \sum_{i=1}^{N_1} p_i |\psi_i\rangle \langle\psi_i| = \sum_{i=1}^{N_1} p_i \left(\sum_{k=1}^{N_3} c_{ik} |a_k\rangle \right) \left(\sum_{k'=1}^{N_3} c_{ik'}^* \langle a_{k'}| \right) = \sum_{i=1}^{N_1} \sum_{k=1}^{N_3} \sum_{k'=1}^{N_3} p_i c_{ik} c_{ik'}^* |a_k\rangle \langle a_{k'}|$$

The operators $|a_k\rangle \langle a_{k'}|$ are all linearly independent, and since

$$\rho = \sum_{k=1}^{N_3} \lambda_k |a_k\rangle \langle a_k| = \sum_{k=1}^{N_3} \sum_{k'=1}^{N_3} \sum_{i=1}^{N_1} (p_i c_{ik} c_{ik'}^*) |a_k\rangle \langle a_{k'}|$$

it must imply $\sum_{i=1}^{N_1} (\sqrt{\lambda_k})^{-1} \sqrt{p_i} c_{ik} c_{ik'}^* \sqrt{p_i} (\sqrt{\lambda_k})^{-1} = \delta_{kk'}$. Let $\tilde{C}_{ik} := (\sqrt{\lambda_k})^{-1} \sqrt{p_i} c_{ik}$ then $\tilde{C}\tilde{C}^\dagger = I$. Extra columns can now be appended to $\tilde{C}$ to make it unitary if $N_1 < N_3$ (the set $\{a_k\}_{k=1}^{N_3}$ are also expanded with zero vectors to gain cardinality N_1). Call this new square matrix V. Repeating the same argument as above for the set $\{|\phi_j\rangle\}_{j=1}^{N_1}$ yielding another unitary matrix W with indices $W_{jk} := (\sqrt{\lambda_k})^{-1} \sqrt{q_j} d_{jk}$. The matrix $U = VW^\dagger$ is then also unitary and

$$\sqrt{p_i} |\psi_i\rangle = \sum_{k=1}^{N_1} \sqrt{p_i} c_{ik} |a_k\rangle = \sum_{k=1}^{N_1} V_{ik} \sqrt{\lambda_k} |a_k\rangle = \sum_{k=1}^{N_1} \sum_{j=1}^{N_1} V_{ik} W_{kj}^\dagger \sqrt{q_j} |\phi_j\rangle = \sum_{j=1}^{N_1} u_{ij} \sqrt{q_j} |\phi_j\rangle . \blacksquare$$

The following example shows two ensembles corresponding to the same density operator, even though they might seem very different at first glance.

Example 2: Generating Sets for the Density Operator
Consider the two ensembles $\{\left(\frac{1}{2}, |0\rangle\right), \left(\frac{1}{2}, |1\rangle\right)\}$ and $\{\left(\frac{1}{2}, \frac{|0\rangle+|1\rangle}{\sqrt{2}}\right), \left(\frac{1}{2}, \frac{|0\rangle-|1\rangle}{\sqrt{2}}\right)\}$. The corresponding density operator is

$$\begin{aligned}
\rho &= \frac{1}{2} \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}} \right) \left(\frac{\langle 0| + \langle 1|}{\sqrt{2}} \right) + \frac{1}{2} \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}} \right) \left(\frac{\langle 0| - \langle 1|}{\sqrt{2}} \right) \\
&= \frac{1}{2} \left(\frac{|0\rangle \langle 0| + |0\rangle \langle 1| + |1\rangle \langle 0| + |1\rangle \langle 1|}{2} \right) + \frac{1}{2} \left(\frac{|0\rangle \langle 0| - |0\rangle \langle 1| - |1\rangle \langle 0| + |1\rangle \langle 1|}{2} \right) \\
&= \frac{1}{2} |0\rangle \langle 0| + \frac{1}{2} |1\rangle \langle 1|
\end{aligned}$$

Showing that both ensembles are associated with the same density operator. This is in correspondance with theorem 2.12 since

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

and

$$\frac{|0\rangle - |1\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}|0\rangle + \frac{-1}{\sqrt{2}}|1\rangle$$

The unitary matrix of theorem 2.12 is therefore

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

That is, the Hadamard matrix.

The density operator can also be used to describe composite quantum systems as the last postulate proposes.

Postulate 2.13
The state space of a composition of N quantum systems, $\mathcal{A} = \mathcal{A}_1 \cdots \mathcal{A}_N$, is the tensor product of the state space of each of the quantum systems, $\mathcal{H}_\mathcal{A} = \bigotimes_{i=1}^{N} \mathcal{H}_{\mathcal{A}_i}$. If system i is in state $\rho^{\mathcal{A}_i}$, the state of the composite quantum system is $\rho = \bigotimes_{i=1}^{N} \rho^{\mathcal{A}_i}$.

[Nielsen and Chuang, 2010, p.102]

According to postulate 2.13, the density operator for a composite system, is given as the tensor product of the density operators for each subsystem. One might wonder, if the composite systems density operator is known, how can one find the density operator for each subsystem. The answer is to "trace out" all other systems besides the one of interest. This is exactly what the reduced density operator is.

Definition 2.14: Reduced Density Operator
Let $\mathcal{A}_1\mathcal{A}_2$ be a composite quantum system with density operator

$$\begin{aligned}
\rho^{\mathcal{A}_1\mathcal{A}_2} &= \sum_{i=1}^{N} p_i |\psi_i\rangle \langle\psi_i| \\
&= \sum_{i=1}^{N} p_i \left(\sum_{j=1}^{N_1} \sum_{k=1}^{N_2} \alpha_{jk} (|a_{1j}\rangle \otimes |a_{2k}\rangle) \right) \left(\sum_{j'=1}^{N_1} \sum_{k'=1}^{N_2} \alpha_{j'k'} (\langle a_{1j'}| \otimes \langle a_{2k'}|) \right) \\
&= \sum_{i=1}^{N} \sum_{j=1}^{N_1} \sum_{k=1}^{N_2} \sum_{j'=1}^{N_1} \sum_{k'=1}^{N_2} \beta_{ijkj'k'} \left(|a_{1j}\rangle \langle a_{1j'}| \otimes |a_{2k}\rangle \langle a_{2k'}| \right),
\end{aligned}$$

where $|a_{1j}\rangle \in \mathcal{H}_{\mathcal{A}_1}$, $|a_{2k}\rangle \in \mathcal{H}_{\mathcal{A}_2}$ and $|\psi_i\rangle \in \mathcal{H}_{\mathcal{A}_1\mathcal{A}_2}$. The reduced density operator of system $\mathcal{A}_1$ is

$$\rho^{\mathcal{A}_1} := \mathrm{tr}_{\mathcal{A}_2}(\rho^{\mathcal{A}_1\mathcal{A}_2}),$$

where

$$\mathrm{tr}_{\mathcal{A}_2} \left(\sum_{i,j,k,j',k'} \beta_{ijkj'k'} \left(|a_{1j}\rangle \langle a_{1j'}| \otimes |a_{2k}\rangle \langle a_{2k'}| \right) \right) := \sum_{i,j,k,j',k'} \beta_{ijkj'k'} |a_{1j}\rangle \langle a_{1j'}| \, \mathrm{tr}(|a_{2k}\rangle \langle a_{2k'}|)$$

$$= \sum_{i,j,k,j',k'} \beta_{ijkj'k'} |a_{1j}\rangle \langle a_{1j'}| \langle a_{2k}|a_{2k'}\rangle .$$

[Nielsen and Chuang, 2010, pp.105-106]

To motivate the use of the partial trace, let $\{p_i, |\psi_i\rangle\}_{i=1}^{N_1}$ and $\{q_j, |\phi_j\rangle\}_{j=1}^{N_2}$ be two ensembles and $\rho^{\mathcal{A}_1}$ and $\rho^{\mathcal{A}_2}$ be the associated density operators respectively. The composite density operator is now given by postulate 2.13, and the reduced density operator can then be found as

$$\begin{aligned}
\mathrm{tr}_{\mathcal{A}_2}(\rho^{\mathcal{A}_1\mathcal{A}_2}) &= \mathrm{tr}_{\mathcal{A}_2}\left(\left(\sum_{i=1}^{N_1} p_i |\psi_i\rangle \langle\psi_i|\right) \otimes \left(\sum_{j=1}^{N_2} q_j |\phi_j\rangle \langle\phi_j|\right)\right) \\
&= \mathrm{tr}_{\mathcal{A}_2}\left(\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} p_i q_j (|\psi_i\rangle \langle\psi_i| \otimes |\phi_j\rangle \langle\phi_j|)\right) \\
&= \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} p_i q_j |\psi_i\rangle \langle\psi_i| \langle\phi_j|\phi_j\rangle \\
&= \sum_{i=1}^{N_1} p_i |\psi_i\rangle \langle\psi_i| = \rho^{\mathcal{A}_1}
\end{aligned}$$

The reduced density operator of system $\mathcal{A}_1$ is exactly the density operator of system $\mathcal{A}_1$. The next examples shows a remarkable result regarding the reduced density operator of a Bell state.

Example 3: Reduced Density Operator
Consider the bell state $(\sqrt{2})^{-1}(|00\rangle + |11\rangle)$, the corresponding density operator is

$$\begin{aligned}
\rho &= \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right)\left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right) \\
&= \frac{|00\rangle \langle 00| + |00\rangle \langle 11| + |11\rangle \langle 00| + |11\rangle \langle 11|}{2}.
\end{aligned}$$

In this case, the density operator of the first system is given by

$$\begin{aligned}
\rho^{\mathcal{A}_1} &= \mathrm{tr}_{\mathcal{A}_2}(\rho) \\
&= \mathrm{tr}_{\mathcal{A}_2}\left(\frac{|00\rangle \langle 00| + |00\rangle \langle 11| + |11\rangle \langle 00| + |11\rangle \langle 11|}{2}\right) \\
&= \frac{\mathrm{tr}_{\mathcal{A}_2}(|00\rangle \langle 00|) + \mathrm{tr}_{\mathcal{A}_2}(|00\rangle \langle 11|) + \mathrm{tr}_{\mathcal{A}_2}(|11\rangle \langle 00|) + \mathrm{tr}_{\mathcal{A}_2}(|11\rangle \langle 11|)}{2} \\
&= \frac{\mathrm{tr}_{\mathcal{A}_2}(|0\rangle \langle 0| \otimes |0\rangle \langle 0|) + \mathrm{tr}_{\mathcal{A}_2}(|0\rangle \langle 1| \otimes |0\rangle \langle 1|) + \mathrm{tr}_{\mathcal{A}_2}(|1\rangle \langle 0| \otimes |1\rangle \langle 0|) + \mathrm{tr}_{\mathcal{A}_2}(|1\rangle \langle 1| \otimes |1\rangle \langle 1|)}{2} \\
&= \frac{|0\rangle \langle 0| \langle 0|0\rangle + |0\rangle \langle 1| \langle 0|1\rangle + |1\rangle \langle 0| \langle 1|0\rangle + |1\rangle \langle 1| \langle 1|1\rangle}{2} \\
&= \frac{|0\rangle \langle 0| + |1\rangle \langle 1|}{2} = \frac{1}{2} I.
\end{aligned}$$

It then follows $\rho\rho = \frac{1}{4} I$ and therefore $\mathrm{tr}(\rho\rho) = \frac{1}{2} < 1$, meaning the first system is in a mixed state, even though the composite system is in a pure state. If the composite state is

entangled, and the reduced density matrices are proportional to the identity, the composite state is said to be maximally entangled [Bergou et al., 2021, pp.31-32].

Example 3 showed that even if two systems are in a mixed state, the composite system can be in a pure state. This is the idea behind purification. Purification is the theoretical act of introducing an auxiliary system, $\mathcal{A}_a$, such that if a system, $\mathcal{A}_1$ is in a mixed state, the composite system, $\mathcal{A}_1\mathcal{A}_a$ is in a pure state. To see this, let $\rho^{\mathcal{A}_1} = \sum_{i=1}^{N} p_i \left|\psi_i\right\rangle \left\langle\psi_i\right|$ be in a mixed state. Let further $\{\left|\phi_i\right\rangle\}_{i=1}^{N}$ be an orthonormal set in $\mathcal{H}_{\mathcal{A}_a}$. Define a state vector by

$$\left|\psi\right\rangle := \sum_{i=1}^{N} \sqrt{p_i}(\left|\psi_i\right\rangle \otimes \left|\phi_i\right\rangle).$$

Suppose the state of the composite system is $\rho = \left|\psi\right\rangle \left\langle\psi\right|$, the composite system is hence in a pure state and

$$\begin{aligned}
\rho^{\mathcal{A}_1} = \mathrm{tr}_{\mathcal{A}_a}(\rho) &= \mathrm{tr}_{\mathcal{A}_a}\left(\sum_{i=1}^{N} \sqrt{p_i}(\left|\psi_i\right\rangle \otimes \left|\phi_i\right\rangle) \sum_{i'=1}^{N} \sqrt{p_{i'}}(\left\langle\psi_{i'}\right| \otimes \left\langle\phi_{i'}\right|)\right) \\
&= \mathrm{tr}_{\mathcal{A}_a}\left(\sum_{i=1}^{N}\sum_{i'=1}^{N} \sqrt{p_i}\sqrt{p_{i'}}(\left|\psi_i\right\rangle \left\langle\psi_{i'}\right| \otimes \left|\phi_i\right\rangle \left\langle\phi_{i'}\right|)\right) \\
&= \sum_{i=1}^{N}\sum_{i'=1}^{N} \sqrt{p_i}\sqrt{p_{i'}}\, \mathrm{tr}_{\mathcal{A}_a}(\left|\psi_i\right\rangle \left\langle\psi_{i'}\right| \otimes \left|\phi_i\right\rangle \left\langle\phi_{i'}\right|) \\
&= \sum_{i=1}^{N}\sum_{i'=1}^{N} \sqrt{p_i}\sqrt{p_{i'}}\, \left|\psi_i\right\rangle \left\langle\psi_{i'}\right| \left\langle\phi_i|\phi_{i'}\right\rangle \\
&= \sum_{i=1}^{N} p_i \left|\psi_i\right\rangle \left\langle\psi_i\right|
\end{aligned}$$

The introduced auxiliary system has no physical meaning, which is why the act of purification is a theoretical tool [Nielsen and Chuang, 2010, pp.110-111].
The density operator of a composite system in a pure state reveals an important property of the reduced density operator of the subsystems. To derive this property, the Schmidt decomposition is a necessary tool.

Theorem 2.15: The Schmidt Decomposition
Let $\mathcal{H}_{\mathcal{A}_1}$ and $\mathcal{H}_{\mathcal{A}_2}$ be the state space of the two quantum systems, $\mathcal{A}_1$ and $\mathcal{A}_2$ respectively and let $\mathcal{H}_{\mathcal{A}} := \mathcal{H}_{\mathcal{A}_1} \otimes \mathcal{H}_{\mathcal{A}_2}$ be the state space of the composite system $\mathcal{A} := \mathcal{A}_1\mathcal{A}_2$. For any state vector $\left|\psi\right\rangle \in \mathcal{H}_{\mathcal{A}}$ there exists orthonormal state vectors $\{a_{1i}\}_{i=1}^{N}$ and $\{a_{2i}\}_{i=1}^{N}$ in $\mathcal{H}_{\mathcal{A}_1}$ and $\mathcal{H}_{\mathcal{A}_2}$ respectively, such

$$\left|\psi\right\rangle = \sum_{i=1}^{N} \lambda_i(\left|a_{1i}\right\rangle \otimes \left|a_{2i}\right\rangle),$$

where $\sum_{i=1}^{N} \lambda_i^2 = 1$ and $\lambda_i \geq 0$ for all i.

[Nielsen and Chuang, 2010, p.109]

Proof
Let $\{a'_{1k}\}_{k=1}^{N_1}$ and $\{a'_{2j}\}_{j=1}^{N_2}$ be orthonormal bases for $\mathcal{H}_{\mathcal{A}_1}$ and $\mathcal{H}_{\mathcal{A}_2}$ respectively. Without loss of generality, assume $N_1 \leq N_2$. Then for any $|\psi\rangle \in \mathcal{H}_{\mathcal{A}}$

$$|\psi\rangle = \sum_{k=1}^{N_1}\sum_{j=1}^{N_2} \alpha_{kj}(|a'_{1k}\rangle \otimes \left|a'_{2j}\right\rangle)$$

Let A be the matrix with indices $A_{kj} = \alpha_{kj}$, then by the singular value decomposition $A = U\Sigma V$ where $U \in \mathbb{C}^{N_1\times N_1}$ and $V \in \mathbb{C}^{N_2\times N_2}$ are unitary, and $\Sigma \in \mathbb{R}^{N_1\times N_2}$ is a diagonal matrix with non-negative entries. It then follows

$$\begin{aligned}|\psi\rangle &= \sum_{k=1}^{N_1}\sum_{j=1}^{N_2}\sum_{i=1}^{N_1} u_{ki}\sigma_{ii}v_{ij}(|a'_{1k}\rangle \otimes \left|a'_{2j}\right\rangle) \\ &= \sum_{i=1}^{N_1}\sigma_{ii}\left(\sum_{j=1}^{N_2} v_{ij}\left|a'_{2j}\right\rangle\right) \otimes \left(\sum_{k=1}^{N_1} u_{ki}\left|a'_{1k}\right\rangle\right)\end{aligned}$$

Defining $\sigma_{ii} := \lambda_i$, $|a_{1i}\rangle := \sum_{k=1}^{N_1} u_{ki}|a'_{1k}\rangle$ and $|a_{2i}\rangle := \sum_{j=1}^{N_2} v_{ij}\left|a'_{2j}\right\rangle$. Then since U and V are unitary it follows

$$\langle a_{1i'}|a_{1i}\rangle = \sum_{k=1}^{N_1}\sum_{k'=1}^{N_1} u^*_{k'i'}u_{ki}\langle a'_{1k'}|a'_{1k}\rangle = \sum_{k=1}^{N_1} u^*_{ki'}u_{ki} = \begin{cases}1, & \text{if } i = i' \\ 0, & \text{else,}\end{cases}$$

and similarly

$$\langle a_{2i'}|a_{2i}\rangle = \begin{cases}1, & \text{if } i' = i \\ 0, & \text{else,}\end{cases}$$

Furthermore

$$\begin{aligned}1 = \langle\psi|\psi\rangle &= \left(\sum_{i'=1}^{N_1}\sigma_{i'i'}(\langle a_{1i'}| \otimes \langle a_{2i'}|)\right)\left(\sum_{i=1}^{N_1}\sigma_{ii}(|a_{1i}\rangle \otimes |a_{2i}\rangle)\right) \\ &= \sum_{i'=1}^{N_1}\sum_{i=1}^{N_1}\sigma_{i'i'}\sigma_{ii}\langle a_{1i'}|a_{1i}\rangle\langle a_{2i'}|a_{2i}\rangle = \sum_{i=1}^{N_1}\sigma_{ii}^2.\end{aligned}$$

■

The non-negative scalars $\{\lambda_i\}_{i=1}^{N}$ of the Schmidt decomposition is called the Schmidt coefficients, and the number of non-zero Schmidt coefficients is called the Schmidt rank. The Schmidt decomposition has an important implication. Let the density operator for a composite system be in a pure state. Then

$$\begin{aligned}\rho = |\psi\rangle\langle\psi| &= \left(\sum_{i=1}^{N}\lambda_i(|a_{1i}\rangle \otimes |a_{2i}\rangle)\right)\left(\sum_{ij=1}^{N}\lambda_j(\langle a_{1j}| \otimes \langle a_{2j}|)\right) \\ &= \sum_{i=1}^{N}\sum_{ij=1}^{N}\lambda_i\lambda_j(|a_{1i}\rangle\langle a_{1j}| \otimes |a_{2i}\rangle\langle a_{2j}|).\end{aligned}$$

Deriving the reduced density operator for the subsystems

$$\rho^{\mathcal{A}_1} = \sum_{i=1}^{N}\sum_{ij=1}^{N}\lambda_i\lambda_j|a_{1i}\rangle\langle a_{1j}|\langle a_{2i}|a_{2j}\rangle = \sum_{i=1}^{N}\lambda_i^2|a_{1i}\rangle\langle a_{1i}| \tag{2.7}$$

and similarly

$$\rho^{\mathcal{A}_2} = \sum_{i=1}^{N} \sum_{ij=1}^{N} \lambda_i \lambda_j \langle a_{1i}|a_{1j}\rangle |a_{2i}\rangle \langle a_{2j}| = \sum_{i=1}^{N} \lambda_i^2 |a_{2i}\rangle \langle a_{2i}| . \tag{2.8}$$

The non-zero eigenvalues for both reduced density operators are therefore $\{\lambda_i^2\}_{i=1}^N$.

From the Schmidt decomposition, it can also be seen that $|\psi\rangle$ is entangled if and only if the Schmidt rank is strictly greater than 1. Suppose the Schmidt rank is $Q > 1$, then according to (2.7), assuming $\rho^{\mathcal{A}_1\mathcal{A}_2} = |\psi\rangle \langle\psi|$

$$\mathrm{tr}\Big(\rho^{\mathcal{A}_1}\rho^{\mathcal{A}_1}\Big) = \mathrm{tr}\left(\sum_{i=1}^{Q}\sum_{j=1}^{Q} \lambda_i^2\lambda_j^2 |a_{1i}\rangle \langle a_{1i}|a_{1j}\rangle \langle a_{1j}|\right) = \sum_{i=1}^{Q} \lambda_i^4 \,\mathrm{tr}(|a_{1i}\rangle \langle a_{1i}|) = \sum_{i=1}^{Q} \lambda_i^4 < 1,$$

where it has been used that $\{\lambda_i^2\}_{i=1}^Q$ sums to 1 and are non-negative, implying $0 < \lambda_i^2 < 1$ and thus $\lambda_i^4 < \lambda_i^2$. Altogether, this shows the state $\rho^{\mathcal{A}_1}$ is mixed if and only if $\rho^{\mathcal{A}_1\mathcal{A}_2}$ is an entangled state [Nielsen and Chuang, 2010, p.109].

The two corollaries of the Schmidt decomposition prove useful when analyzing the entropy of quantum states; unavoidable theory when exploring the limits of quantum channels.

With the general quantum mechanics theory in place, the focus can be laid upon quantum information theory. Describing how to use the quirks of quantum mechanics as a strength to store and/or send information from one place to another.

3 | Quantum Information Theory

Classical information theory quantifies information and studies the object of storing and communicating information. In this regard, two important concepts arrive; the entropy of an information source and communication channels.
The entropy quantifies the uncertainty in an information source, thereby the amount of information present in the source. This, in turn, provides details of how well the source can be compressed; important knowledge when storing or communicating the information in the source.
Communication channels describe the evolution of information when transmitted between two nodes, an important tool in the theoretical study of noise.

The two concepts are also present in the quantum realm and in a similar manner provide useful details about uncertainty in quantum states and a theoretical description of noise when transmitting qubits between two nodes. The chapter is primarily based on [Nielsen and Chuang, 2010].

3.1 Entropy

Before defining the entropy of quantum states, a quick recap of classical entropy is presented.

The Shannon entropy for the discrete random variable $X \sim p$ taking values in the alphabet $\mathcal{X}$ is defined by

$$H(X) := -\sum_{x \in \mathcal{X}} p(x) \log(p(x)), \tag{3.1}$$

with the convention $\log(0) := 0$. The logarithm is usually chosen to be base 2, leading to the entropy having bits as unit. Other bases can be chosen, leading to other units.
The definition can easily be extended to random vectors $[X_1 \cdots X_N]^T \sim p$ where X_i takes values in the alphabet $\mathcal{X}_i$ by

$$H(X_1, \ldots, X_N) := -\sum_{x_N \in \mathcal{X}_N} \cdots \sum_{x_1 \in \mathcal{X}_1} p(x_1, \ldots, x_N) \log(p(x_1, \ldots, x_N))$$

Notice the entropy is a function of the distribution of X only, and not its values, why it is often denoted $H(p)$. The entropy quantifies the amount of information required to describe a random variable and thus the amount of "randomness" contained in it. Let $\mathcal{X} = \{x_0, x_1, \ldots, x_N\}$. Let $p(x) = 0$ for $x \in \mathcal{X} \backslash \{x_0\}$ and $p(x_0) = 1$. The random variable

with such a distribution will be deterministic and have entropy $H(X) = 0$. Conversely if $p(x) = \frac{1}{N+1}$, meaning the corresponding random variable is uniformly distributed, and thus as "random" as possible, its entropy will be

$$H(X) = -\sum_{i=1}^{N+1} \frac{1}{N+1} \log\left(\frac{1}{N+1}\right) = \log(N+1).$$

This can be proven to be the maximum entropy achievable for a discrete random variable with alphabet size $N+1$, intuitively justifying it as a quantification of randomness [Cover and Thomas, 2006, pp.14-16].
Similar to classical alphabets and their associated probability distributions, quantum systems are described by a density operator, which is in a superposition of multiple quantum states. According to postulate 2.9, density operators are described by a set of quantum states with an associated probability distribution, paving the way for a new definition of entropy.

Definition 3.1: Von Neumann Entropy
Let $\mathcal{A}$ be a quantum system with state space $\mathcal{H}_\mathcal{A}$ and associated density operator $\rho : \mathcal{H}_\mathcal{A} \to \mathcal{H}_\mathcal{A}$. The Von Neumann entropy of the density operator ρ is defined by

$$\mathcal{S}(\rho) := -\operatorname{tr}(\rho \log(\rho)),$$

where log is the matrix logarithm.

[Nielsen and Chuang, 2010, p.510]

Let the density operator $\rho : \mathcal{H}_\mathcal{A} \to \mathcal{H}_\mathcal{A}$ with $\dim(\mathcal{H}_\mathcal{A}) = N$ have spectral decomposition $\rho = \sum_{i=1}^N \lambda_i |\psi_i\rangle \langle\psi_i|$ where $\{|\psi\rangle\}_{i=1}^N$ are eigenvectors of ρ forming an orthonormal basis of $\mathcal{H}_\mathcal{A}$ and $\{\lambda_i\}_{i=1}^N$ are the corresponding eigenvalues, then

$$\begin{aligned}
\mathcal{S}(\rho) &= -\operatorname{tr}\left(\sum_{i=1}^N \lambda_i |\psi_i\rangle \langle\psi_i| \log\left(\sum_{j=1}^N \lambda_j |\psi_j\rangle \langle\psi_j|\right)\right) \\
&= -\operatorname{tr}\left(\sum_{i=1}^N \lambda_i |\psi_i\rangle \langle\psi_i| \sum_{j=1}^N \log(\lambda_j) |\psi_j\rangle \langle\psi_j|\right) \\
&= -\operatorname{tr}\left(\sum_{i=1}^N \sum_{j=1}^N \lambda_i \log(\lambda_j) |\psi_i\rangle \langle\psi_i|\psi_j\rangle \langle\psi_j|\right) \\
&= -\operatorname{tr}\left(\sum_{i=1}^N \lambda_i \log(\lambda_i) |\psi_i\rangle \langle\psi_i|\right) \\
&= -\sum_{i=1}^N \lambda_i \log(\lambda_i),
\end{aligned} \tag{3.2}$$

where the second equality is a property of taking the matrix logarithm of a diagonalizable matrix. For $\lambda_i = 0$ the convention $0 \log(0) := 0$ is used. The entropy of a density operator is therefore completely determined by the its eigenvalues [Nielsen and Chuang, 2010, p.510].

The Von Neumann entropy has a lot of useful properties, some of which are expressed in the following theorem.

Theorem 3.2: Properties of Von Neumann Entropy

1. $\mathcal{S}(\rho) \geq 0$ with equality if and only if ρ is a pure state.
2. Let $\dim(\mathcal{H}_\mathcal{A}) = N$, then $\mathcal{S}(\rho) \leq \log(N)$ for $\rho : \mathcal{H}_\mathcal{A} \to \mathcal{H}_\mathcal{A}$ with equality if and only if $\rho = \frac{1}{N} I$.
3. Let $\rho^{\mathcal{A}_1} : \mathcal{H}_{\mathcal{A}_1} \to \mathcal{H}_{\mathcal{A}_1}$ and $\rho^{\mathcal{A}_2} : \mathcal{H}_{\mathcal{A}_2} \to \mathcal{H}_{\mathcal{A}_2}$, then
$$\mathcal{S}(\rho^{\mathcal{A}_1} \otimes \rho^{\mathcal{A}_2}) = \mathcal{S}(\rho^{\mathcal{A}_1}) + \mathcal{S}(\rho^{\mathcal{A}_2}).$$
4. Let $\rho^{\mathcal{A}_1\mathcal{A}_2} : \mathcal{H}_{\mathcal{A}_1\mathcal{A}_2} \to \mathcal{H}_{\mathcal{A}_1\mathcal{A}_2}$ be a pure state, then $\mathcal{S}\left(\rho^{\mathcal{A}_1}\right) = \mathcal{S}\left(\rho^{\mathcal{A}_2}\right)$.
5. Let $\{p_i\}_{i=1}^N$ be such that $p_i \geq 0$ and $\sum_{i=1}^N p_i = 1$. Let $\{\rho_i\}_{i=1}^N$ be a set of density operators with support on orthogonal subspaces. Then
$$\mathcal{S}\left(\sum_{i=1}^N p_i \rho_i\right) = H(p) + \sum_{i=1}^N p_i \mathcal{S}(\rho_i),$$
where $H(p) = -\sum_{i=1}^N p_i \log(p_i)$.
6. Let $\{p_i\}_{i=1}^N$ be such that $p_i \geq 0$ and $\sum_{i=1}^N p_i = 1$. Let $\{|\psi_i\rangle \in \mathcal{H}_{\mathcal{A}_1}\}_{i=1}^N$ be orthogonal states and $\{\rho_i : \mathcal{H}_{\mathcal{A}_2} \to \mathcal{H}_{\mathcal{A}_2}\}_{i=1}^N$. Then
$$\mathcal{S}\left(\sum_{i=1}^N p_i |\psi_i\rangle\langle\psi_i| \otimes \rho_i\right) = H(p) + \sum_{i=1}^N p_i \mathcal{S}(\rho_i),$$
where $H(p) = -\sum_{i=1}^N p_i \log(p_i)$.

[Nielsen and Chuang, 2010, p.513]

Proof

(1)

According to postulate 2.9, the density operator is positive semidefinite and has trace 1. Let $\{\lambda_i\}_{i=1}^N$ be the eigenvalues of the density operator, then these properties together imply $0 \leq \lambda_i \leq 1$ for all $i \in \{1, \ldots, N\}$. The statement then follows from (3.2).

(2)

The proof requires introduction of additional theorems, see [Nielsen and Chuang, 2010, p.513].

(3)

Let $\{\lambda_{1i}\}_{i=1}^N$ and $\{\lambda_{2j}\}_{j=1}^M$ be the eigenvalues of $\rho^{\mathcal{A}_1}$ and $\rho^{\mathcal{A}_2}$, respectively, with corresponding eigenvectors $\{|a_{1i}\rangle\}_{i=1}^N$ and $\{|a_{2j}\rangle\}_{j=1}^M$, then

$$(\rho^{\mathcal{A}_1} \otimes \rho^{\mathcal{A}_2})(|a_{1i}\rangle \otimes |a_{2j}\rangle) = \lambda_{1i}\lambda_{2j}(|a_{1i}\rangle \otimes |a_{2j}\rangle)$$

$\{\lambda_{1i}\lambda_{2j}\}_{1\leq i\leq N,\ 1\leq j\leq M}$ are thus eigenvalues of $\rho^{\mathcal{A}_1}\otimes\rho^{\mathcal{A}_2}$. By (3.2) it follows

$$\begin{aligned}\mathcal{S}(\rho^{\mathcal{A}_1}\otimes\rho^{\mathcal{A}_2}) &= -\sum_{i=1}^{N}\sum_{j=1}^{M}\lambda_{1i}\lambda_{2j}\log(\lambda_{1i}\lambda_{2j})\\ &= -\sum_{i=1}^{N}\lambda_{1i}\log(\lambda_{1i})\sum_{j=1}^{M}\lambda_{2j} - \sum_{j=1}^{M}\lambda_{2j}\log(\lambda_{2j})\sum_{i=1}^{N}\lambda_{1i}\\ &= -\sum_{i=1}^{N}\lambda_{1i}\log(\lambda_{1i}) - \sum_{j=1}^{M}\lambda_{2j}\log(\lambda_{2j})\\ &= \mathcal{S}(\rho^{\mathcal{A}_1})+\mathcal{S}(\rho^{\mathcal{A}_2}).\end{aligned}$$

(4)

According to (2.7) and (2.8) the reduced density operators of a composite system in a pure state will have identical eigenvalues. Since the entropy of a density operator is completely determined by the eigenvalues, the statement follows.

(5)

Let $\{\lambda_{ij}\}_{j=1}^{M}$ be the eigenvalues of ρ_i with corresponding eigenvectors $\{|e_{ij}\rangle\}_{j=1}^{M}$. Assuming the set $\{\rho_i\}_{i=1}^{N}$ have support on orthogonal subspaces, it follows that $\{p_i\lambda_{ij}\}_{1\leq i\leq N,\ 1\leq j\leq M}$ are the eigenvalues of $\sum_{i=1}^{N}p_i\rho_i$ with corresponding eigenvectors $\{|e_{ij}\rangle\}_{1\leq i\leq N,\ 1\leq j\leq M}$. According to (3.2)

$$\begin{aligned}\mathcal{S}\left(\sum_{i=1}^{N}p_i\rho_i\right) &= -\sum_{i=1}^{N}\sum_{j=1}^{M}p_i\lambda_{ij}\log(p_i\lambda_{ij})\\ &= -\sum_{i=1}^{N}p_i\log(p_i)\sum_{j=1}^{M}\lambda_{ij} - \sum_{i=1}^{N}p_i\sum_{j=1}^{M}\lambda_{ij}\log(\lambda_{ij})\\ &= H(p)+\sum_{i=1}^{N}p_i\mathcal{S}(\rho_i)\end{aligned}$$

(6)

Consider $\rho_i^{\mathcal{A}_1\mathcal{A}_2} := |\psi_i\rangle\langle\psi_i|\otimes\rho_i$, then $\rho_i^{\mathcal{A}_1\mathcal{A}_2}$ is a density operator of the composite system $\mathcal{A}_1\mathcal{A}_2$ according to postulate 2.13. Let $\{\lambda_{ij}\}_{j=1}^{M}$ be the eigenvalues of ρ_i and $\{|e_{ij}\rangle\}_{j=1}^{M}$ the corresponding eigenvectors. Then $\{\lambda_{ij}\}_{j=1}^{M}$ are also the eigenvalues of $\rho_i^{\mathcal{A}_1\mathcal{A}_2}$ with eigenvectors $\{|\psi_i\rangle\otimes|e_{ij}\rangle\}_{j=1}^{M}$, since

$$(|\psi_i\rangle\langle\psi_i|\otimes\rho_i)(|\psi_i\rangle\otimes|e_{ij}\rangle) = |\psi_i\rangle\langle\psi_i|\psi_i\rangle\otimes\rho_i|e_{ij}\rangle = \lambda_{ij}(|\psi_i\rangle\otimes|e_{ij}\rangle).$$

Furthermore, these eigenvectors are orthogonal since

$$(\langle\psi_i|\otimes\langle e_{ij}|)(|\psi_{i'}\rangle\otimes|e_{i'j}\rangle) = \langle\psi_i|\psi_{i'}\rangle\langle e_{ij}|e_{i'j}\rangle = \begin{cases}= \||\psi_i\rangle\|^2\||e_{ij}\rangle\|^2, & \text{if } i'=i,\\ = 0, & \text{else.}\end{cases}$$

The operators $\{\rho_i^{\mathcal{A}_1\mathcal{A}_2}\}_{i=1}^{N}$ thus satisfy the conditions of statement 4, implying

$$\mathcal{S}\left(\sum_{i=1}^{N}p_i\rho^{\mathcal{A}_1\mathcal{A}_2}\right) = H(p)+\sum_{i=1}^{N}p_i\mathcal{S}(\rho_i^{\mathcal{A}_1\mathcal{A}_2})$$

According to statement 1 and 3, it follows

$$\mathcal{S}(\rho_i^{\mathcal{A}_1\mathcal{A}_2}) = \mathcal{S}(|\psi_i\rangle\langle\psi_i|) + \mathcal{S}(\rho_i) = \mathcal{S}(\rho_i).$$ ∎

Let $\mathcal{A}_1$ and $\mathcal{A}_2$ with state spaces $\mathcal{H}_{\mathcal{A}_1}$ and $\mathcal{H}_{\mathcal{A}_2}$ respectively. Suppose the composite quantum system $\mathcal{A}_1\mathcal{A}_2$ with state space $\mathcal{H}_{\mathcal{A}_1} \otimes \mathcal{H}_{\mathcal{A}_2}$ has density operator $\rho^{\mathcal{A}_1\mathcal{A}_2} := \rho^{\mathcal{A}_1} \otimes \rho^{\mathcal{A}_2}$, the joint entropy is the entropy of the density operator of the composite system

$$\mathcal{S}\left(\rho^{\mathcal{A}_1\mathcal{A}_2}\right) := -\operatorname{tr}\left(\rho^{\mathcal{A}_1\mathcal{A}_2}\log\left(\rho^{\mathcal{A}_1\mathcal{A}_2}\right)\right).$$

With the joint entropy, conditional entropy and mutual information can be defined.

Definition 3.3: Conditional Entropy and Mutual Information
Let Let $\rho^{\mathcal{A}_1} : \mathcal{H}_{\mathcal{A}_1} \to \mathcal{H}_{\mathcal{A}_1}$, $\rho^{\mathcal{A}_2} : \mathcal{H}_{\mathcal{A}_2} \to \mathcal{H}_{\mathcal{A}_2}$ and $\rho^{\mathcal{A}_1\mathcal{A}_2} : \mathcal{H}_{\mathcal{A}_1} \otimes \mathcal{H}_{\mathcal{A}_2} \to \mathcal{H}_{\mathcal{A}_1} \otimes \mathcal{H}_{\mathcal{A}_2}$, the conditional entropy of $\rho^{\mathcal{A}_1}$ given $\rho^{\mathcal{A}_2}$ is defined by

$$\mathcal{S}\left(\rho^{\mathcal{A}_1}|\rho^{\mathcal{A}_2}\right) := \mathcal{S}\left(\rho^{\mathcal{A}_1\mathcal{A}_2}\right) - \mathcal{S}\left(\rho^{\mathcal{A}_2}\right). \tag{3.3}$$

The mutual information between $\rho^{\mathcal{A}_1}$ and $\rho^{\mathcal{A}_2}$ is defined by

$$\mathcal{S}\left(\rho^{\mathcal{A}_1} : \rho^{\mathcal{A}_2}\right) := \mathcal{S}\left(\rho^{\mathcal{A}_1}\right) + \mathcal{S}\left(\rho^{\mathcal{A}_2}\right) - \mathcal{S}\left(\rho^{\mathcal{A}_1\mathcal{A}_2}\right).$$

[Nielsen and Chuang, 2010, p.514]

Combining the two definitions above, the mutual information can also be expressed by

$$\mathcal{S}\left(\rho^{\mathcal{A}_1} : \rho^{\mathcal{A}_2}\right) = \mathcal{S}\left(\rho^{\mathcal{A}_1}\right) - \mathcal{S}\left(\rho^{\mathcal{A}_1} : \rho^{\mathcal{A}_2}\right) = \mathcal{S}\left(\rho^{\mathcal{A}_2}\right) - \mathcal{S}\left(\rho^{\mathcal{A}_2}|\rho^{\mathcal{A}_1}\right)$$

It follows from the definition of Shannon's entropy that $H(X_1) \leq H(X_1, X_2)$. Given the entropy quantifies the uncertainty in a random variable, this seems reasonable; a set of random variable will always be at least as uncertain, as any single random variable in the set. In the quantum world, the corresponding statement would be $\mathcal{S}\left(\rho^{\mathcal{A}_1}\right) \leq \mathcal{S}\left(\rho^{\mathcal{A}_1\mathcal{A}_2}\right)$. According to (3.3) that is equivalent to $\mathcal{S}\left(\rho^{\mathcal{A}_2}|\rho^{\mathcal{A}_1}\right) \geq 0$. The next theorem will show, that this is not always the case.

Theorem 3.4: Conditional Entropy of Entangled States
Let $\rho^{\mathcal{A}_1\mathcal{A}_2} : \mathcal{H}_{\mathcal{A}_1\mathcal{A}_2} \to \mathcal{H}_{\mathcal{A}_1\mathcal{A}_2}$ be a pure state. Then $\rho^{\mathcal{A}_1\mathcal{A}_2}$ is entangled if and only if $\mathcal{S}\left(\rho^{\mathcal{A}_2}|\rho^{\mathcal{A}_1}\right) < 0$ for $\rho^{\mathcal{A}_2} : \mathcal{H}_{\mathcal{A}_2} \to \mathcal{H}_{\mathcal{A}_2}$ and $\rho^{\mathcal{A}_1} : \mathcal{H}_{\mathcal{A}_1} \to \mathcal{H}_{\mathcal{A}_1}$.

[Nielsen and Chuang, 2010, p.514]

Proof
Assume $\rho^{\mathcal{A}_1\mathcal{A}_2}$ is pure, then according to statement 1 of theorem 3.2 it follows

$$\mathcal{S}\left(\rho^{\mathcal{A}_2}|\rho^{\mathcal{A}_1}\right) := \mathcal{S}\left(\rho^{\mathcal{A}_1\mathcal{A}_2}\right) - \mathcal{S}\left(\rho^{\mathcal{A}_1}\right) = -\mathcal{S}\left(\rho^{\mathcal{A}_1}\right)$$

The conditional entropy is therefore strictly negative if and only if $\rho^{\mathcal{A}_1}$ is in a mixed state. Since $\rho^{\mathcal{A}_1\mathcal{A}_2}$ is pure, it was shown in a corollary to the Schmidt decomposition $\rho^{\mathcal{A}_1}$ is mixed if and only if $\rho^{\mathcal{A}_1\mathcal{A}_2}$ is entangled. ∎

The last theorem introduced pertains to the subadditivity property of the Von Neumann entropy; an important property when examining the quantum singleton bound. The quantum singleton bound will be described more thoroughly in the following chapters on quantum error correction.

Theorem 3.5: Subadditivity of Von Neumann Entropy
Let the composite system $\mathcal{A}_1\mathcal{A}_2$ be in state $\rho^{\mathcal{A}_1\mathcal{A}_2}$, then

$$\mathcal{S}(\rho^{\mathcal{A}_1\mathcal{A}_2}) \leq \mathcal{S}(\rho^{\mathcal{A}_1}) + \mathcal{S}(\rho^{\mathcal{A}_2})$$

[Nielsen and Chuang, 2010, pp.515-516]

Proof
The proof follows directly from Klein's inequality (see [Nielsen and Chuang, 2010, p.511]) and properties of the matrix logarithm.

$$\begin{aligned}
\mathcal{S}(\rho^{\mathcal{A}_1\mathcal{A}_2}) &\leq -\operatorname{tr}\left(\rho^{\mathcal{A}_1\mathcal{A}_2}\log\left(\rho^{\mathcal{A}_1}\otimes\rho^{\mathcal{A}_2}\right)\right)\\
&= -\operatorname{tr}\left(\rho^{\mathcal{A}_1\mathcal{A}_2}\left(\log\left(\rho^{\mathcal{A}_1}\right)\otimes I + I\otimes\log\left(\rho^{\mathcal{A}_2}\right)\right)\right)\\
&= -\operatorname{tr}\left(\rho^{\mathcal{A}_1\mathcal{A}_2}\left(\log\left(\rho^{\mathcal{A}_1}\right)\otimes I\right)\right) - \operatorname{tr}\left(\rho^{\mathcal{A}_1\mathcal{A}_2}\left(I\otimes\log\left(\rho^{\mathcal{A}_2}\right)\right)\right)\\
&= -\operatorname{tr}\left(\rho^{\mathcal{A}_1}\log\left(\rho^{\mathcal{A}_1}\right)\right) - \operatorname{tr}\left(\rho^{\mathcal{A}_2}\log\left(\rho^{\mathcal{A}_2}\right)\right)\\
&= \mathcal{S}(\rho^{\mathcal{A}_1}) + \mathcal{S}(\rho^{\mathcal{A}_2})
\end{aligned}$$

The inequality is an equality if and only if $\rho^{\mathcal{A}_1\mathcal{A}_2} = \rho^{\mathcal{A}_1}\otimes\rho^{\mathcal{A}_2}$, cf. Klein's inequality. ∎

As previously mentioned, present quantum computers are noisy and inter-mediate scale (NISQ), this implies a necessity of exchanging quantum information between quantum computers if more qubits than can be provided by a single quantum computer is needed. A theoretical framework describing the action of noise on quantum states within and between quantum computers are thus vital. This framework is provided by the notion of quantum channels.

3.2 Quantum Channels

The postulates of chapter 2 concerned closed quantum systems. In practice no systems (besides the universe as a whole) can be regarded completely closed. Small interactions between the quantum system of interest (called the principal system) and its environment will appear as noise. This section gives a mathematical foundation of quantum channels and how noise is modelled in such channels.

A quantum channel, $\mathcal{E}$, relates two quantum states by some transformation.

$$\rho' = \mathcal{E}(\rho).$$

The transformation depends on the physical process the channel describes. Whether the physical process is a measurement or a time evolution, the channel behaves linearly and

are given as $\rho' = A\rho A^\dagger$ for some matrix A (possibly requiring some normalisation given by $\mathrm{tr}(A\rho A^\dagger)$). Assuming the principal-environment system is in a product state $\rho \otimes \rho_{\mathrm{env}}$. The state of the principal system after the channel can then be found by tracing out the environment

$$\rho' = \mathrm{tr}_{\mathrm{env}}(U(\rho \otimes \rho_{\mathrm{env}})U^\dagger).$$

This setup is depicted in figure 3.1.

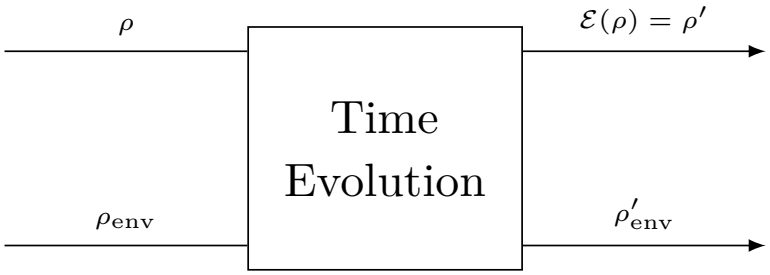


Figure 3.1: Illustration of the time evolution of a quantum system.

Let $\{|\psi_i\rangle\}_{i=1}^N$ be an orthonormal basis for the state space of the environment, then

$$\begin{aligned}
\rho' &= \mathrm{tr}_{\mathrm{env}}\left(U\left(\rho \otimes \left(\sum_{i=1}^N \alpha_i |\psi_i\rangle\langle\psi_i|\right)\right)U^\dagger\right) \\
&= \sum_{i=1}^N \alpha_i \,\mathrm{tr}_{\mathrm{env}}\left(U(\rho \otimes |\psi_i\rangle\langle\psi_i|)U^\dagger\right) \\
&= \sum_{i=1}^N \alpha_i \left(\sum_{j=1}^N (I \otimes \langle\psi_j|)U(\rho \otimes |\psi_i\rangle\langle\psi_i|)U^\dagger(I \otimes |\psi_j\rangle)\right) \\
&= \sum_{i=1}^N \sum_{j=1}^N \alpha_i (I \otimes \langle\psi_j|)U(I \otimes |\psi_i\rangle)\rho(I \otimes \langle\psi_i|)U^\dagger(I \otimes |\psi_j\rangle) \\
&= \sum_{i=1}^N \sum_{j=1}^N E_{ij}\rho E_{ij}^\dagger = \sum_{k=1}^{N^2} E_k \rho E_k^\dagger
\end{aligned}$$

where $E_{ij} := \sqrt{\alpha_i}(I \otimes \langle\psi_j|)U(I \otimes |\psi_i\rangle)$ are called Kraus operators. This way of describing a quantum channel is called the operator-sum representation. To ensure positive probabilities under measurements, the output, ρ', must be positive semidefnite, however, this is always the case since

$$\langle\psi|\,\rho'\,|\psi\rangle = \sum_{k=1}^{N^2} \langle\psi|\,E_k\rho E_k^\dagger\,|\psi\rangle = \sum_{k=1}^{N^2} \langle\psi_k|\,\rho\,|\psi_k\rangle \geq 0, \qquad |\psi_k\rangle := E_k^\dagger\,|\psi\rangle\,.$$

Moreover, the trace of a density operator corresponds to the probability of being in any possible state and must therefore satisfy $\mathrm{tr}(\rho') \leq 1$, thus

$$\mathrm{tr}(\rho') = \mathrm{tr}\left(\rho \sum_{k=1}^{N^2} E_k^\dagger E_k\right) \leq 1 = \mathrm{tr}(\rho) \implies 0 \leq \mathrm{tr}\left(\rho\left(I - \sum_{k=1}^{N^2} E_k^\dagger E_k\right)\right).$$

This implies the operator $I - \sum_{k=1}^{N^2} E_k^\dagger E_k$ must be positive semidefinite. If $\sum_{k=1}^{N^2} E_k^\dagger E_k = I$ the channel is said to be trace-preserving since

$$\mathrm{tr}(\mathcal{E}(\rho)) = \mathrm{tr}\left(\sum_{k=1}^{N^2} E_k \rho E_k^\dagger\right) = \mathrm{tr}\left(\left(\sum_{k=1}^{N^2} E_k^\dagger E_k\right)\rho\right) = \mathrm{tr}(\rho).$$

Such channels are also called completely positive trace-preserving (CPTP) since they map a positive semidefinite operator to a positive semidefinite operator and preserves the trace. Channels for which $\sum_{k=1}^{N^2} E_k^\dagger E_k \neq I$ is called non-trace-preserving (or trace-decreasing) channels. They arise when conditioning on the output of the channel. An example is the channel $\mathcal{E}(\rho) = p\rho$ for some $0 < p < 1$. This channel is obviously non-trace-preserving and can be implemented by discarding the output if it is not equal to the input. M uses of this channel would thus result in approximately Mp outputs yielding ρ and $(1-p)M$ outputs being discarded [Nielsen and Chuang, 2010, pp.357-360].

Similar to the fact that different sets of states, motivates the same density operator, one might wonder if two different sets of Kraus operators could correspond to the same quantum channel. The next theorem provides the condition for when this is the case.

Theorem 3.6: Uniqueness of Quantum Channels
Let $\mathcal{E}$ and $\mathcal{F}$ be two quantum channels with Kraus operators $\{E_i\}_{i=1}^N$ and $\{F_j\}_{j=1}^M$ respectively. Assume $M \leq N$ and extend $\{F_j\}_{j=1}^N$ such $F_j = 0$ for $M < j \leq N$. Then $\mathcal{E} = \mathcal{F}$ if and only if $E_i = \sum_{j=1}^N u_{ij} F_j$ where u_{ij} are the entries of a unitary matrix.

[Nielsen and Chuang, 2010, pp.372-373]

Proof
Assume $\mathcal{E}$ and $\mathcal{F}$ are related by a unitary transformation, then

$$\begin{aligned}
\mathcal{E}(\rho) &= \sum_{i=1}^N E_i \rho E_i^\dagger \\
&= \sum_{i=1}^N \left(\sum_{j=1}^N u_{ij} F_j\right) \rho \left(\sum_{j'=1}^N u_{ij'}^* F_{j'}^\dagger\right) \\
&= \sum_{j=1}^N \sum_{j'=1}^N \sum_{i=1}^N u_{ij} u_{j'i} F_j \rho F_{j'}^\dagger \\
&= \sum_{j=1}^N F_j \rho F_j^\dagger = \mathcal{F}(\rho).
\end{aligned}$$

Conversely, assume $\mathcal{E} = \mathcal{F}$, that is, $\sum_{i=1}^N E_i \rho E_i^\dagger = \sum_{j=1}^N F_j \rho F_j^\dagger$ for any density operator ρ on the state space of the quantum channel, $\mathcal{H}$, where $F_j = 0$ for $M < j \leq N$. Let $\{|\alpha_i\rangle\}_i$ be an orthonormal basis for $\mathcal{H}$. Define

$$|\psi\rangle := \sum_{k=1} |\alpha_k\rangle \otimes |\alpha_k\rangle$$

and thereby

$$|e_i\rangle := (I \otimes E_i)|\psi\rangle = \sum_k |\alpha_k\rangle \otimes E_i |\alpha_k\rangle$$

$$|f_j\rangle := (I \otimes F_j)\,|\psi\rangle = \sum_k |\alpha_k\rangle \otimes F_j\,|\alpha_k\rangle$$

Introduce density operator

$$\begin{aligned}\sigma := (I \otimes \mathcal{E})(|\psi\rangle\langle\psi|) &= \sum_{k,k'} |\alpha_k\rangle\langle\alpha_{k'}| \otimes \mathcal{E}(|\alpha_k\rangle\langle\alpha_{k'}|) \\ &= \sum_{k,k'} |\alpha_k\rangle\langle\alpha_{k'}| \otimes \mathcal{F}(|\alpha_k\rangle\langle\alpha_{k'}|) = (I \otimes \mathcal{F})(|\psi\rangle\langle\psi|)\end{aligned}$$

and notice

$$\begin{aligned}\sum_{i=1}^{N} |e_i\rangle\langle e_i| &= \sum_{i=1}^{N}\sum_{k,k'} |\alpha_k\rangle\langle\alpha_{k'}| \otimes E_i(|\alpha_k\rangle\langle\alpha_{k'}|)E_i^\dagger \\ &= \sigma \\ &= \sum_{j=1}^{N}\sum_{k,k'} |\alpha_k\rangle\langle\alpha_{k'}| \otimes F_j(|\alpha_k\rangle\langle\alpha_{k'}|)F_j^\dagger = \sum_{j=1}^{N} |f_j\rangle\langle f_j|\end{aligned}$$

According to theorem 2.12 it then follows

$$|e_i\rangle = \sum_{j=1}^{N} u_{ij}\,|f_j\rangle$$

where u_{ij} are the entries of some unitary matrix. Denote by $|\phi\rangle = \sum_k \gamma_k\,|\alpha_k\rangle$ an arbitrary vector of $\mathcal{H}$ and define $|\tilde{\phi}\rangle := \sum_k \gamma_k^*\,|\alpha_k\rangle$ then

$$\begin{aligned}E_i\,|\phi\rangle &= (\langle\tilde{\phi}| \otimes I)\,|e_i\rangle \\ &= \sum_{j=1}^{N} u_{ij}(\langle\tilde{\phi}| \otimes I)\,|f_j\rangle \\ &= \sum_{j=1}^{N} u_{ij}F_j\,|\phi\rangle\,.\end{aligned}$$

By the arbitrariness of $|\phi\rangle$ it can be concluded that

$$E_i = \sum_{j=1}^{N} u_{ij}F_j. \qquad \blacksquare$$

The study of quantum channels conclude with an example of a specific quantum channel; the depolarizing channel.

Example 4: The Depolarizing Channel
Let the Kraus operators for a quantum channel be given by $\left\{\sqrt{1-\frac{3p}{4}}I, \frac{\sqrt{p}}{2}X, \frac{\sqrt{p}}{2}Y, \frac{\sqrt{p}}{2}Z\right\}$ for some parameter $0 \leq p \leq 1$. This channel can be verified to be trace-preserving and

$$\mathcal{E}(\rho) = \left(1 - \frac{3p}{4}\right)\rho + \frac{p}{4}X\rho X + \frac{p}{4}Y\rho Y + \frac{p}{4}Z\rho Z \tag{3.4}$$

$$= (1-p)\rho + \frac{p}{4}(\rho + X\rho X + Y\rho Y + Z\rho Z). \tag{3.5}$$

Simple calculations show that

$$\rho + X\rho X + Y\rho Y + Z\rho Z = 2I.$$

Substituting into (3.4) yields

$$\rho' = (1-p)\rho + \frac{p}{2}I.$$

With probability p the channel changes the input state into the completely mixed state $\frac{1}{2}I$ while leaving the input state unchanged with probability p [Nielsen and Chuang, 2010, pp.378-379].

With the foundational concepts of quantum information theory, including entropy, conditional entropy, mutual information, and properties thereof, along with the formalism of quantum channels, rigorously defined and established, the study of quantum error correction can now commence.

4 | Quantum Error-Correcting Codes

Error-correcting codes are a vital tool when transmitting information, classical or quantum, through channels subject to noise. When applying error-correcting codes for such scenarios, the probability of faulty transmission can be drastically reduced; a desirable objective.

This chapter is primarily based on [Nielsen and Chuang, 2010], [Gundersen et al., 2025] and [Bergou et al., 2021].

4.1 Classical Error Correction

Classical computers exchange information between each other using classical channels. A discrete, memoryless classical channel is defined by the input alphabet, the output alphabet and a transition matrix, containing the conditional probability of observing some output in the alphabet given some input in the input alphabet. The most simple of such is the binary symmetric channel, in which a bit is sent through the channel and flipped with probability p. The channel is shown in figure 4.1.

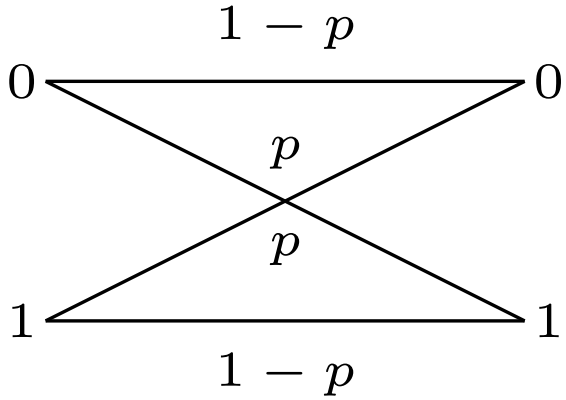


Figure 4.1: The binary symmetric channel

Suppose Alice whishes to send Bob 1 bit of information through the binary symmetric channel. Alice could simply transmit the bit she wishes to send, and accept it is only received successfully with probability $\frac{1}{2} \leq 1 - p \leq 1$. Another strategy is to transmit the bit N times through the channel and tell Bob to choose whichever bit appears the most times as the bit Alice whished to transmit. This strategy is called "majority voting" and

for $N = 3$ the probability of successful transmission is the probability that no more than 1 bit is flipped of the 3 transmitted. This probability is exactly

$$p_s = (1-p)^3 + 3p(1-p)^2 = 1 - 3p^2 + 2p^3$$

implying the probability of error is given by

$$p_e = 1 - p_s = 3p^2 - 2p^3.$$

Sending just 1 bit through the channel would fail with probability p, while sending 3 bits will fail with probability $3p^2 - 2p^3$ which is less than p for $p < \frac{1}{2}$. In the case $p = \frac{1}{2}$ no information can be sent through the channel, as it has capacity 0, making any coding scheme equal [Cover and Thomas, 2006, pp.183-184,187].

Though a simple example of an error-correcting code, majority voting shows the possibility of lowering the probability of error using the channel more times than strictly necessary. This is the foundation of error-correcting codes; add redundant information, such the received signal is less prone to errors. This idea is useful even in the quantum realm, however, some caveats needs to be addressed. Firstly, the errors in the classical world consists of bits being flipped, and are hence discrete. In the quantum realm, a qubit's state is not limited to discrete flips; its phase can vary continuously, forming a continuum of possible states and hence possible errors. Secondly, in the error-correction strategy used above, Bob decided which bit Alice wished to send, based on the output of the channel. However, in the quantum realm this would require a measurement, thus collapsing the quantum state. Thirdly, the repetition strategy in which Alice creates $N-1$ additional copies of the state to be transmitted and sends all N copies through the channel is prohibited in the quantum realm due to the no-cloning theorem: Let $|\psi\rangle \in \mathcal{H}_{\mathcal{A}_1}$ be a qubit one wishes to clone and $|\gamma\rangle \in \mathcal{H}_{\mathcal{A}_2}$ the qubit one wishes to clone $|\psi\rangle$ onto. Since quantum systems evolve unitarily, the existence of such cloning would imply the existence of a unitary U such

$$U(|\psi\rangle \otimes |\gamma\rangle) = |\psi\rangle \otimes |\psi\rangle .$$

The no-cloning theorem shows that no such U can exist.

Theorem 4.1: The No-Cloning Theorem
Let $|\psi\rangle \in \mathcal{H}_{\mathcal{A}_1}$ and $|\gamma\rangle \in \mathcal{H}_{\mathcal{A}_2}$. No unitary operator U exists such that

$$U(|\psi\rangle \otimes |\gamma\rangle) = |\psi\rangle \otimes |\psi\rangle .$$

[Nielsen and Chuang, 2010, p.532]

Proof
Suppose $|\psi\rangle \in \mathcal{H}_{\mathcal{A}_1}$ and $|\phi\rangle \in \mathcal{H}_{\mathcal{A}_1}$ are two arbitrary states, for which a unitary operator U exists such that

$$U(|\psi\rangle \otimes |\gamma\rangle) = |\psi\rangle \otimes |\psi\rangle$$
$$U(|\phi\rangle \otimes |\gamma\rangle) = |\phi\rangle \otimes |\phi\rangle .$$

Taking the inner product of the two equations yield

$$\langle\psi|\phi\rangle^2 = (\langle\psi| \otimes \langle\psi|)(|\phi\rangle \otimes |\phi\rangle) = (\langle\psi| \otimes \langle\gamma|)U^\dagger U(|\phi\rangle \otimes |\gamma\rangle) = \langle\psi|\phi\rangle\,.$$

This implies $\langle\psi|\phi\rangle = 0 \vee 1$, meaning the states $|\psi\rangle$ and $|\phi\rangle$ are orthogonal or identical, a contradiction with the fact that they were chosen arbitrarily. ∎

The three downsides to how the quantum realm operates, however, does not mean no error-correcting codes exists for quantum channels.

4.2 Quantum Error Correction

To show that quantum error-correcting codes exists despite the caveats, consider the following example: A qubit is given by

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle\,.$$

The qubit is then encoded into a larger state space by

$$\begin{aligned} |0\rangle &\xrightarrow{\text{Encode}} |0\rangle_L := |000\rangle \\ |1\rangle &\xrightarrow{\text{Encode}} |1\rangle_L := |111\rangle\,. \end{aligned}$$

The notation emphasizes the fact that one logical qubit is encoded in three physical qubits. The encoded qubits are then of the form

$$|\psi\rangle_{\text{Enc}} := \alpha\,|0\rangle_L + \beta\,|1\rangle_L\,. \tag{4.1}$$

The codewords in this scheme are thus vectors in the subspace spanned by $|0\rangle_L$ and $|1\rangle_L$. Such an encoding can be achieved by the circuit depicted in figure 4.2, using two CNOT gates and five Hadamard gates.

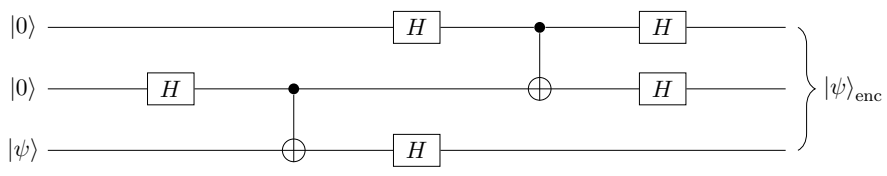


Figure 4.2: The quantum circuit encoding one logical qubit in three physical qubits.

This encoding allows for the correction of one bit flip on any of the three physical qubits by performing a set of orthogonal projective measurements. Consider table 4.1 outlining the possible errors and an associated projector.

Bit flip position	Projector	Error
No bit flip	$P_0 = \|000\rangle\langle 000\| + \|111\rangle\langle 111\|$	$E_0 = I \otimes I \otimes I$
First qubit	$P_1 = \|100\rangle\langle 100\| + \|011\rangle\langle 011\|$	$E_1 = X \otimes I \otimes I$
Second qubit	$P_2 = \|010\rangle\langle 010\| + \|101\rangle\langle 101\|$	$E_2 = I \otimes X \otimes I$
Third qubit	$P_3 = \|001\rangle\langle 001\| + \|110\rangle\langle 110\|$	$E_3 = I \otimes I \otimes X$

Table 4.1: The possible errors and their associated projectors.

Notice that the sum of the projectors equal the identity making them a valid set of measurement operators. Let error $k \in \{0, 1, 2, 3\}$ occur and define

$$E_k |\psi\rangle_{\text{Enc}} = \alpha |\phi_1\rangle + \beta |\phi_2\rangle$$

Then with probability 1, measurement k will occur since

$$\begin{aligned}(\alpha \langle\phi_1| + \beta \langle\phi_2|)P_k^\dagger P_k(\alpha |\phi_1\rangle + \beta |\phi_2\rangle) &= (\alpha \langle\phi_1| + \beta \langle\phi_2|)(|\phi_1\rangle\langle\phi_1| + |\phi_2\rangle\langle\phi_2|)(\alpha |\phi_1\rangle + \beta |\phi_2\rangle) \\ &= 1\end{aligned}$$

Furthermore, the measurement does not collapse the superposition, since

$$P_k(\alpha |\phi_1\rangle + \beta |\phi_2\rangle) = \alpha |\phi_1\rangle + \beta |\phi_2\rangle .$$

This means the measurement reveals the error but not the actual state of the encoded qubit. The error can then be corrected by flipping the corresponding qubit again [Nielsen and Chuang, 2010, pp.427-429].

The error-correcting quantum code introduced above exhibits a remarkable property commonly encountered in the quantum realm but absent from classical error correction. Consider (4.1) and the outcome of performing any two phase flips on the encoded state. It is easy to verify

$$|\psi\rangle_{\text{enc}} = (Z \otimes Z \otimes I) |\psi\rangle_{\text{enc}} = (Z \otimes I \otimes Z) |\psi\rangle_{\text{enc}} = (I \otimes Z \otimes Z) |\psi\rangle_{\text{enc}} . \tag{4.2}$$

Although the three choices of two phase flips are distinct errors, they all yield the same encoded state. In classical coding theory, two distinct errors applied to the same string of bits result in different outputs.

In the general setup, an error is modeled as the result of some quantum channel $\mathcal{E}$. The error-correcting channel $\mathcal{R}$ is then applied to the erroneous state $\mathcal{E}(|\psi\rangle_{\text{enc}} \langle\psi|_{\text{enc}})$ reproducing the original state if the error is correctable, that is

$$\mathcal{R}(\mathcal{E}(|\psi\rangle_{\text{enc}} \langle\psi|_{\text{enc}})) \propto |\psi\rangle_{\text{enc}} \langle\psi|_{\text{enc}} .$$

The quantum channel $\mathcal{E}$ modeling the error may not be trace-preserving. The error-correction channel, however, is required to succeed with probability 1, explaining proportionality rather than equality. In the case where $\mathcal{E}$ is trace-preserving, taking the

trace on both sides of the proportionality yields a proportionality constant of 1, hence an equality.

The three qubit example considered above was able to correct the errors $\{E_i\}_{i=0}^3$ defined in table 4.1. This is only a small subset of the possible errors that may occur. The next theorem provides a method for checking which errors a given code protects against.

Theorem 4.2: The Knill-Laflamme Condition
Let $\mathcal{E}$ be a quantum channel with Kraus operators $\{E_i\}_{i=1}^N$. In addition, let C be some quantum code and P an orthogonal projector onto the code space of C. Then

$$\mathcal{R}(\mathcal{E}(\rho)) \propto \rho$$

if and only if

$$PE_i^\dagger E_j P = h_{ij} P, \tag{4.3}$$

where $\mathcal{R}$ is some error-correcting quantum channel and h_{ij} are the entries of some hermitian matrix H.

[Nielsen and Chuang, 2010, p.436]

Proof
Assume first (4.3) holds for the Kraus operators $\{E_i\}_{i=1}^N$. Since H is hermitian it follows from the spectral theorem it can be decomposed as $H = UDU^\dagger$ or equivalently $D = U^\dagger HU$ where D is a real diagonal matrix and U is unitary. Define new Kraus operators by $\{F_j := \sum_{i=1}^N u_{ij} E_i\}_{j=1}^N$. Since these are related by a unitary matrix, it follows from theorem 3.6 they both correspond to the same quantum channel. For these new Kraus operators it holds;

$$PF_k^\dagger F_{k'} P = \sum_{i=1}^N \sum_{j=1}^N u_{ik}^* u_{jk'} PE_i^\dagger E_j P = \sum_{i=1}^N \sum_{j=1}^N u_{ik}^* h_{ij} u_{jk'} P = d_{kk'} P$$

Since D is real and diagonal it is also hermitian. The Kraus operators $\{F_j\}_{j=1}^N$ thus satisfies a special, more strict, case of (4.3). By the polar decomposition

$$F_k P = U_k \sqrt{P^\dagger F_k^\dagger F_k P} = U_k \sqrt{d_{kk}} P.$$

Define orthogonal projectors by $P_k := U_k P U_k^\dagger = (\sqrt{d_{kk}})^{-1} F_k P U_k^\dagger$. It is evident they project onto orthogonal subspaces since

$$P_k P_{k'} = P_k^\dagger P_{k'} = (U_k P U_k^\dagger)^\dagger U_{k'} P U_{k'}^\dagger = \frac{U_k P F_k^\dagger F_{k'} P U_{k'}^\dagger}{\sqrt{d_{kk} d_{k'k'}}} = \frac{U_k d_{kk'} P U_{k'}^\dagger}{\sqrt{d_{kk} d_{k'k'}}} = \delta_{kk'} U_k P U_{k'}^\dagger.$$

The error-detection and correction channel is then the trace-preserving channel given by $\mathcal{R}(\sigma) := \sum_k U_k^\dagger P_k \sigma P_k^\dagger U_k$, where $\{P_k\}_k$ may be augmented with additional orthogonal projectors to fulfill the completeness relation $\sum_k P_k^\dagger P_k = \sum_k P_k = I$. To realize this channel corrects the errors encountered by $\mathcal{E}$, consider first the following: Let ρ operate on the code space, then $P\rho P = \rho$ and thus

$$U_k^\dagger P_k F_{k'} \rho F_{k'}^\dagger P_k^\dagger U_k = U_k^\dagger P_k^\dagger F_{k'} P \rho P F_{k'}^\dagger P_k U_k$$

$$
\begin{aligned}
&= \frac{U_k^\dagger U_k P F_k^\dagger F_{k'} P \rho P F_{k'}^\dagger F_k P U_k^\dagger U_k}{d_{kk}} \\
&= \frac{d_{kk'} d_{k'k} P \rho P}{d_{kk}} \\
&= \delta_{kk'} d_{kk} \rho.
\end{aligned}
$$

Consequently

$$
\begin{aligned}
\mathcal{R}(\mathcal{E}(\rho)) &= \sum_k U_k^\dagger P_k \left(\sum_{k'=1}^{N} F_{k'} \rho F_{k'}^\dagger \right) P_k^\dagger U_k \\
&= \sum_k \sum_{k'=1}^{N} U_k^\dagger P_k F_{k'} \rho F_{k'}^\dagger P_k^\dagger U_k \\
&= \left(\sum_k d_{kk} \right) \rho \propto \rho.
\end{aligned}
$$

Conversely, assume $\mathcal{R}(\mathcal{E}(\rho)) \propto \rho$. Since P is a projector onto the code space, $P\rho P$ is an operator on the code space for any density operator ρ, it thus follows $\mathcal{R}(\mathcal{E}(P\rho P)) \propto P\rho P$. Expressing $\mathcal{R} \circ \mathcal{E}$ in terms of Kraus operators

$$
\sum_j \sum_{i=1}^{N} R_j E_i P \rho P E_i^\dagger R_j^\dagger = c P \rho P,
$$

for some constant $c \in \mathbb{C}$ independent of ρ and where $\{R_j\}_j$ and $\{E_i\}_{i=1}^N$ are the Kraus operators for $\mathcal{R}$ and $\mathcal{E}$, respectively. This equation reveals that the channel $\mathcal{R} \circ \mathcal{E}$ with Kraus operators $\{R_j E_i\}_{i,j}$ is identical to the channel with a single Kraus operator $\sqrt{c}$. According to theorem 3.6 this implies

$$
R_j E_i = c_{ij} \quad \Leftrightarrow \quad E_i^\dagger R_j^\dagger = c_{ij}^*, \qquad c_{ij} \in \mathbb{C}.
$$

It then follows

$$
P E_i^\dagger R_k^\dagger R_k E_j P = c_{ik}^* c_{jk} P.
$$

Since $\mathcal{R}$ is trace-preserving $\sum_k R_k^\dagger R_k = I$. Addition over k of the above equation then yields

$$
P E_i^\dagger E_j P = \left(\sum_k c_{ik}^* c_{jk} \right) P = \alpha_{ij} P
$$

where $\alpha_{ij} := \sum_k c_{ik}^* c_{jk}$. These are indices of a hermitian matrix since $\alpha_{ij}^* = \sum_k c_{ik} c_{jk}^* = \alpha_{ji}$. ∎

The Kraus operators for the channel $\mathcal{E}$ introducing the noise is called errors and if an error-correcting channel $\mathcal{R}$ exists the Kraus operators are called a correctable set of errors.

As mentioned earlier, the set of possible errors in the quantum realm forms a continuum. The following theorem, which can be seen as an extension of the Knill-Laflamme condition, allows for a discretization of the possible errors a quantum code needs to correct.

Corollary 4.3
Let $\mathcal{E}$ be a quantum channel with Kraus operators $\{E_i\}_{i=1}^N$ and $\mathcal{F}$ be a second quantum channel with Kraus operators $\{F_j = \sum_{i=1}^N \alpha_{ij} E_i\}_{j=1}^N$, where α_{ij} are some complex numbers. Suppose the Kraus operators $\{E_i\}_{i=1}^N$ are correctable errors for the error-correcting quantum channel $\mathcal{R}$ on some quantum code C. Then $\{F_j\}_{j=1}^N$ is also a set of correctable errors for $\mathcal{R}$.

[Nielsen and Chuang, 2010, p.438]

Proof
Let the noise introducing channel $\mathcal{E}$ have Kraus operators $\{E_i\}_{i=1}^N$ and assume this is a correctable set of errors, thus satisfying (4.3). According to the proof of theorem 4.2, the hermitian matrix H of (4.3) can be chosen to be real and diagonal without loss of generality. Let P be the projector onto the code space C, and U_k be the unitary matrix from the polar decomposition of $E_k P$, that is, $E_k P = U_k \sqrt{P^\dagger E_k^\dagger E_k P} = U_k \sqrt{h_{kk}} P$. Moreover, let P_k be the orthogonal projectors defined by $P_k := U_k P U_k^\dagger$. From the proof of theorem 4.2 it follows that $U_k^\dagger P_k$ are the Kraus operators of the error-correction channel $\mathcal{R}$ for $\mathcal{E}$ and satisfies

$$U_k^\dagger P_k E_i \rho E_i^\dagger P_k^\dagger U_k = \delta_{ki} d_{kk} \rho.$$

Since $F_j = \sum_{i=1}^N \alpha_{ij} E_i$, multiplication by α_{ij} followed by addition over i reveals

$$U_k P_k F_j \rho F_j^\dagger P_k^\dagger U_k = \sum_{i=1}^N \alpha_{ij} \delta_{ki} d_{kk} \rho = \alpha_{kj} d_{kk} \rho.$$

Altogether, this implies the quantum channel $\mathcal{F}$ with Kraus operators $\{F_j\}_{j=1}^N$ is correctable by $\mathcal{R}$ as

$$\begin{aligned} \mathcal{R}(\mathcal{F}(\rho)) &= \sum_k U_k^\dagger P_k \left(\sum_{j=1}^N F_j \rho F_j^\dagger \right) P_k^\dagger U_k \\ &= \left(\sum_k d_{kk} \sum_{j=1}^N \alpha_{kj} \right) \rho \propto \rho. \end{aligned}$$

■

It is fairly simple to show that any 2×2 complex matrix A can be expressed as a linear combination of the three Pauli matrices and the identity. The Y Pauli matrix can further be expressed as the product of the X and Z matrix. Combined this means there exists complex numbers α, β, γ and δ such

$$A = \alpha X + \beta XZ + \gamma Z + \delta I.$$

Taking into account corollary 4.3, it is sufficient to consider the Pauli errors when examining the correctability properties of a quantum error-correcting code. It is furthermore seen that if one can show a code protects against any single bit flip, phase flip, and the combination of both, the code protects against any arbitrary single qubit error.

4.3 Stabilizer Codes

This section makes use of group-theoretic definitions and results presented in appendix A. An important group used in stabilizer codes is the N-fold Pauli; an extension of the Pauli group defined in the appendix, given by

$$\mathcal{P}_N := \{\alpha \bigotimes_{i=1}^{N} A_i \mid A_i \in \{I, X, Y, Z\},\ \alpha \in \{\pm 1, \pm i\}\}.$$

Consider the example of section 4.2, which encodes one logical qubit in three physical qubits, and especially (4.2). The equations show that any two phase flips leave the encoded state unchanged; obviously, this is also the case for the 3-fold identity matrix. Consequently, one says the set $\{I \otimes I \otimes I, Z \otimes Z \otimes I, Z \otimes I \otimes Z, I \otimes Z \otimes Z\}$ stabilizes $|\psi\rangle_{\text{enc}}$. For a given operator, the vectors it stabilizes are exactly the eigenvectors corresponding to eigenvalue 1. This principle is fundamental to stabilizer codes. However, prior to their formal definition, several preliminary concepts require introduction.

Firstly, assume the set $\{A_i\}_{i=1}^{N}$ consists of matrices, forms a group under matrix multiplication, and stabilizes some vector $|x\rangle$. Let the set be generated by $\{B_i\}_{i=1}^{M}$, that is, $\langle B_1, \ldots, B_M\rangle = \{A_i\}_{i=1}^{N}$. Since the set stabilizes $|x\rangle$, all generators stabilize $|x\rangle$ as they belong to the set. Consider the converse, i.e. $|x\rangle$ is stabilized by all B_i. Then for any $A_i = B_1^{r_{i,1}} \cdots B_M^{r_{i,M}}$ with $r_{i,j} \in \mathbb{Z}$, but since the generators stabilize $|x\rangle$, it follows $A_i |x\rangle = B_1^{r_{i,1}} \cdots B_M^{r_{i,M}} |x\rangle = |x\rangle$. This means a set can be verified to stabilize some vector simply by considering a generating set; a time-saving fact as there exists a generating set for which $M \leq \log(N)$. Considering this, it is of interest to find a generating set such

$$B_1^{r_1} \cdots B_{i-1}^{r_{i-1}} B_{i+1}^{r_{i+1}} \cdots B_M^{r_M} = B_i^{r_i}, \qquad r_1, \ldots, r_M \in \mathbb{Z}$$

implies $r_i = 0$ for all $i \in \{1, \ldots, M\}$, since the converse would imply

$$\langle B_1, \ldots, B_{i-1}, B_i, B_{i+1}, \ldots, B_M\rangle = \langle B_1, \ldots, B_{i-1}, B_{i+1}, \ldots, B_M\rangle.$$

Generating sets for which $r_i = 0$ for all $i \in \{1, \ldots, M\}$ is said to be independent.

Checking if a generating set for a subgroup of $\mathcal{P}_N$ is independent may seem infeasible, as it possibly involves large sets of huge matrices. The introduction of the check matrix presents a straightforward method involving simple linear algebra. Consider the example of section 4.2. The code space is stabilized by the matrices of figure 4.3a. These matrices forms a group under matrix multiplication and is generated by the elements in boldface.

For a general code space with stabilizers $G = \langle g_1, \ldots, g_M\rangle \subseteq \mathcal{P}_N$, the check matrix is an $M \times 2N$ matrix with entries in $\mathbb{F}_2$ where row i corresponds to generator g_i of G. Since $g_i \in \mathcal{P}_N$ it is of the form $\alpha \bigotimes_{k=1}^{N} A_k$ for some $A_k \in \{I, X, Y, Z\}$ and $\alpha \in \{\pm 1, \pm i\}$. Let $j \in \{1, \ldots, N\}$ the $ij'th$ entry of the check matrix is then a 1 if $A_j = X \vee Y$ and 0 otherwise. Similarly, the $n + j'th$ entry of row i is 1 if $A_j = Z \vee Y$. Notice the check matrix does not encode information regarding the phase α. Since $Y = iXZ$ any element of $\mathcal{P}_N$ can be expressed as

$$\alpha \bigotimes_{i=1}^{N} X^{a_i} Z^{b_i}, \qquad a_i, b_i \in \{0, 1\}. \tag{4.4}$$

A row in the check matrix is therefore given by the map $r : \mathcal{P}_N \to \mathbb{F}_2^{2N}$ defined by

$$r\left(\alpha \bigotimes_{i=1}^{N} X^{a_i} Z^{b_i}\right) := [a_1 \ \ldots \ a_N \ b_1 \ \ldots \ b_N]. \tag{4.5}$$

The check matrix corresponding to the three qubit code of section 4.2 is the 4×6 matrix shown in figure 4.3b [Nielsen and Chuang, 2010, pp-456-457].

g_1	$I \otimes I \otimes I$
$\mathbf{g_2}$	$\mathbf{Z \otimes Z \otimes I}$
g_3	$Z \otimes I \otimes Z$
$\mathbf{g_4}$	$\mathbf{I \otimes Z \otimes Z}$

(a) Stabilizers for the three-qubit code.

$$\left[\begin{array}{ccc|ccc} 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right]$$

(b) Corresponding 2×6 check matrix.

Figure 4.3: Stabilizers for the three qubit code and corresponding check matrix. The generators of the stabilizer group is marked in boldface.

As postulated, the check matrix is an important object in determining the independence of a generating set. The following theorem connects the independence of a generating set to the linear independence of the rows of the associated check matrix.

Theorem 4.4
Let $G = \langle g_1, \ldots, g_M \rangle \subseteq \mathcal{P}_N$ be a group under matrix multiplication. Assume G is abelian and $-I \notin G$. The generators of G are then independent if and only if the rows of the check matrix are linearly independent.

[Bergou et al., 2021, p.180]

Proof
Let $g \in \mathcal{P}_N$, it then follows $gg = I \vee -I$. Since $G = \langle g_1, \ldots, g_M \rangle \subseteq \mathcal{P}_N$ it specifically applies to $g = g_i$ for any $i \in \{1, \ldots, M\}$, however, by assumption $-I \notin G$ implying $g_i g_i = I$. Notice the statement of the theorem is equivalent to the rows of the check matrix being linearly dependent if and only if the generators of G are dependent. Thus, assume the rows of the check matrix are linearly dependent and denote by $r(g_i)$ the row corresponding to generator g_i. Then there exists $\{\alpha_i \in \{0, 1\}\}_{i=1}^{M}$ not all zero such

$$\sum_{i=1}^{M} \alpha_i r(g_i) = r\left(\prod_{i=1}^{M} g_i^{\alpha_i}\right) = 0.$$

Considering how the check matrix encodes information, the element of G (expressed as a product of elements of G) must be the identity possibly with an additional phase

$$\prod_{i=1}^{M} g_i^{\alpha_i} = \pm I \vee \pm iI.$$

Since $-I \notin G$ it follows $\pm iI \notin G$. The above equation must therefore evaluate to I. Assume $\alpha_k = 1$ for some $k \in \{1, \dots, M\}$. Multiplication by g_k^{-1} on both sides of the above, paired with the fact that G is assumed abelian it follows

$$g_k = g_k^{-1} = \prod_{\substack{i=1 \\ i \neq k}}^{M} g_i^{\alpha_i}.$$

Conversely, assume the generating set is dependent, then there exists $k \in \{1, \dots, M\}$ and $\{\alpha_i \in \mathbb{Z}\}_{i=1}^{M}$ not all zero such

$$\prod_{\substack{i=1 \\ i \neq k}}^{M} g_i^{\alpha_i} = g_k^{\alpha_k} \qquad \Leftrightarrow \qquad \prod_{i=1}^{M} g_i^{\alpha_i} = I, \quad \alpha_i \in \{0, 1\},$$

where the bi-implication follows from the fact that G is assumed abelian and $g_i g_i = I$. Now consider the rows of the check matrix

$$\sum_{i=1}^{M} \alpha_i r(g_i) = r\left(\prod_{i=1}^{M} g_i^{\alpha_i}\right) = 0.$$

Since α_i are not all zero, the rows of the check matrix are linearly dependent. ∎

Not only does the above theorem connect independence of generators to the check matrix, it also provides useful information necessary to prove the next important theorem.

Theorem 4.5
Let $G = \langle g_1, \dots, g_{N-k} \rangle \subseteq \mathcal{P}_N$ be a group under matrix multiplication. Assume G is abelian, $-I \notin G$ and its generators are independent. The vector space stabilized by the elements of G is then of dimension 2^k.

[Bergou et al., 2021, p.181]

For proof see appendix B.

The next theorem is vital when defining stabilizer codes and can be viewed as an extension of the above theorem.

Theorem 4.6
Let $G = \langle g_1, \dots, g_{N-k} \rangle$ be a subgroup of $\mathcal{P}_N$ under matrix multiplication and assume the generators are independent. The vector space stabilized by the elements of G is a non-trivial subspace if and only if $-I \notin G$ and G is abelian.

[Nielsen and Chuang, 2010, p.455]

Proof
Assume G stabilizes a non-trivial vector space $V \neq \{0\}$ and let $|x\rangle \neq [0 \ \cdots \ 0]^T \in V$. If $-I \in G$ then $-|x\rangle = -I|x\rangle = |x\rangle$ a contradiction since $|x\rangle \neq [0 \ \cdots \ 0]^T$. Furthermore, since all Pauli matrices commute or anti-commute it suffices to show the elements of G

cannot anti-commute. Let $A, B \in G$ anti-commute, then $|x\rangle = AB|x\rangle = -BA|x\rangle = -|x\rangle$ a contradiction.
The converse implication, that any abelian group G for which $-I \notin G$ stabilizes a non-trivial vector space, follows directly from theorem 4.5. ∎

4.3.1 Construction and Error Correction

Keeping the above theorems and notation in mind, stabilizer codes can now be formally defined. Let $G = \langle g_1, \ldots, g_{N-k} \rangle \subseteq \mathcal{P}_N$ be an abelian group with independent generators and $-I \notin G$. The vector space stabilized by G, denoted V_G, is then, by definition, a stabilizer code space, and the stabilizer code is denoted $C(G)$. The code space V_G is the intersection of the $+1$ eigenspace of the operators of G. According to theorem 4.5, V_G is of dimension 2^k, thus encoding k logical qubits in N physical qubits. Such a code is called an $[[N, k]]$ code with the double parenthesis indicating a quantum code, thereby distinguishing it from a classical code.

Two important questions arise when defining an error-correcting code. Firstly, what errors can the code detect and correct? And secondly, what scheme can correct the correctable errors? Consider the first question. Any error $E \in \mathcal{P}_N$ can be divided into three categories according to theorem A.8. Either the error belongs to the stabilizer group G, to the N-fold Pauli group but not the normalizer of G in $\mathcal{P}_N$ or to the normalizer of G in $\mathcal{P}_N$ but not to G. In the first case, no error has occurred since $E \in G$ implies $E|\psi\rangle = |\psi\rangle$ for any $|\psi\rangle \in V_G$. In the second case, since $N_{\mathcal{P}_N}(G) = C_{\mathcal{P}_N}(G)$ for any stabilizer group G, not belonging to the normalizer implies at least one element of G does not commute with the error E. Any two elements of the N-fold Pauli group either commutes or anti-commutes, thus at least one stabilizer anti-commutes with the error. Denote this stabilizer by g, then for any $|\psi\rangle, |\phi\rangle \in V_G$,

$$\langle\phi| E |\psi\rangle = \langle\phi| g^\dagger E g |\psi\rangle = -\langle\phi| E |\psi\rangle,$$

which would imply $\langle\phi| E |\psi\rangle = 0$. The error thus maps any vector in V_G to an orthogonal subspace. The erroneous codeword can therefore be reliably distinguished from any codeword. In the third case, the error commutes with every stabilizer, but is not in the stabilizer group, that is, $Eg = gE$ for every $g \in G$. Let $|\psi\rangle \in V_G$, then

$$E|\psi\rangle = Eg|\psi\rangle = gE|\psi\rangle$$

which implies $E|\psi\rangle \in V_G$. The error thus maps a codeword into a new codeword, making it undetectable [Nielsen and Chuang, 2010, p.465].

The following theorem summarizes the above, highlighting the fact that the critical errors lies in the third category considered.

Theorem 4.7: Error-Correction Condition for Stabilizer Codes
Let $G \subseteq \mathcal{P}_N$ be a stabilizer group for some stabilizer code $C(G)$. Let $\{E_i \in \mathcal{P}_N\}_{i=1}^M$ be a set of Pauli operators such $E_i^\dagger E_j \notin N_{\mathcal{P}_N}(G) \backslash G$ for all $i, j \in \{1, \ldots, M\}$. Then $\{E_i\}_{i=1}^M$ is a correctable set of errors for the code $C(G)$.

[Nielsen and Chuang, 2010, p.466]

Proof

Let P be the orthogonal projector onto the code space V_G of the code $C(G)$ and assume $E_i^\dagger E_j \notin N_{\mathcal{P}_N}(G)\backslash G$. According to theorem A.8 $G \subseteq N_{\mathcal{P}_N}(G) \subseteq \mathcal{P}_N$, thus two cases are possible. Either $E_i^\dagger E_j \in G$ or $E_i^\dagger E_j \in \mathcal{P}_N \backslash N_{\mathcal{P}_N}(G)$. In the first case, since $E_i^\dagger E_j \in G$, they stabilize any vector in V_G and it thus follows $PE_i^\dagger E_j P = P = PE_j^\dagger E_i P$. Next, consider the second case. Since $E_i^\dagger E_j \in \mathcal{P}_N \backslash N_{\mathcal{P}_N}(G)$ and $N_{\mathcal{P}_N}(G) = Z(G)$ there must be at least one element in G that does not commute with $E_i^\dagger E_j$, however, all elements of $\mathcal{P}_N$ either commute or anti-commute, thus there exist $g \in G$ such $E_i^\dagger E_j g = -gE_i^\dagger E_j$. Let $\{g_1 = g, \ldots, g_{N-k}\}$ be a generating set for G. The projector P can then be expressed as

$$P = \frac{\prod_{i=1}^{N-k}(I + g_i)}{2^{N-k}}$$

Since $E_i^\dagger E_j$ anti-commutes with g_1 it follows

$$E_i^\dagger E_j P = (I - g_1)\frac{\prod_{i=2}^{N-k}(I + g_i)}{2^{N-k}}.$$

By definition of G, g_i commutes and thus $I + g_i$ also commutes. Since $(I + g_1)(I - g_1) = 0$ it is the case that $P(I - g_1) = 0$, hence $PE_i^\dagger E_j P = 0 = PE_j^\dagger E_i P$. The errors $\{E_i\}_{i=1}^M$ are therefore correctable according to theorem 4.2, with the matrix H having entries $h_{ij} = h_{ji} \in \{0, 1\}$. ∎

With the first important question answered, focus can be laid upon the second. However, before showing a correction strategy for the correctable errors, consider first the N-fold Pauli group. Based on (4.4), the weight of any element $E \in \mathcal{P}_N$ is defined by

$$w(E) = \sum_{i=1}^{N} \mathbb{1}_{[a_i=1 \ \vee \ b_i=1]}. \tag{4.6}$$

The weight of E, is the number of entries in the tensor product of E not equal to the identity. Notice the phase α has no effect on the weight.

Suppose an error $E \in \mathcal{P}_N$ occurred, transforming the codeword $|\psi\rangle \in V_G$ into the erroneous state $E|\psi\rangle$. Define orthogonal projectors onto the $+1$ and -1 eigenspace, respectively, of each generator of the stabilizer group G by

$$P_i^{(+)} := \frac{I + g_i}{2} \qquad \text{and} \qquad P_i^{(-)} := \frac{I - g_i}{2}.$$

Since the generators commute, so does the projectors $\left\{P_i^{(+)}, P_i^{(-)}\right\}_{i=1}^{N-k}$. Furthermore, for each i these projectors satisfy the completeness relation

$$\left(P_i^{(+)}\right)^\dagger P_i^{(+)} + \left(P_i^{(-)}\right)^\dagger P_i^{(-)} = I,$$

and thus define valid measurement operators. Measuring these projectors for each i reveals the syndrome given by $[f_{g_1}(E), \ldots, f_{g_{N-k}}(E)]^T$ where

$$f_{g_i}(E) := \begin{cases} 1, & \text{if } [g_i, E] := g_i E - E g_i = 0, \\ -1, & \text{if } \{g_i, E\} := g_i E + E g_i = 0. \end{cases}$$

Suppose the syndrome is -1 for entries $S \subseteq \{1, \ldots, N-k\}$. Define the set

$$T := \{A \in \mathcal{P}_N \mid g_k A = -A g_k \wedge g_l A = A g_l, \quad k \in S, \quad l \in \{1, \ldots, N-k\} \backslash S\},$$

the error, E, then belongs to T. Since T is a subset of the N-fold Pauli group, all elements are unitary and hence invertible, the inverse being the hermitian conjugate. Suppose an error-correction scheme identifies the error as some minimum weight element of T, denoted $E_{\min}$. Then since both the identified error and E have the same syndrome, it must be the case that $E_{\min} g_i E_{\min}^\dagger = \pm g_i = E g_i E^\dagger$ and thus $E_{\min}^\dagger E g_i E^\dagger E_{\min} = g_i$, implying $E_{\min}^\dagger E$ belongs to the normalizer of G in $\mathcal{P}_N$. Two cases could now occur, either $E_{\min}^\dagger E \in G$, in which case $E_{\min}^\dagger E |\psi\rangle = |\psi\rangle$ hence restoring the original state. If, on the other hand, $E_{\min}^\dagger E \in N_{\mathcal{P}_N}(G) \backslash G$ then $E_{\min}^\dagger E |\psi\rangle = |\phi\rangle \in V_G$ for some $|\phi\rangle \neq |\psi\rangle$, that is, the error correction incorrectly maps the erroneous state back into the code space [Gottesman, 1997, pp.19-20].

The distance, d, of an $[[N, k]]$ stabilizer code is defined to be the minimum weight of any element of $N_{\mathcal{P}_N}(G) \backslash G$. Such a code is called an $[[N, k, d]]$ code. According to the above paragraph and theorem 4.7, the error E is correctable if $E_{\min}^\dagger E \notin N_{\mathcal{P}_N}(G) \backslash G$, since any element of $N_{\mathcal{P}_N}(G) \backslash G$ has weight at least d, the condition is guaranteed if $w(E_{\min}^\dagger E) \leq d - 1 < d$. Let $E = \alpha_1 \bigotimes_{i=1}^N X^{a_i} Z^{b_i}$ and $E_{\min} = \alpha_2 \bigotimes_{i=1}^N X^{c_i} Z^{d_i}$ then

$$w(E) = \sum_{i=1}^N \mathbb{1}_{[a_i=1 \ \vee \ b_i=1]} := t \qquad \text{and} \qquad w(E_{min}^\dagger) = \sum_{i=1}^N \mathbb{1}_{[c_i=1 \ \vee \ d_i=1]}.$$

Furthermore, $E_{\min}^\dagger E = \alpha_2 \alpha_1 \bigotimes_{i=1}^N Z^{d_i} X^{c_i} X^{a_i} Z^{b_i}$ and thus

$$\begin{aligned} w(E_{\min}^\dagger E) = \sum_{i=1}^N \mathbb{1}_{[a_i \neq c_i \ \vee \ b_i \neq d_i]} &\leq \sum_{i=1}^N \mathbb{1}_{[(a_i=1 \ \vee \ b_i=1) \ \vee \ (c_i=1 \ \vee \ d_i=1)]} \\ &\leq \sum_{i=1}^N \mathbb{1}_{[a_i=1 \ \vee \ b_i=1]} + \sum_{i=1}^N \mathbb{1}_{[c_i=1 \ \vee \ d_i=1]} \\ &= w(E) + w(E_{\min}^\dagger). \end{aligned}$$

Since $E_{\min}$ was chosen as the minimum weight element of T and $E \in T$ it follows

$$w(E_{\min}^\dagger E) \leq w(E_{\min}^\dagger) + w(E) \leq 2w(E) = 2t.$$

If $2t \leq d - 1$ it is then guaranteed the error is correctable.

The above derivation does not, however, imply an error of weight greater than t is not correctable as the next example will show.

Example 5: Correctable Errors

Consider the stabilizer code introduced in section 4.2. The stabilizer group was found to be $\{I \otimes I \otimes I, Z \otimes Z \otimes I, Z \otimes I \otimes Z, I \otimes Z \otimes Z\}$. The operator Z commutes with itself and it is thus easy to verify that $\{Z \otimes I \otimes I, I \otimes Z \otimes I, I \otimes I \otimes Z\}$ belongs to the normalizer of the stabilizer group in $\mathcal{P}_N$, but not the stabilizer group itself. Since all three of these elements have weight 1, the lowest possible weight of an element not in the stabilizer, the code has distance 1. It can therefore correct any arbitrary error of weight $\lfloor \frac{1-1}{2} \rfloor = 0$, which is only

satisfied by the identity operator, an operator belonging to the stabilizer and thus not an error. However, a simple derivation shows the code can detect an arbitrary single bit flip error, since such an error anti-commutes with either the first, second or both generators.

As mentioned previously, in quantum coding theory, distinct errors could lead to the same state, a property not encountered in classical coding theory. This implies that distinct errors may be corrected by a single inverse operation. Based on this, quantum error-correcting codes can be divided into two groups; degenerate and nondegenerate. Denote by $\mathcal{A} := \{E_j \in \mathcal{P}_N\}_j$ the intended correctable set of errors for some code. If the code indeed corrects $\mathcal{A}$ and, in addition, multiple errors in $\mathcal{A}$ are mapped to the same syndrome, then the code is called degenerate. This also means a code that is degenerate with respect to $\mathcal{A}$ may be nondegenerate with respect to $\mathcal{B} \subset \mathcal{A}$ [Smmith and Smolin, 2006, p.1].

Considering the error-correcting abilities of stabilizer codes derived above, an important connection between the number of physical qubits, the number of logical qubits, and the distance of a stabilizer code can be derived. The connection is known as the quantum singleton bound, a quantum version of the classical connection under the same name.

Theorem 4.8: Quantum Singleton Bound
Let a stabilizer code with parameters $[[N, k, d]]$ be given. Then

$$N - k \geq 2(d-1)$$

[Nielsen and Chuang, 2010, pp.568-569]

Proof
Let an $[[N, k, d]]$ stabilizer code be given and denote by $\mathcal{H}_\mathcal{A}$ the state space to which the N physical qubits belong. The coding space is then a 2^k dimensional subspace of $\mathcal{H}_\mathcal{A}$ denoted by $\mathcal{H}_{\mathcal{A}_c}$ according to theorem 4.5. Denote by $\{|\psi_i\rangle\}_i^{2^k}$ an orthonormal basis of $\mathcal{H}_{\mathcal{A}_c}$. Let $\mathcal{H}_\mathcal{B}$ be a 2^k dimensional state space with orthonormal basis identical to that of $\mathcal{H}_{\mathcal{A}_c}$. Consider the entangled state of $\mathcal{H}_{\mathcal{A}_c} \otimes \mathcal{H}_\mathcal{B}$ given by

$$|\phi\rangle = \frac{1}{\sqrt{2^k}} \sum_i^{2^k} |\psi_i\rangle \otimes |\psi_i\rangle .$$

Thus, $\mathcal{B}$, is a reference system that keeps track of the logical information encoded in $\mathcal{A}$. The density operator $\rho^{\mathcal{B}\mathcal{A}_c} = |\phi\rangle\langle\phi|$ can be verified to be pure. Furthermore, the reduced density operator of system $\mathcal{A}_c$ is given by $\rho^{\mathcal{A}_c} = \frac{1}{2^k} I$, the maximally mixed state. By the second property of theorem 3.2 it then follows $\mathcal{S}(\rho^{\mathcal{A}_c}) = \log\left(2^k\right) = k$. Divide the system $\mathcal{A}$ consisting of N qubits into three. The first two, $\mathcal{A}_1$ and $\mathcal{A}_2$, containing $d-1$ qubits, and the third system, $\mathcal{A}_3$, containing the remaining $N - 2(d-1)$ qubits. Since the stabilizer has distance d, any located error of weight $d-1$ can be corrected. Thus, if either $\mathcal{A}_1$ or $\mathcal{A}_2$ are corrupted, the logical information can be recovered. This would imply that no logical information reside in $\mathcal{A}_1$ or $\mathcal{A}_2$ alone. Since $\mathcal{B}$ is a reference system keeping track of the logical information, the mutual information between $\mathcal{B}$ and $\mathcal{A}_1$ or $\mathcal{A}_2$ must be zero, by the definition of mutual information, it then follows $\mathcal{S}(\rho^{\mathcal{B}\mathcal{A}_1}) = \mathcal{S}(\rho^\mathcal{B}) + \mathcal{S}(\rho^{\mathcal{A}_1})$ and

$\mathcal{S}(\rho^{\mathcal{B}\mathcal{A}_2}) = \mathcal{S}(\rho^{\mathcal{B}}) + \mathcal{S}(\rho^{\mathcal{A}_2})$. Additionally, by the purity of $\rho^{\mathcal{B}\mathcal{A}} = \rho^{\mathcal{B}\mathcal{A}_1\mathcal{A}_2\mathcal{A}_3}$ according to property 4 of theorem 3.2, $\mathcal{S}(\rho^{\mathcal{B}\mathcal{A}_1}) = \mathcal{S}(\rho^{\mathcal{A}_2\mathcal{A}_3})$ and $\mathcal{S}(\rho^{\mathcal{B}\mathcal{A}_2}) = \mathcal{S}(\rho^{\mathcal{A}_1\mathcal{A}_3})$ in conjunction with the subadditivity of theorem 3.5, two equations have been derived

$$\mathcal{S}(\rho^{\mathcal{B}}) + \mathcal{S}(\rho^{\mathcal{A}_1}) = \mathcal{S}(\rho^{\mathcal{B}\mathcal{A}_1}) = \mathcal{S}(\rho^{\mathcal{A}_2\mathcal{A}_3}) \leq \mathcal{S}(\rho^{\mathcal{A}_2}) + \mathcal{S}(\rho^{\mathcal{A}_3})$$
$$\mathcal{S}(\rho^{\mathcal{B}}) + \mathcal{S}(\rho^{\mathcal{A}_2}) = \mathcal{S}(\rho^{\mathcal{B}\mathcal{A}_2}) = \mathcal{S}(\rho^{\mathcal{A}_1\mathcal{A}_3}) \leq \mathcal{S}(\rho^{\mathcal{A}_1}) + \mathcal{S}(\rho^{\mathcal{A}_3}).$$

Adding the two equations and simplifying yields

$$2\mathcal{S}(\rho^{\mathcal{B}}) + \mathcal{S}(\rho^{\mathcal{A}_1}) + \mathcal{S}(\rho^{\mathcal{A}_2}) \leq \mathcal{S}(\rho^{\mathcal{A}_1}) + \mathcal{S}(\rho^{\mathcal{A}_2}) + 2\mathcal{S}(\rho^{\mathcal{A}_3}) \qquad \Rightarrow \qquad \mathcal{S}(\rho^{\mathcal{B}}) \leq \mathcal{S}(\rho^{\mathcal{A}_3})$$

Since system $\mathcal{A}_3$ consists of $N - 2(d-1)$ qubits, by property 2 of theorem 3.2, $\mathcal{S}(\rho^{\mathcal{A}_3}) \leq N - 2(d-1)$, thus $k \leq N - 2(d-1)$ or $N - k \geq 2(d-1)$ ∎

Suppose that Alice wishes to send 1 logical qubit encoded in 5 physical qubits to Bob. The quantum singleton bound then implies $d \leq 3$. For given N and k the quantum singleton bound thus provides an upper bound on the achievable distance.

Having established the definition of a stabilizer code and characterized its error-correction capabilities, the vital question remains; given k logical qubits, how to encode the information they contain in N physical qubits? An encoding circuit for a general stabilizer is derived in the following.

4.3.2 Encoding and Decoding

In classical linear coding theory, codes are defined by their generator matrix. Not only does this matrix give insight into the parameters of the code, it also provides an encoding strategy; simple matrix multiplication between the message and generator matrix yields the encoded message. In the quantum case, stabilizer codes are defined in terms of their check matrix; however, this matrix does not directly provide the encoded state by matrix multiplication. Instead, row and column operation can be performed on the check matrix to reveal the encoding circuit. The encoding circuits are derived using the theory of [Wilde, 2008, Sec.3.5].

Let $|\psi\rangle$ belong to some subspace stabilized by the operator A. Then $A|\psi\rangle = |\psi\rangle$, suppose that B is some unitary operator, then

$$BAB^\dagger B|\psi\rangle = BA|\psi\rangle = B|\psi\rangle.$$

Thus, if A stabilizes $|\psi\rangle$, then $BAB^\dagger$ stabilizes $B|\psi\rangle$. This is the foundation behind the encoding strategy for a general stabilizer derived below. Consider the action of the 3 Clifford gates; the Hadamard gate H, the phase gate S, and the CNOT gate

$$HX^aZ^bH^\dagger = X^bZ^a$$
$$SX^aZ^bS^\dagger = X^aZ^{a+b}$$
$$\text{CNOT}(X^aZ^b \otimes X^cZ^d)\text{CNOT}^\dagger = X^aZ^{b+d} \otimes X^{a+c}Z^d$$

The actions are summarized below

H: Performed on qubit i interchanges the Z and X column on qubit i.

S: Performed on qubit i add the X column to the Z column.

CNOT: Performed between qubit i and qubit j adds the X column on qubit i to the X column on qubit j and adds the Z column on qubit j to the Z column on qubit i.

Three CNOT gates in connection can also be used to swap two qubits. This is depicted in figure 4.4.

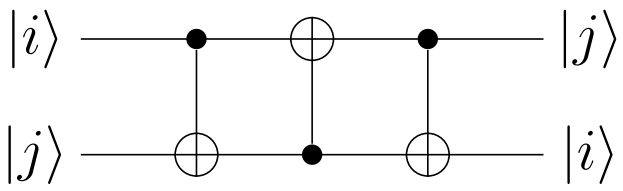


Figure 4.4: Circuit consisting of 3 CNOT gates able to swap two qubits.

To encode a general stabilizer, one simply finds a stabilizer with known coding space and encoding scheme, this stabilizer will be referred to as a stabilizer in standard form, then transforms the associated check matrix of the general stabilizer into the check matrix of the standard form stabilizer using the Clifford operators and swap gate. The operators, in reverse order, then transform the standard coding space into the general coding space. The standard form stabilizer can be found as follows: Let the general stabilizer be defined by $N-k$ independent generators operating on state space $\bigotimes_{i=1}^{N}\mathcal{H}_i$. Define operators on state space $\bigotimes_{i=1}^{N}\mathcal{H}_i$ by

$$Z_i = \bigotimes_{j=1}^{i-1} I \otimes Z \otimes \bigotimes_{k=i+1}^{N} I, \qquad i \in \{1,\ldots,N-k\}. \tag{4.7}$$

These operators then defined a stabilizer with coding space

$$\bigotimes_{i=1}^{N-k} |0\rangle \otimes \bigotimes_{j=1}^{k} |\psi_j\rangle\,.$$

where $|\psi_j\rangle \in \mathcal{H}_{N-k+j}$ is an arbitrary state. The information is thus encoded in $\bigotimes_{j=1}^{k}|\psi_j\rangle$ and $\bigotimes_{i=1}^{N-k}|0\rangle$ are ancilla qubits used for error correction.

Algorithm 1 provides a general method for transforming the general check matrix into the check matrix associated with the stabilizer of (4.7).

Algorithm 1 Encoding Algorithm for Stabilizers

Let $G = \langle g_1, \ldots, g_{N-k} \rangle$ be a stabilizer group.
Denote by $r_i := r(g_i)$ for $i \in \{1, \ldots, N-k\}$ row i of the check matrix cf. (4.5)
for $i = 1, \ldots, N-k$ **do**
    if $r_{i,i} = 0 \wedge r_{i,i+N} = 0$ **then**
        Apply swap gate between qubit i and qubit l for some $i < l \in \{i+1, \ldots, N\}$ with $r_{i,l} = 1 \vee r_{i,l+N} = 1$
    end if
    for $j = i, \ldots, N$ **do**
        if $r_{i,j} = 1 \wedge r_{i,j+N} = 0$ **then**
            Continue
        else if $r_{i,j} = 0 \wedge r_{i,j+N} = 1$ **then**
            Apply Hadamard gate on qubit j
        else if $r_{i,j} = 1 \wedge r_{i,j+N} = 1$ **then**
            Apply phase gate to qubit j
        end if
    end for
    for $j = i+1, \ldots, N$ **do**
        if $r_{i,j} = 1$ **then**
            Apply CNOT gate from qubit i to qubit j
        end if
    end for
    Row i are now of the form $r_{i,j} = \begin{cases} 1, & \text{if } j = i \\ 0, & \text{else} \end{cases}$ for $j \in \{1, \ldots, 2N\}$
    for $j = i+1, \ldots, N-k$ **do**
        if $r_{j,i} = 1$ **then**
            Add row i to row j
        end if
    end for
    Apply Hadamard gate to qubit i
    for $j = i+1, \ldots, N$ **do**
        if $r_{j,N+i} = 1$ **then**
            Add row i to row j
        end if
    end for
end for
The check matrix has now been transformed into standard form

In the following example, the encoding algorithm is used on the stabilizer of section 4.2.

Example 6
Consider the check matrix associated with the stabilizer example of section 4.2 given by

$$\left[\begin{array}{ccc|ccc} 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right].$$

Denote by H_i a Hadamard gate on qubit i and by CNOT(i, j) a CNOT gate from qubit i to qubit j. Following algorithm 1, the matrix is transformed in the following way

$$\left[\begin{array}{ccc|ccc} 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right] \xrightarrow{H_1 H_2} \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow{\text{CNOT}(1,2)} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow{H_1 H_3} \left[\begin{array}{ccc|ccc} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{array}\right] \xrightarrow{\text{CNOT}(2,3)} \left[\begin{array}{ccc|ccc} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{H_2} \left[\begin{array}{ccc|ccc} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right]$$

The encoder is shown in figure 4.5.

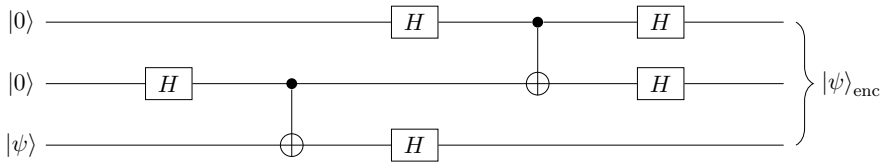


Figure 4.5: The encoder of the stabilizer code of section 4.2.

It should be noted that the encoder is not unique. Several encoders may achieve the same result by using different numbers of gates. Finding an optimal encoder is thus a challenge in itself.

Having established the important concepts of stabilizer codes, the notion of puncturing is examined.

4.3.3 Puncturing

Puncturing is an important concept that has been thoroughly studied for classical error-correcting codes. In a classical error-correcting code, k bits are encoded into N bits, which are then transmitted. Let such a code have distance d, then it is denoted an $[N, k, d]$

code. The rate of such a code is given by $R = k/N$. Suppose one wishes to increase this rate. One way of achieving this is to decrease N while maintaining k; a task made easy by puncturing. Mathematically, this is done by removing one or more columns from the generator matrix, specified by an index set T. Denoting by v_i row i of the generator matrix, such a procedure can be achieved by applying the map π defined by

$$\pi_T(v_i) := [v_{i,j}]_{j \notin T}, \tag{4.8}$$

to each row of the generator matrix. In practice, the encoder can be kept and one simply removes the bits in the positions specified by T. The choice of T must be made carefully, as it can significantly affect d and therefore the error-correcting capability of the code. In a quantum code, the idea is fundamentally the same; decrease the number of physical qubits while maintaining the number of logical qubits. However, the math is a bit more nuanced. Consider the check matrix of a stabilizer code. Since the stabilizer group is abelian $r(g_i)\Lambda r(g_j)^T = 0$ for any two elements g_i and g_j of the stabilizer group where

$$\Lambda = \begin{bmatrix} 0 & I_N \\ I_N & 0 \end{bmatrix}.$$

Denote a row of the check matrix by $r(g_i) = [a_{i,1} \ \dots \ a_{i,N} \ b_{i,1} \ \dots \ b_{i,N}] = [\overline{a}_i \ \overline{b}_i]$, then

$$r(g_i)\Lambda r(g_j)^T = \sum_{k=1}^{N}(a_{i,k}b_{j,k} + b_{i,k}a_{j,k}) = \left\langle \overline{a}_i \middle| \overline{b}_j \right\rangle + \left\langle \overline{b}_i \middle| \overline{a}_j \right\rangle = 0.$$

Define

$$([\overline{a}_i \ \overline{b}_i], [\overline{a}_j \ \overline{b}_j])_S := \left\langle \overline{a}_i \middle| \overline{b}_j \right\rangle + \left\langle \overline{b}_i \middle| \overline{a}_j \right\rangle. \tag{4.9}$$

By this definition $(\cdot,\cdot)_S : \mathbb{F}_2^{2N} \times \mathbb{F}_2^{2N} \to \mathbb{F}_2$ is a symplectic bilinear form, and the stabilizer group is self-orthogonal with respect to this form. Since the symplectic form vanishes for any two commuting elements, $C_{\mathcal{P}_N}(G)$ is orthogonal to G with respect to the symplectic form. Denote by $G_2 \subseteq \mathbb{F}_2^{2N}$ and $G_2^{\perp_S} \subseteq \mathbb{F}_2^{2N}$ the image of G and $C_{\mathcal{P}_N}(G)$ under r, defined by (4.5), respectively. Both sets are vector spaces over $\mathbb{F}_2$ and the image of the generating set of the respective group under r is a basis of the corresponding space. Since $G \subseteq C_{\mathcal{P}_N}(G)$ it implies the inclusion $G_2 \subseteq G_2^{\perp_S} \subseteq \mathbb{F}_2^{2N}$. The check matrix, denoted $C \in \mathbb{F}_2^{N-k\times 2N}$, is the matrix which rows consists of the image of the generating set of the stabilizer under r and the rows thus constitutes a basis of G_2. This basis can be extended to a basis of $G_2^{\perp_S}$ by adding an extra $2k$ elements to the generating set of the stabilizer, such it becomes a generating set for $C_{\mathcal{P}_N}(G)$, then taking the image of this new generating set under r. Denote the matrix which rows consists of the image of the added elements by $C^{\perp_S}\backslash C \in \mathbb{F}_2^{2k\times 2N}$, the centralizer matrix, is then the matrix $C_{\text{ext}} \in \mathbb{F}_2^{N+k\times 2N}$ defined by

$$C_{\text{ext}} := \begin{bmatrix} C \\ C^{\perp_S}\backslash C \end{bmatrix}.$$

This matrix plays a great role when puncturing a stabilizer code. To begin puncturing, choose some $[\alpha\ \beta] \in \mathbb{F}_2^2 \backslash \{[0\ 0]\}$. The stabilizers after puncturing are then given by

$$G_2^{[\alpha\ \beta]} := \{\pi_{\{m\}}^Q([\overline{a}\ \overline{b}]) \mid [\overline{a}\ \overline{b}] \in G_2,\ ([a_m\ b_m], [\alpha\ \beta])_S = 0\}, \qquad m \in \{1, \ldots, N\}, \quad (4.10)$$

where $\pi_{\{m\}}^Q([\overline{a}\ \overline{b}]) := [\pi_{\{m\}}(\overline{a})\ \pi_{\{m\}}(\overline{b})]$ for $\overline{a}, \overline{b} \in \mathbb{F}_2^N$ and π is defined by (4.8). The punctured stabilizers thus corresponds to the stabilizers for which the m'th entry of their G_2 representation is symplectic orthogonal to $[\alpha\ \beta]$, then removing the operator acting on the m'th qubit of these stabilizers.

To show that this indeed corresponds to stabilizers for a stabilizer code with parameters $[[N-1, k, d']]$ it remains to be argued that $G_2^{[\alpha\ \beta]}$ is the image under r of an abelian subgroup of $\mathcal{P}_{N-1}$, denoted $G^{[\alpha\ \beta]}$, not containing $-I$ and which can be generated by an independent generating set of $(N-1)-k$ elements.

The fact that $G^{[\alpha\ \beta]}$ is a subset of $\mathcal{P}_{N-1}$ is directly evident from (4.10), since the m'th operator of the stabilizers are removed. Furthermore, $-I \notin G^{[\alpha\ \beta]}$ can be satisfied by choice. The puncturing is done to the G_2 representation of the stabilizers which contains no information on the phase, thus, after puncturing one simply chooses the phase on the punctured stabilizers. Since any stabilizer is hermitian, only ± 1 is a valid phase choice, simply choosing $+1$ for the element representing I satisfies the condition. As for the other elements, the choice only affects the subspace containing the logical states, not the properties of the code.
Regarding the last two statements, recall that $G^{[\alpha\ \beta]}$ is abelian if and only if $G_2^{[\alpha\ \beta]}$ is self-orthogonal with respect to the symplectic form of (4.9). Moreover, the image under r of a generating set for $G^{[\alpha\ \beta]}$ forms a basis for $G_2^{[\alpha\ \beta]}$. Consequently, if $G^{[\alpha\ \beta]}$ is a subspace of dimension $(N-1)-k = \dim(G_2) - 1$, then $G^{[\alpha\ \beta]}$ has a generating set consisting of $(N-1)-k$ elements.
Before proving the first statement, consider the following extension of an element $\overline{e} \in \mathbb{F}_2^M$ for some $M \in \mathbb{N}$ given by

$$[\overline{e}|e_m]_i = \begin{cases} [\overline{e}]_i, & 1 \leq i < m, \\ e_m, & i = m, \\ [\overline{e}]_{i-1}, & m < i \leq M. \end{cases}$$

Then $[\overline{e}|e_m] \in \mathbb{F}_2^{M+1}$ with the added element being e_m in the m'th entry.

Theorem 4.9
Let $G_2^{[\alpha\ \beta]}$ be defined by (4.10). Then $G_2^{[\alpha\ \beta]}$ is a self-orthogonal with respect to the symplectic form of (4.9).

[Gundersen et al., 2025, p.4]

Proof
Let $[\overline{a}\ \overline{b}], [\overline{a}'\ \overline{b}'] \in G_2^{[\alpha\ \beta]}$. Then by definition of $G_2^{[\alpha\ \beta]}$ there exist $a_m, b_m, a'_m, b'_m \in \mathbb{F}_2$ such

$$\pi_{\{m\}}^Q([\overline{a}|a_m\ \overline{b}|b_m]) = [\overline{a}\ \overline{b}] \quad \wedge \quad ([a_m\ b_m], [\alpha\ \beta])_S = 0$$

and

$$\pi^Q_{\{m\}}([\overline{a}'|a'_m\ \overline{b}'|b'_m]) = [\overline{a}'\ \overline{b}'] \quad \wedge \quad ([a'_m\ b'_m],[\alpha\ \beta])_S = 0.$$

It can be verified that $([e_1\ e_2],[\alpha\ \beta])_S = 0$ if and only if $[e_1\ e_2] = c[\alpha\ \beta]$ for some $[e_1\ e_2],[\alpha\ \beta] \in \mathbb{F}_2^2$ and $c \in \mathbb{F}_2$. It must therefore be the case that $[a_m\ b_m] = c_1[\alpha\ \beta]$ and $[a'_m\ b'_m] = c_2[\alpha\ \beta]$, hence $([a_m\ b_m],[a'_m\ b'_m])_S = 0$. It then follows

$$\begin{aligned}([\overline{a}\ \overline{b}],[\overline{a}'\ \overline{b}'])_S &= ([a_m\ b_m],[a'_m\ b'_m])_S + ([\overline{a}\ \overline{b}],[\overline{a}'\ \overline{b}'])_S \\ &= ([\overline{a}|a_m\ \overline{b}|b_m],[\overline{a}'|a'_m\ \overline{b}'|b'_m])_S = 0.\end{aligned}$$

The last equality follows from $[\overline{a}|a_m\ \overline{b}|b_m],[\overline{a}'|a'_m\ \overline{b}'|b'_m] \in G_2$. ∎

Before proving the second statement, consider the linear map $\sigma_{[\alpha\ \beta]} : G_2 \to \mathbb{F}_2$ defined by

$$\sigma_{[\alpha\ \beta]}([\overline{a}|a_m\ \overline{b}|b_m]) := ([a_m\ b_m],[\alpha\ \beta])_S, \qquad [\overline{a}|a_m\ \overline{b}|b_m] \in G_2.$$

Puncturing can then be expressed by $\pi^Q_{\{m\}}(\ker(\sigma_{[\alpha\ \beta]}))$. The second statement now depends on d as the following theorem shows.

Theorem 4.10
Assume $d \geq 2$ and $[\alpha\ \beta] \neq [0\ 0]$. Then $\dim\left(G_2^{[\alpha\ \beta]}\right) = \dim(G_2) - 1$.

[Gundersen et al., 2025, p.4]

Proof
Define the restriction of $\pi^Q_{\{m\}}$ to $\ker(\sigma_{[\alpha\ \beta]})$ by $\tilde{\pi}^Q_{\{m\}} := \pi^Q_{\{m\}} \restriction_{\ker(\sigma_{[\alpha\ \beta]})}$. Then according to the rank-nullity theorem for the restricted map

$$\begin{gathered}\dim(\mathrm{im}(\tilde{\pi}^Q_{\{m\}})) = \dim(\mathrm{domain}(\tilde{\pi}^Q_{\{m\}})) - \dim(\ker(\tilde{\pi}^Q_{\{m\}})) \\ \Leftrightarrow \\ \dim(G_2^{[\alpha\ \beta]}) = \dim(\ker(\sigma_{[\alpha\ \beta]})) - \dim(\ker(\tilde{\pi}^Q_{\{m\}})).\end{gathered}$$

Since $[\overline{a}\ \overline{b}] \in \ker(\sigma_{[\alpha\ \beta]})$ implies $([a_m\ b_m],[\alpha\ \beta])_S = 0$ which is the case if and only if $[a_m\ b_m] = c[\alpha\ \beta]$ for some $c \in \mathbb{F}_2$ thus any $[\overline{a}\ \overline{b}] \in \ker(\tilde{\pi}^Q_{\{m\}})$ must be of the form $[\overline{0}|a_m\ \overline{0}|b_m] = c[\overline{0}|\alpha_m\ \overline{0}|\beta_m]$, where $\overline{0}$ is the zero-vector. Consequently, $\ker(\tilde{\pi}^Q_{\{m\}})$ is at most 1 dimensional. Define by $G_2^m := \{[a_m\ b_m] \mid [\overline{a}\ \overline{b}] \in G_2\}$, that is, the restriction of G_2 to the m'th coordinate. Two mutually exclusive cases now arise; either $\mathrm{span}(G_2^m) \subseteq \mathrm{span}(\{[\alpha\ \beta]\})$ or $\mathrm{span}(G_2^m) \nsubseteq \mathrm{span}(\{[\alpha\ \beta]\})$.
In the first case $[a_m\ b_m] = c[\alpha\ \beta]$ for all $[\overline{a}\ \overline{b}] \in G_2$ implying $\ker(\sigma_{[\alpha\ \beta]}) = G_2$, but then $[\overline{0}|\alpha_m\ \overline{0}|\beta_m] \in G_2^{\perp_S}$. Obviously, $[\overline{0}|\alpha_m\ \overline{0}|\beta_m]$ corresponds to a weight 1 Pauli operator and can thus by assumption not be in $G_2^{\perp_S} \backslash G_2$ meaning it belongs to G_2. Since $[\overline{0}|\alpha_m\ \overline{0}|\beta_m] \in \ker(\tilde{\pi}^Q_{\{m\}})$ it is of dimension 1.
In the second case there exist at least one element $[\overline{a}\ \overline{b}] \in G_2$ such $[a_m\ b_m] \neq c[\alpha\ \beta]$ implying $([a_m\ b_m],[\alpha\ \beta])_S \neq 0$. This means $\sigma_{[\alpha\ \beta]}$ is not the zero-map on G_2. Applying the

rank-nullity theorem to the linear functional $\sigma_{[\alpha\ \beta]}$ yields $\dim(\ker(\sigma_{[\alpha\ \beta]}) = \dim(G_2) - 1$. It thus only remains to show that the kernel of $\tilde{\pi}^Q_{\{m\}}$ is trivial. It has already been established that elements in $\ker(\tilde{\pi}^Q_{\{m\}})$ are of the form $c[\overline{0}|\alpha_m\ \overline{0}|\beta_m]$. Pick an element $[\overline{a}\ \overline{b}] \in G_2$ such $[a_m\ b_m] \neq c[\alpha\ \beta]$ then $([a_m\ b_m], [\alpha\ \beta])_S \neq 0$. Since G_2 is symplectic self-orthogonal

$$0 = ([\overline{a}\ \overline{b}], [\overline{0}|c\alpha_m\ \overline{0}|c\beta_m])_S = c([a_m\ b_m], [\alpha\ \beta])_S.$$

This can only be satisfied if $c = 0$ thus implying $\ker(\tilde{\pi}^Q_{\{m\}}) = \{\overline{0}\}$. ∎

Based on the previous two theorems, the method of (4.10) indeed describes a valid stabilizer code. The punctured stabilizer encodes k logical qubits in $N - 1$ physical qubits. However, it remains to study the effect of puncturing on the distance of the code. This information lies in $(G_2^{[\alpha\ \beta]})^{\perp_S} \backslash G_2^{[\alpha\ \beta]}$, that is, the centralizer excluding the stabilizer of the code. The next theorem eases the study of this set.

Theorem 4.11
Let $G_2^{[\alpha\ \beta]}$ be defined by (4.10). Then

$$(G_2^{[\alpha\ \beta]})^{\perp_S} = (G_2^{\perp_S})^{[\alpha\ \beta]}.$$

[Gundersen et al., 2025, p.4]

For proof see [Gundersen et al., 2025, p.4].

The centralizer of the punctured code is found simply by puncturing the centralizer of the unpunctured code. With this in mind, the dimension of the punctured centralizer can be determined.

Corollary 4.12
Let $d \geq 2$ and $[\alpha\ \beta] \neq [0\ 0]$. Then

$$\dim((G_2^{\perp_S})^{[\alpha\ \beta]}) = \dim(G_2^{\perp_S}) - 1.$$

[Gundersen et al., 2025, p.5]

Proof
The proof follows directly from theorem 4.10 and theorem 4.11, together with the dimension identity for a subspace and its symplectic complement.

$$\begin{aligned}\dim((G_2^{\perp_S})^{[\alpha\ \beta]}) &= \dim((G_2^{[\alpha\ \beta]})^{\perp_S}) \\ &= 2(N-1) - \dim(G_2^{[\alpha\ \beta]}) \\ &= 2N - 2 - \dim(G_2) + 1 \\ &= \dim(G_2^{\perp_S}) - 1.\end{aligned}$$

∎

From theorem 4.11 it is evident no new elements are introduced into the centralizer when puncturing. Some are removed and those kept have their operator on the m'th qubit removed. According to theorem 4.10 and corollary 4.12 it is also the case that all elements of $(G_2^{\perp_S})^{[\alpha\ \beta]} \backslash G_2^{[\alpha\ \beta]}$ is a punctured version of some element from $G_2^{\perp_S} \backslash G_2$. The distance of the punctured code, d', can thus at most decrease by 1, i.e. $d' \geq d-1$, where d is the distance of the unpunctured code. The lower bound is attained if and only if an element of weight d belonging to $G_2^{\perp_S} \backslash G_2$ survives the puncturing and acts non-trivially on the m'th qubit. The next theorem shows that choosing $[\alpha\ \beta]$ carefully may lead to a non-decrease in distance when puncturing.

Lemma 4.13
Let G_2 define an $[[N,k,d]]$ stabilizer code with $d \geq 2$. Assume $[a_m\ b_m] \neq [\alpha\ \beta]$ for all $[\overline{a}|a_m\ \overline{b}|b_m] \in G_2^{\perp_S} \backslash G_2$. Then puncturing on the m'th qubit with respect to $[\alpha\ \beta]$ yields a code with distance $d' \geq d$.

[Gundersen et al., 2025, p.6]

Proof
Denote by $[\overline{a}|a_m\ \overline{b}|b_m] \in G_2^{\perp_S} \backslash G_2$ a weight d element. Two cases can now occur. Either $a_m = b_m = 0$ or $a_m = 1 \vee b_m = 1$. In the first case $[\overline{a}\ \overline{b}] \in G_2^{[\alpha\ \beta]}$ since $([a_m\ b_m], [\alpha\ \beta])_S = 0$ for any $[\alpha\ \beta] \neq [0\ 0]$, but puncturing this element does not change its weight. Such an element does therefore not decrease the distance of the punctured code. In the second case, let $[\overline{a}'|a'_m\ \overline{b}'|b'_m] \in \ker(\sigma_{[\alpha\ \beta]}) \subseteq G_2$, that is, $([a'_m\ b'_m], [\alpha\ \beta])_S = 0$, then since $[a_m\ b_m] \neq [\alpha\ \beta]$ it follows $([a_m\ b_m], [a'_m\ b'_m])_S \neq 0$. But $([\overline{a}|a_m\ \overline{b}|b_m], [\overline{a}'|a'_m\ \overline{b}'|b'_m])_S = 0$ and thus by linearity of the symplectic form it follows $([\overline{a}\ \overline{b}], [\overline{a}'\ \overline{b}'])_S \neq 0$. However, $[\overline{a}'\ \overline{b}'] \in \pi^Q_{\{m\}}(\ker(\sigma_{[\alpha\ \beta]})) = G_2^{[\alpha\ \beta]}$ and it must therefore be the case that $[\overline{a}\ \overline{b}] \notin (G_2^{[\alpha\ \beta]})^{\perp_S}$. This means all elements of $G_2^{\perp_S} \backslash G_2$ of weight d with $[a_m\ b_m] \neq [0\ 0]$ are removed during puncturing. Combining the two cases it can be concluded that $d' \geq d$. ∎

The method of (4.10) requires evaluating all of G_2, to determine whether they are removed or kept. A preferable method would involve only the generators of the stabilizer group and would then be performed directly on the check- or centralizer matrix. A simple example shows such a method is not as straight forward as one would hope. Let $[\overline{a}|a_m\ \overline{b}|b_m], [\overline{a}'|a'_m\ \overline{b}'|b'_m] \in G_2$ and assume $([a_m\ b_m], [\alpha\ \beta])_S = ([a'_m\ b'_m], [\alpha\ \beta])_S = 1$ for some $[\alpha\ \beta] \neq [0\ 0]$. Puncturing with respect to $[\alpha\ \beta]$ would discard $[\overline{a}\ \overline{b}], [\overline{a}'\ \overline{b}']$, that is, $[\overline{a}\ \overline{b}], [\overline{a}'\ \overline{b}'] \notin G_2^{[\alpha\ \beta]}$, however,

$$([a_m + a'_m\ b_m + b'_m], [\alpha\ \beta])_S = ([a_m\ b_m], [\alpha\ \beta])_S + ([a'_m\ b'_m], [\alpha\ \beta])_S = 1 + 1 = 0.$$

This means $[\overline{a} + \overline{a}'\ \overline{b} + \overline{b}'] \in G_2^{[\alpha\ \beta]}$. Simply applying the method of (4.10) directly to a generating set does not, in general, produce a generating set for the punctured code, as this may remove generators whose linear combinations satisfy the puncturing condition. Instead, since the rows of the centralizer matrix constitutes a basis of $G_2^{\perp_S}$ exactly one row should be removed according to theorem 4.10 and corollary 4.12. Performing appropriate row operations yields a basis in which all but a single stabilizer generator satisfy the puncturing condition. This procedure is summarized in the following algorithm.

Algorithm 2 Puncturing on the Check Matrix

Denote centralizer matrix by $C_{\text{ext}} \in \mathbb{F}_2^{N+k\times 2N}$
Denote row i of C_{ext} by $r_i = [\overline{a}^{(i)}\ \overline{b}^{(i)}]$ for $i \in \{1,\dots,N+k\}$
Initialize $l = 0$ and $j = 0$
for $i = 1,\dots,N-k$ **do**
    Compute $\sigma(r_i) = ([a_1^{(i)}\ b_1^{(i)}], [\alpha\ \beta])_S$
    if $\sigma(r_i) = 1$ **then**
        $l \leftarrow l + 1$
        if $l = 1$ **then**
            Remove row i from C_{ext} and set $j = i$
        else
            $r_i \leftarrow r_i + r_j$
        end if
    else
        Continue
    end if
end for
Return $C_{\text{ext}} \in \mathbb{F}_2^{N+k-1\times 2N}$

The puncturing algorithm is applied in a simple example below.

Example 7: Puncturing the Five Qubit Code
Consider the $[[5,1,3]]$ code given by the stabilizers and corresponding centralizer matrix shown below.

g_1	$X \otimes Z \otimes Z \otimes X \otimes I$
g_2	$I \otimes X \otimes Z \otimes Z \otimes X$
g_3	$X \otimes I \otimes X \otimes Z \otimes Z$
g_4	$Z \otimes X \otimes I \otimes X \otimes Z$

(a) Stabilizers for the $[[5,1,3]]$ code.

$$\left[\begin{array}{ccccc|ccccc} 1&0&0&1&0&0&1&1&0&0\\ 0&1&0&0&1&0&0&1&1&0\\ 1&0&1&0&0&0&0&0&1&1\\ 0&1&0&1&0&1&0&0&0&1\\ \hline 1&0&0&0&1&0&0&1&0&0\\ 0&0&1&0&0&0&1&0&1&0 \end{array}\right]$$

(b) Corresponding 6×10 centralizer matrix.

Let $[\alpha\ \beta] = [0\ 1]$ and $\{m\} = \{1\}$, that is, puncturing is performed on the first qubit. Then

according to algorithm 2, the centralizer matrix of the punctured code is found by

$$\left[\begin{array}{ccccc|ccccc} 1&0&0&1&0&0&1&1&0&0\\ 0&1&0&0&1&0&0&1&1&0\\ 1&0&1&0&0&0&0&0&1&1\\ 0&1&0&1&0&1&0&0&0&1\\ \hline 1&0&0&0&1&0&0&1&0&0\\ 0&0&1&0&0&0&1&0&1&0 \end{array}\right] \xrightarrow{\text{Algorithm 2}} \left[\begin{array}{ccccc|ccccc} 1&0&0&1&0&0&1&1&0&0\\ 0&1&0&0&1&0&0&1&1&0\\ 0&0&1&1&0&0&1&1&1&1\\ 0&1&0&1&0&1&0&0&0&1\\ \hline 0&0&0&1&1&0&1&0&0&0\\ 0&0&1&0&0&0&1&0&1&0 \end{array}\right]$$

The new centralizer matrix and corresponding stabilizer group are summarized below.

g_2	$X \otimes Z \otimes Z \otimes X$
g_3	$Z \otimes Y \otimes Y \otimes Z$
g_4	$X \otimes I \otimes X \otimes Z$

(a) Stabilizers for the punctured code.

$$\left[\begin{array}{cccc|cccc} 1&0&0&1&0&1&1&0\\ 0&1&1&0&1&1&1&1\\ 1&0&1&0&0&0&0&1\\ \hline 0&0&1&1&1&0&0&0\\ 0&1&0&0&1&0&1&0 \end{array}\right]$$

(b) Corresponding 3×8 check matrix.

The punctured code has parameters $[[4, 1, 2]]$. The distance has thus been reduced by 1 and while the unpunctured code could protect against any weight 1 error, there exist weight 1 errors against which the punctured code cannot protect. An example is $E = Z \otimes I \otimes I \otimes I$ and $E' = I \otimes I \otimes Y \otimes I$. Both of these anti-commutes with the first and third generator, but commutes with the second generator and thus produces the same syndrome. However, $r(E) + r(E') \in (G_2^{[0\ 1]})^{\perp_S} \backslash G_2^{[0\ 1]}$ meaning an error-correction-strategy correcting E would fail to correct E' and vice versa.

The puncturing procedure described above removes a single physical qubit. To remove multiple qubits, the procedure can be applied iteratively: After puncturing the first qubit, one can apply the puncturing method to the already punctured code. Repeating this procedure j times yields a stabilizer code on $N - j$ physical qubits while preserving the number of logical qubits, k. For this iterative method to make the transformation

$$[[N, k, d]] \text{ code} \quad \xrightarrow{j \text{ puncturings}} \quad [[N - j, k, d']] \text{ code},$$

it relies on $d' \geq 2$. If performing the puncturing procedure on a code for which $d = 1$, the parameters are not as easily determined. In such a case it may lead to the loss of not only physical qubits but also logical qubits. The reader is referred to [Gundersen et al., 2025] for a more extensive analysis of such cases.

5 | Entanglement-Assisted Quantum Error-Correcting Codes

Stabilizer codes can be constructed directly from classical linear codes with an added constraint of self-orthogonality. This added constraint greatly limits the number of classical linear codes admissible in the construction of stabilizer codes. The constraint, however, can be overcome by the introduction of entanglement. This chapter is primarily based on [Brun et al., 2006].

As previously mentioned, a stabilizer code is defined by an abelian subgroup, $G = \langle g_1, \ldots, g_{N-k} \rangle$, of $\mathcal{P}_N$ with independent generators and such $-I \notin G$. A stabilizer code can similarly be defined by its corresponding check matrix. The check matrix, $C \in \mathbb{F}_2^{N-k \times 2N}$, has linearly independent rows defined by (4.5) and is self-orthogonal with respect to the symplectic form of (4.9). Considering this, constructing a stabilizer code simplifies to constructing a matrix $C \in \mathbb{F}_2^{N-k \times 2N}$ with the desired properties of linear independence and self-orthogonality. One way of finding such a matrix is through the use of classical linear codes. These are defined by a generator matrix, or equivalently by a parity check matrix. For a classical linear code with parameters $[N, k, d]$, the parity check matrix is of size $N - k \times N$ with full rank. The parity check matrix of a $[2N, N + k]$ classical linear code thus defines a stabilizer code with parameters $[[N, k]]$ if the rows of the parity check matrix are self-orthogonal. The search for a parity check matrix that meets this requirement among all classical linear codes with parameters $[2N, N + k]$ is very impractical. However, introducing a number of entangled bits between the sender and the receiver makes it possible to remove this constraint.

To facilitate a rigorous justification of how the distribution of entanglement between a sender and a receiver relaxes the self-orthogonality constraint, the notion of symplectic vector spaces is introduced. Let the set $\{\overline{a}_1, \ldots, \overline{a}_n, \overline{b}_1, \ldots, \overline{b}_m\}$ be a basis of some vector space V. The basis is called symplectic if

$$(\overline{a}_i, \overline{a}_j)_S = (\overline{b}_i, \overline{b}_j)_S = 0 \qquad \text{and} \qquad (\overline{a}_i, \overline{b}_j)_S = \delta_{ij}.$$

The following theorem provides a procedure for transforming any basis of a vector space into a symplectic basis.

Theorem 5.1
Let V be an $N-k'$ dimensional subspace of $\mathbb{F}_2^{2N}$ for $N, k' \in \mathbb{N}$ with $N \geq k'$. Then there exists a symplectic basis of $\mathbb{F}_2^{2N}$ given by $\{\overline{a}_1, \ldots, \overline{a}_N, \overline{b}_1, \ldots, \overline{b}_N\}$ such the subset $\{\overline{a}_1, \ldots, \overline{a}_{s+c}, \overline{b}_1, \ldots, \overline{b}_c\}$ with $2c + s = N - k'$ for some $c, s \in \mathbb{N}_0$ forms a basis of V. The subspace $\mathrm{symp}(V) := \mathrm{span}\{\overline{a}_1, \ldots, \overline{a}_c, \overline{b}_1, \ldots, \overline{b}_c\}$ is symplectic and $\mathrm{iso}(V) := \mathrm{span}\{\overline{a}_{1+c}, \ldots, \overline{a}_{s+c}\}$ is isotropic.

Proof
Let $\{\overline{c}_1, \ldots, \overline{c}_{N-k'}\}$ be an arbitrary basis for V. Extend the basis to a basis of $\mathbb{F}_2^{2N}$ denoted $\{\overline{c}_1, \ldots, \overline{c}_{N-k'}, \overline{c}_{N-k'+1}, \ldots, \overline{c}_{2N}\}$. This basis of $\mathbb{F}_2^{2N}$ is then inductively transformed into a symplectic basis in N cycles; one pair at a time.
To initialize, set $i = 1$, $m = N - k'$ and $\mathcal{A} = \mathcal{B} = \varnothing$. The i'th cycle then goes as follows: From the previous cycles, the basis has been transformed into
$\{\overline{c}_1, \ldots, \overline{c}_{2(N-i+1)}, \overline{a}_1, \ldots, \overline{a}_{i-1}, \overline{b}_1, \ldots, \overline{b}_{i-1}\}$. For this set it holds true that

1. $\{\overline{c}_1, \ldots, \overline{c}_{2(N-i+1)}, \overline{a}_1, \ldots, \overline{a}_{i-1}, \overline{b}_1, \ldots, \overline{b}_{i-1}\}$ is a basis of $\mathbb{F}_2^{2N}$
2. $(\overline{a}_j, \overline{c}_q)_S = (\overline{b}_j, \overline{c}_q)_S = 0$ for $j \in \{1, \ldots, i-1\}$ and $q \in \{1, \ldots, 2(N-i+1)\}$
3. $V = \mathrm{span}\{\overline{c}_j \mid j \in \{1, \ldots, m\}\} \oplus \mathrm{span}\{\overline{a}_j \mid j \in \mathcal{A}\} \oplus \mathrm{span}\{\overline{b}_j \mid j \in \mathcal{B}\}$

Obviously, these are all satisfied for $i = 1$ when initializing the induction, thus let $i \in \{1, \ldots, N\}$. The proof is then a matter of showing that a pair of elements of $\{\overline{c}_1, \ldots, \overline{c}_{2(N-i+1)}\}$ can be transformed into $(\overline{a}_i, \overline{b}_i)$ while still satisfying the three items. Define $\overline{a}_i := \overline{c}_1$. If $m \geq 1$ add i to $\mathcal{A}$. Since the symplectic form is non-degenerate and $\overline{a}_i$ is not the zero vector, there exists $\overline{z} \in \mathbb{F}_2^{2N}$ such $(\overline{a}_i, \overline{z})_S = 1$. Express this $\overline{z}$ in the basis given in item 1, then

$$\begin{aligned}(\overline{a}_i = \overline{c}_1, \overline{z})_S &= \sum_{q=1}^{2(N-i+1)} \alpha_q(\overline{c}_1, \overline{c}_q)_S + \sum_{q=1}^{i-1} \beta_q(\overline{c}_1, \overline{a}_q)_S + \sum_{q=1}^{i-1} \gamma_q(\overline{c}_1, \overline{b}_q)_S \\ &= \sum_{q=1}^{2(N-i+1)} \alpha_q(\overline{c}_1, \overline{c}_q)_S = 1, \qquad \alpha_q, \beta_q, \gamma_q \in \{0, 1\}.\end{aligned}$$

This implies the existence of at least a single $q \geq 2$ such $(\overline{a}_i, \overline{c}_q)_S = 1$. Denote by j the smallest of such q's and define $\overline{b}_i := \overline{c}_j$. Two cases can now occur.

If $j \leq m$, then add i to $\mathcal{B}$ and swap $\overline{c}_2$ with $\overline{c}_j$. For $q = 3, \ldots, 2(N-i+1)$ define

$$\overline{c}'_{q-2} := \overline{c}_q - (\overline{b}_i, \overline{c}_q)_S \overline{a}_i - (\overline{a}_i, \overline{c}_q)_S \overline{b}_i. \tag{5.1}$$

After this linear transformation

$$(\overline{c}'_{q-2}, \overline{a}_i)_S = (\overline{c}_q, \overline{a}_i)_S - (\overline{b}_i, \overline{c}_q)_S(\overline{a}_i, \overline{a}_i)_S - (\overline{a}_i, \overline{c}_q)_S(\overline{b}_i, \overline{a}_i)_S = 0 = (\overline{c}'_{q-2}, \overline{b}_i)_S. \tag{5.2}$$

Furthermore, for $l \in \{1, \ldots, i-1\}$

$$(\overline{c}'_{q-2}, \overline{a}_l)_S = (\overline{c}_q, \overline{a}_l)_S - (\overline{b}_i, \overline{c}_q)_S(\overline{a}_i, \overline{a}_l)_S - (\overline{a}_i, \overline{c}_q)_S(\overline{b}_i, \overline{a}_l)_S = 0 = (\overline{c}'_{q-2}, \overline{b}_i)_S.$$

Item 2 is therefore still satisfied for $\{\overline{a}_1, \ldots, \overline{a}_i, \overline{b}_1, \ldots, \overline{b}_i\}$ and $\{\overline{c}'_1, \ldots, \overline{c}'_{2(N-i)}\}$. The linear transform of (5.1) from $\{\overline{c}_3, \ldots, \overline{c}_{2(N-i+1)}\}$ to $\{\overline{c}'_1, \ldots, \overline{c}'_{2(N-i)}\}$ is invertible and thus

$\text{span}\{\overline{c}_1, \overline{c}_2, \overline{c}_3, \ldots, \overline{c}_{2(N-i+1)}\} = \text{span}\{\overline{a}_i = \overline{c}_1, \overline{b}_i = \overline{c}_2, \overline{c}'_1, \ldots, \overline{c}'_{2(N-i)}\}$ implying item 1 is satisfied for $\{\overline{c}'_1, \ldots, \overline{c}'_{2(N-i)}, \overline{a}_1, \ldots, \overline{a}_i, \overline{b}_1, \ldots, \overline{b}_i\}$. For this case set $m := m - 2$.

If $j > m$, then swap $\overline{c}_j$ with $\overline{c}_{2(N-i+1)}$ and for $q = 2, \ldots, 2(N-i)+1$ define

$$\overline{c}'_{q-1} := \overline{c}_q - (\overline{b}_i, \overline{c}_q)_S \overline{a}_i - (\overline{a}_i, \overline{c}_q)_S \overline{b}_i \tag{5.3}$$

By similar argumentation to (5.2) it follows

$$(\overline{c}'_{q-1}, \overline{a}_i)_S = (\overline{c}'_{q-1}, \overline{b}_i)_S = 0$$

and

$$(\overline{c}'_{q-1}, \overline{a}_l)_S = (\overline{c}'_{q-1}, \overline{b}_l)_S = 0, \qquad l \in \{1, \ldots, i-1\}.$$

Together these equations imply item 2 is satisfied for $\{\overline{a}_1, \ldots, \overline{a}_i, \overline{b}_1, \ldots, \overline{b}_i\}$ and $\{\overline{c}_1, \ldots, \overline{c}_{2(N-i)}\}$. With the same argumentation as in the first case, item 1 is also satisfied for
$\{\overline{c}'_1, \ldots, \overline{c}'_{2(N-i)}, \overline{a}_1, \ldots, \overline{a}_i, \overline{b}_1, \ldots, \overline{b}_i\}$. In this case if $m \geq 1$ set $m := m - 1$ otherwise leave $m = 0$.

It remains to argue that item 3 is satisfied for $\{\overline{c}'_1, \ldots, \overline{c}'_{2(N-i)}, \overline{a}_1, \ldots, \overline{a}_i, \overline{b}_1, \ldots, \overline{b}_i\}$ in both cases. If $m = 0$ in the beginning of the cycle, $\mathcal{A}$ and $\mathcal{B}$ remains unchanged, item 3 are thus satisfied from the previous cycle. If $m \geq 1$, then in the first case the linear transformation between $\{\overline{c}_1, \ldots, \overline{c}_m\}$ and $\{\overline{a}_i, \overline{b}_i, \overline{c}'_1, \ldots, \overline{c}'_{m-2}\} \subset V$ is invertible thus satisfying item 3. In the second case, since $(\overline{b}_i, \overline{c}_q)_S = 0$ for $q \in \{1, \ldots, m\}$, the linear transformation of (5.3) restricted to $\{\overline{c}_2, \ldots, \overline{c}_m\}$ has codomain $\text{span}\{\overline{a}_i = \overline{c}_1, \overline{c}_2, \ldots, \overline{c}_m\} \subset V$ and is invertible thus satisfying item 3.

By defining $\overline{a}_i$ and $\overline{b}_i$ this way it is the case that

$$(\overline{a}_q, \overline{a}_l)_S = (\overline{b}_q, \overline{b}_l)_S = 0 \qquad \text{and} \qquad (\overline{a}_q, \overline{b}_l)_S = \delta_{ql}, \qquad q, l \in \{1, \ldots, i\}$$

This means $\{\overline{a}_1, \ldots, \overline{a}_i, \overline{b}_1, \ldots, \overline{b}_i\}$ is a symplectic basis for $\text{span}\{\overline{a}_1, \ldots, \overline{a}_i, \overline{b}_1, \ldots, \overline{b}_i\}$.

It is clear that from the beginning $m \leq 2N$, where m denotes the number of elements in the basis of $\mathbb{F}_2^{2N}$ also part of the basis of V. Thus $0 \leq m \leq 2(N-i)$ meaning after at most N cycles $V = \text{span}\{\overline{a}_j \mid j \in \mathcal{A}\} \oplus \text{span}\{\overline{b}_j \mid j \in \mathcal{B}\}$ after reordering the elements of $\{\overline{a}_1, \ldots, \overline{a}_N\}$ and $\{\overline{b}_1, \ldots, \overline{b}_N\}$ the theorem is satisfied.

Regarding the isotropic statement of the theorem. The subspace $\text{iso}(V) := \text{span}\{\overline{a}_{1+c}, \ldots, \overline{a}_{s+c}\}$ is isotropic if and only if $(\cdot,\cdot)_S \restriction_{\text{iso}(V)}$: $\text{iso}(V) \times \text{iso}(V) \to \mathbb{F}_2$ is the zero-map. Since $\{\overline{a}_1, \ldots, \overline{a}_N, \overline{b}_1, \ldots, \overline{b}_N\}$ is a symplectic basis $(\overline{a}_q, \overline{a}_l)_S = 0$ for $q, l \in \{1, \ldots, N\}$, the statement then follows by linearity of the symplectic form.
For the symplectic part. The subspace $\text{symp}(V) := \text{span}\{\overline{a}_1, \ldots, \overline{a}_c, \overline{b}_1, \ldots, \overline{b}_c\}$ is symplectic if and only if $(\cdot,\cdot)_S \restriction_{\text{symp}(V)}$: $\text{symp}(V) \times \text{symp}(V) \to \mathbb{F}_2$ is non-degenerate. Let $\overline{x} = \sum_{l=1}^{c} \alpha_l \overline{a}_l + \sum_{l=1}^{c} \beta_l \overline{b}_l$ and assume $(\overline{y}, \overline{x})_S = 0$ for all $\overline{y} \in \text{symp}(V)$. Thus for $q \in \{1, \ldots, c\}$ let $\overline{y} = \overline{b}_q$ then

$$0 = (\overline{x}, \overline{y} = \overline{b}_q)_S = \sum_{l=1}^{c} \alpha_l (\overline{a}_l, \overline{b}_q)_S + \sum_{l=1}^{c} \beta_l (\overline{b}_l, \overline{b}_q)_S = \alpha_q.$$

Similarly for $q \in \{1, \dots, c\}$ let $\overline{y} = \overline{a}_q$

$$0 = (\overline{x}, \overline{y} = \overline{a}_q)_S = \sum_{l=1}^{c} \alpha_l (\overline{a}_l, \overline{a}_q)_S + \sum_{l=1}^{c} \beta_l (\overline{b}_l, \overline{a}_q)_S = \beta_q.$$

Together, these equations imply $\overline{x}$ is the zero vector. ∎

Let $G = \langle g_1, \dots, g_{N-k'} \rangle \subseteq \mathcal{P}_N$ be a group not containing $-I$. Theorem 5.1 implies the existence of a generating set $\{\tilde{Z}_1, \dots, \tilde{Z}_{s+c}, \tilde{X}_1, \dots, \tilde{X}_c\}$ such that $G = \langle \tilde{Z}_1, \dots, \tilde{Z}_{s+c}, \tilde{X}_1, \dots, \tilde{X}_c \rangle$ for some $c, s \in \mathbb{N}_0$ with $2c + s = N - k'$ and with the commutation relations:

$$\begin{aligned} [\tilde{Z}_i, \tilde{Z}_j] &= 0, \qquad i, j \in \{1, \dots, s+c\}, \\ [\tilde{X}_i, \tilde{X}_j] &= 0, \qquad i, j \in \{1, \dots, c\}, \\ [\tilde{Z}_i, \tilde{X}_j] &= 0, \qquad i \neq j, \\ \{\tilde{Z}_i, \tilde{X}_i\} &= 0, \qquad i \in \{1, \cdots, c\}. \end{aligned} \tag{5.4}$$

The subgroup $\langle \tilde{Z}_1, \dots, \tilde{Z}_c, \tilde{X}_1, \dots, \tilde{X}_c \rangle$ is called the symplectic subgroup and $\langle \tilde{Z}_{c+1}, \dots, \tilde{Z}_{c+s} \rangle$ the isotropic subgroup. Considering the commutation relations, the isotropic subgroup is abelian and hence a stabilizer group. The symplectic subgroup (and hence G) is, however, not abelian and can therefore not be considered a stabilizer group. The problem lies in the last commutation relation. This relation states that c pairs of the symplectic generating set anti-commutes. Define operators by

$$Z_i = \left(\bigotimes_{j=1}^{i-1} I \right) \otimes Z \otimes \left(\bigotimes_{j=i+1}^{s+c+k'} I \right) \qquad \text{and} \qquad X_i := \left(\bigotimes_{j=1}^{i-1} I \right) \otimes X \otimes \left(\bigotimes_{j=i+1}^{s+c+k'} I \right) \tag{5.5}$$

Then the group $G = \langle \tilde{Z}_1, \dots, \tilde{Z}_{s+c}, \tilde{X}_1, \dots, \tilde{X}_c \rangle$ has similar commutation relations to the group $H = \langle Z_1, \dots, Z_{s+c}, X_1, \dots, X_c \rangle$ in the sense that a group isomorphism exists between them. The groups are thus said to be isomorphic, denoted $G \cong H$. A group isomorphism is a bijective map $\phi : G \to H$ satisfying

$$\phi(g_1 g_2) = \phi(g_1)\phi(g_2), \qquad \forall g_1, g_2 \in G. \tag{5.6}$$

The commutation relations between two elements $g_1, g_2 \in G$ can be expressed by

$$g_1 g_2 g_1^{-1} g_2^{-1} = \begin{cases} I, & \text{if } g_1 g_2 = g_2 g_1 \\ -I, & \text{if } g_1 g_2 = -g_2 g_1 \end{cases}$$

A group isomorphism preserves these commutation relations since

$$\alpha I = \phi(\alpha I) = \phi(g_1 g_2 g_1^{-1} g_2^{-1}) = \phi(g_1)\phi(g_2)\phi(g_1)^{-1}\phi(g_2)^{-1}, \qquad \alpha \in \{1, -1\}.$$

A group isomorphism can therefore only be defined between two groups with similar commutation relations. With that in mind, define $\phi : \{\tilde{Z}_1, \dots, \tilde{Z}_{s+c}, \tilde{X}_1, \dots, \tilde{X}_c\} \to \{Z_1, \dots, Z_{s+c}, X_1, \dots, X_c\}$ by

$$\phi(\tilde{Z}_i) := Z_i \qquad \text{and} \qquad \phi(\tilde{X}_i) := X_i. \tag{5.7}$$

The function can then be extended to all of G by requiring the property of (5.6). This, in turn, makes the codomain all of H. It is fairly simple to show the image of ϕ under G is exactly H, making ϕ a surjection. Since G and H contain the same number of elements, this suffices to argue ϕ is a bijection and hence a group isomorphism. Though abstractly defined, the isomorphism defined above can be realized by a unitary transformation, as the next theorem shows.

Theorem 5.2
Let $G = \langle \tilde{Z}_1, \ldots, \tilde{Z}_{s+c}, \tilde{X}_1, \ldots, \tilde{X}_c \rangle$ and $H = \langle Z_1, \ldots, Z_{s+c}, X_1, \ldots, X_c \rangle$ be subgroups of $\mathcal{P}_N$, where Z_i and X_i are defined by (5.5). Assume $G \cong H$ with the group isomorphism, ϕ, satisfying $\phi(\tilde{Z}_i) = Z_i$ and $\phi(\tilde{X}_i) = X_i$. Then there exists a unitary operator U such that

$$\forall g \in G \exists h \in H : \quad g = UhU^\dagger.$$

Proof
The generating set for H satisfies the commutation relations:

$$\begin{aligned} [Z_i, Z_j] &= 0, \qquad i,j \in \{1,\ldots,s+c\} \\ [X_i, X_j] &= 0, \qquad i,j \in \{1,\ldots,c\} \\ [Z_i, X_j] &= 0, \qquad i \neq j \\ \{Z_i, X_i\} &= 1, \qquad i \in \{1,\ldots,c\} \end{aligned}$$

Since $\phi(\tilde{Z}_i) = Z_i$ and $\phi(\tilde{X}_i) = X_i$, the generating set for G satisfies the commutation relation of (5.4). By theorem 5.1 the generating set $\langle \tilde{Z}_1, \ldots, \tilde{Z}_{s+c}, \tilde{X}_1, \ldots, \tilde{X}_c \rangle$ can be extended to a generating set for $\mathcal{P}_N / \langle I, -I, iI, -iI \rangle$ such

$$\begin{aligned} [\tilde{Z}_i, \tilde{Z}_j] &= 0, \qquad i,j \in \{1,\ldots,N\} \\ [\tilde{X}_i, \tilde{X}_j] &= 0, \qquad i,j \in \{1,\ldots,N\} \\ [\tilde{Z}_i, \tilde{X}_j] &= 0, \qquad i \neq j \\ \{\tilde{Z}_i, \tilde{X}_i\} &= 1, \qquad i \in \{1,\ldots,N\}. \end{aligned} \tag{5.8}$$

This is achieved by adding an additional s elements $\{\tilde{X}_{c+1}, \ldots, \tilde{X}_{c+s}\}$ and an additional $N-(s+c)$ pairs of elements $\{\tilde{Z}_{s+c+i}, \tilde{X}_{s+c+i}\}$ for $i \in \{1, \ldots, N-(s+c)\}$ to the generating set of G. Extending the generating set for H by $\{Z_1, \ldots, Z_N, X_1, \ldots, X_N\}$ yields a generating set for $\mathcal{P}_N / \langle I, -I, iI, -iI \rangle$ having similar commutation relations to that of (5.8). Multiplying the extended generating set of G or H by a phase $\alpha \in \{\pm 1, \pm i\}$ leaves the commutation relations and the group they generate unchanged. Both sets can therefore, without loss of generality, be assumed hermitian. Since the generators of both G and H belong to $\mathcal{P}_N$ they are also unitary. Together, these properties imply the eigenvalues of both generating sets are ± 1 and eigenvectors corresponding to distinct eigenvalues are orthogonal. Let $|\tilde{0}\rangle$ be an eigenvector corresponding to the $+1$ eigenvalue for all generators of G. Furthermore, let $\alpha \in \mathbb{F}_2^N$ and define $\left|\tilde{\psi}_\alpha\right\rangle := \tilde{X}_1^{\alpha_1} \cdots \tilde{X}_N^{\alpha_N} |\tilde{0}\rangle$. Then for

$i \in \{1, \ldots, N\}$

$$\begin{aligned} \tilde{Z}_i \left|\tilde{\psi}_\alpha\right\rangle &= \tilde{Z}_i \tilde{X}_1^{\alpha_1} \cdots \tilde{X}_N^{\alpha_N} \left|\tilde{0}\right\rangle \\ &= (-1)^{\delta_{\alpha_i,1}} \tilde{X}_1^{\alpha_1} \cdots \tilde{X}_N^{\alpha_N} \tilde{Z}_i \left|\tilde{0}\right\rangle \\ &= (-1)^{\delta_{\alpha_i,1}} \tilde{X}_1^{\alpha_1} \cdots \tilde{X}_N^{\alpha_N} \left|\tilde{0}\right\rangle \\ &= (-1)^{\delta_{\alpha_i,1}} \left|\tilde{\psi}_\alpha\right\rangle. \end{aligned} \tag{5.9}$$

This means $\left|\tilde{\psi}_\alpha\right\rangle$ is an eigenvector of Z_i corresponding to eigenvalue $+1$ if $\alpha_i = 0$ and -1 if $\alpha_i = 1$. Furthermore, if $\alpha_1, \alpha_2 \in \mathbb{F}_2^N$, then $\left|\tilde{\psi}_{\alpha_1}\right\rangle$ and $\left|\tilde{\psi}_{\alpha_2}\right\rangle$ are orthogonal if $\alpha_1 \neq \alpha_2$. Since Z_i operates on $\mathbb{C}^{2^N}$, the eigenvectors $\left|\tilde{\psi}_\alpha\right\rangle$ belongs to $\mathbb{C}^{2^N}$, paired with the fact they are orthogonal and there are 2^N of them, they constitutes an orthonormal basis of $\mathbb{C}^{2^N}$. Similar results to (5.9) can be expressed for the extended generating set of H, simply replace $\tilde{Z}_i$ by Z_i and $\tilde{X}_i$ by X_i. Denote the corresponding eigenvector basis of this case by $|\psi_\alpha\rangle$. The operator

$$U := \sum_{\alpha \in \mathbb{F}_2^N} |\psi_\alpha\rangle \left\langle \tilde{\psi}_\alpha \right|$$

is then a unitary operator on $\mathbb{C}^{2^N}$. Consider now the effect of U on the extended generating set of G. For $i \in \{1, \ldots, N\}$

$$\begin{aligned} U \tilde{Z}_i U^\dagger &= \left(\sum_{\alpha \in \mathbb{F}_2^N} |\psi_\alpha\rangle \left\langle \tilde{\psi}_\alpha \right| \right) \tilde{Z}_i \left(\sum_{\alpha' \in \mathbb{F}_2^N} \left|\tilde{\psi}_{\alpha'}\right\rangle \langle \psi_{\alpha'}| \right) \\ &= \sum_{\alpha, \alpha' \in \mathbb{F}_2^N} (-1)^{\delta_{\alpha'_i,1}} |\psi_\alpha\rangle \left\langle \tilde{\psi}_\alpha \middle| \tilde{\psi}_{\alpha'} \right\rangle \langle \psi_{\alpha'}| \\ &= \sum_{\alpha \in \mathbb{F}_2^N} (-1)^{\delta_{\alpha_i,1}} |\psi_\alpha\rangle \langle \psi_\alpha| \\ &= \sum_{\alpha \in \mathbb{F}_2^N} Z_i |\psi_\alpha\rangle \langle \psi_\alpha| = Z_i \end{aligned}$$

For the effect of U on $\tilde{X}_i$ consider first

$$\tilde{X}_i \left|\tilde{\psi}_\alpha\right\rangle = \tilde{X}_i \tilde{X}_1^{\alpha_1} \cdots \tilde{X}_N^{\alpha_N} \left|\tilde{0}\right\rangle = \tilde{X}_1^{\alpha_1} \cdots \tilde{X}_{i-1}^{\alpha_{i-1}} \tilde{X}_i^{\alpha_i + 1} \tilde{X}_{i+1}^{\alpha_{i+1}} \cdots \tilde{X}_N^{\alpha_N} \left|\tilde{0}\right\rangle = \left|\tilde{\psi}_{\alpha + e_i}\right\rangle$$

where $(e_i)_j = \delta_{ij}$. Analogues results hold for X_i. For $i \in \{1, \ldots, N\}$ it follows

$$\begin{aligned} U \tilde{X}_i U^\dagger &= \left(\sum_{\alpha \in \mathbb{F}_2^N} |\psi_\alpha\rangle \left\langle \tilde{\psi}_\alpha \right| \right) \tilde{X}_i \left(\sum_{\alpha' \in \mathbb{F}_2^N} \left|\tilde{\psi}_{\alpha'}\right\rangle \langle \psi_{\alpha'}| \right) \\ &= \sum_{\alpha, \alpha' \in \mathbb{F}_2^N} |\psi_\alpha\rangle \left\langle \tilde{\psi}_\alpha \middle| \tilde{\psi}_{\alpha' + e_i} \right\rangle \langle \psi_{\alpha'}| \\ &= \sum_{\alpha' \in \mathbb{F}_2^N} |\psi_{\alpha' + e_i}\rangle \langle \psi_{\alpha'}| \\ &= \sum_{\alpha' \in \mathbb{F}_2^N} X_i |\psi_{\alpha'}\rangle \langle \psi_{\alpha'}| = X_i. \end{aligned}$$

The effects of U on the extended generating set of G determines the effect of U for any $g \in \langle \tilde{Z}_1, \ldots, \tilde{Z}_N, \tilde{X}_1, \ldots, \tilde{X}_N \rangle$ as for any $g = \tilde{Z}_1^{\gamma_1} \cdots \tilde{Z}_N^{\gamma_N} \tilde{X}_1^{\gamma_{N+1}} \cdots \tilde{X}_N^{\gamma_{2N}}$ with $\gamma_1, \ldots, \gamma_{2N} \in \mathbb{Z}$ it is the case that

$$\begin{aligned} UgU^\dagger &= U\tilde{Z}_1^{\gamma_1}U^\dagger U \cdots U^\dagger U\tilde{Z}_N^{\gamma_N}U^\dagger U\tilde{X}_1^{\gamma_{N+1}}U^\dagger U \cdots U^\dagger U\tilde{X}_N^{\gamma_{2N}}U^\dagger \\ &= Z_1^{\gamma_1} \cdots Z_N^{\gamma_N} X_1^{\gamma_{N+1}} \cdots X_N^{\gamma_{2N}}. \end{aligned}$$

Since U maps G to H by conjugation, $U^\dagger$ is unitary and maps H to G by conjugation and the proof is concluded. ∎

Suppose both H and thus G of theorem 5.2 define a stabilizer group. As derived earlier, the error-correction capabilities of a stabilizer group depends solely on the commutation relations concerning the group: The stabilizer defined by H corrects the set of errors $\mathcal{A} := \{E_i \in \mathcal{P}_N\}_i$ if and only if $E_j^\dagger E_i \notin N_{\mathcal{P}_N}(H) \backslash H$ for all i, j. Thus, if the product between any two elements of $\mathcal{A}$ and $\mathcal{A}^{-1} := \{E_i^{-1}\}_i$ either anti-commutes with at least one stabilizer or belongs to the stabilizer group, the set $\mathcal{A}$ is a correctable set of errors.

Consider two arbitrary elements of $\langle Z_1, \ldots, Z_N, X_1, \ldots, X_N \rangle$ and $\langle \tilde{Z}_1, \ldots, \tilde{Z}_N, \tilde{X}_1, \ldots, \tilde{X}_N \rangle$, respectively, defined by $h_i := Z_1^{\alpha_{i,1}} \cdots Z_N^{\alpha_{i,N}} X_1^{\alpha_{i,N+1}} \cdots X_N^{\alpha_{i,2N}}$ and similarly $g_i := \tilde{Z}_1^{\alpha_{i,1}} \cdots \tilde{Z}_N^{\alpha_{i,N}} \tilde{X}_1^{\alpha_{i,N+1}} \cdots \tilde{X}_N^{\alpha_{i,2N}}$ for $i = 1, 2$. By the proof of theorem 5.2 $Ug_iU^\dagger = h_i$ and thus $Ug_i^{-1}U^\dagger = h_i^{-1}$. The commutation relation between g_1 and g_2 are captured by the product

$$h_1 h_2 h_1^{-1} h_2^{-1} = \begin{cases} I, & \text{if } h_1 h_2 = h_2 h_1 \\ -I, & \text{if } h_1 h_2 = -h_2 h_1. \end{cases}$$

This relation is preserved by the unitary since

$$\alpha I = U\alpha I U^\dagger = U h_1 h_2 h_1^{-1} h_2^{-1} U^\dagger = g_1 g_2 g_1^{-1} g_2^{-1}. \tag{5.10}$$

As a consequence, since the stabilizer of H (and only the stabilizer) is mapped to the stabilizer of G according to theorem 5.2, if $\mathcal{A} = \{E_i \in \mathcal{P}_N\}_i$ is a correctable set of errors for H then $U\mathcal{A}U^\dagger := \{UE_iU^\dagger\}_i$ is a correctable set of errors for G. Furthermore, if $\mathcal{A}$ is the largest correctable set of errors possible for H, then $U\mathcal{A}U^\dagger$ will be the largest possible set of correctable errors for G. The codespace of the two stabilizers are also related by the unitary U, since if $|\psi\rangle$ is an eigenvector corresponding to the simultaneous $+1$ eigenspace for all $h \in H$, then for any $g = UhU^\dagger \in G$

$$gU|\psi\rangle = UhU^\dagger U|\psi\rangle = Uh|\psi\rangle = U|\psi\rangle$$

Conversely, if $|\phi\rangle$ belongs to the codespace of G, then $U^\dagger|\phi\rangle$ belongs to the codespace of H.

The problem with the above is that neither H nor G generally defines a stabilizer, since neither is abelian if $c > 0$. However, inspecting the generating set $\langle Z_1, \ldots, Z_{s+c}, X_1, \ldots, X_c \rangle$ for H reveals a method of converting H into an abelian group by embedding the generators

in operators on a larger state space. Define new operators by

$$\begin{aligned}
\hat{Z}_i &:= Z_i \otimes \left(\bigotimes_{j=1}^{i-1} I\right) \otimes Z \otimes \left(\bigotimes_{j=i+1}^{c} I\right), \qquad i \in \{1,\ldots,c\},\\
\hat{Z}_i &:= Z_i \otimes \left(\bigotimes_{j=1}^{c} I\right), \qquad i \in \{c+1,\ldots,s+c\},\\
\hat{X}_i &:= X_i \otimes \left(\bigotimes_{j=1}^{i-1} I\right) \otimes X \otimes \left(\bigotimes_{j=i+1}^{c} I\right), \qquad i \in \{1,\ldots,c\}.
\end{aligned} \tag{5.11}$$

The group defined by $\hat{H} := \langle \hat{Z}_1,\ldots,\hat{Z}_{s+c},\hat{X}_1,\ldots,\hat{X}_c\rangle$ is then abelian and thus defines a stabilizer. An analogue extension of the generating set for G can be defined by replacing Z with $\tilde{Z}$ in (5.11). Define a function by

$$\phi(\hat{\tilde{Z}}_i) = \hat{Z}_i \qquad \text{and} \qquad \phi(\hat{\tilde{X}}_i) = \hat{X}_i,$$

and require (5.6) is satisfied, it is evident that the newly obtained groups are still isomorphic. It thus suffices to consider $\hat{H}$ when studying the properties of the stabilizer defined by $\hat{G} := \langle \hat{\tilde{Z}}_1,\ldots,\hat{\tilde{Z}}_{s+c},\hat{\tilde{X}}_1,\ldots,\hat{\tilde{X}}_c\rangle$; a helpful fact when trying to identify the codespace. Denote by $\bigotimes_{i=1}^{N}\mathcal{H}_{\mathcal{A}_i}$ and $\left(\bigotimes_{i=1}^{N}\mathcal{H}_{\mathcal{A}_i}\right)\otimes\left(\bigotimes_{j=1}^{c}\mathcal{H}_{\mathcal{B}_j}\right)$ the state space which the operators of H and $\hat{H}$ operates on, respectively. The embedding of H into $\hat{H}$ necessary to satisfy the abelian property thus adds qubits residing in $\bigotimes_{j=1}^{c}\mathcal{H}_{\mathcal{B}_j}$. The codespace of $\hat{H}$ can now be identified and expressed as

$$V_{\hat{H}} = \left(\bigotimes_{j=1}^{c}\left|\Phi^+\right\rangle_{\mathcal{A}_j,\mathcal{B}_j}\right)\otimes\left(\bigotimes_{j=c+1}^{c+s}|0\rangle_{\mathcal{A}_j}\right)\otimes\left(\bigotimes_{j=1}^{k}|\psi_j\rangle_{\mathcal{A}_{c+s+j}}\right) \tag{5.12}$$

where $|\psi_i\rangle$ is an arbitrary state in state space $\mathcal{H}_{\mathcal{A}_{s+c+i}}$, $k = N-(c+s)$ and $\left|\Phi^+\right\rangle_{\mathcal{A}_i,\mathcal{B}_i} \in \mathcal{H}_{\mathcal{A}_i}\otimes\mathcal{H}_{\mathcal{B}_i}$ is the Bell state given by

$$\left|\Phi^+\right\rangle_{\mathcal{A}_i,\mathcal{B}_i} := \frac{|0\rangle_{\mathcal{A}_i}\otimes|0\rangle_{\mathcal{B}_i} + |1\rangle_{\mathcal{A}_i}\otimes|1\rangle_{\mathcal{B}_i}}{\sqrt{2}}.$$

Since the added qubits are entangled, such a setup can be realized by the sender and the receiver sharing c Bell pairs before communication. The sender then encodes k logical qubits into k physical qubits, s ancilla qubits and their half of the c pre-shared entangled qubits. The $k+c+s$ qubits are then transmitted to the receiver who uses the received qubits in conjunction with their half of the c Bell pairs to error correct and decode. Since the receiver's half of the Bell pairs are never transmitted over the noisy channel, these are assumed noiseless. The possible errors are therefore of the form $E\otimes I$ where E operates on $\bigotimes_{j=1}^{N}\mathcal{H}_{\mathcal{A}_j}$ and I on $\bigotimes_{j=1}^{c}\mathcal{H}_{\mathcal{B}_j}$. For these errors to be correctable they must at the minimum anti-commute with at least one generator or belong to the stabilizer group, $\hat{H}$. Consider the first case, the error $E\otimes I$ anti-commutes with some element $A\otimes B\in\hat{H}$ if and only if E anti-commutes with $A\in H$. This is the case if and only if $E\in\mathcal{P}_N\backslash N_{\mathcal{P}_N}(H)$ where $N_{\mathcal{P}_N}(H) = C_{\mathcal{P}_N}(H)$. In the second case $E\otimes I\in\hat{H}$ if and only if $E\in\text{iso}(H) = \langle Z_{c+1},\ldots,Z_{c+s}\rangle$. Altogether, $\mathcal{D}\subset P_N$ is a correctable set of errors if $E_i^\dagger E_j\in\text{iso}(H)\cup\mathcal{P}_N\backslash N_{\mathcal{P}_N}(H)$ for any two $E_i,E_j\in\mathcal{D}$.

An entanglement-assisted quantum error-correcting code such as the above is characterized by four parameters. k; the number of logical qubits containing the information wanting to be transmitted, N; the number of physical qubits transmitted over the noisy channel, d; the minimum weight of any uncorrectable error, that is, the minimum weight of $N_{\mathcal{P}_N}(H)\backslash \text{iso}(H)$ and lastly, c; the number of entangled qubit pairs utilized in the code. A code of these parameters is called an $[[N, k, d; c]]$ entanglement-assisted quantum error-correcting code. The rate of such a code can be defined in 3 ways depending on the perspective of the entanglement resource. In the first case, the resource is considered free, therefore, the rate is defined as $\frac{k}{N}$, the fraction of logical qubits transmitted to the total number of qubits transmitted per channel use. In the second case, the resource is not free and the rate is defined as the pair $(\frac{k}{N}, \frac{c}{N})$. The first entry is the number of logical qubits transmitted to the total number of qubits transmitted per channel use, the second entry is the number of entangled qubits consumed per transmission, to the total number of qubits transmitted per channel use. In the last case, the resource comes at the expense of transmitted qubits, and thus the rate is given by $\frac{k-c}{N}$; the fraction of the difference between the number of logical qubits transmitted and the number of entangled qubits consumed to the total number of qubits transmitted per channel use. In the last case, the rate can be negative if the code requires more entangled qubits than transmitted logical qubits per channel use.

The following example demonstrates how generators of a simple non-abelian subgroup of $\mathcal{P}_4$ that are not initially categorized into symplectic and isotropic parts can be transformed into generators that are categorized accordingly by applying the method of theorem 5.1. This process, in turn, reveals how to extend the group into an abelian group.

Example 8
Consider the group generated by the following generators and the corresponding check matrix.

$$\begin{aligned} g_1 &:= Z \otimes X \otimes Z \otimes I, \\ g_2 &:= Z \otimes Z \otimes I \otimes Z, \\ g_3 &:= X \otimes Y \otimes X \otimes I, \\ g_4 &:= X \otimes X \otimes I \otimes X. \end{aligned} \quad \longrightarrow \quad \left[\begin{array}{cccc|cccc} 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \end{array}\right]$$

Denote the group by G and the corresponding subspace whose basis are the rows of the check matrix by G_2. The rows of the check matrix are easily extended to a basis of $\mathbb{F}_2^8$ using Gaussian elimination. By the transformation explained in the proof of theorem 5.1, this basis is then converted into a symplectic basis, keeping track of the rows constituting

a basis of G_2. The convertion are depicted below.

$$\begin{array}{c|cccc|cccc} \overline{c}_1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ \overline{c}_2 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ \overline{c}_3 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ \overline{c}_4 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline \overline{c}_5 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ \overline{c}_6 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ \overline{c}_7 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ \overline{c}_8 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \end{array} \rightarrow \begin{array}{cccc|cccc|c} 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & \overline{a}_1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & \overline{b}_1 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & \overline{a}_2 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \overline{b}_2 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & \overline{a}_3 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \overline{b}_3 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & \overline{a}_4 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & \overline{b}_4 \end{array}$$

The elements of $\{\overline{a}_i\}_{i=1}^4$ and $\{\overline{b}_i\}_{i=1}^4$ which belongs to G_2 are given by indices $\mathcal{A} = \{1, 2, 3\}$ and $\mathcal{B} = \{1\}$, respectively. The operators

$$\begin{aligned} \tilde{Z}_1 &:= Z \otimes X \otimes Z \otimes I, \\ \tilde{X}_1 &:= Z \otimes Z \otimes I \otimes Z, \\ \tilde{Z}_2 &:= Y \otimes X \otimes X \otimes Z, \\ \tilde{Z}_3 &:= Z \otimes Y \otimes Y \otimes X, \end{aligned}$$

thus constitutes a generating set for G. The symplectic subgroup is given by $\langle \tilde{Z}_1, \tilde{X}_1 \rangle$ and the isotropic subgroup is given by $\langle \tilde{Z}_2, \tilde{Z}_3 \rangle$. The group can now be extended to an abelian group by

$$\begin{aligned} \hat{\tilde{Z}}_1 &:= Z \otimes X \otimes Z \otimes I \otimes Z, \\ \hat{\tilde{X}}_1 &:= Z \otimes Z \otimes I \otimes Z \otimes X, \\ \hat{\tilde{Z}}_2 &:= Y \otimes X \otimes X \otimes Z \otimes I, \\ \hat{\tilde{Z}}_3 &:= Z \otimes Y \otimes Y \otimes X \otimes I. \end{aligned}$$

The above generators can be verified to define a valid $[[5, 1, 3]]$ stabilizer or a $[[4, 1, 3; 1]]$ entanglement-assisted stabilizer depending on the view of origin.

Entanglement-assisted stabilizers have now been established along with their motivation and error-correcting abilities. However, it remains to examine their encoding schemes; given the required Bell pairs, the logical qubits containing the information of interest and the ancilla qubits, how to encode the information into the set of qubits consisting of the logical qubits, the ancilla qubits and the transmitters half of the Bell pairs.

5.1 Encoding and Decoding

The encoding scheme for an entanglement-assisted stabilizer follows a construction similar to that of non-entanglement-assisted stabilizers derived in section 4.3.2. In the algorithm

providing the encoding circuit for an entanglement-assisted stabilizer, the three Clifford gates and the swap gate lie the foundation. Furthermore, the entanglement-assisted stabilizer of interest is transformed into an entanglement-assisted stabilizer with known coding space and encoding circuit denoted the standard entanglement-assisted stabilizer. A suitable choice of standard entanglement-assisted stabilizer is the one with the coding space provided in (5.12).

Following the above notation, denote two non-abelian isomorphic subgroups of $\mathcal{P}_N$ by $H := \langle Z_1, \ldots, Z_{s+c}, X_1, \ldots, X_c \rangle$ and $G := \langle \tilde{Z}_1, \ldots, \tilde{Z}_{s+c}, \tilde{X}_1, \ldots, \tilde{X}_c \rangle$ with isomorphism defined by (5.7). Extend both subgroups using the method of (5.11) into abelian subgroup. The extension transforms H into the standard entanglement-assisted stabilizer denoted $\hat{H} := \langle \hat{Z}_1, \ldots, \hat{Z}_{s+c}, \hat{X}_1, \ldots, \hat{X}_c \rangle$ and G into a general entanglement-assisted stabilizer denoted $\hat{G} := \langle \hat{\tilde{Z}}_1, \ldots, \hat{\tilde{Z}}_{s+c}, \hat{\tilde{X}}_1, \ldots, \hat{\tilde{X}}_c \rangle$. Since the last c entries in the tensor product of each generator coincide, the encoder only considers the transformation of the first N entries, that is, the non-abelian subgroups pre-extensions. The last c qubits are thus not transformed by the encoding circuit in agreement with the fact that these correspond to the receivers half of the shared Bell states. Since these are shared before the encoding begins, they are not possessed by the transmitter during encoding.

In algorithm 3 it is assumed that the pre-extended entanglement-assisted stabilizer of interest follows the isomorphic structure described above. If that is not the case, theorem 5.1 should be applied to the non-abelian subgroup before following the encoding algorithm.

Algorithm 3 Encoding Algorithm for Entanglement-Assisted Stabilizers

Let $G = \langle \tilde{Z}_1, \ldots, \tilde{Z}_{s+c}, \tilde{X}_1, \ldots, \tilde{X}_c \rangle$ be a subgroup of $\mathcal{P}_N$

Denote by $r_i := \begin{cases} r(\tilde{Z}_i), & i \in \{1, \ldots, s+c\}, \\ r(\tilde{X}_i), & i \in \{s+c+1, \ldots, s+2c\}, \end{cases}$ row i of the check matrix of G cf. (4.5)

Perform row swaps such $r_i \leftarrow \begin{cases} r_{\frac{i+1}{2}}, & i \in \{1, 3, 5, \ldots, 2c-1\} \\ r_{s+c+\frac{i}{2}}, & i \in \{2, 4, 6, \ldots, 2c\} \\ r_{i-c}, & i \in \{2c+1, \ldots, 2c+s\} \end{cases}$

for $i = 1, \ldots, c$ **do**
  if $r_{2i-1,i} = 0 \wedge r_{2i-1,i+N} = 0$ **then**
    Apply swap gate between qubit i and qubit l for some $i < l \in \{i+1, \ldots, N\}$ with $r_{i,l} = 1 \vee r_{i,l+N} = 1$
  end if
  for $j = i, \ldots, N$ **do**
    if $r_{2i-1,j} = 1 \wedge r_{2i-1,j+N} = 0$ **then**
      Continue
    else if $r_{2i-1,j} = 0 \wedge r_{2i-1,j+N} = 1$ **then**
      Apply Hadamard gate on qubit j
    else if $r_{2i-1,j} = 1 \wedge r_{2i-1,j+N} = 1$ **then**
      Apply phase gate to qubit j
    end if
  end for
  for $j = i+1, \ldots, N$ **do**
    if $r_{2i-1,j} = 1$ **then**
      Apply CNOT gate from qubit i to qubit j
    end if
  end for
  Row $2i-1$ are now of the form $r_{i,j} = \begin{cases} 1, & \text{if } j = i \\ 0, & \text{else} \end{cases}$ for $j \in \{1, \ldots, 2N\}$
  Apply Hadamard gate on qubit i
  for $j = i, \ldots, N$ **do**
    if $r_{2i,j} = 1 \wedge r_{2i-1,j+N} = 0$ **then**
      Continue
    else if $r_{2i,j} = 0 \wedge r_{2i-1,j+N} = 1$ **then**
      Apply Hadamard gate on qubit j
    else if $r_{2i,j} = 1 \wedge r_{2i-1,j+N} = 1$ **then**
      Apply phase gate to qubit j
    end if
  end for

Algorithm 3 Encoding Algorithm for Entanglement-Assisted Stabilizers (continued)

for $j = 2i+1, 2i+2, \ldots, 2c+s$ **do**
  if $r_{j,i} = 1$ **then**
    Add row $2i-1$ to row j
  end if
  if $r_{j,N+i} = 1$ **then**
    Add row $2i$ to row j
  end if
end for
end for
for $i = 1, \ldots, s$ **do**
  if $r_{2c+i,i} = 0 \wedge r_{2c+i,i+N} = 0$ **then**
    Apply swap gate between qubit i and qubit l for some $i < l \in \{i+1, \ldots, N\}$ with $r_{i,l} = 1 \vee r_{i,l+N} = 1$
  end if
  for $j = i, \ldots, N$ **do**
    if $r_{i,j} = 1 \wedge r_{i,j+N} = 0$ **then**
      Continue
    else if $r_{i,j} = 0 \wedge r_{i,j+N} = 1$ **then**
      Apply Hadamard gate on qubit j
    else if $r_{i,j} = 1 \wedge r_{i,j+N} = 1$ **then**
      Apply phase gate to qubit j
    end if
  end for
  for $j = i+1, \ldots, N$ **do**
    if $r_{i,j} = 1$ **then**
      Apply CNOT gate from qubit i to qubit j
    end if
  end for
  Row i are now of the form $r_{i,j} = \begin{cases} 1, & \text{if } j = i \\ 0, & \text{else} \end{cases}$ for $j \in \{1, \ldots, 2N\}$
  for $j = i+1, \ldots, N-k$ **do**
    if $r_{j,i} = 1$ **then**
      Add row i to row j
    end if
  end for
  Apply Hadamard gate to qubit i
  for $j = i+1, \ldots, N$ **do**
    if $r_{j,N+i} = 1$ **then**
      Add row i to row j
    end if
  end for
end for
The check matrix of G has now been transformed into the check matrix of H

Following algorithm 3, the encoding circuit for the pre-extended entanglement assisted stabilizer presented in example 8 can be found. The pre-extended entanglement-assisted stabilizer and corresponding check matrix is shown figure 5.1.

$$\begin{aligned}
\tilde{Z}_1 &:= Z \otimes X \otimes Z \otimes I, \\
\tilde{X}_1 &:= Z \otimes Z \otimes I \otimes Z, \\
\tilde{Z}_2 &:= Y \otimes X \otimes X \otimes Z, \\
\tilde{Z}_3 &:= Z \otimes Y \otimes Y \otimes X.
\end{aligned}$$

(a) Pre-extended entanglement-assisted stabilizer.

$$\left[\begin{array}{cccc|cccc}
0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\
1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\
0 & 1 & 1 & 1 & 1 & 1 & 1 & 0
\end{array}\right]$$

(b) Corresponding 4×8 check matrix.

Figure 5.1: Pre-extended entanglement-assisted stabilizer and corresponding check matrix of example 8.

The encoding circuit for the extended entanglement-assisted stabilizer is depicted in figure 5.2. As previously stated, the part of the Bell state in possession of the receiver is untouched by the encoder.

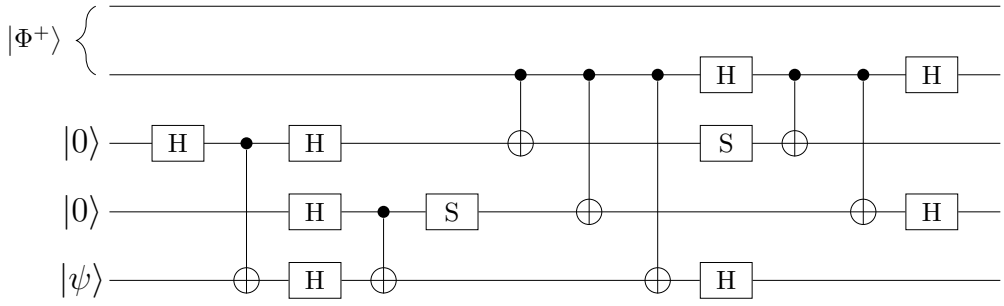


Figure 5.2: Encoding circuit of the new stabilizer group.

The encoding scheme of entanglement-assisted stabilizers reveal one of their great advantages. The following section, examining the communication setup, elaborates on this advantage.

5.2 Communication Framework

Entanglement-assisted stabilizers arise from non-abelian subgroups of $\mathcal{P}_N$. As has been established, such groups can be extended by the use of Bell pairs effectively embedding the non-abelian group in an abelian group operating on a larger set of qubits. The result is a stabilizer on $N + c$ qubits with c being the number of additional qubits needed to make the non-abelian group abelian. Since the abelian group on $N + c$ qubits defines a stabilizer, it may be implemented as such, yielding a $[[N + c, k, d]]$ stabilizer. However, recognizing and implementing it as an entanglement-assisted stabilizer grants abilities beyond what the

stabilizer may achieve. In section 5.1 it was revealed that the encoder of an entanglement-assisted stabilizer leaves the added c qubits untouched. This facilities the possibility of pre-distributing the added qubits, prior to encoding. Such a framework is depicted in figure 5.3: A base station prepares the c needed Bell pairs for the entanglement-assisted stabilizer. The Bell pairs are then distributed between the transmitter and receiver. Once c Bell pairs have been distributed, the Bell pairs on the transmitter's side are passed to the encoder along with the ancilla and logical qubits. The encoded qubits are then transmitted to the receiver, who, in conjunction with its half of the Bell pairs, performs error correction and decoding. This framework obtains a communication rate of $\frac{N}{k}$ in the communication phase. The stabilizer, not leveraging entanglement, may only obtain a rate of $\frac{N+c}{k}$.

The distribution may require the use of distillation before entering the communication phase, as the Bell pairs may undergo decoherence in the transmission and/or waiting phase. However, such a procedure can be performed in parallel to the communication phase. Suppose that the transmitter wants to send k qubits of logical information at time T. Then c Bell pairs must be distributed at time T. However, assuming no decoherence while idle in memory, the Bell pairs can be distributed several time steps prior to time T [Nielsen and Chuang, 2010, pp.578-580].

From the framework it is evident why the receiver's half of the Bell states are assumed error free; they are never transmitted once being passed to the communication phase. They may still undergo decoherence while idle in memory. The communication phase should thus be time constrained to ensure that the decoherence, from distillation to decoding, is negligible.

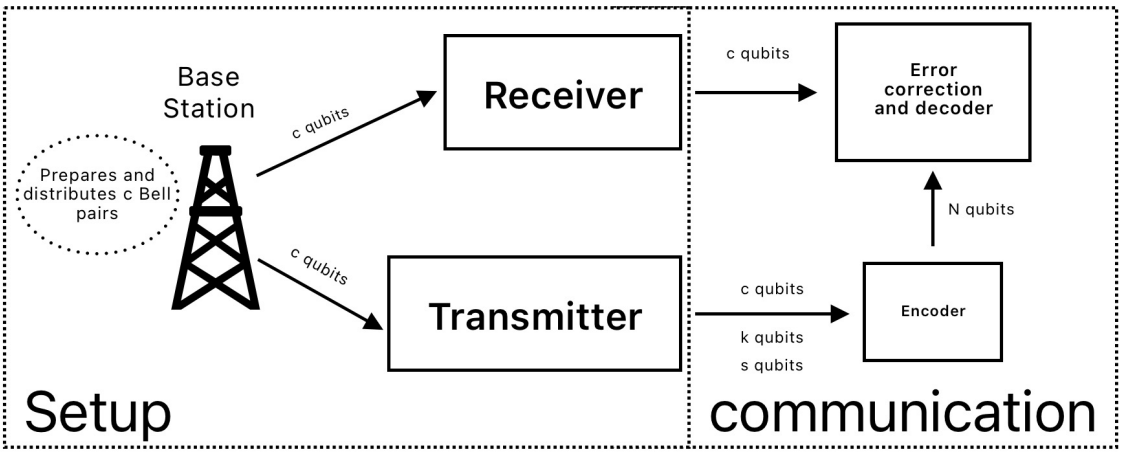


Figure 5.3: The communication framework when using entanglement-assisted stabilizers. The framework is split in two parts; setup and communication. The first part concerns entanglement distribution, the second concerns encoding, error correction and decoding.

5.2.1 Imperfect Bell Pairs and Latency Considerations

The error correction and decoder in the communication phase are based on the assumption that the c qubits input from the receiver are error-free. As mentioned, this is an impractical assumption, as the qubits may undergo decoherence while idle in memory, waiting for the

transmitter to encode and transmit their N qubits. This assumption may be relaxed by introducing a stabilizer code to protect the receivers qubits. After the qubits are passed to the communication phase, the transmitter encodes the logical information and transmits it. In the meantime, the receiver could encode the c qubits in its possession, using a number of ancilla qubits. The error correction and decoding of this scheme should then be done in a timely fashion such that the transmitter's qubits are received shortly after decoding, otherwise the decoded qubits are prone to decoherence in this (although shorter) waiting time. Let T denote the start time of the communication phase, at which time both the receiver and the transmitter begin the encoding of their respective qubits. Assume that the encoding on the transmitter's side ends at time $t_{\text{enc,trans}}$ after which the encoded qubits are transmitted to the receiver. The receiver is then in possession of all qubits at time t_{trans}. Denote by $t_{\text{enc,rec}}$, $t_{\text{dec,1}}$ and $t_{\text{dec,2}}$, the end of the encoding, the beginning of the error correction/decoding and the end of the error correction/decoding of the receiver's error-correction scheme, respectively. To ensure that the additional error-correction scheme introduced at the receiver's end does not increase communication latency, it is required $t_{\text{dec,2}} < t_{\text{enc,trans}}$. Although this alleviates the decoherence errors that emerge from qubits lying idle in memory, the scheme is only designed to protect against errors in the time span $[t_{\text{enc,rec}}; t_{\text{dec,1}}]$, the error-correction scheme should therefore be implemented as to minimize the difference $t_{\text{trans}} - t_{\text{dec,2}} > 0$. The timeline described above are illustrated in figure 5.4 [Lai and Brun, 2012].

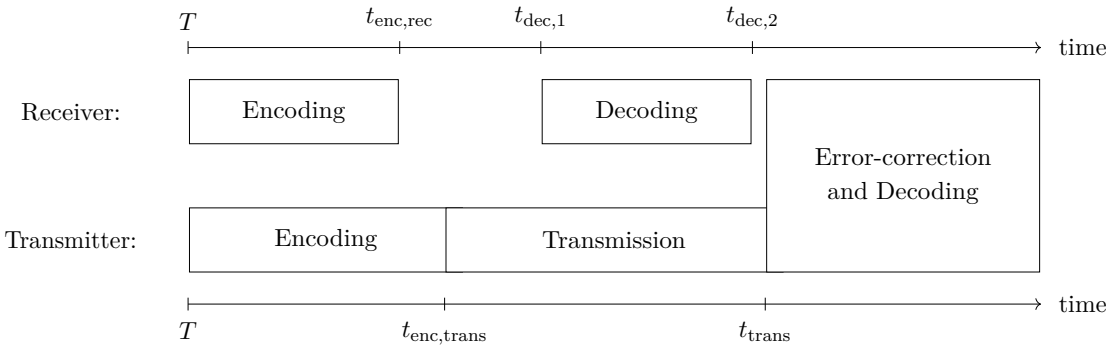


Figure 5.4: The timeline with an additional error-correction scheme introduced at the receiver's end to correct decoherence errors emerging as qubits lie idle in memory. The timeline requires synchronization between transmitter and receiver to minimize the communication latency.

Although the above introduction of an error-correction code to protect the receiver's c qubits mitigates the error-free assumption, the communication framework still requires c perfect Bell pairs passed to the communication phase. Not only does this require the base station to create perfect Bell pairs, but it also requires that there be no errors in the transmission of the Bell pairs to the transmitter/receiver, and that no decoherence occurs from the creation of the first Bell pair to the last Bell pair has been distributed. As mentioned, distillation may be used to obtain the perfect Bell pairs, thereby alleviating the requirements of the setup; however, this, in turn, requires a large number of imperfect Bell

pairs to be distributed. While this may be feasible given no time-constraints, practical settings typically demand a latency upper bound. In such a setting, distributing a large number of imperfect Bell pairs to be distilled, or simply waiting for the base station to supply c perfect Bell pairs consecutively, is unfeasible. Instead, a choice must be made; use c imperfect Bell pairs or puncture the error-correcting code to match the number of supplied perfect Bell pairs. The following section investigates the latter. It should be noted that a combination of both choices may be the best practical option. By using a combination of the two methods, one could require Bell pairs with a given fidelity. These Bell pairs are then supplied through distillation, and the error-correcting code is punctured to match the outcome of the distillation.

5.3 Puncturing

The notion of puncturing qubits in non-entanglement-assisted stabilizers has already been established in section 4.3.3. This section extends the method to include entanglement-assisted stabilizers, in particular, puncturing the qubits possessed by the receiver prior to the communication phase and examining its effect on the parameters of such codes. An entanglement-assisted stabilizer of parameters $[[N, k, d; c]]$ is a quantum error-correcting code on $N + c$ qubits where $N = s + c + k$ qubits are possessed by the transmitter. The remaining c qubits, corresponding to the receiver's half of the required Bell states, are possessed by the receiver only. Such codes arise from non-abelian groups on N qubits, extended to an abelian group by adding an additional c qubits. Since the code on $N + c$ qubits is abelian, it may be viewed as a stabilizer with parameters $[[N+c, k, d']]$ rather than an entanglement-assisted stabilizer with parameters $[[N, k, d; c]]$. The notion of puncturing already established can thus be used directly on the entanglement-assisted setup with a few remarks. Assuming the distance of the code prior to puncturing is greater than or equal to 2, puncturing on one of the c qubits possessed by the receiver transforms the parameters of the code by

$$[[N + c, k, d']] \quad \xrightarrow{\text{puncturing}} \quad [[N + c - 1, k, d'_p]].$$

The corresponding code is on $N + c - 1$ qubits, the missing qubit is exactly one of the c qubits possessed by the receiver. Since $d \geq 2$ the number of logical qubits after puncturing remains k. For an entanglement-assisted stabilizer, such a puncturing would imply N and k remain unchanged while c decreases by 1. Thus,

$$[[N, k, d; c]] \quad \xrightarrow{\text{puncturing}} \quad [[N, k, d_p; c - 1]]. \tag{5.13}$$

Taking into account $N = c + s + k$, it implies that s increases by 1. The punctured qubit is part of a Bell state with a qubit possessed by the transmitter. This corresponding qubit thus changes the characteristic from an entangled qubit to an ancilla qubit under puncturing.

The procedure is shown in the following example. An entanglement-assisted code is punctured on one of the receiver's qubits. The resulting code requires one less pre-shared Bell state. The previously required Bell state is instead transformed into an ancilla qubit given to the transmitter before encoding.

Example 9: Puncturing an EAQECC
Consider the non-abelian group of example 8 given by

$$\begin{aligned}\tilde{Z}_1 &:= Z \otimes X \otimes Z \otimes I,\\ \tilde{X}_1 &:= Z \otimes Z \otimes I \otimes Z,\\ \tilde{Z}_2 &:= Y \otimes X \otimes X \otimes Z,\\ \tilde{Z}_3 &:= Z \otimes Y \otimes Y \otimes X.\end{aligned}$$

The first and second generators anti-commute, making the coding space trivial. Adding a qubit, the group can be made abelian by

$$\begin{aligned}\hat{\tilde{Z}}_1 &:= Z \otimes X \otimes Z \otimes I \otimes Z,\\ \hat{\tilde{X}}_1 &:= Z \otimes Z \otimes I \otimes Z \otimes X,\\ \hat{\tilde{Z}}_2 &:= Y \otimes X \otimes X \otimes Z \otimes I,\\ \hat{\tilde{Z}}_3 &:= Z \otimes Y \otimes Y \otimes X \otimes I.\end{aligned}$$

This new group thus defines a valid stabilizer with parameters $[[5, 1, 3]]$ or as an entanglement-assisted stabilizer $[[4, 1, 3; 1]]$. The encoder of the extended group is shown in figure 5.2. The check matrix of the new group is given by

$$\begin{array}{c} \\ \hat{\tilde{Z}}_1 \\ \hat{\tilde{X}}_1 \\ \hat{\tilde{Z}}_2 \\ \hat{\tilde{Z}}_3 \end{array} \left[\begin{array}{cccc:c|cccc:c} 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \end{array}\right].$$

Puncture the code wrt. (10) on the last qubit.

$$\left[\begin{array}{cccc:c|cccc:c} 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{cccc|cccc} 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{array}\right].$$

The punctured stabilizer is given by

$$\begin{aligned}g_2 &:= Z \otimes Z \otimes I \otimes Z,\\ g_3 &:= Y \otimes X \otimes X \otimes Z\\ g_4 &:= Z \otimes Y \otimes Y \otimes X.\end{aligned}$$

This can be encoded by the circuit shown in figure 5.5.

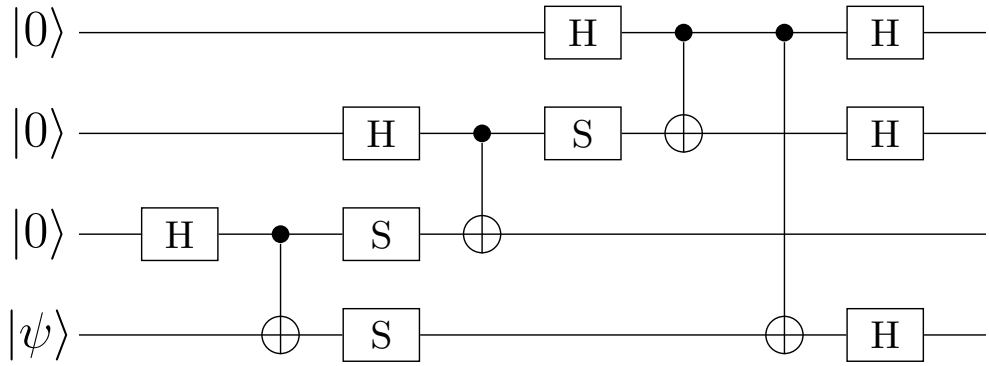


Figure 5.5: Encoding circuit of punctured stabilizer.

This code utilizes only the qubits possessed by the transmitter and thus does not require any pre-distributed Bell states. The code has parameters $[[4, 1, 2]]$.

The puncturing algorithm allows for a dynamic setup, solving the problem of non-provided Bell states explained in section 5.2: Rather than defining codes relying on c' Bell states for all $c' \in \{1, \ldots, c\}$ in case the base station is unable to provide c Bell states, one simply punctures the code requiring c Bell states (possibly recursively).

It remains to examine the change in distance under puncturing. For a non-entanglement-assisted stabilizer, the distance may decrease by at most 1. In the entanglement-assisted case, this cannot be guaranteed for all codes. As mentioned previously, the qubits possessed by the receiver are assumed to be error-free. This implies that the set of valid logical errors contains identities on the c entries corresponding to the receiver's qubits. The set of valid logical errors for an $[[N, k, d; c]]$ entanglement-assisted stabilizer is thus a subset of the logical errors for the corresponding $[[N + c, k, d']]$ non-entanglement-assisted stabilizer. In case the subset does not contain the minimum weight elements of the set, the entanglement-assisted stabilizer achieves a greater distance than the corresponding non-entanglement-assisted stabilizer. When using the puncturing strategy developed for non-entanglement-assisted stabilizers on entanglement-assisted stabilizers, part of the process requires viewing the entanglement-assisted stabilizers as non-entanglement-assisted stabilizer before puncturing and vice versa post puncturing. These steps introduce a possible change in distance in addition to the change in distance occurring under puncturing. The steps to follow when puncturing an entanglement-assisted stabilizer are illustrated in figure 5.6. The figure highlights the possible changes in distance that may occur under a puncturing procedure.

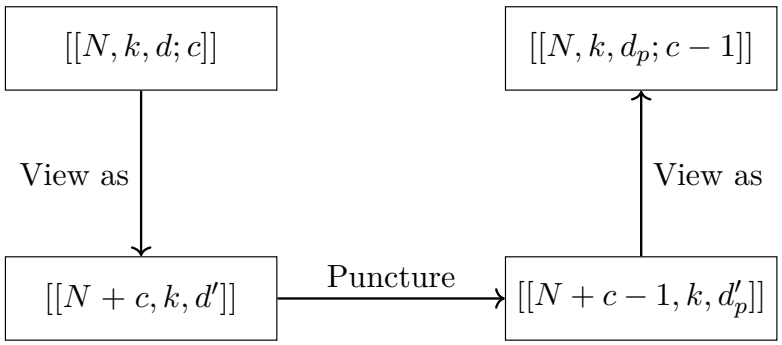


Figure 5.6: The steps necessary to puncture an entanglement-assisted stabilizer. Each step may change the distance of the examined code.

The following example provides a specific case of an entanglement-assisted stabilizer that achieves a greater distance than its non-entanglement-assisted counterpart. The distance of such codes under puncturing is highly volatile.

Example 10
Consider the non-abelian group given by

$$\begin{aligned}
\tilde{Z}_1 &:= Z \otimes Z \otimes I \otimes I \otimes I \\
\tilde{X}_1 &:= I \otimes X \otimes X \otimes X \otimes X \\
\tilde{Z}_2 &:= I \otimes Z \otimes Z \otimes I \otimes I \\
\tilde{X}_2 &:= X \otimes X \otimes I \otimes I \otimes I \\
\tilde{Z}_3 &:= I \otimes I \otimes Z \otimes Z \otimes I \\
\tilde{X}_3 &:= I \otimes I \otimes I \otimes X \otimes X \\
\tilde{Z}_4 &:= I \otimes I \otimes I \otimes Z \otimes Z \\
\tilde{X}_4 &:= X \otimes X \otimes X \otimes X \otimes I.
\end{aligned}$$

Extending the group by the introduction of four additional qubits yields the abelian group

$$\begin{aligned}
\tilde{Z}_1 &:= Z \otimes Z \otimes I \otimes I \otimes I \otimes Z \otimes I \otimes I \otimes I \\
\tilde{X}_1 &:= I \otimes X \otimes X \otimes X \otimes X \otimes X \otimes I \otimes I \otimes I \\
\tilde{Z}_2 &:= I \otimes Z \otimes Z \otimes I \otimes I \otimes I \otimes Z \otimes I \otimes I \\
\tilde{X}_2 &:= X \otimes X \otimes I \otimes I \otimes I \otimes I \otimes X \otimes I \otimes I \\
\tilde{Z}_3 &:= I \otimes I \otimes Z \otimes Z \otimes I \otimes I \otimes I \otimes Z \otimes I \\
\tilde{X}_3 &:= I \otimes I \otimes I \otimes X \otimes X \otimes I \otimes I \otimes X \otimes I \\
\tilde{Z}_4 &:= I \otimes I \otimes I \otimes Z \otimes Z \otimes I \otimes I \otimes I \otimes Z \\
\tilde{X}_4 &:= X \otimes X \otimes X \otimes X \otimes I \otimes I \otimes I \otimes I \otimes X
\end{aligned}$$

The resulting entanglement-assisted stabilizer has parameters $[[5, 1, 5; 4]]$ while the underlying stabilizer has parameters $[[9, 1, 2]]$

In example 10 an entanglement-assisted stabilizer achieved a distance of 5, while the underlying stabilizer only has a distance of 2. The added qubits possessed by the receiver correspond to qubits 6 through 9. Puncturing qubit 9 on the non-entanglement-assisted stabilizer yields the following relations:

$$\begin{aligned} [[9,1,2]] &\xrightarrow{\text{qubit 9 wrt. } X} [[8,1,1]] \\ [[9,1,2]] &\xrightarrow{\text{qubit 9 wrt. } Y} [[8,1,2]] \\ [[9,1,2]] &\xrightarrow{\text{qubit 9 wrt. } Z} [[8,1,2]]. \end{aligned}$$

The stabilizer is thus able to maintain its distance when puncturing with respect to Y or Z. Considering the entanglement-assisted version, with the same puncturing scenarios, the following relations are obtained;

$$\begin{aligned} [[5,1,5;4]] &\xrightarrow{\text{qubit 9 wrt. } X} [[5,1,1;3]] \\ [[5,1,5;4]] &\xrightarrow{\text{qubit 9 wrt. } Y} [[5,1,2;3]] \\ [[5,1,5;4]] &\xrightarrow{\text{qubit 9 wrt. } Z} [[5,1,3;3]]. \end{aligned}$$

In this case, the distance cannot be maintained for any puncturing scenario. In the best case, the distance is decreased by 2, while in the worst case, the distance is decreased by 4. Following the notion of figure 5.6, the 4 stages of codes under puncturing is summarized in table 5.1.

Original code	Puncturing scenario	View as	Puncture	View as
[[5,1,5;4]]	qubit 9 wrt. X	[[9,1,2]]	[[8,1,1]]	[[5,1,1;3]]
[[5,1,5;4]]	qubit 9 wrt. Y	[[9,1,2]]	[[8,1,2]]	[[5,1,2;3]]
[[5,1,5;4]]	qubit 9 wrt. Z	[[9,1,2]]	[[8,1,2]]	[[5,1,3;3]]

Table 5.1: The change of parameters when following the puncturing method of figure 5.6.

In this case, the greatest source of distance decrease is found when viewing the entanglement-assisted stabilizer as its non-entanglement-assisted counterpart to initiate the puncturing. In all cases, this causes a decrease of 3 in distance. The puncturing itself causes at most a decrease of 1, in fact, the distance is maintained in two of three scenarios. Lastly, viewing the punctured stabilizer as an entanglement-assisted stabilizer has the potential for an increase in distance. However, this is only experienced in one of the three puncturing scenarios.

While the upper bound for distance decrease is 1 in the non-entanglement-assisted case, such a bound cannot be satisfied in the entanglement-assisted case, as shown by the

example above. However, define $l := d - d' \geq 0$, where d is the distance of the unpunctured entanglement-assisted stabilizer and d' is the distance of the corresponding non-entanglement-assisted stabilizer, both given by figure 5.6. Puncturing an entanglement-assisted stabilizer of parameters $[[N, k, d; c]]$ can then at most decrease the distance by $l+1$. Although this provides an upper bound on the possible distance decrease, the bound may be vacuous for practical purposes. For the above example $l + 1 = 4$, which was encountered when puncturing qubit 9 with respect to X. Before puncturing, the code could protect against an arbitrary weight 2 error or any positional weight 4 error. After puncturing, the code cannot even protect against a single positional error in the worst case. Choosing the correct puncturing scenario is thus vital in the entanglement-assisted setup.

Consider a scenario in which errors occur independently on each qubit with probability p and uniformly in $\{X, Y, Z\}$. An error resulting from an outcome of this scenario can be corrected if it belongs to the correctable set of errors for the underlying error-correction code. Figure 5.7 illustrates the probability that the underlying code can correct errors resulting from an experiment of 100,000 independent trials for each probability $p \in \{0.01, 0.02, \ldots, 0.99, 1.0\}$ on the entanglement-assisted stabilizers of table 5.1. The figure highlights the importance of finding an optimal puncturing scenario. Puncturing qubit 9 with respect to Z yields an entanglement-assisted stabilizer with error-correcting abilities close to the unpunctured version. These abilities change drastically when puncturing with respect to X or Y, although all scenarios solve the problem of inadequately supplied Bell pairs described in section 5.2. Denote by $\mathcal{A}$ the correctable set of errors for any of the error-correcting codes examined in figure 5.7. The probability of an error, E, from the independent trials being identical to some $E' \in \mathcal{A}$ is given by

$$\mathbb{P}(E = E') = \left(\frac{p}{3}\right)^{w(E')} (1-p)^{5-w(E')},$$

where w is the weight map defined in (4.6). The probability that E is correctable, meaning the probability that $E \in \mathcal{A}$, is therefore

$$\mathbb{P}(E \in \mathcal{A}) = \sum_{E' \in \mathcal{A}} \left(\frac{p}{3}\right)^{w(E')} (1-p)^{5-w(E')},$$

In all four cases examined, the maximum absolute deviation from the theoretical probability across all codes and all p is 0.00368. The average absolute deviation is 0.00081.

Table 5.2 is an extension of table 5.1 providing parameters for all puncturing scenarios involving the receiver's half of required Bell pairs for the $[[5, 1, 5; 4]]$ entanglement-assisted stabilizer. Puncturing qubit 9 experiences almost all possible distance change scenarios except one. Puncturing qubits 7 or 8 with respect to Y yields a distance decrease of only one. Either of these two scenarios are thus the optimal puncturing scenario with respect to distance for solving the inadequately supplied Bell pairs problem.

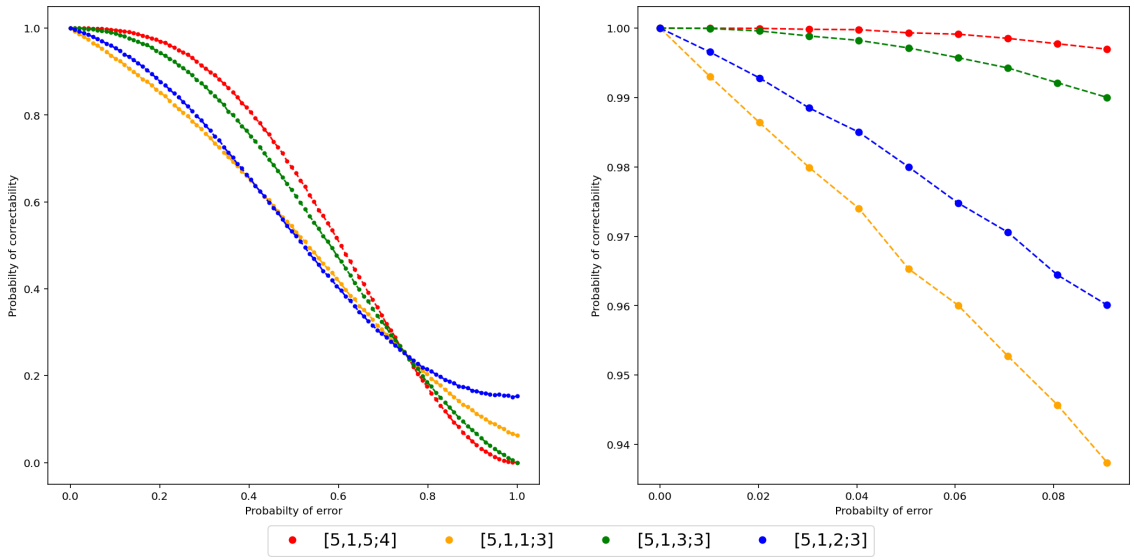


Figure 5.7: Probability of correctability as a function of probability of error on an entanglement-assisted stabilizer and three of its punctured variants.

Original code	Puncturing scenario	View as	Puncture	View as
[[5,1,5;4]]	qubit 9 wrt. X	[[9,1,2]]	[[8,1,1]]	[[5,1,1;3]]
[[5,1,5;4]]	qubit 9 wrt. Y	[[9,1,2]]	[[8,1,2]]	[[5,1,2;3]]
[[5,1,5;4]]	qubit 9 wrt. Z	[[9,1,2]]	[[8,1,2]]	[[5,1,3;3]]
[[5,1,5;4]]	qubit 8 wrt. X	[[9,1,2]]	[[8,1,2]]	[[5,1,3;3]]
[[5,1,5;4]]	qubit 8 wrt. Y	[[9,1,2]]	[[8,1,2]]	[[5,1,4;3]]
[[5,1,5;4]]	qubit 8 wrt. Z	[[9,1,2]]	[[8,1,2]]	[[5,1,3;3]]
[[5,1,5;4]]	qubit 7 wrt. X	[[9,1,2]]	[[8,1,2]]	[[5,1,3;3]]
[[5,1,5;4]]	qubit 7 wrt. Y	[[9,1,2]]	[[8,1,2]]	[[5,1,4;3]]
[[5,1,5;4]]	qubit 7 wrt. Z	[[9,1,2]]	[[8,1,2]]	[[5,1,3;3]]
[[5,1,5;4]]	qubit 6 wrt. X	[[9,1,2]]	[[8,1,1]]	[[5,1,1;3]]
[[5,1,5;4]]	qubit 6 wrt. Y	[[9,1,2]]	[[8,1,2]]	[[5,1,2;3]]
[[5,1,5;4]]	qubit 6 wrt. Z	[[9,1,2]]	[[8,1,2]]	[[5,1,3;3]]

Table 5.2: The change of parameters when following the puncturing method of figure 5.6.

According to the quantum Singleton bound, the best possible distance for a stabilizer encoding 1 logical qubits in 9 physical qubits is 5. This is achieved in the pre-punctured entanglement-assisted stabilizer. After puncturing, the stabilizer encodes 1 logical qubits in 8 physical qubits. In this case, the best possible distance is 4. The Singleton bound is

thus saturated for this setup, if and only if the puncturing is performed with respect to Y on qubit 7 or 8.

A method to puncture the qubits possessed by the receiver in an entanglement-assisted stabilizer setup has been derived. The method rely on existing methods for non-entanglement-assisted stabilizers and is a step towards a dynamical communication framework for entanglement-assisted stabilizers. However, the method possess some caveats, namely the unpredictable change in distance. While it may be feasible to explore every puncturing scenario in the above example, for entanglement-assisted stabilizers involving a greater number of qubits, that may not be the case.

6 | Conclusion and Discussion

Puncturing of error-correcting codes has been extensively studied in the classical setting, and more recently, a corresponding framework has been developed for quantum stabilizer codes. In this report, the framework is extended to encompass entanglement-assisted (ea) stabilizer codes, restricting attention to the puncturing of qubits held by the receiver prior to the communication phase. This extension allows for a dynamic coding strategy in a situation where the base station may be unable to provide the necessary Bell pairs required for the ea stabilizer.

Any ea stabilizer may be viewed as a non-ea stabilizer, possibly at the cost of a decrease in distance. The developed framework takes advantage of this fact; rather than developing a strategy directly for ea stabilizers, it transforms the ea stabilizer into a non-ea stabilizer, performs the puncturing, then transforms the punctured stabilizer back to an ea stabilizer. The developed strategy transforms the coding parameters by

$$[[N = c + s + k, k, d; c]] \quad \xrightarrow{\text{puncturing}} \quad [[N = (c-1) + (s+1) + k, k, d'; c-1]],$$

assuming $d \geq 2$. If $d = 1$, logical information may be lost under puncturing; it is therefore generally not advised to puncture codes with distance 1. The developed method decreases the number of required Bell pairs by 1 while maintaining N and k. This implies that the qubit on the transmitter's end, previously part of the removed Bell pair, is transformed into an ancilla qubit. The parameters N, k, and c are thus mapped in a controlled manner under puncturing. This is unfortunately not the case for the distance d. Generally, the distance d undergoes 3 transformations when performing puncturing in the developed method. The first is when viewing the ea stabilizer as a non-ea stabilizer. This may lead to a decrease of $l \in \mathbb{N}_0$, since the logical errors for an ea stabilizer are a subset of those for the corresponding non-ea stabilizer, as the receiver's qubits in an ea setup are treated as error-free. Secondly, the puncturing may decrease the distance of the stabilizer by at most 1. Thirdly, when viewing the punctured non-ea stabilizer as an ea stabilizer, a distance increase may be experienced. The distance loss is thus bounded by $l + 1$.
The puncturing method may be recursively executed yielding the transformation

$$[[N = c + s + k, k, d; c]] \quad \xrightarrow{\text{puncturing}} \quad [[N = (c-q) + (s+q) + k, k, d'; c-q]],$$

for $q \in \{1, \ldots, c\}$ assuming the distance is at least 2 throughout.
A $[[5, 1, 5; 4]]$ code was examined. The code was transformed in the following way:

$$[[5, 1, 5; 4]] \quad \xrightarrow{\text{view as}} \quad [[9, 1, 2]] \quad \xrightarrow{\text{puncturing}} \quad [[8, 1, d]] \quad \xrightarrow{\text{view as}} \quad [[5, 1, d'; 3]].$$

where $d \in \{1, 2\}$ and $d' \in \{1, 2, 3, 4\}$. This suggest the distance is very volatile under puncturing. In this case $l = 3$ the distance decrease is thus bounded by 4 which was achieved in 2 out of 12 puncturing scenarios. Such a decrease is detrimental to the error correction, as the code transforms from being able to protect against 2 arbitrary error or 4 positional errors, to not being able to protect against a single positional errors.

The distance under puncturing for an ea stabilizer is very volatile. In future work, it should be explored whether an optimal puncturing scenario with respect to distance could be derived as a function of the examined ea stabilizer. In this report, all possible scenarios to puncture a single qubit on the receiver's end were explored. While this was feasible when the receiver was in possession of only 4 qubits, this may not be the case generally. For a receiver in possession of c qubits, the number of puncturing scenarios is $3c$.

It also remains to examine the transformation of encoders under puncturing. In the classical setting, puncturing a bit corresponds to removing a column of the generator matrix (or removing the corresponding bit after encoding). This allows for a dynamic encoding strategy. In the quantum setting, encoding is not as simple. For the developed puncturing strategy to be implementable, there should exist a simple transformation of the encoding circuit; transforming the unpunctured encoding circuit into the punctured encoding circuit. Preferably by the removal of a number of gates. Unfortunately, this is not satisfied for the encoders examined in this report. However, encoders are not unique, and perhaps a transformation of encoders could reveal a satisfiable relationship. If no such relationship exist, encoder for all possible puncturing scenarios should be implemented. This eliminates the dynamic advantage in puncturing as one could equivalently store optimized ea stabilizers requiring $c - q$ Bell pairs for $q \in \{0, \ldots, c\}$.

Lastly, it remains to examine the effect of utilizing imperfect Bell pairs in an ea stabilizer setup. In practice, one may be forced to accept Bell pairs with fidelity less than one, despite the application of a distillation protocol. How this affects the error-correction capabilities of an ea stabilizer should be studied in depth before practical deployment.

Bibliography

Frank Arute et al. Quantum supremacy using a programmable superconducting processor. *Nature*, 574, 2019. ISSN 1476-4687.

Charles Averill. A brief analysis of the apollo guidance computer. *arXiv preprint arXiv:2201.08230*, 2022.

János A. Bergou, Mark Hillery, and Mark Saffman. *Quantum Information Processing*. Springer, 2nd edition, 2021.

Todd Brun, Igor Devetak, and Min-Hsiu Hsieh. Correcting quantum errors with entanglement. *Science*, 314:436–439, 2006. ISSN 0036-8075.

Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. John Wiley & Sons, 2nd edition, 2006.

David S. Dummit and Richard M. Foote. *Abstract Algebra*. John Wiley & Sons, 3rd edition, 2004.

Craig Gidney and Martin Ekerå. How to factor 2048 bit rsa integers in 8 hours using 20 million noisy qubits. *Quantum*, 5, 2021. ISSN 2521-327X.

Daniel Gottesman. Stabilizer codes and quantum error correction. *Ph.D. Thesis, arXiv:9705052v1*, 1997.

Jaron S. Gundersen, René Bødker Christensen, Markus Grassl, Petar Popovski, and Rafał Wisniewski. Puncturing quantum stabilizer codes. *Journal on Selected Areas in Information Theory*, 6:74 – 84, 2025. ISSN 2641-8770.

Johan P. Hansen et al. *Matematiske Mysterier: Historien, forklaringerne og løsningerne*. Aarhus Universitetsforlag, 1st edition, 2013.

Jørn Justesen and Tom Høholdt. *A Course In Error-Correcting Codes*. European Mathematical Society, 1st edition, 2004.

Ching-Yi Lai and Todd A. Brun. Entanglement-assisted quantum error-correcting codes with imperfect ebits. *Phys. Rev. A*, 86:032319, 2012. ISSN 2469-9934.

LLNL. El capitan. URL `https://hpc.llnl.gov/hardware/compute-platforms/el-capitan`. Accessed: 23 January 2026.

Merriam-Webster. Computer definition. URL https://www.merriam-webster.com/dictionary/computer. Accessed: 10 February 2026.

Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge, 10th edition, 2010.

Gerard O'Regan. *A Brief History of Computing*. Springer, 4th edition, 2026.

John Preskill. Quantum computing 40 years later. *arXiv preprint:2106.10522*, 2021.

Live Science. Scientists hit quantum computer error rate of 0.000015 URL https://www.livescience.com/technology/computing/scientists-hit-quantum-computer-error-rate-of-0-000015-percent-a-world-record-achievem Accessed: 3 February 2026.

Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26:1484–1512, 1997. ISSN 0097-5397.

Graeme Smmith and John A. Smolin. Degenerate quantum codes for pauli channels. *Physical Review Letters*, 98, 2006. ISSN 1079-7114.

IEEE Spectrum. Quantinuum claims key step toward scaling up quantum computers. URL https://spectrum.ieee.org/quantinuum-fault-tolerant-quantum-computing. Accessed: 3 February 2026.

SpinQ. Types of quantum computers you need to know. URL https://www.spinquanta.com/news-detail/types-of-quantum-computers-you-need-to-know-in20250226071709. Accessed: 29 January 2026.

Top500. Super computers ranked. URL https://top500.org/. Accessed: 10 February 2026.

Mark McMahon Wilde. Quantum coding with entanglement. *Ph.D. Thesis, arXiv:0806.4214*, 2008.

Appendices

A | Group Theory

This appendix provides the necessary group theory used in section 4.3. The appendix is mainly based on [Dummit and Foote, 2004].

Group theory is the study of the algebraic structure of groups. From these simple definitions, powerful concepts can be developed that apply to a wide variety of sets. Thus, the study begins with the definition of a group.

Definition A.1: Group
The pair $(G, *)$ consisting of a set G and a binary operation $* : G \times G \to G$ is called a group if the following axioms are satisfied:

1. $(a * b) * c = a * (b * c), \qquad \forall a, b, c \in G \qquad$ (Associative Axiom),
2. $\exists e \in G : \ a * e = e * a = a, \qquad \forall a \in G \qquad$ (Identity Axiom),
3. $\forall a \in G \exists a^{-1} \in G : \ a * a^{-1} = a^{-1} * a = e \qquad$ (Inverse Axiom).

[Dummit and Foote, 2004, pp.16-17]

Let $(G, *)$ be a group, we say G is a group under $*$. If the elements of a group $(G, *)$ commute, that is, $a * b = b * a$ for all $a, b \in G$, the group is called abelian. In the following appendix, the definition $a^k := \underbrace{a * \cdots * a}_{k \text{ times}}$ for some $k \in \mathbb{N}$ is used.

The size of a group, $(G, *)$, is the number of elements in the set G, i.e. the sets cardinality. The size is denoted $|G|$. A group is called finite if $|G| < \infty$.

Instead of listing every element of a group, one can specify a small subset of the elements, from which all the other elements can be derived using the binary operation $*$. This is concept of generating sets, providing a compact description of a group.

Definition A.2: Generating Sets
Let $(G, *)$ be a group. A generating set for $(G, *)$ is a subset $\{g_1, \ldots, g_N\} \subseteq G$ such that $a = g_1^{r_1} * \cdots * g_N^{r_N}$ for all $a \in G$ with $r_i \in \mathbb{Z}$. The elements g_i are called generators of $(G, *)$.

[Dummit and Foote, 2004, p.26]

If $(G, *)$ is a group and $\{g_1, \ldots, g_N\}$ is a generating set for the group we write $\langle g_1, \ldots, g_N \rangle = G$.

Theorem A.3: Size of a Generator
Let $(G, *)$ be a group. Then there exists a generating set for the group containing at most $\log_2(|G|)$ elements.

[Nielsen and Chuang, 2010, p.611]

Proof
Let $G = \langle g_1, \ldots, g_N \rangle$ be a finite group. For every $n \in \{1, \ldots, N\}$, define $G_n := \langle g_1, \ldots, g_n \rangle$. Assume $g \notin G_n$ for some $g \in G$ and let $i \in \{1, \ldots, n\}$, then this implies $g * g_i \notin G_n$, as the converse would imply $g = (g * g_i) * g_i^{-1} \in G_n$ which is false by assumption. Letting $g = g_{n+1}$ the new group G_{n+1} thus has at least the elements of G_n and the additional $g * g_i$ elements, implying $2|G_n| \leq |G_{n+1}|$. Recursively this means

$$2^N \leq 2^{N-1}|G_1| \leq 2^{N-2}|G_2| \leq \cdots \leq 2^2|G_{N-2}| \leq 2|G_{N-1}| \leq |G_N| = |G|,$$

implying $N \leq \log_2(|G|)$. ∎

Let G be a set of matrices and the binary operation $*$ be matrix multiplication. Matrix multiplication satisfies the associative axiom for any three matrices. The set G is hence a group under matrix multiplication if it is closed under matrix multiplication, contains the identity matrix and every element in G is invertible, with the inverse contained in G. A group consisting of matrices under matrix multiplication is called a matrix group.

Example 11: A Simple Group
Consider the set

$$G = \{\pm I, \pm X\}.$$

Under matrix multiplication this is a group. It contains the identity and since both I and X are unitary and hermitian, they are their own inverses. This implies the set contains the inverses of every element. It only remains to show the set is closed under matrix multiplication. This can be done by the use of a multiplication table as shown in table A.1.

$*$	I	$-I$	X	$-X$
I	I	$-I$	X	$-X$
$-I$	$-I$	I	$-X$	X
X	X	$-X$	I	$-I$
$-X$	$-X$	X	$-I$	I

Table A.1: Multiplication table for the Pauli group.

A more commonly encountered group when studying quantum error-correcting codes is the Pauli group consisting of the set;

$$\mathcal{P}_1 := \{\pm I, \pm X, \pm Y, \pm Z, \pm iI, \pm iX, \pm iY, \pm iZ\},$$

paired with matrix multiplication. It is evident that for the set of example 11, $G \subseteq \mathcal{P}_1$. The group of example 11 is therefore called a subgroup of the Pauli group.

Definition A.4: Subgroup
Let $(G, *)$ be a group. If $H \subseteq G$ and $(H, *)$ is a group, then $(H, *)$ is a subgroup of $(G, *)$.

[Dummit and Foote, 2004, p.46]

Given a subset of a group, the next theorem provides a simple method of verifying if the subset is also a subgroup.

Theorem A.5: The Subgroup Criterion
Let $(G, *)$ be a group and $H \subseteq G$. Then $(H, *)$ is a group if and only if $H \neq \varnothing$ and for all $a, b \in H$ it follows $a * b^{-1} \in H$.

[Dummit and Foote, 2004, p.47]

Proof
Assume $(H, *)$ is a subgroup of $(G, *)$, then H contains at least the identity element e implying $H \neq \varnothing$. Furthermore, if $b \in H$, then $b^{-1} \in H$ and H is closed under $*$. Thus if $a, b \in H$ then $a * b^{-1} \in H$.
Conversely, assume $H \subseteq G$ is non-empty and if $a, b \in H$ then $a * b^{-1} \in H$. Let $a \in H$ and choose $b = a$, then $e = a * a^{-1} \in H$. Since H contains the identity, choose $a = e$ and $b = a$, then $a^{-1} = e * a^{-1} \in H$. It only remains to show that H is closed under $*$. Suppose $a, b \in H$ since H is closed under inverses, it follows $b^{-1} \in H$ and thus $a * (b^{-1})^{-1} = a * b \in H$. ∎

If $(G, *)$ is group, $H \subseteq G$ and H is finite. It suffices to check that H is non-empty and closed under $*$ to conclude that $(H, *)$ is a subgroup of $(G, *)$. To realize this, let H be finite, closed under multiplication and non-empty. Now let $a \in H$ and consider the sequence $h,\ h^2,\ h^3, \ldots$. Since the set is finite and closed under $*$, the sequence lies in H and must therefore eventually repeat, meaning there exist $n, m \in N$ with $n < m$ such $h^n = h^m$, setting $l = m - n > 0$ it follows $h^l = h^{m-n} = e$. Furthermore, $h^{-1} = h^l * h^{-1} = h^{l-1} \in H$.

The study of group theory ends with a short introduction of some important groups, beginning with the centralizer group.

Definition A.6: Centralizer
Let $(G, *)$ be a group and $A \subseteq G$ be a non-empty subset. The centralizer of A in G is the set defined by

$$C_G(A) := \{g \in G \mid g * a * g^{-1} = a, \quad \forall a \in A\}.$$

[Dummit and Foote, 2004, pp.49-50]

Since $g * a * g^{-1} = a \Leftrightarrow g * a = a * g$ the centralizer of A in G is exactly the elements of G which commutes with every element of A. Choosing $A = G$ one obtains the set of all elements in G commuting with all elements of G, this set is called the center of G denoted $Z(G) := C_G(G)$. For an abelian group, $G = Z(G)$. Conversely, if $G = Z(G)$ then G is an abelian group.

The centralizer is a group under $*$ since $e * a * e = a$ for all $a \in A$ thus $e \in C_G(A) \implies C_G(A) \neq \varnothing$. Furthermore, assuming $b, c \in C_G(A)$ it follows $c*a*c^{-1} = a \Leftrightarrow a*c^{-1} = c^{-1}*a$ implying $c^{-1} \in C_G(A)$ thus $(b * c^{-1}) * a = b * (a * c^{-1}) = (a * b) * c^{-1} = a * (b * c^{-1})$ hence $b * c^{-1} \in C_G(A)$, the statement now follows from the subgroup criterion.

The next group of importance is the normalizer group.

Definition A.7: Normalizer
Let $(G, *)$ be a group and $A \subseteq G$ be non-empty. Define $g*A*g^{-1} := \{g*a*g^{-1} \mid a \in A\}$. The normalizer of A in G is the set $N_G(A) := \{g \in G \mid g * A * g^{-1} = A\}$.

[Dummit and Foote, 2004, p.50]

For any $g \in N_G(A)$, $g*A*g^{-1}$ induces a permutation of the elements of A; that is, it maps each $a \in A$ to another element of A. In the special case where the induced permutation is the identity permutation the corresponding elements form the centralizer. Consequently, $C_G(A) \subseteq N_G(A)$.

The normalizer is also a group under $*$ since $e*A*e = A$ thus $e \in N_G(A) \implies N_G(A) \neq \varnothing$ and if $b, c \in N_G(A)$ then $(b * c^{-1}) * A * (c * b^{-1}) = b * (c^{-1} * A * c) * b^{-1} = b * A * b^{-1} = A$ thus $b * c^{-1} \in N_G(A)$. The statement then follows from the subgroup criterion.

Two important properties of the normalizer, when considering the N-fold Pauli group, are presented in the following theorem.

Theorem A.8: Properties of the Normalizer
Let $G \subseteq \mathcal{P}_N$ be a subgroup of the N-fold Pauli group.

1. $G \subseteq N_{\mathcal{P}_N}(G)$,
2. If $-I \notin G$ then $N_{\mathcal{P}_N}(G) = C_{\mathcal{P}_N}(G)$.

[Nielsen and Chuang, 2010, pp.465-466]

Proof
(1)
Let $g \in G$, since G is a group, it follows by the inverse axiom $g^{-1} \in G$. For any $a \in G$, closure under $*$ entails, $g * a * g^{-1} \in G$. Since a was arbitrary, it must hold $g * G * g^{-1} \subseteq G$. Substitution of g by its inverse in the above reasoning analogously reveals $g^{-1} * G * g \subseteq G$

and thus $G \subseteq g * G * g^{-1}$. In conjunction of the two derivations it can be concluded $g * G * g^{-1} = G$, from the definition of the normalizer it then follows $g \in N_{\mathcal{P}_N}(G)$, and since g was arbitrary $G \subseteq N_{\mathcal{P}_N}(G)$.

(2)

It has already been established that $C_{\mathcal{P}_N}(G) \subseteq N_{\mathcal{P}_N}(G)$, thus it remains to show the reverse inclusion. Any two Pauli operators either commute or anti-commute, that is for any two $a, b \in \mathcal{P}_N$ it holds $a * b = \pm b * a$ and thus $a * b * a^{-1} = \pm b$. Let $g \in N_{\mathcal{P}_N}(G) \subseteq \mathcal{P}_N$, by the definition of the normalizer it then follows that for every $a \in G \subseteq \mathcal{P}_N$; $g * a * g^{-1} = \pm a \in G$. Assume $g * a * g^{-1} = -a \in G$. Because $a \in G$, $a^{-1} \in G$. Consequently, due to closure under $*$, $G \ni (-a) * a^{-1} = -I$, which is false by assumption, thus it must be the case $g * a * g^{-1} = a$, and hence $g \in C_{\mathcal{P}_N}(G)$, which implies $N_{\mathcal{P}_N}(G) \subseteq C_{\mathcal{P}_N}(G)$. ∎

For any subgroup G of $\mathcal{P}_N$ with $-I \notin G$, the Venn diagram of figure A.1 depicts the relationsship between G, $N_{\mathcal{P}_N}(G)$ and $\mathcal{P}_N$.

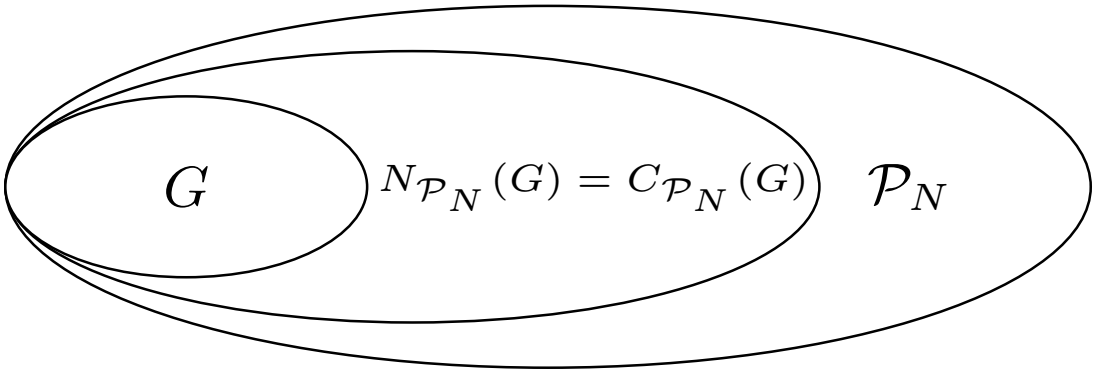


Figure A.1: The relationship between a group, G, the corresponding normalizer and the N-fold Pauli group.

Before concluding the appendix, the notion of normal groups and quotient groups are introduced, starting with normal groups.

Definition A.9: Normal Subgroups

Let $(N, *)$ be a subgroup of some group $(G, *)$. $(N, *)$ is then called a normal subgroup of G if for all $g \in G$, $gNg^{-1} = N$.

[Dummit and Foote, 2004, p.82]

Let $(H, *)$ be a subgroup of $(G, *)$ and g an element of G. The sets defined by $gH := \{gh \mid h \in H\}$ ($Hg := \{hg \mid h \in H\}$) is called a left (right) coset of H in G. For a subgroup $(N, *)$ of $(G, *)$, an equivalent definition to definition A.9, is that $(N, *)$ is normal if the left and right coset of N in G are equal for all $g \in G$ since $gNg^{-1} = N \Leftrightarrow gN = Ng$ [Dummit and Foote, 2004, p.77].

The normal subgroups are used to define quotient groups.

Definition A.10: Quotient Groups
Let $(N, *)$ be a normal subgroup of $(G, *)$. The set defined by

$$G/N := \{g * N \mid g \in G\}$$

paired with the binary operation defined by

$$*_{\text{quo}} : (a * N) *_{\text{quo}} (b * N) \mapsto (a * b) * N, a, b \in G$$

is called the quotient group.

[Dummit and Foote, 2004, p.77]

The quotient group G/N is thus the set of all left cosets of N in G. Although the definition defines a set and a binary operation on the set, it strictly does not justify calling the pair a group. In fact, it has not even been proven that the binary operation $*_{\text{quo}}$ is well defined. Consider the four elements $g, g', h, h' \in G$. Assume $g * N = g' * N$ and $h * N = h' * N$. For $*_{\text{quo}}$ to be well defined it is required that $(g * N) *_{\text{quo}} (h * N) = (g' * N) *_{\text{quo}} (h' * N)$. However, since N is normal, this is satisfied

$$\begin{aligned}(g * N) *_{\text{quo}} (h * N) &= ((g * h) * N) = (g * (h * N)) = (g * (h' * N)) = (g * (N * h')) \\ &= ((g * N) * h') = ((g' * N) * h') = (g' * (N * h')) = (g' * (h' * N)) \\ &= ((g' * h') * N) = (g' * N) *_{\text{quo}} (h' * N).\end{aligned}$$

With the group operation being well defined, it can be justified that G/N is in fact a group. Since G is closed under $*$ and $*$ is associative, it follows that G/N is closed under $*_{\text{quo}}$ and $*_{\text{quo}}$ is associative. Furthermore, the element $e * N$ is the identity element of G/N where e is the identity element of G and for any element $g \in G$, the inverse element of $g * N$ is the element $g^{-1} * N$. Both $e * N$ and $g^{-1} * N$ belong to G/N since $e, g^{-1} \in G$.

The center of $\mathcal{P}_1$ are given by the normal subgroup $Z(\mathcal{P}_1) = \{I, -I, iI, -iI\}$ of $\mathcal{P}_1$. An important quotient group encountered when studying quantum error correction is the phase independent Pauli group given by

$$\mathcal{P}_1/Z(\mathcal{P}_1) = \{[I], [X], [Y], [Z]\},$$

where

$$\begin{aligned}[I] &= \{I, -I, iI, -iI\}, & [X] &= \{X, -X, iX, -iX\}, \\ [Y] &= \{Y, -Y, iY, -iY\}, & [Z] &= \{Z, -Z, iZ, -iZ\}.\end{aligned}$$

B | Proof of Theorem 4.5

This appendix proves the statement of theorem 4.5. The appendix is primarily based on [Bergou et al., 2021, pp.184-186].

Before proceeding to the proof, a brief overview of necessary theory is introduced.

Let $g_1, g_2 \in \mathcal{P}_1$, they can then be expressed by

$$g_1 = \alpha_1 X^{a_1} Z^{b_1} \qquad \text{and} \qquad g_2 = \alpha_2 X^{a_2} Z^{b_2}, \qquad \alpha_i \in \{\pm 1, \pm i\}.$$

Then $g_1 g_2 = g_2 g_1$ if and only if $a_1 b_2 + b_1 a_2 = 0$ where addition is modulo 2. For $g_1, g_2 \in \mathcal{P}_N$, the expression is extended as

$$g_1 = \alpha_1 \bigotimes_{i=1}^{N} X^{a_{1,i}} Z^{b_{1,i}} \qquad \text{and} \qquad g_2 = \alpha_2 \bigotimes_{i=1}^{N} X^{a_{2,i}} Z^{b_{2,i}}, \qquad \alpha_i \in \{\pm 1, \pm i\}.$$

Define $g_{1,i} := X^{a_{1,i}} Z^{b_{1,i}}$ and $g_{2,i} := X^{a_{2,i}} Z^{b_{2,i}}$, then g_1 and g_2 commute if and only if all $g_{1,i}$ commutes with their corresponding $g_{2,i}$ or and even number of pairs anti-commute. Thus, $g_1 g_2 = g_2 g_1$ if and only if $\sum_{i=1}^{N} (a_{1,i} b_{2,i} + b_{1,i} a_{2,i}) = 0$ where addition is modulo 2. With this in mind, a necessary lemma can be introduced and proved.

Lemma B.1
Let $g_1, g_2 \in \mathcal{P}_N$. Define

$$\Lambda := \begin{bmatrix} 0 & I_N \\ I_N & 0 \end{bmatrix}.$$

Then g_1 and g_2 commute if and only if $r(g_1) \Lambda r(g_2)^T = 0$ where $r(g_i)$ denotes the row of the check matrix corresponding to g_i and addition is modulo 2.

[Bergou et al., 2021, pp.184-185]

Proof
Let $g_j = \alpha_i \bigotimes_{i=1}^{N} X^{a_{j,i}} Z^{b_{j,i}}$, the corresponding row of the check matrix is then given by $r(g_j) = [a_{j,1} \ \dots \ a_{j,N} \ b_{j,1} \ \dots \ b_{j,N}]$. Now, consider the expression;

$$\begin{aligned} r(g_1) \Lambda r(g_2)^T &= [a_{1,1} \ \dots \ a_{1,N} \ b_{1,1} \ \dots \ b_{1,N}] \, [b_{2,1} \ \dots \ b_{2,N} \ a_{2,1} \ \dots \ a_{2,N}]^T \\ &= \sum_{i=1}^{N} (a_{1,i} b_{2,i} + b_{1,i} a_{2,i}). \end{aligned}$$

With the paragraph leading to the lemma in mind, the proof is concluded. ∎

Since any two $g_1, g_2 \in \mathcal{P}_N$ either commutes or anti-commutes, it follows directly from the lemma that $g_1 g_2 = -g_2 g_1$ if and only if $r(g_1)\Lambda r(g_2)^T = 1$.

The theory necessary to prove theorem 4.5 has now been discussed and the proof can commence.

Proof (Proof of Theorem 4.5)

Let $G = \langle g_1, \ldots, g_{N-k} \rangle$ be an abelian subgroup of $\mathcal{P}_N$ such $-I \notin G$ and $g_1, \ldots, g_{N-k}$ are independent. Then g_i operates on the 2^N dimensional state space $\mathcal{H} = \mathbb{C}^{2^N}$. Let further $x \in \mathbb{F}_2^{N-k}$ and define

$$P_x := \prod_{i=1}^{N-k} \frac{I + (-1)^{x_i} g_i}{2}.$$

Since $I \pm g_i$ project onto the ± 1 eigenspace of g_i, P_x projects onto the intersection of the $(-1)^{x_i}$ eigenspaces. Define $P_0 := P_{[0 \ \ldots \ 0]}$, then P_0 projects onto the code space V_G. Furthermore, it is evident that there exists 2^{N-k} projectors and these are mutually orthogonal. The orthogonality comes from the fact that g_i commute which implies $I + (-1)^{x_i} g_i$ commute and thus if $x \neq x'$, that is they differ in at least the $j'th$ entry, then $P_x P_{x'}$ will contain a factor $(I + g_j)(I - g_j) = 0$ hence $P_x P_{x'} = 0$. It is also the case that

$$\sum_{x \in \mathbb{F}_2^{N-k}} P_x = I \tag{B.1}$$

To realize this, define

$$Q_m := \sum_{x \in \mathbb{F}_2^m} \prod_{i=1}^{m} \frac{I + (-1)^{x_i} g_i}{2}.$$

Obviously $Q_1 = I$, furthermore

$$\begin{aligned} Q_{m+1} &= \sum_{x \in \mathbb{F}_2^{m+1}} \prod_{i=1}^{m+1} \frac{I + (-1)^{x_i} g_i}{2} \\ &= \sum_{x_{m+1} \in \{0,1\}} \left(\frac{I + (-1)^{x_{m+1}} g_{m+1}}{2} \right) \sum_{x \in \mathbb{F}_2^m} \prod_{i=1}^{m} \frac{I + (-1)^{x_i} g_i}{2} \\ &= \frac{Q_m + g_{m+1} Q_m}{2} + \frac{Q_m - g_{m+1} Q_m}{2} = Q_m, \end{aligned}$$

which proves $Q_{N-k} = I$, the assertion of (B.1). Denote by $P_x \mathcal{H}$ the subspace of $\mathcal{H}$ for which P_x projects onto. Since $P_x \mathcal{H}$ are mutually orthogonal and collectively span $\mathcal{H}$, the union of any basis of $P_x \mathcal{H}$ is a basis of $\mathcal{H}$, consequently

$$\dim(\mathcal{H}) = \sum_{x \in \mathbb{F}_2^{N-k}} \dim(P_x \mathcal{H}) = 2^N. \tag{B.2}$$

Since G is abelian and any two member of $\mathcal{P}_N$ either commute or anti-commute,

$$r(g_i)\Lambda r(y)^T = \begin{cases} 0, & \text{if } g_i y = y g_i, \\ 1, & \text{if } g_i y = -y g_i, \end{cases} \tag{B.3}$$

for g_i any generator of G and $y \in \mathcal{P}_N$ according to lemma B.1. Furthermore, since $-I \notin G$ and its generators are independent, it follows that the rows of the check matrix are linearly independent, according to theorem 4.4. This implies the check matrix, denoted C, has full rank. The same applies to Λ, meaning $\text{rank}(C\Lambda) = \text{rank}(C) = N - k$. The map $C\Lambda : \mathbb{F}_2^{2N} \to \mathbb{F}_2^{N-k}$ is therefore onto, and thus for any $x \in \mathbb{F}_2^{N-k}$ there exists $y \in \mathcal{P}_N$ such

$$C\Lambda r(y)^T = x.$$

Based on (B.3), this equation corresponds to $yg_i = (-1)^{x_i} g_i y$, hence

$$yP_0y^\dagger = y\left(\prod_{i=1}^{N-k} \frac{I + g_i}{2}\right) y^\dagger = \prod_{i=1}^{N-k} \frac{yy^\dagger + yg_iy^\dagger}{2} = \prod_{i=1}^{N-k} \frac{I + (-1)^{x_i} g_i}{2} = P_x.$$

This implies

$$\dim(P_0\mathcal{H}) = \text{rank}(P_0) = \text{rank}(yP_0y^\dagger) = \text{rank}(P_x) = \dim(P_x\mathcal{H}),$$

and thus with (B.2) in mind

$$\dim(P_x\mathcal{H}) = 2^N 2^{-(N-k)} = 2^k$$

proving that $\dim(V_G = P_0\mathcal{H}) = 2^k$. ∎

C | Software

Throughout the project, 3 Python scripts have been developed and are attached to the report. The scripts are described in detail below.

`stabilizer_group` : This scripts contains a class by the name `examine_group`. The class takes as input a subgroup of the N-fold Pauli group or its associated check matrix and is able to return a number of properties of the group. Including but not limited to; whether the group is abelian, the extended group (if the group is not abelian), the centralizer matrix, the parameters of the stabilizer generated by the group (after extension if the group is not abelian) and it can puncture the stabilizer generated by the group, returning the punctured centralizer matrix of the punctured code.

`examine_stabilizer` : This script serves as an interface for the constructed class in `stabilizer_group`. When running the script, the user can provide a group it wishes to examine. The script then returns the parameters of the associated stabilizer if the group is abelian. If not the group add Bell pairs and return the parameters of the entanglement-assisted stabilizer. The user also has the ability to puncture the provided stabilizer and get the parameters of the punctured stabilizer in return. If the stabilizer is entanglement-assisted, only the added qubits on the receiver's end can be punctured. The tables of section 5.3 have been constructed using this script.

`error_properties` : This script takes as input a group and examines the error-correcting abilities of the underlying stabilizer before and after puncturing. If the provided group is not abelian, the script adds Bell pairs, making the underlying stabilizer entanglement-assisted. Figure 5.7 have been constructed using this script.