Characterization of Permutation Gates in the Third Level of the Clifford Hierarchy

Zhiyang He
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, USA

Luke Robitaille
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, USA

Xinyu Tan
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, USA

Abstract

The Clifford hierarchy is a fundamental structure in quantum computation whose mathematical properties are not fully understood. In this work, we characterize permutation gates – unitaries which permute the 2^n basis states – in the third level of the hierarchy. 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 by Clifford permutations. We then present necessary and sufficient 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.

2012 ACM Subject Classification Theory of computation → Quantum information theory

Keywords and phrases Quantum fault-tolerance, Clifford hierarchy, permutation gates

Digital Object Identifier 10.4230/LIPIcs.TQC.2026.2

Related Version *Full Version*: `https://arxiv.org/pdf/2510.04993` [24]

Funding *Zhiyang He*: National Science Foundation Graduate Research Fellowship under Grant No. 2141064.
Xinyu Tan: 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.

Acknowledgements 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 Jonas Anderson, Jeongwan Haah, Aram Harrow, Greg Kahanamoku-Meyer, Andrey Khesin and Anirudh Krishna for helpful discussions.

1 Introduction

The Clifford hierarchy is a ubiquitous structure in quantum computation which classifies unitary operations based on their conjugate actions on the Pauli group. Precisely, the first level $\mathcal{C}_1$ of the hierarchy $\mathcal{CH}$ is the Pauli group $\mathcal{P}$, and its subsequent levels are defined recursively: $\mathcal{C}_k$ is the set of all unitaries U such that $UPU^\dagger \in \mathcal{C}_{k-1}$ for all Pauli operators P. Within $\mathcal{CH}$, gates in the third level are of unique importance to the study of fault-tolerant quantum computation (FTQC) [33, 34, 19], as established by several foundational works. The Gottesman-Knill theorem [20] states that any circuits made of Pauli and Clifford gates, which are the first two levels of $\mathcal{CH}$, can be efficiently simulated by a classical computer. In contrast,

adding any non-Clifford gate to the Clifford group forms a universal gate set. Among the non-Clifford unitaries, gates in $\mathcal{CH}$ can be implemented fault-tolerantly by gate teleportation [21], where higher-level gates are performed using resource states and lower-level gates. From this perspective, the non-Clifford gates that are in $\mathcal{C}_3$ are the "easiest-to-implement" non-Clifford gates. Consequently, $\mathcal{C}_3$ gates have been at the center of study for FTQC, with decades of extensive research studying their fault-tolerant implementations [34, 8], synthesis into unitaries [15], algorithmic resource costs [14] and more.

Despite its importance, the mathematical structure of the Clifford hierarchy is not fully understood. Notably, $\mathcal{C}_k$ no longer forms a group for any $k \geq 3$, as it is neither closed under multiplication nor closed under inverse. Ample work has been done to elucidate structures within the hierarchy, which we briefly discuss in Section 1.3.

In this paper, we study the permutation gates, which are n-qubit unitaries that permute the 2^n computational basis states, in $\mathcal{C}_3$. We denote these gates by $\mathcal{C}_3^{\text{sym}}$, and remark that they are both interesting and important for a few reasons. First of all, permutation gates correspond to all reversible classical computations on n bits. Gates in $\mathcal{C}_3$, on the other hand, can be implemented fault-tolerantly using resource states and Clifford gates. $\mathcal{C}_3^{\text{sym}}$ therefore captures classical gates and computational subroutines which are relatively low-cost to implement for FTQC. The canonical example is the Toffoli gate TOF, which is classically universal and ubiquitous in quantum circuits.

Unfortunately, products of Toffoli gates (which capture all permutations) are notoriously unruly. For instance, while a single Toffoli gate is in $\mathcal{C}_3$, a product of as few as two Toffoli gates can leave the hierarchy.[1] Prior work by Anderson [1] conjectured that $\mathcal{C}_3^{\text{sym}}$ are precisely products of pairwise commuting Toffoli gates (up to multiplying on both sides by Clifford permutations), and that $\mathcal{C}_k^{\text{sym}}$ is closed under inverse for all k. We disprove both of these conjectures in this work.

Furthermore, permutation gates are important components for gates in $\mathcal{CH}$. Beigi and Shor [5] proved that all gates in $\mathcal{C}_3$ are *generalized semi-Clifford*, which means they can be written as $\phi_1 \pi d \phi_2$ for Clifford gates ϕ_1, ϕ_2, a diagonal gate d, and a permutation gate π. The conjecture that all gates in $\mathcal{CH}$ are generalized semi-Clifford remains open [36]. In [13], Cui, Gottesman, and Krishna fully characterized all the diagonal gates in $\mathcal{CH}$. Therefore, characterizing the permutation gates is a crucial step towards characterizing all gates in $\mathcal{C}_3$ and potentially $\mathcal{CH}$.

In this work, we present a full characterization of $\mathcal{C}_3^{\text{sym}}$ (Result 1, Result 2), elucidating an essential substructure of $\mathcal{C}_3$. We then present a family of gates in $\mathcal{C}_3^{\text{sym}}$, which disproves Anderson's conjectures and challenges prior understanding of $\mathcal{C}_3$ (Result 3). Finally, we derive lower bounds on the number of qubits a gate in $\mathcal{C}_3^{\text{sym}}$ must be supported on given the "complexity" of the function it implements, showing that our construction is optimal (Result 4).

1.1 Main results and techniques

To characterize the gates in $\mathcal{C}_3^{\text{sym}}$, we study structured products of Toffoli gates. We define a product of distinct Toffoli gates to be in *staircase form* if each gate $\text{TOF}_{i,j,k}$ in the product has $i, j < k$ and the target qubit indices are nondecreasing in the order that the gates are applied. For example, the gate in Figure 1 is in staircase form.

[1] Specifically, $\text{TOF}_{1,2,3}\text{TOF}_{1,3,2}$ is not in $\mathcal{CH}$; see equation (E.2) of [1].

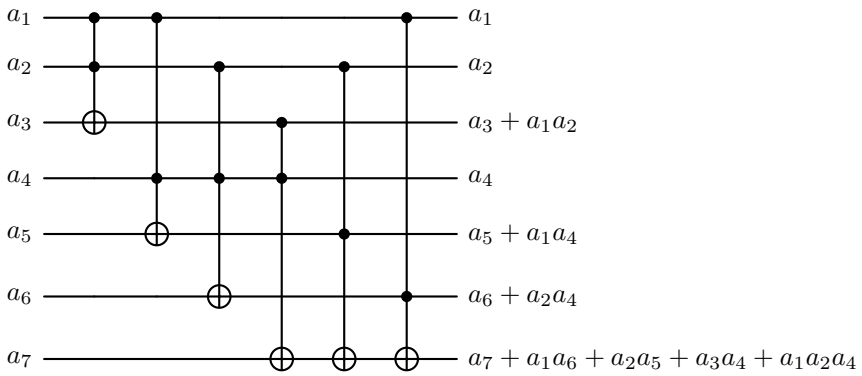

Figure 1 A permutation gate in staircase form, consisting of 6 Toffoli gates.

Our first main result, which was presented in an earlier version of this paper [23], states that every permutation in $\mathcal{C}_3^{\text{sym}}$ can be written in staircase form.

▶ **Result 1** (Theorem 25). *If $\pi \in \mathcal{C}_3$ is a permutation gate, then there exist Clifford permutations ϕ_1, ϕ_2 and a product μ of Toffoli gates in* staircase form *such that $\mu \in \mathcal{C}_3$ and $\pi = \phi_1\mu\phi_2$.*

We remark that there are permutations in staircase form which are not in $\mathcal{C}_3$. To derive a full characterization, we consider bilinear products defined over vectors of $\mathbb{F}_2^n$, which we denote by juxtaposition. For a commutative and associative bilinear product, we call it a *descending multiplication* if for standard basis vectors $\{e_i\}_{i\in[n]}$, it holds that $e_ie_i = 0$, and e_ie_j (for $i < j$) is in the span of $\{e_k : k > j\}$ (see Definition 33). Our second main result identifies a one-to-one correspondence between gates in $\mathcal{C}_3^{\text{sym}}$ and descending multiplications.

▶ **Result 2** (Theorem 34). *Every staircase form permutation gate in $\mathcal{C}_3$ induces a descending multiplication over $\mathbb{F}_2^n$. In correspondence, every descending multiplication over $\mathbb{F}_2^n$ induces a staircase form permutation gate in $\mathcal{C}_3$.*

Using this characterization, we construct a novel family of permutation gates in $\mathcal{C}_3$. For each integer $k \geq 3$, we take $n = 2^k - 1$ and label each standard basis vector of $\mathbb{F}_2^n$ by e_S for an nonempty subset $S \subseteq [k]$. We define a bilinear product operation by setting $e_Se_T = e_{S\cup T}$ if $S \cap T = \varnothing$, and $e_Se_T = 0$ if $S \cap T \neq \varnothing$, and extending linearly. This gives a descending multiplication, which induces a staircase form $\mathcal{C}_3$ permutation gate U_k.

▶ **Result 3** (Theorem 42). *For all $k \geq 3$, we have $U_k \in \mathcal{C}_3$ but $U_k^\dagger \notin \mathcal{C}_k$.*

We note that the permutation in Figure 1 is precisely U_3. This construction brings new understanding to the study of $\mathcal{C}_3$ gates in a few ways. First of all, $\{U_k\}_{k\geq 3}$ are the first examples of permutation gates in $\mathcal{C}_3$ whose inverses leave $\mathcal{C}_3$ (in fact, leave any fixed level of $\mathcal{CH}$). As a result, they disprove both conjectures of Anderson [1]. Previously, Gottesman and Mochon presented a gate in $\mathcal{C}_3$ (which is not a permutation; see Lemma 10 of main text) whose inverse is not in $\mathcal{C}_3$. Our U_3 is actually equivalent to their example up to conjugation by a Clifford gate.

The permutation implemented by U_k is also interesting. For a permutation π on n bits, we represent it by n Boolean polynomials in the n input variables, one polynomial for each output bit. For example, $\text{CNOT}_{1,2}$ maps $(a_1, a_2) \mapsto (a_1, a_2 + a_1)$ and $\text{TOF}_{1,2,3}$ maps $(a_1, a_2, a_3) \mapsto (a_1, a_2, a_3 + a_1a_2)$. For U_k, if we consider qubits $1, 2, 4, 8, \ldots, 2^{k-1}$

as "controls", and set the other $2^k - 1 - k$ "ancilla" input variables to be zero, then the $2^k - 1$ output polynomials representing U_k are precisely all the non-constant monomials of $a_1, a_2, a_4, a_8, \ldots, a_{2^{k-1}}$. In other words, arbitrarily powerful classical computation can be implemented using one $\mathcal{C}_3$ gate followed by CNOT gates, at the expense of a large number of ancilla qubits. This extra ancilla cost is, in fact, optimal: using the characterization with descending multiplications, we show the following lower bound.

▶ **Result 4** (Theorem 47). *Any permutation gate in $\mathcal{C}_3$ with an output bit represented by a polynomial of degree at least k must be supported on at least $2^k - 1$ qubits.*

1.2 Future directions

Combining these four results, our characterization discovers and delineates a novel design space within $\mathcal{C}_3^{\text{sym}}$. By constructing descending multiplications, which are easier to reason with than Toffoli circuits, we can derive permutation gates in $\mathcal{C}_3$ which encode different classical computations. These computations can then be implemented in FTQC via gate teleportation, where the majority of the cost is offloaded to resource state preparation. It would be interesting to explore whether the cost of important FTQC primitives, such as arithmetic, can be reduced with this approach.

In addition to gate design and teleportation, it would be interesting to explore whether our results can be generalized to higher levels of $\mathcal{CH}$. Evidently, characterizing more or all of the permutation gates in $\mathcal{CH}$ is consequential to our understanding of $\mathcal{CH}$. Our results mark a concrete step in this direction.

Another important direction of research is to understand the product of permutation and diagonal gates. Significant progress was made in [1], where Anderson showed several results which, in the case of $\mathcal{C}_3$, imply that if $\pi d \in \mathcal{C}_3$, then $\pi \in \mathcal{C}_3$ and $d \in \mathcal{CH}$. Since every gate in $\mathcal{C}_3$ is generalized semi-Clifford, given our characterization of permutation gates in $\mathcal{C}_3$ and the characterization of diagonal gates in $\mathcal{CH}$ by [13], can we precisely characterize all of $\mathcal{C}_3$? Such a result would have significant implications for our study of FTQC, given the unique importance of $\mathcal{C}_3$ gates.

Finally, our Result 4 raises an interesting complexity theory question. Specifically, while $\{U_k\}_{k\geq 3}$ demonstrate that polynomials of arbitrarily high degree can be encoded into a single $\mathcal{C}_3$ gate, such encodings necessarily incur an ancilla cost exponential in k. In contrast, implementing a degree k monomial in $\mathcal{C}_{k+1}$ takes only one k-controlled NOT gate, supported on $k+1$ qubits. How does this tradeoff between polynomial degree and ancilla cost transform through the different levels of $\mathcal{CH}$? What implications does this tradeoff have on the fundamental cost of gate teleportation, which is ubiquitous in FTQC?

1.3 Prior works

Since its introduction by Gottesman and Chuang in 1999 [21], the Clifford hierarchy has been studied by many prior works. Early works defined and studied the (generalized) semi-Clifford gates [37, 18, 22, 36, 5]. Subsequent works have elucidated various structures within the qubit Clifford hierarchy [6, 30, 32, 1, 4, 2] and the qudit Clifford hierarchy [13, 16, 17]. As the original motivation comes from FTQC, prior works have studied characterizations of and improved gate teleportation protocols for semi-Clifford gates [37, 36, 16, 12]; efficient resource state preparation methods, notably magic state distillation [8, 7, 10, 28, 35], for various non-Clifford gates; and constructions and limitations of quantum codes which admit transversal implementation of gates in different levels of the hierarchy [9, 3, 27, 26, 25].

1.4 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 2.2 discusses representation of 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 correspondence 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 Theorem 42 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. All omitted proofs can be found in the full version of this paper [24].

2 Preliminaries

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 $\mathcal{X} = \{I_2, X\}^{\otimes n}$. 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).

The action of CNOT can be viewed 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 $\mathrm{CNOT}_{3,1}$ is a CNOT gate with the third qubit as control and the first qubit as target.

2.1 Gates in the Clifford hierarchy

The Clifford hierarchy is defined recursively, with the first level being the Pauli gates.

▶ **Definition 5** (The Clifford hierarchy). *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 $\mathcal{CH} := \mathcal{C}_1 \cup \mathcal{C}_2 \cup \mathcal{C}_3 \cup \ldots$ is called the* Clifford hierarchy*; we refer to $\mathcal{C}_k$ as the k-th layer of $\mathcal{CH}$.*

We list a few standard facts of the Clifford hierarchy.

▶ **Fact 6.**

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. Note that $\mathcal{X}$ is exactly the set of all permutation gates in $\mathcal{P}_n$. A gate is called *diagonal* if its associated matrix is diagonal.

▶ **Definition 7** (Semi-Clifford and generalized semi-Clifford gates). *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 $\mathcal{P}_n$, denote by $\mathrm{span}(A)$ its linear span with complex coefficients. The lemma below shows equivalent definitions of (generalized) semi-Clifford gates.

▶ **Lemma 8.** *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 $U\,\mathrm{span}(A_1)U^\dagger = \mathrm{span}(A_2)$.*

We note that in literature, semi-Clifford and generalized semi-Clifford are usually defined as in Lemma 8 whereas Definition 7 is proved as a proposition [18, 36, 5, 1]. For proofs of this equivalence, we refer readers to Appendix A of [1].

▶ **Proposition 9** (Semi-Clifford gates are closed under taking inverses). *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 7, we can write $U = \phi_1 d \phi_2$ for some Clifford gates ϕ_1, ϕ_2 and a diagonal gate d. Using Fact 6 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 $U^{-1} = \phi_2^{-1} d^{-1} \phi_1^{-1} \in \mathcal{C}_k$. ◀

It was conjectured in [36] that all gates in $\mathcal{C}_3$ are semi-Clifford. This was disproved by Gottesman and Mochon [5] with the following 7-qubit unitary.

▶ **Lemma 10.** *For $n = 7$, $\mathcal{C}_3$ contains a non-semi-Clifford element.*

Proof. Let G be the 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 2 for a circuit diagram.

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 9. However, a computer calculation shows that $G^{-1} \notin \mathcal{C}_3$ (in particular, $G^{-1} X_7 G \notin \mathcal{C}_2$). Thus, G is not semi-Clifford. ◀

This 7-qubit operator is the smallest known example of a non-semi-Clifford operator in $\mathcal{C}_3$. [36] showed that for $n \leq 3$, all elements of $\mathcal{C}_3$ are semi-Clifford. Recently, [2] showed that for $n = 4$, all elements of $\mathcal{C}_3$ are semi-Clifford. For $n = 5$ or 6, it is an open problem whether there is a $\mathcal{C}_3$ operator that is non-semi-Clifford. We make partial progress on this problem in Appendix A of the full version of this paper [24] by showing that all permutation gates in $\mathcal{C}_3$ on at most six qubits are semi-Clifford.

While the semi-Clifford characterization of $\mathcal{C}_3$ was disproved, Beigi and Shor [5] proved that every gate in $\mathcal{C}_3$ is generalized semi-Clifford.

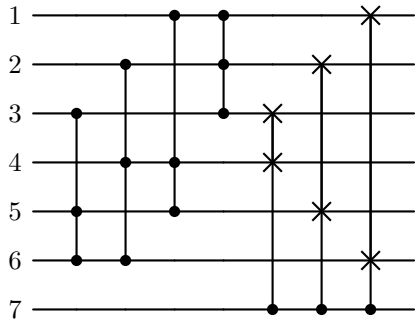

Figure 2 Circuit diagram for the Gottesman–Mochon seven-qubit gate G (with time flowing from left to right).

▶ **Theorem 11.** *Every element of $\mathcal{C}_3$ is generalized semi-Clifford.*

The conjecture that every gate in $\mathcal{CH}$ is generalized semi-Clifford [36] remains open. Partial progress was made in [1].

▶ **Conjecture 12.** *Every element of the Clifford hierarchy is generalized semi-Clifford.*

▶ Remark 13. A generalized semi-Clifford gate takes the form $\phi_1 \pi d \phi_2$. A similar form of $\pi d \phi$ is considered in the context of approximate unitary designs or pseudorandom unitaries in [29, 11], where ϕ and π are sampled uniformly at random from their respective groups, and d is a diagonal gate with random ± 1 entries.

2.2 Polynomial representations of permutations

To study the permutation gates in the Clifford hierarchy, we represent them as collections of Boolean polynomials.

▶ **Fact 14.** *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.*

▶ **Lemma 15.** *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 [13, See Eq. (1)]. ◀

▶ **Definition 16** (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 Fact 14 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 $(a_1, a_2, a_3) \mapsto (a_1, a_2, a_3 + a_1 a_2)$. We prove a few useful lemmas regarding the polynomial representations of permutation gates.

▶ **Lemma 17.** *For any integer $k \geq 1$ 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

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

By Lemma 15, one has that π_i^{-1}, the i-th coordinate of π^{-1}, must have degree at most k. ◀

▶ Remark 18. For $\pi \in \mathcal{C}_3$, Lemma 17 tells us that every coordinate of π^{-1} has degree at most 2; however, as we will see in Section 5, 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 $\mathbb{F}_2^n$.

▶ **Proposition 19** (Clifford permutations). *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.*

Recall that $\mathcal{X}$ is the set of all the n-qubit Pauli X operators.

▶ **Proposition 20** (Pauli permutations). *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 $\nu|0^n\rangle = |0^n\rangle$ and $\nu X_i \nu^{-1} = X'_i$ for all $i \in [m]$.*

The proofs of Propositions 19 and 20 can be found at the full version of this paper [24].

2.3 Anderson's conjectures

Besides polynomials, we consider the more operational representation of permutation gates as products of multi-controlled NOT gates, which we denote by C^kX for $k \geq 0$ (note that Toffoli is C^2X). In [1], Anderson considered *mismatch-free* circuits, which are products of pairwise commuting multi-controlled NOT gates.

▶ **Theorem 21** (Theorem D.4 of [1]). *A mismatch-free permutation circuit is in $\mathcal{CH}$ at the level of the highest-level gate in the circuit.*

Following this classification, Anderson gave two conjectures on permutation gates in $\mathcal{CH}$.

▶ **Conjecture 22.** *A permutation is in $\mathcal{C}_3$ if and only if it can be written as a circuit of commuting Toffoli gates, up to left and right multiplication by Clifford gates.*

▶ **Conjecture 23.** *For a permutation $\pi \in \mathcal{C}_k$, we have $\pi^\dagger \in \mathcal{C}_k$.*

In Appendix B of the full version of this paper [24], we show that mismatch-free permutation circuits are precisely the semi-Clifford permutation gates. However, as we will see in the next few sections, the permutation gates in $\mathcal{C}_3$ form a much richer space, which we characterize in this work. In particular, we disprove both Conjecture 22 and Conjecture 23.

3 Staircase form representations of $\mathcal{C}_3$ permutations

We begin by presenting the most important definition of our work.

▶ **Definition 24** (Staircase form Toffoli circuits). *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 3 for an example.*

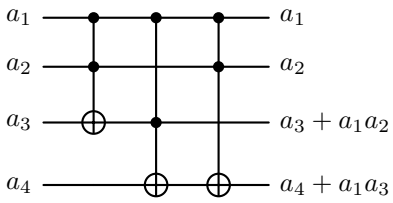

Figure 3 This circuit for $\mathrm{TOF}_{1,2,4}\mathrm{TOF}_{1,3,4}\mathrm{TOF}_{1,2,3}$ is in staircase form but not mismatch-free, as qubit 3 is used as a control for $\mathrm{TOF}_{1,3,4}$ and a target for $\mathrm{TOF}_{1,2,3}$.

Note that a staircase form Toffoli circuit is unique up to ordering of gates with the same target. We say that a permutation is in staircase form if it can be written as a Toffoli circuit in staircase form.

Our main result of this section is the following.

▶ **Theorem 25.** *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 $\pi = \phi_1\mu\phi_2$.*

Since $\pi \in \mathcal{C}_3$ is a permutation, each $\pi X_j \pi^{-1}$ is a Clifford permutation. Then by Proposition 19, we can find a binary matrix A_j and a vector b_j such that $\pi X_j \pi^{-1}$ implements the permutation $|v\rangle \mapsto |v + A_j v + b_j\rangle$. To prove Theorem 25, we construct a sequence of Clifford operators to reduce A_j, b_j to a specific form, which will help us build a staircase form representation of π.

We caution that Theorem 25 is not if-and-only-if, because there exist permutations in staircase form that are not in $\mathcal{C}_3$. For example, $\pi' = \mathrm{TOF}_{3,4,5}\mathrm{TOF}_{1,2,3}$ is in staircase form, but $\pi' X_1 \pi'^{-1} = X_1\mathrm{CNOT}_{2,3}\mathrm{TOF}_{2,4,5} \notin \mathcal{C}_2$.

3.1 Reducing $\mathcal{C}_3$ permutations

The goal of this subsection is to reduce a $\mathcal{C}_3$ permutation to a much more friendly form, which serves as the first step in proving Theorem 25:

▶ **Lemma 26** (Reducing $\mathcal{C}_3$ permutations)**.** *For any permutation gate $\pi \in \mathcal{C}_3$, there exist Clifford permutations ϕ_1, ϕ_2 such that the following is true. Let $\tau = \phi_1\pi\phi_2$. Then $\tau|0^n\rangle = |0^n\rangle$ and for each i, we have $\tau|e_i\rangle = |e_i\rangle$ and $\tau X_i \tau^{-1}$ is given by $|v\rangle \mapsto |v + A_i v + e_i\rangle$ for some strictly lower triangular matrix A_i over $\mathbb{F}_2$.*

To prove Lemma 26, we will use a series of helper lemmas which can be proved via standard linear algebra arguments.

The first helper lemma stated below is essentially the same as standard results on simultaneous triangularization of commuting nilpotent matrices; see, for example, [31]. We include a proof in the full version of this paper [24].

▶ **Lemma 27.** *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*.*

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

▶ **Lemma 28.** *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 28 will be used tacitly throughout what follows.

▶ **Proposition 29.** *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 $A_i + A_j + A_iA_j$ and update b_i to be $b_i + b_j + A_ib_j$.*

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.

Using Propositions 19, 20, and 29 and Lemma 27, we can prove Lemma 26. The omitted proofs of Proposition 29 and Lemma 26 can be found at the full version of this paper [24].

3.2 Proof of Theorem 25

Now we only need one more ingredient to prove Theorem 25: the polynomial representations of staircase form permutations, whose proof follows intuitively from the circuit.

▶ **Lemma 30** (Characterization of staircase form permutation via the polynomial representation of its inverse)**.** *A permutation gate π is staircase form if and only if, in the polynomial representation of π^{-1}, for all k, the k-th coordinate is a_k plus a (possibly empty) sum of terms of the form a_ia_j with $i < j < k$.*

Furthermore, given a permutation π written as a staircase form Toffoli circuit, for any $i < j < k$, we have that $\mathrm{TOF}_{i,j,k}$ appears in the product if and only if the k-th coordinate in the polynomial form of π^{-1} contains an a_ia_j term.

Proof. 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. In that product, whenever a gate is applied, its controls have never been targeted so far, and thus are unchanged from the input. This means that, when performing the gates of π^{-1} in order, a gate $\mathrm{TOF}_{i,j,k}$ adds a term of a_ia_j to the k-th coordinate of the output. All parts of the desired result now can be easily shown. ◀

We are now ready to prove Theorem 25, restated below for convenience.

▶ **Theorem 25.** *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 $\pi = \phi_1\mu\phi_2$.*

Proof. By Lemma 26, we can assume without loss of generality that $\pi\,|0^n\rangle = |0^n\rangle$ and for each $i \in [n]$, $\pi\,|e_i\rangle = |e_i\rangle$ and $\pi X_i \pi^{-1}$ is given by $|v\rangle \mapsto |v + A_iv + e_i\rangle$ for some strictly lower triangular matrix A_i over $\mathbb{F}_2$.

We now 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

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

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

We now build a polynomial representation (see Definition 16) for π^{-1}. By Lemma 17, 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

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

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 30 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 Fact 6, since $\phi_1^{-1}, \phi_2^{-1} \in \mathcal{C}_2$ and $\pi \in \mathcal{C}_3$. ◀

3.3 A consequence of Theorem 25

Below, in Corollary 32, we strengthen the result of Beigi and Shor [5], who showed that any $\mathcal{C}_3$ element is generalized semi-Clifford (restated in Theorem 11). Our corollary sharpens this by showing that the underlying permutation can be taken to be in staircase form.

▶ **Lemma 31.** *For any permutation gate π and diagonal gate d, if $\pi d \in \mathcal{C}_3$, then $\pi \in \mathcal{C}_3$.*

Proof. This is a special case of Corollary A.11.2 from [1]. ◀

▶ **Corollary 32.** *For any $\psi \in \mathcal{C}_3$, there exist $\phi_1, \phi_2 \in \mathcal{C}_2$, a diagonal gate d, and a staircase form permutation $\pi \in \mathcal{C}_3$ with $\psi = \phi_1 \pi d \phi_2$.*

Proof. By Theorem 11, we can write $\psi = \phi_3 \pi_0 d_0 \phi_4$ for $\phi_3, \phi_4 \in \mathcal{C}_3$, a permutation gate π_0, and a diagonal gate d_0. Now $\pi_0 d_0 = \phi_3^{-1} \psi \phi_4^{-1} \in \mathcal{C}_3$, so $\pi_0 \in \mathcal{C}_3$ by Lemma 31. Then, by Theorem 25, we can write $\pi_0 = \phi_5 \mu \phi_6$ for Clifford permutations ϕ_5, ϕ_6 and a staircase form permutation $\mu \in \mathcal{C}_3$. Now $\phi_6 d_0 \phi_6^{-1}$ is diagonal as ϕ_6 is a permutation and d_0 is diagonal. Let $\phi_1 = \phi_3 \phi_5$, $\pi = \mu$, $d = \phi_6 d_0 \phi_6^{-1}$, $\phi_2 = \phi_6 \phi_4$. Then $\phi_1, \phi_2 \in \mathcal{C}_2$, π is a staircase form permutation in $\mathcal{C}_3$, and d is diagonal, and

$$\phi_1 \pi d \phi_2 = \phi_3 \phi_5 \mu \phi_6 d_0 \phi_6^{-1} \phi_6 \phi_4 = \phi_3 \phi_5 \mu \phi_6 d_0 \phi_4 = \phi_3 \pi_0 d_0 \phi_4 = \psi.$$ ◀

4 Descending multiplications

As we proved that every permutation in $\mathcal{C}_3$ can be written in staircase form, in this section we show sufficient conditions for a staircase form permutation to be in $\mathcal{C}_3$. Previously we have been working with $\mathbb{F}_2^n$ as a vector space without any multiplicative structure. Our next definition captures bilinear products defined over $\mathbb{F}_2^n$ which encode staircase form Toffoli circuits in $\mathcal{C}_3$.

▶ **Definition 33** (Descending multiplications). *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 will also write 0 and 0^n interchangeably.

We now state the main theorem of this section.

▶ **Theorem 34** (A bijection between descending multiplications and $\mathcal{C}_3$ permutations in staircase form). *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*

$$\forall S \subseteq [n], \quad \pi \left| \sum_{i \in S} e_i \right\rangle = \left| \sum_{T \subseteq S, T \neq \varnothing} \prod_{i \in T} e_i \right\rangle . \tag{2}$$

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

$$\pi \, |e_i + e_j\rangle = |e_i + e_j + e_i e_j\rangle \, . \tag{3}$$

The multiplication is then extended linearly.

We break the proof of this theorem into several propositions. We refer readers to the full version [24] of this paper for the proofs of Propositions 35–38.

▶ **Proposition 35.** *For any descending multiplication, the resulting π from Equation (2) is indeed a staircase form permutation in $\mathcal{C}_3$.*

▶ **Proposition 36.** *For any staircase form permutation π in $\mathcal{C}_3$, the resulting operation from Equation (3) is indeed a descending multiplication.*

▶ **Proposition 37.** *Given a descending multiplication, applying the procedure in Equation (2) to get a permutation π, then applying the procedure in Equation (3) to that permutation, yields the original multiplication.*

▶ **Proposition 38.** *Given a staircase form permutation π in $\mathcal{C}_3$, applying the procedure in Equation (3) to get a descending multiplication, then applying the procedure in Equation (2) to that multiplication, yields the original π.*

Proof of Theorem 34. Combining Propositions 35–38 yields the theorem. ◀

5 A family of non-semi-Clifford Permutations

Utilizing our characterization of $\mathcal{C}_3$ permutations as descending multiplications, we construct an infinite family of permutation gates $\{U_k\}_{k \geq 3}$ in Definition 39 and prove in Theorem 42 that every U_k is non-semi-Clifford in $\mathcal{C}_3$. This family of gates disproves both of Anderson's conjectures (restated in Conjecture 22 and Conjecture 23 for convenience). We study the smallest case of $k = 3$ in Section 5.3, and show that this 7-qubit permutation U_3 is in fact conjugate to the Gottesman–Mochon example by a Clifford operator.

▶ **Definition 39** (A family of permutation gates in $\mathcal{C}_3$ from descending multiplications). *For each integer $k \geq 3$, let $n = 2^k - 1$. We label a basis of $\mathbb{F}_2^n$ as e_S for nonempty subsets $S \subseteq [n]$. Define a multiplication by setting $e_S e_T = e_{S \sqcup T}$ if $S \cap T = \varnothing$, and $e_S e_T = 0$ if $S \cap T \neq \varnothing$, 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 34.*

▶ **Proposition 40.** *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 $\mathrm{TOF}_{S,T,S \sqcup T}$; specifically, apply these gates in nondecreasing order of target.*

▶ Remark 41. From the perspective of labeling gates with integers instead of sets, Proposition 40 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.

The main feature of these permutations is captured by the following theorem.

▶ **Theorem 42** (Inverses of $\mathcal{C}_3$ permutations may lie outside $\mathcal{C}_k$). *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.*

5.1 Proof of Proposition 40

▶ **Proposition 43.** *Given a staircase form permutation π in $\mathcal{C}_3$ and $i < j$, let v_{ij} be such that $\pi^{-1}(e_i + e_j) = e_i + e_j + v_{ij}$. Then for all $i < j < k$, $\mathrm{TOF}_{i,j,k}$ 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 30. ◀

Proof of Proposition 40. Take $U_k = \pi$ in Proposition 43. We know from the proof of Proposition 35 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 43 implies the desired, by the definition of $e_S e_T$. ◀

5.2 Proof of Theorem 42

▶ **Proposition 44.** *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 Equation (2), we have that $\sum_{T \subseteq S; T \neq \varnothing} p_T = \sum_{T \subseteq S, T \neq \varnothing} \prod_{i \in T} e_i$ 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. ◀

We prove Theorem 42 through the polynomial representations of U_k and U_k^{-1}.

▶ **Lemma 45.** *For each integer $k \geq 3$, let us denote the polynomial representations (Definition 16) of U_k and U_k^{-1} respectively as $(\pi_S)_{S\subseteq[k],S\neq\varnothing}$ and $(\pi'_S)_{S\subseteq[k],S\neq\varnothing}$. Then*

$$\pi_S(a) = \sum_{m=1}^{|S|} \sum_{\substack{\bigsqcup_{i=1}^{m} T_i = S,\\ T_i \neq \varnothing}} a_{T_1} a_{T_2} \dots a_{T_m}, \tag{4}$$

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

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

where the sum is over unordered pairs T_1, T_2.

Proof. The polynomial form of π^{-1} follows directly from Proposition 40 and Lemma 30. 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 $e_{T_1 \sqcup \dots \sqcup T_m}$ if the T_i are pairwise disjoint, and 0 if not, so the desired follows from Proposition 44. ◀

Proof of Theorem 42. We know $U_k \in \mathcal{C}_3$ by definition. Also, by Lemma 45, we know that $\pi_{[k]}$ is a polynomial of degree k because it contains the monomial $a_{\{1\}}a_{\{2\}}\dots a_{\{k\}}$. This implies that $U_k^{-1} \notin \mathcal{C}_k$ using Lemma 17. In particular, $U_k^{-1} \notin \mathcal{C}_3$, so U_k is not semi-Clifford by Proposition 9. ◀

5.3 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 1):

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

Recall the Gottesman–Mochon 7-qubit gate (see Figure 2 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 10. Now 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 46.** *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 42. Checking that $FGF^{-1} = U_3$ is a direct computation. ◀

In particular, this implies that for all $n \geq 7$, $\mathcal{C}_3$ contains a non-semi-Clifford permutation (see Lemma A.3 in the full version [24]).

5.4 Lower bound on the number of qubits required for a $\mathcal{C}_3$ permutation

From Theorem 42, we know that U_k is a $\mathcal{C}_3$ permutation on $2^k - 1$ qubits that contains a degree-k monomial in its polynomial representation. Our next result shows that any such permutation in $\mathcal{C}_3$ must be supported on at least $2^k - 1$ qubits.

▶ **Theorem 47** (Lower bound on the number of qubits required for a $\mathcal{C}_3$ permutation)**.** *If π is a $\mathcal{C}_3$ permutation such that there is a degree-k monomial somewhere in the polynomial representation of π, then $n \geq 2^k - 1$; this bound is sharp for all $k \geq 3$.*

We prove Theorem 47 by taking advantage of the one-to-one correspondence between descending multiplications (Proposition 48) and $\mathcal{C}_3$ permutations in staircase form (Proposition 49).

▶ **Proposition 48.** *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 33, 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 $(\sum_{T\in F} p_T)p_V = 0$. 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 49.** *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 44; then Proposition 48 yields the desired. ◀

We are now ready to prove Theorem 47.

Proof of Theorem 47. Assume without loss of generality that $k \geq 2$. By Theorem 25, 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 49 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. ◀

5.5 Rejection of Anderson's conjectures

We now disprove Anderson's two conjectures.

▶ **Lemma 50.** *Two multi-controlled NOT 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. Assume for the sake of contradiction that 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 $BA|1^n\rangle \neq AB|1^n\rangle$, so they do not commute. ◀

▶ **Lemma 51.** *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 8. ◀

▶ **Proposition 52.** *Conjecture 22 and Conjecture 23 are false.*

Proof. Conjecture 23 is false by Theorem 42. For Conjecture 22, suppose it holds. Then by Lemma 50, every permutation in $\mathcal{C}_3$ is a mismatch-free product of Toffoli gates, up to Clifford permutations on the left and right. This implies that every permutation in $\mathcal{C}_3$ is semi-Clifford by Lemma 51, which results in a contradiction, as we know U_3 is a non-semi-Clifford permutation in $\mathcal{C}_3$ for $n = 7$. ◀

References

1 Jonas T. Anderson. On groups in the qubit Clifford hierarchy. *Quantum*, 8:1370, June 2024. doi:10.22331/q-2024-06-13-1370.

2 Jonas T Anderson and Andrew Connelly. Affine equivalence in the clifford hierarchy. *arXiv preprint*, 2025. arXiv:2507.14370.

3 Jonas T Anderson and Tomas Jochym-O'Connor. Classification of transversal gates in qubit stabilizer codes. *arXiv preprint arXiv:1409.8320*, 2014. doi:10.48550/arxiv.1409.8320.

4 Jonas T Anderson and Matthew Weippert. Controlled gates in the clifford hierarchy. *arXiv preprint*, 2024. arXiv:2410.04711.

5 Salman Beigi and Peter W. Shor. C_3, semi-Clifford and generalized semi-Clifford operations, 2009. arXiv:0810.5108.

6 Ingemar Bengtsson, Kate Blanchfield, Earl Campbell, and Mark Howard. Order 3 symmetry in the clifford hierarchy. *Journal of Physics A: Mathematical and Theoretical*, 47(45):455302, 2014.

7 Sergey Bravyi and Jeongwan Haah. Magic-state distillation with low overhead. *Physical Review A*, 86(5):052329, November 2012. doi:10.1103/physreva.86.052329.

8 Sergey Bravyi and Alexei Kitaev. Universal quantum computation with ideal Clifford gates and noisy ancillas. *Physical Review A*, 71(2):022316, 2005. URL: https://journals.aps.org/pra/abstract/10.1103/PhysRevA.71.022316.

9 Sergey Bravyi and Robert König. Classification of topologically protected gates for local stabilizer codes. *Physical review letters*, 110(17):170503, 2013.

10 Earl T Campbell, Hussain Anwar, and Dan E Browne. Magic-state distillation in all prime dimensions using quantum reed-muller codes. *Physical Review X*, 2(4):041021, 2012.

11 Chi-Fang Chen, Jeongwan Haah, Jonas Haferkamp, Yunchao Liu, Tony Metger, and Xinyu Tan. Incompressibility and spectral gaps of random circuits. *arXiv preprint arXiv:2406.07478*, 2024. doi:10.48550/arXiv.2406.07478.

12 Imin Chen and Nadish de Silva. Characterising semi-clifford gates using algebraic sets. *Communications in Mathematical Physics*, 405(9):201, 2024.

13 Shawn X Cui, Daniel Gottesman, and Anirudh Krishna. Diagonal gates in the Clifford hierarchy. *Physical Review A*, 95(1):012329, 2017. URL: S.X.Cui,D.Gottesman,andA.Krishna,DiagonalgatesintheCliffordhierarchy,PhysicalReviewA.

14 Alexander M Dalzell, Sam McArdle, Mario Berta, Przemyslaw Bienias, Chi-Fang Chen, András Gilyén, Connor T Hann, Michael J Kastoryano, Emil T Khabiboulline, Aleksander Kubica, et al. Quantum algorithms: A survey of applications and end-to-end complexities. *arXiv preprint*, 2023. arXiv:2310.03011.

15 Christopher M Dawson and Michael A Nielsen. The solovay-kitaev algorithm. *arXiv preprint quant-ph/0505030*, 2005.

16 Nadish de Silva. Efficient quantum gate teleportation in higher dimensions. *Proceedings of the Royal Society A*, 477(2251):20200865, 2021.

17 Nadish de Silva and Oscar Lautsch. The clifford hierarchy for one qubit or qudit. *arXiv preprint*, 2025. arXiv:2501.07939.

18 Jeroen Dehaene and Bart De Moor. Clifford group, stabilizer states, and linear and quadratic operations over gf (2). *Physical Review A*, 68(4):042318, 2003.

19 D. Gottesman. Stabilizer Codes and Quantum Error Correction. *arXiv:quant-ph/9705052*, 1997. doi:10.48550/arxiv.quant-ph/9705052.

20 Daniel Gottesman. The heisenberg representation of quantum computers, 1998. arXiv:quant-ph/9807006.

21 Daniel Gottesman and Isaac L. Chuang. Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations. *Nature*, 402(6760):390–393, November 1999. doi:10.1038/46503.

22 David Gross and Maarten Nest. The lu-lc conjecture, diagonal local operations and quadratic forms over gf (2). *arXiv preprint*, 2007. arXiv:0707.4000.

23 Zhiyang He, Luke Robitaille, and Xinyu Tan. Permutation gates in the third level of the clifford hierarchy. *arXiv preprint*, 2024. arXiv:2410.11818.

24 Zhiyang He, Luke Robitaille, and Xinyu Tan. Characterization of permutation gates in the third level of the clifford hierarchy. *arXiv preprint*, 2025. arXiv:2510.04993.

25 Zhiyang He, Vinod Vaikuntanathan, Adam Wills, and Rachel Yun Zhang. Quantum Codes with Addressable and Transversal Non-Clifford Gates. *arXiv preprint arxiv:2502.01864*, 2025. doi:10.48550/arxiv.2502.01864.

26 Jingzhen Hu, Qingzhong Liang, and Robert Calderbank. Climbing the diagonal clifford hierarchy, 2021. arXiv:2110.11923.

27 Tomas Jochym-O'Connor, Aleksander Kubica, and Theodore J Yoder. Disjointness of stabilizer codes and limitations on fault-tolerant logical gates. *Physical Review X*, 8(2):021047, 2018.

28 Anirudh Krishna and Jean-Pierre Tillich. Towards Low Overhead Magic State Distillation. *Physical Review Letters*, 123(7):070507, August 2019. doi:10.1103/physrevlett.123.070507.

29 Tony Metger, Alexander Poremba, Makrand Sinha, and Henry Yuen. Simple constructions of linear-depth t-designs and pseudorandom unitaries. *arXiv preprint arXiv:2404.12647*, 2024. doi:10.48550/arXiv.2404.12647.

30 Tefjol Pllaha, Narayanan Rengaswamy, Olav Tirkkonen, and Robert Calderbank. Un-weyl-ing the clifford hierarchy. *Quantum*, 4:370, 2020. doi:10.22331/Q-2020-12-11-370.

31 Heydar Radjavi and Peter Rosenthal. *Simultaneous Triangularization*. Springer Science & Business Media, 2000. doi:10.1007/978-1-4612-1200-3.

32 Narayanan Rengaswamy, Robert Calderbank, and Henry D Pfister. Unifying the clifford hierarchy via symmetric matrices over rings. *Physical Review A*, 100(2):022304, 2019.

33 Peter W Shor. Scheme for reducing decoherence in quantum computer memory. *Physical review A*, 52(4):R2493, 1995.

34 P.W. Shor. Fault-tolerant quantum computation. In *Proceedings of 37th Conference on Foundations of Computer Science*, pages 56–65, 1996. doi:10.1109/SFCS.1996.548464.

35 Adam Wills, Min-Hsiu Hsieh, and Hayata Yamasaki. Constant-overhead magic state distillation. *arXiv preprint*, 2024. arXiv:2408.07764.

36 Bei Zeng, Xie Chen, and Isaac L Chuang. Semi-Clifford operations, structure of C_k hierarchy, and gate complexity for fault-tolerant quantum computation. *Physical Review A*, 77(4):042313, 2008. URL: https://journals.aps.org/pra/abstract/10.1103/PhysRevA.77.042313.

37 Xinlan Zhou, Debbie W Leung, and Isaac L Chuang. Methodology for quantum logic gate construction. *Physical Review A*, 62(5):052316, 2000.