PERMUTATION GATES IN THE THIRD LEVEL OF THE CLIFFORD HIERARCHY

ZHIYANG HE, LUKE ROBITAILLE, AND XINYU TAN

Abstract. The Clifford hierarchy is a fundamental structure in quantum computation, classifying unitary operators based on their commutation relations with the Pauli group. Despite its significance, the mathematical structure of the hierarchy is not well understood at the third level and higher. In this work, we study permutation gates in the hierarchy – unitaries which permute the 2^n basis states – and present several classification results. In particular, we prove that any permutation gate in the third level must be a product of Toffoli gates in what we define as *staircase form*, up to left and right multiplications of Clifford permutations. We then present sufficient and necessary conditions for a staircase form permutation gate to be in the third level of the Clifford hierarchy. As a corollary, we construct a family of non-semi-Clifford permutation gates $\{U_k\}_{k\geq 3}$ in staircase form such that each U_k is in the third level but its inverse is *not* in the k-th level.

We remark that this is a preliminary version of our manuscript. We intend to polish the discussions further before posting to arXiv.

1. Introduction

The study of the Clifford hierarchy originated from the quest for universal and fault-tolerant quantum computation. It is well-known that the Clifford group is not universal, and any circuit made completely of Clifford gates can be efficiently simulated by a classical computer [Got98]. However, adding any non-Clifford gate to the Clifford group forms a universal gate set. These fundamental results place non-Clifford gates in an unique position: to realize useful quantum computation in practice, we must have fault-tolerant implementation of non-Clifford gates.

The easiest way to implement a single gate fault-tolerantly is to use a quantum error-correcting code for which the corresponding logical operation is transversal. However, the Eastin–Knill theorem states that no quantum-error correcting code can implement a universal gate set transversally [EK09]. This no-go result prompts the search of new fault-tolerance techniques. In a seminal work in 1999 [GC99], Gottesman and Chuang defined the *Clifford hierarchy* and proposed a simple protocol—*gate teleportation*—to fault-tolerantly implement gates at any level of this hierarchy. It remains one of the leading approaches to universal fault-tolerant quantum computation.

Mathematically, the Clifford hierarchy has a simple recursive definition. The first two levels $\mathcal{C}_1$ and $\mathcal{C}_2$ are the Pauli and Clifford groups respectively. For $k \geq 3$, the k-th level $\mathcal{C}_k$ is the set of all unitaries U such that $UPU^\dagger \in \mathcal{C}_{k-1}$ for all $P \in \mathcal{P}_n$. As a result, $\mathcal{C}_{k-1} \subseteq \mathcal{C}_k$. Gates at higher levels require more resources to implement via gate teleportation, so $\mathcal{C}_3$ contains the "cheapest" non-Clifford gates. This includes many commonly used examples, such as the T, Toffoli, and CCZ gates. Motivated by both its applications and the simplicity of its definition, the Clifford hierarchy has been the subject of many attempts to understand its mathematical structure [ZLC00; CGK17; BS09; And24; AC25]. We refer readers to Section 1.2 for a brief review of these prior works. However, $\mathcal{C}_k$ remains largely mysterious for all $k \geq 3$. There is no known closed-form characterization; $\mathcal{C}_k$ no longer forms a group[1] and its precise set of gates is unknown.

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
E-mail address: szhe,lrobitai,norahtan@mit.edu.

[1]Note that $\mathcal{C}_2$, the Clifford group, is the largest finite subgroup (up to phase) of the unitary group that contains the Clifford group [NRS00, Theorem 6.5]. Any other subgroups containing the Clifford group must be dense in the

In this paper, we characterize permutation gates in the Clifford hierarchy at level three. A permutation gate on n qubits permutes the 2^n computational basis states. Intuitively, $\mathcal{C}_k$ permutation gates are simpler to study but also sufficiently general, since permutation gates represent all classical reversible circuits. Our main results are as follows.

1.1. **Main results.** A Toffoli gate denoted as $\mathrm{TOF}_{i,j,k}$ is a NOT gate targeting at qubit k controlled by qubits i and j. We define a product of Toffoli gates to be in *staircase form* if each gate $\mathrm{TOF}_{i,j,k}$ in the product has $i, j < k$ and the target qubits are in nondecreasing order in the order that the gates are applied. For example, the gates in Figures 2 and 3 are in staircase form. We first show that all $\mathcal{C}_3$ permutation gates can be written in this staircase form up to left and right multiplications of Clifford permutations.

Result 1 (Theorem 3.2)**.** For any $\mathcal{C}_3$ permutation gate π, there exist Clifford permutations ϕ_1, ϕ_2 and a product μ of Toffoli gates in *staircase form* such that $\pi = \phi_1 \cdot \mu \cdot \phi_2$.

Next, we show that a staircase form permutation naturally induces a multiplication operation on vectors in $\mathbb{F}_2^n$. If this multiplication satisfy certain conditions (see Definition 4.1), we call it a *descending multiplication* and show that such operations are in one-to-one correspondence with staircase form permutations in $\mathcal{C}_3$.

Result 2 (Theorem 4.2)**.** There is a one-to-one correspondence between descending multiplications and $\mathcal{C}_3$ permutations in staircase form.

Utilizing this characterization, we construct a family of permutation gates $\{U_k\}_{k\geq 3}$ such that $U_k \in \mathcal{C}_3$ but $U_k^\dagger \notin \mathcal{C}_k$. This is not only mathematically interesting but also has an operational consequence. In gate teleportation, gates at higher levels of the Clifford hierarchy are more resource-intensive. Thus what we show is that there exist gates that are cheap to implement fault-tolerantly (level-three), yet whose inverses are expensive (not level-three). Moreover, the inverse can be made arbitrarily costly: our constructions yield $\mathcal{C}_3$ gates whose inverses lie at any prescribed level of the Clifford hierarchy.

Result 3 (Definition 5.2 and Theorem 5.5)**.** We construct a family of non-semi-Clifford $\mathcal{C}_3^{\mathrm{sym}}$ permutations $\{U_k\}_{k\geq 3}$ where each U_k acts on $2^k - 1$ qubits and $U_k^\dagger \notin \mathcal{C}_k$. We also show that its qubit count is optimal – any permutations which satisfy these properties must act on at least $2^k - 1$ qubits.

We give a computer-assisted proof in the appendix that all 6-qubit $\mathcal{C}_3^{\mathrm{sym}}$ gates are semi-Clifford and hence U_3 which acts on 7 qubits is the smallest non-semi-Clifford $\mathcal{C}_3^{\mathrm{sym}}$ gate.

1.2. **Prior works.** Since its first proposal by Gottesman and Chuang in 1999 [GC99], there have been many attempts in understanding the mathematical structure of the Clifford hierarchy.

In [ZLC00], Zhou, Leung, and Chuang proposed to study the diagonal gates in the hierarchy, which do form a group at every level. Gates in the form $\phi_1 d \phi_2$, where ϕ_1, ϕ_2 are Clifford gates and d is a diagonal gate, are later referred to as *semi-Clifford* operations. Cui, Gottesman, and Krishna fully characterized all the diagonal gates in $\mathcal{C}_k$ [CGK17], and thus all the semi-Clifford unitaries in the hierarchy as well.

It was once believed that $\mathcal{C}_3$ should behave as nicely as the diagonal gates in the hierarchy: Zeng, Chen, and Chuang conjectured in [ZCC08, Conjecture 1] that all gates in $\mathcal{C}_3$ are semi-Clifford. However, this conjecture was shown to be false by Gottesman and Mochon with a carefully constructed counterexample on seven qubits, which consists of three controlled SWAP gates and four CCZ gates. Before our paper, this counterexample was the only known $\mathcal{C}_3$ gate that is not

unitary group. So if we force a definition where each level of the hierarchy must form a finite group up to phase, then $\mathcal{C}_k$ can only be the Clifford group for all $k \geq 2$, which is less interesting.

semi-Clifford.[2] Beigi and Shor later proved that all gates in $\mathcal{C}_3$ are in fact *generalized semi-Clifford* gates [BS09], which adopt the form $\phi_1 \pi d \phi_2$ for some Clifford gates ϕ_1, ϕ_2, a diagonal gate d, and a permutation π. The conjecture that all gates in $\mathcal{C}_k$ are generalized semi-Clifford remains open [ZCC08, Conjecture 2].

In this paper, we focus on characterizing the permutation gates in the Clifford hierarchy, denoted as $\mathcal{C}_k^{\text{sym}}$ for each level k. Anderson studied $\mathcal{C}_k^{\text{sym}}$ in [And24] and made a conjecture (Conjecture D.1), which, if true, would imply that all gates in $\mathcal{C}_3^{\text{sym}}$ are semi-Clifford. We disprove Anderson's conjecture by construct an infinite family of non-semi-Clifford permutation $\mathcal{C}_3$ gates.

1.3. **Organization.** The content of the paper is divided into sections as follows. Section 2 gives background results and lemmas for later use. In particular, Section 2.3 discusses our perspective of seeing permutation gates as polynomials over $\mathbb{F}_2$. In Section 3, we present the definition for staircase form, in which all permutations in $\mathcal{C}_3$ can be written (up to Clifford permutations). In Section 4, we define descending multiplications and prove their one-to-one correspondance with staircase form permutations in $\mathcal{C}_3$. In Section 5, we construct our family of non-semi-Clifford gates $\{U_k\}_{k\geq 3}$ where each permutation U_k is in staircase form. We prove in Section 5.1 that each $U_k \in \mathcal{C}_3$ but $U_k^\dagger \notin \mathcal{C}_k$. In particular, the smallest example in this family U_3 is a 7-qubit permutation, and it is conjugate to the Gottesman–Mochon example by a Clifford operator (Section 5.2).

Appendix A classifies semi-Clifford permutations and where they appear in the Clifford hierarchy. In Appendix B, we detail a computer search of six-qubit permutations, assisted by the staircase form of Section 3, to show that 7 is the smallest number of qubits for which there exists a non-semi-Clifford permutation in $\mathcal{C}_3$.

2. Preliminaries

The single-qubit Pauli gates I_2, X, Y, and Z are given by

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

The *Pauli group* on n qubits, denoted as $\mathcal{P}_n$, is the collection of all gates of the form $cP_1 \otimes P_2 \otimes \cdots \otimes P_n$ for $c \in \{\pm 1, \pm i\}$ and one-qubit gates $P_1, \ldots, P_n \in \{I_2, X, Y, Z\}$. In particular, we denote the set of all the n-qubit Pauli X operators as $\mathcal{X} = \{I_2, X\}^{\otimes n}$. Note that $\mathcal{X}$ is exactly the set of all permutation gates in $\mathcal{P}_n$.

We frequently use the *Hadamard*, *controlled* NOT, and *Toffoli* gates throughout the paper:

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \ \text{CNOT}_{1,2} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \ \text{TOF}_{1,2,3} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

We can view the action of CNOT as $|a_1\rangle \otimes |a_2\rangle \mapsto |a_1\rangle \otimes |a_1 + a_2\rangle$ where the first qubit is the *control* and the second qubits is the *target*. Similarly, we can view the Toffoli gate as $|a_1\rangle \otimes |a_2\rangle \otimes |a_3\rangle \mapsto |a_1\rangle \otimes |a_2\rangle \otimes |a_3 + a_1 a_2\rangle$ where the first two qubits are controls and the third is the target. We use subscripts to denote the qubits that a gate acts upon. For example, Y_4 is a Pauli Y gate acting on the fourth qubit, and $\text{CNOT}_{3,1}$ is a CNOT gate with the third qubit as control and the first qubit as target.

[2]Here we consider the trivial generalizations of this counterexample, such as multiplication by a Clifford, as the same counterexample.

The *Clifford group* on n qubits is the normalizer of the Pauli group in the unitary group. It can be generated by the Pauli group, the Hadamard and phase gate on each qubit, the CNOT gate on each ordered pair of distinct qubits, and $\{cI : |c| = 1\}$. Henceforth we refer to elements of $\{cI : |c| = 1\}$ as *phases* (not to be confused with the phase gate).

2.1. **The Clifford hierarchy, semi-Clifford, and generalized semi-Clifford.**

Definition 2.1. Let n be the number of qubits. Let $\mathcal{C}_1 = \mathcal{P}_n$. For $k \geq 2$, inductively define $\mathcal{C}_k$ to be the set of all unitaries U such that $UPU^\dagger \in \mathcal{C}_{k-1}$ for all $P \in \mathcal{P}_n$. Note that $\mathcal{C}_2$ is the Clifford group. The set $\mathcal{CH} := \mathcal{C}_1 \cup \mathcal{C}_2 \cup \mathcal{C}_3 \cup \cdots$ is called the *Clifford hierarchy*; we refer to $\mathcal{C}_k$ as the k-th layer of $\mathcal{CH}$.

The following propositions about the Clifford hierarchy are standard.

Proposition 2.2.

(1) For any k, $\mathcal{C}_k$ is finite up to phase and $\mathcal{C}_k \subseteq \mathcal{C}_{k+1}$.
(2) For $k \geq 2$, $\mathcal{C}_k$ is closed under left and right multiplication of Clifford gates.
(3) For $k \geq 3$, $\mathcal{C}_k$ is not a group.
(4) For any k, $\mathcal{C}_k$ is closed under complex conjugation.

We say that a gate is a *permutation gate* if it corresponds to a $2^n \times 2^n$ permutation matrix. Note that this is different from only permuting the qubits. A gate is called *diagonal* if its associated matrix is diagonal.

Definition 2.3. A gate is *semi-Clifford* if it can be written as $\phi_1 d \phi_2$ for some Clifford gates ϕ_1, ϕ_2 and a diagonal gate d. A gate is *generalized semi-Clifford* if it can be written as $\phi_1 \pi d \phi_2$ for some Clifford gates ϕ_1, ϕ_2, a permutation gate π, and a diagonal gate d.

Observe that the inverse of a semi-Clifford gate is semi-Clifford. The inverse of a generalized semi-Clifford gate is generalized semi-Clifford, as we can write $(\phi_1 \pi d \phi_2)^{-1} = \phi_2^{-1} \pi^{-1} (\pi d^{-1} \pi^{-1}) \phi_1^{-1}$, and $\pi d^{-1} \pi^{-1}$ is diagonal. If we multiply a semi-Clifford (resp. generalized semi-Clifford) element on the left or right by a Clifford gate, the resulting operator is still semi-Clifford (resp. generalized semi-Clifford).

For a maximal abelian subgroup A of $\mathcal{P}_n$, let $\operatorname{span}(A)$ denote its linear span with complex coefficients.

Lemma 2.4. *An operator U is semi-Clifford if and only if there exist maximal abelian subgroups A_1 and A_2 of $\mathcal{P}_n$ such that $UA_1U^\dagger = A_2$. An operator U is generalized semi-Clifford if and only if there exist maximal abelian subgroups A_1 and A_2 of $\mathcal{P}_n$ such that $U \operatorname{span}(A_1) U^\dagger = \operatorname{span}(A_2)$.*[3]

2.2. **Not all gates in $\mathcal{C}_3$ are semi-Clifford.** Gottesman and Mochon showed that there exists a 7-qubit unitary in $\mathcal{C}_3$ that is not semi-Clifford [BS09].

Proposition 2.5. *For any k, the inverse of any semi-Clifford element of $\mathcal{C}_k$ is in $\mathcal{C}_k$.*

Proof. For any $U \in \mathcal{C}_k$ that is semi-Clifford, by Definition 2.3, we can write $U = \phi_1 d \phi_2$ for some Clifford gates ϕ_1, ϕ_2 and a diagonal gate d. Using Proposition 2.2 repeatedly, we know that $d = \phi_1^{-1} U \phi_2^{-1} \in \mathcal{C}_k$. Hence, $d^{-1} = d^\dagger = \overline{d} \in \mathcal{C}_k$ and thus $U^{-1} = \phi_2^{-1} d^{-1} \phi_1^{-1} \in \mathcal{C}_k$. □

Lemma 2.6. *For $n = 7$, $\mathcal{C}_3$ contains a non-semi-Clifford element.*

Proof. Let G be a 7-qubit gate given by

$$G = \mathrm{CSWAP}_{7,1,6} \mathrm{CSWAP}_{7,2,5} \mathrm{CSWAP}_{7,4,3} \cdot \mathrm{CCZ}_{1,2,3} \mathrm{CCZ}_{1,4,5} \mathrm{CCZ}_{2,4,6} \mathrm{CCZ}_{3,5,6},$$

where CSWAP denotes the controlled SWAP gate and CCZ denotes the controlled controlled Z gate. See Figure 1 for a circuit diagram.

[3] In most literature, semi-Clifford and generalized semi-Clifford are usually defined as in Lemma 2.4 whereas Definition 2.3 is proved as a proposition.

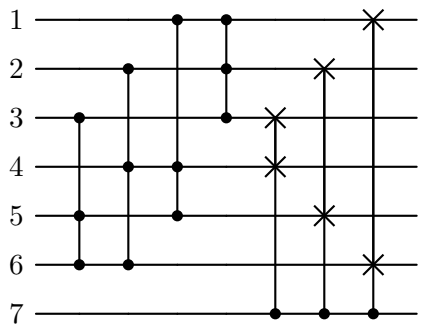

FIGURE 1. Circuit diagram for Gottesman–Mochon seven-qubit gate G (with time flowing from left to right).

It can be verified with a computer program that $G \in \mathcal{C}_3$. If G were semi-Clifford, then we would have $G^{-1} \in \mathcal{C}_3$ by Proposition 2.5. However, a computer calculation shows that $G^{-1} \notin \mathcal{C}_3$ (in particular, $U^{-1}X_7U \notin \mathcal{C}_2$). Thus, G is not semi-Clifford. □

The above 7-qubit operator is the smallest known example of a non-semi-Clifford operator in $\mathcal{C}_3$. It was shown in [ZCC08] that for $n = 3$, all elements of $\mathcal{C}_3$ are semi-Clifford. For $n = 4, 5, 6$, it is an open problem whether there is a $\mathcal{C}_3$ operator that is non-semi-Clifford.

Beigi and Shor proved in [BS09] the following theorem on $\mathcal{C}_3$.

Theorem 2.7. *Every element of $\mathcal{C}_3$ is generalized semi-Clifford.*

The following conjecture made in [ZCC08] remains open.

Conjecture 2.8. *Every element of the Clifford hierarchy is generalized semi-Clifford.*

Remark 2.9. Recall that a generalized semi-Clifford gate can be written as $\phi_1\pi d\phi_2$ where ϕ_1 and ϕ_2 are Cliffords, π is a permutation, and d is diagonal. A similar form of $\pi d\phi$ is considered in the context of approximate unitary designs or pseudorandom unitaries in [MPSY24; CHH+24], where ϕ and π are sampled uniformly at random from their respective groups, and d is a diagonal gate with random ± 1 entries.

2.3. **The polynomial viewpoint.** The lemmas in this section are not hard to prove. Nevertheless, they provide a crucial perspective for studying permutation gates in the Clifford hierarchy.

Lemma 2.10. *Any function from $\mathbb{F}_2^n$ to $\mathbb{F}_2$ can be uniquely written as an n-variable polynomial that has degree at most 1 in each variable.*

Theorem 2.11. *For any function $f : \mathbb{F}_2^n \to \mathbb{F}_2$, the diagonal gate $\sum_{a\in\mathbb{F}_2^n}(-1)^{f(a)}|a\rangle\langle a|$ is in $\mathcal{C}_k$ if and only if f, considered as a polynomial, has degree at most k.*

Proof. This is a special case of the main theorem in [CGK17, See Eq. (1)]. □

Definition 2.12 (Polynomial representation)**.** Given a permutation gate $\pi : \mathbb{F}_2^n \to \mathbb{F}_2^n$, let $\pi_i : \mathbb{F}_2^n \to \mathbb{F}_2$ denote the function π restricted to the i-th output bit, i.e.

$$\pi = \sum_{a\in\mathbb{F}_2^n} |\pi_1(a), \ldots, \pi_n(a)\rangle\langle a|.$$

From Lemma 2.10 we know that each π_i can be written as a polynomial in the input bits. We refer to $(\pi_1, \ldots, \pi_n)$ as the polynomial representation of π and π_i as the i-th coordinate of π.

As an example, $\mathrm{TOF}_{1,2,3}$ can be represented as $(a_1, a_2, a_3) \mapsto (a_1, a_2, a_3 + a_1a_2)$.

Lemma 2.13. *For any positive integer k and permutation gate $\pi \in \mathcal{C}_{k+1}$, each coordinate of π^{-1} has degree at most k.*

Proof. For each $i \in [n]$, we have

$$\mathcal{C}_k \ni \pi Z_i \pi^{-1} = \sum_{a \in \mathbb{F}_2^n} (-1)^{a_i} |\pi(a)\rangle\langle\pi(a)| = \sum_{a \in \mathbb{F}_2^n} (-1)^{(\pi^{-1}(a))_i} |a\rangle\langle a| = \sum_{a \in \mathbb{F}_2^n} (-1)^{\pi_i^{-1}(a)} |a\rangle\langle a|.$$

It follows from Theorem 2.11 that π_i^{-1}, the i-th coordinate of π^{-1}, must have degree at most k. □

Remark 2.14. For $\pi \in \mathcal{C}_3$, Lemma 2.13 tells us that every coordinate of π^{-1} has degree at most 2; however, as we will see in Remark 5.7, the coordinates of π themselves do not necessarily have degree at most 2.

We denote by $e_1, \ldots, e_n$ the standard basis of $\mathbb{F}_2^n$.

Lemma 2.15. *For any $n \times n$ invertible matrix M over $\mathbb{F}_2$ and any vector w, the permutation gate sending $|v\rangle \mapsto |Mv + w\rangle$ is Clifford. Conversely, any Clifford permutation is of this form for some M and w.*

Proof. For the first part, it is clear that $|v\rangle \mapsto |Mv + w\rangle$ is a permutation. Call it π. To show that π is Clifford, it suffices to show that $\pi X_i \pi^{-1}, \pi Z_i \pi^{-1} \in \mathcal{P}_n$. For $\pi X_i \pi^{-1}$, it sends

$$|v\rangle \mapsto |M^{-1}(v - w)\rangle \mapsto |M^{-1}(v - w) + e_i\rangle \mapsto |M(M^{-1}(v - w) + e_i) + w\rangle = |v + Me_i\rangle,$$

which is a product of X gates. For $\pi Z_i \pi^{-1}$, it sends

$$|v\rangle \mapsto |M^{-1}(v - w)\rangle \mapsto (-1)^{e_i^\top M^{-1}(v-w)} |M^{-1}(v - w)\rangle \mapsto (-1)^{e_i^\top M^{-1}(v-w)} |v\rangle.$$

We can rewrite this as $|v\rangle \mapsto (-1)^{-e_i^\top M^{-1} w} (-1)^{((M^{-1})^\top e_i)^\top v} |v\rangle$, so this is a product of Z operators up to a phase of ± 1. Hence, we have $\pi \in \mathcal{C}_2$.

For the second part, we have π^{-1} is a permutation in $\mathcal{C}_2$ (as $\mathcal{C}_2$ is a group). Using Lemma 2.13 with $k = 1$, every coordinate of $(\pi^{-1})^{-1} = \pi$ has degree at most 1. This directly yields a matrix M and vector w so that π can be written as $|v\rangle \mapsto |Mv + w\rangle$. Since π is a permutation, M must be invertible. □

2.4. **Abelian subgroups of $\mathcal{P}_n$.** Suppose an n-qubit unitary $U \in \mathcal{C}_k$ is not semi-Clifford. It is trivial to see the $(n+1)$-qubit unitary $U \otimes I_2$ is in $\mathcal{C}_k$, but it is not completely trivial to conclude that $U \otimes I_2$ is also not semi-Clifford. It is sometimes glossed over in the literature (e.g. the proof of [ZCC08, Theorem 3]). We here provide a more careful treatment of this fact. Some lemmas will be used again in later sections.

Lemma 2.16. *Any maximal abelian subgroup of $\mathcal{P}_n$ is isomorphic to $(\mathbb{Z}/2\mathbb{Z})^n$ up to phase.*

Lemma 2.17. *Let A be a maximal abelian subgroup of $\mathcal{P}_n$, and let B be a (not necessarily maximal) abelian subgroup of $\mathcal{P}_n$. Then there exists a maximal abelian subgroup A' of $\mathcal{P}_n$ such that $B \subseteq A' \subseteq \langle A, B\rangle$.*

Proof. We first consider the case where B is generated by a single operator b (up to phase). Let $\{a_1, \ldots, a_n\}$ be the generators of A up to phase. Without loss of generality, suppose $a_1, \ldots, a_k$ are the generators of A which anti-commute with b. Consider sequential pairwise products of the form $a_1a_2, a_2a_3, a_3a_4, \ldots, a_{k-1}a_k$, and let A' be the group generated by $\{b, a_1a_2, \ldots, a_{k-1}a_k, a_{k+1}, \ldots, a_n\}$ (up to $\{\pm 1, \pm i\}$ phase). We see that A' is a maximal abelian subgroup of $\mathcal{P}_n$ and $B \subseteq A' \subseteq \langle A, B\rangle$.

In the case where B is generated by operators $b_1, \ldots, b_k$, we iteratively update A' for every generator of B with the above procedure. Note that the update procedure can be seen to preserve all elements of A that commute with b. Thus, at every update, we keep all generators of B that were already added (as B is abelian); thus we obtain the desired subgroup. □

Lemma 2.18. *Suppose U is an n-qubit non-semi-Clifford gate. Then $U \otimes I_{2^m}$ is also not semi-Clifford on $n+m$ qubits for any positive integer m.*

Proof. We show the contrapositive. Suppose $U' = U \otimes I_{2^m}$ is semi-Clifford. Consider the subgroup G of $\mathcal{P}_{n+m}$ consisting of all P such that $U'P(U')^{-1} \in \mathcal{P}_{n+m}$. We know by Lemma 2.4 that G contains a maximal abelian subgroup A of $\mathcal{P}_{n+m}$. Let $B = \langle Z_{n+1}, \ldots, Z_{n+m} \rangle \subseteq G$. Using Lemma 2.17, we get a maximal abelian subgroup A' of $\mathcal{P}_{n+m}$ such that $B \subseteq A' \subseteq \langle A, B \rangle \subseteq G$. Thus, there exists a subgroup A_1 of $\mathcal{P}_n$ with $A' = \langle A_1, B \rangle$. We can see that A_1 must have at least 2^n elements up to phase, so it must be a maximal abelian subgroup of $\mathcal{P}_n$. So $UA_1U^{-1} \subseteq \mathcal{P}_n$, which means that U is semi-Clifford. □

It follows from Lemmas 2.6 and 2.18 that $\mathcal{C}_3$ contains a non-semi-Clifford element for any $n \geq 7$.

3. Every $\mathcal{C}_3$ permutation is in staircase form up to Clifford permutations

Definition 3.1. A product of pairwise distinct Toffoli gates is said to be in *staircase form* if each gate $\mathrm{TOF}_{i,j,k}$ in the product has $i < j < k$ and the target qubits are in nondecreasing order in the order that the gates are applied. (See Figure 2 for an example.)

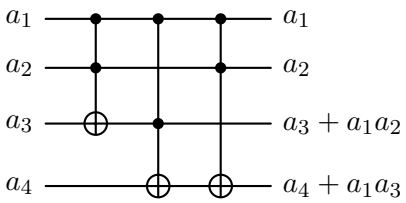

FIGURE 2. This circuit for $\mathrm{TOF}_{1,2,4}\mathrm{TOF}_{1,3,4}\mathrm{TOF}_{1,2,3}$ is in staircase form but not mismatch-free, as qubit 3 is used as a control for $\mathrm{TOF}_{1,3,4}$ and a target for $\mathrm{TOF}_{1,2,3}$.

The main result of this section is the following.

Theorem 3.2. *Suppose $\pi \in \mathcal{C}_3$ is a permutation gate. Then there exist Clifford permutations ϕ_1, ϕ_2 and a staircase form permutation $\mu \in \mathcal{C}_3$ such that $\pi = \phi_1 \mu \phi_2$.*

We caution that the statement of Theorem 3.2, unlike Theorem A.6, is not an if-and-only-if statement. In particular, note that $\mathrm{TOF}_{3,4,5}\mathrm{TOF}_{1,2,3}$ is a product of Toffoli gates in staircase form but is not in $\mathcal{C}_3$, as $\mathrm{TOF}_{3,4,5}\mathrm{TOF}_{1,2,3}X_1\mathrm{TOF}_{1,2,3}\mathrm{TOF}_{3,4,5} = X_1\mathrm{CNOT}_{2,3}\mathrm{TOF}_{2,4,5}$.

To construct such a representation of π, we consider the operators $\pi X_j \pi^{-1}$. Since these operators are Clifford permutations, by Lemma 2.15, we have binary matrices A_j and vectors b_j such that $\pi X_j \pi^{-1}$ implements the permutation $|v\rangle \mapsto |v + A_j v + b_j\rangle$. The high level idea of our proof is clear: using a sequence of Clifford operators, we will reduce A_j, b_j to a specific form, which will help us build a staircase form representation of π. Let us begin by proving several useful lemmas.

The following lemma is essentially the same as standard results on simultaneous triangularization of commuting nilpotent matrices; see, for example, [RR00]. We include a proof for completeness.

Lemma 3.3. *Suppose $A_1, \ldots, A_k$ are linear transformations of an n-dimensional vector space V over a field F such that $A_j^2 = 0$ and $A_iA_j = A_jA_i$. Then there exists a basis of F^n in which all the A_i are strictly lower triangular. Recall that a matrix is* strictly lower triangular *if it is lower triangular, and all diagonal elements are* 0.

Proof. First we show that the intersections of kernels of A_i, namely $\cap_i \ker(A_i)$, is non-empty. Assume for the sake of contradiction it is empty, and let v be a non-zero vector which maximizes the number of indices i for which $A_iv = 0$. Take j with $A_jv \neq 0$, we see that $A_i(A_jv) = A_jA_iv = 0$

for any i with $A_i v = 0$, and $A_j(A_j v) = A_j^2 v = 0$. Therefore, $A_j v$ is in more kernels $\ker(A_i)$ than v is, contradicting our assumption on v. Hence, there must exist non-zero v such that for all i, $A_i v = 0$.

We now induct on n. Consider the $(n-1)$-dimensional vector space $V/\{v\}$. Since $v \in \cap_i \ker(A_i)$, all linear transformations A_i are well-defined on $V/\{v\}$ and satisfy $A_i A_j = A_j A_i$ and $A_i^2 = 0$. So there exists a basis $v_1+\{v\}, \ldots, v_{n-1}+\{v\}$ of $V/\{v\}$ in which all the A_i are strictly lower triangular. Now take the basis $v_1, v_2, \ldots, v_{n-1}, v$ (in that order) on V, one can check that all A_i are strictly lower triangular, as desired. □

For any nonzero column vector v over $\mathbb{F}_2$, let $\alpha(v)$ denote the index of its first nonzero component. Let $\alpha(0) = \infty$ by convention.

Lemma 3.4. *Suppose A is an $n \times n$ strictly lower triangular matrix over $\mathbb{F}_2$, and b is a nonzero column vector in $\mathbb{F}_2^n$. Then $\alpha(Ab) > \alpha(b)$.*

Proof. This follows directly from the definition of strictly lower triangular. □

Lemma 3.4 will be used tacitly throughout what follows.

Proposition 3.5. *Suppose we have a list of tuples $(A_1, b_1), \ldots, (A_n, b_n)$, where each A_i is an $n \times n$ strictly lower triangular matrix over $\mathbb{F}_2$ and each b_i is a column vector in $\mathbb{F}_2^n$. Suppose we can perform the following operations:*

(1) "Swap": swap the indices of two pairs (A_i, b_i) and (A_j, b_j), or
(2) "Compose": choose two distinct indices i and j, and update A_i to be $A_i + A_j + A_i A_j$ and update b_i to be $b_i + b_j + A_i b_j$.

Then it is possible to perform operations either to reach a state where $b_i = e_i$ for all i, or to reach a state where some b_i is 0. Recall that e_i denote the standard basis vectors of $\mathbb{F}_2^n$.

Proof. First note that the new matrix given by "compose" is always strictly lower triangular. Let us assume without loss of generality that we cannot reach $b_i = 0$ for any i. We describe a two-phase procedure which will reach the state $b_i = e_i$ for all i.

For the first phase of the process, we will reach a state with $\alpha(b_i) = i$ for all i, as follows. There are finitely many reachable states, so we can reach a state maximizing the value of $\sum_{i=1}^n \alpha(b_i)$ over all reachable states. In this state, the values of $\alpha(b_i)$ must be pairwise distinct. To see this, suppose $\alpha(b_i) = \alpha(b_j) = k$ for some $i \neq j$. Then note that $\alpha(b_i + b_j) > k$ and $\alpha(A_i b_j) > k$, so $\alpha(b_i + b_j + A_i b_j) > k = \alpha(b_i)$. This means if we compose (A_i, b_i) with (A_j, b_j) to obtain $(A_i + A_j + A_i A_j, b_i + b_j + A_i b_j)$, we will increase the value of $\sum_{i=1}^n \alpha(b_i)$, which is a contradiction. Therefore $\alpha(b_1), \ldots, \alpha(b_n)$ are pairwise distinct, so they must equal $1, 2, \ldots, n$ in some order. Perform swaps so that $\alpha(b_i) = i$ for all i, this completes the first phase.

The second phase of our procedure is simply row reduction. Suppose there exists $b_i \neq e_i$ and let $\alpha(b_i + e_i) = k > i$. Then we can compose (A_i, b_i) with (A_k, b_k) to get the new vector $b_i + b_k + A_i b_k$. Observe that

$$\alpha(b_i + e_i + b_k) > k, \alpha(A_i b_k) > k \Rightarrow \alpha(b_i + b_k + A_i b_k + e_i) > k.$$

Therefore we can repeat this procedure until $\alpha(b_i + e_i) > n$, which means $b_i = e_i$. Repeating this for all i leads to our desired state. □

Lemma 3.6. *A permutation gate π is staircase form if and only if, in the polynomial form of π^{-1}, for all k, the kth coordinate is a_k plus some (possibly empty) sum of terms of the form $a_i a_j$ with $i < j < k$. Furthermore, given a permutation π written as a product of Toffoli gates in staircase form, for any $i < j < k$, we have that $TOF_{i,j,k}$ appears in the product if and only if the kth coordinate in the polynomial form of π^{-1} contains an $a_i a_j$ term.*

Proof. Observe that, if π is a product of Toffoli gates in staircase form, then π^{-1} is the product of those same Toffoli gates in reverse order, and in that product, whenever a gate is applied, its controls have never been targeted so far, and thus are unchanged from the input. All parts of the desired result now can be easily shown. □

Lemma 3.7. *Given a staircase form permutation π, its representation as a product of Toffoli gates in staircase form is unique up to reordering gates with the same target.*

Proof. Lemma 3.6 yields that the polynomial form of π^{-1} determines exactly which Toffoli gates appear in any representation of π as a product of Toffoli gates in staircase form; the desired directly follows. □

To prove Theorem 3.2, one final ingredient is the following lemma, whose proof can be found in Appendix A.

Lemma 3.8. *Suppose $X'_1, X'_2, \ldots, X'_m \in \mathcal{X}$ are independent (that is, no nontrivial product of them yields the identity). Then there exists some Clifford permutation ν such that $\nu X_i \nu^{-1} = X'_i$ for all $i \in [m]$.*

We are now ready to prove Theorem 3.2.

Proof of Theorem 3.2. By multiplying π by suitable X's on the left, we assume without loss of generality that $\pi|0^n\rangle = |0^n\rangle$.

As discussed above, each $\pi X_j \pi^{-1}$ is a Clifford permutation, so by Lemma 2.15 we can write it as $|v\rangle \mapsto |v + A_j v + b_j\rangle$ for some matrix A_j and vector b_j over $\mathbb{F}_2$. Since $X_j^2 = I$ and $X_i X_j = X_j X_i$, we have $A_j^2 = 0$ and $A_i A_j = A_j A_i$. By Lemma 3.3, these conditions imply that there is some basis in which the A_j are simultaneously strictly lower triangular, so we can take some matrix M such that, for all i, MA_iM^{-1} is strictly lower triangular. Let ψ be the permutation gate $|v\rangle \mapsto |Mv\rangle$, which is Clifford by Lemma 2.15. The map $(\psi\pi)X_j(\psi\pi)^{-1}$ sends

$$|v\rangle \mapsto |M^{-1}v\rangle \mapsto |M^{-1}v + A_j M^{-1}v + b_j\rangle \mapsto |v + MA_jM^{-1}v + Mb_j\rangle.$$

Therefore, by replacing π with $\psi\pi$, we can assume without loss of generality that all matrices A_j are strictly lower triangular.

We now apply Proposition 3.5 to reduce b_i to e_i. Note that the map $\pi X_i X_j \pi^{-1} = (\pi X_i \pi^{-1})(\pi X_j \pi^{-1})$ sends

$$\begin{aligned}|v\rangle \mapsto |v + A_j v + b_j\rangle \mapsto &|(v + A_j v + b_j) + A_i(v + A_j v + b_j) + b_i\rangle \\ =&|v + (A_i + A_j + A_iA_j)v + b_i + b_j + A_ib_j\rangle,\end{aligned}$$

which corresponds to the compose operation. Therefore, by Proposition 3.5, there exists a sequence of swaps and multiplications which transform the generators $X_1, \ldots, X_n$ to $X'_1, \ldots, X'_n$, where each X'_i is a product of X gates, such that either $\pi X'_i \pi^{-1}$ sends $|v\rangle \mapsto |v + A'_i v + e_i\rangle$ for all i, or there exists i such that $\pi X'_i \pi^{-1}$ sends $|v\rangle \mapsto |v + A'_i v\rangle$. However, the latter case cannot happen, as otherwise $\pi X'_i \pi^{-1}$ sends $|0^n\rangle \mapsto |0^n + A'_i 0^n\rangle = |0^n\rangle$, which contradicts the assumption that $\pi\,|0^n\rangle = |0^n\rangle$.

Since $X'_1, \ldots, X'_n$ form a basis for $\mathcal{X}$, by Lemma A.4 there exists a Clifford permutation ν such that $\nu X_i \nu^{-1} = X'_i$ for all i. Therefore, if we replace π with $\pi\nu$, we get $(\pi\nu)X_i(\pi\nu)^{-1} = \pi X'_i \pi^{-1}$, which means we can assume without loss of generality that $b_i = e_i$ for all i. In particular, we have $\pi|e_i\rangle = (\pi X_i \pi^{-1})|0^n\rangle = |e_i\rangle$.

Next we show that for any v, if $\pi|v\rangle = |w\rangle$, then $\alpha(v) = \alpha(w)$. Suppose for the sake of contradiction that this is false. We know it is true for $v = 0^n$ or e_n, so $\alpha(v) < n$ in any counterexample. Take the largest k for which there exists a counterexample with $\alpha(v) = k$. We know $\pi|e_k\rangle = |e_k\rangle$, so $v \neq e_k$. Let $u \neq 0^n$ be such that $\pi|v + e_k\rangle = |u\rangle$. Then $\alpha(v + e_k) > k$, so $\alpha(u) = \alpha(v + e_k) = m$ for some $m > k$. We have

$$|w\rangle = \pi|v\rangle = \pi X_k|v + e_k\rangle = \pi X_k \pi^{-1}|u\rangle = |u + A_k u + e_k\rangle\,.$$

Since $\alpha(u) = m > k$, and $\alpha(A_k u) > m > k$, we must have $\alpha(w) = \alpha(e_k + u + A_k u) = k = \alpha(v)$, which is a contradiction.

We now build a polynomial representation (see Definition 2.12) for π^{-1}. By Lemma 2.13, every coordinate of π^{-1} has degree at most 2. Since $\pi^{-1}|0^n\rangle = |0^n\rangle$ and $\pi^{-1}|e_i\rangle = |e_i\rangle$ for all i, we have that the constant term of every coordinate is 0 and the linear term of the i-th coordinate is a_i for all i. Thus we can write π^{-1} as

$$|a_1, \ldots, a_n\rangle \mapsto |a_1 + q_1, \ldots, a_n + q_n\rangle,$$

where each q_k is a sum of some (possibly zero) monomials of the form $a_i a_j$ with $i < j$.

For any $i < j$, we have $\pi^{-1}|e_i + e_j\rangle = |e_i + e_j + w_{ij}\rangle$, where w_{ij} has ones exactly at the positions k for which q_k contains the monomial $a_i a_j$. Let v be such that $\pi|e_j + w_{ij}\rangle = |v\rangle$, we have

$$\begin{aligned}\pi X_i \pi^{-1}|v\rangle &= \pi X_i |e_j + w_{ij}\rangle = \pi|e_i + e_j + w_{ij}\rangle = |e_i + e_j\rangle \\ &= |v + A_i v + e_i\rangle\,,\end{aligned}$$

which means $v + A_i v = e_j$, and $j = \alpha(v + A_i v) = \alpha(v)$. Since $\pi|e_j + w_{ij}\rangle = |v\rangle$, we must also have $\alpha(e_j + w_{ij}) = \alpha(v)$. Thus $\alpha(e_j + w_{ij}) = j$, which means $\alpha(w_{ij}) > j$. In other words, any appearance of an $a_i a_j$ term can only be in a q_k with $k > j$. Then Lemma 3.6 implies that π is in staircase form.

Thus, unraveling our without-loss-of-generality assumptions, we can write $\pi = \phi_1 \mu \phi_2$ for Clifford permutations ϕ_1 and ϕ_2 and a staircase form permutation μ. Finally, $\mu = \phi_1^{-1} \pi \phi_2^{-1}$ is in $\mathcal{C}_3$ by Proposition 2.2, since $\phi_1^{-1}, \phi_2^{-1} \in \mathcal{C}_2$ and $\pi \in \mathcal{C}_3$.

□

4. Staircase Form Permutations in $\mathcal{C}_3$

In this section, we give a classification of when a permutation in staircase form is in $\mathcal{C}_3$. Previously we have been considering $\mathbb{F}_2^n$ as a vector space without any multiplicative structure. In this section we will explain how certain ways to define a multiplicative structure on $\mathbb{F}_2^n$ are in one-to-one correspondence with staircase form $\pi \in \mathcal{C}_3$.

Definition 4.1. A map $\mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{F}_2^n$, denoted by juxtaposition, is called a *descending multiplication* if

- it is linear in each coordinate (distributive), associative, and commutative,
- for all $i \in [n]$, we have $e_i e_i = e_i^2 = 0$, and
- for all $i < j \in [n]$, we have $e_i e_j$ is in the span of $\{e_k : k > j\}$.

Observe that, for any descending multiplication, we have $v^2 = 0$ for all v. Henceforth, for any permutation gate π, we will also interpret π as a permutation of $\mathbb{F}_2^n$, so that whenever $\pi|v\rangle = |w\rangle$, we can write $\pi(v) = w$.

We now state the main theorem of this section.

Theorem 4.2. *There is a one-to-one correspondence of descending multiplications to $\mathcal{C}_3$ permutations in staircase form, which we describe as follows. For each descending multiplication, the corresponding $\mathcal{C}_3$ permutation π is given by*

$$\forall S \subseteq [n], \pi \left| \sum_{i \in S} e_i \right\rangle = \left| \sum_{T \subseteq S, T \neq \emptyset} \prod_{i \in T} e_i \right\rangle. \tag{4.1}$$

Each permutation $\pi \in \mathcal{C}_3$ in staircase form induces a multiplication operation where each $e_i e_j$ is given by

$$\pi\,|e_i + e_j\rangle = |e_i + e_j + e_i e_j\rangle\,. \tag{4.2}$$

The multiplication is then extended linearly.

We break the proof of this theorem into several propositions.

Proposition 4.3. *For any descending multiplication, the resulting π from Eq. (4.1) is indeed a staircase form permutation in $\mathcal{C}_3$.*

Proposition 4.4. *For any staircase form permutation π in $\mathcal{C}_3$, the resulting operation from Eq. (4.2) is indeed a descending multiplication.*

Proposition 4.5. *Given a descending multiplication, applying the procedure in Eq. (4.1) to get a permutation π, then applying the procedure in Eq. (4.2) to that permutation, yields the original multiplication.*

Proposition 4.6. *Given a staircase form permutation π in $\mathcal{C}_3$, applying the procedure in Eq. (4.2) to get a descending multiplication, then applying the procedure in Eq. (4.1) to that multiplication, yields the original π.*

Proof of Theorem 4.2. Combining Propositions 4.3 to 4.6 yields the theorem. □

The following technical lemma is helpful for proving Propositions 4.3 to 4.6.

Lemma 4.7. *Let $\pi \in \mathcal{C}_3$ be a permutation. Then π is in staircase form if and only if the following conditions hold:*

- $\pi|0\rangle = |0\rangle$,
- $\pi|e_i\rangle = |e_i\rangle$ *for all* $i \in [n]$, *and*
- *for any vector v with at least two 1s, the indices of the first two 1s in v and $\pi(v)$ are the same.*

Proof. We use the equivalent characterization of staircase form permutations in Lemma 3.6. For the "if" direction, consider the polynomial representation $(b_1, \ldots, b_n)$ of π^{-1}. We know from Lemma 2.13 that b_i has degree at most 2 for all i. Now $\pi^{-1}(0) = 0$ yields that the constant term of b_i is 0 for all i. Moreover, for all i, $\pi^{-1}(e_i) = e_i$ yields that the linear term of b_i is a_i. Thus we can write $b_k = a_k + q_k$ where q_k is a sum of monomials of the form $a_i a_j$ with $i < j$.

Now for any $i < j$, $\pi^{-1}(e_i + e_j)$ cannot be 0 or any e_k (since π is a permutation). Therefore, $\pi^{-1}(e_i + e_j)$ considered as a vector has at least two 1s. Since $\pi(\pi^{-1}(e_i + e_j)) = e_i + e_j$, we invoke the third condition and conclude that $\pi^{-1}(e_i + e_j) = e_i + e_j + x_{ij}$ where x_{ij} is a (possibly empty) sum of terms of the form e_ℓ for $\ell > j$. But now note that, for any k, x_{ij} contains an e_k term if and only if q_k contains an $a_i a_j$ term; in other words, if q_k contains an $a_i a_j$ term, then $k > j$. Therefore the conditions from Lemma 3.6 hold, which means π is staircase form, as desired.

For the "only if" direction, suppose π is staircase form. We know from Lemma 3.6 that we can write the polynomial representation of π^{-1} as $(a_1 + q_1, \ldots, a_n + q_n)$ where q_k is a (possibly empty) sum of terms of the form $a_i a_j$ with $i < j < k$. This implies that that $\pi(0) = 0$ and $\pi(e_i) = e_i$ for all i. To prove the third condition, consider v containing at least two 1s. Let $w = \pi(v)$, then $w \neq 0$ and $w \neq e_i$ for all i, so it contains at least two 1s. Let $I < J$ be the positions of the first two 1s in w. If w and $\pi^{-1}(w) = v$ differ on the k coordinate, then $q_k(w) \neq 0$, which means a term of q_k, say $a_i a_j$ with $i < j < k$, is nonzero when evaluated at w. In other words $a_i = a_j = 1$ in w, which means $j \geq J$ and $k > J$. Thus w and v agree on the first J coordinates, in particular they agree on the positions of the first two 1s. □

Lemma 4.8. *Eq. (4.1) implies that*

$$\pi(v + w) = \pi(v) + \pi(w) + \pi(v)\pi(w) \tag{4.3}$$

for all v and w.

Proof. Let $v = \sum_{i\in V} e_i$ and $w = \sum_{i\in W} e_i$. Let $A = V \cap W, B = V\backslash W, C = W\backslash V$. Let $a = \sum_{i\in A} e_i, b = \sum_{i\in B} e_i, c = \sum_{i\in C} e_i$. Now observe that $A \cup B = V$ and $A \cap B = \emptyset$, so

$$\begin{aligned}\pi(a) + \pi(b) + \pi(a)\pi(b) &= \sum_{T\subseteq A, T\neq\emptyset} \prod_{i\in T} e_i + \sum_{T\subseteq B, T\neq\emptyset} \prod_{i\in T} e_i + \left(\sum_{T\subseteq A, T\neq\emptyset} \prod_{i\in T} e_i\right)\left(\sum_{T\subseteq B, T\neq\emptyset} \prod_{i\in T} e_i\right) \\ &= \sum_{T\subseteq A\cup B, T\neq\emptyset} \prod_{i\in T} e_i = \pi(v).\end{aligned}$$

We can similarly show that $\pi(a)+\pi(c)+\pi(a)\pi(c) = \pi(w)$ and $\pi(b)+\pi(c)+\pi(b)\pi(c) = \pi(v+w)$. Further, since $\pi(c)\pi(c) = 0$, we have

$$\begin{aligned}\pi(v) + \pi(w) + \pi(v)\pi(w) &= \pi(a) + \pi(b) + \pi(a)\pi(b) + \pi(a) + \pi(c) + \pi(a)\pi(c) \\ &\quad + (\pi(a) + \pi(b) + \pi(a)\pi(b))(\pi(a) + \pi(c) + \pi(a)\pi(c)) \\ &= \pi(b) + \pi(a)\pi(b) + \pi(c) + \pi(a)\pi(c) \\ &\quad + (\pi(a)\pi(b) + \pi(a)\pi(c) + \pi(b)\pi(c)) \\ &= \pi(b) + \pi(c) + \pi(b)\pi(c) \\ &= \pi(v + w).\end{aligned}$$

□

We are now ready to prove Propositions 4.3 to 4.6.

Proof of Proposition 4.3. We first note that Eq. (4.1) implies the conditions in Lemma 4.7. In particular, the fact that e_ie_j is a descending multiplication implies the third condition. Therefore, it suffices for us to show that π is a permutation gate in $\mathcal{C}_3$.

Since $\pi(v)$ is nonzero for $v \neq 0$, π is injective: if $\pi(v) = \pi(w)$, then Eq. (4.3) implies that $\pi(v + w) = 0$, which means $v = w$. Thus π is a valid permutation of $\mathbb{F}_2^n$.

To show $\pi \in \mathcal{C}_3$ we consider its conjugation of single qubit Pauli operators. For any v and any index i, if we let $\pi^{-1}(v) = w$, then

$$\pi X_i\pi^{-1}(v) = \pi(e_i + w) = \pi(e_i) + \pi(w) + \pi(e_i)\pi(w) = e_i + v + e_iv.$$

The map $v \mapsto e_i + v + e_iv$ is invertible since it is its own inverse; thus $\pi X_i\pi^{-1} \in \mathcal{C}_2$ by Lemma 2.15.

We now prove that $\pi Z_i\pi^{-1} \in \mathcal{C}_2$ for all i. For each pair of indices $i < j$, define v_{ij} to be such that $\pi(e_i + e_j + v_{ij}) = e_i + e_j$. Observe that

$$\pi(v_{ij}) = \pi((e_i + e_j + v_{ij}) + (e_i + e_j)) = (e_i + e_j)(e_i + e_j + e_ie_j) + (e_i + e_j) + (e_i + e_j + e_ie_j) = e_ie_j.$$

We claim that, for any set S of indices,

$$\pi\left(\sum_{i\in S} e_i + \sum_{i,j\in S; i<j} v_{ij}\right) = \sum_{i\in S} e_i. \tag{4.4}$$

The proof is by induction on $|S|$. The base case $|S| \leq 2$ is clear. For the inductive step, let i be the smallest element of S, and let $T = S\backslash\{i\}$. Let $w = \sum_{j\in T} e_j$, and let $x = e_i + \sum_{j\in T} v_{ij}$. By repeatedly using Eq. (4.3) and the facts that $e_i^2 = 0$ and $\pi(v_{ij}) = e_ie_j$, we can see that

$$\pi(x) = e_i + \sum_{j\in T} e_ie_j = e_i + e_iw.$$

Also, by the inductive hypothesis, $\pi(w + \sum_{j,k\in T;j<k} v_{jk}) = w$. Thus

$$\begin{aligned}\pi\left(\sum_{i\in S} e_i + \sum_{i,j\in S;i<j} v_{ij}\right) &= \pi\left(\left(w + \sum_{j,k\in T;j<k} v_{jk}\right) + x\right)\\ &= w + (e_i + e_i w) + w(e_i + e_i w)\\ &= w + e_i = \sum_{i\in S} e_i.\end{aligned}$$

This completes the inductive step, so Eq. (4.4) holds. Note that Eq. (4.4) implies that every coordinate in the polynomial representation of π^{-1} has degree at most 2, since each vector v_{ij} appears in the sum when both i, j are in S. Thus $\pi Z_i \pi^{-1} \in \mathcal{C}_2$ for all i, and we conclude that $\pi \in \mathcal{C}_3$. Lemma 4.7 then completes our proof. □

Proof of Proposition 4.4. The distributive property holds by definition. The commutative property clearly holds, since $e_i e_j = e_j e_i$. We have $e_i^2 = 0$ since $\pi|0\rangle = |0\rangle$. The fact that $e_i e_j$ is in the span of $\{e_k : k > j\}$ for $i < j$ follows from Lemma 4.7. It remains for us to prove associativity.

For each i, we have that $\pi X_i \pi^{-1}$ is a Clifford permutation, so by Lemma 2.15, there is a matrix A_i and a vector b_i so that $\pi X_i \pi^{-1}|v\rangle = |v + A_i v + b_i\rangle$ for all v. Setting $v = 0$ yields that $b_i = e_i$. Then setting $v = e_j$ yields that $\pi|e_i + e_j\rangle = |e_i + e_j + A_i e_j\rangle$, so $A_i e_j = e_i e_j$. Since we defined the multiplication to be linear in each coordinate, this implies $A_i v = e_i v$ for all v.

Now, since X_i and X_k commute, so do $\pi X_i \pi^{-1}$ and $\pi X_k \pi^{-1}$. Therefore the maps $v \mapsto v + A_i v + e_i$ and $v \mapsto v + A_k v + e_k$ commute, which means A_i and A_k commute. Thus $A_i(A_k e_j) = A_k(A_i e_j)$ for all i, j, k, so $e_i(e_j e_k) = e_i(e_k e_j) = e_k(e_i e_j) = (e_i e_j)e_k$, by commutativity of the multiplication. This yields the desired associativity. □

Proof of Proposition 4.5. For $i \neq j$, by setting $S = \{i, j\}$ in Eq. (4.1), we see that the new multiplication has the same value of $e_i e_j$ as the original multiplication; also, setting $j = i$ in Eq. (4.2) gives that the new multiplication has $e_i^2 = 0$. Thus the original and new multiplications coincide on the value of $e_i e_j$ for all i, j, and both are linear in each input. This means they are the same multiplication. □

Proof of Proposition 4.6. Let π denote the original permutation. For any set S of indices, we have

$$\pi\left|\sum_{i\in S} e_i\right\rangle = \pi\left(\prod_{i\in S} X_i\right)\pi^{-1}|0\rangle = \prod_{i\in S}(\pi X_i \pi^{-1})|0\rangle.$$

From the proof of Proposition 4.4, $\pi X_i \pi^{-1}|v\rangle = |v + e_i v + e_i\rangle$ for all v, i. It can be easily verified that applying the maps $v \mapsto v + e_i v + e_i$ sequentially for all $i \in S$, starting with $v = 0$, yields $\sum_{T\subseteq S, T\neq\emptyset} \prod_{i\in T} e_i$ as desired. □

Proposition 4.9. *Given a staircase form permutation π in $\mathcal{C}_3$, taking v_{ij} defined by $\pi^{-1}(e_i + e_j) = e_i + e_j + v_{ij}$ for all $i < j$, we have, for all $i < j < k$, that $TOF_{i,j,k}$ appears in the staircase form of π if and only if there is an e_k term in v_{ij}.*

Proof. Note that v_{ij} is the vector of the positions containing a $a_i a_j$ monomial in the polynomial form of π^{-1}, so this follows directly from Lemma 3.6. □

Proposition 4.10. *Given a descending multiplication and its corresponding $\mathcal{C}_3$ permutation π, for any nonempty set $S \subseteq [n]$, the value of $\Pi_{i\in S} e_i$ is exactly the vector corresponding to the positions at which an $\Pi_{i\in S} a_i$ term appears in the polynomial representation of π.*

Proof. For any S, let p_S be the vector corresponding to the positions at which an $\Pi_{i\in S} a_i$ term appears in the polynomial representation of π. Then we can see, for any set S, that $\pi(\sum_{i\in S} e_i) =$

$\sum_{T\subseteq S} p_T$. From Eq. (4.1), we have that $\sum_{T\subseteq S;T\neq\emptyset} p_T = \sum_{T\subseteq S,T\neq\emptyset} \prod_{i\in T} e_i$ for any nonempty set S. It then easily follows by strong induction on $|S|$ that $p_S = \prod_{i\in S} e_i$ for all nonempty S. □

5. A Family of Non-Semi-Clifford Permutations

In this section, we construct an infinite family of permutation gates $\{U_k\}_{k\geq 3}$ in Definition 5.2 and prove in Theorem 5.5 that every U_k is non-semi-Clifford in $\mathcal{C}_3$. This rejects Anderson's conjectures. To better understand these gates, we study the smallest case of $k = 3$ closely in Section 5.2. We show that this 7-qubit permutation U_3 is in fact conjugate to the Gottesman–Mochon example by a Clifford operator.

The following perspective of labeling qubits by subsets is crucial for our construction and the proof of Theorem 5.5.

Definition 5.1 (Labeling qubits)**.** Given a positive integer k and $2^k - 1$ qubits, qubit $x \in \{1, 2, \ldots, 2^k - 1\}$ can be labeled equivalently by (1) its binary representation $\mathrm{bin}(x) \in \{0,1\}^k \setminus \{0^k\}$, or (2) a non-empty subset $S_x \subseteq [k]$ where $i \in S_x$ if and only if the i-th rightmost bit of $\mathrm{bin}(x)$ is 1. (Note that this correspondence of nonempty sets to integers in binary yields a total ordering of the set of nonempty subsets of $[k]$.)

Definition 5.2 (A family of permutations)**.** For each integer $k \geq 3$, taking $n = 2^k - 1$, label a basis of $\mathbb{F}_2^n$ as e_S for nonempty subsets $S \subseteq [n]$, analogously to Definition 5.1. Then define a multiplication by setting $e_S e_T = e_{S\cup T}$ if $S \cap T = \emptyset$, and $e_S e_T = 0$ if $S \cap T \neq \emptyset$, and extending linearly; it is easy to check that this yields a descending multiplication. Let U_k be the staircase form $\mathcal{C}_3$ permutation corresponding to this descending multiplication, as in Theorem 4.2.

Proposition 5.3. *We can express U_k as a product of Toffoli gates in staircase form as follows: for each pair S, T of nonempty disjoint subsets of $[n]$ with $S < T$, apply the gate $\mathrm{TOF}_{S,T,S\cup T}$; specifically, apply these gates in nondecreasing order of target.*

Remark 5.4. From the perspective of labeling gates with integers instead of sets, Proposition 5.3 states that we can express U_k as a product of Toffoli gates in staircase form as follows: for each pair of indices $i < j$ that do not have any 1s in the same place as each other in binary, apply $\mathrm{TOF}_{i,j,i+j}$; specifically, apply these gates in nondecreasing order of target.

Proof. Take $U_k = \pi$ in Proposition 4.9. We know from the proof of Proposition 4.3 that $v_{ij} = \pi^{-1}(e_i e_j)$ for all $i < j$. In the notation of indexing qubits as sets, this becomes $v_{ST} = \pi^{-1}(e_S e_T)$ for $S < T$. Now note that $e_S e_T$ is $e_{S\cup T}$ or 0, and in either case $\pi^{-1}(e_S e_T) = e_S e_T$. Thus $v_{ST} = e_S e_T$ in any case, so Proposition 4.9 implies the desired, by the definition of $e_S e_T$. □

5.1. All U_k's are non-semi-Clifford.

Theorem 5.5. *For any integer $k \geq 3$, we have $U_k \in \mathcal{C}_3$ but $U_k^{-1} \notin \mathcal{C}_k$. Thus U_k is not semi-Clifford.*

Lemma 5.6 (Polynomial representations of U_k and U_k^{-1})**.** *For each integer $k \geq 3$, let us denote the polynomial representations (Definition 2.12) of U_k and U_k^{-1} respectively as $(\pi_S)_{S\subseteq[k],S\neq\emptyset}$ and $(\pi'_S)_{S\subseteq[k],S\neq\emptyset}$. Then*

$$\pi_S(a) = \sum_{m=1}^{|S|} \sum_{\substack{\bigsqcup_{i=1}^m T_i = S,\\ T_i\neq\emptyset}} a_{T_1} a_{T_2} \cdots a_{T_m}, \tag{5.1}$$

where the sum is over unordered non-empty subsets $T_1, \ldots, T_m$, in other words, all partitions of S, and

$$\pi'_S(a) = a_S + \sum_{T_i\neq\emptyset, T_1\sqcup T_2 = S} a_{T_1} a_{T_2},$$

where the sum is over unordered pairs T_1, T_2.

Proof. The polynomial form of π^{-1} follows directly from Proposition 5.3 and Lemma 3.6. For the polynomial form of π, observe that, for any $T_1, \dots, T_m$, we have $e_{T_1} \dots e_{T_m}$ is $e_{T_1 \cup \dots \cup T_m}$ if the T_i are pairwise disjoint, and 0 if not, so the desired follows from Proposition 4.10. □

Proof of Theorem 5.5. We know $U_k \in \mathcal{C}_3$ by definition. Also, by Lemma 5.6, we know that $\pi_{[k]}$ is a polynomial of degree k because it contains the monomial $a_{\{1\}}a_{\{2\}} \cdots a_{\{k\}}$. This implies that $U_k^{-1} \notin \mathcal{C}_k$ using Lemma 2.13. In particular, $U_k^{-1} \notin \mathcal{C}_3$, so U_k is not semi-Clifford by Proposition 2.5. □

5.2. **Example: $k = 3$.** Let us examine the case of $k = 3$, the simplest gate in the family. U_3 is a 7-qubit permutation gate with 6 Toffoli gates in staircase form (see circuit in Figure 3):

$$U_3 = \mathrm{TOF}_{1,6,7}\mathrm{TOF}_{2,5,7}\mathrm{TOF}_{3,4,7}\mathrm{TOF}_{2,4,6}\mathrm{TOF}_{1,4,5}\mathrm{TOF}_{1,2,3}.$$

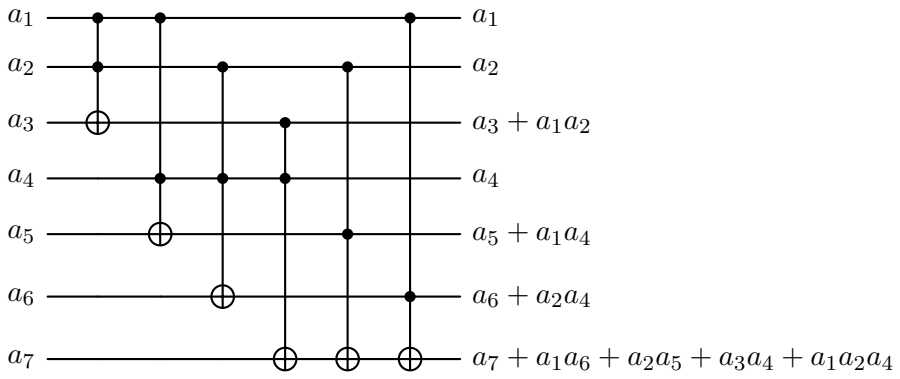

FIGURE 3. The non-semi-Clifford permutation gate $U_3 \in \mathcal{C}_3$.

Remark 5.7. In Figure 3, note that the seventh coordinate of U_3 is of degree 3, so this is an example of the form mentioned in Remark 2.14.

Recall the Gottesman–Mochon 7-qubit gate (see Figure 1 for the circuit)

$$\mathcal{C}_3 \ni G = \mathrm{CSWAP}_{7,1,6}\mathrm{CSWAP}_{7,2,5}\mathrm{CSWAP}_{7,4,3} \cdot \mathrm{CCZ}_{1,2,3}\mathrm{CCZ}_{1,4,5}\mathrm{CCZ}_{2,4,6}\mathrm{CCZ}_{3,5,6},$$

which is not semi-Clifford, as in the proof of Lemma 2.6. Let us define a 7-qubit Clifford gate F:

$$F = H_3H_5H_6\mathrm{CNOT}_{6,1}\mathrm{CNOT}_{5,2}\mathrm{CNOT}_{3,4}H_7.$$

Proposition 5.8. *U_3 is a non-semi-Clifford permutation in $\mathcal{C}_3$ on 7 qubits and $FGF^{-1} = U_3$.*

Proof. The fact that U_3 is a non-semi-Clifford permutation on 7 qubits is a special case of Theorem 5.5. Checking that $FGF^{-1} = U_3$ is a direct computation.

□

Note that, by Lemma 2.18, this implies that, for all $n \geq 7$, $\mathcal{C}_3$ contains a non-semi-Clifford permutation.

5.3. **High-Degree Monomials in $\mathcal{C}_3$ Permutations.** We now know that U_k is a $\mathcal{C}_3$ permutation on $2^k - 1$ qubits that contains a degree-k monomial somewhere in its polynomial representation (as in the proof of Theorem 5.5). In this subsection we will show that this is the smallest possible number of qubits for this to happen.

Proposition 5.9. *Given any descending multiplication, if $\Pi_{i \in S} e_i$ is nonzero for some k-element set S, then $n \geq 2^k - 1$.*

Proof. Suppose $\Pi_{i\in S}e_i$ is nonzero. For any nonempty subset T of S, let $p_T = \Pi_{i\in T}e_i$. We shall show that the vectors p_T, over all nonempty subsets T of S, are linearly independent.

Suppose for contradiction they are linearly dependent. Take a family F of nonempty subsets of S such that $\sum_{T\in F} p_T = 0$. Take a minimal element U of F (i.e., such that no proper subset of U is in F). Let $V = S\backslash U$. Note that for any $T \in F$ with $T \neq U$, we have $T \not\subseteq U$, which means $T \cap V \neq \varnothing$. From Definition 4.1, we see that $p_T p_V = 0$. Therefore,

$$\sum_{T\in F} p_T p_V = p_U p_V = p_S.$$

On the other hand, $\sum_{T\in F} p_T = 0$ implies $(\sum_{T\in F} p_T = 0)p_V = 0$. Thus $p_S = 0$, which is a contradiction. We conclude that the vectors p_T must be linearly independent. Since there are $2^k - 1$ such vectors in the vector space $\mathbb{F}_2^n$, we must have $n \geq 2^k - 1$. □

Proposition 5.10. *If π is a staircase form $\mathcal{C}_3$ permutation such that there is a degree-k monomial somewhere in the polynomial form of π, then $n \geq 2^k - 1$.*

Proof. Suppose $\Pi_{i\in S}a_i$ appears somewhere in the polynomial form of π, for some k-element set S. Then, for the descending multiplication corresponding to π, we have $\Pi_{i\in S}e_i$ is nonzero, by Proposition 4.10; then Proposition 5.9 yields the desired. □

Theorem 5.11. *If π is a $\mathcal{C}_3$ permutation such that there is a degree-k monomial somewhere in the polynomial form of π, then $n \geq 2^k - 1$; this bound is sharp for all $k \geq 3$.*

Proof. Assume WLOG that $k \geq 2$. By Theorem 3.2, we can write $\pi = \phi_1\mu\phi_2$, for Clifford permutations ϕ_1 and ϕ_2 and staircase form $\mu \in \mathcal{C}_3$. Now all terms in the polynomial forms of ϕ_1 and ϕ_2 are degree at most 1; then if all terms in the polynomial form of μ have degree less than k, then all terms in the polynomial form of $\phi_1\mu\phi_2 = \pi$ have degree less than k, a contradiction. Thus there exists some term in the polynomial form of μ with degree $d \geq k$. Then Proposition 5.10 applied to μ yields that $n \geq 2^d - 1 \geq 2^k - 1$, as desired. The example of U_k shows that the bound is sharp. □

5.4. **Rejection of Anderson's conjectures.**

Conjecture 5.12 ([And24, Conjecture D.1])**.** *Any permutation in $\mathcal{C}_3$ is a product of pairwise commuting Toffoli gates, up to left and right multiplications of Clifford permutations.*

Conjecture 5.13 ([And24, Conjecture D.2])**.** *For any permutation π and any positive integer k, if $\pi \in \mathcal{C}_k$, then $\pi^\dagger \in \mathcal{C}_k$.*

We use the following lemma to disprove Anderson's two conjectures.

Lemma 5.14. *Two C^*X gates commute if and only if they have no mismatch (that is, there is no qubit that is used as a target in one and a control in the other).*

Proof. The "if" direction is clear. Let us prove the "only if" direction. Suppose for contradiction they have mismatch. Without loss of generality, let the gates be A, with qubit 1 as a control and qubit 2 as target, and B, with qubit 1 as target. Then $AB|1^n\rangle = A|01^{n-1}\rangle = |01^{n-1}\rangle$, while $BA|1^n\rangle = B|101^{n-2}\rangle$, which is either $|101^{n-2}\rangle$ or $|001^{n-2}\rangle$; in either case $BA|1^n\rangle \neq AB|1^n\rangle$, therefore they do not commute. □

Theorem 5.15. *Conjecture 5.12 and Conjecture 5.13 are false.*

Proof. Suppose Conjecture 5.12 is true. By Lemma 5.14, every permutation in $\mathcal{C}_3$ is a mismatch-free product of Toffoli gates, up to Clifford permutations on the left and right. So every permutation in $\mathcal{C}_3$ is semi-Clifford by Lemma A.1. This is a contradiction, as we know R is a non-semi-Clifford permutation in $\mathcal{C}_3$ for $n = 7$.

For Conjecture 5.13, we have R is a permutation in $\mathcal{C}_3$, while $R^\dagger \notin \mathcal{C}_3$, as $R^\dagger = R^{-1}$, and we showed in the proof of Proposition 5.8 that $R^{-1} \notin \mathcal{C}_3$. Thus Conjecture 5.13 is false. □

6. Conclusion

6.1. **Open problems.** We believe that the mismatch-free and staircase form definitions, alongside the polynomial viewpoint, offer the right perspectives for studying permutation gates in the Clifford hierarchy. Below are some open problems that we think worth exploring further:

(1) Are all products of Toffoli gates in staircase form in the Clifford hierarchy at some level?
(2) Is the inverse of any $\mathcal{C}_3$ permutation in the Clifford hierarchy?
(3) A natural follow-up question to our Theorem B.2: are all $\mathcal{C}_3$ gates on $n \leq 6$ qubits semi-Clifford?

Acknowledgments

The work was conducted as a part of the 2024 Summer Program for Undergraduate Research (SPUR) and 2024-2025 Undergraduate Research Opportunities Program (UROP) at MIT. We thank Jonathan Bloom, Isaac Chuang, David Jerison, and Peter Shor for their mentorship. X. Tan would like to thank Robert Calderbank for introducing the problem of characterizing the third-level Clifford hierarchy to her in 2022 and many insightful discussions afterwards. We thank Jeongwan Haah, Aram Harrow, Greg Kahanamoku-Meyer, Andrey Khesin, and Anirudh Krishna for helpful discussions.

Z. He is supported by National Science Foundation Graduate Research Fellowship under Grant No. 2141064. X. Tan is supported by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers, Co-design Center for Quantum Advantage (C2QA) under contract number DE-SC0012704.

References

[AC25] Jonas T Anderson and Andrew Connelly. "Affine Equivalence in the Clifford Hierarchy". In: *arXiv preprint arXiv:2507.14370* (2025).

[And24] Jonas T. Anderson. "On Groups in the Qubit Clifford Hierarchy". In: *Quantum* 8 (June 2024), p. 1370. DOI: 10.22331/q-2024-06-13-1370.

[BCP97] Wieb Bosma, John Cannon, and Catherine Playoust. "The Magma algebra system. I. The user language". In: *J. Symbolic Comput.* 24.3-4 (1997). Computational algebra and number theory (London, 1993), pp. 235–265. DOI: 10.1006/jsco.1996.0125.

[BS09] Salman Beigi and Peter W. Shor. *C_3, Semi-Clifford and Generalized Semi-Clifford Operations*. 2009. arXiv: 0810.5108 [quant-ph].

[CGK17] Shawn X. Cui, Daniel Gottesman, and Anirudh Krishna. "Diagonal gates in the Clifford hierarchy". In: *Phys. Rev. A* 95 (1 Jan. 2017), p. 012329. DOI: 10.1103/PhysRevA.95.012329.

[CHH+24] Chi-Fang Chen, Jeongwan Haah, Jonas Haferkamp, Yunchao Liu, Tony Metger, and Xinyu Tan. "Incompressibility and spectral gaps of random circuits". In: *arXiv preprint arXiv:2406.07478* (2024).

[EK09] Bryan Eastin and Emanuel Knill. "Restrictions on Transversal Encoded Quantum Gate Sets". In: *Physical Review Letters* 102.11 (Mar. 2009). DOI: 10.1103/physrevlett.102.110502.

[GC99] Daniel Gottesman and Isaac L. Chuang. "Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations". In: *Nature* 402.6760 (Nov. 1999), pp. 390–393. DOI: 10.1038/46503.

[Got98] Daniel Gottesman. *The Heisenberg Representation of Quantum Computers*. 1998. arXiv: quant-ph/9807006 [quant-ph].

[MPSY24] Tony Metger, Alexander Poremba, Makrand Sinha, and Henry Yuen. "Simple constructions of linear-depth t-designs and pseudorandom unitaries". In: *arXiv preprint arXiv:2404.12647* (2024).

[NRS00] Gabriele Nebe, E. M. Rains, and N. J. A. Sloane. *The invariants of the Clifford groups*. 2000. arXiv: math/0001038 [math.CO].

[RR00] Heydar Radjavi and Peter Rosenthal. *Simultaneous Trianglularization*. Springer Science & Business Media, 2000. DOI: 10.1007/978-1-4612-1200-3.

[ZCC08] Bei Zeng, Xie Chen, and Isaac L. Chuang. "Semi-Clifford operations, structure of C_k hierarchy, and gate complexity for fault-tolerant quantum computation". In: *Phys. Rev. A* 77 (4 Apr. 2008), p. 042313. DOI: 10.1103/PhysRevA.77.042313.

[ZLC00] Xinlan Zhou, Debbie W. Leung, and Isaac L. Chuang. "Methodology for quantum logic gate construction". In: *Physical Review A* 62.5 (Oct. 2000). DOI: 10.1103/physreva.62.052316.

Appendix A. Semi-Clifford Permutations

Consider a permutation gate written as a product of Toffoli gates. This representation is said to be *mismatch-free* if no qubit is used as both a control and target. The notion of mismatch-free extends more generally (see Appendix D of [And24]). Specifically, a Toffoli gate can be considered as a controlled controlled X gate, and this can be generalized to any number of control bits: the C^kX gate sends $|a_1\rangle \otimes \cdots \otimes |a_k\rangle \otimes |a_{k+1}\rangle \mapsto |a_1\rangle \otimes \cdots \otimes |a_k\rangle \otimes |a_{k+1} + a_1 \cdots a_k\rangle$. We refer to all these collectively as C^*X gates. The CNOT and X gates are special cases of C^*X gates with one and zero controls, respectively. Then the definition of mismatch-free naturally extends to any product of C^*X gates: a representation of a permutation as a product of C^*X gates is mismatch-free if no qubit is used both as control and target. Observe that in any such mismatch-free product, every two gates commute.

Lemma A.1. *Any mismatch-free product μ of C^*X gates is semi-Clifford.*

Proof. Consider an X gate on every target qubit and a Z gate on every non-target qubit. These gates generate a maximal abelian subgroup of $\mathcal{P}_n$ up to phase, and μ will commute with every element of this subgroup. The claim follows from Lemma 2.4. □

Theorem A.2. *For any permutation gate π that is semi-Clifford, there exist Clifford permutations ϕ_1, ϕ_2 and a mismatch-free product μ of C^*X gates such that $\pi = \phi_1 \mu \phi_2$.*

Before we prove the theorem, we prove some lemmas. Given a vector $u \in \mathbb{F}_2^n$, we use the notation X^u to denote the operator $X^{u[1]} \otimes X^{u[2]} \otimes \cdots \otimes X^{u[n]}$, where $u[i]$ denote the i-th index of u, $X^1 = X$ and $X^0 = I$. Define Z^u similarly. Any Pauli operator $P \in \mathcal{P}_n$ has a decomposition as $P = cX^uZ^v$ for some phase c and $u, v \in \mathbb{F}_2^n$.

Lemma A.3. *Every Pauli gate can be uniquely written as the product of a permutation gate and a diagonal gate; furthermore, the permutation gate and diagonal gate are each individually Pauli.*

Proof. Any Pauli operator P can be written as $P = cX^uZ^v$, where $p = X^u$ is a permutation gate and $d = cZ^v$ is a diagonal gate. It remains to show uniqueness of the representation. To see this, suppose $P = p'd'$ for permutation p' and diagonal d'. We have $(p')^{-1}p = (p')^{-1}Pd^{-1} = d'd^{-1}$. Since $(p')^{-1}p$ is a permutation matrix, $d'd^{-1}$ is a diagonal matrix, and the only diagonal permutation matrix is the identity, we must have $(p')^{-1}p = d'd^{-1} = I$. Therefore $p' = p$ and $d' = d$, as desired. □

Lemma A.4. *Suppose $X'_1, X'_2, \ldots, X'_m \in \mathcal{X}$ are independent (that is, no nontrivial product of them yields the identity). Then there exists some Clifford permutation ν such that $\nu X_i \nu^{-1} = X'_i$ for all $i \in [m]$.*

Proof. Note that we can view each X'_i as a map $|v\rangle \mapsto |v + v_i\rangle$. The independence property gives that $v_1, \ldots, v_m$ are linearly independent, i.e. there exists an invertible matrix M such that $Me_i = v_i$

for all $i \in [m]$. Let ν be $|v\rangle \mapsto |Mv\rangle$ which is a Clifford permutation by Lemma 2.15. Then $\nu X_i \nu^{-1}$ sends

$$|v\rangle \mapsto |M^{-1}v\rangle \mapsto |M^{-1}v + e_i\rangle \mapsto |M(M^{-1}v + e_i)\rangle = |v + Me_i\rangle = |v + v_i\rangle,$$

which means that $\nu X_i \nu^{-1} = X_i'$, as desired. □

The following lemma is a special case of Theorem A.2.

Lemma A.5. *Suppose π is a permutation gate, and $m \le n$ is a nonnegative integer such that $\pi X_1 \pi^{-1}, \dots, \pi X_m \pi^{-1}, \pi Z_{m+1} \pi^{-1}, \dots, \pi Z_n \pi^{-1} \in \mathcal{P}_n$. Then there exist Clifford permutations ϕ_1, ϕ_2 and a mismatch-free product μ of C^*X gates such that $\pi = \phi_1 \mu \phi_2$.*

Proof. Let $X_i' = \pi X_i \pi^{-1}$. By Lemma A.4 we can take a Clifford permutation ν such that $X_i' = \nu X_i \nu^{-1}$. Replacing π with $\nu^{-1}\pi$, which preserves the property that $\pi Z_j \pi^{-1} \in \mathcal{P}_n$ for $m+1 \le j \le n$, we can assume without loss of generality that π commutes with $X_1, \dots, X_m$.

For $m+1 \le j \le n$, $\pi Z_j \pi^{-1}$ is a diagonal gate in the Pauli group, i.e. $\epsilon_j Z^{w_j}$ for some vector w_j and $\epsilon_j = \pm 1$. Since $\pi Z_j \pi^{-1}$ commutes with $\pi X_i \pi^{-1} = X_i$ for $i \in [m]$, we must have w_j is zero on the first m components for all $m+1 \le j \le n$. Let χ be the product of X_j over all j with $\epsilon_j = -1$. By replacing π with $\pi\chi$, we can assume without loss of generality that $\epsilon_j = 1$ for all j, while preserving the property that π commutes with $X_1, \dots, X_m$, and without changing $w_{m+1}, \dots, w_n$.

Since Z_j are independent, the vectors w_j are also linearly independent. Since the first m components of each w_j are zeros, $e_1, \dots, e_m, w_{m+1}, \dots, w_n$ forms a linear basis. Hence, there exists an invertible matrix M with $Me_i = e_i$ for $i \in [m]$ and $Me_j = w_j$ for $m+1 \le j \le n$. Consider the map ϖ defined as $|v\rangle \mapsto |M^\top v\rangle$ which is a Clifford permutation by Lemma 2.15. Then, $\varpi(\pi Z_j \pi^{-1})\varpi^{-1} = \varpi Z^{Me_j} \varpi^{-1}$ sends

$$|v\rangle \mapsto |(M^\top)^{-1}v\rangle \mapsto (-1)^{v^\top M^{-1} M e_j} |(M^\top)^{-1}v\rangle \mapsto (-1)^{v^\top e_j} |v\rangle,$$

so $\varpi\pi Z_j \pi^{-1}\varpi^{-1} = Z_j$. Also, since the first m components of w_j are zeros for $m+1 \le j \le n$, we have $M^\top e_i = e_i$ for $i = 1, \dots, m$. Therefore, $\varpi\pi X_i \pi^{-1}\varpi^{-1} = \varpi X_i \varpi^{-1}$ sends

$$|v\rangle \mapsto |(M^\top)^{-1}v\rangle \mapsto |(M^\top)^{-1}v + e_i\rangle \mapsto |M^\top((M^\top)^{-1}v + e_i)\rangle = |v + e_i\rangle,$$

so $\varpi\pi X_i \pi^{-1}\varpi^{-1} = X_i$. Since ϖ is a Clifford permutation and $\varpi\pi$ commutes with $X_1, \dots, X_m$ and $Z_{m+1}, \dots, Z_n$, by replacing π with $\varpi\pi$, we can assume without loss of generality that π commutes with $X_1, \dots, X_m, Z_{m+1}, \dots, Z_n$.

Now consider the polynomial representation $(\pi_1, \dots, \pi_n)$ of π. For $m+1 \le j \le n$, $\pi Z_j = Z_j \pi$ implies that $\pi_j(v) = v_j$ for $v \in \mathbb{F}_2^n$. For $1 \le j \le m$, $\pi X_j = X_j \pi$ implies that $\pi(v + e_j) = \pi(v) + e_j$, i.e. $\pi_i(v + e_j) = \pi_i(v)$ for $i \ne j$, and $\pi_j(v + e_j) = \pi_j(v) + 1$. So π_j is v_j plus a polynomial p_j in terms of only $v_{m+1}, \dots, v_n$.

Note that every monomial in p_j corresponds to a C^*X gate with qubit j as target and a subset of qubits in $\{m+1, \dots, n\}$ as controls. Now π is the product of all these C^*X gates and is mismatch-free, since qubits $1, \dots, m$ are used only as targets and qubits $m+1, \dots, n$ are never used as targets. □

We now prove Theorem A.2 by reducing to the case of Lemma A.5.

Proof of Theorem A.2. Let G be the subgroup of $\mathcal{P}_n$ of all elements P with $\pi P \pi^{-1} \in \mathcal{P}_n$, and let M be the set consisting of all permutations in G; now $M = G \cap \mathcal{X}$, so M is an abelian subgroup of G. By Lemma A.4 we can find a Clifford permutation ν such that $M = \nu\langle X_1, \dots, X_m\rangle\nu^{-1}$ for some m. If we replace π by $\pi\nu$, we would replace G with $\nu^{-1}G\nu$ and replace M with

$$\nu^{-1}G\nu \cap \mathcal{X} = \nu^{-1}G\nu \cap \nu^{-1}\mathcal{X}\nu = \nu^{-1}M\nu = \langle X_1, \dots, X_m\rangle.$$

Therefore, let us assume without loss of generality that $M = \langle X_1, \dots, X_m\rangle$ for some m.

We know G contains a maximal abelian subgroup A of $\mathcal{P}_n$, since π is semi-Clifford. Applying Lemma 2.17 on A and M, we get a maximal abelian subgroup A' of $\mathcal{P}_n$ with $M \subseteq A' \subseteq$

$\langle A, M\rangle \subseteq G$. We claim that $A' = \langle X_1, \ldots, X_m, Z_{m+1}, \ldots, Z_n\rangle$ up to phase. To see this, take a basis $X_1, \ldots, X_m, W_{m+1}, \ldots, W_n$ for A'. Decompose $W_i = c_i X^{u_i} Z^{v_i}$. We can assume the first m indices of u_i are zeros. Since W_i commutes with $X_1, \ldots, X_m$, the first m indices of v_i must be zeros. It now suffices to show that $u_i = 0$ for all $m < i \leq n$.

Let $p = X^{u_i}$ and $d = c_i Z^{v_i}$. Since $W_i \in A' \subseteq G$, we have $W_i' = \pi W_i \pi^{-1} \in \mathcal{P}_n$. Note that

$$W_i' = \pi p d \pi^{-1} = (\pi p \pi^{-1})(\pi d \pi^{-1}),$$

where $\pi p \pi^{-1}$ is a permutation and $\pi d \pi^{-1}$ is diagonal. It follows from Lemma A.3 that this decomposition is unique and $\pi p \pi^{-1}, \pi d \pi^{-1} \in \mathcal{P}_n$. So $p, d \in G$. Since $p \in \mathcal{X}$, we have $p \in G \cap \mathcal{X} = M = \langle X_1, \ldots, X_m\rangle$. In other words, $u_i[j] = 0$ for all $j > m$. Therefore, we have $u_i = 0$ and $A' = \langle X_1, \ldots, X_m, Z_{m+1}, \ldots, Z_n\rangle$ up to phase. The theorem then follows from Lemma A.5. □

The following is the main theorem of this section.

Theorem A.6. *For any positive integer k, a permutation gate π is a semi-Clifford gate in $\mathcal{C}_{k+1}$ if and only if there exist Clifford permutations ϕ_1, ϕ_2 and a mismatch-free product μ of C^*X gates such that $\pi = \phi_1 \mu \phi_2$ and, in μ, each gate has at most k controls.*

Proof. For the "if" direction, π being semi-Clifford follows from Lemma A.1, and $\pi \in \mathcal{C}_{k+1}$ follows directly from [And24, Theorem D.3] (along with part 2 of Proposition 2.2).

For the "only if" direction, we apply Theorem A.2 to get a representation $\phi_1 \mu \phi_2$ where ϕ_1 and ϕ_2 are Clifford permutations and μ is a mismatch-free product of C^*X gates. As any two gates in such a product commutes, and every such gate is its own inverse, we can assume without loss of generality that no gate is repeated. By part 2 of Proposition 2.2, $\mu \in \mathcal{C}_{k+1}$. By Lemma 2.13, in the polynomial representation of μ^{-1}, every coordinate has degree at most k. Note that $\mu^{-1} = \mu$. If there is a gate in μ with $m > k$ controls, this would yield a monomial of degree m in μ which would not be canceled out. Therefore every gate in μ has at most k controls, as desired. □

Our result has two immediate corollaries.

Corollary A.7. *A permutation gate π is a semi-Clifford gate in $\mathcal{C}_3$ if and only if there exist Clifford permutations ϕ_1, ϕ_2 and a mismatch-free product μ of Toffoli gates such that $\pi = \phi_1 \mu \phi_2$.*

Corollary A.8. *Every semi-Clifford permutation gate is in $\mathcal{C}_n$. (Thus every semi-Clifford permutation gate is in $\mathcal{CH}$.)*

Proof. The claim follows from Theorems A.2 and A.6 and the fact that a C^*X gate on n qubits has at most $n - 1$ controls. □

Appendix B. The Smallest Non–Semi-Clifford Permutation

Lemma B.1. *All permutations in $\mathcal{C}_3$ on at most six qubits are semi-Clifford.*

We give a computer-assisted proof for Lemma B.1 using `C++` and Magma [BCP97]. Our code can be found at `github.com/Likable-outlieR/clifford-hierarchy`.

Proof. By Lemma 2.18, it suffices to show that all permutations on exactly six qubits in $\mathcal{C}_3$ are semi-Clifford.[4] Then, by Theorem 3.2, it suffices to show that any permutation in $\mathcal{C}_3$ on six qubits that is written as the product of Toffoli gates in staircase form is semi-Clifford.

Now note that, given a permutation gate in staircase form, any two Toffoli gates with the same target commute; then we can cancel out any repeat appearances of a gate (note that $\mathrm{TOF}_{i,j,k}$ and $\mathrm{TOF}_{j,i,k}$ are the same, so we assume all the gates have $i < j < k$ here and for the rest of this

[4]Note that the process of adding inert qubits to a gate (that is, qubits on which the gate acts as the identity) preserves the property of being in $\mathcal{C}_k$, for any k, by induction on k. Thus, if we had a non-semi-Clifford permutation in $\mathcal{C}_3$ on fewer than six qubits, adding inert qubits to it yields a six-qubit permutation gate in $\mathcal{C}_3$ that is not semi-Clifford, by Lemma 2.18.

theorem's proof); let us do so. Then we can see that a gate in staircase form only depends on which gates appear, since the order of gates is fixed by the staircase form up to reordering gates with the same target, and such gates commute anyway. Thus the number of permutations to consider is upper bounded by 2^{20} (since now there are $\binom{6}{3} = 20$ relevant Toffoli gates).

At this point a computer search is viable; we show by checking all 2^{20} options that for any six-qubit permutation π in staircase form, if π is in $\mathcal{C}_3$ and is not mismatch-free, then there must exist a maximal abelian subgroup A of $\mathcal{P}_6$ such that $\pi A \pi^{-1}$ is a maximal abelian subgroup of $\mathcal{P}_6$. In fact, we show a stronger result (by exhaustively checking) that A can always be chosen by the following list:

- $\langle Z_1, Z_2, Z_3Z_4, X_3X_4, X_5, X_6\rangle$,
- $\langle Z_1, Z_2, Z_3, Z_4Z_5, X_4X_5, X_6\rangle$,
- $\langle Z_1, Z_2, Z_3Z_5, Z_4, X_3X_5, X_6\rangle$,
- $\langle Z_1, Z_2, Z_3Z_5, Z_4Z_5, X_3X_4X_5, X_6\rangle$, and
- $\langle Z_1, Z_2, Z_3Z_4, Z_5, X_3X_4, X_6\rangle$. □

The following result is immediate with Proposition 5.8 and Lemma B.1.

Theorem B.2. *The smallest number of qubits for which there exists a non-semi-Clifford permutation in $\mathcal{C}_3$ is 7.*