Optimization of Entanglement Distillation Protocols via Group Theoretical Tools

by

Sarah Jansen

To obtain the degree of Bachelor of Science in Applied Mathematics and Applied Physics at the Delft University of Technology

Student number:	4682513
Supervisors:	Dr. D. Elkouss Coronas
	Dr. D. C. Gijswijt
Other committee members:	Dr. M. Blaauboer
	Dr. J. G. Spandaw

Delft, July 2020

ABSTRACT

This thesis deals with n-to-1 entanglement distillation protocols for bipartite quantum systems in a group theoretical setting. The local operations in the protocols are restricted to Clifford operations. An efficient representation of the elements of the Clifford group in terms of binary matrices is presented. Moreover, a characterization of the subgroup of the Clifford group which leaves the fidelity and the success probability invariant is given. Based on this characterization, a novel approach for the optimization of distillation protocols is presented: instead of checking every element of the Clifford group individually, it is sufficient to consider one element of every right coset of this subgroup. These new insights are used to find optimal protocols for $n = 2, 3, 4$ and 5.

CONTENTS

1

INTRODUCTION

Entanglement is one of the fundamental concepts in quantum mechanics. It has been the topic of many studies since it was first described by Einstein et al., 1935. Not only is entanglement interesting from a theoretical point of view, but it also has a wide range of applications. Within the field of quantum information theory, entanglement is broadly used in for instance quantum cryptography (Deutsch et al., 1996) and quantum teleportation (Bennett et al., 1996c).

Many of these applications exploit pure maximally entangled states. In practice, however, setting up entangled states never results in pure maximally entangled states. Instead, mixed states are created that have a certain probability of being the intended maximally entangled state, but also have a nonzero probability of being a different, unwanted state. This is referred to as noise. A natural question that arises is how to deal with this noise. This problem has first been considered by Bennett et al., 1996c. In their paper, they established a framework for the distillation (or purification) of entanglement. The starting point is a number of copies (n) of a mixed state, which is shared by a number of parties. Using only local operations and classical communication, these noisy states can be transformed into a smaller number of copies (m) of a state close to a maximally entangled state. Such a set of local operations and classical communication is called an n-to-m entanglement distillation protocol.

In this thesis n-to-1 entanglement distillation protocols for bipartite systems are studied. In Deutsch et al., 1996, a 2-to-1 protocol was published, which has been proved to be optimal for various input states, see for instance Dehaene et al., 2003b and Rozpędek et al., 2018. Much is unknown, however, about n-to-1 protocols for larger values of n. One of the difficulties of studying these protocols lies in the number of possible local operations that can be performed, because this number increases rapidly with the number of copies in the input state. Therefore, the main goal of this thesis is to decrease the number of operations that needs to be considered when looking for an optimal protocol.

To this end, entanglement distillation protocols are studied in a group theoretical setting. In this setting, quantum states are described in terms of Pauli matrices. These states are acted upon by elements of the Clifford group, an important specific class of quantum operations. Special attention is put to the equivalent description of the quantum states and Clifford operations in terms of binary linear algebra. This description is based on Dehaene and Moor, 2003a and Hostens, 2007, but is reformulated here in the context of entanglement distillation protocols.

To decrease the number of operations that need to be considered, we search for elements of the Clifford group that do not affect the outcome of the protocol. These elements form a subgroup of the Clifford group that is referred to as the *subgroup that preserves the distillation statistics*. The main novel contribution of this thesis is a characterization of this subgroup in terms of its order and a generating set of operations. Moreover, it is shown that every element of the Clifford group can be decomposed into an element of the subgroup that preserves distillation statistics and an (other) element of the Clifford group. As a result, in the optimization of the distillation protocols, it is sufficient to consider only one element of every right coset of the subgroup in the Clifford group, instead of checking every element of the Clifford group individually. These new insights are used to find optimal protocols for $n = 2, 3, 4$ and 5.

In Chapter 2 and Chapter 3 the quantum mechanical and mathematical frameworks for the remainder of this thesis are given. Firstly, in Chapter 2 the necessary concepts of quantum information theory are introduced. Then, in Chapter 3 the group theoretical framework to describe distillation protocols is given. This includes the representation of the relevant group in terms of binary matrix algebra, which we reformulated in a way that is more convenient to work with in the context of distillation protocols.

In Chapter 4 a general description of entanglement distillation protocols is given. This includes an introduction to the concept of quantum entanglement and a description of the structure of distillation protocols. As an example, the protocol published by Deutsch et al., 1996 is considered. Starting from the general description of protocols, Chapter 5 describes the restrictions made to arrive at the bi-local Clifford circuits, which are the protocols of interest in this thesis. We introduce the new concepts of *base* and *pillars*, which allow for a visual way of thinking about distillation protocols. Keeping this visualization in mind, we present and prove two theorems about the structure of the subgroup that preserves the distillation statistics. Firstly, we come up with a generating set of operations for the subgroup, and then we provide a proof for the order of the subgroup. The latter indicates that for optimizing the distillation protocols it is indeed far more efficient to consider only one element of every right coset of the subgroup

In Chapter 6 we provide an algorithm for finding a transversal of the cosets. Moreover, it is explained how, using this transversal, optimal protocols can be found, based on a variety of criteria. The results of this optimization for various input states are shown in Chapter 7. Finally, Chapter 8 provides a conclusion on the main results achieved in this thesis and a discussion on the implications of the research in this thesis.

This thesis has been written as part of the double bachelor's degree Applied Mathematics and Applied Physics at the Delft University of Technology.

2

INTRODUCTION TO QUANTUM INFORMATION THEORY

In this chapter an introduction to the main concepts from quantum information theory used in this thesis is provided. Firstly, in Section 2.1 the framework that is needed to describe quantum systems is given. Then, in Section 2.2, the quantum circuit model is explained. The quantum circuit model is a model to describe quantum computations. It is used later on in this thesis to describe entanglement distillation protocols. This chapter is mainly based on Nielsen and Chuang, 2016.

2.1. DESCRIBING A QUANTUM STATE

In this section the framework to describe the state of a quantum system is given. After a general introduction, including the definition of a qubit, four different aspects of describing a quantum system are explained. Firstly in Section 2.1.1 the density operator formulation is treated. In Section 2.1.2 the tensor product is introduced as a way to describe composite quantum systems. In Section 2.1.3 the knowledge from these two sections is combined to arrive at the reduced density operator. Finally, in Section 2.1.4 measurements on quantum systems are considered.

In quantum information theory the possible states of a quantum system are often represented by unit vectors in a Hilbert space.

Definition 2.1. (Driver, 2003) A (complex) *Hilbert space* $(\mathscr{H}, \langle\cdot,\cdot\rangle)$ is a (complex) vector space with inner product $\langle\cdot,\cdot\rangle$ such that the induced norm $\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}$ is complete.

In this thesis, if $\mathscr{H}$ is used, a Hilbert space is implied. We will confine ourselves to finite dimensional Hilbert spaces. Every finite dimensional Hilbert space of dimension n is isomorphic to $\mathbb{C}^n$ (see for instance Triebel, 1986). In quantum information theory, the 'bra-ket' notation is often used to denote vectors. In this notation a vector is denoted by $|\cdot\rangle \in \mathscr{H}$ (the 'ket'). Its conjugate transpose is denoted by $\langle\cdot| \in \mathscr{H}^*$ (the 'bra'). In the

'bra-ket' notation, inner products can be written as $\langle\cdot|\cdot\rangle$.

As mentioned before, a quantum system can be represented by a normalized vector $|\Psi\rangle \in \mathscr{H}$. The system that we will be most concerned with in this thesis is the qubit.

Definition 2.2. A *qubit* is a quantum state that can be represented by a vector $|\Psi\rangle \in \mathbb{C}^2$,

$$|\Psi\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \qquad \alpha, \beta \in \mathbb{C},\ |\alpha|^2 + |\beta|^2 = 1$$

Transformations of a closed quantum system (i.e., a system that is not interacting with its surroundings) can be described by unitary operators $U : \mathscr{H} \to \mathscr{H}$, $|\psi\rangle \to U|\psi\rangle$.

2.1.1. DENSITY OPERATOR

The notation we have used so far can be used to describe *pure states*. However, sometimes the state of a quantum system is not exactly known. For example, a system is prepared in state $|\psi_1\rangle$ with probability p_1 an in state $|\psi_2\rangle$ with probability p_2. The state of such a system is called a *mixed state*. To describe a mixed state, a density operator can be used. Density operators can also be used to describe an entangled state that is shared by two parties that are separated in distance. Entanglement will be discussed in more detail in Chapter 4.

Definition 2.3. Suppose a quantum system is in one of the states $|\psi_i\rangle$, each with probability p_i. Then the *density operator* is given by

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$$

The density operator for a pure state $|\psi\rangle$ simply equals $|\psi\rangle\langle\psi|$. Theorem 2.4 gives a characterization of general density operators.

Theorem 2.4. *(Nielsen and Chuang, 2016) An operator ρ is a density operator associated to an ensemble $\{p_i, |\psi_i\rangle\}$ if and only if the following two conditions are satisfied.*

1. $\mathrm{Tr}(\rho) = 1$
2. *ρ is a positive semidefinite operator.*

Proof. Suppose $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ is a density operator. Then

$$\mathrm{Tr}(\rho) = \mathrm{Tr}\left(\sum_i p_i |\psi_i\rangle\langle\psi_i|\right) = \sum_i p_i \mathrm{Tr}(|\psi_i\rangle\langle\psi_i|) = \sum_i p_i \langle\psi_i|\psi_i\rangle = \sum_i p_i = 1.$$

To prove that ρ is positive semidefinite, first note that ρ is Hermitian. Indeed,

$$\rho^\dagger = \left(\sum_i p_i |\psi_i\rangle\langle\psi_i|\right)^\dagger = \sum_i p_i (|\psi_i\rangle\langle\psi_i|)^\dagger = \sum_i p_i |\psi_i\rangle\langle\psi_i| = \rho.$$

Moreover, let $|\phi\rangle \in \mathscr{H}$ be an arbitrary vector. Then

$$\langle\phi|\rho|\phi\rangle = \sum_i p_i \langle\phi|\psi_i\rangle\langle\psi_i|\phi\rangle = \sum_i p_i |\langle\phi|\psi_i\rangle|^2.$$

Since the p_i are probabilities, they are larger than or equal to 0, so $\langle\phi|\rho|\phi\rangle \geq 0$. Thus, ρ is positive semidefinite.

Conversely, suppose ρ is a positive semidefinite operator satisfying $\mathrm{Tr}(\rho) = 1$. Since ρ is positive semidefinite, it follows from the spectral decomposition theorem that

$$\rho = \sum_j \lambda_j |j\rangle\langle j|,$$

where the vectors $|j\rangle$ are orthogonal and λ_j are real and non-negative eigenvalues of ρ. Because $\mathrm{Tr}(\rho) = 1$, it follows that $\sum_j \lambda_j = 1$. Thus, ρ can be seen as the density operator for the ensemble $\{\lambda_j, |j\rangle\}$. □

Suppose that the evolution of a closed quantum system in time is given by U. If the system was originally in state $|\psi_i\rangle$ with probability p_i, then after the transformation it is in state $U|\psi_i\rangle$ with the same probability p_i. Hence, the evolution in time of a density operator is described by

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \xrightarrow{U} \sum_i p_i U|\psi_i\rangle\langle\psi_i|U^\dagger = U\rho U^\dagger \tag{2.1}$$

2.1.2. COMBINING SYSTEMS: TENSOR PRODUCT

In order to combine multiple quantum systems, the tensor product is widely used in quantum mechanics. The definition that is used, was provided by Fulton and Harris, 2004. Before stating the definition, firstly a definition of a bilinear map is given.

Definition 2.5. Let V, W and U be vector spaces over the same field K. A function $\beta : V \times W \to U$ is a *bilinear map* if for all $w \in W$ the map $v \mapsto \beta(v, w)$ is linear and for all $v \in V$ the map $w \mapsto \beta(v, w)$ is linear.

A bilinear map $\beta : V \times W \to U$ thus satisfies the following two properties for all $a, b \in K$:

1. For all $v_1, v_2 \in V, w \in W$: $\beta(av_1 + bv_2, w) = a\beta(v_1, w) + b\beta(v_2, w)$.
2. For all $v \in V, w_1, w_2 \in W$: $\beta(v, aw_1 + bw_2) = a\beta(v, w_1) + b\beta(v, w_2)$.

Using this definition, the tensor product can now be defined as follows.

Definition 2.6. The *tensor product* of two vector spaces V and W (over a field K) is a vector space $V \otimes W$ equipped with a bilinear map $\beta : V \times W \to V \otimes W, \quad (v, w) \mapsto v \otimes w$ that is universal: for every bilinear map $\gamma : V \times W \to U$ to a vector space U, there is a unique linear map $\tilde{\gamma} : V \otimes W \to U$ that takes $v \otimes w$ to $\gamma(v, w)$.

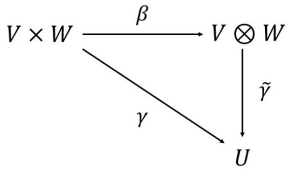


Figure 2.1: β is a universal map: for every bilinear map $\gamma : V \times W \to U$, there is a unique linear map $\tilde{\gamma} : V \otimes W \to U$ such that $(\tilde{\gamma} \circ \beta)(v, w) = \gamma(v, w)$.

The picture to keep in mind for this rather abstract definition is shown in Figure 2.1. If $\{e_i\}$ is a basis for V and $\{f_j\}$ is a basis for W, then $\{e_i \otimes f_j\}$ is a basis for $V \otimes W$. The tensor product is associative: $(U \otimes V) \otimes W = U \otimes (V \otimes W) = U \otimes V \otimes W$. The *tensor power* $V^{\otimes n}$ of a vector space V is defined as the n-fold tensor product of V with itself: $V^{\otimes n} = V \otimes \cdots \otimes V$.

For linear operators acting on V and W, the tensor product is defined as follows.

Definition 2.7. Let A and B be linear operators that act on V and W, respectively. Then the *tensor product of* those *operators* is defined as

$$(A \otimes B)\left(\sum_i a_i v_i \otimes w_i\right) = \sum_i a_i A(v_i) \otimes B(v_i), \quad \forall v_i \in V,\ \forall w_i \in W,\ \forall a_i \in K$$

This rather abstract discussion can be made more concrete by looking at an example: the Kronecker product. This is the form of the tensor product that will be used throughout the rest of this thesis. The definition of the Kronecker product is based on Horn and Johnson, 1991.

Definition 2.8. Let $A = [a_{ij}] \in M_{m\times n}(K)$ $B = [b_{ij}] \in M_{p\times q}(K)$. Then the *Kronecker product* is defined as

$$A \otimes B \equiv \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1}B & a_{n2}B & \cdots & a_{nn}B \end{bmatrix}$$

A useful property of the Kronecker product is the mixed-product property, which gives a relation between the Kronecker product and the standard matrix multiplication. It is proved in Horn and Johnson, 1991.

Theorem 2.9. *Let $A \in M_{m\times n}(K)$, $B \in M_{p\times q}(K)$, $C \in M_{n\times k}(K)$ and $D \in M_{q\times l}(K)$ Then*

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$$

2.1.3. REDUCED DENSITY OPERATOR

Suppose now that we have a composite system, consisting of the subsystems A and B, whose state is described by the density operator ρ. To go back to the description of either system A or system B, we can use the reduced density operator.

Definition 2.10. Let ρ be the density operator describing a composite quantum system. The *reduced density operator* for subsystem A is defined by

$$\rho^A = \mathrm{Tr}_B(\rho),$$

where $\mathrm{Tr}_B(\rho)$ denotes the partial trace over system B. For a pure tensor product $|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2|$, it is defined as $\mathrm{Tr}_B(|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2|) = |a_1\rangle\langle a_2|\mathrm{Tr}_B(|b_1\rangle\langle b_2|)$. This can be extended linearly for a general tensor product.

The partial trace is a trace-preserving completely positive map, meaning that, if ρ is positive semidefinite, then $\mathrm{Tr}_B \rho$ is positive semidefinite too and that $\mathrm{Tr}(\mathrm{Tr}_B \rho) = \mathrm{Tr}(\rho)$ (Hayashi, 2006). Consequently, ρ_A satisfies the conditions from Theorem 2.4, thus ρ_A is a density operator. Of course, the roles of A and B can be switched in the discussion above to derive a similar statement for ρ_B.

2.1.4. QUANTUM MEASUREMENTS

Suppose that we want to perform measurements on a quantum system. Measurements imply interactions of the quantum system with the environment, which makes the system no longer closed. Hence, these interactions are not necessarily described by unitary operations. In this section a way to describe quantum measurements is provided.

A quantum measurement is described by a collection of operators $\{M_a\}$ acting on $\mathcal{H}$. Here a refers to the possible outcomes of the measurement. If the system is in state $|\Psi\rangle$ before the measurement, the probability of obtaining outcome a is given by

$$p(a) = \langle\psi| M_a^\dagger M_a |\psi\rangle. \tag{2.2}$$

If the measurement yields the outcome a, then the new state (the state after the measurement) is described by

$$|\psi^a\rangle = \frac{M_a|\psi\rangle}{\sqrt{\langle\psi| M_a^\dagger M_a |\psi\rangle}}. \tag{2.3}$$

The measurement operators must satisfy the completeness relation

$$\sum_a M_a^\dagger M_a = I. \tag{2.4}$$

This relation ensures the fact that the probabilities associated to the measurement outcomes sum to one:

$$\sum_a p(a) = \sum_a \langle\psi| M_a^\dagger M_a |\psi\rangle = \langle\psi|\psi\rangle = 1. \tag{2.5}$$

An important example of a measurement is the measurement of a qubit in the computational basis. This measurement is defined by the measurement operators $M_0 = |0\rangle\langle 0|$

and $M_1 = |1\rangle\langle 1|$. Suppose that the state that is being measured is $|\phi\rangle = \alpha|0\rangle + \beta|1\rangle$. The probability of obtaining the outcome 0 then equals

$$p(0) = \langle\psi| M_0^\dagger M_0 |\psi\rangle = |\alpha|^2. \tag{2.6}$$

If the outcome 0 is obtained, the state after the measurement is

$$\frac{M_0|\phi\rangle}{|\alpha|} = \frac{\alpha}{|\alpha|}|0\rangle, \tag{2.7}$$

which is equal to the state $|0\rangle$ up to a phase factor. Similarly, the probability of obtaining the outcome 1 equals $p(1) = |\beta|^2$ and the state after the measurement is $\frac{\beta}{|\beta|}|1\rangle$.

Measurements can also be expressed in terms of density operators. If the system was originally in state $|\psi_i\rangle$, then it follows from equation (2.2) that the probability of measuring the outcome a is equal to

$$p(a|i) = \langle\psi_i| M_a^\dagger M_a |\psi_i\rangle. \tag{2.8}$$

Hence, the probability of obtaining outcome a is given by

$$\begin{aligned} p(a) &= \sum_a p(a|i)p_i \\ &= \sum_a p_i \langle\psi_i| M_a^\dagger M_a |\psi_i\rangle \\ &= \sum_a p_i \operatorname{Tr}\left(M_a^\dagger M_a |\psi_i\rangle\langle\psi_i|\right) \\ &= \operatorname{Tr}\left(M_a M_a^\dagger \rho\right). \end{aligned} \tag{2.9}$$

Using equation (2.3) it can be derived that the state after the measurement can be described by

$$\begin{aligned} \rho_a &= \sum_i p(i|a) |\psi_i^a\rangle\langle\psi_i^a| \\ &= \sum_i p(i|a) \frac{M_a|\psi_i\rangle\langle\psi_i| M_a^\dagger}{\langle\psi| M_a^\dagger M_a |\psi\rangle}. \end{aligned} \tag{2.10}$$

It follows from Bayes' theorem that $p(i|a) = p(i,a)/p(a) = p(a|i)p_i/p(a)$. By equations (2.8) and (2.9) we can thus conclude that

$$\begin{aligned} \rho_a &= \sum_i p_i \frac{M_a|\psi_i\rangle\langle\psi_i| M_a^\dagger}{\operatorname{Tr}\left(M_a M_a^\dagger \rho\right)} \\ &= \frac{M_a \rho M_a^\dagger}{\operatorname{Tr}\left(M_a M_a^\dagger \rho\right)}. \end{aligned} \tag{2.11}$$

2.2. QUANTUM CIRCUITS

In this section a model to describe quantum computations, the quantum circuit model, is explained. This section includes the introduction of important transformations of single qubit states (Section 2.2.1) and of multiple qubit states (Section 2.2.2). The quantum circuit model will be used later on to describe entanglement distillation protocols.

Classical computation is based upon the concept of a bit, which can take the value '0' or '1'. Similarly, quantum computation is based on the concept of a qubit. As we saw in Definition 2.2, a qubit can have many different states. It is customary to associate the classical bits with qubits as follows

$$0 \to |0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad 1 \to |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Note that $|0\rangle$ and $|1\rangle$ form a basis for the Hilbert space $\mathbb{C}^2$. This basis is referred to as the *standard basis* or *computational basis*. The state of a general qubit can now be described by

$$\alpha|0\rangle + \beta|1\rangle, \quad \alpha, \beta \in \mathbb{C}, \ |\alpha|^2 + |\beta|^2 = 1. \tag{2.12}$$

2.2.1. SINGLE QUBIT OPERATIONS

We start off with operations on a single qubit. In Section 2.1 we saw that transformations of quantum states can be described by unitary operations. For a one-qubit system, these transformations are thus represented by 2×2 unitary matrices. In the context of quantum computation, we will often refer to the transformation operators as *gates*. For a general gate U, the gate can visually be represented as shown in Figure 2.2.

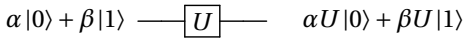


Figure 2.2: Visual representation of a single qubit gate.

Important examples of such gates are the Pauli gates:

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$
$$Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \qquad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{2.13}$$

The action of the X and Z gate on a qubit is shown in Figure 2.3. Note that $Y = iXZ$, so the action of Y can be derived from the actions of X and Z, up to a factor i. This factor i only results in a multiplication of the whole state with i, which will show to be irrelevant in the context of entanglement distillation protocols, as explained in more detail in Chapter 5.

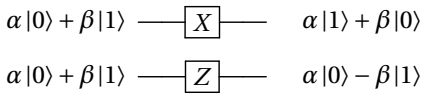


Figure 2.3: Action of the X and the Z gate on a single qubit.

Another gate that we will encounter more often is the Hadamard gate.

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{2.14}$$

The action of the Hadamard gate on a single qubit is shown in Figure 2.4.

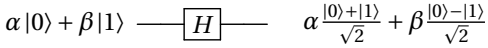


Figure 2.4: Action of the Hadamard gate on a single qubit.

The final single qubit gate that we discuss here is the phase gate.

$$S = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix} \tag{2.15}$$

The action of the phase gate on a single qubit is shown in Figure 2.5.

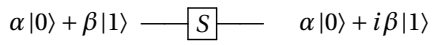


Figure 2.5: Action of the phase gate on a single qubit.

In the rest of this thesis, we will use the notation U_i to indicate a U gate that acts on qubit i.

2.2.2. MULTIPLE QUBIT OPERATIONS

Now let us continue with generalizing to operations on multiple qubits. One of the most used operations in classical computing is the controlled operation: *If A is true, then do B.* In quantum circuits this type of operations can be implemented as well. For an arbitrary unitary operation U we can define the controlled-U operation as an operation on two qubits, one of which will be referred to as the control qubit and the other one as the target qubit. If the control qubit is in state $|0\rangle$, then the target qubit is left alone. If the control qubit is in state $|1\rangle$, then U is applied to the target qubit.

An example of a controlled-U gate is the controlled-X or controlled-NOT gate (CNOT). In terms of the computational basis it acts as follows: if the control qubit is in the state $|1\rangle$, the target qubit is flipped. Otherwise, the target qubit is left alone. The matrix

representation (in the computational basis) is given by

$$\text{CNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{2.16}$$

We will use the notation CNOT_{ij} to indicate a CNOT gate with control qubit i and target qubit j. The visual representation of the CNOT gate together with its action on a two-qubit state is shown in Figure 2.6.

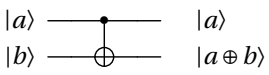


Figure 2.6: Action of the CNOT gate on a two-qubit system. The qubits are originally is states $|a\rangle$ and $|b\rangle$. After applying the CNOT gate, the target qubit is in state $|a \oplus b\rangle$, where $\oplus$ denotes addition modulo 2.

Another important example of a two-qubit gate is the controlled-Z gate (CZ). In the computational basis, this gate can be represented as follows.

$$\text{CZ} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \tag{2.17}$$

The visual representation of a CZ gate is shown in Figure 2.7a. If the control qubit is in state $|0\rangle$, then the target qubit does not change. If the control qubit is in state $|1\rangle$, then the CZ gate acts as a Z gate on the target qubit. As a result, the CZ gate acts as the identity on $|00\rangle$, $|10\rangle$ and $|01\rangle$ and changes $|11\rangle$ to $-|11\rangle$. Thus, the CZ gate is symmetric in the control and the target qubit. Therefore, it is often represented by two controls, as is shown in Figure 2.7b.

(a) (b)

Figure 2.7: Two visual representations of the CZ gate.

We will use the notation CZ_{ij} to indicate a CZ gate between qubits i and j. We will end this section with two useful circuit identities that will be used frequently in the rest of this thesis. Firstly, by noting that $X = HZH$, it follows that the CNOT gate can be rewritten into a CZ gate and vice versa. This equivalence is shown in Figure 2.8.

(a) (b)

Figure 2.8: CNOT gate rewritten in terms of a CZ gate and two H gates on the target qubit (a) and vice versa (b).

Secondly, using 3 CNOT gates a SWAP gate can be made. The SWAP gates permutes the two qubits that it acts on. The SWAP gate is represented by a cross on the two qubits. It can be built from CNOT gates as shown in Figure 2.9.

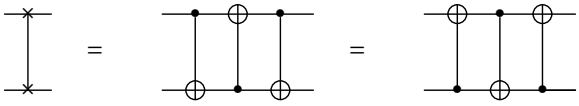


Figure 2.9: SWAP gate written in terms of 3 CNOT gates.

3

GROUP THEORETICAL FRAMEWORK

This chapter describes the group theoretical framework for the description of distillation protocols. In Section 3.1, we start with the introduction of the Pauli group, which will be used to describe the state of a quantum system, as will be explained in Section 4.1.2. The other group that plays an important role in this thesis is the Clifford group. This is the group that consists of all local operations in distillation protocols that are considered in this thesis (see Chapter 5).

After these groups are introduced, in Section 3.2 an alternative representation of the Pauli group and the Clifford group in terms of binary matrices is presented. It is shown that every Clifford operation can be represented as an element of the symplectic group $Sp(2n, \mathbb{Z}_2)$. The main result of this section is the characterization of the homomorphism from the Clifford group to the symplectic group. In Theorem 3.12 it is shown that the homomorphism is surjective. That is, all symplectic matrices correspond to a Clifford operation. Moreover, in Theorem 3.13, a description of the kernel of the homomorphism is provided.

3.1. THE PAULI GROUP AND THE CLIFFORD GROUP

In this section, the Pauli group and the Clifford group are introduced and some of their properties are discussed.

3.1.1. PAULI GROUP

We begin this section with a recap of the definition of the Pauli matrices

$$\begin{aligned} \sigma_0 = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad & \sigma_1 = X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \\ \sigma_2 = Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \qquad & \sigma_3 = Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \end{aligned} \tag{3.1}$$

which are widely used in quantum mechanics. The Pauli matrices, including the multiplicative factors ± 1, $\pm i$, form a group under matrix multiplication:

$$\mathcal{P}_1 = \{\pm I, \pm iI, \pm X, \pm iX, \pm Y, \pm iY, \pm Z, \pm iZ\}. \tag{3.2}$$

The group $\mathcal{P}_1$ is referred to as the Pauli group on one qubit and has order $|\mathcal{P}_1| = 16$. The group $\mathcal{P}_1$ can be extended to $\mathcal{P}_n$, the Pauli group on n qubits. The group $\mathcal{P}_n$ contains all $2^n \times 2^n$ matrices of the form $\lambda P_1 \otimes \cdots \otimes P_n$ with $\lambda \in \{\pm 1, \pm i\}$ and $P_i \in \{I, X, Y, Z\}$ for all $i \in \{1, ..., n\}$. An element of this form will be referred to as a *Pauli string of length n* or a *Pauli string* if n is clear from the context. The Pauli strings of length n form a group under standard matrix multiplication, which can be carried out using the mixed-product property (see Theorem 2.9). The order of $\mathcal{P}_n$ equals $|\mathcal{P}_n| = 4^{n+1}$.

The commutator of two matrices is defined as

$$[A, B] = AB - BA, \tag{3.3}$$

and the anti-commutator is defined as

$$\{A, B\} = AB + BA. \tag{3.4}$$

It can be easily verified that the identity matrix σ_0 commutes with all Pauli matrices. Furthermore, the Pauli matrices σ_1, σ_2 and σ_3 obey the following commutation relations:

$$[\sigma_j, \sigma_k] = 2i \sum_{l=1}^{3} \epsilon_{jkl} \sigma_l, \qquad j, k, l \in \{1, 2, 3\}, \tag{3.5}$$

where ϵ_{jkl} is the Levi-Civita symbol for three indices:

$$\epsilon_{jkl} = \begin{cases} 1, & (j,k,l) \in \{(1,2,3), (2,3,1), (3,1,2)\}, \\ -1, & (j,k,l) \in \{(3,2,1), (2,1,3), (1,3,2)\}, \\ 0, & \text{otherwise.} \end{cases} \tag{3.6}$$

Moreover, they obey the following anti-commutation relations:

$$\{\sigma_j, \sigma_k\} = 2\delta_{jk} \sigma_0, \qquad j, k \in \{1, 2, 3\}, \tag{3.7}$$

where δ_{jk} is the Kronecker delta:

$$\delta_{jk} = \begin{cases} 1, & j = k, \\ 0, & j \neq k. \end{cases} \tag{3.8}$$

Using the mixed-product property of the Kronecker product (Theorem 2.9), these commutation relations can be extended to general Pauli strings.

3.1.2. CLIFFORD GROUP

In this section the Clifford group, which will be extensively used in the remainder of this thesis, is introduced. The Clifford group is a group of unitary matrices that maps the set of the Pauli group to itself under conjugation. Let $U(n)$ denote the group of all unitary $n \times n$ matrices. Observe that the Pauli group $\mathcal{P}_n$ is contained in the unitary group $U(2^n)$. The normalizer of $\mathcal{P}_n$ in $U(2^n)$ is defined as

$$N_{U(2^n)}(\mathcal{P}_n) = \{C \in U(2^n) : \forall P \in \mathcal{P}_n \;\; CPC^\dagger \in \mathcal{P}_n\}. \tag{3.9}$$

Note that $N_{U(2^n)}(\mathcal{P}_n)$ has an infinite center $\{e^{2\pi i\theta} I, \theta \in \mathbb{R}\}$. Although quantum mechanically, a factor $e^{2\pi i\theta}$ makes no difference ($e^{2\pi i\theta} P e^{-2\pi i\theta} = P$ for all $\theta \in \mathbb{R}$, $P \in \mathcal{P}_n$), we prefer to work with a finite group. For this reason, in some literature the Clifford group is defined as $\{C \in U(2^n) : \forall P \in \mathcal{P}_n \;\; CPC^\dagger \in \mathcal{P}_n\}/\{e^{2\pi i\theta} I, \theta \in \mathbb{R}\}$ (see for instance Ozols, 2008). However, we prefer to work with matrices instead of equivalence classes of matrices, which is why a different definition is adopted here. In accordance with the definition provided by Calderbank et al., 1998, the Clifford group is defined as follows:

Definition 3.1. The *Clifford group* on n qubits is the subgroup of $N_{U(2^n)}(\mathcal{P}_n)$ whose elements are matrices with entries from the field $\mathbb{Q}(\eta)$, with $\eta = \frac{1+i}{\sqrt{2}}$. [1]

The field $\mathbb{Q}(\eta)$ is the smallest field which contains $\mathbb{Q}$ and η. It consists of the elements $\{a + b\eta + c\eta^2 + d\eta^3 : a, b, c, d \in \mathbb{Q}\}$. The motivation behind this restriction to entries from $\mathbb{Q}(\eta)$ is that it is the smallest field such that $\mathcal{C}_n$ contains the Pauli matrices, the phase gate, the Hadamard gate and, for $n > 1$, the CNOT gate. The Clifford group from Definition 3.1 is known to have order $|\mathcal{C}_n| = 2^{n^2+2n+3} \prod_{j=1}^{n} \left(4^j - 1\right)$ (see Calderbank et al., 1998). Moreover, it is proved in Section 5.6 and 5.8 of Gottesman, 1997 that $\mathcal{C}_n$ is generated by the set of phase and Hadamard gates on every qubit and CNOT gates on every pair of qubits.

The next theorem tells us something about the structure of these Clifford operations.

Theorem 3.2. *Let $C \in \mathcal{C}_n$ be a Clifford operation. Then $\sigma : \mathcal{P}_n \to \mathcal{P}_n$ defined as $\sigma(P) = CPC^\dagger$ is an automorphism.*

Proof. Let $P, Q \in \mathcal{P}_n$. By definition, C is unitary, so $CC^\dagger = C^\dagger C = I$. Thus, $\sigma(PQ) = CPQC^\dagger = CPC^\dagger CQC^\dagger = \sigma(P)\sigma(Q)$, so σ is a homomorphism. This directly implies that σ is an endomorphism. To show that σ is an automorphism, it is now enough to show that σ is injective, since $\mathcal{P}_n$ is finite. We have

$$\begin{aligned} \sigma(P) = I &\implies CPC^\dagger = I \\ &\implies C^\dagger CPC^\dagger C = C^\dagger IC \\ &\implies P = I. \end{aligned}$$

Thus, $\ker(\sigma) = \{\mathrm{I}\}$ and σ is indeed injective. □

[1] In the definition from Calderbank et al., 1998 the ring $\mathbb{Q}[\eta]$ is used, which is the smallest ring that contains $\mathbb{Q}$ and η. However, because η is an algebraic number, we have $\mathbb{Q}[\eta] = \mathbb{Q}(\eta)$.

It follows from Theorem 3.2 that conjugation by Clifford operations preserves the commutation relations of the Pauli strings. Indeed, suppose that $P, Q \in \mathcal{P}_n$ commute. Then $\sigma(P)\sigma(Q) = \sigma(PQ) = \sigma(QP) = \sigma(Q)\sigma(P)$, so $\sigma(P)$ and $\sigma(Q)$ commute. Similarly, if P and Q anti-commute, then $\sigma(P)\sigma(Q) = \sigma(PQ) = \sigma(-QP) = -\sigma(QP) = -\sigma(Q)\sigma(P)$, so $\sigma(P)$ and $\sigma(Q)$ anti-commute.

Since σ is a homomorphism, it is fully determined by the image of a generating set of $\mathcal{P}_n$. In Theorem 3.4 it is proved that this determines C up to a phase factor, but first a preliminary lemma is proved.

Lemma 3.3. *The elements from the group $\mathcal{P}_n$ span the matrix space $M(2^n, \mathbb{C})$.*

Proof. Firstly, we show that the Pauli matrices I, X, Y and Z form a basis for the matrix space $M_{2\times 2}(\mathbb{C})$. Because $\dim(M_{2\times 2}(\mathbb{C})) = 4$, it is sufficient to show that the four Pauli matrices are linearly independent. Let c_0, c_1, c_2, $c_3 \in \mathbb{C}$ such that

$$c_0 I + c_1 X + c_2 Y + c_3 Z = 0.$$

Then

$$\begin{bmatrix} c_0 + c_3 & c_1 - ic_2 \\ c_1 + ic_2 & c_0 - c_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

so $c_0 = c_1 = c_2 = c_3 = 0$. Thus indeed, the set $\{I, X, Y, Z\}$ is a basis for $M_{2\times 2}(\mathbb{C})$. As a consequence, the set of n-fold tensor products of I, X, Y and Z is a basis for $M_{2^n\times 2^n}(\mathbb{C})$. Because $\mathcal{P}_n$ contains these tensor products, it follows that $\mathcal{P}_n$ spans $M_{2^n\times 2^n}(\mathbb{C})$ too. □

Theorem 3.4. *Let $C \in \mathcal{C}_n$ such that $CPC^\dagger = P$ for all $P \in \mathcal{P}_n$. Then $C = \lambda I$ with $\lambda \in \mathbb{C}$.*

Proof. Note that $CPC^\dagger = P$ implies that $CP = PC$ for all $P \in \mathcal{P}_n$. Thus, C commutes with all elements of $\mathcal{P}_n$. From Lemma 3.3 it follows that C must commute with all elements of $M(2^n, \mathbb{C})$. We show that this implies that $C = \lambda I$ with $\lambda \in \mathbb{C}$.

Let $E_{ij} \in M_{2^n\times 2^n}(\mathbb{C})$ denote the matrix with a 1 on position (i, j) and 0 elsewhere. Note that every matrix in $M_{2^n\times 2^n}(\mathbb{C})$ is a linear combination of matrices E_{ij} with $i, j \in \{1, ..., 2^n\}$. Hence, C must commute with E_{ij} for all $i, j \in \{1, ..., 2^n\}$.

Note that we can write

$$C = \sum_{k=1}^{2^n} \sum_{l=1}^{2^n} c_{kl} E_{kl}.$$

It follows that

$$CE_{ij} = \sum_{k=1}^{2^n} = c_{ki} E_{kj},$$

and that

$$E_{ij} C = \sum_{l=1}^{2^n} = c_{jl} E_{il}.$$

To satisfy $CE_{ij} = E_{ij}C$ it must thus hold that $c_{ii} = c_{jj}$ and that $c_{ki} = 0$ and $c_{lj} = 0$ for all $k, l \neq i, j$. Since this must hold for all $i, j \in \{1, ..., n\}$, it follows that $C = \lambda I$ for a $\lambda \in \mathbb{C}$. □

From Theorem 3.4 it follows that C is fixed up to a scalar once the image of a generating set of Pauli matrices under conjugation by C is known.

3.2. BINARY REPRESENTATION

As an alternative to the notation used above, Pauli strings and Clifford operations can be described efficiently in terms of binary matrices. A disadvantage of this binary representation is that we lose the information about any multiplicative factor in front of a Pauli string. However, as we will discuss in more detail in Chapter 5, this information is not important in the context of entanglement distillation protocols. In this section it is shown how Pauli strings and Clifford operations can be represented as binary matrices. This representation is adjusted from Dehaene et al., 2003b.

Firstly, the Pauli matrices defined in equation (3.1) can be represented by binary vectors. Let

$$\tau_{00} = I, \quad \tau_{10} = X, \quad \tau_{11} = iY, \quad \tau_{01} = Z. \tag{3.10}$$

Neglecting the multiplicative factor $\pm 1, \pm i$, a Pauli string can be denoted by

$$\tau_a = \tau_{v_1 w_1} \otimes \cdots \otimes \tau_{v_n w_n}, \quad a = \begin{bmatrix} v \\ w \end{bmatrix}, \ v, w \in \mathbb{Z}_2^n. \tag{3.11}$$

For example, the Pauli string $I \otimes X \otimes Y \otimes Z$ is equal to $\tau_{00} \otimes \tau_{10} \otimes \tau_{11} \otimes \tau_{01}$ up to a factor i. For this Pauli string, we thus have

$$v = \begin{bmatrix} 0 & 1 & 1 & 0 \end{bmatrix}^T, \qquad w = \begin{bmatrix} 0 & 0 & 1 & 1 \end{bmatrix}^T.$$

So, in the binary notation, the Pauli string $I \otimes X \otimes Y \otimes Z$ is represented by the vector

$$a = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}^T.$$

Recall that the Pauli strings form a group under matrix multiplication. What does this matrix multiplication look like in the binary representation? To determine this, suppose that we have two Pauli strings $\tau_{a_1}, \tau_{a_2} \in \mathbb{Z}_2^{2n}$, with $a_1 = \begin{bmatrix} v_1 \\ w_1 \end{bmatrix} = \begin{bmatrix} v_{11} \dots v_{1n} \ w_{11} \dots w_{1n} \end{bmatrix}^T$ and $a_2 = \begin{bmatrix} v_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} v_{21} \dots v_{2n} \ w_{21} \dots w_{2n} \end{bmatrix}^T$. Then

$$\begin{aligned} \tau_{a_1} \tau_{a_2} &= (\tau_{v_{11} w_{11}} \otimes \cdots \otimes \tau_{v_{1n} w_{1n}})(\tau_{v_{21} w_{21}} \otimes \cdots \otimes \tau_{v_{2n} w_{2n}}) \\ &= \bigotimes_{k=1}^{n} \tau_{v_{1k} w_{1k}} \tau_{v_{2k} w_{2k}}. \end{aligned} \tag{3.12}$$

For all $k \in \{1, \dots, n\}$ we have

$$\begin{aligned} \tau_{v_{1k} w_{1k}} \tau_{v_{2k} w_{2k}} &= X^{v_{1k}} Z^{w_{1k}} X^{v_{2k}} Z^{w_{2k}} \\ &= X^{v_{1k}} (-1)^{v_{2k} w_{1k}} X^{v_{2k}} Z^{w_{1k}} Z^{w_{2k}} \\ &= (-1)^{v_{2k} w_{1k}} X^{v_{1k} + v_{2k}} Z^{w_{1k} + w_{2k}} \\ &= (-1)^{v_{2k} w_{1k}} \tau_{v_{1k} + v_{2k}, w_{1k} + w_{2k}}. \end{aligned} \tag{3.13}$$

As a result,

$$\begin{aligned}\tau_{a_1}\tau_{a_2} &= \bigotimes_{k=1}^{n}(-1)^{v_{2k}w_{1k}}\tau_{v_{1k}+v_{2k},w_{1k}+w_{2k}}\\ &= (-1)^{\sum_{k=1}^{n}v_{2k}w_{1k}}\,\tau_{v_1+v_2,w_1+w_2}\\ &= (-1)^{v_2\cdot w_1}\tau_{v_1+v_2,w_1+w_2}.\end{aligned} \tag{3.14}$$

Here $v_2 \cdot w_1$ is the standard vector dot product. We can rewrite this dot product in terms of the vectors a_1 and a_2:

$$v_2 \cdot w_1 = a_2^T \Xi a_1, \qquad \Xi = \begin{bmatrix} 0 & I_n \\ 0 & 0 \end{bmatrix}. \tag{3.15}$$

Hence, the product of two Pauli strings is given by

$$\tau_{a_1}\tau_{a_2} = (-1)^{a_2^T\Xi a_1}\tau_{a_1+a_2}. \tag{3.16}$$

In the binary representation, the matrix product of Pauli strings can thus be calculated by adding the corresponding binary vectors, where the addition is performed modulo 2. This completes the binary representation of Pauli strings.

Next, let us look at the Clifford operations. Let $C \in \mathcal{C}_n$ be a Clifford operation and $\sigma : \mathcal{P}_n \to \mathcal{P}_n$, $\sigma(p) = CpC^\dagger$ be the corresponding automorphism. We are looking for a representation $\pi : \mathbb{Z}_2^{2n} \to \mathbb{Z}_2^{2n}$ of this automorphism in the binary picture. Let $a, b \in \mathbb{Z}_2^{2n}$. Then we know that $C\tau_{a+b}C^\dagger = (-1)^{b^T\Xi a}C\tau_a\tau_b C^\dagger = (-1)^{b^T\Xi a}C\tau_a C^\dagger C\tau_b C^\dagger$. In the binary representation, the prefactor $(-1)^{b^T\Xi a}$ does not make a difference. Thus, $\pi(a+b) = \pi(a)+\pi(b)$, so π is a linear map. Because the vector space over $\mathbb{Z}_2^{2n}$ has finite dimension, it follows that there is a (binary) $2n \times 2n$-matrix M such that $\pi(a) = Ma$ for all $a \in \mathbb{Z}_2^{2n}$.

Furthermore, by Theorem 3.2 we know that σ preserves the commutation relations between Pauli strings. Thus, π must preserve the commutation relations too. By equation (3.16) we have $\tau_a\tau_b = (-1)^{b^T\Xi a}\tau_{a+b}$ and $\tau_b\tau_a = (-1)^{a^T\Xi b}\tau_{a+b}$, so that

$$\tau_a\tau_b = (-1)^{b^T\Xi a + a^T\Xi b}\tau_b\tau_a. \tag{3.17}$$

Note that $a^T\Xi b = (a^T\Xi b)^T = b^T\Xi^T a$. Thus

$$\tau_a\tau_b = (-1)^{b^T\Xi a + b^T\Xi^T a}\tau_b\tau_a = (-1)^{b^T\Omega a}\tau_b\tau_a, \quad \Omega = \Xi + \Xi^T = \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix}. \tag{3.18}$$

Similarly, it can be derived that

$$\tau_{Ma}\tau_{Mb} = (-1)^{(Mb)^T\Omega Ma}\tau_{Mb}\tau_{Ma}. \tag{3.19}$$

In order to preserve the (anti-)commutation relations, it must hold that $b^T M^T \Omega M a = b^T \Omega a$ for all $a, b \in \mathbb{Z}_2^{2n}$, so $M^T\Omega M = \Omega$. The matrices M that satisfy this condition form a group known as the symplectic group over $\mathbb{Z}_2$. The symplectic group can be defined over an arbitrary ring R.

Definition 3.5. Let $n \in \mathbb{N}$, let R be a ring and $\Omega = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$. The *symplectic group* is defined as the set $Sp(2n,R) = \{M \in M_{2n\times 2n}(R) : M^T \Omega M = \Omega\}$ under matrix multiplication.

The elements of a symplectic group are referred to as symplectic matrices. The following proposition gives a useful characterization of the symplectic matrices.

Proposition 3.6. *Let $n \in \mathbb{N}$ and let R be a ring. If we write $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ for $M \in Sp(2n,R)$ with $A,B,C,D \in M_{n\times n}(R)$, then A,B,C and D must obey the following relations:*

$$\begin{aligned} B^T D - D^T B &= 0, \\ A^T C - C^T A &= 0, \\ A^T D - C^T B &= I_n, \\ B^T C - D^T A &= -I_n. \end{aligned}$$

Proof. The proposition follows immediately from substitution of $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ into the equation $M^T \Omega M = \Omega$ and working out the matrix multiplications. □

Note that, if M is symplectic, then M is invertible. Indeed, from $M^T \Omega M = \Omega$ it follows that $\det(\Omega) = \det(M^T)\det(\Omega)\det(M)$. Because $\det(\Omega) = 1$, we have $\det(M)^2 = 1$, so $\det(M) \neq 0$ and M is invertible. Proposition 3.7 provides a way to calculate the inverse of a symplectic matrix.

Proposition 3.7. *Let $n \in \mathbb{N}$ and let R be a ring. Let $M \in Sp(2n,R)$, $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$. Then*

$$M^{-1} = \begin{bmatrix} D^T & -B^T \\ -C^T & A^T \end{bmatrix}.$$

Proof. From $M^T \Omega M = \Omega$ it follows that

$$M^{-1} = \Omega^{-1} M^T \Omega = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix} \begin{bmatrix} A^T & C^T \\ B^T & D^T \end{bmatrix} \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} = \begin{bmatrix} D^T & -B^T \\ -C^T & A^T \end{bmatrix}.$$

□

Note that it follows from Proposition 3.7 that a symplectic group can indeed by defined over any ring and not only over fields, as one might expect in the first place, since the entries of the matrices do not necessarily need to have a multiplicative inverse in R.

As mentioned earlier, Clifford operations can be represented as binary matrices, so $R = \mathbb{Z}_2$. The order of the symplectic group over $\mathbb{Z}_2$ is well-known to be equal to

$$|Sp(2n, \mathbb{Z}_2)| = 2^{n^2} \prod_{j=1}^{n} (4^j - 1). \tag{3.20}$$

A proof of this formula can be found in Artin, 1957.

Let $\phi : \mathcal{C}_n \to Sp(2n, \mathbb{Z}_2)$ denote the function that maps $C \in \mathcal{C}_n$ to the symplectic matrix M that satisfies $\tau_{Ma} = C\tau_a C^\dagger$ for all $a \in \mathbb{Z}_2^{2n}$. In Theorem 3.9 it is proved that ϕ is a homomorphism, but first it is shown in Lemma 3.8 that ϕ is well-defined.

Lemma 3.8. *The function $\phi : \mathcal{C}_n \to Sp(2n, \mathbb{Z}_2)$ that maps $C \in \mathcal{C}_n$ to the symplectic matrix M that satisfies $\tau_{Ma} = C\tau_a C^\dagger$ for all $a \in \mathbb{Z}_2^{2n}$ is well-defined.*

Proof. Let $C_1, C_2 \in \mathcal{C}_n$, with $\phi(C_1) = M_1$ and $\phi(C_2) = M_2$. Suppose that $C_1 = C_2$. Then $\tau_{M_1 a} = C_1\tau_a C_1^\dagger = C_2\tau_a C_2^\dagger = \tau_{M_2 a}$ for all $a \in \mathbb{Z}_2^{2n}$. Using equations (3.10) and (3.11) it follows that $M_1 a = M_2 a$ for all $a \in \mathbb{Z}_2^{2n}$. In particular, it holds that $M_1 e_i = M_2 e_i$ for all standard basis vectors $e_i \in \mathbb{Z}_2^{2n}$. Because $M_1 e_i$ and $M_2 e_i$ are equal to the ith column of the matrices M_1 and M_2, respectively, and the equality must hold for all $i \in \{1, ..., n\}$, it follows that $M_1 = M_2$. So indeed, ϕ is well-defined. □

Theorem 3.9. *The function $\phi : \mathcal{C}_n \to Sp(2n, \mathbb{Z}_2)$ that maps $C \in \mathcal{C}_n$ to the symplectic matrix M that satisfies $\tau_{Ma} = C\tau_a C^\dagger$ for all $a \in \mathbb{Z}_2^{2n}$ is a homomorphism.*

Proof. Let $C_1, C_2 \in \mathcal{C}_n$ with $\phi(C_1) = M_1$, $\phi(C_2) = M_2$. Because $C_1 C_2 \tau_a C_2^\dagger C_1^\dagger \sim C_1 \tau_{M_2 a} C_1^\dagger \sim \tau_{M_1 M_2 a}$, where $\sim$ indicates equality up a phase factor, it follows that $\phi(C_1 C_2) = M_1 M_2 = \phi(C_1)\phi(C_2)$. So, ϕ is indeed a homomorphism. □

3.2.1. IMPORTANT CLIFFORD OPERATIONS IN THEIR SYMPLECTIC FORM

As proved in Section 3.1.2, an element $C \in \mathcal{C}_n$ is fixed, up to an overall phase factor, once we know the image of a generating set of $\mathcal{P}_n$. This notion is used in this section to compute the binary representation of the Clifford operations H, S, CNOT and CZ. Moreover, two results to represent combinations of those Clifford operations in the binary picture are presented.

We start off by determining the images of the generators H, S, CNOT and CZ under ϕ. As generating set we use the union of the Pauli strings $\tilde{X}_i = I_1 \otimes \cdots \otimes I_{i-1} \otimes X_i \otimes I_{i+1} \otimes \cdots \otimes I_n$, with one X matrix on the ith position and identity matrices on every other position, and similar Pauli strings $\tilde{Z}_i = I_1 \otimes \cdots \otimes I_{i-1} \otimes Z_i \otimes I_{i+1} \otimes \cdots \otimes I_n$ with one Z matrix. Thus as generating set, the following union is used:

$$\{\tilde{X}_i : i \in \{1, ..., n\}\} \cup \{\tilde{Z}_i : i \in \{1, ..., n\}\}. \tag{3.21}$$

From equation (3.11) it follows that the string $\tilde{X}_i$ is represented by the standard basis vector $e_i \in \mathbb{Z}_2^{2n}$ and that $\tilde{Z}_i$ is represented by e_{n+i}. Note that since we are interested in the binary representation and we cannot represent an overall phase there, it is indeed enough to consider the generators from equation (3.21).

Firstly, let us consider the Hadamard gate. This is a single qubit gate, so $\phi(H)$ is a 2×2 matrix and we only need to consider X and Z. Using equation (2.14) it can be easily verified that $HXH^\dagger = Z$ and $HZH^\dagger = X$. In the binary representation, this boils down to

$$\phi(H)\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix}, \quad \phi(H)\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix}. \tag{3.22}$$

From equation (3.22) it can be concluded that

$$\phi(H) = \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}. \tag{3.23}$$

For the phase gate, which again is a single qubit gate, it can be verified using equation (2.15) that $SXS^\dagger = Y$ and $SZS^\dagger = Z$. Thus we have

$$\phi(S)\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}1\\1\end{bmatrix}, \quad \phi(S)\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix}, \tag{3.24}$$

from which it follows that

$$\phi(S) = \begin{bmatrix}1 & 0\\1 & 1\end{bmatrix}. \tag{3.25}$$

The CNOT gate is a two qubit gate, thus $\phi(\text{CNOT})$ is a 4×4 matrix. Using equation (2.16) the following relations can be verified:

$$\begin{aligned}
&\text{CNOT}_{12}(X\otimes I)\text{CNOT}_{12}{}^\dagger = (X\otimes X),\\
&\text{CNOT}_{12}(I\otimes X)\text{CNOT}_{12}{}^\dagger = (I\otimes X),\\
&\text{CNOT}_{12}(Z\otimes I)\text{CNOT}_{12}{}^\dagger = (Z\otimes I),\\
&\text{CNOT}_{12}(I\otimes Z)\text{CNOT}_{12}{}^\dagger = (Z\otimes Z).
\end{aligned} \tag{3.26}$$

Thus, $\phi(\text{CNOT})$ must satisfy

$$\begin{aligned}
&\phi(\text{CNOT}_{12})\begin{bmatrix}1\\0\\0\\0\end{bmatrix} = \begin{bmatrix}1\\1\\0\\0\end{bmatrix}, \quad \phi(\text{CNOT}_{12})\begin{bmatrix}0\\1\\0\\0\end{bmatrix} = \begin{bmatrix}0\\1\\0\\0\end{bmatrix},\\
&\phi(\text{CNOT}_{12})\begin{bmatrix}0\\0\\1\\0\end{bmatrix} = \begin{bmatrix}0\\0\\1\\0\end{bmatrix}, \quad \phi(\text{CNOT}_{12})\begin{bmatrix}0\\0\\0\\1\end{bmatrix} = \begin{bmatrix}0\\0\\1\\1\end{bmatrix}.
\end{aligned} \tag{3.27}$$

From equation (3.27) it follows that

$$\phi(\text{CNOT}_{12}) = \begin{bmatrix}1 & 0 & 0 & 0\\1 & 1 & 0 & 0\\0 & 0 & 1 & 1\\0 & 0 & 0 & 1\end{bmatrix}. \tag{3.28}$$

Similary, it can be derived that

$$\phi(\mathrm{CNOT}_{21}) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}. \tag{3.29}$$

Finally, for the CZ gate it can be derived from equation (2.17) that the following relations hold.

$$\begin{aligned} \mathrm{CZ}_{12}(X \otimes I)\mathrm{CZ}_{12}{}^{\dagger} &= (X \otimes Z), \\ \mathrm{CZ}_{12}(I \otimes X)\mathrm{CZ}_{12}{}^{\dagger} &= (Z \otimes X), \\ \mathrm{CZ}_{12}(Z \otimes I)\mathrm{CZ}_{12}{}^{\dagger} &= (Z \otimes I), \\ \mathrm{CZ}_{12}(I \otimes Z)\mathrm{CZ}_{12}{}^{\dagger} &= (I \otimes Z). \end{aligned} \tag{3.30}$$

In a similar way as for the CNOT gate it can be derived that

$$\phi(\mathrm{CZ}_{12}) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}. \tag{3.31}$$

Using equation (3.11) the binary representation of the Kronecker product can be derived.

Proposition 3.10. *(Hostens, 2007) Let $C_1, C_2 \in \mathcal{C}_n$ such that*

$$\phi(C_1) = \begin{bmatrix} M_{11} & M_{21} \\ M_{12} & M_{22} \end{bmatrix}$$

$$\phi(C_2) = \begin{bmatrix} \tilde{M}_{11} & \tilde{M}_{21} \\ \tilde{M}_{12} & \tilde{M}_{22} \end{bmatrix}$$

Then the Kronecker product $C_1 \otimes C_2$ is represented by

$$\phi(C_1 \otimes C_2) = \begin{bmatrix} M_{11} & 0 & M_{21} & 0 \\ 0 & \tilde{M}_{11} & 0 & \tilde{M}_{21} \\ M_{12} & 0 & M_{22} & 0 \\ 0 & \tilde{M}_{12} & 0 & \tilde{M}_{22} \end{bmatrix}$$

Note that if C is the identity Clifford operation, then $\phi(C)$ is the identity matrix. Thus the following statement is an immediate consequence of Proposition 3.10.

Corollary 3.11. *(Hostens, 2007) A Clifford operation $C \in \mathcal{C}_n$ with binary representation $\phi(C)$ that acts on a subset of the qubits $S \subseteq \{1, \ldots, n\}$ is represented by the $2n \times 2n$ identity matrix, with the rows and columns with indices in $S \cup (S + n)$ replaced by the rows and columns of $\phi(C)$.*

3.2.2. CHARACTERIZATION OF THE HOMOMORPHISM ϕ

In the previous section it was explained how the image of an element $C \in \mathcal{C}_n$ under the homomorphism $\phi : \mathcal{C}_n \to Sp(2n, \mathbb{Z}_2)$ can be determined. In this section, these images are used to obtain more information about the homomorphism ϕ.

Suppose that $M \in Sp(2n, \mathbb{Z}_2)$. Using Corollary 3.11 and equations (3.25), (3.28) and (3.29) the action of left multiplication by $\phi(H)$, $\phi(S)$ and $\phi(\text{CNOT})$ gates on M can be derived:

- A Hadamard gate on qubit i results in swapping row i and $n+i$.
- A phase gate on qubit i results in the addition of row i to row $n+i$.
- A CNOT gate from qubit i to qubit j results in the addition of row i to row j and of row $n+j$ to row $n+i$.

In Theorem 3.12 these notions are used to prove that the homomorphism that maps the Clifford group $\mathcal{C}_n$ to the symplectic group $Sp(2n, \mathbb{Z}_2)$ is surjective. Although used to prove a slightly different statement, the structure of this proof is based on Hostens, 2007.

Theorem 3.12. *The homomorphism $\phi : \mathcal{C}_n \to Sp(2n, \mathbb{Z}_2)$ that maps $C \in \mathcal{C}_n$ to the symplectic matrix M that satisfies $\tau_{Ma} = C\tau C^\dagger$ for all $a \in \mathbb{Z}_2^{2n}$ is surjective.*

Proof. We know that $\mathcal{C}_n$ is generated by the set of H and S gates on all qubits and CNOT gates on every pair of qubits. Thus $\phi[\mathcal{C}_n]$ is generated by the images of H, S and CNOT under ϕ. We prove that any $M \in Sp(2n, \mathbb{Z}_2)$ can be written in terms of these generators by showing that M can be transformed to the identity through left multiplication by (a part of) the generators. M is then equal to the product of the inverses of these generators in reverse order, so $M \in \phi[\mathcal{C}_n]$.

Let $M \in Sp(2n, \mathbb{Z}_2)$. We first look at the first column M_1 of M. We take the following steps, where we keep on referring to intermediate stages as M.

1. Make sure that $M_{11} = 1$. Suppose that $M_{11} = 0$. If there is a qubit $i > 1$ such that $M_{i1} = 1$, then we apply a CNOT gate from qubit i to qubit 1. If the upper half of M_1 contains only zeros, then we first apply a Hadamard gate to a qubit i with $M_{(n+i)1} = 1$ and then apply the CNOT gate from i to 1. Note that there must always be at least one entry of M_1 which is equal to 1, otherwise M is not invertible.
2. For all i with $M_{i1} = 1$, apply a CNOT gate from qubit 1 to qubit i. This ensures that all entries M_{i1} with $i \in \{2, \ldots, n\}$ are equal to zero.
3. If $M_{(n+1)1} = 1$, apply a phase gate to qubit 1. As a result, $M_{(n+1)1} = 0$.
4. Apply a Hadamard gate to qubit 1.
5. For all i with $M_{(n+i)1} = 1$, apply a CNOT gate from qubit i to qubit 1.

Steps 1 to 5 are schematically summarized below.

$$M_1 \xrightarrow{1} \left[\begin{array}{c}1\\\cdot\\\cdot\\\hline\cdot\\\cdot\\\cdot\end{array}\right] \xrightarrow{2} \left[\begin{array}{c}1\\0\\0\\\hline\cdot\\\cdot\\\cdot\end{array}\right] \xrightarrow{3} \left[\begin{array}{c}1\\0\\0\\\hline0\\\cdot\\\cdot\end{array}\right] \xrightarrow{4} \left[\begin{array}{c}0\\0\\0\\\hline1\\\cdot\\\cdot\end{array}\right] \xrightarrow{5} \left[\begin{array}{c}0\\0\\0\\\hline1\\0\\0\end{array}\right]$$

After steps 1 to 5, M_1 is thus transformed to the standard basis vector e_{n+1}. Since M is a symplectic matrix and it is multiplied by symplectic matrices only, the transformed matrix is still symplectic. As a consequence, it must satisfy $M_1^T \Omega M_{n+1} = 1$. Because $M_1 = e_{n+1}$, it follows that $M_{1(n+1)} = 1$. We proceed with the transformation as follows:

6. For all i with $M_{i(n+1)} = 1$, apply a CNOT gate from qubit 1 to qubit i. This results in all entries $M_{i(n+1)}$ with $i \in \{2, ..., n\}$ being equal to zero, without changing M_1.
7. If $M_{(n+1)(n+1)} = 1$, apply a phase gate to qubit 1. This ensures that $M_{(n+1)(n+1)} = 0$.
8. Apply a Hadamard gate to qubit 1 to switch rows 1 and $n+1$.
9. For all i with $M_{(n+i)(n+1)} = 1$, apply a CNOT gate from qubit i to qubit 1.

Steps 6 to 9 are schematically depicted below.

$$M_1 = \left[\begin{array}{c}0\\0\\0\\\hline1\\0\\0\end{array}\right] \xrightarrow{6} \left[\begin{array}{c}0\\0\\0\\\hline1\\0\\0\end{array}\right] \xrightarrow{7} \left[\begin{array}{c}0\\0\\0\\\hline1\\0\\0\end{array}\right] \xrightarrow{8} \left[\begin{array}{c}1\\0\\0\\\hline0\\0\\0\end{array}\right] \xrightarrow{9} \left[\begin{array}{c}1\\0\\0\\\hline0\\0\\0\end{array}\right]$$

$$M_{n+1} = \left[\begin{array}{c}1\\\cdot\\\cdot\\\hline\cdot\\\cdot\\\cdot\end{array}\right] \xrightarrow{6} \left[\begin{array}{c}1\\0\\0\\\hline\cdot\\\cdot\\\cdot\end{array}\right] \xrightarrow{7} \left[\begin{array}{c}1\\0\\0\\\hline0\\\cdot\\\cdot\end{array}\right] \xrightarrow{8} \left[\begin{array}{c}0\\0\\0\\\hline1\\\cdot\\\cdot\end{array}\right] \xrightarrow{9} \left[\begin{array}{c}0\\0\\0\\\hline1\\0\\0\end{array}\right]$$

The columns M_1 and M_{n+1} now equal the corresponding columns of the $2n \times 2n$ identity matrix. Using Proposition 3.6 it can be derived that also the rows 1 and $n+1$ of M are

equal to the corresponding rows of the identity matrix. That is,

$$M = \left[\begin{array}{c|ccc|c|ccc} 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \hline 0 & & & & 0 & & & \\ \vdots & & M_A & & \vdots & & M_B & \\ 0 & & & & 0 & & & \\ \hline 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \hline 0 & & & & 0 & & & \\ \vdots & & M_C & & \vdots & & M_D & \\ 0 & & & & 0 & & & \end{array}\right].$$

Let

$$M' = \begin{bmatrix} M_A & M_B \\ M_C & M_D \end{bmatrix},$$

then we can repeat the procedure described above for M'. Finally, this results in the identity matrix. So indeed, $M \in \phi[\mathcal{C}_n]$ and $\phi[\mathcal{C}_n] = Sp(2n, \mathbb{Z}_2)$. □

Thus, the homomorphism $\phi : \mathcal{C}_n \to Sp(2n, \mathbb{Z}_2)$ is surjective. From the first isomorphism theorem it now follows that $\mathcal{C}_n / \ker(\phi) \cong Sp(2n, \mathbb{Z}_2)$. This isomorphism is used in Theorem 3.13 to characterize the kernel of ϕ.

Theorem 3.13. *Let $\phi : \mathcal{C}_n \to Sp(2n, \mathbb{Z}_2)$ be the homomorphism that maps $C \in \mathcal{C}_n$ to the symplectic matrix M that satisfies $\tau_{Ma} = C\tau C^\dagger$ for all $a \in \mathbb{Z}_2^{2n}$. Then $\ker(\phi) = \langle \mathcal{P}_n, \eta I \rangle$.*

Proof. Firstly, we show that $\langle \mathcal{P}_n, \eta I \rangle \subseteq \ker(\phi)$. Note that $C \in \mathcal{C}_n$ is contained in $\ker(\phi)$ if and only if $CPC^\dagger \sim P$ for all $P \in \mathcal{P}_n$, where $\sim$ denotes equality up to an overall phase.

Recall from Section 3.1.1 that the Pauli group on n qubits consists of all $2^n \times 2^n$ matrices of the form $\lambda P_1 \otimes \cdots \otimes P_n$ with $\lambda \in \{\pm 1, \pm i\}$ and $P_i \in \{I, X, Y, Z\}$. It can easily be verified that $\langle \mathcal{P}_n, \eta I \rangle$ consists of all $2^n \times 2^n$ of the form $\kappa P_1 \otimes \cdots \otimes P_n$ with $\kappa \in \{\pm 1, \pm i, \pm \eta, \pm i\eta\}$ and $P_i \in \{I, X, Y, Z\}$.

Let $C \in \langle \mathcal{P}_n, \eta I \rangle$, then $C = \kappa \tilde{C}$ with $\kappa \in \{\pm 1, \pm i, \pm \eta, \pm i\eta\}$ and $\tilde{C} = P_1 \otimes ... \otimes P_n$ with $P_i \in \{I, X, Y, Z\}$. It follows that, for all $P \in \mathcal{P}_n$,

$$\kappa \tilde{C} P (\kappa \tilde{C})^\dagger = \kappa \kappa^\dagger \tilde{C} P \tilde{C}^\dagger = \tilde{C} P \tilde{C}^\dagger.$$

Moreover, by equation (3.7) it follows that

$$\tilde{C} P = \pm P \tilde{C},$$

so that

$$\tilde{C} P \tilde{C}^\dagger = \pm P. \tag{3.32}$$

Thus indeed, $C = \kappa \tilde{C} \in \ker(\phi)$. It follows that $\langle \mathcal{P}_n, \eta I \rangle \subset \ker(\phi)$.

To prove the equality of $\langle \mathcal{P}_n, \eta I \rangle$ and $\ker(\phi)$, recall that $|\mathcal{C}_n| = 2^{n^2+2n+3} \prod_{j=1}^{n} (4^j - 1)$ and that $|Sp(2n, \mathbb{Z}_2)| = 2^{n^2} \prod_{j=1}^{n} (4^j - 1)$. Because $\mathcal{C}_n / \ker(\phi) \cong Sp(2n, \mathbb{Z}_2)$ it follows that $|\ker(\phi)| = \frac{|\mathcal{C}_n|}{|Sp(2n,\mathbb{Z}_2)|} = 2^{2n+3}$.

By counting the elements of the form $\kappa P_1 \otimes \cdots \otimes P_n$ with $\kappa \in \{\pm 1, \pm i, \pm \eta, \pm i\eta\}$ and $P_i \in \{I, X, Y, Z\}$ for $i = 1, ..., n$, it follows that the order of $\langle \mathcal{P}_n, \eta I \rangle$ is equal to $8 \cdot 4^n = 2^{2n+3}$, which is the same as the order of the kernel. Thus indeed, $\ker(\phi) = \langle \mathcal{P}_n, \eta I \rangle$. □

From Theorem 3.12 and Theorem 3.13 it can be concluded that $\mathcal{C}_n / \langle \mathcal{P}_n, \eta I \rangle \cong Sp(2n, \mathbb{Z}_2)$. This is the group of operations that is considered in the distillation protocols in this thesis, as will be explained in more detail in Section 5.3.

Finally, observe that the steps in the proof of Theorem 3.12 can be used to find a circuit of Clifford gates that corresponds to a given symplectic matrix. These steps will be used in Chapter 7 to find Clifford operations that correspond to optimal symplectic matrices.

4

ENTANGLEMENT DISTILLATION

This chapter describes the framework for entanglement distillation. The system that is considered, is a bipartite quantum system that is shared by two parties. These parties are referred to as *Alice* and *Bob* or A and B.

In Section 4.1 the basic characterizations of bipartite entanglement that are relevant in the context of distillation protocols are given. These include the concepts of entanglement and fidelity, based on the definitions provided by Nielsen and Chuang, 2016. In this section the Bell states are introduced and it is argued why we focus on Bell diagonal states in the remainder of this thesis. In Section 4.2 the general structure of entanglement distillation protocols is explained. This explanation is illustrated in Section 4.3 by the description of a protocol known as the DEJMPS protocol, which was published by Deutsch et al., 1996.

4.1. CHARACTERIZATION OF BIPARTITE ENTANGLEMENT

In this section a characterization of bipartite entanglement is given. In Section 4.1.1 an introduction to the concept of entanglement is given. After introducing the Bell states as maximally entangled states, the concept of fidelity is introduced. In this thesis, fidelity of an arbitrary state and a Bell state is used to quantify the entanglement of the arbitrary state. In Section 4.1.2, a correspondence between the Bell states and the Pauli matrices is introduced. This correspondence gives rise to an alternative description of our system, which will be used in the remainder of this thesis.

4.1.1. INTRODUCTION TO QUANTUM ENTANGLEMENT

A central topic in this thesis is the topic of entanglement. Quantum entanglement is the phenomenon that two or more systems are connected in such a way that they cannot be described independently from each other. For instance, if we have two entangled photons and measure the polarization of one photon, then immediately the polarization of the second photon is known, even if the photons are physically separated. Albert

Einstein called this phenomenon 'spooky action at a distance'. He believed that this apparent non-local behaviour could be explained by classical correlations between local hidden variables (Einstein et al., 1935). These local hidden variables are unobservable hypothetical variables that satisfy the concept of local realism, which is the principle that a system is only influenced by its immediate surroundings. This principle implies that an event at one point cannot simultaneously cause a result at another point, since information cannot travel faster than the speed of light.

However, in Bell, 1964 an inequality known as Bell's inequality was presented. It was shown that this inequality cannot be violated if there are local hidden variables in the system. This inequality and variations of it were experimentally tested and violated many times, for instance by Aspect et al., 1982, hereby contradicting Einstein.

Entanglement can be seen as 'the amount of non-locality' in a quantum system consisting of multiple subsystems. It is usually defined as the negation of separability.

Definition 4.1. A state ρ_{AB} is said to be *separable* with respect to the parties A and B if it can be written as

$$\rho_{AB} = \sum_i p_i \rho_A^{(i)} \otimes \rho_B^{(i)},$$

where $\rho_A^{(i)}$ and $\rho_B^{(i)}$ are the density matrices for the systems A and B and $\{p_i\}$ is a probability distribution. If a state is not separable, it is said to be *entangled*.

If the parties are only allowed to use *local operations and classical communication* (LOCC), it is not possible to create an entangled state from a separable state. Moreover, the amount of entanglement does not increase under LOCC. This gives rise to a certain ordering of entanglement: one state contains at least as much entanglement as another state if it can be transformed into the other state under LOCC (Plenio and Virmani, 2007).

The entanglement of a pure state can be quantified by its entropy of entanglement.

Definition 4.2. The *entropy of entanglement* of a pure state $|\psi\rangle$ is defined as

$$E(|\psi\rangle) = -\operatorname{Tr}(\rho_A \log_2 \rho_A) = -\operatorname{Tr}(\rho_B \log_2 \rho_B).$$

Here ρ_A and ρ_B are reduced density operators (see Definition 2.10). Because ρ_A is a density matrix, it is Hermitian and thus diagonalizable. Hence, we can write $\rho_A = VAV^{-1}$, where V is a matrix that consists of the eigenvectors of ρ_A and $A = V^{-1}\rho_A V$ is a diagonal matrix. Because A is a diagonal matrix, $\log_2 A$ can be obtained by replacing each diagonal element of A by the $\log_2$ of this element. The quantity $\log_2 \rho_A$ is then equal to

$$\log_2 \rho_A = V \log_2 AV^{-1}. \tag{4.1}$$

Of course, $\log_2 \rho_B$ can be calculated in a similar manner. The quantity $S(\rho) = -\operatorname{Tr}(\rho \log_2 \rho)$ is known as the *Von Neumann entropy*. The entropy of entanglement ranges from zero for separable states to one for so-called maximally entangled states (Bennett et al., 1996a).

Well-known examples of such maximally entangled states for a bipartite system are the *Bell states:*

$$\begin{aligned}
\left|\Phi^+\right\rangle &= \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right), \\
\left|\Psi^+\right\rangle &= \frac{1}{\sqrt{2}}\left(|01\rangle + |10\rangle\right), \\
|\Psi^-\rangle &= \frac{1}{\sqrt{2}}\left(|01\rangle - |10\rangle\right), \\
|\Phi^-\rangle &= \frac{1}{\sqrt{2}}\left(|00\rangle - |11\rangle\right).
\end{aligned} \tag{4.2}$$

To illustrate the computation of the entropy of entanglement and show that the Bell states are indeed maximally entangled, $E(\left|\Phi^+\right\rangle)$ is calculated below. From Definition 2.3 it follows that the density matrix corresponding to $\left|\Phi^+\right\rangle$ is equal to

$$\rho = \left|\Phi^+\right\rangle\left\langle\Phi^+\right| = \frac{1}{2}\left(|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|\right).$$

The reduced density operator ρ_A is thus equal to

$$\begin{aligned}
\rho_A &= \frac{1}{2}\left(|0\rangle\langle 0|\,\mathrm{Tr}_B(|0\rangle\langle 0|) + |0\rangle\langle 1|\,\mathrm{Tr}_B(|0\rangle\langle 1|) + |1\rangle\langle 0|\,\mathrm{Tr}_B(|1\rangle\langle 0|) + |1\rangle\langle 1|\,\mathrm{Tr}_B(|1\rangle\langle 1|)\right) \\
&= \frac{1}{2}\left(|0\rangle\langle 0|\langle 0|0\rangle + |0\rangle\langle 1|\langle 1|0\rangle + |1\rangle\langle 0|\langle 0|1\rangle + |1\rangle\langle 1|\langle 1|1\rangle\right) \\
&= \frac{1}{2}\left(|0\rangle\langle 0| + |1\rangle\langle 1|\right) \\
&= \frac{1}{2}I_2.
\end{aligned}$$

Because ρ_A is a diagonal matrix, it follows that $\log_2\rho_A$ can be obtained by taking the logarithm of the elements on the diagonal. Thus

$$E\left(\left|\Phi^+\right\rangle\right) = -\mathrm{Tr}\left(\frac{1}{2}I_2\log_2\left(\frac{1}{2}I_2\right)\right) = -\mathrm{Tr}\left(\begin{bmatrix}1/2 & 0\\ 0 & 1/2\end{bmatrix}\begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix}\right) = -\left(-\frac{1}{2}-\frac{1}{2}\right) = 1.$$

So indeed, we see that $\left|\Phi^+\right\rangle$ is maximally entangled. Similar calculations can be performed for the other three Bell states.

Any other two-qubit state, including the Bell states themselves, can be created from a Bell-state under LOCC. It therefore is a legitimate choice to use the Bell states as standard unit of entanglement for a bipartite system. The measure of entanglement that will be used in this thesis is the fidelity to one of the Bell states. The fidelity of two arbitrary states is defined as follows.

Definition 4.3. Let σ and ρ be two density operators. Then the *fidelity* is defined as

$$F(\sigma,\rho) = \mathrm{Tr}\left(\sqrt{\sigma^{1/2}\rho\sigma^{1/2}}\right)^2.$$

The fidelity is a measure for how close two arbitrary quantum states are. It yields a number between 0 and 1, where $F(\sigma,\rho)=1$ corresponds to the situation where $\sigma=\rho$. Note that for a pure state $\sigma=|\Psi\rangle\langle\Psi|$ we have $\sigma^{1/2}=\sigma$. Thus

$$F(|\Psi\rangle\langle\Psi|,\rho)=\mathrm{Tr}\left(\sqrt{|\Psi\rangle\langle\Psi|\rho|\Psi\rangle\langle\Psi|}\right)^2=\langle\Psi|\rho|\Psi\rangle\,\mathrm{Tr}\left(\sqrt{\langle\Psi|\Psi\rangle}\right)^2=\langle\Psi|\rho|\Psi\rangle. \tag{4.3}$$

4.1.2. CORRESPONDENCE BETWEEN BELL STATES AND PAULI MATRICES

The Bell states, which were introduced in equation (4.2), form a basis for the Hilbert space of a two-qubit system. In this thesis we will focus on initial states of the protocol that are diagonal in the Bell basis:

$$\rho=A\left|\Phi^+\right\rangle\left\langle\Phi^+\right|+B\left|\Psi^+\right\rangle\left\langle\Psi^+\right|+C|\Psi^-\rangle\langle\Psi^-|+D|\Phi^-\rangle\langle\Phi^-|, \tag{4.4}$$

where $A=\left\langle\Phi^+\right|\rho\left|\Phi^+\right\rangle$, $B=\left\langle\Psi^+\right|\rho\left|\Psi^+\right\rangle$, $C=\langle\Psi^-|\rho|\Psi^-\rangle$ and $D=\langle\Phi^-|\rho|\Phi^-\rangle$.

A two-qubit state that is not Bell diagonal can be be transformed into a Bell diagonal state through twirling. *Twirling* is a technique in which each of the unitary operations in a set $T=\{U_i\}$ is applied with equal probability to an input state. In the case that T is finite, this results in the transformation

$$\rho\rightarrow\frac{1}{|T|}\sum_{U_i\in T}U_i\rho U_i^\dagger. \tag{4.5}$$

To transform an arbitrary two-qubit state into a Bell diagonal state, twirling over the set $T=\{I\otimes I,X\otimes X,Y\otimes Y,Z\otimes Z\}$ can be applied (Bennett et al., 1996b).

Using equations (3.1) and (4.2), it can be derived that

$$\begin{aligned}\left|\Psi^+\right\rangle&=(I\otimes X)\left|\Phi^+\right\rangle,\\ |\Psi^-\rangle&=(I\otimes Y)\left|\Phi^+\right\rangle,\\ |\Phi^-\rangle&=(I\otimes Z)\left|\Phi^+\right\rangle.\end{aligned} \tag{4.6}$$

This results in an alternative description for 4.4:

$$\begin{aligned}\rho=&p_I\left|\Phi^+\right\rangle\left\langle\Phi^+\right|+p_X(I\otimes X)\left|\Phi^+\right\rangle\left\langle\Phi^+\right|(I\otimes X)+p_Y(I\otimes Y)\left|\Phi^+\right\rangle\left\langle\Phi^+\right|(I\otimes Y)\\ &+p_Z(I\otimes Z)\left|\Phi^+\right\rangle\left\langle\Phi^+\right|(I\otimes Z).\end{aligned} \tag{4.7}$$

Note that the coefficients p_I,p_X,p_Y and p_Z are equal to A,B,C and D in 4.4, respectively, but are renamed here for convenience. One way of interpreting this formula is that Alice prepares a qubit pair in the state $\left|\Phi^+\right\rangle$. Then she transmits one qubit of each pair to Bob, possibly introducing errors. These errors on Bob's qubits can be modelled as letting a Pauli operator P act on the qubit with a certain probability p_P.

Note that Pauli operations are performed on Bob's qubits only. We will shorten the notation by writing P instead of $I\otimes P$ and writing P_1P_2 for the tensor product $(I\otimes P_1)\otimes(I\otimes P_2)$.

If we have n independent two-qubit states, their combined state can be fully described by the probabilities $p_{P_1P_2\ldots P_n}=p_{P_1}\cdot\ldots\cdot p_{P_n}$ that the system is in the state obtained by

applying $P_1P_2...P_n$ to the initial state $|\Phi^+\rangle^{\otimes n}$. Note that for states that are not independent, the joint probabilities are in general not equal to products of the probabilities of the separate states. In this thesis, however, we will focus on protocols with independent input states.

For a system that consists of two qubit pairs, for example, these probabilities are given in Figure 4.1a. Vertically written is the Pauli operation that acts on Bob's qubit of the first pair, horizontally the operation on the second pair. This can be extended to a cube for three dimensions (Figure 4.1b) and a hypercube for higher dimensions.

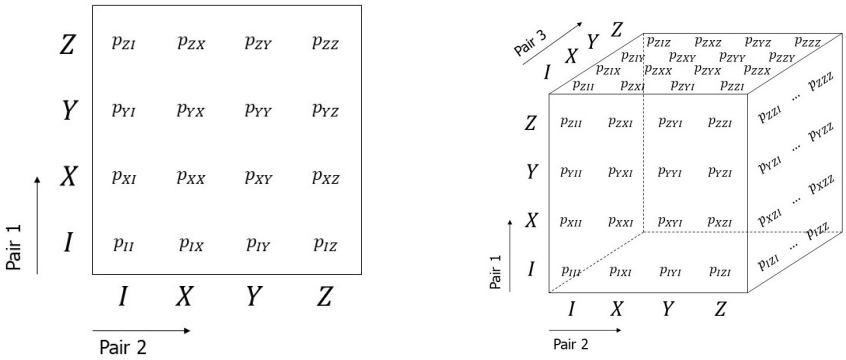


Figure 4.1: Probabilities that describe the state of a 2-pair system (a) and a 3-pair system (b). Each dimension corresponds with one qubit pair.

4.2. STRUCTURE OF DISTILLATION PROTOCOLS

This section gives a general description of the structure of distillation protocols that are considered in this thesis. As mentioned before, we will consider a system that consists of two parties, Alice and Bob. Alice and Bob have access to n independent, known input states. These states may be mixed or pure states, and are entangled. Of each pair, Alice and Bob possess one qubit. In the protocol, only LOCC is used by Alice and Bob. Firstly, Alice and Bob perform local (unitary) operations on their qubits, which we denote by U_A and U_B. Then, they perform a measurement on $n-m$ of the qubit pairs, with $n-m$ strictly smaller than n. Alice and Bob report their measurement results to each other via a classical communication channel. Based on the outcomes, they either keep or discard the remaining m qubit pairs. In Figure 4.2, this procedure is visualised for $n=2$ and $m=1$.

If Alice and Bob keep the qubit pairs, the protocol is called *successful*. The precise details of this decision depend on the chosen protocol, but the result is always that successful measurements result in states that have entanglement greater than or equal to the entanglement of the input states.

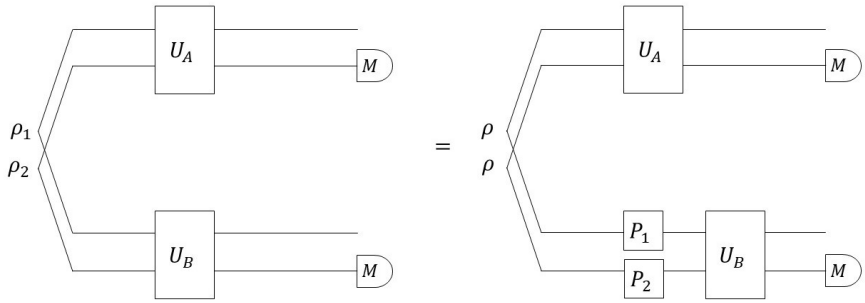


Figure 4.2: Structure of distillation protocols for $n = 2$ and $m = 1$. We start with two independent two-qubit states described by density operators ρ_1 and ρ_2 (left). Alternatively, the states can be described in terms of Pauli matrices on Bob's qubits that act on the state $\rho = |\Psi^+\rangle\langle\Psi^+|$, as explained in Section 4.1.2 (right). Alice and Bob both perform unitary operations on their qubits, given by U_A and U_B respectively. Finally, one of the qubit pairs is measured. The outcomes are communicated via classical communication.

Since the protocol starts with n qubit pairs and it outputs m qubit pairs, it is called $n \to m$ distillation. Figure 4.3 gives a schematic overview of $n \to m$ distillation.

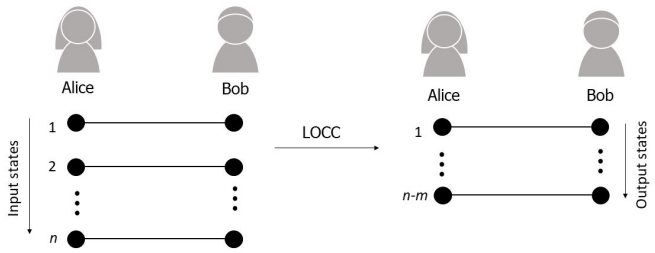


Figure 4.3: Quantum entanglement distillation for two parties, Alice and Bob. They measure $n - m$ of the qubits. If the protocol is successful, Alice and Bob are left with m pairs with higher entanglement than the input states.

To further increase the entanglement, it is possible to use a recurrence scheme. For instance, for arbitrary $n > 1$ and $m = 1$, the protocol is performed n times. The n independent output states are used as input for the next iteration. Of course, it is possible to do more than two iterations. Recurrence schemes that use a protocol with $m \neq 1$ are also possible, but then the input for the second iterations no longer consists of n independent pairs.

In the limit of high output fidelity, however, the yield of using a recurrence scheme approaches zero, because in each iteration at least half of the qubit pairs is lost: the pairs that are measured are always discarded and the pairs that are not measured are discarded in case of unsuccessful measurements. The yield can be improved by switching from a recurrence scheme to a hashing protocol, as soon as doing so results in more good qubit pairs then doing another iteration of the recurrence scheme.

Suppose that we have n copies of a Bell diagonal state described by the four probabilities $p = \{p_I, p_X, p_Y, p_Z\}$. In the limit of large n, the m output states approach perfect maximally entangled states and the yield $\frac{m}{n}$ approaches $1 - H(p)$. The quantity $H(p)$ is known as the *Shannon entropy* and is given by

$$H(p) = -\sum_{p_i \in p} p_i \log_2 p_i. \tag{4.8}$$

The details of the hashing method can be found in Bennett et al., 1996b. The yield in a hashing protocol can be used as a measure to compare distillation protocols. For this, we first perform one round of the n-to-1 distillation protocol and then we use the output of this distillation protocol as input for the hashing protocol. For a fair comparison of the distillation protocols, the number of copies used in the n-to-1 protocol and the probability of success of this protocol should be taken into account. Altogether, this results in a measure that we will refer to as the rate.

Definition 4.4. The *rate* of an n-to-1 protocol with success probability p_{suc} and output state described by $p = \{p_I, p_X, p_Y, p_Z\}$ is defined as

$$r = \frac{(1 - H(p))\, p_{suc}}{n}.$$

4.3. DEJMPS PROTOCOL

An example of a distillation protocol is the DEJMPS-protocol, which was published by Deutsch et al., 1996. The DEJMPS protocol is a protocol for $2 \to 1$ distillation. The measure of entanglement used in the DEJMPS protocol is the fidelity $F\left(\left|\Phi^+\right\rangle\left\langle\Phi^+\right|, \rho\right)$ of the Bell state $\left|\Phi^+\right\rangle$ and an arbitrary state ρ. It can be proved that the DEJMPS-protocol, yields the best possible fidelity that can be achieved through one iteration for input states with $p_I > p_X \geq p_Z \geq p_Y$. Such an ordering of the coefficients can be achieved using single-qubit operations. In a recurrence scheme, this reordering of the probabilities can be done after every iteration (Dehaene et al., 2003b). In this section we give a description of the DEJMPS protocol.

We start with two qubit pairs, whose density operators we denote by ρ and ρ'. The whole system is assumed to be in the state $\rho \otimes \rho'$. Firstly, Alice performs a unitary operation given by

$$\begin{aligned} U_A : |0\rangle &\mapsto \frac{1}{\sqrt{2}}\left(|0\rangle - i\,|1\rangle\right), \\ |1\rangle &\mapsto \frac{1}{\sqrt{2}}\left(|1\rangle - i\,|0\rangle\right). \end{aligned} \tag{4.9}$$

Bob performs the inverse of this operation, which is given by

$$\begin{aligned} U_B : |0\rangle &\mapsto \frac{1}{\sqrt{2}}\left(|0\rangle + i\,|1\rangle\right), \\ |1\rangle &\mapsto \frac{1}{\sqrt{2}}\left(|1\rangle + i\,|0\rangle\right). \end{aligned} \tag{4.10}$$

Then Alice and Bob both perform a CNOT operation on their qubits. Here, the two control qubits belong to the same pair ρ. Finally Alice and Bob both measure the target qubit in the Z basis. This is equivalent to the measurement in the computational basis, which was described in Section 2.1.4. If the outcomes of Alice and Bob coincide, they keep the control pair. The control pair can be used in further iterations of the protocol. However, if the outcomes do not coincide, the protocol was unsuccessful and the control pair is disregarded. The target pair is disregarded in both cases. The DEJMPS protocol is graphically summarized in Figure 4.4.

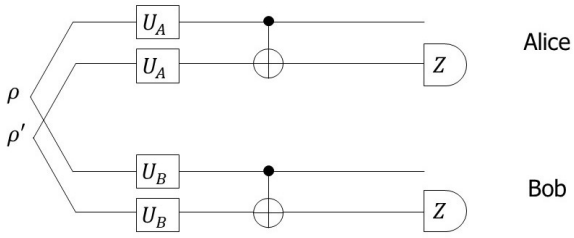


Figure 4.4: Quantum circuit of the DEJMPS protocol. Alice and Bob share two qubit pairs ρ and ρ'. They first perform a unitary operation given by equations (4.9) and (4.10), respectively. Then Alice and Bob both apply a CNOT operation and finally they perform a measurement of the target qubits in the σ_z basis.

Using equations (4.2) and (4.6) the result of the operations U_A, U_B and the CNOT operations can be translated to permutations of the probabilities from Figure 4.1a. These permutations are shown in Figure 4.5. In each step the probabilities in bold are the probabilities that have changed in that step.

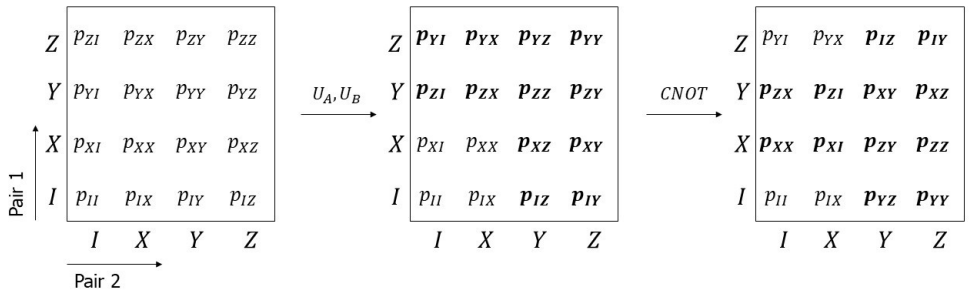


Figure 4.5: Overview of the permutations of the probabilities due to U_A and U_B (first to second table) and the CNOT operations (second to third table). The probabilities in bold are the probabilities that have changed in that step.

Keeping in mind that the protocol is successful if the measurements of Alice and Bob coincide, it can be derived from equation (4.2) that the protocol is successful if the measured qubit pair (pair 2) is in state $|\Phi^+\rangle$ or $|\Phi^-\rangle$. From (4.6) we know that these states correspond to the Pauli I and Z matrices, respectively. The probability that qubit pair 2 is described by either I or Z after applying the protocol can be read from the third table

of Figure 4.5: it is equal to the sum of the elements of the first and the last column. That is, the success probability is equal to

$$p_{suc} = p_{II} + p_{XX} + p_{ZX} + p_{YI} + p_{YY} + p_{ZZ} + p_{XZ} + p_{IY}. \tag{4.11}$$

Moreover, the fidelity of the remaining control pair is equal to the probability that the remaining pair (pair 1) is in state $|\Phi^+\rangle$, which is described by I. Taking into account that pair 1 is discarded if pair 2 is in the X or Y column, it follows that the fidelity is proportional to the sum of the probabilities in the bottom corners. Finally, since all probabilities of the remaining state must add up to 1, the probability should be normalized by dividing by the probabilities of all possible states. That is,

$$F = \frac{p_{II} + p_{YY}}{p_{suc}}. \tag{4.12}$$

5

BILOCAL CLIFFORD CIRCUITS

In this thesis we restrict ourselves to Clifford circuits. Clifford circuits are circuits that are composed of the Hadamard gate (H), the Phase gate (S) and the controlled-NOT gate (CNOT). The Hadamard and Phase gate are allowed on every qubit and the CNOT gate is allowed on every set of two qubits. Note that this set of operations is exactly equal to the Clifford set that was defined in Section 3.1.2. It is a well-known result, known as the Gottesman-Knill theorem, that these Clifford circuits can be simulated in polynomial time on a classical computer (Gottesman, 1998).

In Section 5.1 the structure of the Clifford circuits is explained in more detail. The starting point for this discussion is the general structure that was explained in Section 4.2. From here, Clifford circuits are characterized by putting restrictions on the operations and the measurements.

Then, in Section 5.2 the concepts *base* and *pillars* are introduced. The introduction of these concepts yields a more visual and intuitive way to think about distillation protocols.

Finally, in Section 5.3 a characterization of the subgroup of all Clifford operations that preserves the distillation statistics (the fidelity and the success probability) is given. In this section, two important new results are proved. Firstly, a generating set (Theorem 5.9) for the subgroup is given and proved and then, the order of the subgroup is determined (Theorem 5.13).

5.1. CHARACTERIZATION OF BILOCAL CLIFFORD CIRCUITS

This section covers the structure of the bilocal Clifford circuits. The general structure of these protocols is the same as the structure explained in Section 4.2, but restrictions are put on the operations and the measurements.

Firstly, Alice and Bob perform only Clifford operations. These are operations that can

be decomposed into Hadamard gates (H), Phase gates (S) and Controlled-NOT gates (CNOT). The Clifford gates that Alice and Bob use are strongly correlated. Suppose that Alice uses a gate C, then Bob uses the Hermitian conjugate of this gate, C^*. Note that C^* is again a Clifford operation, because $H^* = H$, CNOT* =CNOT and $S^* = S^3$. For two qubit pairs, the structure of a bilocal Clifford circuit is shown schematically in Figure 5.1a. In this figure, the representation of Bell diagonal states in terms of Pauli matrices is used (see Section 4.1.2).

The following theorem yields a way to rewrite a bilocal Clifford circuit in terms of operations on Bob's qubits only. Although in this thesis only qubits ($d = 2$) are considered, the result holds for general $d \geq 2$.

Theorem 5.1. *Let* $|\Psi\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |ii\rangle$ *be a maximally entangled quantum state, where* $\{|i\rangle\}_{i=0}^{d-1}$ *denotes the standard computational basis. For every* $d \times d$ *matrix* A *we have* $A \otimes I |\Psi\rangle = I \otimes A^T |\Psi\rangle$.

Proof. Let A be a $d \times d$ matrix. Then we can write $A = \sum_{i=0}^{d-1} \sum_{j=0}^{d-1} a_{ij} |i\rangle \langle j|$. Let $k, l \in \{0, ...d-1\}$. Then

$$(|k\rangle \langle l| \otimes I) |\Psi\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |k\rangle \langle l|i\rangle \otimes |i\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} \delta_{li} |k\rangle \otimes |i\rangle = \frac{1}{\sqrt{d}} |k\rangle \otimes |l\rangle .$$

By linearity of A it follows that $A \otimes I |\Psi\rangle = \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} \sum_{l=0}^{d-1} a_{kl} |k\rangle \otimes |l\rangle$.

On the other hand, note that $A^T = \sum_{i=0}^{d-1} \sum_{j=0}^{d-1} a_{ij} |j\rangle \langle i|$. Again, let $k, l \in \{0, ..., d-1\}$. Then

$$(I \otimes |l\rangle \langle k|) |\Psi\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i\rangle \otimes |l\rangle \langle k|i\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i\rangle \otimes \delta_{ki} |l\rangle = |k\rangle |l\rangle ,$$

and thus $I \otimes A^T |\Psi\rangle = \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} \sum_{l=0}^{d-1} a_{kl} |k\rangle \otimes |l\rangle$. So indeed, $A \otimes I |\Psi\rangle = I \otimes A^T |\Psi\rangle$. □

Note that we can indeed apply this theorem if we use the Pauli description of the Bell states, because the two-qubit states are then equal to the maximally entangled state $\rho = |\Phi^+\rangle \langle \Phi^+|$ acted on by some Pauli operations on Bob's qubits. Thus if Alice performs the Clifford operation C and Bob performs the operation C^*, then this is effectively the same as Bob performing $C^T(\cdot)C^*$. Note that C^T is again a Clifford operation itself, because $H^T = H$, CNOT^T =CNOT and $S^T = S$, so instead of C^T we can write $\tilde{C}$ or just C. As a result, $C^* \to C^\dagger$. The resulting circuit is shown in Figure 5.1b.

Note that the operations on Bob's qubits, $C(\cdot)C^\dagger$, are exactly the automorphisms on the Pauli group that were discussed in theorem 3.2.

Now that we have discussed the unitary operations performed by Alice and Bob, we arrive at the measurements. In this thesis, the choice is made to aim at increasing the fidelity of the input state and the $|\Phi^+\rangle$ state. Note that it does not matter which of the Bell states we chose, since the amount of entanglement of all four Bell states is equal and

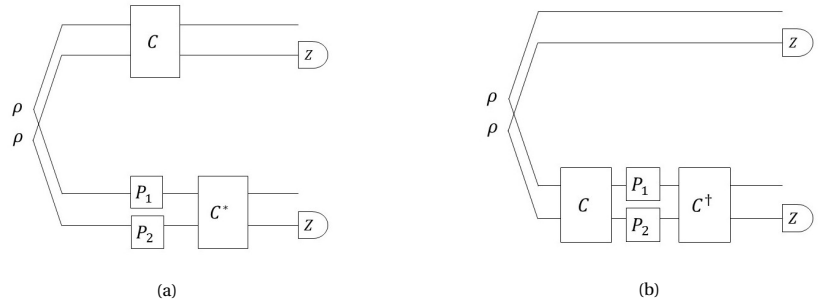


Figure 5.1: Bilocal Clifford circuit for two-qubits. Figure (a) shows the original description of the circuit, figure (b) shows the description after applying Theorem 5.1. Alice and Bob both perform a measurement in the computational basis on one qubit.

maximal. However, a choice should be made, because this influences the measurements that should be performed.

In our protocols, measurements in the computational basis are performed. The possible outcomes of the measurements thus are 1 (corresponding to $|0\rangle$) or -1 (corresponding to $|1\rangle$). Hence, the measurements are successful if the measured state is either $\left|\Phi^+\right\rangle$ or $|\Phi^-\rangle$.

5.2. BASE AND PILLARS

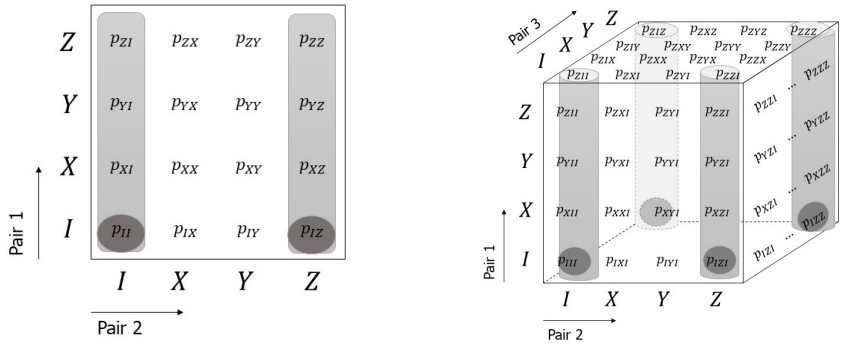
To illustrate the characterization of bilocal Clifford circuits in the previous section, in this section we consider the examples of two and three qubit pairs. These examples give rise to the introduction of the definitions of the base and pillars, which will allow for a more efficient description of the distillation protocols.

Firstly, we will consider the simplest non-trivial situation, namely the situation with two qubit pairs. We look at the results of the measurements in terms the probabilities described in Section 4.1.2. In Figure 4.1a a visualisation of these probabilities was given in the form a square. Which elements of the square do now correspond to a successful protocol? Keeping in mind that the protocol is successful if the second qubit pair is in state $\left|\Phi^+\right\rangle$ or $|\Phi^-\rangle$, that correspond with the Pauli matrices I and Z, respectively, we see that the protocol is successful if we are in the first or in the last column. These columns are highlighted in Figure 5.2a.

We see that the probability of success is equal to the sum of the probabilities in the highlighted columns. So in this case,

$$p_{suc} = p_{II} + p_{XI} + p_{YI} + p_{ZI} + p_{IZ} + p_{XZ} + p_{YZ} + p_{ZZ}. \tag{5.1}$$

Note that the probabilities in Figure 5.2a are not permuted compared to the initial situation. Of course, due to performed Clifford operations, the probabilities may be permuted, resulting in different probabilities occupying the highlighted columns. In each



(a) (b)

Figure 5.2: Probabilities that describe the state of a 2-pair system (a) and a 3-pair system (b). Each dimension corresponds with one qubit pair. The light grey rectangles/cylinders highlight the probabilities that correspond to the success probability (pillars).The darker circles highlight the probabilities that correspond to the fidelity (base).

5

case, the success probability is equal to the sum of the probabilities occupying the highlighted columns.

Following as similar argumentation as above, it can be derived that in general the success probability can be calculated as follows:

Lemma 5.2. *Let $\mathscr{S}$ be an n-qubit bipartite quantum system. The success probability of an n-to-1 bi-local Clifford protocol is given by*

$$p_{suc} = \sum_{P\in\{I,X,Y,Z\},\ Q_j\in\{I,Z\}} p_{PQ_1Q_2\ldots Q_{n-1}}.$$

For a three-qubit pair system, these probabilities are highlighted in light grey in Figure 5.2b.

Next to the success probability of the protocol, the fidelity of the remaining state is also of great importance. Since we are interested in obtaining the $|\Phi^+\rangle$ state, the fidelity is given by the probability that pair 1 is in the I-row. Because pair 1 is only kept if pair 2 is in the I- or Z-column, this probability is determined by the sum of p_{II} and p_{IZ}, the probabilities highlighted by darker circles in Figure 5.2a. (Again, this only holds for the situation with no permutations.) Moreover, since the probabilities of all possible states of pair 1 must add up to 1, we have to normalize $p_{II}+p_{IZ}$ by dividing it by the sum of all relevant probabilities. Thus for this situation, the fidelity is given by

$$F\left(\rho', |\Phi^+\rangle\right) = \frac{p_{II}+p_{IZ}}{p_{suc}}. \tag{5.2}$$

Again, this can be expanded to the general situation with n qubits, as is done in the Lemma 5.3.

Lemma 5.3. *Let $\mathscr{S}$ be an n-qubit bipartite quantum system. The fidelity of the remaining qubit pair and the $|\Psi^+\rangle$-state in an n-to-1 bi-local Clifford protocol is given by*

$$F\left(\rho', |\Phi^+\rangle\right) = \frac{\sum_{Q_j \in \{I,Z\}} p_{IQ_1Q_2...Q_{n-1}}}{p_{suc}}$$

We will refer to the fidelity and the success probability as the *distillation statistics*. Moreover, we will refer to the light grey columns as the *pillars* and to the darker circles as the *base*. For an arbitrary number of qubit pairs, this implies the following definitions for the base and the pillars.

Definition 5.4. The *base* of an n-qubits bipartite quantum system is given by

$$\mathscr{B} = \{IQ_1Q_2...Q_{n-1} \in \mathcal{P}_n : Q_j \in \{I,Z\} \forall\, j \in \{1,...,n-1\}\}.$$

From equation (3.11) it follows that, in the binary representation, an element in the base $\mathscr{B}$ is represented by a vector

$$b = \begin{bmatrix} v \\ w \end{bmatrix}, \quad v, w \in \mathbb{Z}_2^n,\ v = 0, w_1 = 0. \tag{5.3}$$

Definition 5.5. The *pillars* of an n-qubits bipartite quantum system are given by

$$\mathscr{P} = \{PQ_1Q_2...Q_{n-1} \in \mathcal{P}_n : P \in \{I,X,Y,Z\}, Q_j \in \{I,Z\} \forall\, j \in \{1,...,n-1\}\}.$$

An element in the pillars $\mathscr{P}$ can be represented by a binary vector

$$p = \begin{bmatrix} v \\ w \end{bmatrix}, \quad v, w \in \mathbb{Z}_2^n,\ v_i = 0 \quad \forall\, i \in \{2,...,n\}. \tag{5.4}$$

Note that every element of the pillars can be seen as a combination of a base element and a Pauli string $\tilde{P}_1$ with a Pauli matrix P on the first position and identity matrices on every other position. In the binary representation, p can thus be written as a linear combination of a binary base vector b and a vector with zeros on every position, except for, possibly, position 1 and $n+1$.

It is not hard to see that the probabilities that contribute to the success probability are exactly those probabilities that correspond to the elements of the pillars. This minor, but useful result is summarized in the following lemma.

Lemma 5.6. *Let $\mathscr{S}$ be an n-qubit bipartite quantum system with pillars $\mathscr{P}$, then the success probability of an n-to-1 distillation protocol on this system is given by*

$$p_{suc} = \sum_{P \in \mathscr{P}} p_P.$$

Similarly, the following relation between the base, the pillars and the fidelity can easily be seen.

Lemma 5.7. *Let $\mathscr{S}$ be an n-qubit bipartite quantum system with base $\mathscr{B}$ and pillars $\mathscr{P}$, then the fidelity of the remaining state after the distillation protocol is given by*

$$F\left(\rho', |\Phi^+\rangle\right) = \frac{\sum_{P\in\mathscr{B}} p_P}{\sum_{P\in\mathscr{P}} p_P}.$$

The Clifford operators performed in the protocol cause a permutation of the probabilities $p_{P_1P_2...P_n}$. By choosing the right Clifford operators, we can thus increase the fidelity or the success probability. However, as can be seen from Lemma 5.7, there is a trade-off between the two quantities: a high success probability leads to a lower fidelity and vice versa. Moreover, the probabilities of the base and pillar elements cannot be changed independently, as can be seen from Theorem 5.8.

Theorem 5.8. *Let $\mathscr{S}$ be an n-qubit bipartite quantum system with base $\mathscr{B}$ and pillars $\mathscr{P}$. Let $\sigma : P_n \to P_n$, $\sigma(P) = CPC^\dagger$, with $C \in \mathcal{C}_n$, such that $\sigma[\mathscr{B}] = \mathscr{B}$. Then $\sigma[\mathscr{P}] = \mathscr{P}$.*

Proof. Let $\tilde{P} \in \mathscr{P}$. Then $\tilde{P} = PQ_1...Q_{n-1}$ with $P \in \{I, X, Y, Z\}$ and $Q_j \in \{I, Z\}$ for all $j \in \{1, ..., n-1\}$. Note that P commutes with I and that for all j, Q_j commutes with both I and Z. As a result, $\tilde{P}$ commutes with all elements of $\mathscr{B}$.

Now let $\tilde{R} \in P_n \backslash \mathscr{P}$. Write $\tilde{R} = PR_1R_2...R_{n-1}$, where $P, R_j \in \{I, X, Y, Z\}$. Note that there is an $i \in \{1, ..., n-1\}$ such that $R_i \notin \{I, Z\}$. As a result, $\tilde{R}$ does not commute with $II_1..I_{i-1}Z_iI_{i+1}..I_{n-1} \in \mathscr{B}$. Hence, $\tilde{R}$ does not commute with all elements of $\mathscr{B}$.

We see that the elements of the pillars are exactly those elements of P_n that commute with all elements of the base. By Theorem 3.2 we know that σ is an automorphism on P_n. Thus σ preserves the commutation relations between the Pauli strings. Let $\tilde{B} \in \mathscr{B}$ and $\tilde{P} \in \mathscr{P}$. Then $\tilde{B}$ and $\tilde{P}$ commute. Thus $\sigma(\tilde{B})$ and $\sigma(\tilde{P})$ must commute too. Thus by the statement above $\sigma(\tilde{P}) \in \mathscr{P}$. Finally, since σ is an automorphism, we know that it is bijective and thus $\sigma[\mathscr{P}] = \mathscr{P}$. □

From Theorem 5.8 we can conclude that operations that leave the base invariant also leave the pillars invariant and thus preserve the distillation statistics.

5.3. PRESERVATION OF DISTILLATION STATISTICS

Recall from Section 3.1.2 that the Clifford group has order $|\mathcal{C}_n| = 2^{n^2+2n+3}\prod_{j=1}^{n}(4^j - 1)$. The order thus increases really fast with n. A better understanding of the permutations that preserve the distillation statistics significantly limits our search space of protocols. Therefore, in this section a characterization of the subset $\mathcal{D}_n \subset \mathcal{C}_n$, that leaves the distillation statistics invariant, is given. Firstly, note that $\mathcal{D}_n$ actually is a subgroup of $\mathcal{C}_n$. Indeed, if two operations $D_1, D_2 \in \mathcal{D}_n$ leave the base invariant, then their product D_1D_2 also leaves the base invariant. Moreover, the identity operation of course leaves the base invariant. And finally, the inverse of an operation $D \in \mathcal{D}_n$, which is an element

of $\mathcal{C}_n$, again leaves the base invariant, since D yields a permutation of the base elements.

In the trivial case that $n = 1$, we have $\mathcal{D}_1 = \mathcal{C}_1$. In this case, the only base element is the identity I, which is always mapped to itself under an automorphism. For all $n > 1$, however, $\mathcal{D}_n$ is a proper subgroup of $\mathcal{C}_n$. Consider for instance the Hadamard gate on the second qubit, which is an element of $\mathcal{C}_n$. This gate induces the swap of X_2 and Z_2 and hereby changes the base.

The aim of this section is to prove two theorems about the structure of $\mathcal{D}_n$. Firstly, in Section 5.3.1 a generating set of quantum gates is given. Then, in Section 5.3.2 a formula for the order of $\mathcal{D}_n$ in the binary representation is given and proved.

5.3.1. GENERATORS OF $\mathcal{D}_n$

Theorem 5.9. *Consider an n-to-1 distillation protocol with $n > 1$. The subgroup $\mathcal{D}_n$ of $\mathcal{C}_n$ that preserves the distillation statistics is generated by the union of the sets*

$$\begin{aligned} \{H_1\} &\quad \textit{(Hadamard gate on first qubit),} \\ \{S_i : i = 1,...,n\} &\quad \textit{(Phase gate on every qubit),} \\ \{CNOT_{ij} : i,j = 2,...,n,\ i \neq j\} &\quad \textit{(CNOT gate with not the first qubit as control or target qubit),} \\ \{CZ_{1i} : i = 2,...,n\} &\quad \textit{(CZ gate between the first qubit and every other qubit).} \end{aligned}$$

Throughout the rest of this section, the set $\{H_1\} \cup \{S_i : i = 1,...,n\} \cup \{\mathrm{CNOT}_{ij} : i,j = 2,...,n,\ i \neq j\} \cup \{\mathrm{CZ}_{1i} : i = 2,...,n\}$ is abbreviated as S^D. Theorem 5.9 is proved at the end of this section, but firstly some preliminary results are discussed and proved.

First of all, observe that $\langle S^D \rangle$ is a subgroup of $\mathcal{D}_n$. This can easily be checked by considering the permutations induced by the elements of S^D. To prove equality of the two subgroups $\langle S^D \rangle$ and $\mathcal{D}_n$, we claim that it is enough to show that $\langle S^D \rangle$ is a maximal subgroup of $\mathcal{C}_n$. Indeed, suppose that $\langle S^D \rangle$ is a maximal subgroup of $\mathcal{C}_n$. Then there does not exist a proper subgroup $\mathcal{F} \subset \mathcal{C}_n$, such that $\langle S^D \rangle \subset \mathcal{F}$ is a proper subgroup. We already showed that $\langle S^D \rangle$ is a subgroup of $\mathcal{D}_n$ and that $\mathcal{D}_n$ is a proper subgroup of $\mathcal{C}_n$. Thus if $\langle S^D \rangle$ is maximal, then it must be equal to $\mathcal{D}_n$. To prove that $\langle S^D \rangle$ is a maximal subgroup, we use the following lemma, which holds for general groups.

Lemma 5.10. *Let G be a group and let $H = \langle S^H \rangle$ be a subgroup of G. The subgroup H is a maximal subgroup of G if and only if for all $g \in G$, $\langle S^H, g \rangle = H$ or $\langle S^H, g \rangle = G$.*

Proof. Suppose $H = \langle S^H \rangle$ is a maximal subgroup of G. Let $g \in G$. Then $g \in H$ or $g \in G \backslash H$. If $g \in H$, then $\langle S^H, g \rangle = H$. If $g \in G \backslash H$ then $H = \langle S^H \rangle$ is a proper subgroup of $\langle S^H, g \rangle$. Because H is maximal, it then must hold that $\langle S^H, g \rangle = G$.

On the contrary, suppose H is not a maximal subgroup of G. Then there exists a proper subgroup F of G such that $H \subset F \subset G$. Define $T = F \backslash H$. Let $t \in T \subseteq G$, then $\langle S^H \rangle \subset \langle S^H, t \rangle \subseteq F \subset G$. □

The idea of the proof of Theorem 5.9 thus is to show that for every element $C \in \mathcal{C}_n$, we either have $\langle S^D, C\rangle = \langle S^D\rangle$ or $\langle S^D, C\rangle = \mathcal{C}_n$. As mentioned before, every element in $\mathcal{C}_n$ can be written in terms of Hadamard gates, phase gates and CNOT gates. Yet, this still leaves many possible combinations of gates to consider. Therefore, before going into the proof of Theorem 5.9, firstly a more efficient description of the Clifford gates is introduced.

Our starting point for this description is the Bruhat decomposition of the Clifford group. The Clifford group can be written as a disjoint union

$$\mathcal{C}_n = \bigsqcup_{h\in\{0,1\}^n} \bigsqcup_{\sigma\in\mathcal{S}_n} \mathcal{B}_n \left(\prod_{i=1}^{n} H_i^{h_i}\right) \sigma \mathcal{B}_n. \tag{5.5}$$

Here, $\mathcal{B}_n$ is the Borel subgroup of $\mathcal{C}_n$, which is generated by the gates S, CZ and $\text{CNOT}^{\downarrow}$. The notation $\text{CNOT}^{\downarrow}$ is used to indicate all CNOT_{ij} gates with $i < j$. From equation (5.5) it follows that every $C \in \mathcal{C}_n$ can be written as $C = FWF'$, with $F, F' \in \mathcal{B}_n$ and W a layer of Hadamard gates followed by a permutation $\sigma \in \mathcal{S}_n$ (Bravyi and Maslov, 2020; Maslov and Roetteler, 2018).

Next we define the alternative Borel subgroup $\hat{\mathcal{B}}_n$ as the group generated by S, CZ and $\text{CNOT}^{\uparrow}$, where $\text{CNOT}^{\uparrow}$ denotes the set of al CNOT_{ij} gates with $i > j$. In the following lemma, it is shown that $\mathcal{B}_n$ and $\hat{\mathcal{B}}_n$ are isomorphic.

Lemma 5.11. *Let $\tau \in \mathcal{S}_n$, $\tau = (1\ n)(2\ n-1)\ldots\left(\lceil \frac{n}{2}\rceil \lfloor \frac{n}{2}+1\rfloor\right)$. Let $f : \mathcal{B}_n \to \mathcal{C}_n,\ B \mapsto \tau B \tau$. Then f is a homomorphism and $f[\mathcal{B}_n] = \hat{\mathcal{B}}_n$. Moreover, $f : \mathcal{B}_n \to \hat{\mathcal{B}}_n$ is an isomorphism.*

Proof. The permutation τ reverses the order of all qubits. That is, qubit n becomes the first qubit, qubit $n-1$ the second qubit, and so on. Let $B_1, B_2 \in \mathcal{B}_n$. Because $\tau = \tau^{-1}$, we have $f(B_1B_2) = \tau B_1 B_2 \tau = \tau B_1 \tau \tau B_2 \tau = f(B_1) f(B_2)$. So indeed, f is an homomorphism.

To prove that $f[\mathcal{B}_n] = \hat{\mathcal{B}}_n$, we show that the images of the generators of $\mathcal{B}_n$ are exactly the generators of $\hat{\mathcal{B}}_n$. Firstly, let us look at the phase gates. It can easily be checked that for all $i \in \{1,\ldots,n\}$ we have $f(S_i) = S_{n+1-i}$. As a consequence, $f[\{S_i : i \in \{1,\ldots,n\}\}] = \{S_i : i \in \{1,\ldots,n\}\}$.

Now, let $B = \text{CZ}_{ij}$. Observe that $f(\text{CZ}_{ij}) = \text{CZ}_{ji} = \text{CZ}_{ij}$. So obviously, $f[\{\text{CZ}_{ij} : i, j \in \{1,\ldots,n\}, i \neq j\}] = \{\text{CZ}_{ij} : i, j \in \{1,\ldots,n\}, i \neq j\}$.

Finally, for $B = \text{CNOT}_{ij}$ we have $f(\text{CNOT}_{ij}) = \text{CNOT}_{ji}$. Thus, $f[\text{CNOT}^{\downarrow}] = \{\text{CNOT}_{ji} : i, j \in \{1,\ldots,n\}, i > j\} = \{\text{CNOT}_{ij} : i, j \in \{1,\ldots,n\}, i < j\} = \text{CNOT}^{\uparrow}$. So indeed $f[\mathcal{B}_n] = \hat{\mathcal{B}}_n$.

To show that $f : \mathcal{B}_n \to \hat{\mathcal{B}}_n$ is an isomorphism, we have to show that it is injective and surjective. Since $\tau B \tau = I$ implies that $B\tau = \tau$, which only holds for $B = I$, it follows that f indeed is injective. Now let $B' \in \hat{\mathcal{B}}_n$. Let $B = \tau B' \tau$. Then $B \in \mathcal{B}_n$ and $f(B) = B'$. Thus f is surjective. □

From Lemma 5.11 it follows that $\mathcal{B}_n$ and $\hat{\mathcal{B}}_n$ are isomorphic. As a result, we can rewrite equation (5.5) as

$$\mathcal{C}_n = \bigsqcup_{h\in\{0,1\}^n} \bigsqcup_{\sigma\in\mathcal{S}_n} \tau\hat{\mathcal{B}}_n\tau \left(\prod_{i=1}^n H_i^{h_i}\right)\sigma\tau\hat{\mathcal{B}}_n\tau, \tag{5.6}$$

or equivalently,

$$\tau\mathcal{C}_n\tau = \bigsqcup_{h\in\{0,1\}^n} \bigsqcup_{\sigma\in\mathcal{S}_n} \hat{\mathcal{B}}_n\tau \left(\prod_{i=1}^n H_i^{h_i}\right)\sigma\tau\hat{\mathcal{B}}_n. \tag{5.7}$$

To further simplify this expression, observe that $\mathcal{C}_n = \tau\mathcal{C}_n\tau$. Indeed, we know that $\tau\mathcal{C}_n\tau$ is the image of the function defined as $f : \mathcal{C}_n \to \mathcal{C}_n$, $G \mapsto \tau G\tau$. Because $\tau = \tau^{-1}$, this map is equal to conjugation of $\mathcal{C}_n$ by τ, which is an isomorphism. Thus it follows that $\tau\mathcal{C}_n\tau = \mathcal{C}_n$.

Regarding the middle part $\tau\left(\prod_{i=1}^n H_i^{h_i}\right)\sigma\tau$, note that the set $\left\{\left(\prod_{i=1}^n H_i^{h_i}\right)\sigma : h \in \{0,1\}^n, \sigma \in \mathcal{S}_n\right\}$ is symmetric in all qubits. Thus this set is left invariant under a relabeling of the qubits, as a consequence, the τ's can be removed from the expression.

The result of the above discussion is that we can rewrite equation (5.5) in terms of the alternative Borel subgroup $\hat{\mathcal{B}}_n$:

$$\mathcal{C}_n = \bigsqcup_{h\in\{0,1\}^n} \bigsqcup_{\sigma\in\mathcal{S}_n} \hat{\mathcal{B}}_n \left(\prod_{i=1}^n H_i^{h_i}\right)\sigma\hat{\mathcal{B}}_n. \tag{5.8}$$

It follows from equation (5.8) that every element $C \in \mathcal{C}_n$ can be written as $C = GWG'$, with $G, G' \in \hat{\mathcal{B}}_n$. The importance of this representation of $\mathcal{C}_n$ in terms of the alternative Borel subgroup $\hat{\mathcal{B}}_n$ instead of the standard Borel subgroup $\mathcal{B}_n$ becomes clear from Lemma 5.12.

Lemma 5.12. *$\hat{\mathcal{B}}_n$ is a subgroup of $\langle S^D\rangle$*

Proof. Since $\hat{\mathcal{B}}_n$ and $\langle S^D\rangle$ are both groups, it is enough to show that all generators of $\hat{\mathcal{B}}_n$ are contained in $\langle S^D\rangle$. Recall that $\hat{\mathcal{B}}_n$ is generated by $\{S_i : i = 1,...,n\} \cup$ $\{\mathrm{CZ}_{ij} : i, j = 1,...,n, i \neq j\} \cup \{\mathrm{CNOT}_{ij} : i, j = 1,...,n, i > j\}$. Since $S_i \in S^D$ for all i, it is clear that $S_i \in \langle S^D\rangle$.

Next, we show that $\mathrm{CZ}_{ij} \in \langle S^D\rangle$ for $i, j \neq 1$. We use that $\mathrm{CZ}_{1i}, \mathrm{CZ}_{1j}, H_1 \in S^D$. Then we can rewrite CZ_{ij} as shown in Figure 5.3. An extended proof of this equivalence can be found in the Appendix A.

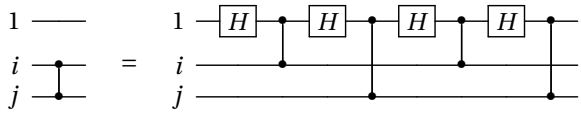


Figure 5.3: CZ_{ij} gate rewritten in terms of CZ_{1i}, CZ_{1j} and H_1.

To conclude, we show that $\text{CNOT}_{ij} \in S^D$ for all $i > j$. The set S^D already contains the CNOT gates between all qubits except for the first qubit, so the only thing left to show is that CNOT_{i1} is contained in S^D for all $i > 1$. That this is the case, follows from the fact that $\text{CZ}_{1i}, H_1 \in S^D$ and the equivalence of the ciruits in Figure 5.4.

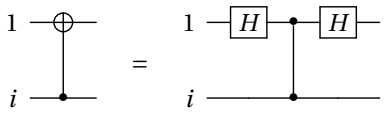


Figure 5.4: CNOT_{i1} gate rewritten in terms of CZ_{1i} and H_1.

Thus indeed, all generators of $\hat{\mathcal{B}}_n$ are contained in S^D and $\hat{\mathcal{B}}_n$ is a subgroup of $\langle S^D \rangle$. □

Note that unlike $\hat{\mathcal{B}}_n$, $\mathcal{B}_n$ is not a subgroup of $\langle S^D \rangle$. A CNOT_{12} gate, for instance, is contained in $\mathcal{B}_n$, but it is not contained in $\langle S^D \rangle$. This motivates the reformulation of equation (5.5) to equation (5.8).

Finally, we have covered all tools needed to complete the proof of Theorem 5.9.

Proof. Let $C \in \mathcal{C}_n$. Then we can write $C = GWG'$ with $G, G' \in \hat{\mathcal{B}}_n$ and

$$W = \left(\prod_{i=1}^{n} H_i^{h_i} \right) \sigma,$$

with $h \in \{0,1\}^n$, $\sigma \in \mathcal{S}_n$. We consider the group $\langle S^D,\ C \rangle$. Because $G, G' \in \hat{\mathcal{B}}_n$, it follows from Lemma 5.12 that $G, G' \in \langle S^D \rangle$. Thus $G^{-1}, G'^{-1} \in \langle S^D \rangle$ and $W = G^{-1} C G'^{-1}$. As a result, $\langle S^D,\ C \rangle = \langle S^D,\ W \rangle$.

It is left to show that for all W either $\langle S^D,\ W \rangle = \langle S^D \rangle$ or $\langle S^D,\ W \rangle = \mathcal{C}_n$. That is, we need to show that either $W \in \langle S^D \rangle$ or that by adding W to the generators we can 'build' a Hadamard gate on every qubit $i > 2$ and a CNOT_{1i} gate with $i > 1$. To prove this, it is enough to show that we can a SWAP_{1i} gate for one $i > 1$, because H_i can be built from H_1 and SWAP_{1i} (Figure 5.5a) and CNOT_{1i} can be built from CNOT_{i1} and SWAP_{1i} (Figure 5.5b). Moreover, for all qubits $j > 1$, H_j and CNOT_{1j} can be obtained by swapping qubits i and j. The SWAP operation of qubits i and j can be composed from CNOT gates that are already in S^D (see Figure 2.9).

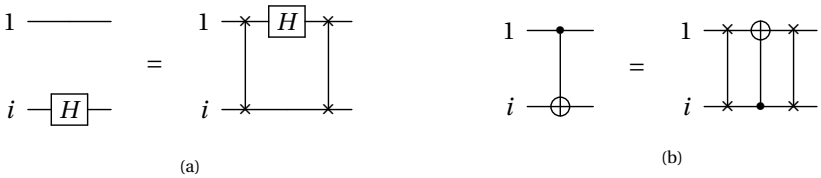


Figure 5.5: H_i gate rewritten in terms of two SWAP_{1i} gate and an H_i gate (a) and a CNOT_{1i} gate rewritten in terms of two SWAP_{12} gates and a CNOT_{i1} gate (b).

Firstly, we consider the case where σ is the identity permutation, so that we only have a column of Hadamard gates. Trivially, if $h_i = 0$ for all i, then $\langle S^D,\ W\rangle = \langle S^D\rangle$. Suppose that $h_1 = 1$ and $h_i = 0$ for all $i > 1$, then we have have the $W = H_1$, so $\langle S^D, W\rangle = \langle S^D\rangle$. Now suppose that h_1 is either 0 or 1, that there are k qubits $q_1, ..., q_k > 1$ such that $h_{q_i} = 1$ for all $i \in \{1, ..., k\}$ and that $h_j = 0$ for all other elements. In Figure 5.6 it is shown that the SWAP_{1q_1} gate can be composed from W, H_1 and CZ_{1q_1}. It follows that for $W = H_1^{h_1} H_{q_1} ... H_{q_k}$ we indeed have $\langle S^D,\ W\rangle = \mathcal{C}_n$.

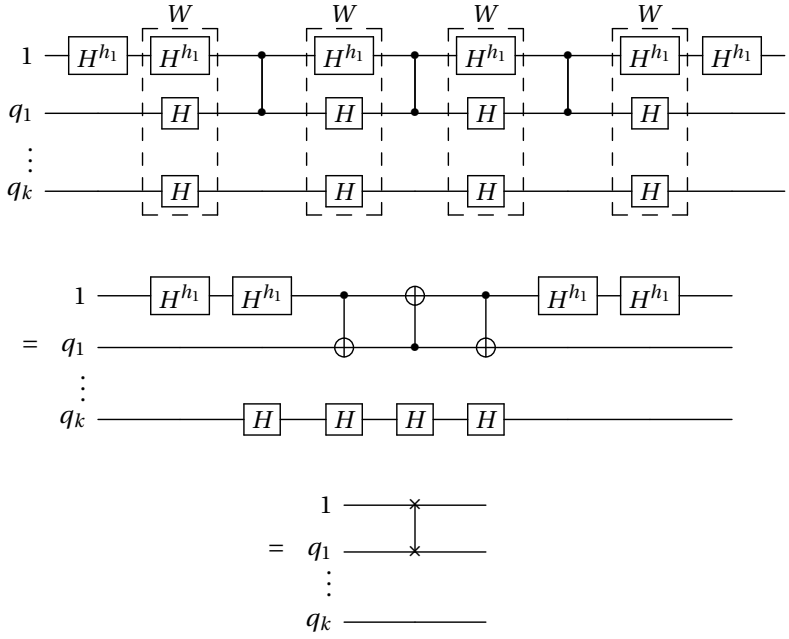


Figure 5.6: Composing the SWAP_{1q_1} from H_1, $W = H_1^{h_1} H_{q_1} ... H_{q_k}$ and CZ_{12}.

Next, consider the case that $h_i = 0$ for all i, but let σ now be a non-identity permutation. Then σ can be written as one cycle or as a product of disjoint cycles. We distinguish two types of cycles, namely cycles that include the first qubit, $(1\ q_1\ ...\ q_k)$, and cycles that do not include the first qubit, $(q_1\ ...\ q_k)$. Here $q_1,...,q_k$ denote different, arbitrary qubits other than 1. It is well-known that any such cycle can be written as a product of transpositions (see for instance Rotman, 1995). In particular, $(q_1\ ...\ q_k) = (q_1\ q_2)(q_2\ q_3)...(q_{k-1}\ q_k)$. In terms of quantum gates, we can thus rewrite any permutation in terms of SWAP gates. Now for cycles of the form $(q_1\ ...\ q_k)$ this only includes SWAP gates between qubits $i > 2$, which are already included in $\langle S^D \rangle$. On the other hand, if the permutation includes a cycle of the form $(1\ q_1\ ...\ q_k)$, then it contains one SWAP_{1q_1} gate. All other SWAP gates that are part of the permutation can be removed by multiplying with elements in $\langle S^D \rangle$, leaving only the SWAP_{1q_1} gate. Thus in this case, the whole group $\mathcal{C}_n$ can be generated.

Finally, let us consider elements W that are combinations of permutations and Hadamard gates. Note that if the permutation does not include a SWAP_{1i} gate, we can rewrite the element to a column of Hadamard gates, which was discussed earlier. Similarly, if the Hadamard column only contains H_1, then we are in the same situation as when there is no Hadamard column. Now suppose that our element contains a SWAP_{1i} gate and a Hadamard gate H_j, $j > 1$. By the arguments above, we can reduce such an element to a single SWAP_{12} gate an a column of Hadamard gates. We distinguish two cases, with or without a Hadamard gate on qubit i. Firstly, suppose that W has no H_i gate, but there is an H_j gate and possibly other H_k gates. This gate can be rewritten to a SWAP_{1j} gate by adding H_1 gates and a SWAP_{ij} gate, as shown in Figure 5.7.

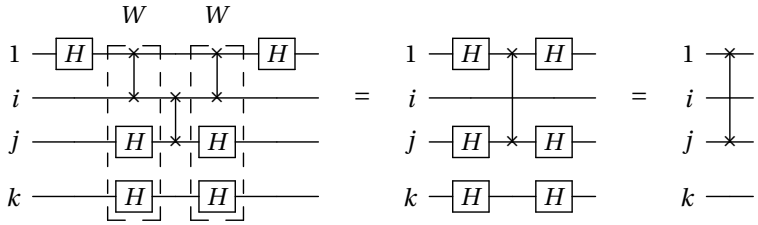


Figure 5.7: The combination of a SWAP_{1i} gate and a column of Hadamard gates $H_j H_k$ without H_i can be rewritten to SWAP_{1j} by adding two H_1 gates and a SWAP_{ij} gate.

Note that if W contains additional Hadamard gates on other qubits that $1, i, j$ and k, that these will vanish, just as the H_k gates do.

Secondly, suppose W has a Hadamard gate on qubit i. Then by adding a H_1 gate, we can rewrite it as a SWAP_{1i} gate, as shown in Figure 5.8. Note that if there are also Hadamard gates on other qubits $i > 2$, that after rewriting as shown in Figure 5.8 this case reduces to the first case (no H_i gate).

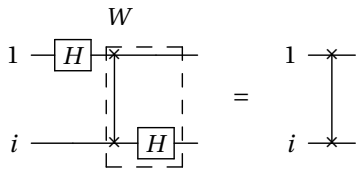


Figure 5.8: The combination of a SWAP_{1i} gate and a H_i gate can be rewritten to SWAP_{1i} by adding a H_1 gate.

The results of the discussion above are summarized in Table 5.1.

Table 5.1: For all possible elements W that are a combination of a Hadamard column $\prod_{i=1}^{n} H_i^{h_i}$ and a permutation σ either $\langle S^D, W\rangle = \langle S^D\rangle$ or $\langle S^D = \mathcal{C}_n$.

Hadamard	Permutation	$\langle S^D, W\rangle$
None	None	$\langle S^D\rangle$
None	Does not permute qubit 1	$\langle S^D\rangle$
None	Permutes qubit 1	$\mathcal{C}_n$
Only H_1	None	$\langle S^D\rangle$
Only H_1	Does not permute qubit 1	$\langle S^D\rangle$
Only H_1	Permutes qubit 1	$\mathcal{C}_n$
Includes H_i, $i \neq 1$	None	$\mathcal{C}_n$
Includes H_i, $i \neq 1$	Does not permute qubit 1	$\mathcal{C}_n$
Includes H_i, $i \neq 1$	Permutes qubit 1	$\mathcal{C}_n$

We see that for all possible $W = \left(\prod_{i=1}^{n} H_i^{h_i}\right)\sigma$, indeed $\langle S^D, W\rangle = \langle S^D\rangle$ or $\langle S^D, W\rangle = \mathcal{C}_n$. This completes the proof of Theorem 5.9. □

We have thus found a characterization of the subgroup of $\mathcal{C}_n$ that preserves the distillation statistics. At the beginning of this section it was claimed that a better understanding of this subgroup $\mathcal{D}_n$ limits the search space of protocols. And indeed, consider a protocol $C \in \mathcal{C}_n$. Then we can freely add or remove elements from $\mathcal{D}_n$ at the end of this protocol, without changing the fidelity and the success probability. That is, all elements in the right coset $\mathcal{D}_n C = \{DC : D \in \mathcal{D}_n\}$ yield the same distillation statistics. Instead of optimizing over all possible Clifford circuits it thus suffices to optimize over the right cosets of $\mathcal{D}_n$ in $\mathcal{C}_n$.

5.3.2. ORDER OF $\phi[\mathcal{D}_n]$

Recall from Section 3.2 that there is a homomorphism ϕ from the Clifford group $\mathcal{C}_n$ to the symplectic group $Sp(2n, \mathbb{Z}_2)$ that maps every Clifford operation to a binary matrix. In Section 3.2.1 it was proved that this homomorphism is surjective. The Clifford operations that are in the kernel $\langle \mathcal{P}_n, \eta I\rangle$ of ϕ map every Pauli string to itself, up to a phase factor. However, since we are only interested in the outcomes of measurements and we are working in the density operator formalism, this phase factor does not make

a difference for the distillation statistics (see Section 2.1.1 and Section 2.1.4). For the optimization of entanglement distillation protocols, it thus is sufficient to consider the elements of the Clifford group in the binary representation.

Because ϕ is a group homomorphism, the image of the subgroup $\mathcal{D}_n$, $\phi[\mathcal{D}_n]$, is a subgroup of $\phi[\mathcal{C}_n] = Sp(2n, \mathbb{Z}_2)$. At the end of Section 5.3.1 it was shown that is is enough to consider right cosets of $\mathcal{D}_n$ in $\mathcal{C}_n$ for the optimization of distillation protocols. Thus, in the binary representation, it is sufficient to optimize over cosets of $\phi[\mathcal{D}_n]$ in $\phi[\mathcal{C}_n]$. Note that the right cosets of $\mathcal{D}_n$ translate to left cosets in the binary picture.

To see how much looking at (left) cosets of $\phi[\mathcal{D}_n]$ in $\phi[\mathcal{C}_n]$ limits the search space of protocols, in this section a formula to compute the order of $\phi[\mathcal{D}_n]$ is presented and proved. As mentioned earlier, in the trivial case that $n = 1$ we have $\mathcal{D}_1 = \mathcal{C}_1$, so also $\phi[\mathcal{D}_1] = \phi[\mathcal{C}_1]$ and thus $|\phi[\mathcal{D}_1]| = |\phi[\mathcal{C}_1]| = 6$. For $n > 2$ the order of $\phi[\mathcal{D}_n]$ is given in Theorem 5.13.

Theorem 5.13. *For an n-to-1 distillation protocol, with $n > 1$, the order of $\phi[\mathcal{D}_n]$ is given by*

$$|\phi[\mathcal{D}_n]| = 6 \cdot 2^{n^2-1} \prod_{j=1}^{n-1} (2^j - 1).$$

5

Proof. Recall from Section 3.2.1 that $\tilde{X}_i$, $i \in \{1, ..., n\}$ denotes a Pauli string with an X-matrix on the ith position and I matrices at all other positions and $\tilde{Z}_i$ denotes the similar Pauli string with a Z-matrix on position i. Recall that an element $D \in \mathcal{D}_n$ is fixed up to a phase factor by the image of the generating Pauli strings $\tilde{X}_i$ and $\tilde{Z}_i$ under conjugation by D. Since this phase factor vanishes in the binary representation, $\phi(D)$ is fully determined by this image.

We count how many transformations of $\tilde{X}_i$ and $\tilde{Z}_i$ are possible. Let us start by looking at $\tilde{Z}_n$. The Pauli string $\tilde{Z}_n$ is a base element, thus it must again be transformed to a base element, because D preserves the distillation statistics. There are 2^{n-1} base elements, but the identity element $I^{\otimes n}$ is always mapped to itself under conjugation. Thus there are $2^{n-1}-1$ possibilities for the transformation of $\tilde{Z}_n$. That all transformations are indeed possible, is proved by giving a construction in a similar way as was done in the proof of Theorem 3.12. Recall that $\tilde{Z}_n$ is represented in the binary representation by a vector $a^Z \in \mathbb{Z}_2^{2n}$ with $a^Z_{2n} = 1$ and zeros on every other position. Suppose that $\tilde{Z}_n$ is mapped to a base element $b \in \mathscr{B}$. We show that b can be transformed to a^Z through left multiplication by elements of $\phi[\mathcal{D}_n]$. The transformation from a^Z to b can then be obtained by taking the product of the inverses of these generators in reverse order.

Recall that the action of left multiplication by $\phi(S)$ and $\phi(\text{CNOT})$ gates is as follows:

- A phase gate on qubit i results in the addition of row i to row $n+i$.
- A CNOT gate from qubit i to qubit j results in the addition of row i to row j and of row $n+j$ to row $n+i$.

Moreover, using equation (3.31) it can be derived that the action of left multiplication by a CZ gate is as follows.

- A CZ gate on qubit i and j results in the addition of row i to row $n+j$ and the addition of row j to row $n+i$.

Note that $b_1, ..., b_{n+1} = 0$. The vector b can be transformed to a by taking the following steps.

1. If $b_{2n} = 0$, apply a CNOT_{ni} gate with i chosen such that $b_{n+i} = 1$. Note that there always is a $i > 2$ such that this is possible, because otherwise b is the zero vector, which corresponds to the identity element $I^{\otimes n}$. This ensures that $b_{2n} = 1$.
2. For all $i \in \{2, ..., n\}$ with $b_{n+i} = 1$, apply a CNOT_{in} gate. As a result, $b_{n+i} = b_{n+i} + b_{2n} = 1 + 1 = 0$.

Steps 1 and 2 are visually summarized below.

$$b = \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ \hline 0 \\ \cdot \\ \cdot \end{array}\right] \xrightarrow{1} \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ \hline 0 \\ \cdot \\ 1 \end{array}\right] \xrightarrow{1} \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ \hline 0 \\ 0 \\ 1 \end{array}\right] = a^Z$$

Regarding $\tilde{X}_n$, we know that conjugation preserves the commutation relations between Pauli strings, so $D\tilde{X}_n D^\dagger$ must anti-commute with $D\tilde{Z}_n D^\dagger$. Observe that every element $P \in \mathcal{P}_n \backslash \{I^{\otimes n}\}$ anti-commutes with exactly half of the elements of $\mathcal{P}_n$ [1]. Thus there are $\frac{|\mathcal{P}_n|}{2} = \frac{4^n}{2} = 2^{2n-1}$ possibilities for the transformation of $\tilde{X}_n$.

Recall that $\tilde{X}_n$ is represented in the binary representation by a vector $a^X \in \mathbb{Z}_2$ with $a^X_n = 1$ and zeros on every other position. Suppose that D maps $\tilde{X}_n$ to a Pauli string whose binary representation is equal to c. Because $\tilde{X}_n$ and $\tilde{Z}_n$ anti-commute, it follows that steps 1 and 2 map c to a string that anti-commutes with a^Z. By equation (3.19) it follows that then $c_n = 1$. Now c can be transformed to a^X without affecting a^Z by taking the following steps.

3. For all i with $c_i = 1$ apply a CNOT_{ni} gate. This ensures that the upper half of c contains only zeros.
4. For all $i \neq n$ with $c_{n+i} = 1$, apply a CZ_{in} gate.
5. If $c_{2i} = 1$, apply an S gate on qubit n.

[1] Let $P \in \mathcal{P}_n \backslash \{I^{\otimes n}\}$. Let k be a position where P does not have an identity matrix. Then the Pauli strings that anti-commute with P can be constructed by selecting random Pauli matrices for all positions but position k and then choosing the Pauli matrix on position k such that the string anti-commutes with P.

Steps 3 to 5 are visually summarized below.

$$c = \begin{bmatrix} \cdot \\ \cdot \\ 1 \\ \hline \cdot \\ \cdot \\ \cdot \end{bmatrix} \xrightarrow{3} \begin{bmatrix} 0 \\ 0 \\ 1 \\ \hline \cdot \\ \cdot \\ \cdot \end{bmatrix} \xrightarrow{4} \begin{bmatrix} 0 \\ 0 \\ 1 \\ \hline 0 \\ 0 \\ \cdot \end{bmatrix} \xrightarrow{5} \begin{bmatrix} 0 \\ 0 \\ 1 \\ \hline 0 \\ 0 \\ 0 \end{bmatrix} = a^X$$

The gates used in this construction are CNOT_{ni}, CNOT_{in} with $i > 1$, CZ_{in} and S_n, which are indeed all contained in $\mathcal{D}_n$. Thus in total there are $2^{2n-1}(2^{n-1}-1)$ possible transformations for $\tilde{X}_n$ and $\tilde{Z}_n$.

The elements of $\phi[\mathcal{D}_n]$ that leave $\tilde{X}_n$ and $\tilde{Z}_n$ invariant form a subgroup that is isomorphic to $\phi[\mathcal{D}_{n-1}]$, with the number of cosets in $\phi[\mathcal{D}_n]$ equal to $2^{2n-1}(2^{n-1}-1)$. Thus

$$|\phi[\mathcal{D}_n]| = 2^{2n-1}(2^{n-1}-1)|\phi[\mathcal{D}_{n-1}]|.$$

By induction on n it follows that

$$\begin{aligned} |\phi[\mathcal{D}_n]| &= |\phi[\mathcal{D}_1]| \prod_{j=2}^{n} 2^{2j-1}(2^{j-1}-1) \\ &= 6 \cdot 2^{\sum_{j=2}^{n}(2j-1)} \prod_{j=2}^{n} (2^j - 1) \\ &= 6 \cdot 2^{n^2-1} \prod_{j=1}^{n-1} (2^j - 1). \end{aligned}$$

□

Corollary 5.14. *The index of $\phi[\mathcal{D}_n]$ in $\phi[\mathcal{C}_n]$ is given by*

$$\left[\phi[\mathcal{C}_n] : \phi[\mathcal{D}_n]\right] = \frac{1}{3}(2^n - 1) \prod_{j=1}^{n} (2^j + 1).$$

Proof. Recall that $|\phi[\mathcal{C}_n]| = 2^{n^2}\prod_{j=1}^{n}(4^j-1)$. As a consequence,

$$\begin{aligned}\big[\phi[\mathcal{C}_n]:\phi[\mathcal{D}_n]\big] &= \frac{|\phi[\mathcal{C}_n]|}{|\phi[\mathcal{D}_n]|}\\ &= \frac{2^{n^2}\prod_{j=1}^{n}(4^j-1)}{6\cdot 2^{n^2-1}\prod_{j=1}^{n-1}(2^j-1)}\\ &= \frac{\prod_{j=1}^{n}(2^j-1)(2^j+1)}{3\prod_{j=1}^{n-1}(2^j-1)}\\ &= \frac{\prod_{j=1}^{n}(2^j-1)(2^j+1)}{3\prod_{j=1}^{n}(2^j-1)}(2^n-1)\\ &= \frac{1}{3}(2^n-1)\prod_{j=1}^{n}(2^j+1).\end{aligned}$$

□

5

5.3.3. FURTHER REDUCTION FOR SYMMETRIC INPUT STATES

In Section 5.3.1 it was explained that all elements in a right coset of $\mathcal{D}_n$ in $\mathcal{C}_n$ yield the same distillation statistics when applied to a general input state. In literature on quantum information theory, often input states that possess some sort of symmetry are considered. A frequently used input state, for instance, is the n-fold tensor product $\rho^{\otimes n}$ of a single qubit pair ρ. Because all qubit pairs are initially equal, we can freely choose which qubits we measure and which one we keep, without changing the distillation statistics. In other words, we can apply a permutation to the qubit pairs at the beginning of the protocol without changing the results. As a result, all distillation protocols in a left coset $C\mathcal{S}_n$ in $\mathcal{C}_n$, with $C \in \mathcal{C}_n$ and $\mathcal{S}_n$ the group of permutations of $\{1,...,n\}$, yield the same distillation statistics.

Another type of state that is often considered in quantum information theory is the isotropic state. Isotropic states are states that are invariant under any unitary operation of the form $U \otimes U^*$:

$$\rho_{AB} = (U\otimes U^*)\rho_{AB}(U^\dagger \otimes (U^*)^\dagger). \tag{5.9}$$

An isotropic state has the form

$$\begin{aligned}\rho_{AB} &= p\left|\Phi^+\right\rangle\left\langle\Phi^+\right| + \frac{1-p}{4}I\\ &= \frac{3p-1}{4}\left|\Phi^+\right\rangle\left\langle\Phi^+\right| + \frac{1-p}{4}\left(\left|\Psi^+\right\rangle\left\langle\Psi^+\right| + |\Psi^-\rangle\langle\Psi^-| + |\Phi^-\rangle\langle\Phi^-|\right).\end{aligned} \tag{5.10}$$

An isotropic state is invariant under the permutation of $|\Phi^-\rangle$, $\left|\Psi^+\right\rangle$ and $|\Psi^-\rangle$. Again, it follows that all distillation protocols in a left coset $C\mathcal{S}$, with $\mathcal{S}$ a permutation group, yield the same distillation statistics.

Suppose that the input state indeed possesses some form of symmetry. Let $\mathcal{S}$ be the permutation group that leaves the input state invariant. Then it follows from the previous

section and the discussion above that all distillation protocols in the double coset $\mathcal{D}_n C \mathcal{S}$ in $\mathcal{C}_n$ yield the same distillation statistics. As a consequence, we only need to optimize over the double cosets. Especially for input states that possess a lot of symmetry, such as the n-fold tensor product of an isotropic state, this may result in a large reduction of the search space of protocols.

However, it is not clear yet how much this would limit the search space of protocols. In contrast to right (or left) cosets, double cosets do not necessarily all have the same order. Because $\mathcal{C}_n$, $\mathcal{D}_n$ and $\mathcal{S}_n$ are finite, the order of $\mathcal{D}_n C \mathcal{C}_n$ is equal to

$$|\mathcal{D}_n C \mathcal{S}_n| = \frac{|\mathcal{D}_n||\mathcal{S}_n|}{|\mathcal{D}_n \cap C \mathcal{S}_n C^{-1}|}. \tag{5.11}$$

This expression depends on C and is thus not necessarily equal for all double cosets.

6

ALGORITHMS FOR OPTIMIZATION

This chapter provides algorithms for the optimization over cosets of $\mathcal{C}_n$. In Section 1 it is described how a tranversal for the cosets can be found. Using this transversal, in Section 6.2 it is explained how the distillation statistics can be calculated. An implementation in SageMath can be found in Appendix B.

6.1. ALGORITHM FOR FINDING A TRANSVERSAL

In this section an algorithm for finding a transversal of the right cosets of $\mathcal{D}_n$ in $\mathcal{C}_n$ is described. This algorithm uses the binary representation of the Clifford operations that was explained in Section 3.2. The transversal is found by sampling random elements from the symplectic group $Sp(2n, \mathbb{Z}_2)$, which is already implemented in SAGEMATH. [1] For every new sampled element, it is checked whether or not the coset that this element belongs to is already represented in the transversal. If the coset is not represented yet, then the sampled element is added to the transversal.

For this approach to work, it is thus necessary to be able to check if two elements belong to the same coset. Recall that two elements belong to the same coset if and only if they result in the same distillation statistics (for a general input state). This is the case if the same Pauli strings are mapped to the base and the pillars (see Section 5.2). By Theorem 5.8 it is enough to check if the same Pauli strings are mapped to the base.

More formally, consider an n-qubit pairs bipartite system with base $\mathscr{B}$. Let M_1, $M_2 \in Sp(2n, \mathbb{Z}_2)$. Let σ_1, σ_2 denote the permutations of the Pauli strings induced by M_1 and M_2, respectively. Let $\mathscr{V}$ denote the set of Pauli strings that are mapped to the base under the permutation σ_1 and let $\mathscr{W}$ denote the set of Pauli strings that are mapped to the base

[1] The definition of the symplectic group that is implemented in SAGEMATH slightly differs from the definition used in this thesis. In SAGEMATH the symplectic matrices are defined as the matrices $M \in M_{2n\times 2n}(\mathbb{Z}_2)$ such that $M^T P M = P$ with P the anti-diagonal identity matrix. These matrices M can be transformed to matrices from the symplectic group from definition 3.5 via the transformation $M \to TMT^{-1}$ with T chosen such that $P = T^T \Omega T$.

under σ_2. Then M_1 and M_2 belong to the same coset if and only if $\mathcal{V} = \mathcal{W}$. Because σ_1 and σ_2 are permutations, this is equivalent to $\sigma_1^{-1}[\mathcal{B}] = \sigma_2^{-1}[\mathcal{B}]$. To check if M_1 and M_2 belong to the same coset, it thus suffices to check if the image $\mathcal{B}$ under σ_1^{-1} and σ_2^{-1} is the same. In terms of the binary representation, we thus need to check that left multiplication of the binary base vectors by M_1^{-1} and M_2^{-1} results in the same set of vectors.

Of course, this can be calculated by simply carrying out the matrix multiplications. We propose a more efficient manner below. Let $M \in Sp(2n, \mathbb{Z}_2)$. Firstly, using Proposition 3.7, the matrix M^{-1} can be efficiently calculated. Note that the base element $I^{\otimes n}$, the zero vector in the binary representation, is always mapped to itself under left multiplication by a matrix. Since the image of the zero vector is the same for all symplectic matrices, it can be omitted when comparing the images of the base under left multiplication by different symplectic matrices. The image of the other binary base vectors under left multiplication by M^{-1} can be calculated as follows. Firstly, the image of simplest non-trivial base element is calculated: a Pauli string $\tilde{Z}_i$ with $i \in \{2, ..., n\}$, with a Z matrix on position i and identity matrices at every other position. From equation (3.11) it follows that, in the binary representation, $\tilde{Z}_i$ is represented by a vector $a \in \mathbb{Z}_2^{2n}$ with $a_{n+i} = 1$ and zeros at every other position. It follows that $M^{-1}a$ is equal to the $(n+i)$th column of M^{-1}. By linearity of M^{-1}, the image of a base element with a Z matrix on position i and j can be obtained by adding the image of $\tilde{Z}_i$ and $\tilde{Z}_j$ modulo 2. This can be extended to calculate the image of an arbitrary base element. Thus the image of the binary base vectors under M^{-1} is equal to the set that consists of the zero vector of length $2n$, the columns $n+1, ..., 2n$ of M^{-1} and all linear combinations of these columns.

The algorithm used to find a transversal is summarized in Algorithm 1.

Algorithm 1: Finding a transversal

input : Number of copies n in the input state of the distillation protocol.
output: A transversal for the right cosets of $\mathcal{D}_n$ in $\mathcal{C}_n$ in the binary representation.

```
Transversal = {}
while length(Transversal) < [φ[C_n] : φ[D_n]] do
    Sample a random element M from Sp(2n, Z_2).
    Calculate M^{-1} using Proposition 3.7.
    Calculate the image M_base of the base under M^{-1}. This is equal to the set
      containing columns n+1, ..., 2n of M^{-1} and all linear combinations of these
      columns.
    if there is no N ∈ Transversal with N_base = M_base, where N_base is the image of the
      base under N^{-1} then
        Add M to Transversal.
end
```

Note that, when calculating the distillation statistics, the inverse of M is needed again. Therefore, in the implementation of this algorithm in SageMath, the matrix M^{-1} is added to the transversal instead of M.

6.2. ALGORITHM FOR CALCULATING DISTILLATION STATISTICS

In this section an algorithm to calculate the distillation statistics of a specific protocol applied to an initial state is described. Recall from Lemma 5.6 that the success probability of a distillation protocol can be calculated from the pillars $\mathscr{P}$ as follows:

$$p_{suc} = \sum_{P \in \mathscr{P}} p_P. \tag{6.1}$$

Similarly, by Lemma 5.7, the fidelity can be calculated from the base $\mathscr{B}$ and the pillars $\mathscr{P}$ as follows:

$$F\left(\rho', |\Phi^+\rangle\right) = \frac{\sum_{P \in \mathscr{B}} p_P}{\sum_{P \in \mathscr{P}} p_P}. \tag{6.2}$$

The first step towards calculating the fidelity and the success probability thus is to calculate which Pauli strings are mapped to the base and to the pillars by application of the operation with binary representation M. Recall from Section 6.1 that this is the same as calculating where the binary vectors corresponding to the base and the pillars are mapped to by the inverse matrix M^{-1}. In Section 6.1 it was described how this image can be determined for the base elements. For the pillars, recall from Section 5.2 that the binary representation of each pillar element is a linear combination of a binary base vector and a vector with zeros on every position, except for, possibly, position 1 and $n+1$. Thus, once the image of the base is known, the image of the pillars can be calculated by taking all linear combinations that consist of the image of one base vector and the image one vector $\begin{bmatrix} x_1 & 0 & \dots & 0 & z_1 & 0 & \dots & 0 \end{bmatrix}^T$ with $x_1, z_1 \in \mathbb{Z}_2$. The latter is equal to $x_1 M_1^{-1} + z_1 M_{n+1}^{-1}$, where M_1^{-1} and M_{n+1}^{-1} denote the first and $(n+1)$th column of M^{-1}, respectively.

The input for this algorithm is a Bell diagonal input state, specified by the four probabilities p_I, p_X, p_Y, p_Z. Firstly, the n-fold tensor product is calculated to obtain the hypercube of coefficients (see Figure 4.1). For a given protocol M, the success probability can now be found by adding the coefficients from the array that correspond to the image of the pillars. The fidelity can be found in a similar way, namely by adding the coefficients from the array that correspond to the image of the base and then dividing this by the success probability.

In Section 4.2 another measure of the quality of entanglement distillation statistics was discussed, namely the rate. To calculate the rate, not only the fidelity is needed, but also the other three coefficients that describe the output state are needed. To calculate these coefficients, observe that the pillars can be seen as four 'layers': one that corresponds to the I matrix on the first qubit, one that corresponds to X on the first qubit, one that corresponds to Y and one that corresponds to Z. The coefficients of the output state can be obtained by adding the coefficients in each layer of the pillars and dividing it by the success probability. Then, the rate can be calculated using Definition 4.4.

Using the algorithm described above, either a list of all possible distillation statistics can be obtained or the best achievable fidelity, success probability or rate can be calculated. If it is desired, the symplectic matrices can be translated back to a Clifford circuit by following the steps from the proof of Theorem 3.12.

7

RESULTS OF OPTIMIZATION

In this chapter the results of the optimization using the algorithms described in the previous chapter are shown.

7.1. ISOTROPIC STATES

Let us start by considering isotropic input states. Recall that an isotropic state has the form

$$\rho_{AB} = F_{in} \left|\Phi^+\right\rangle\left\langle\Phi^+\right| + \frac{1-F_{in}}{3}\left(\left|\Psi^+\right\rangle\left\langle\Psi^+\right| + \left|\Psi^-\right\rangle\left\langle\Psi^-\right| + \left|\Phi^-\right\rangle\left\langle\Phi^-\right|\right), \tag{7.1}$$

where F_{in} is the fidelity of the input state. Because of their symmetry, isotropic states are widely considered in quantum information theory.

As input states we consider n copies of an isotropic state. In Figure 7.1 the fidelity of the output state is plotted against the success probability for different input fidelities.

Note that for larger n the maximal achievable fidelity increases, but the success probability decreases. The latter is due to the fact that a protocol is called successful if and only if all measurements yield the same outcome. Of course, when more states are measured, the probability of all measurement outcomes being the same decreases. If a higher success probability is desired, one can always perform a protocol on only a part of the qubits instead of all of them. For $n = 4$, for instance, a 3-to-1 protocol can be performed, leaving the fourth qubit pair untouched.

In Figure 7.1 it can also be seen that there is only a small number of possible outcomes. Recall that the distillation statistics are calculated for $\frac{1}{3}(2^n-1)\prod_{j=1}^{n}(2^j+1)$ cosets. For $n = 2$, 3, 4 and 5 this is equal to 15, 315, 11475 and 782595 cosets, respectively. However, as can be seen in Figure 7.1, there are only 2, 4, 12 and 31 different outcomes for $n = 2$, 3, 4 and 5. This supports the discussion in Section 5.3.3 and is worth further investigation.

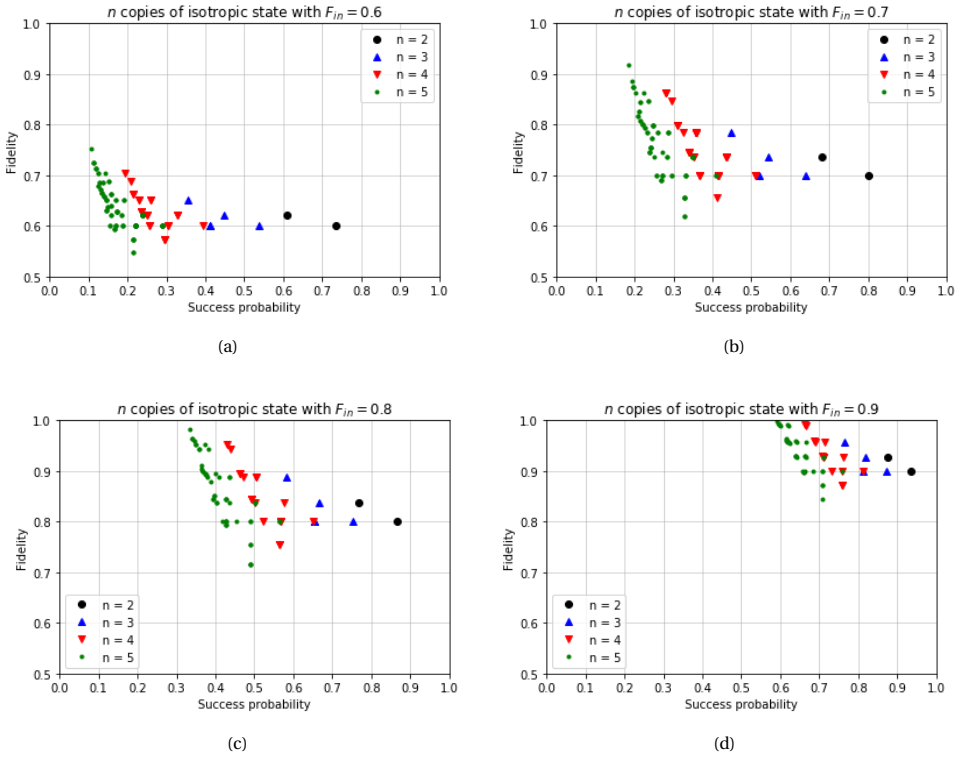


Figure 7.1: Possible distillation statistics of an n-to-1 distillation protocol applied to the n-fold tensor product of an isotropic state with input fidelity F_{in}.

In Figure 7.1 it can be seen that there are many protocols that are not optimal: they are dominated by other protocols that yield a higher fidelity and a higher success probability. In the end, however, we are interested in the best achievable distillation statistics. These distillation statistics are shown in Figure 7.2, where the Pareto front (the set of points that are not dominated by other points) is plotted in combination with linear interpolation between the points. The distillation statistics of the DEJMPS protocol are plotted as well. It can be seen that DEJMPS yields the best achievable fidelity for an $n = 2$ protocol and obtains this fidelity with the highest possible success probability.

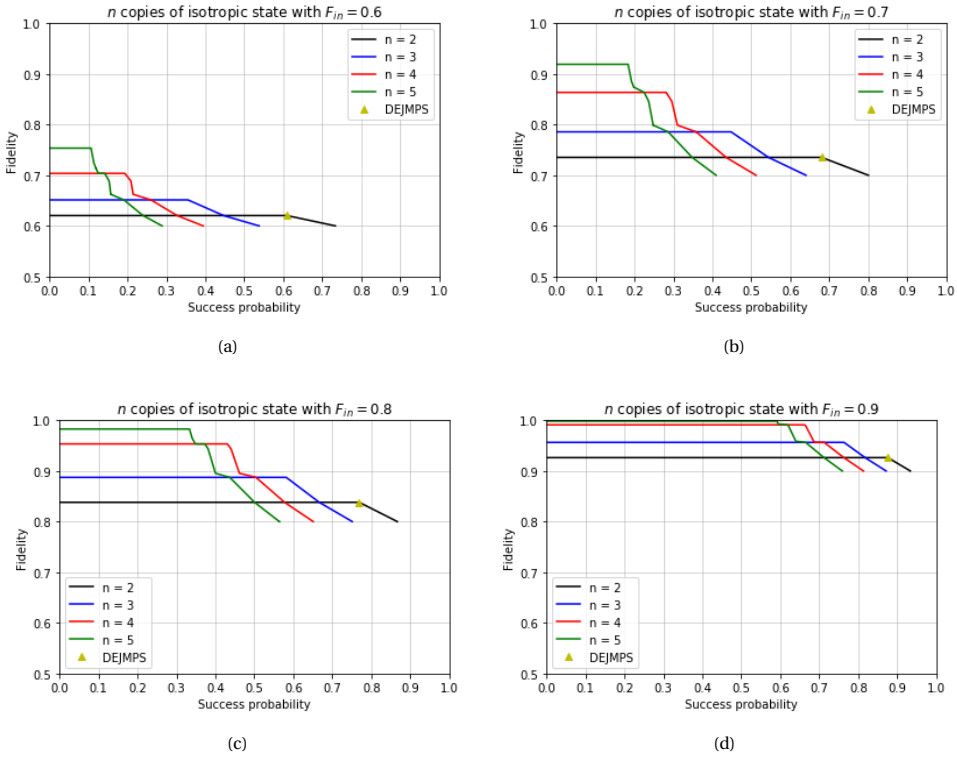


Figure 7.2: Best achievable distillation statistics of an n-to-1 distillation protocol applied to the n-fold tensor product of an isotropic state with input fidelity F_{in}.

In Figure 7.3 the highest achievable rate in a hashing protocol is plotted for isotropic states with different input fidelities. The results for $n = 5$ are not shown here, because the calculations were too computationally intensive. It can be seen that for low input fidelity ($F_{in} \lesssim 0.77$), distillation protocols for larger n result in a higher rate than a 2-to-1 protocol or no protocol. For higher input fidelities, however, the best rate can be obtained by not performing any n-to-1 protocol at all. Apparently, the larger number of copies needed in the n-to-1 protocol and the lower success probability for higher values of n significantly decrease the rate.

The rate achieved by applying the DEJMPS protocol is plotted as well. It can be seen that the DEJMPS protocol yields the highest achievable rate for $n = 2$.

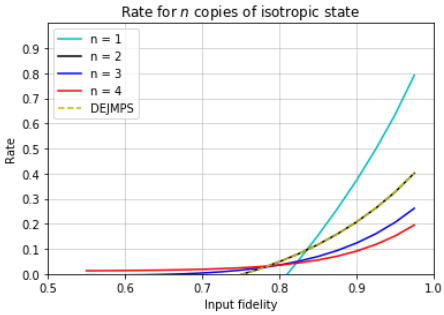


Figure 7.3: Best achievable rate in a hashing protocol for different input fidelities in an n-to-1 protocol with $n = 2$, 3 and 4. The rate achieved by applying DEJMPS is plotted as well.

7.2. BELL DIAGONAL STATES

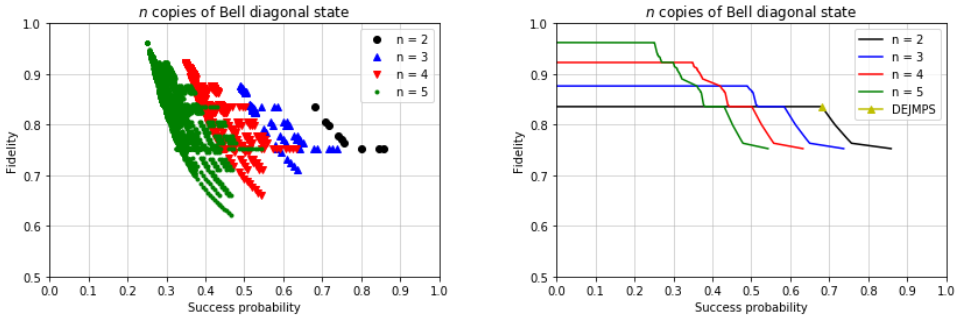
Next, let us consider general Bell diagonal states. As mentioned earlier, any two-qubit state can be brought into this form by twirling. We consider Bell diagonal states of the form

$$\rho_{AB} = p_1 \left|\Phi^+\right\rangle\left\langle\Phi^+\right| + p_2 \left|\Psi^+\right\rangle\left\langle\Psi^+\right| + p_3 \left|\Phi^-\right\rangle\left\langle\Phi^-\right| + (1 - p_1 - p_2 - p_3) \left|\Psi^-\right\rangle\left\langle\Psi^-\right|, \quad (7.2)$$

with $p_1 > 0.5$ and $p_1 \geq p_2 \geq p_3 \geq 1 - p_1 - p_2 - p_3$. Any Bell diagonal state with one coefficient larger than 0.5 can be transformed into this form by using local operations only. This ordering of the Bell coefficients allows us to obtain the best achievable fidelity (see Dehaene et al., 2003b).

The input state was randomly generated, which resulted in the coefficients $p_1 = 0.7526$, $p_2 = 0.1057$, $p_3 = 0.0938$. In Figure 7.4a the possible outcomes of distillation protocols are shown. It can be seen that there are a lot more different outcomes compared to the isotropic states in Figure 7.1. However, it cannot be said with certainty how many different outcomes there are, because it is difficult to determine which outcomes are really different and which outcomes are only different due to the computations of the computer.

In Figure 7.4b the best achievable distillation statistics are plotted. The DEJMPS protocol is indicated here as well. It can be seen that for $n = 2$, DEJMPS yields the best achievable fidelity ($F_{out} = 0.8356$) and that it is obtained with the highest possible success probability ($p_{suc} = 0.6806$).



(a) (b)

Figure 7.4: Possible outcomes for distillation protocols applied to an initial state that is the n-fold tensor product of a Bell diagonal state with $p_1 = 0.7526$, $p_2 = 0.1057$, $p_3 = 0.0938$ (a) and the best achievable distillation statistics for this input state (b).

For $n = 3$ the best achievable fidelity for this input states is equal to $F_{out} = 0.8763$. It can be obtained with success probability $p_{suc} = 0.4895$. One of the protocols that results in this fidelity and success probability is

$$\hat{M}_{n=3} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{7.3}$$

In terms of Clifford gates, it can be implemented as

$$\begin{aligned} C =& \mathrm{CNOT}_{13} S_1 H_1 \mathrm{CNOT}_{21} \mathrm{CNOT}_{31} \mathrm{CNOT}_{12} S_1 H_1 \mathrm{CNOT}_{21} \mathrm{CNOT}_{12} H_1 \mathrm{CNOT}_{21} \\ & S_1 H_1 \mathrm{CNOT}_{21} S_1. \end{aligned} \tag{7.4}$$

For $n = 4$ the best achievable fidelity equals $F_{out} = 0.9228$. It can be obtained with success probability $p_{suc} = 0.3494$. One of the corresponding protocols is

$$\hat{M}_{n=4} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \end{bmatrix}. \tag{7.5}$$

It can be implemented as

$$\begin{aligned} C = &\text{CNOT}_{21}\text{CNOT}_{12}\text{CNOT}_{13}\text{CNOT}_{14}H_1\text{CNOT}_{41}\text{CNOT}_{12}\text{CNOT}_{14}H_1\text{CNOT}_{41}H_2 \\ &\text{CNOT}_{21}\text{CNOT}_{12}\text{CNOT}_{13}S_1H_1\text{CNOT}_{21}\text{CNOT}_{31}\text{CNOT}_{12}\text{CNOT13}S_1H_1H_2 \\ &\text{CNOT}_{21}\text{CNOT}_{12}H_1\text{CNOT}_{21}\text{CNOT}_{12}H_1\text{CNOT}_{21}H_1. \end{aligned} \tag{7.6}$$

Finally, for $n = 5$ the best achievable fidelity is $F_{out} = 0.9621$, which can be achieved with success probability $p_{suc} = 0.2515$. A Clifford circuit with which these results can be obtained, can be found in the same way as the $n = 3$ and $n = 4$ circuits were found. This circuit, however, consists of many Clifford gates and does not provide any new insights, thus it is omitted here.

8

CONCLUSION

In this thesis bilocal Clifford circuits were studied from a group theoretical point of view. The main goal was to limit the search space for an optimal protocol by getting a better understanding of the structure of the Clifford group. This was achieved by finding a characterization of the subgroup of the Clifford group that preserves the distillation statistics. Starting from the Bruhat decomposition of the Clifford group, a generating set of gates for the subgroup was found. It was shown that, when optimizing the distillation statistics, it is sufficient to consider only one element of every right coset of this subgroup in the Clifford group, instead of every element of the Clifford group, which is an improvement upon previous work on the optimization of entanglement distillation protocols. By establishing a formula for the order of the subgroup, it was proved that this indeed significantly limits the search space.

Special emphasis was put to the representation of Clifford operations in terms of binary matrices from the symplectic group. This representation was proved to be 'onto': every element from the symplectic group corresponds to an element from the Clifford group. This result was used to develop an algorithm that calculates a transversal of the right cosets. For $n \geq 5$, however, this algorithm requires a very high random access memory. As an alternative, a probabilistic approach could be used, were protocols are sampled randomly and directly applied to an initial state. If the number of samples is high enough, an optimal protocol can be obtained with a certain probability. Time-wise, this takes about the same time as first calculating a transversal and then applying it to one initial state, but it requires less memory. However, once the transversal is calculated, the application to an initial state can be performed rather fast. Because various initial states were used in this thesis, the approach of first calculating the transversal was chosen.

For isotropic input states and a Bell diagonal state, the possible distillation statistics were calculated. Remarkably, it was found that for isotropic states the number of possible outcomes is much lower than the number of cosets. Although it is not very surpris-

ing that for a symmetric input state the number of outcomes decreases, the degree to which this happens indicates that there is great potential in further investigation of this problem. A better understanding of the consequences of this symmetry for distillation protocols can lead to a more efficient optimization for isotropic input states and may enable optimization for larger values on n. For the Bell diagonal state, a decrease in possible outcomes was found as well. However, this decrease was a lot less significant than the decrease that was found for isotropic states.

Lastly, the best achievable distillation statistics were calculated. For $n = 2$ it was found that the DEJMPS protocol yields the highest achievable fidelity with the best success probability possible. For $n = 3$ and $n = 4$ protocols that result in the highest achievable fidelity with the best success probability possible are given in both their symplectic form and as a Clifford circuit. Possibly, for these protocols a shorter Clifford circuit can be found, for instance using the Bruhat decomposition, but this is not further investigated in this thesis.

ACKNOWLEGDEMENTS

First and foremost, I would like to thank both my supervisors Dr. D. Elkouss Coronas and Dr. D. C. Gijswijt for the opportunity to work on this project and their guidance and supervision throughout the project. Additionally, I wish to pay my special regards to Kenneth Goodenough and Sébastian de Bone for their help and the insights and passion they shared with me. Thank you, also, to my parents and my sister for their unconditional support. Finally, I wish to thank Jort de Groot for his love and encouragement and for lending me his laptop, which has a larger random access memory than my laptop and made the calculations for $n = 5$ possible.

REFERENCES

Artin, E. (1957). In *Geometric algebra* (1st ed., pp. 143–147). Interscience Publishers New York.

Aspect, A., Grangier, P., & Roger, G. (1982). Experimental realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A new violation of Bell's inequalities. *Physical Review Letters*, *49*(2), 91–94.

Bell, J. S. (1964). On the Einstein Podolsky Rosen paradox. *Physics Physique*, *1*(3), 195–200.

Bennett, C. H., Bernstein, H. J., Popescu, S., & Schumacher, B. (1996a). Concentrating partial entanglement by local operations. *Physical Review A*, *53*(4), 2046–2052.

Bennett, C. H., Brassard, G., Popescu, S., Schumacher, B., Smolin, J. A., & Wootters, W. K. (1996c). Purification of noisy entanglement and faithful teleportation via noisy channels. *Physical Review Letters*, *76*(5), 722–725.

Bennett, C. H., Divincenzo, D. P., Smolin, J. A., & Wootters, W. K. (1996b). Mixed-state entanglement and quantum error correction. *Physical Review A*, *54*(5), 3824–3851.

Bravyi, S., & Maslov, D. (2020). Hadamard-free circuits expose the structure of the clifford group. *arXiv preprint arXiv:2003.09412v1 [quant-ph]*.

Calderbank, A. R., Rains, E. M., Shor, P. M., & Sloane, N. J. A. (1998). Quantum error correction via codes over $GF(4)$. *IEEE Transactions on Information Theory*, *44*(4), 1369–1387.

Dehaene, J., den Nest, M. V., Moor, B. D., & Verstraete, F. (2003b). Local permutations of products of bell states and entanglement distillation. *Physical Review A*, *67*(2).

Dehaene, J., & Moor, B. D. (2003a). Clifford group, stabilizer states, and linear and quadratic operations over GF(2). *Physical Review A*, *68*(4).

Deutsch, D., Ekert, A., Jozsa, R., Macchiavello, C., Popescu, S., & Sanpera, A. (1996). Quantum privacy amplification and the security of quantum cryptography over noisy channels. *Physical Review Letters*, *77*(13), 2818–2821.

Driver, B. K. (2003). *Analysis tools with applications* [Lecture notes]. University of California.

Einstein, A., Podolsky, B., & Rosen, N. (1935). Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, *47*(10), 777–780.

Fulton, W., & Harris, J. (2004). *Representation theory*. Springer New York.

Garcia-Escartin, J. C., & Chamorro-Posada, P. (2008). Equivalent quantum circuits, arXiv:1110.2998.

Gottesman, D. (1997). *Stabilizer codes and quantum error correction* (PhD Thesis). California Institute of Technology.

Gottesman, D. (1998). The Heisenberg representation of quantum computers, arXiv quant-ph/9807006.

Hayashi, M. (2006). In *Quantum information: An introduction* (pp. 119–121). Springer Berlin Heidelberg New York.

Horn, R., & Johnson, C. (1991). *Topics in matrix analysis*. Cambridge University Press.

Hostens, E. (2007). *Quantum entanglement distillation in the stabilizer formalism* (PhD Thesis). Katholieke Universiteit Leuven.

Maslov, D., & Roetteler, M. (2018). Shorter stabilizer circuits via bruhat decomposition and quantum circuit transformations. *IEEE Transactions on Information Theory*, *64*(7), 4729–4738.

Nielsen, M. A., & Chuang, I. L. (2016). *Quantum computation and quantum information*. Cambridge University Press.

Ozols, M. (2008). Clifford group [Online; accessed 6 June 2020]. http://home.lu.lv/~sd20008/papers/essays/Clifford%5C%20group%5C%20[paper].pdf

Plenio, M. B., & Virmani, S. (2007). An introduction to entanglement measures. *Quantum Information & Computation, 7*, 1–51.

Rotman, J. J. (1995). In *An Introduction to the Theory of Groups* (4th ed., pp. 7–8). Springer New York.

Rozpędek, F., Schiet, T., Thinh, L. P., Elkouss, D., Doherty, A. C., & Wehner, S. (2018). Optimizing practical entanglement distillation. *Physical Review A, 97*(6).

Triebel, H. (1986). In *Analysis and mathematical physics* (1st ed., p. 216). D. Reidel Publishing Company.

A

EQUIVALENT QUANTUM CIRCUITS

A.1. PROOF THAT $CZ_{ij} \in \langle S^D \rangle$

It was claimed in Figure 5.3 that a CZ_{ij} gate can be rewritten in terms of CZ_{1i}, CZ_{1j} and H_1 gates. The equivalence of the circuits of Figure 5.3 is proved here. In the proof the distributed CNOT operation (Figure A.1a) is used. A proof of this equality can be found in Garcia-Escartin and Chamorro-Posada, 2008. Moreover the CNOT-CZ transformation rule (Figure A.1b) is used. Although this rule has been used before in this thesis, it is repeated here for convenience.

(a) (b)

Figure A.1: Distributed CNOT operation (a) and a CNOT gate rewritten in terms of two Hadamard gates and a CZ gate (b).

A proof of the equivalence in Figure 5.3 is shown in Figure A.2.

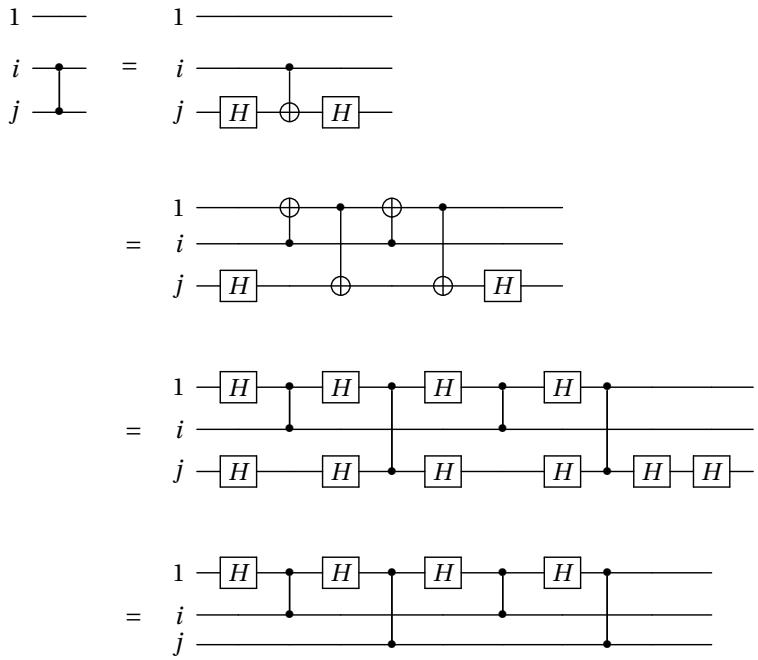


Figure A.2: Proof that CZ_{ij} can be rewritten in terms of H_1, CZ_{1i} and CZ_{1j}.

B

IMPLEMENTATIONS IN SAGEMATH

In this appendix implementations in SageMath of the algorithms described in Chapter 6 can be found. All implementations, including the code used for the figures of Chapter 7, can also be found at https://github.com/sarahjansen08/optimization-entanglement-distillation.

B.1. FINDING A TRANSVERSAL

```
import numpy as np
import itertools as it

def base(M):
    # calculate the image of the base under a matrix M
    s = list([M[0:2*n, i] for i in range(n+1, 2*n)])
    powerset = it.chain.from_iterable(it.combinations(s, r) for r in range
        (1, len(s)+1))
    base_map = set()
    for i in powerset:
        v = vector(sum(i))
        v.set_immutable()
        base_map.add(v)
    return frozenset(base_map)

# number of qubits (n > 1)
n = 2
# calculate number of cosets
prod = 1
for i in range(1,n+1):
    prod = prod * (2**i + 1)
index = 1/3 * (2**n - 1) * prod
# load symplectic group
Cn = Sp(2*n, GF(2))
```

```
# calculate matrix needed for transformation from 'sage symplectic' to \\ '
    literature symplectic'
anti_identity = matrix.identity(n)
for i in range(0, (n/2).ceil()):
    anti_identity.swap_rows(i, n-i-1)
T = block_matrix(GF(2), [[matrix.identity(n), zero_matrix(n,n)], [
    zero_matrix(n,n), anti_identity]], subdivide = False)

transversal_inv = {}
while len(transversal_inv)<index:
    # generate random element of symplectic group
    M = T * Cn.random_element() * T
    # calculating M inverse
    A = M[0:n, 0:n]
    B = M[0:n, n:2*n]
    C = M[n:2*n, 0:n]
    D = M[n:2*n, n:2*n]
    M_inv = block_matrix([[D.transpose(),-B.transpose()],[-C.transpose(),A.
        transpose()]], subdivide=False)
    M_basecol = base(M_inv)
    # check whether coset is already in transversal
    if M_basecol not in transversal_inv:
        transversal_inv[M_basecol] = M_inv
```

B.2. CALCULATING DISTILLATION STATISTICS

```
import numpy as np
import itertools as it

m = 2
# Load transversal
transversal_inv = load('2_transversal_inv.sobj')

# FUNCTIONS

def base(M, n):
    # calculate the image of the base under a matrix M
    s = []
    for i in range(n+1, 2*n):
        s.append(M[0:2*n, i])
    powerset = it.chain.from_iterable(it.combinations(s, r) for r in range
        (1, len(s)+1))
    res = [vector(GF(2),2*n)]
    for i in powerset:
        v = vector(sum(i))      # calculate the sum of the elements of each
            combination (e.g IZZ = IZI + IIZ)
        res.append(v)
    return res

def pillars(M, n):
```

```
    # calculate the image of the pillars under a matrix M
    X1 = vector(M[0:2*n, 0])
    Z1 = vector(M[0:2*n, n])
    Y1 = X1 + Z1
    pI = base(M, n)
    pX = [(X1 + b) for b in pI]
    pY = [(Y1 + b) for b in pI]
    pZ = [(Z1 + b) for b in pI]
    return [pI, pX, pY, pZ]

def tensor(A, n):
    # calculate the n fold tensor product of a matrix A
    kron = A
    count = 1
    while count < n:
        kron = np.kron(kron,A)
        count = count + 1
    if n == 2:
        res = np.reshape(kron, (4,4))
    elif n == 3:
        res = np.reshape(kron, (4,4,4))
    elif n == 4:
        res = np.reshape(kron, (4,4,4,4))
    elif n == 5:
        res = np.reshape(kron, (4,4,4,4,4))

def dist_stat(initial, M, n):
    # returns the success probability, fidelity and rate
    pil = pillars(M, n)
    out = []
    for layer in pil:
        coef = 0
        for elt in layer:
            if n == 2:
                coef = coef + initial[int(elt[0]) + 2*int(elt[n]), int(elt
                    [1]) + 2*int(elt[n+1])]
            if n == 3:
                coef = coef + initial[int(elt[0]) + 2*int(elt[n]), int(elt
                    [1]) + 2*int(elt[n+1]), \
                                int(elt[2]) + 2*int(elt[n+2])]
            if n == 4:
                coef = coef + initial[int(elt[0]) + 2*int(elt[n]), int(elt
                    [1]) + 2*int(elt[n+1]), \
                                int(elt[2]) + 2*int(elt[n+2]), int(elt
                                    [3]) + 2*int(elt[n+3])]
            if n == 5:
                coef = coef + initial[int(elt[0]) + 2*int(elt[n]), int(elt
                    [1]) + 2*int(elt[n+1]), \
                                int(elt[2]) + 2*int(elt[n+2]), int(elt
```

```
                                                [3]) + 2*int(elt[n+3]), \
                                              int(elt[4]) + 2*int(elt[n+4])]
            out.append(coef)
        sp = sum(out)
        fid = out[0]/sp
        r = 1
        for i in out:
            r = r + float(i * log(i) / log(2))
        r = r*sp/n
        return sp, fid, r
        return res

def best_protocol(initial, transversal_inv, n, measure = "fidelity"):
        # calculates the best protocol from a dictionary of inverses of
            protocols (transversal_inv) applied to an initial state; as quality
            measures the fidelity ("fidelity"), success probability ("sucprob")
            or rate ("rate") can be chosen
        if measure == "sucprob":
            res = 0
            for key, M in transversal_inv.items():
                s = (dist_stat(initial, M, n))[0]
                if s > res:
                    res = s
                    opt_inv = M
        if measure == "fidelity":
            res = 0
            for key, M in transversal_inv.items():
                f = (dist_stat(initial, M, n))[1]
                if f > res:
                    res = f
                    opt_inv = M
        if measure == "rate":
            res = -100
            for key, M in transversal_inv.items():
                r = (dist_stat(initial, M, n))[2]
                if r > res:
                    res = r
                    opt_inv = M
        # Calculate inverse of optimal protocol
        A = opt_inv[0:n, 0:n]
        B = opt_inv[0:n, n:2*n]
        C = opt_inv[n:2*n, 0:n]
        D = opt_inv[n:2*n, n:2*n]
        opt = block_matrix([[D.transpose(),-B.transpose()],[-C.transpose(),A.
            transpose()]], subdivide=False)
        return res, opt

def sucprob_fid_lists(initial, transversal_inv, n):
        # calculate the possible distillation statistics (success probability &
```

```
        fidelity) of the protocols in a transversal applied to an initial
        state
    fid = []
    sp = []
    fslist = []
    for key, M in transversal_inv.items():
        s = dist_stat(initial, M, n)[0]
        f = dist_stat(initial, M, n)[1]
        if (s,f) not in fslist:
            sp.append(s)
            fid.append(f)
            fslist.append((s,f))
    return sp, fid

# INPUT STATES

# isotropic input state
fid_in = [0.55, 0.575, 0.6, 0.625, 0.65, 0.675, 0.7, 0.725, 0.75, 0.775,
    0.8, 0.825, 0.85, 0.875, 0.9, 0.925, 0.95, 0.975]
init_iso = []
for f in fid_in:
    init_iso.append(vector([f, (1-f)/3, (1-f)/3, (1-f)/3]))

# Bell diagonal state
p0 = np.random.uniform(0.5, 1)
p1 = np.random.uniform(0, 1-p0)
p2 = np.random.uniform(0, 1-p0-p1)
p3 = 1 - p0 - p1 - p2
lst = [p0, p1, p2, p3]
lst.sort()
init_bell = vector([lst[3], lst[2], lst[1], lst[0]])
print(init_bell)
#init_bell = vector([0.752604752074101, 0.10567340022267524,
    0.0938043735551225, 0.04791747414810121])

for i in init_iso:
    print(sucprob_fid_lists(tensor(i, m), transversal_inv, m))

print(sucprob_fid_lists(tensor(init_bell, m), transversal_inv, m))

r_list = []
for i in init_iso:
    r_list.append((best_protocol(tensor(i, m), transversal_inv, m, measure =
        "rate"))[0])
print(r_list)
```