Constructions and Applications of W-States

A Major Qualifying Project Submitted to the Faculty of

Worcester Polytechnic Institute

in partial fulfillment of the requirements for the Degree in Bachelor of Science in

Physics

and

Mathematical Sciences

By

James McClung

Date

May 17, 2020

Advisors

Professor P. K. Aravind

Professor Umberto Mosco

Abstract

After introducing the fundamentals of quantum circuits for a general audience, this report evaluates two means of constructing W-states: the probabilistic naive construction (PNC) and the optimized splitter construction (OSC). These methods are compared in terms of asymptotic gate cost and other parameters. Furthermore, a proof of Bell's theorem is explored with new results concerning the probability of contradicting elements of reality using W-states.

Acknowledgements

Thanks to Connor Anderson for discovering the connection between (96) and Eulerian numbers. Thanks also to Ethan Washock for helping with the same.

Thanks to all my friends in the math lounge, especially Kyle Dituro, for making my time there brighter, and to Mr. Greg Aubin for motivating me on late nights spent working.

Thanks to my parents and to Emily Staknis for their constant support.

Most of all I thank Professor P. K. Aravind for putting me to work on Qudoku over a year ago, and guiding my studies in quantum information theory all the way up to this point.

Contents

Abstract i

Acknowledgements ii

1 Introduction 1

2 Quantum Circuit Preliminaries 2

2.1 What Is a Qubit? . . . 2

2.1.1 The Bloch Sphere . . . 3

2.1.2 Matrix Representation . . . 4

2.1.3 Multiple Qubits and Entanglement . . . 4

2.2 Measurement and The Born Rule . . . 5

2.3 Special Quantum States . . . 6

2.3.1 Bell States . . . 6

2.3.2 GHZ States . . . 7

2.3.3 W-states . . . 7

2.4 Quantum Gates . . . 7

2.4.1 Single Qubit Gates . . . 8

2.4.2 Multi-Qubit Gates . . . 9

2.4.3 Properties of Gates . . . 12

2.4.4 Universality of Gates and Arbitrary Approximation . . . 14

2.5 Quantum Circuits . . . 15

2.5.1 How to Read a Circuit Diagram . . . 15

2.5.2 Constructions of Arbitrary Controlled Gates . . . 16

3 Constructions of W-states 18

3.1 On 2^k Qubits . . . 18

3.2 On n Qubits, Sometimes . . . 19

3.2.1 Optimized Probabilistic Naive Construction . . . 21

3.3 On n Qubits, Always but Inefficiently . . . 21

3.4 On n Qubits, Always and Efficiently . . . 23

4 Proof of Bell's Theorem using W-states, Without Inequalities 25

4.1 Elements of Reality . . . 25

4.2 W-states in X-Basis 25
4.3 Proof Using Three Qubits 26
4.4 Proof Using More Qubits 27

5 W-states and LOCC **29**
5.1 What Is LOCC? 29
5.2 LOCC Classes 29

6 Conclusion **31**

References **32**

Appendix A Expected Cost of Optimized Probabilistic Naive Construction **35**

List of Figures

1 Construction of $|\Psi^+\rangle$, the simplest W-state. 15

2 Construction of $|\Psi^+\rangle$using $\overline{\mathtt{C}}_1\mathtt{X}_0$. 15

3 Circuit identity for (62). 15

4 Circuit that outputs $|0\rangle$ in the lower qubit with probability $\frac{1}{2}$ and $|1\rangle$ otherwise. 16

5 Construction of arbitrary `CU` gate. 16

6 Construction of arbitrary `CCU` gate, where $\mathtt{V}^2 = \mathtt{U}$. 17

7 Construction of reverse-controlled gates. The dotted box indicates where two `X` gates canceled each other out. 17

8 Construction of $|\mathrm{W}_4\rangle$ using naive method. Left to right, the enclosed subcircuits split $|0000\rangle$ into 4 equally likely states; map $|0000\rangle$ to $|0100\rangle$; and map $|0011\rangle$ to $|1000\rangle$. 18

9 Construction of $|\mathrm{W}_8\rangle$ using naive method. First block: splitting $|0_8\rangle$ into 8-state superposition. Second block: placing 1s in desired locations. Third block: removing unwanted 1s. 19

10 Construction of $|\mathrm{W}_5\rangle$ in bottom five qubits via PNC. If any of the measured qubits return 1, the register is reset to 0s and the circuit is run again. 20

11 Construction of $|\mathrm{W}_5\rangle$ via OPNC. Block 1: splitting initial state into superposition of 8 states. Block 2: filtering out three extraneous states from superposition. Block 3: mapping remaining states to one-hot states. 21

12 Sub-optimal construction of $|\mathrm{W}_5\rangle$ using splitter gates. 22

13 Optimal construction of $|\mathrm{W}_5\rangle$ using splitter gates. 23

List of Tables

1 Some notation used in this report. 2

2 Basis states along x-, y-, and z-axes expressed as non-normalized superpositions of each other. 4

3 Properties of tensor product. 5

4 Summary of common single-qubit gates. 8

5 Input and output table of controlled-`Z` gate with different configurations of control and target. 12

6 Comparison of construction methods. "Standard gates" are those gates introduced in §2.4.1 and the `CX` gate. 24

7 Representative states for each of the six LOCC classes on three qubits. 29

1 Introduction

Quantum computing currently receives a lot of attention due to its ability to dramatically outperform ordinary (classical) computing at many tasks. Quantum cryptography, in particular, is a still-developing field with major ramifications for cybersecurity. Peter Shor's seminal algorithm [1] makes use of quantum phenomena to factor products of two large primes, which is virtually impossible for classical computers and forms the basis of modern encryption.

Quantum computers harness the power of superposition to outperform classical computers. A set of classical bits, represented by a string of 0s and 1s, represents a definite state and can be read and written to at will. By contrast, a set of quantum bits, or qubits, can be in a superposition of several bit strings at the same time. Although it is impossible to determine the overall state of a qubit, since a direct measurement "collapses" the superposition into one of its component states, it is nevertheless possible to linearly transform the entire superposition without causing collapse. These linear operations, which must be reversible, enable parallel processing of bit strings. Since adding a qubit to a register of qubits doubles the number of bit strings that can be in a superposition, the power of quantum computers grows exponentially with the number of qubits.

W-states are one of two main classes of multipartite entangled states, the other being GHZ states.[1] A W-state is an equally-weighted superposition of "one-hot" states, i.e. states in which a single qubit is in the state 1 and all the others are 0. For example, the 3-qubit W-state is

$$|\mathrm{W}_3\rangle = \frac{1}{\sqrt{3}}|001\rangle + \frac{1}{\sqrt{3}}|010\rangle + \frac{1}{\sqrt{3}}|100\rangle\,. \tag{1}$$

W-states have many applications. Jha et al. employ W-states to solve the n-queens problem [3]. W-states are also used in quantum communication [4] and cryptography [5] protocols. More generally, W-states are more robust than GHZ states in some ways [6, 7] and their form can be exploited to disprove locality [8].

W-states are difficult to construct with elementary gates, however. Most constructions are complicated and tend to be specific to the underlying quantum computer [9–12]. In §3 of this report I analyze two approaches to constructing W-states, neither of which utilizes advanced gates common to other constructions. The two approaches are compared in terms of asymptotic gate cost and other parameters. A generalization of Cabello's proof of Bell's theorem using W-states [8] is given in §4. Finally, a brief explanation of LOCC classes is given in §5.

Before presenting the new results obtained in this study, we first provide an introduction to the basics of quantum information theory, including quantum circuits, for a reader who is new to this field. An understanding of the material in §2 should be sufficient for an understanding of the rest this report. For a more thorough introduction to quantum information and circuits, the reader is referred to the excellent works of Nielsen and Chuang [13] and Mermin [14].

[1] For more than 3 qubits, other classes include the cluster states, which have their own interesting properties [2].

2 Quantum Circuit Preliminaries

This section introduces some of the basic notions of quantum computer science. A reader who is already proficient in this subject may skip to §3. This section also covers the notation that I use throughout the report, and thus may be worth skimming. Table 1 summarizes the essentials of the notation used in this report.

Symbol(s)	Meaning
$\|0\rangle$	The state represented by $\begin{bmatrix}1\\0\end{bmatrix}$ in the Z-basis (the "computational" basis)
$\|1\rangle$	The state represented by $\begin{bmatrix}0\\1\end{bmatrix}$ in the Z-basis
$\|+\rangle$	The state represented by $\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}$ in the Z-basis
$\|-\rangle$	The state represented by $\begin{bmatrix}1\\-1\end{bmatrix}$ in the Z-basis
$\|\wedge\rangle$	The state represented by $\frac{1}{\sqrt{2}}\begin{bmatrix}1\\i\end{bmatrix}$ in the Z-basis
$\|\vee\rangle$	The state represented by $\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-i\end{bmatrix}$ in the Z-basis
$\|\mathrm{W}_n\rangle$	The W-state comprising n qubits
$\mathtt{U}_i$	The gate $\mathtt{U}$ applied to qubit i
$\mathtt{C}_i\mathtt{U}_j$	The gate $\mathtt{U}$ applied to qubit j if and only if qubit i is in the state $\|1\rangle$
$\overline{\mathtt{C}}_i\mathtt{U}_j$	The gate $\mathtt{U}$ applied to qubit j if and only if qubit i is in the state $\|0\rangle$
$x := y$	x is defined to be equal to y
$x \approx y$	x is approximately equal to y
$f(n) \sim g(n)$	$f(n) \in \Theta(g(n))$, or $f(n)$ is on the order of $g(n)$

Table 1: Some notation used in this report.

2.1 What Is a Qubit?

Just as an ordinary bit (or cbit, for "classical bit") can be 0 or 1, a qubit (pronounced q-bit, for "quantum bit") can be in two distinct (i.e. orthogonal) states denoted by $|0\rangle$ and $|1\rangle$. However a qubit, unlike a cbit, can be in a superposition of two orthogonal states. Mathematically speaking, a qubit is described by a unit vector in a two-dimensional Hilbert space over the complex numbers. What this means is that the most

general state of a qubit is of the form

$$|\psi\rangle = c_0 |0\rangle + c_1 |1\rangle , \tag{2}$$

$$\text{where } |c_0|^2 + |c_1|^2 = 1. \tag{3}$$

The latter equation is called the normalization condition, and it ensures that the quantities $|c_0|^2$ and $|c_1|^2$ can be interpreted as the probabilities of finding the qubit in the state $|0\rangle$ or $|1\rangle$, respectively, if it is initially prepared in the state $|\psi\rangle$. All states will be assumed to be normalized unless stated otherwise.

The inner product of two vectors $|\psi\rangle$ and $|\varphi\rangle$, written $\langle\psi|\varphi\rangle$, is a complex number whose magnitude is at most 1. The larger the magnitude of this number, the more similar the two states are. To evaluate the inner product of the states $|\psi\rangle = c_0 |0\rangle + c_1 |1\rangle$ and $|\varphi\rangle = d_0 |0\rangle + d_1 |1\rangle$, one must first determine the "bra" of ψ, given by

$$\langle\psi| := |\psi\rangle^\dagger \tag{4}$$

$$= c_0^* \langle 0| + c_1^* \langle 1| \tag{5}$$

where * denotes complex conjugation. Then, using the fact that $\langle 0|0\rangle = \langle 1|1\rangle = 1$ and $\langle 0|1\rangle = \langle 1|0\rangle = 0$, one finds that

$$\langle\psi|\varphi\rangle = c_0^* d_0 + c_1^* d_1. \tag{6}$$

Note that if two states $|\psi\rangle$ and $|\varphi\rangle$ are normalized by (3), then $|\langle\psi|\varphi\rangle|^2 \leq 1$. In fact, (3) is sometimes expressed as $|\langle\psi|\psi\rangle| = 1$. This form of the normalization condition applies to higher dimensions, i.e. states representing more than one qubit. The physical meaning of c_0 and c_1 will be discussed further in §2.2.

2.1.1 The Bloch Sphere

The Bloch sphere provides a useful visualization of single-qubit states. Consider the unit sphere in three dimensions. Every point on the surface of the sphere corresponds to a unique pure quantum state. This may seem impossible at first, since c_0 and c_1 have a real part and an imaginary part, making a total of 4 unknowns, whereas a point on the sphere is specified by only two parameters, namely its spherical coordinates (θ, ϕ). However, the overall phase of a quantum state has no physical meaning, so c_0 can be assumed to be a real number. Furthermore, (3) allows for one of the remaining 3 unknowns to be solved for in terms of the other two. It can be shown that c_0 and c_1 can be expressed by $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi]$ as

$$c_0 = \cos\frac{\theta}{2} \tag{7}$$

$$c_1 = e^{i\phi} \sin\frac{\theta}{2}. \tag{8}$$

Points on opposite sides of the sphere are orthogonal to each other. For example, $|0\rangle$ corresponds to $\theta = 0$ (the north pole), and $|1\rangle$ corresponds to $\theta = \pi$ (the south pole). Of course, at these points, ϕ becomes meaningless.

Note that in 3-d space, $|0\rangle$ and $|1\rangle$ lie along the z-axis. Other notable antipodes are those that lie along the x-axis, here denoted $|+\rangle$ and $|-\rangle$ and those on the y-axis, denoted $|\wedge\rangle$ and $|\vee\rangle$. Any pair of antipodes is a basis for the Hilbert space in which qubits live. Table 2 shows how various basis kets can be expressed in terms of each other.

State	(θ, ϕ)	In X-Basis	In Y-Basis	In Z-Basis
$\|+\rangle$	$(\pi/2, 0)$	$\|+\rangle$	$(1-i)\|\wedge\rangle + (1+i)\|\vee\rangle$	$\|0\rangle + \|1\rangle$
$\|-\rangle$	$(\pi/2, \pi)$	$\|-\rangle$	$(1+i)\|\wedge\rangle + (1-i)\|\vee\rangle$	$\|0\rangle - \|1\rangle$
$\|\wedge\rangle$	$(\pi/2, \pi/2)$	$(1+i)\|+\rangle + (1-i)\|-\rangle$	$\|\wedge\rangle$	$\|0\rangle + i\|1\rangle$
$\|\vee\rangle$	$(\pi/2, 3\pi/2)$	$(1-i)\|+\rangle + (1+i)\|-\rangle$	$\|\vee\rangle$	$\|0\rangle - i\|1\rangle$
$\|0\rangle$	$(0, \phi)$	$\|+\rangle + \|-\rangle$	$\|\wedge\rangle + \|\vee\rangle$	$\|0\rangle$
$\|1\rangle$	(π, ϕ)	$\|+\rangle - \|-\rangle$	$-i\|\wedge\rangle + i\|\vee\rangle$	$\|1\rangle$

Table 2: Basis states along x-, y-, and z-axes expressed as non-normalized superpositions of each other.

2.1.2 Matrix Representation

Qubits are readily represented by vectors. The general 1-qubit state $c_0|0\rangle + c_1|1\rangle$ can be written as $\begin{bmatrix} c_0 \\ c_1 \end{bmatrix}$. Note that the ket became a *column* vector. The corresponding bra $c_0^*\langle 0| + c_1^*\langle 1|$ would be written as the row vector $\begin{bmatrix} c_0^* & c_1^* \end{bmatrix}$. In this way, the "braket" $\langle\psi|\varphi\rangle$ is standard matrix multiplication, and the normalization condition $\langle\psi|\psi\rangle = 1$ is quite clearly equivalent to (3).

2.1.3 Multiple Qubits and Entanglement

To deal with quantum circuits, one needs to work with the states of multiple qubits. A two-qubit state in which qubit 0 is $|0\rangle$ and qubit 1 is $|1\rangle$ can be expressed as $|10\rangle$. This convention generalizes to multi-qubit states, where the qubit states are written right-to-left. The state of qubit 0 appears at the far right of the ket; the state of qubit 1 appears to the left of qubit 0; and so on.

The notation $|10\rangle$ is really a shorthand for $|1\rangle \otimes |0\rangle$, where $\otimes$ is the tensor product. A brief summary of the properties of the tensor product is given in Table 3. For two qubits, we have the following definition, which generalizes to more qubits:

$$|ij\rangle := |i\rangle \otimes |j\rangle \tag{9}$$

$$\equiv |n_2\rangle \tag{10}$$

where $i, j \in \{0, 1\}$ and $n = i2^1 + j2^0$. Note the subscript in (10) denotes the number of qubits described by the state (clearly $|0_2\rangle \equiv |00\rangle$ is not the same as $|0_3\rangle \equiv |000\rangle$). One must not confuse $|n_q\rangle$ with $|n\rangle_q$; the former is a state of q qubits, while the latter describes the state of the qth qubit.

Property	Identity
Scalar multiplication	$c(\|\psi\rangle \otimes \|\varphi\rangle) = (c\,\|\psi\rangle) \otimes \|\varphi\rangle = \|\psi\rangle \otimes (c\,\|\varphi\rangle)$
Left distributive	$\|\psi\rangle \otimes (\|\varphi\rangle + \|\mu\rangle) = \|\psi\rangle \otimes \|\varphi\rangle + \|\psi\rangle \otimes \|\mu\rangle$
Right distributive	$(\|\psi\rangle + \|\varphi\rangle) \otimes \|\mu\rangle = \|\psi\rangle \otimes \|\mu\rangle + \|\varphi\rangle \otimes \|\mu\rangle$
Inner product	$(\langle\psi_1\| \otimes \langle\psi_2\|)(\|\varphi_1\rangle \otimes \|\varphi_2\rangle) = \langle\psi_1\|\varphi_1\rangle\langle\psi_2\|\varphi_2\rangle$

Table 3: Properties of tensor product.

If the kets are replaced with matrices, the tensor product has clearer meaning. For two arbitrary qubits $|\psi\rangle = a_0\,|0\rangle + a_1\,|1\rangle$ and $\varphi = b_0\,|0\rangle + b_1\,|1\rangle$,

$$|\varphi\rangle \otimes |\psi\rangle = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} \otimes \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} \tag{11}$$

$$= \begin{bmatrix} b_0 a_0 \\ b_0 a_1 \\ b_1 a_0 \\ b_1 a_1 \end{bmatrix}. \tag{12}$$

States such as (12) that can be broken down into a tensor product of single qubit states are called product states. Manipulating one qubit of a product state does not affect the other qubits.

Now consider the following state of a system of two qubits:

$$|\psi\rangle = c_{01}\,|01\rangle + c_{10}\,|10\rangle\,. \tag{13}$$

There is no way to write $|\psi\rangle$ as a product of single-qubit states $|\psi_1\rangle$ and $|\psi_2\rangle$, so $|\psi\rangle$ is not a product state. It an example of an entangled state. If qubit 0 were measured to be up, then with absolute certainty the other qubit would be down. If qubit 0 were measured to be down, then qubit 1 would have to be up. This is true regardless of the physical distance between the qubits! This is a manifestation of quantum nonlocality, or as Einstein once put it in a letter to Max Born, *spukhafte Fernwirkungen*—"spooky action at a distance" [15].

2.2 Measurement and The Born Rule

Having dismissed two parameters describing a qubit as nonphysical in §2.1, it is may not be entirely clear what *is* physical about qubits. A partial explanation of this has been given under (3), but that point will be expanded upon further here.

When a qubit is "measured" along an axis it is forced to take on a definite value on that axis. For example, measuring an arbitrary qubit $|\psi\rangle = c_0\,|0\rangle + c_1\,|1\rangle$ along the z-axis will result in either $|0\rangle$ or $|1\rangle$, forcing the qubit to take on the resulting state. The probability $\mathcal{P}_0$ of taking on $|0\rangle$ is $|c_0|^2$, and the probability $\mathcal{P}_1$ of

taking on $|1\rangle$ is $|c_1|^2$. In general, if $|\uparrow\rangle$ and $|\downarrow\rangle$ are basis states,

$$\mathcal{P}_\uparrow = |\langle\uparrow|\psi\rangle|^2 \tag{14}$$

$$\mathcal{P}_\downarrow = |\langle\downarrow|\psi\rangle|^2. \tag{15}$$

This principle is known Born's Rule. When measuring a state consisting of multiple qubits, an orthogonal basis must be chosen to measure along, and each basis state has a probability associated with it given by a formula akin to (14) and (15).

Now suppose that only a subset of all qubits in an entangled superposition $|\psi\rangle$ are measured. Let m be the number of qubits measured out of n total qubits. If $\left|e^i\right\rangle$ is the ith basis vector of some basis of an m-qubit space, then $|\psi\rangle$ can be expressed as

$$|\psi\rangle = \sum_{i=0}^{2^m-1} c_i \left|e^i\right\rangle \otimes \left|\varphi^i\right\rangle \tag{16}$$

where $\left|\varphi^i\right\rangle$ describes the state of the remaining $n-m$ qubits. In this form, it is evident that the probability of finding $\left|e^i\right\rangle$ is $|c_i|^2$. Another outcome of such a measurement is that the non-measured qubits collapse to the state $\left|\varphi^i\right\rangle$. This rule has been called the generalized Born rule [14], and it is a powerful technique that will be used in §3.2 and §4.

2.3 Special Quantum States

2.3.1 Bell States

The Bell states are maximally entangled 2-qubit states. They are mutually orthogonal, and form a basis for two qubit-systems.

$$\left|\Phi^+\right\rangle := \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle \tag{17}$$

$$\left|\Phi^-\right\rangle := \frac{1}{\sqrt{2}}|00\rangle - \frac{1}{\sqrt{2}}|11\rangle \tag{18}$$

$$\left|\Psi^+\right\rangle := \frac{1}{\sqrt{2}}|01\rangle + \frac{1}{\sqrt{2}}|10\rangle \tag{19}$$

$$\left|\Psi^-\right\rangle := \frac{1}{\sqrt{2}}|01\rangle - \frac{1}{\sqrt{2}}|10\rangle \tag{20}$$

2.3.2 GHZ States

Greenberger-Horne-Zeilinger states [16], or GHZ states, are an infinite class of states. It is a long name for a simple concept: the qubits are either all up or all down. A GHZ state on n qubits is defined as

$$|\mathrm{GHZ}_n\rangle := \frac{1}{\sqrt{2}}|0_n\rangle + \frac{1}{\sqrt{2}}|(2^n-1)_n\rangle \tag{21}$$

$$= \frac{1}{\sqrt{2}}|00\ldots0\rangle + \frac{1}{\sqrt{2}}|11\ldots1\rangle\,. \tag{22}$$

For example, $|\Phi^+\rangle$ is GHZ states of two qubits, and $|\Phi^-\rangle$ is a GHZ state up to a difference in phase.

2.3.3 W-states

Like GHZ states, W-states are a class of multi-qubit states. In a W-state, precisely one of the qubits is on, and the rest are off. A W-state for an n-qubit system is described by

$$|\mathrm{W}_n\rangle := \sum_{i=0}^{n-1} \frac{1}{\sqrt{n}}\left|(2^i)_n\right\rangle \tag{23}$$

$$= \frac{1}{\sqrt{n}}|00\ldots01\rangle + \frac{1}{\sqrt{n}}|0\ldots010\rangle + \ldots + \frac{1}{\sqrt{n}}|010\ldots0\rangle + \frac{1}{\sqrt{n}}|10\ldots00\rangle\,. \tag{24}$$

The Bell state $|\Psi^+\rangle$ is a W-state on two qubits, and $|\Psi^-\rangle$ is a W-state up to phase.

2.4 Quantum Gates

A quantum gate changes the state of a qubit. I will refer to qubits as "passing through" gates, although this is not necessarily what physically happens in a quantum computer. It is also mathematically appropriate to speak of a gate "acting" on one or more qubits, or "mapping" one state to another, since gates are merely linear transformations.

All quantum gates are unitary; that is, if $\mathtt{U}$ is a gate represented by a matrix U, then $U^\dagger U = I$ (where $U^\dagger$ is the conjugate-transpose[2] of U). This property is necessary to ensure that states remain normalized after passing through gates.

In this report, I use typewriter font to denote gates. However, for all intents and purposes, a quantum gate is just a unitary matrix and I will treat them as such. In particular, I will write "$\mathtt{U}^\dagger$" without hesitation from here on.

[2]Also known as the Hermitian conjugate, or adjoint.

2.4.1 Single Qubit Gates

The most conceptually simple gate is the `NOT` gate. When a qubit passes through a `NOT` gate, its $|0\rangle$ component flips to $|1\rangle$ and vice versa. That is, $\texttt{NOT}(c_0 |0\rangle + c_1 |1\rangle) = c_0 |1\rangle + c_1 |0\rangle$. Because the `NOT` gate corresponds to a 180° rotation around the x-axis on the Bloch sphere, it is usually referred to as the `X` gate. Its effect on qubits is repeated below:

$$\texttt{X}(c_0 |0\rangle + c_1 |1\rangle) = c_0 |1\rangle + c_1 |0\rangle . \tag{25}$$

Unlike `X`, the next two gates have no natural analog in classical computing. They correspond to 180° rotations around the y- and z-axes. Respectively, they are `Y` and `Z`. They produce the following effects:

$$\texttt{Y}(c_0 |0\rangle + c_1 |1\rangle) = i\, c_0 |1\rangle - i\, c_1 |0\rangle , \tag{26}$$

$$\texttt{Z}(c_0 |0\rangle + c_1 |1\rangle) = c_0 |0\rangle - c_1 |1\rangle . \tag{27}$$

Note that these two gates affect the coefficients as well—in fact, that is all `Z` does. None of the gates introduced thus far have been able to take $|0\rangle$ into a superposition of $|0\rangle$ and $|1\rangle$. The Hadamard gate, `H`, defined by

$$\texttt{H}(c_0 |0\rangle + c_1 |1\rangle) = \frac{c_0 + c_1}{\sqrt{2}} |0\rangle + \frac{c_0 - c_1}{\sqrt{2}} |1\rangle , \tag{28}$$

is unique among the common gates in this respect.

For many purposes, `H` is the only gate needed to "split" states into superpositions. I will later introduce a generalized class of "splitter" gates that are useful for constructing W-states. Table 4 summarizes commonly used gates, including the identity gate `I` which does nothing to qubits that pass through it.

Gate	Matrix Representation	Maps $\|0\rangle$ to	Maps $\|1\rangle$ to
`I`	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\|0\rangle$	$\|1\rangle$
`X`	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$	$\|1\rangle$	$\|0\rangle$
`Y`	$\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$	$i\|1\rangle$	$-i\|0\rangle$
`Z`	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	$\|0\rangle$	$-\|1\rangle$
`H`	$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$	$\frac{\|0\rangle+\|1\rangle}{\sqrt{2}}$	$\frac{\|0\rangle-\|1\rangle}{\sqrt{2}}$

Table 4: Summary of common single-qubit gates.

More gates remain to be discussed. Two simple ones are the phase gate $\mathtt{S}$ and the T gate $\mathtt{T}$,

$$\mathtt{S} := \begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{bmatrix} \tag{29}$$

$$\mathtt{T} := \begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/8} \end{bmatrix} \tag{30}$$

where clearly $\mathtt{T}^4 = \mathtt{S}^2 = \mathtt{Z}$.

Perhaps the most important gates—or really, infinite family of gates—don't often appear explicitly in circuits, but are used in the construction of so many other gates as discussed in §2.5.2. The gates correspond to general rotations about the x-, y-, and z-axes on the Bloch sphere.

$$\mathtt{R}_x(\theta) := \exp\left(-i\,\mathtt{X}\,\frac{\theta}{2}\right) = \mathtt{I}\cos\frac{\theta}{2} - i\,\mathtt{X}\sin\frac{\theta}{2} = \begin{bmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix} \tag{31}$$

$$\mathtt{R}_y(\theta) := \exp\left(-i\,\mathtt{Y}\,\frac{\theta}{2}\right) = \mathtt{I}\cos\frac{\theta}{2} - i\,\mathtt{Y}\sin\frac{\theta}{2} = \begin{bmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix} \tag{32}$$

$$\mathtt{R}_z(\theta) := \exp\left(-i\,\mathtt{Z}\,\frac{\theta}{2}\right) = \mathtt{I}\cos\frac{\theta}{2} - i\,\mathtt{Z}\sin\frac{\theta}{2} = \begin{bmatrix} \exp\left(-i\,\frac{\theta}{2}\right) & 0 \\ 0 & \exp\left(i\,\frac{\theta}{2}\right) \end{bmatrix} \tag{33}$$

Since any gate corresponds to a unique rotation on the Bloch sphere, any gate can be uniquely decomposed into three rotation gates according to Euler angles. Rotation gates can be used to prove the following theorem [17], which has relevance later in §2.5.2.

Theorem 1. *For any gate $\mathtt{G}$ there exist gates $\mathtt{A}$, $\mathtt{B}$, and $\mathtt{C}$ such that $\mathtt{G} = \mathtt{AXBXC}$ and $\mathtt{ABC} = \mathtt{I}$. In particular, if $\mathtt{G} = \mathtt{R}_z(\alpha)\mathtt{R}_y(\theta)\mathtt{R}_z(\beta)$, then $\mathtt{A} = \mathtt{R}_z(\alpha)\mathtt{R}_z(\theta/2)$, $B = \mathtt{R}_y(\theta/2)\mathtt{R}_z(-(\alpha+\beta)/2)$, and $C = \mathtt{R}_z((\beta-\alpha)/2)$.*

Finally, I conclude with what I refer to as the "overall-phase gate," $\Phi(\phi)$. It is a family of gates similar to the rotation gates. Like the identity gate, it is is useful in discussing mathematical relations between gates.

$$\Phi(\phi) := \begin{bmatrix} e^{i\phi} & 0 \\ 0 & e^{i\phi} \end{bmatrix} \tag{34}$$

2.4.2 Multi-Qubit Gates

A gate for a single qubit affects only that qubit and no others. When there are $n > 1$ qubits, it is convenient to specify which qubit a gate acts on with a subscript. For a gate (or any linear transformation) $\mathtt{G}$, let

$$\mathtt{G}_i := \mathtt{I}^{\otimes\, n-i-1} \otimes \mathtt{G} \otimes \mathtt{I}^{\otimes\, i}. \tag{35}$$

Observe that (35) describes a multi-qubit gate that acts with $\mathtt{G}$ on qubit i ($0 \leq i < n$) and leaves all other qubits untouched. The notation introduced above is invaluable for describing single-qubit gates in multi-qubit systems, but single-qubit gates offer no means of entangling (or disentangling) a set of qubits. To this end, it is necessary to use multi-qubit gates.

There are surprisingly few standard gates that act on multi-qubit systems. One reason for this is simple: it is difficult to physically implement such gates. The most important gate, perhaps out of all quantum gates, is the controlled-NOT gate, abbreviated as CNOT or CX. It is analogous to a conditional statement in classical computing. CX acts on two qubits, referred to as the "control" and the "target". The target is flipped (as if passed through X) if and only if the control is in the state $|1\rangle$, or "on." This manifests as the following formula:

$$\mathtt{CX}(c_{00}|00\rangle + c_{01}|01\rangle + c_{10}|10\rangle + c_{11}|11\rangle) = (c_{00}|00\rangle + c_{01}|01\rangle + c_{10}|11\rangle + c_{11}|10\rangle) \tag{36}$$

where qubit 0 is the target and qubit 1 is the control.

To be clearer about which qubit is the control and which is the target, I will apply subscripts to the symbol. $\mathtt{C}_i\mathtt{X}_j$ will mean that qubit i is the control and qubit j is the target. For example, in (36), CX could be replaced with $\mathtt{C}_1\mathtt{X}_0$ for enhanced clarity.

There are many ways of generalizing the concept of controlled gates.

More Control Qubits To specify more control qubits, I will increase the number of Cs. For example, $\mathtt{CCX} = \mathtt{C}_2\mathtt{C}_1\mathtt{X}_0 = \mathtt{C}_1\mathtt{C}_2\mathtt{X}_0$ applies X to qubit 0 if and only if qubits 2 and 3 are both in the state $|1\rangle$. The CCX gate is also known as the Toffoli gate.

Different Controlled Gates It is also useful to have other gates controlled by a qubit. Most common is CZ, which applies the Z gate to qubit 0 if and only if qubit 1 is in the state $|1\rangle$. Extrapolating this, define $\mathtt{C}_i\mathtt{G}_j$ to apply gate G to qubit j if and only if qubit i is in the state $|1\rangle$. Formally, in a system with n qubits and for $i \neq j$,

$$\mathtt{C}_i\mathtt{G}_j := |0\rangle\langle 0|_i + |1\rangle\langle 1|_i \, \mathtt{G}_j \tag{37}$$

using the notation introduced in (35).

In the special case where a system only has two qubits, a $\mathtt{C}_1\mathtt{G}_0$ gate has the following form:

$$\mathtt{C}_1\mathtt{G}_0 = \begin{bmatrix} \mathtt{I} & \mathtt{0} \\ \mathtt{0} & \mathtt{G} \end{bmatrix}, \tag{38}$$

where 0 is the zero matrix.

More Target Qubits If a collection of qubits i, j, k are controls for multiple controlled gates $\mathtt{A}_a, \mathtt{B}_b, \mathtt{D}_d$, it is convenient to write

$$\mathtt{C}_i\mathtt{C}_j\mathtt{C}_k\mathtt{A}_a\mathtt{B}_b\mathtt{D}_d := (\mathtt{I}^{\otimes n} - |111\rangle\langle 111|_{ijk}) + |111\rangle\langle 111|_{ijk}\,\mathtt{A}_a\mathtt{B}_b\mathtt{D}_d \tag{39}$$

$$= \mathtt{I}^{\otimes n} + (\mathtt{A}_a\mathtt{B}_b\mathtt{D}_d - \mathtt{I}^{\otimes n})\,|111\rangle\langle 111|_{ijk}\,. \tag{40}$$

Also note that $\mathtt{C}_i\mathtt{C}_j\mathtt{C}_k\mathtt{A}_a\mathtt{B}_b\mathtt{D}_d = (\mathtt{C}_i\mathtt{C}_j\mathtt{C}_k\mathtt{A}_a)(\mathtt{C}_i\mathtt{C}_j\mathtt{C}_k\mathtt{B}_b)(\mathtt{C}_i\mathtt{C}_j\mathtt{C}_k\mathtt{D}_d)$, and all the factors commute.

Even more generally, if C is a set of control qubits and T is a set of target qubits with $\mathtt{G}(t)$ being the gate applied to qubit $t \in T$,

$$\mathtt{C}_c\mathtt{G}(t)_T := \left(\mathtt{I}^{\otimes n} - \prod_{c\in C}|1\rangle\langle 1|_c\right) + \left(\prod_{c\in C}|1\rangle\langle 1|_c\right)\prod_{t\in T}\mathtt{G}(t)_t \tag{41}$$

$$= \mathtt{I}^{\otimes n} + \left(\prod_{t\in T}\mathtt{G}(t)_t - \mathtt{I}^{\otimes n}\right)\prod_{c\in C}|1\rangle\langle 1|_c \tag{42}$$

where as usual n is the total number of qubits.

The forms of (41) and (42) should be intuitive. The first term of (41) maps a state that does not contain the all-1s state of the qubits in C to itself; the second term acts with $\mathtt{G}(t)_t$ for all $t \in T$ on any other state (i.e., a state in which all qubits $c \in C$ are $|1\rangle$). (42) is easily derived from (41) using the bilinearity of the tensor product. Its intuitive meaning is different, but equivalent: the first term maps every state to itself, and the second term acts on any state in which the control qubits are all $|1\rangle$ by acting on the target qubits and canceling out the output of the first term.

Different Activation Requirements If the previous generalization is not general enough, it is possible to go one step further. We can specify exactly what state each control qubit must be in for the controlled gates to activate. Extending (42) and letting $|\psi(c)\rangle$ be the state that qubit c must be in for activation,

$$\mathtt{C}_C^{|\psi(c)\rangle}\mathtt{G}(t)_T := \mathtt{I}^{\otimes n} + \left(\prod_{t\in T}\mathtt{G}(t)_t - \mathtt{I}^{\otimes n}\right)\prod_{c\in C}|\psi(c)\rangle\langle\psi(c)|_c\,. \tag{43}$$

For convenience, let

$$\overline{\mathtt{C}}_C\mathtt{G}(t)_T := \mathtt{C}_C^{|0\rangle}\mathtt{G}(t)_T. \tag{44}$$

I will often adopt the notation of (44) in simpler schemes, such as $\overline{\mathtt{CX}}$, which negates the target qubit if and only if the control qubit is off.

Swap Gate There is one other ubiquitous multi-qubit gate: the SWAP gate. It and its variants play important roles, although perhaps not as fundamentally as CNOT. The swap gate does exactly as it claims to do: it swaps the states of qubits 0 and 1. When applying a swap gate to a system with more than two qubits, it is necessary to use subscripts to specify which qubits to swap. Its most general form is thus $\mathtt{SWAP}_{ij}$.

Of course, $\texttt{SWAP}_{ij} = \texttt{SWAP}_{ji}$.

2.4.3 Properties of Gates

A reader familiar with quantum mechanics may have already noticed that `X`, `Y`, and `Z` are the Pauli matrices, and thus have all the same properties. Most importantly, they are their own inverses:

$$\texttt{XX} = \texttt{YY} = \texttt{ZZ} = \texttt{I}. \tag{45}$$

Some readers may also have noticed that the Hadamard gate is a Hadamard matrix scaled to have determinant -1, so it follows that

$$\texttt{HH} = \texttt{I}. \tag{46}$$

Obviously, the same holds for the controlled versions of those gates:

$$(\texttt{CX})^2 = (\texttt{CY})^2 = (\texttt{CZ})^2 = (\texttt{CH})^2 = \texttt{I}^{\otimes 2}. \tag{47}$$

The controlled-`Z` gate has a particularly interesting property. In matrix form, along the lines of (38), we have

$$\texttt{C}_1\texttt{Z}_0 = \begin{bmatrix} \texttt{I} & \texttt{0} \\ \texttt{0} & \texttt{Z} \end{bmatrix} \tag{48}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \tag{49}$$

$$= \texttt{C}_0\texttt{Z}_1. \tag{50}$$

Table 5 also demonstrates this equality, since a gate is entirely determined by its action on a basis.

$\lvert\psi\rangle$	$\texttt{C}_1\texttt{Z}_0 \lvert\psi\rangle$	$\texttt{C}_0\texttt{Z}_1 \lvert\psi\rangle$
$\lvert 00\rangle$	$\lvert 00\rangle$	$\lvert 00\rangle$
$\lvert 01\rangle$	$\lvert 01\rangle$	$\lvert 01\rangle$
$\lvert 10\rangle$	$\lvert 10\rangle$	$\lvert 10\rangle$
$\lvert 11\rangle$	$-\lvert 11\rangle$	$-\lvert 11\rangle$

Table 5: Input and output table of controlled-`Z` gate with different configurations of control and target.

It is also important to know how these common gates interact with each other. Since there is no way of "adding" two gates together in a quantum circuit,[3] only multiplicative relations are strictly necessary to

[3] In some sense it actually is possible to add gates together using controlled gates where the control qubit is in a superposition of $|0\rangle$ and $|1\rangle$. Chiribella et al. take this a step further and describe superpositions of entire circuits [18].

understand. However, additive relations still offer insights into the multiplicative relations, so we will briefly explore some.

Additive Relations By inspection, it is clear from the matrix representations of $\mathtt{X}$, $\mathtt{Z}$, and $\mathtt{H}$ that

$$\mathtt{H} = \frac{\mathtt{X} + \mathtt{Z}}{\sqrt{2}}. \tag{51}$$

It is in some sense fair to say that the Hadamard gate is a superposition of the $\mathtt{X}$ and $\mathtt{Z}$ gates. One could just as easily define superpositions of the other two pairs of Pauli matrices:

$$\mathtt{G}_{XY} := \frac{\mathtt{X} + \mathtt{Y}}{\sqrt{2}} \tag{52}$$

$$\mathtt{G}_{YZ} := \frac{\mathtt{Y} + \mathtt{Z}}{\sqrt{2}}. \tag{53}$$

To my knowledge, these last two gates do not appear in the literature. Nor need they—the Hadamard gate alone covers most of their use cases. I will mention them later in this section and then promptly discard them.

Multiplicative Relations The first thing to note is that, like the quaternions, the Pauli matrices are anticommutative (e.g. $\mathtt{XY} = -\mathtt{YX} = \Phi(\pi)\mathtt{YX}$). It immediately follows that conjugating one Pauli matrix by another yields the negation of the former. However, unlike the quaternions, we have $\mathtt{XYZ} = \Phi(\pi/2)$.

The anticommutativity of the Pauli matrices and the form of the Hadamard gate given in (51) can be used to prove an important pair of relations between $\mathtt{H}$, $\mathtt{X}$, and $\mathtt{Z}$:

$$\mathtt{HXH} = \mathtt{Z} \tag{54}$$

$$\mathtt{HZH} = \mathtt{X} \tag{55}$$

In words, conjugating $\mathtt{X}$ by $\mathtt{H}$ yields $\mathtt{Z}$, and vice versa. Note that (55) also follows immediately from (54) and (46). The same sort of relations also hold between $\mathtt{G}_{XY}$, $\mathtt{X}$, and $\mathtt{Y}$, and between $\mathtt{G}_{YZ}$, $\mathtt{Y}$, and $\mathtt{Z}$.[4] Relations (54) and (55) will soon be used to show a remarkable result that distinguishes quantum computing from classical computing.

One more observation is needed. Let $\mathtt{G}$ be an involutory gate (that is, $\mathtt{G}^2 = \mathtt{I}$), and consider the actions of $\mathtt{G}_0(\mathtt{C}_1\mathtt{X}_0)\mathtt{G}_0$. If the control qubit is off, the target qubit is acted on by $\mathtt{G}$ twice in a row, thereby canceling itself out. If the control qubit is on, the target qubit is acted on by $\mathtt{G}$, then $\mathtt{X}$, then $\mathtt{G}$ again. It follows that $\mathtt{G}_0(\mathtt{C}_1\mathtt{X}_0)\mathtt{G}_0$ is actually a controlled-$\mathtt{GXG}$ gate:

$$\mathtt{G}^2 = \mathtt{I} \implies \mathtt{G}_0(\mathtt{C}_1\mathtt{X}_0)\mathtt{G}_0 = \mathtt{C}_1(\mathtt{GXG})_0. \tag{56}$$

The above claims lead to the identity given in (62). This identity demonstrates that the roles of control and target qubit in a $\mathtt{CX}$ gate can be switched by applying single-qubit gates before and after. It goes without

[4] I shall now discard $\mathtt{G}_{XY}$ and $\mathtt{G}_{YZ}$.

saying that classical computers are incapable of this.

$$\begin{aligned}
\mathtt{H}^{\otimes 2}(\mathtt{C}_1\mathtt{X}_0)\mathtt{H}^{\otimes 2} &= \mathtt{H}_1(\mathtt{C}_1(\mathtt{HXH})_0)\mathtt{H}_1 && (57)\\
&= \mathtt{H}_1(\mathtt{C}_1\mathtt{Z}_0)\mathtt{H}_1 \qquad \text{via (54)} && (58)\\
&= \mathtt{H}_1(\mathtt{C}_0\mathtt{Z}_1)\mathtt{H}_1 \qquad \text{via (50)} && (59)\\
&= \mathtt{H}_1(\mathtt{C}_0(\mathtt{HXH})_1)\mathtt{H}_1 \qquad \text{via (54)} && (60)\\
&= (\mathtt{H}_1)^2(\mathtt{C}_0\mathtt{X}_1)(\mathtt{H}_1)^2 && (61)\\
&= \mathtt{C}_0\mathtt{X}_1 \qquad \text{via (46)} && (62)
\end{aligned}$$

Rational Radicals Sometimes it is necessary to find a gate that, when applied many times to a single qubit, is equivalent to the application of some other gate. For example, one implementation of the Toffoli gate uses controlled-$\sqrt{\mathtt{X}}$ gates. The usual way of finding a matrix raised to a rational power is by diagonalizing it, i.e. $G^k = UD^kU^{-1}$. For involutary gates, the following simpler formula holds (where n is the number of qubits):

$$\sqrt{\mathtt{G}} = \frac{\sqrt{i}\,\mathtt{I}^{\otimes n} + \sqrt{-i}\,\mathtt{G}}{\sqrt{2}} \tag{63}$$

2.4.4 Universality of Gates and Arbitrary Approximation

A significant difference between quantum computing and classical computing is that any quantum circuit can be realized using only two-qubit gates [19], whereas a classical circuit may require one or more three-bit gates. In this sense, two-qubit gates are "universal" for quantum computation. In fact, it has been shown that CX and single-qubit gates are universal [17].

Perhaps even more impressively, if a pair of gates is chosen in a particular way, *any* single-qubit gate can be simulated to arbitrary accuracy via finitely many applications of those two gates. The Bloch sphere aids in reasoning here. If two gates correspond to rotations of θ_1 and θ_2 around different axes, and θ_1/π and θ_2/π are irrational, then any rotation (ergo any gate) can be approximated to arbitrary accuracy [20]. This fundamental result is known as the Solovay-Kitaev theorem.

The universality of CX and single-qubit gates and the Solovay-Kitaev theorem together imply that only three well-picked gates can approximate any other gate to arbitrary accuracy. Therefore, if a quantum computer can execute three specific gates to high accuracy, it can execute any quantum algorithm.

2.5 Quantum Circuits

2.5.1 How to Read a Circuit Diagram

A circuit diagram is a graphical representation of a quantum algorithm. A diagram consists of horizontal lines representing qubits, and blocks with symbols on those lines representing gates. Time passes from left to right. The qubits are ordered bottom to top, analogous to how the qubits in a ket are ordered right to left.

Figure 1 shows a circuit that constructs $|\Psi^+\rangle$ when its "input register" is $|00\rangle$. The circuit is equivalent to $\mathtt{X}_1(\mathtt{C}_1\mathtt{X}_0)\mathtt{H}_1|00\rangle$. Note that because gates are prepended to quantum states mathematically, the order in which they appear is flipped compared to the circuit. The vertical line joining a black dot and an $\mathtt{X}$ gate denotes $\mathtt{C}_1\mathtt{X}_0$. Black dots are placed on control qubits. If the dot were white instead as in Figure 2, the gate would represent $\overline{\mathtt{C}}_1\mathtt{X}_0$.

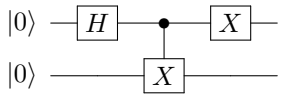

Figure 1: Construction of $|\Psi^+\rangle$, the simplest W-state.

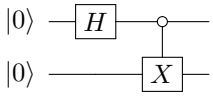

Figure 2: Construction of $|\Psi^+\rangle$using $\overline{\mathtt{C}}_1\mathtt{X}_0$.

The input register is almost always initialized to zeros. Sometimes it is not specified. For example, Figure 3 is a circuit identity equivalent to (62).

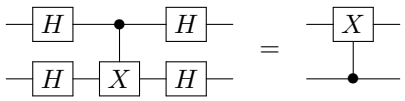

Figure 3: Circuit identity for (62).

Quantum algorithms might also involve measurement. The meter symbol in Figure 4 indicates measurement of a qubit along the z-axis. Since such a measurement always results in a definite 0 or 1, the information from a measurement is encoded in a bit. The double line in Figure 4 indicates the transport of classical information which, in this case, serves to activate the $\mathtt{X}$ gate below only if the result of the measurement is a 1.

It should be noted that the output of Figure 4 is *not* an entangled state. The bottom qubit is not even in a superposition of $|0\rangle$ and $|1\rangle$. It is instead called a "mixed state," which are in general described by probability distributions. Mixed states fill the interior of the Bloch sphere.

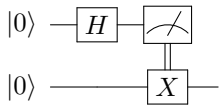

Figure 4: Circuit that outputs $|0\rangle$ in the lower qubit with probability $\frac{1}{2}$ and $|1\rangle$ otherwise.

There are a few other notations used in quantum circuits, such as depictions of general multiqubit gates, swap gates, more general measurements, and multi-qubit wires. The notation covered in this section is sufficient for the remainder of this report, so rather than introduce yet more notation, I will move on to applying these ideas to the problems of interest.

2.5.2 Constructions of Arbitrary Controlled Gates

Some constructions of W-states will make use of multiply-controlled gates. In practice, these gates can be difficult to physically implement because they typically consist of many other gates. Since qubits are so delicate, increasing the number of gates generally increases the probability of unwanted collapse, so it is of utmost importance to minimize the number of gates in any given algorithm.

In this report, the controlled-NOT gate will be taken as the only fundamental multi-qubit gate. All single-qubit gates will also be considered fundamental. I will refer to `CX` and single-qubit gate `U` as "atomic."

The construction of `CU` for a general single-qubit gate `U` takes inspiration from the Bloch sphere and Euler angles. By Theorem 1, `U` can be expressed as `AXBXC` for some gates `A`, `B` and `C` with `ABC` = `I`. The circuit identity for `CU` in Figure 5 immediately follows. Under this construction, a general `CU` gate consists of three single qubit operations and two controlled-NOTs, for a total of five atomic gates.

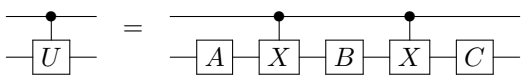

Figure 5: Construction of arbitrary `CU` gate.

A multiply-controlled NOT gate can also be broken down into several `CX` gates. Consider Figure 6, which shows a more general gate construction. One can find `V` such that $\texttt{V}^2 = X$ using (63): $\texttt{V} = \sqrt{\texttt{X}} = \frac{1}{\sqrt{2}}\begin{bmatrix} \sqrt{i} & \sqrt{-i} \\ \sqrt{-i} & \sqrt{i} \end{bmatrix}$. Figure 6 shows that the worst-case scenario for a doubly-controlled gate requires five controlled gates, two of which are `CX`. Combined with the result from Figure 5, a general `CCU` gate needs up to eight `CX` gates and nine single-qubit gates, totalling 17 atomic gates.

Despite the ostensibly rapid increase in gate cost from 5 to 17 by adding another control qubit, the number of atomic gates needed to construct a controlled gate can be made linear with the number of control qubits. A proof of Theorem 2 is given by Barenco et al. [17] (see Corollary 7.10).

Theorem 2. *For n qubits, a general $(n-2)$-controlled gate requires on the order of n atomic gates.*

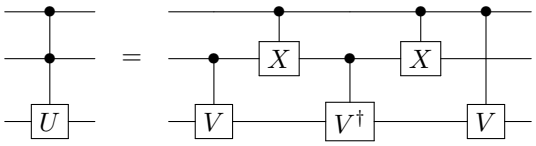

Figure 6: Construction of arbitrary `CCU` gate, where $\texttt{V}^2 = \texttt{U}$.

When precision is desirable, I will use $\mathcal{C}_{\texttt{U}}(k)$ to denote the number of atomic gates needed to simulate a k-controlled `U` gate.

The above constructions have all been of controlled gates that activate when each control qubit is 1. While arbitrary control requirements are possible as discussed in §2.4.2, `CU` and $\overline{\texttt{CU}}$ gates are the most common. Normal controlled gates can be easily converted to reverse-controlled gates by conjugating the control qubit with NOTs. Since two consecutive NOT gates cancel each other out, a single qubit acting as a control-on-0 multiple times in a row only needs to pass through a total of two NOT gates. The top qubit of Figure 7 acts twice in succession as a control-on-0, and thus the inner `X` gates cancel.

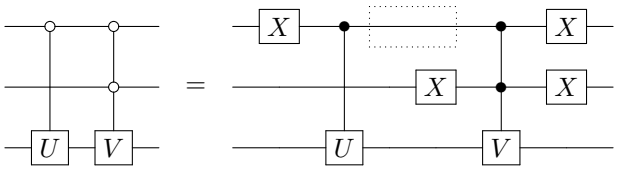

Figure 7: Construction of reverse-controlled gates. The dotted box indicates where two `X` gates canceled each other out.

3 Constructions of W-states

Achieving a superposition of n one-hot states on n qubits is actually very simple using controlled Hadamard gates. The difficulty is in controlling the probabilities to make each state equally likely.

3.1 On 2^k Qubits

When the number of qubits is a power of two, there are two simple tricks that can be used to make a W-state. The key behind both tricks is to realize that the Hadamard gate, when applied to a qubit in a Z-basis state, doubles the number of terms in the superposition describing the system.

For instance, take the initial state $|00\rangle$. This state has a single term. Now apply $\mathtt{H}_0$ to transform the state into $\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|01\rangle$. The state of the system now has two terms, both equally likely to be the result of a measurement along z. Applying $\mathtt{H}_0$ again would collapse the system back to one term, but applying $\mathtt{H}_1$ would double the number of terms up to 4. No other gate introduced in §2.4.1 is able to "split" a state into more states in this way.

Let an input register of $n = 2^k$ qubits all be initialized to 0. Apply $\mathtt{H}^{\otimes k}$ to the first k qubits. After k Hadamards there are 2^k states in the overall superposition. Furthermore, each state is equally likely. One can simply map each of the 2^k states to a unique one-hot state. Figure 8 shows the construction of $|\mathrm{W}_4\rangle$ using this method, with emphasis on subcircuits devoted to mapping individual states. Figure 9 shows the construction of $|\mathrm{W}_8\rangle$ using the same method, but organized differently.

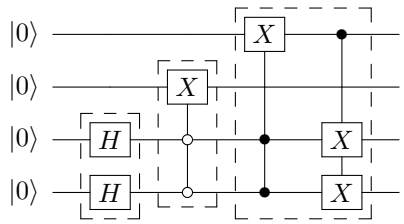

Figure 8: Construction of $|\mathrm{W}_4\rangle$ using naive method. Left to right, the enclosed subcircuits split $|0000\rangle$ into 4 equally likely states; map $|0000\rangle$ to $|0100\rangle$; and map $|0011\rangle$ to $|1000\rangle$.

This method has two shortcomings, earning it the title "naive construction." First of all, it does not nicely generalize to $n \neq 2^k$ qubits, although §3.2 discusses a means of getting around this limitation. Second, it uses more gates than theoretically needed; §3.3 explores an alternative. The following theorem states the asymptotic gate cost of the naive construction. I use $\mathcal{C}_{\mathtt{U}}(n)$ to denote precisely the minimum number of atomic gates needed to simulate an n-controlled $\mathtt{U}$ gate.

Theorem 3. *The naive construction of $|W_n\rangle$ where $n = 2^k$ requires on the order of $\mathcal{C}_X(k)n \in \Theta(n \log n)$ atomic gates.*

Proof. The number of Hadamards goes as $k = \log_2 n$, which is negligible. We may suppose that the the

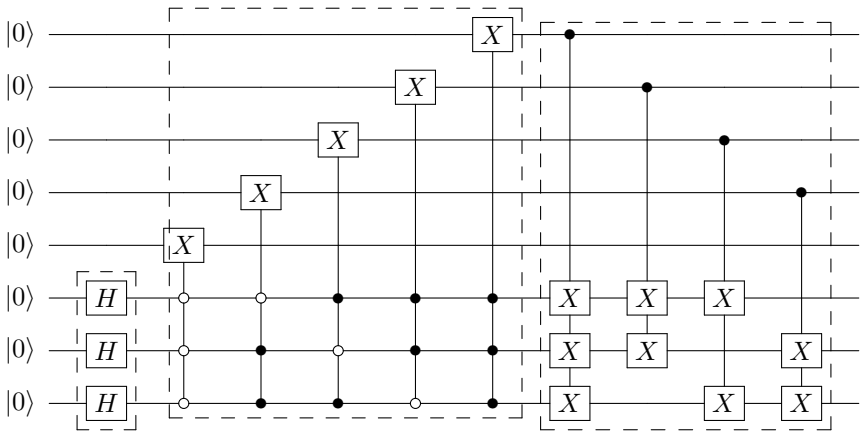

Figure 9: Construction of $|\mathrm{W}_8\rangle$ using naive method. First block: splitting $|0_8\rangle$ into 8-state superposition. Second block: placing 1s in desired locations. Third block: removing unwanted 1s.

system is in the state

$$\frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} |i_n\rangle \,. \tag{64}$$

The number of these k states in which precisely l qubits are $|1\rangle$ is $\binom{k}{l}$. So $\binom{k}{1} = k$ states are already one-hot and do not need further processing. The remaining states do need processing, which consists of (a) a handful of `X` gates implicitly applied to the split qubits to make reverse-controlled gates; (b) a k-controlled NOT gate to set the desired qubit to $|1\rangle$; and later (c) l `CX` gates to correct for undesired $|1\rangle$s. Neglecting (a) and letting $\mathcal{C}_{\mathtt{X}}(k)$ denote the number of atomic operations needed to implement a k-controlled NOT gate,

$$C_n \approx \sum_{l=0,\neq 1}^{k} (\mathcal{C}_{\mathtt{X}}(k) + l)\binom{k}{l} \tag{65}$$

$$\approx \sum_{l=0}^{k} \mathcal{C}_{\mathtt{X}}(k)\binom{k}{l} \tag{66}$$

$$= \mathcal{C}_{\mathtt{X}}(k) 2^k \tag{67}$$

$$= \mathcal{C}_{\mathtt{X}}(k) n \tag{68}$$

By Theorem 2, $\mathcal{C}_{\mathtt{X}}(k) \in \Theta(k)$. So $C_n \sim kn = n \log_2(n)$. □

3.2 On n Qubits, Sometimes

If $n \neq 2^k$, the naive method of W-state construction can be extended using auxiliary qubits. The necessary extension consists of constructing the W-state for the lowest power of 2 greater than n and filtering out unwanted states by measuring the auxiliary qubits. Relying on measurement makes the construction

nondeterministic, so I will refer to the method as the probabilistic naive construction (PNC). Figure 10 demonstrates this procedure to produce $|\mathrm{W}_5\rangle$.

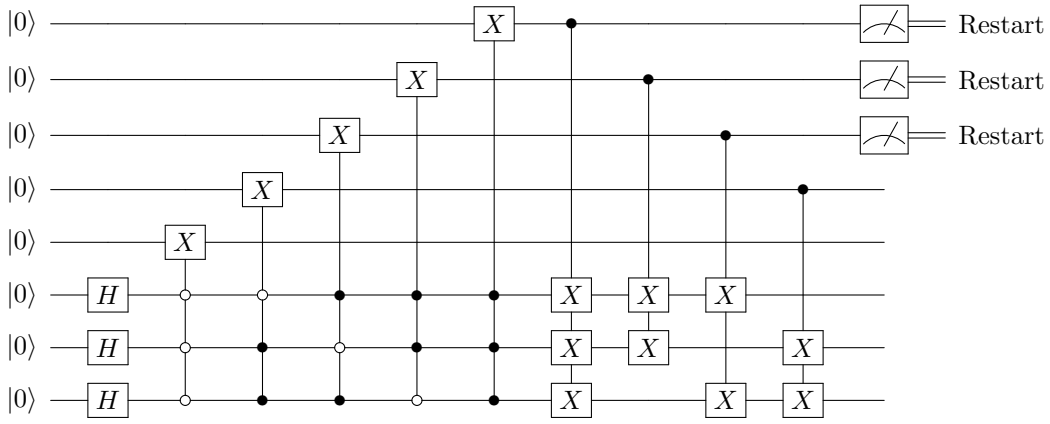

Figure 10: Construction of $|\mathrm{W}_5\rangle$ in bottom five qubits via PNC. If any of the measured qubits return 1, the register is reset to 0s and the circuit is run again.

PNC may seem very inefficient, especially when n is slightly greater than a power of 2. As Theorem 4 shows, however, is that even the worst case scenario has a cost on the order of $n \log n$. Intuitively, this is because the chance of the $|1\rangle$ being found in the auxiliary qubits is always less than $1/2$, regardless of n.

Theorem 4. *The probabilistic naive construction of $|W_n\rangle$ requires on the order of $n \log n$ atomic gates.*

Proof. Let $k \in \mathbb{N}$ such that $2^{k-1} < n < 2^k =: N$, and define $\delta := N - n$. From Theorem 3, the cost to construct $|\mathrm{W}_N\rangle$ is on the order of $N \log N$. After constructing $|\mathrm{W}_N\rangle$ and performing a joint measurement on the δ auxiliary qubits, there is a $\frac{\delta}{N}$ of finding a qubit in the state $|1\rangle$. The expected number of attempts needed to find all 0s is thus $\frac{1}{1-\delta/N}$ so the expected total cost for n qubits is

$$EC(n) \sim (N \log N)\left(\frac{1}{1-\delta/N}\right) \tag{69}$$

$$= \frac{N^2}{n} \log N. \tag{70}$$

Also note that the input register must be set to 0s before each attempt, necessitating more gates. However, since the average number of attempts needed is essentially constant, and wiping the register requires on the order of n gates, the cost of the resets is relatively negligible.

In the worst case, $N \approx 2n$ so (70) becomes

$$EC(n) \sim 4n \log 2n. \tag{71}$$

The best case, where $N = n$, is of course non-probabilistic and is on the order of $n \log n$. So $EC(n) \in \Theta(n \log n)$. □

Several optimizations to PNC are possible. Of the last four gates in Figure 10, all but the very last

one are completely unnecessary. Furthermore, each auxiliary qubit can be measured directly after passing through the triple-controlled NOT gate. The circuit can be immediately reset as soon as a 1 is detected, bypassing the remainder of the protocol. This optimization drastically cuts down on gate cost, but cannot reduce asymptotic cost. Theorem 4 shows that even the unoptimized construction has the same asymptotic cost as the deterministic construction of $|\mathrm{W}_{2^k}\rangle$. This motivates the following corollary to Theorem 4:

Corollary 5. *No optimization to PNC can reduce its asymptotic expected cost to below* $n \log n$.

Proof. No matter how optimized, PNC of $|\mathrm{W}_n\rangle$ can never be more efficient than the naive construction of $|\mathrm{W}_m\rangle$ where m is the greatest power of 2 less than n. By Theorem 3, the cost to construct $|\mathrm{W}_m\rangle$ is on the order of $m \log m$. Since $n/2 < m < n$, the asymptotic cost to construct $|\mathrm{W}_n\rangle$ must be bounded below by $(n/2)\log(n/2)$, or equivalently, $n \log n$. □

3.2.1 Optimized Probabilistic Naive Construction

The probabilistic naive construction can be optimized by checking for 1s in auxiliary qubits as soon as possible, rather than after the full W-state has been constructed. Doing so eliminates the need for auxiliary qubits entirely, since the same qubit can be reused for each check and then again as part of the final W-state. Figure 11 shows the optimized probabilistic naive construction (OPNC) of $|\mathrm{W}_5\rangle$.

The asymptotic gate cost of OPNC is $n \log n$, in compliance with Corollary 5. A more precise approximation of its expected cost is derived in Theorem 9, given in Appendix A.

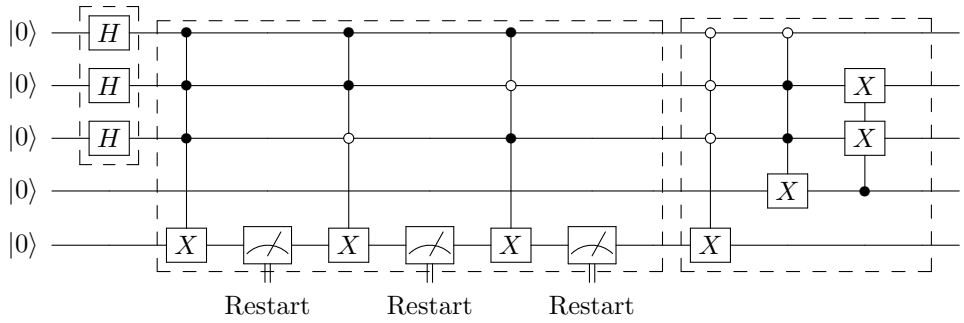

Figure 11: Construction of $|\mathrm{W}_5\rangle$ via OPNC. Block 1: splitting initial state into superposition of 8 states. Block 2: filtering out three extraneous states from superposition. Block 3: mapping remaining states to one-hot states.

3.3 On n Qubits, Always but Inefficiently

As previously claimed, there is no way of achieving even weights in an arbitrary w_n state using only Hadamards and controlled Hadamards. So although it may seem like cheating, one solution is to invent a new gate. In this case, an entire new class of gates is needed. I will refer to them as "splitter gates," denoted $\mathtt{S}_n$ (not to be confused with the phase gate, $\mathtt{S}$).

$$\mathtt{S}_n := \frac{1}{\sqrt{n}} \begin{bmatrix} \sqrt{n-1} & 1 \\ 1 & -\sqrt{n-1} \end{bmatrix} \tag{72}$$

Splitter gates are designed to have the property that measuring $\mathtt{S}_n |0\rangle$ yields $|1\rangle$ with probability $\frac{1}{n}$, and otherwise yields $|0\rangle$.[5] An interesting consequence of this is that $\mathtt{X} \equiv \mathtt{S}_1$ and $\mathtt{H} \equiv \mathtt{S}_2$.

Also note that $\mathtt{S}_n^2 = \mathtt{I}$, so (63) applies to splitter gates.

Controlled splitter gates make the construction of arbitrary $|\mathrm{W}_n\rangle$ trivial. Figure 12 demonstrates the idea.

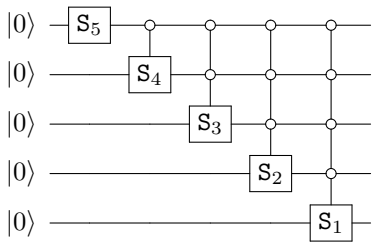

Figure 12: Sub-optimal construction of $|\mathrm{W}_5\rangle$ using splitter gates.

Theorem 6. *The inefficient construction of $|W_n\rangle$ using splitter gates requires on the order of n^2 atomic operations.*

Proof. By inspection of 12, the construction of $|\mathrm{W}_n\rangle$ contains n controlled gates. Supposing $\mathcal{C}_{\mathtt{S}_i}(k) = \mathcal{C}_{\mathtt{S}_j}(k) =: \mathcal{C}_{\mathtt{S}}(i)$, the total cost is simply

$$C(n) = \sum_{i=0}^{n-1} \mathcal{C}_{\mathtt{S}}(i) + 2(n-1) \tag{73}$$

where $\mathcal{C}_{\mathtt{U}}(0) = 1$ for any single-qubit gate $\mathtt{U}$, since $\mathcal{C}_{\mathtt{U}}(0)$ is simply the cost of $\mathtt{U}$. The second term of (73) accounts for the $\mathtt{X}$ gates to convert control-on-1 qubits to control-on-0, as shown in Figure 12. The optimization illustrated by Figure 7 significantly cuts down on the number of $\mathtt{X}$ gates needed.

By Theorem 2, $\mathcal{C}_{\mathtt{S}}(i) \in \Theta(i)$. So from (73),

$$C(n) \sim \sum_{i=0}^{n-1} i + 2(n-1) \tag{74}$$

$$= \frac{n(n-1)}{2} + 2(n-1) \tag{75}$$

$$\sim n^2. \tag{76}$$

[5] There are other classes of gates with similar properties. Craig Gidney and one other user gave similar gates on StackExchange [21]. Gidney called his gate the "odds gate," a generalization of splitter gates that differ in phase. I have seen neither splitter gates nor odds gates in literature on quantum circuits. Gidney took inspiration from the classical problem of reservoir sampling.

□

Comparing Theorem 4 with Theorem 6 implies that PNC is asymptotically better than the inefficient construction using splitter gates. The problem with the latter method is that it uses controlled gates with as many as $n-1$ control qubits, whereas PNC never uses more than $\lceil \log_2 n \rceil$-controlled gates. The next section solves this problem.

3.4 On n Qubits, Always and Efficiently

A simple observation can improve the asymptotic efficiency of splitter-gate constructions [21]. In Figure 12, the splitter for each qubit is explicitly controlled by the state of all previous qubits. It is possible to automatically "check" the state of all previous qubits by only checking the immediate predecessor. I will call this construction the optimized splitter construction (OSC). Figure 13 demonstrates this notion on 5 qubits.

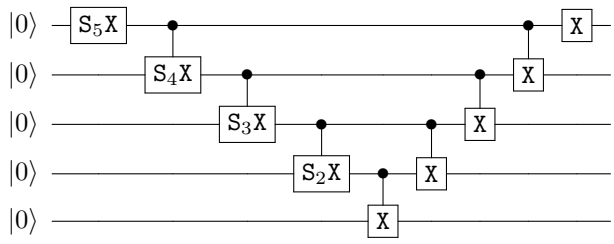

Figure 13: Optimal construction of $|\mathrm{W}_5\rangle$ using splitter gates.

Theorem 7. *The efficient construction of $|W_n\rangle$ using splitter gates requires on the order of n atomic operations.*

Proof. According to Figure 5, $\mathcal{C}_{\mathtt{S}}(1) = 5$. The number of atomic gates needed to construct $|\mathrm{W}_n\rangle$ $(n \geq 2)$ is then precisely

$$C(n) = 1 + (n-2)\mathcal{C}_{\mathtt{S}}(1) + n \tag{77}$$

$$= 1 + 5(n-2) + n \tag{78}$$

$$= 6n - 9 \tag{79}$$

which is linear in n. □

This improved construction beats the $\Theta(n \log n)$ cost of PNC and OPNC. The pros and cons of the three methods are compared in Table 6.

It is impossible to discern which of the properties in Table 6 are most important in general. The importance of each property depends on the quantum computer itself. That said, determinability is probably the least consequential parameter. Unless the underlying machine does not handle measurement well, overall cost is more important than determinibility of cost.

Method	Asymptotic Cost	Deterministic?	Uses Standard Gates?	Auxiliary Qubits
PNC	$\Theta(n \log n)$	No	Yes	δ
OPNC	$\Theta(n \log n)$	No	Yes	0
OSC	$\Theta(n)$	Yes	No	0

Table 6: Comparison of construction methods. "Standard gates" are those gates introduced in §2.4.1 and the `CX` gate.

The asymptotic difference in cost between the methods is a factor of $\log n$, which is not much. So unless n is very large,[6] the biggest factor in which method to use may come down to what basic gates the machine natively has and how many qubits it has access to.

It may be unreasonable to expect a machine to natively support splitter gates. A typical machine would have to approximate the splitters, as discussed in §2.4.4. This introduces several complications, including possibly increasing the asymptotic cost. It also means that any W-state produced by splitter gates would be imperfect, a downside that could be fatal for some applications. On the other hand, qubits are not cheap. In the worst case, PNC requires nearly $2n$ qubits to construct $|\mathrm{W}_n\rangle$. Fortunately, the auxiliary qubits always end in the state $|0\rangle$, and can be reused for other purposes. Regardless, OPNC requires no auxiliary qubits. Ultimately, of the options presented in this section, OPNC may be the best for the majority of the first generation of quantum computers.

[6] As of the time of this writing, the highest number of qubits achieved on a quantum computer is 53 [22]. So n is not very large.

4 Proof of Bell's Theorem using W-states, Without Inequalities

Bell's theorem is a famous result first published by John Bell in 1964 [23]. It was the first proof of the nonexistence of elements of reality and has profound implications for the nature of reality.

4.1 Elements of Reality

"Elements of physical reality" are a concept introduced by Einstein, Podolsky, and Rosen (collectively "EPR") as an argument against the "completeness" of quantum mechanics in 1935. They define elements of reality as follows [24]:

> If, without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity.

In brief, some physicists at the time were unhappy that quantum mechanics was a nondeterministic theory. EPR presented their paradox and it went unresolved for nearly 30 years. Bell's original proof used inequalities that could be experimentally tested. The proofs in this section, based on the properties of W-states, use no inequalities and are interesting alternatives to the usual proof.

4.2 W-states in X-Basis

The proofs will involve measuring qubits of W-states along both the x- and z-axes. For future reference, I will derive expressions for $|\mathrm{W}_2\rangle$ and $|\mathrm{W}_3\rangle$ in the XX- and XXX-bases using the definitions from Table 2 and properties of the tensor product.

Beginning with $|\mathrm{W}_2\rangle$, note that

$$|01\rangle_{ij} = |0\rangle_i \otimes |1\rangle_j \tag{80}$$

$$= \frac{1}{\sqrt{2}}(|+\rangle_i + |-\rangle_i) \otimes \frac{1}{\sqrt{2}}(|+\rangle_j - |-\rangle_j) \tag{81}$$

$$= \frac{1}{2}\left(|{+}{+}\rangle_{ij} - |{+}{-}\rangle_{ij} + |{-}{+}\rangle_{ij} - |{-}{-}\rangle_{ij}\right). \tag{82}$$

Similarly (omitting subscripts), $|10\rangle = \frac{1}{2}\left(|{+}{+}\rangle + |{+}{-}\rangle - |{-}{+}\rangle - |{-}{-}\rangle\right)$. So in taking a superposition of the two expressions, cross terms cancel, showing that $|\mathrm{W}_2\rangle$ is actually $|\mathrm{GHZ}_2\rangle$ if measured along x:

$$|\mathrm{W}_2\rangle := \frac{1}{\sqrt{2}}|01\rangle + \frac{1}{\sqrt{2}}|10\rangle \tag{83}$$

$$= \frac{1}{\sqrt{2}}|{+}{+}\rangle + \frac{1}{\sqrt{2}}|{-}{-}\rangle. \tag{84}$$

For $n \geq 3$ qubits, W- and GHZ states are truly inequivalent. The following derivation of $|\mathrm{W}_3\rangle$ expressed

in the XXX-basis results in a superposition with many terms. The important detail to note is that $|{+}{+}{+}\rangle$ and $|{-}{-}{-}\rangle$ are each nine times as likely to be the result of a measurement than any other particular state.

$$|001\rangle_{ijk} = |0\rangle_i \otimes |0\rangle_j \otimes |1\rangle_k \tag{85}$$
$$= \frac{1}{\sqrt{8}}(|+\rangle_i + |-\rangle_i) \otimes (|+\rangle_j + |-\rangle_j) \otimes (|+\rangle_k - |-\rangle_k) \tag{86}$$
$$= \frac{1}{\sqrt{8}}(|{+}{+}{+}\rangle - |{+}{+}{-}\rangle + |{+}{-}{+}\rangle - |{+}{-}{-}\rangle + |{-}{+}{+}\rangle - |{-}{+}{-}\rangle + |{-}{-}{+}\rangle - |{-}{-}{-}\rangle) \tag{87}$$
$$|010\rangle = \frac{1}{\sqrt{8}}(|{+}{+}{+}\rangle + |{+}{+}{-}\rangle - |{+}{-}{+}\rangle - |{+}{-}{-}\rangle + |{-}{+}{+}\rangle + |{-}{+}{-}\rangle - |{-}{-}{+}\rangle - |{-}{-}{-}\rangle) \tag{88}$$
$$|100\rangle = \frac{1}{\sqrt{8}}(|{+}{+}{+}\rangle + |{+}{+}{-}\rangle + |{+}{-}{+}\rangle + |{+}{-}{-}\rangle - |{-}{+}{+}\rangle - |{-}{+}{-}\rangle - |{-}{-}{+}\rangle - |{-}{-}{-}\rangle) \tag{89}$$
$$|\mathrm{W}_3\rangle = \frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle) \tag{90}$$
$$= \frac{1}{\sqrt{24}}(3|{+}{+}{+}\rangle + |{+}{+}{-}\rangle + |{+}{-}{+}\rangle - |{+}{-}{-}\rangle + |{-}{+}{+}\rangle - |{-}{+}{-}\rangle - |{-}{-}{+}\rangle - 3|{-}{-}{-}\rangle) \tag{91}$$

4.3 Proof Using Three Qubits

Let three qubits q_1, q_2, and q_3 be entangled in a W-state. Alice, Bob, and Charlie are far apart and each have one qubit. Let x_i and z_i represent the resulting state $(+, -, 0, 1)$ of measuring σ_x and σ_z on the ith qubit.

Observation 1: Suppose Alice and Bob measure z_1 and z_2. This is sufficient information to determine z_3 without having to measure it, so z_3 is an element of reality. By the same reasoning, z_1 and z_2 are also elements of reality.

What about x_i? It is helpful to express $|\mathrm{W}_3\rangle$ in the ZXX-basis:

$$|W_3\rangle = \frac{1}{\sqrt{3}}\left(|001\rangle + |010\rangle + |100\rangle\right) \tag{92}$$
$$= \frac{1}{\sqrt{3}}\left(|0\rangle_1 \otimes (|01\rangle_{23} + |10\rangle_{23}) + |1\rangle_1 \otimes |00\rangle_{23}\right) \tag{93}$$
$$= \frac{1}{\sqrt{3}}\left(|0\rangle_1 \otimes (|{+}{+}\rangle_{23} + |{-}{-}\rangle_{23}) + |1\rangle_1 \otimes |00\rangle_{23}\right) \tag{94}$$
$$\tag{95}$$

Observation 2: Suppose Alice measures $z_1 = 0$, and Bob measures x_2. From the expression above, Alice and Bob can conclude that $x_3 = x_2$ without ever measuring q_3. So x_3 is an element of reality, provided $z_1 = 0$. In general, if $z_i = 0$, then $x_j = x_k$ where i, j, and k are distinct.

In $|\mathrm{W}_3\rangle$, precisely two z values will be 0. By Observation 1, without loss of generality, suppose $z_1 = z_2 = 0$. Since $z_1 = 0$, we know that $x_2 = x_3$ by Observation 2. Similarly, since $z_2 = 0$, we know that $x_1 = x_3$. So

in any W-state, the x values are all $+$ or all $-$. This contradicts the form of $|\mathrm{W}_3\rangle$ given in (91), which allows for every possible combination of $+$ and $-$ with various probabilities. If Alice, Bob, and Charlie were each to measure σ_x of their respective qubits, the results would be inconsistent with elements of reality $1 - 2|\langle +++|\mathrm{W}_n\rangle|^2 = 25\%$ of the time.

4.4 Proof Using More Qubits

Is it possible to do better? Consider $|\mathrm{W}_4\rangle$. As before, each qubit will be in the possession of one person who are all at space-like separations from each other.

Observation 1: Measuring σ_z of any three qubits will identify the fourth, so the z_i's are elements of reality.

Observation 2: Now suppose Alice and Bob measure $z_0 = z_1 = 0$. Then the remaining two qubits, in the state proportional to $|01\rangle + |10\rangle$, have equal σ_x values by (84). In general, for unique i, j, k, l, if $z_i = z_j = 0$, then $x_k = x_l$.

The argument proceeds exactly as it did for the 3-qubit W-state. By Observation 1, without loss of generality identify $z_0 = z_1 = z_2 = 0$. Now by several applications of Observation 2, $x_0 = x_1 = x_2 = x_3$. But actually measuring σ_x for each qubit of $|\mathrm{W}_4\rangle$ reveals that the x_i's are all equal only 50% of the time, corresponding to a 50% contradiction rate.

This reasoning generalizes to n qubits in the following manner: immediately note that z_i is an element of reality, and that any pair of σ_x values must be equal. Therefore all σ_x values must be the same.

The following theorem generalizes the above results.

Theorem 8. *The probability $\mathcal{P}_n$ of a measurement on $|W_n\rangle$ contradicting the prediction of elements of reality is*

$$\mathcal{P}_n = 1 - \frac{n}{2^{n-1}} \tag{96}$$

$$= 2^{-n} \left\langle\!\!\left\langle {n-1 \atop 1} \right\rangle\!\!\right\rangle, \tag{97}$$

where $\left\langle\!\!\left\langle {n \atop m} \right\rangle\!\!\right\rangle$ are the second-order Eulerian numbers [25].

Proof. It should be clear that a contradiction is obtained if and only if the n qubits do not all yield the same result when measured in the X-basis. In other words,

$$\mathcal{P}_n = 1 - 2|\langle +|^{\otimes n} |\mathrm{W}_n\rangle|^2,$$

so it suffices to show that

$$|\langle +|^{\otimes n} |\mathrm{W}_n\rangle|^2 = \frac{n}{2^n}. \tag{98}$$

Rather than express $|\mathrm{W}_n\rangle$ in the X-basis, as was done above, consider the $\langle +|^{\otimes n}$ expressed in the z-basis:

$$\langle +|^{\otimes n} \equiv \left(\frac{\langle 0| + \langle 1|}{\sqrt{2}}\right)^{\otimes n} \tag{99}$$

$$= \frac{1}{2^{n/2}}(\langle 0| + \langle 1|)^{\otimes n}. \tag{100}$$

If the tensor product were commutative, one could apply the binomial expansion directly to (100). Unfortunately, the tensor product is not commutative—but it is distributive, so it is still possible to indirectly use the binomial expansion. To do so, note that the effect of taking the inner product of (100) with $|\mathrm{W}_n\rangle$ is twofold: first, the coefficient of $\frac{1}{\sqrt{n}}$ can be factored out. Second, and more importantly, the remaining sum of one-hot states simply counts the number of terms in (100) that are one-hot.

The binomial expansion of $(a+b)^n$ asserts that the number of terms where a has multiplicity k is $\binom{n}{k}$. We are interested in the number of terms where $|1\rangle$ has multiplicity 1. According to the binomial expansion, there are precisely $\binom{n}{1}$ such terms. So the left-hand side of (98) evaluates to

$$\left|\frac{1}{2^{n/2}\sqrt{n}}\binom{n}{1}\right|^2 = \left(\frac{\sqrt{n}}{2^{n/2}}\right)^2 \tag{101}$$

$$= \frac{n}{2^n} \tag{102}$$

as claimed. □

Since $\mathcal{P}_n$ approaches 1 exponentially fast, for a fixed number of qubits it is significantly more efficient to use a single $|\mathrm{W}_n\rangle$ than two copies of $|\mathrm{W}_{n/2}\rangle$.

There are many proofs of Bell's theorem [26] using many different techniques. The proof using three qubits, which generalized to more qubits, was adapted from [8].

5 W-states and LOCC

5.1 What Is LOCC?

LOCC stands for "local operations with classical communication." If a set of entangled qubits is distributed between two or more distant observers, the observers can perform arbitrary operations on their own qubits (LO), perhaps conditioned on classical information they exchange with each other (CC). LOCC is thus a realistic description of current technological limitations on qubit manipulation.

LOCC techniques are especially appealing when Alice and Bob share a large number of partially entangled qubit pairs. Using LOCC, Alice and Bob can transform a fraction of the pairs into maximally entangled states, which can be used for interesting protocols such as teleportation [27–29].

5.2 LOCC Classes

The set of states that can be converted into each other via LOCC constitutes an LOCC class. LOCC classes are an interesting concept because they represent the entire space of possible quantum states available to observers who share mutually entangled qubits.

LOCC classes are closely related to entanglement. For instance, there is no way to entangle two arbitrary qubits using just LOCC. So for any number of qubits, the unentangled states form an LOCC class. LOCC classes represent fundamentally different resources for observers who share entangled states.

For a system of three qubits, there are a total of six LOCC classes [30] [31] (see Table 7). First is the unentangled set, as usual. Then there are three bipartite entangled sets—i.e., classes whose elements contain a pair of entangled qubits and and unentangled third qubit. The final two classes are both tripartite entangled. One class contains GHZ states, and the other contains W-states. The GHZ and W classes are the most interesting classes with respect to quantum information protocols. The fact that GHZ states and W-states are not equivalent to each other means that, under LOCC, there are tasks that can be performed with each that cannot be duplicated with the other.

Class	Example State
Unentangled	$\|000\rangle$
Bipartite entangled: qubits 0 & 1	$\frac{1}{\sqrt{2}}\|000\rangle + \frac{1}{\sqrt{2}}\|011\rangle$
Bipartite entangled: qubits 0 & 2	$\frac{1}{\sqrt{2}}\|000\rangle + \frac{1}{\sqrt{2}}\|101\rangle$
Bipartite entangled: qubits 1 & 2	$\frac{1}{\sqrt{2}}\|000\rangle + \frac{1}{\sqrt{2}}\|110\rangle$
Tripartite entangled, GHZ	$\frac{1}{\sqrt{2}}\|000\rangle + \frac{1}{\sqrt{2}}\|111\rangle$
Tripartite entangled, W	$\frac{1}{\sqrt{3}}\|001\rangle + \frac{1}{\sqrt{3}}\|010\rangle + \frac{1}{\sqrt{3}}\|100\rangle$

Table 7: Representative states for each of the six LOCC classes on three qubits.

For four or more qubits, the number of distinct LOCC classes goes up considerably and the problem of classifying them and establishing that they are inequivalent is a complex problem that has not yet been fully solved. For further information on this point, see [32].

6 Conclusion

This report has carried out a detailed investigation of two problems involving W-states. The first problem is a study of two different methods of constructing these states, and the second is an exploration of how they can be used to prove Bell's theorem.

Of the two methods of constructing W-states presented in this report, the optimized probabilistic naive construction seems most promising since it can be implemented using elementary gates, a property I have not seen of other methods as in [9], [10], or [21]. I believe the method's simplicity is an added benefit as well.

The proof of Bell's theorem given in this report is a generalization of the proof given by Cabello [8] from three to n qubits. It shows that the probability of experimentally establishing Bell's Theorem approaches 1 as more qubits are incorporated into the W-state. Of course, the scheme would be more difficult to physically implement, but that does not deprive it of its formal interest. It is also interesting that Eulerian numbers of the second kind make their appearance in the proof of the theorem.

This report does not come close to describing all the useful properties that W-states have. A demonstration of their use in a quantum algorithm that solves the n-queens problem is given in [3]. Their robustness of entanglement [7] is one general property, and a few more properties are examined in [33].

To capture all the properties of W-states and compare all the construction methods in one comprehensive text would be a Herculean effort. This report, if nothing else, has been a small step in that direction.

References

[1] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, Oct 1997.

[2] Michael A. Nielsen. Cluster-state quantum computation. *Reports on Mathematical Physics*, 57(1):147–161, Feb 2006.

[3] Rounak Jha, Debaiudh Das, Avinash Dash, Sandhya Jayaraman, Bikash K. Behera, and Prasanta K. Panigrahi. A novel quantum n-queens solver algorithm and its simulation and application to satellite communication using IBM quantum experience, 2018.

[4] Wang Jian, Zhang Quan, and Tang Chao-Jing. Quantum secure communication scheme with W state. *Communications in Theoretical Physics*, 48(4):637–640, oct 2007.

[5] Wen Liu, Yong-Bin Wang, and Zheng-Tao Jiang. An efficient protocol for the quantum private comparison of equality with W state. *Optics Communications*, 284(12):3160 – 3163, 2011.

[6] W. Dür. Multipartite entanglement that is robust against disposal of particles. *Phys. Rev. A*, 63:020303, Jan 2001.

[7] Rafael Chaves and Luiz Davidovich. Robustness of entanglement as a resource. *Physical Review A*, 82(5), Nov 2010.

[8] Adán Cabello. Bell's theorem with and without inequalities for the three-qubit Greenberger-Horne-Zeilinger and W states. *Physical Review A*, 65(032108), 2002.

[9] Diogo Cruz, Romain Fournier, Fabien Gremion, Alix Jeannerot, Kenichi Komagata, Tara Tosic, Jarla Thiesbrummel, Chun Lam Chan, Nicolas Macris, Marc-André Dupertuis, and Clément Javerzac-Galy. Efficient quantum algorithms for GHZ and w states, and implementation on the IBM quantum computer. *Advanced Quantum Technologies*, 2(1900015), 2019.

[10] Firat Diker. Deterministic construction of arbitrary W states with quadratically increasing number of two-qubit gates. *arXiv e-prints*, page arXiv:1606.09290, June 2016.

[11] Xue-Ping Zang, Ming Yang, Fatih Ozaydin, Wei Song, and Zhuo-Liang Cao. Generating multi-atom entangled W states via light-matter interface based fusion mechanism. *Scientific Reports*, 5, 2015.

[12] Fatih Ozaydin, Sinan Bugu, Can Yesilyurt, Azmi Ali Altintas, Mark Tame, and Şahin Kaya Özdemir. Fusing multiple w states simultaneously with a Fredkin gate. *Phys. Rev. A*, 89:042311, Apr 2014.

[13] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, USA, 10th edition, 2011.

[14] N. David Mermin. *Quantum Computer Science: An Introduction*. Cambridge University Press, 2007.

[15] M Born. *The Born-Einstein Letters 1916-1955*. Macmillan Press, New York, 1971.

[16] Daniel M. Greenberger, Michael A. Horne, and Anton Zeilinger. Going beyond Bell's theorem, 2007.

[17] Adriano Barenco, Charles H. Bennett, Richard Cleve, David P. DiVincenzo, Norman Margolus, Peter Shor, Tycho Sleator, John A. Smolin, and Harald Weinfurter. Elementary gates for quantum computation. *Physical Review A*, 52(5), 1995.

[18] Giulio Chiribella, Giacomo Mauro D'Ariano, Paolo Perinotti, and Benoit Valiron. Quantum computations without definite causal structure. *Physical Review A*, 88(2), Aug 2013.

[19] David P. DiVincenzo. Two-bit gates are universal for quantum computation. *Physical Review A*, 51(2), 1995.

[20] Christopher M. Dawson and Michael A. Nielsen. The Solovay-Kitaev algorithm, 2005.

[21] Craig Gidney. General construction of Wn-state, 2018. `https://quantumcomputing.stackexchange.com/questions/4350/general-construction-of-w-n-state`.

[22] Stephen Shankland. IBM's new 53-qubit quantum computer is its biggest yet, 2019. `https://www.cnet.com/news/ibm-new-53-qubit-quantum-computer-is-its-biggest-yet/`.

[23] J. S. Bell. On the Einstein Podolsky Rosen paradox. *Physics Publishing Co.*, 1(3):195–200, November 1964.

[24] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, May 1935.

[25] The on-line encyclopedia of integer sequences, May 2020. `https://oeis.org/A005803`.

[26] Mordecai Waegell and P.K. Aravind. GHZ paradoxes based on an even number of qubits. *Physics Letters A*, 377(7):546 – 549, 2013.

[27] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.*, 70:1895–1899, Mar 1993.

[28] Charles H. Bennett, Herbert J. Bernstein, Sandu Popescu, and Benjamin Schumacher. Concentrating partial entanglement by local operations. *Phys. Rev. A*, 53:2046–2052, Apr 1996.

[29] Jaewoo Joo, Young-Jai Park, Sangchul Oh, and Jaewan Kim. Quantum teleportation via a W state. *New Journal of Physics*, 5:136–136, Oct 2003.

[30] W. Dür, G. Vidal, and J. I. Cirac. Three qubits can be entangled in two inequivalent ways. *Physical Review A*, 62(062314), 2000.

[31] C. Spee, J. I. de Vicente, and B. Kraus. The maximally entangled set of 4-qubit states. *Journal of Mathematical Physics*, 57(052201), 2016.

[32] Eric Chitambar, Debbie Leung, Laura Mančinska, Maris Ozols, and Andreas Winter. Everything you always wanted to know about LOCC (but were afraid to ask). *Communications in Mathematical Physics*, 328(1):303–326, Mar 2014.

[33] Manoranjan Swain, Amit Rai, M. Karthick Selvan, and Prasanta K. Panigrahi. Single photon generation and non-locality of perfect W-state, 2020.

[34] Márcio M. Cunha, Alejandro Fonseca, and Edilberto O. Silva. Tripartite Entanglement: Foundations and Applications. *Universe*, 5(10):209, October 2019.

Appendices

A Expected Cost of Optimized Probabilistic Naive Construction

Theorem 9. *The optimized probabilistic naive construction of* $|W_n\rangle$ *where* $2^{k-1} < n < 2^k =: N$ *requires on average approximately*

$$\frac{N^2 + N + n^2 - n}{2n} \mathcal{C}_{\mathtt{X}}(k) \tag{103}$$

atomic gate applications, where $\mathcal{C}_{\mathtt{X}}(k)$ *is the number of atomic gates needed to simulate a* k*-controlled NOT gate.*

Proof. From Theorem 3, constructing $|\mathrm{W}_N\rangle$ requires about $\mathcal{C}_{\mathtt{X}}(k)N$ atomic gates. A total of δ checks are necessary to filter out the δ surplus states, where each check involves a k-controlled NOT gate and a measurement.

It is optimal to perform the δ checks one at a time, and restart the procedure the first time a $|1\rangle$ is measured. Figure 11 demonstrates this protocol. Let $\mathcal{P}_{\mathrm{s}}$ be the probability that all δ checks return 0 (i.e., the probability of success), and let $\mathcal{P}_{\mathrm{f},m}$ be the probability that the mth check returns 1 (i.e., the probability that it fails on the mth check).

$$\mathcal{P}_{\mathrm{s}} = \frac{n}{N} \tag{104}$$

$$\mathcal{P}_{\mathrm{f},m} = \frac{1}{N}. \tag{105}$$

Let EC_{filter} be the expected total cost of the filtering routine in the construction of $|\mathrm{W}_n\rangle$. Since each check uses about $\mathcal{C}_{\mathtt{X}}(k)$ atomic operations, we have

$$EC_{\text{filter}} = \mathcal{P}_{\text{s}}\delta\mathcal{C}_{\mathtt{X}}(k) + \sum_{m=1}^{\delta} \mathcal{P}_{\text{f},(}m\mathcal{C}_{\mathtt{X}}(k) + EC_{\text{filter}}) \tag{106}$$

$$= \frac{n}{N}\delta\mathcal{C}_{\mathtt{X}}(k) + \sum_{m=1}^{\delta} \frac{1}{N}m\mathcal{C}_{\mathtt{X}}(k) + \delta\frac{1}{N}EC_{\text{filter}} \tag{107}$$

$$= \frac{n}{N}\delta\mathcal{C}_{\mathtt{X}}(k) + \frac{\delta(\delta+1)}{2N}\mathcal{C}_{\mathtt{X}}(k) + \frac{\delta}{N}EC_{\text{filter}} \tag{108}$$

$$= \left(\frac{n}{N} + \frac{\delta+1}{2N}\right)\delta\mathcal{C}_{\mathtt{X}}(k)\left(1 - \frac{\delta}{N}\right)^{-1} \tag{109}$$

$$= \frac{2n+\delta+1}{2N}\delta\mathcal{C}_{\mathtt{X}}(k)\frac{N}{n} \tag{110}$$

$$= \frac{N+n+1}{2n}(N-n)\mathcal{C}_{\mathtt{X}}(k) \tag{111}$$

$$= \frac{N^2+N-n^2-n}{2n}\mathcal{C}_{\mathtt{X}}(k) \tag{112}$$

After the filtering routine, there is still the matter of mapping remaining states to one-hot states. Suppose the δ filtered states were chosen to as in Figure 11, since none of these states are one-hot. Then taking (66) and subtracting off extraneous mappings, the remaining cost is

$$C_{\text{map}} \approx \sum_{l=0}^{k} \mathcal{C}_{\mathtt{X}}(k)\binom{k}{l} - \delta\mathcal{C}_{\mathtt{X}}(k) \tag{113}$$

$$= \mathcal{C}_{\mathtt{X}}(k)N - \delta\mathcal{C}_{\mathtt{X}}(k) \tag{114}$$

$$= n\mathcal{C}_{\mathtt{X}}(k) \tag{115}$$

as one might expect. The expected total cost of OPNC of $|\mathrm{W}_n\rangle$ is then about

$$EC_{\text{total}} = EC_{\text{filter}} + C_{\text{map}} \tag{116}$$

$$\approx \frac{N^2+N-n^2-n}{2n}\mathcal{C}_{\mathtt{X}}(k) + n\mathcal{C}_{\mathtt{X}}(k) \tag{117}$$

$$= \frac{N^2+N+n^2-n}{2n}\mathcal{C}_{\mathtt{X}}(k) \tag{118}$$

where again $k = \lceil \log_2 n \rceil$, $N = 2^k$, and $\delta = N - n$. □

In the worst case scenario when $N \approx 2n$, (118) becomes

$$EC_{\text{total}} \approx \frac{4n^2+2n+n^2-n}{2n}\mathcal{C}_{\mathtt{X}}(k) \tag{119}$$

$$= \frac{5n+1}{2}\mathcal{C}_{\mathtt{X}}(k). \tag{120}$$

Taking $\mathcal{C}_{\mathtt{X}}(k) \sim k \sim \log n$ according to Theorem 2, the expected total cost of OPNC is on the order of $n \log n$.