Third-level Clifford-hierarchy gates on at most six qubits are semi-Clifford

Yuxuan Zhang

Institute of Physics, École Polytechnique Fédérale de Lausanne (EPFL),
CH-1015 Lausanne, Switzerland

Department of Physics, Princeton University,
Princeton, New Jersey 08544, USA

26 September 2026

Abstract

A unitary is semi-Clifford if it is a diagonal gate up to Clifford gates on each side. Every gate in the third level of the Clifford hierarchy is semi-Clifford on at most three qubits (Zeng, Chen and Chuang) and on four qubits (Anderson and Connelly), while a seven-qubit gate of Gottesman and Mochon shows that the property fails in general; five and six qubits were open. We prove that every third-level gate on at most six qubits is semi-Clifford, so seven is the least number of qubits on which the third level contains a gate that is not semi-Clifford. More precisely, every such gate equals $C_1 D_1 \, \pi \, D_h D_2 C_2$ with Clifford gates C_1, C_2, diagonal Clifford gates D_1, D_2, a single layer π of Toffoli gates whose control and target qubits are disjoint, and a product D_h of T, controlled-S and doubly controlled-Z gates on the non-target qubits. The only non-textbook published results the proof uses are the theorem of Beigi and Shor that third-level gates are generalized semi-Clifford and an elementary factorization of Zeng, Chen and Chuang. From these, the correspondence between abelian regular subgroups of the affine group and nilpotent algebras shows that the permutation part is a single Toffoli layer whenever $n \le 6$, because a nonzero triple product forces seven dimensions; an exact computation then determines the admissible diagonal phases for each of the ten resulting classes. A standalone verification program using only the Python standard library is included.

1 Introduction

Let $\mathcal{P}_n$ be the n-qubit Pauli group with arbitrary global phases, consisting of the operators $e^{i\theta} P_1 \otimes \cdots \otimes P_n$ with $\theta \in \mathbb{R}$ and $P_j \in \{I, X, Y, Z\}$. The Clifford hierarchy of Gottesman and Chuang [1] is defined by

$$\mathcal{C}_1(n) := \mathcal{P}_n, \qquad \mathcal{C}_{k+1}(n) := \{U \in U(2^n) : \; UPU^\dagger \in \mathcal{C}_k(n) \text{ for every } P \in \mathcal{P}_n\}, \qquad k \ge 1, \quad (1)$$

so that $\mathcal{C}_2(n)$ is the Clifford group. A unitary is *semi-Clifford* if it can be written as $C_L D C_R$ with $C_L, C_R \in \mathcal{C}_2(n)$ and D diagonal in the computational basis; we write $\mathrm{SC}(n)$ for the set of such unitaries. Semi-Clifford gates can be implemented with a simpler teleportation circuit than general hierarchy gates, which is why the relation between $\mathcal{C}_3(n)$ and $\mathrm{SC}(n)$ has been studied since Zeng, Chen and Chuang [2]. They proved that every gate of every level is semi-Clifford for $n \le 2$ and that $\mathcal{C}_3(3) \subseteq \mathrm{SC}(3)$, and they conjectured that every third-level gate is semi-Clifford. Beigi and Shor [3] proved the weaker statement that every third-level gate is *generalized* semi-Clifford, and reported in their Section 1.1 a seven-qubit gate of Gottesman and Mochon in $\mathcal{C}_3(7) \setminus \mathrm{SC}(7)$. Anderson and Connelly [4] proved $\mathcal{C}_3(4) \subseteq \mathrm{SC}(4)$. The quantity

$$n_{\min} := \min\{n \ge 1 : \; \mathcal{C}_3(n) \setminus \mathrm{SC}(n) \ne \varnothing\} \quad (2)$$

was therefore known to lie in $\{5, 6, 7\}$; He, Robitaille and Tan [5] state explicitly that it is open whether a third-level gate on five or six qubits can fail to be semi-Clifford. The question is recorded as an open problem in [9], whose definitions (1)–(2) we follow.

Theorem 1. *Let $n \le 6$ and $U \in \mathcal{C}_3(n)$. There are Clifford gates C_1, C_2, diagonal Clifford gates D_1, D_2, a partition of the qubits into control qubits C, target qubits T and idle qubits I, a permutation gate $\pi_f : |x_\mathsf{C}, x_\mathsf{T}, x_\mathsf{I}\rangle \mapsto |x_\mathsf{C}, x_\mathsf{T} + f(x_\mathsf{C}), x_\mathsf{I}\rangle$ whose components f_τ are sums of products of two control bits, and a diagonal gate D_h in the group generated by T, controlled-S and doubly controlled-Z gates acting on $\mathsf{C} \cup \mathsf{I}$, such that*

$$U = C_1 D_1 \, \pi_f D_h \, D_2 C_2 = (C_1 D_1 H_\mathsf{T}) \, E \, (H_\mathsf{T} D_2 C_2), \qquad E = \operatorname{diag}((-1)^{\sum_{\tau \in \mathsf{T}} x_\tau f_\tau(x_\mathsf{C})}) \, D_h, \tag{3}$$

where H_T is the product of Hadamard gates on T. In particular $\mathcal{C}_3(n) \subseteq \mathrm{SC}(n)$ for every $n \le 6$.

Corollary 2. $n_{\min} = 7$.

Proof. Theorem 1 gives $n_{\min} \ge 7$. Proposition 17 exhibits a gate in $\mathcal{C}_3(7) \setminus \mathrm{SC}(7)$, so $n_{\min} \le 7$; this reproves the bound of Gottesman and Mochon [3, §1.1] without relying on it. □

Outline. By the theorem of Beigi and Shor and a factorization of Zeng, Chen and Chuang, a third-level gate is $C_L \Pi D C_R$ with Π a permutation and D diagonal (Section 3). We observe that every conjugate $\Pi X^b \Pi^{-1}$ is then an affine permutation, without needing Π itself to lie in the third level. Permutations with this property correspond to commutative nilpotent algebra structures on $\mathbb{F}_2^n$ with all squares zero; such an algebra has a nonzero triple product only if $n \ge 7$ (Section 4). Hence for $n \le 6$ the permutation is a single Toffoli layer up to affine maps, and the layers fall into ten classes (Section 5). For each class we determine the group of admissible diagonal phases exactly (Section 6); its structure yields the factorization (3) (Section 7). The computer-assisted parts are an orbit census for one step of the classification and the sizes of the ten phase groups; both are checked by the included program (Section 9).

2 Preliminaries

Computational basis states are $|x\rangle$, $x \in \mathbb{F}_2^n$, and $X^b Z^a |x\rangle = (-1)^{a \cdot x} |x + b\rangle$ for $a, b \in \mathbb{F}_2^n$. For a permutation s of $\mathbb{F}_2^n$ let $P_s|x\rangle = |s(x)\rangle$, and let $\tau_b(x) = x + b$. For a function $\theta : \mathbb{F}_2^n \to \mathbb{R}$ let $D_\theta = \operatorname{diag}(e^{i\theta(x)})$ and $(\partial_b \theta)(x) = \theta(x + b) - \theta(x)$; then

$$D_\theta X^b D_\theta^\dagger = X^b \operatorname{diag}(e^{i \, \partial_b \theta}). \tag{4}$$

Write $\omega = e^{i\pi/4}$. For $\lambda : \mathbb{F}_2^n \to \mathbb{Z}_8$ we write $D_\lambda = \operatorname{diag}(\omega^{\lambda(x)})$; every such λ has a unique expansion $\lambda = \sum_{S \subseteq [n]} c_S x^S$ with $c_S \in \mathbb{Z}_8$ and $x^S = \prod_{i \in S} x_i$, the $x_i \in \{0, 1\}$ being read as integers. Replacing x_i by $1 - x_i$ shows that the coefficient of x^R in $\partial_{e_i} \lambda$ is

$$c_{R \cup \{i\}} \text{ if } i \notin R, \qquad -2c_R \text{ if } i \in R. \tag{5}$$

Lemma 3. (i) *For $k \in \{2, 3\}$, $\mathcal{C}_k(n)$ is closed under left and right multiplication by Clifford gates, and so is $\mathrm{SC}(n)$. A unitary U lies in $\mathcal{C}_3(n)$ if and only if $UX_iU^\dagger, UZ_iU^\dagger \in \mathcal{C}_2(n)$ for $i = 1, \dots, n$.* (ii) *If $M = P_s D'$ with D' diagonal is a Clifford gate, then s is affine. Conversely P_s is a Clifford gate for every affine bijection s.* (iii) *If D_θ is a Clifford gate, then $\theta(x) - \theta(x') \in \frac{\pi}{2}\mathbb{Z}$ for all x, x'. The diagonal Clifford gates with phases in $\omega^{\mathbb{Z}_8}$ are exactly the D_λ with $\lambda \in \mathrm{Cliff}_n$, where* $\mathrm{Cliff}_n := \{\lambda : \ c_S = 0 \ (|S| \ge 3), \ c_S \in 4\mathbb{Z}_8 \ (|S| = 2), \ c_S \in 2\mathbb{Z}_8 \ (|S| = 1)\}$.

Proof. (i) $(CU)P(CU)^\dagger = C(UPU^\dagger)C^\dagger$ and $(UC)P(UC)^\dagger = U(CPC^\dagger)U^\dagger$ with $CPC^\dagger \in \mathcal{P}_n$; and $\mathcal{C}_2(n)$ is a group containing all global phases, so the images of products and phase multiples of the generators X_i, Z_i lie in $\mathcal{C}_2(n)$ once the images of the generators do. (ii) $MZ^aM^\dagger = P_s Z^a P_s^\dagger = \operatorname{diag}_y((-1)^{a \cdot s^{-1}(y)})$, since D' commutes with Z^a. This is a Pauli operator up to phase only if $y \mapsto a \cdot s^{-1}(y)$ is affine; as this holds for every a, s^{-1} and hence s is affine. If $s(x) = Lx + c$,

then $P_s X^b P_s^\dagger = X^{Lb}$ and $P_s Z^a P_s^\dagger$ is a Pauli operator. (iii) By (4), $D_\theta X_j D_\theta^\dagger$ is a Pauli operator up to phase iff $\partial_{e_j}\theta(y) - \partial_{e_j}\theta(0) \in \pi\mathbb{Z}$ for all y and $y \mapsto (\partial_{e_j}\theta(y) - \partial_{e_j}\theta(0))/\pi$ is linear modulo 2. In particular $\partial_{e_j}\theta(y) \equiv \kappa_j \pmod{\pi}$ with κ_j independent of y; since $\partial_{e_j}\theta(y+e_j) = -\partial_{e_j}\theta(y)$ we get $2\kappa_j \in \pi\mathbb{Z}$, so every single-bit phase difference, and hence every phase difference, lies in $\frac{\pi}{2}\mathbb{Z}$. For $\theta = \frac{\pi}{4}\lambda$ the condition reads: $\partial_{e_j}\lambda - \partial_{e_j}\lambda(0)$ takes values in $4\mathbb{Z}_8$ and is linear modulo 8 after division by 4, for every j. By (5) this is equivalent to the stated conditions on the c_S. □

The following classification of third-level diagonal gates is the qubit, level-three case of the characterization of Cui, Gottesman and Krishna [7]; we include the short proof because the argument recurs below.

Lemma 4. *A diagonal unitary D_θ lies in $\mathcal{C}_3(n)$ if and only if, after removing a global phase, $\theta = \frac{\pi}{4}\lambda$ with $\lambda \in \mathrm{Cliff}_n + H_n$, where $H_n := \langle x_i,\ 2x_ix_j,\ 4x_ix_jx_k\rangle$ is the group of phase functions of products of T, controlled-S and doubly controlled-Z gates. Equivalently, $c_S = 0$ for $|S| \geq 4$, $c_S \in 4\mathbb{Z}_8$ for $|S| = 3$ and $c_S \in 2\mathbb{Z}_8$ for $|S| = 2$.*

Proof. $D_\theta Z_i D_\theta^\dagger = Z_i$, and $D_\theta X_i D_\theta^\dagger = X_i \,\mathrm{diag}(e^{i\partial_{e_i}\theta})$ lies in $\mathcal{C}_2(n)$ iff $\mathrm{diag}(e^{i\partial_{e_i}\theta})$ does. By Lemma 3(iii) the values of $\partial_{e_i}\theta$ differ by elements of $\frac{\pi}{2}\mathbb{Z}$; comparing y and $y+e_i$ gives $\theta(y+e_i) - \theta(y) \in \frac{\pi}{4}\mathbb{Z}$, so $\theta - \theta(0) \in \frac{\pi}{4}\mathbb{Z}$. For $\lambda \in \mathbb{Z}_8^{\mathbb{F}_2^n}$ the condition is $\partial_{e_i}\lambda \in \mathrm{Cliff}_n$ for all i, and by (5) this is exactly the stated coefficient condition (the case $i \notin R$ bounds $c_{R\cup\{i\}}$, the case $i \in R$ is then automatic). □

For a unitary U let $S_U := \{(b,a) \in \mathbb{F}_2^{2n} :\ UX^bZ^aU^\dagger \in \mathcal{P}_n\}$, a subspace because images of products are products. Equip $\mathbb{F}_2^{2n}$ with the symplectic form $\langle (b,a),(b',a')\rangle = a\cdot b' + a'\cdot b$, let $S^\perp$ be the orthogonal complement and $\mathrm{rad}\, S = S \cap S^\perp$. A *Lagrangian* is an n-dimensional isotropic subspace.

Lemma 5. *$U \in \mathrm{SC}(n)$ if and only if S_U contains a Lagrangian, if and only if $\dim S_U + \dim \mathrm{rad}\, S_U = 2n$.*

Proof. If $U = C_L D C_R$ then S_U contains the Lagrangian $C_R^\dagger\langle Z_1,\dots,Z_n\rangle C_R$ (read modulo phases). Conversely let $L \subseteq S_U$ be a Lagrangian. Its image $ULU^\dagger$ consists of Pauli operators that commute pairwise and are independent, so it is again a Lagrangian. Choose Cliffords with $C_R L C_R^\dagger = \langle Z\rangle$ and $C_L\langle Z\rangle C_L^\dagger = ULU^\dagger$. Then $V = C_L^\dagger U C_R^\dagger$ maps each Z_i to $\pm Z^{a_i}$ with independent a_i; a circuit W of CNOT and X gates does the same, so $W^\dagger V$ commutes with every Z_i, is diagonal, and $U = (C_L W)(W^\dagger V)C_R$. Finally, a maximal isotropic subspace of S has dimension $(\dim S + \dim \mathrm{rad}\, S)/2$, and $\dim S + \dim \mathrm{rad}\, S \leq \dim S + \dim S^\perp = 2n$. □

3 Reduction to a permutation and a diagonal

Beigi and Shor proved: *"Every gate in $\mathcal{C}_3$ is generalized semi-Clifford"* [3, Theorem 1.1]. Zeng, Chen and Chuang proved: *"If R is a generalized semi-Clifford operation, then there exist Clifford operations L_1, L_2, and a classical permutation operator P such that PL_1RL_2 is diagonal"* [2, Proposition 2]. Both papers call U generalized semi-Clifford if conjugation by U maps the linear span of some maximal abelian subgroup of the Pauli group onto the span of another [3, Definition 1.3], [2, Definition 3]. Since $\mathcal{C}_k(n)$ is unchanged if $\mathcal{P}_n$ carries only the phases $\pm1, \pm i$, their hierarchy agrees with (1).

Proposition 6. *Every $U \in \mathcal{C}_3(n)$ can be written $U = C_L P_\sigma D C_R$ with $C_L, C_R \in \mathcal{C}_2(n)$, D diagonal, and σ a permutation of $\mathbb{F}_2^n$ such that $\sigma\tau_b\sigma^{-1}$ is affine for every $b \in \mathbb{F}_2^n$.*

Proof. The two quoted results give $U = L_1^\dagger P^{-1} D L_2^\dagger$; put $P_\sigma = P^{-1}$, $C_L = L_1^\dagger$, $C_R = L_2^\dagger$. By Lemma 3(i), $V := P_\sigma D = C_L^\dagger U C_R^\dagger \in \mathcal{C}_3(n)$, so $VX^bV^\dagger \in \mathcal{C}_2(n)$. By (4), $VX^bV^\dagger =$

$P_\sigma X^b \operatorname{diag}(e^{i\partial_b\theta})P_\sigma^\dagger = P_{\sigma\tau_b\sigma^{-1}}D'$ with D' diagonal, and Lemma 3(ii) shows that $\sigma\tau_b\sigma^{-1}$ is affine. □

Remark 7. He, Robitaille and Tan [5, Corollary 3.9] give a factorization in which the permutation itself lies in $\mathcal{C}_3(n)$ and is in their staircase form. Proposition 6 needs only the affine-conjugation property, which already follows from the refereed results of [3, 2]; the proof below uses nothing else.

4 A single Toffoli layer

For $q = (q_1, \dots, q_r)$ with each $q_k : \mathbb{F}_2^m \to \mathbb{F}_2$ a sum of products $w_i w_j$ $(i < j)$, let $\pi_q(w, u) = (w, u + q(w))$ on $\mathbb{F}_2^m \times \mathbb{F}_2^r$. This is a layer of Toffoli gates with controls w and targets u; it is an involution, and $P_{\pi_q} \in \mathcal{C}_3(m+r)$: indeed $\pi_q\tau_b\pi_q(w, u) = (w+b_w,\ u+b_u+q(w+b_w)+q(w))$ is affine because $q(w+b_w)+q(w)$ is affine in w, and $P_{\pi_q}Z^aP_{\pi_q} = \operatorname{diag}((-1)^g)$ with $g = a_w \cdot w + a_u \cdot (u+q(w))$ of degree at most two, which is a diagonal Clifford gate: $(-1)^g = \omega^{4g}$, and since $4(u \oplus v) \equiv 4u + 4v \pmod 8$ the $\mathbb{Z}_8$-expansion of $4g$ is 4 times the algebraic normal form of g, which lies in Cliff_{m+r} by Lemma 3(iii).

Lemma 8. *Let $n \le 6$ and let σ be a permutation of $\mathbb{F}_2^n$ such that $\sigma\tau_b\sigma^{-1}$ is affine for every b. Then $\sigma = \alpha \circ \pi_q \circ \beta$ with affine bijections α, β and some $m + r = n$ and q as above.*

Proof. (a) Replacing σ by $\tau_{\sigma(0)} \circ \sigma$ keeps every conjugate affine, so assume $\sigma(0) = 0$. Put $T_b = \sigma\tau_b\sigma^{-1}$. Then $b \mapsto T_b$ is an injective homomorphism, and $E = \{T_b\}$ is an elementary abelian group of affine maps acting regularly, because $T_b(\sigma(x)) = \sigma(x + b)$.

(b) For $y \in \mathbb{F}_2^n$ let g_y be the element of E with $g_y(0) = y$, and write $g_y(x) = M_y x + y$. Define $x \cdot y := (M_y + I)x = g_y(x) + x + y$. Applying $g_y g_x = g_x g_y$ to 0 gives $g_y(x) = g_x(y)$, so $x \cdot y = y \cdot x$; the product is linear in x, hence bilinear. Since $g_y g_x$ and $g_{g_y(x)}$ both lie in E and send 0 to $g_y(x)$, they are equal, so $M_y M_x = M_{x+y+x\cdot y}$. Writing $L_y z = z \cdot y$ this reads $(I + L_y)(I + L_x) = I + L_x + L_y + L_{x\cdot y}$, i.e. $(z \cdot x) \cdot y = z \cdot (x \cdot y)$. Finally $g_y^2 = 1$ gives $g_y(y) = 0$, i.e. $y \cdot y = 0$. Thus $N = (\mathbb{F}_2^n, +, \cdot)$ is a commutative associative algebra in which every square vanishes. (This is the correspondence of Caranti, Dalla Volta and Sala between abelian regular subgroups of the affine group and radical rings [6], re-derived here in the form needed.)

(c) *If $xyz \neq 0$ for some $x, y, z \in N$, then $n \ge 7$.* Every product with a repeated factor vanishes. If $ax + by + cz + d\,xy + e\,xz + f\,yz + g\,xyz = 0$, multiplying by yz, xz, xy gives $a = b = c = 0$; then multiplying by z, y, x gives $d = e = f = 0$; then $g = 0$. So x, y, z, xy, xz, yz, xyz are linearly independent. Hence $N^3 = 0$ for $n \le 6$.

(d) Since $T_c = g_{\sigma(c)}$, we have $\sigma(b + c) = \sigma(b) + \sigma(c) + \sigma(b) \cdot \sigma(c)$. With $u_i = \sigma(e_i)$ and $N^3 = 0$, induction on the support of b gives $\sigma(b) = \sum_i b_i u_i + \sum_{i<j} b_i b_j\, u_i u_j$. As σ is onto, $N = \operatorname{span}(u_i) + N^2$, hence $N^2 = \operatorname{span}(u_i u_j)$ (because $N \cdot N^2 = 0$), and $b \mapsto \sum_i b_i u_i$ maps $\mathbb{F}_2^n$ onto N/N^2. Let $r = \dim N^2$, $m = n - r$, and choose a basis $\beta_1, \dots, \beta_n$ of $\mathbb{F}_2^n$ with $\beta_{m+1}, \dots, \beta_n$ spanning the kernel of that map. Replacing σ by $\sigma \circ B$ with $B : e_i \mapsto \beta_i$ replaces T_b by T_{Bb}, so E and N are unchanged. Now $u_{m+1}, \dots, u_n \in N^2$, which annihilates N; they are linearly independent, since a relation $\sum_j c_j u_j = 0$ would give $\sigma(c) = 0 = \sigma(0)$; so they form a basis of N^2, and $u_1, \dots, u_m$ are independent modulo N^2. In the basis $u_1, \dots, u_n$ of $N = \mathbb{F}_2^n$ the formula for σ becomes $(w, u) \mapsto (w, u + q(w))$, where $q(w)$ collects the N^2-coordinates of $\sum_{i<j\le m} w_i w_j u_i u_j$. Undoing the normalizations gives $\alpha = \tau_{\sigma(0)} \circ C^{-1}$ and $\beta = B^{-1}$, with Γ the coordinate map. □

Remark 9. Part (c) is sharp: the free commutative algebra with vanishing squares on three generators has dimension seven, and it produces the seven-qubit gate of Proposition 17.

5 Classification of the layers

By Proposition 6, Lemma 8 and Lemma 3(i)–(ii), every $U \in \mathcal{C}_3(n)$ with $n \le 6$ equals $C'_L P_{\pi_q} D'' C'_R$ with $C'_L = C_L P_\alpha$, $C'_R = P_\beta C_R$ Clifford and $D'' = P_\beta D P_\beta^{-1}$ diagonal. We now normalise q.

For $G_C \in \mathrm{GL}(m,2)$ and $G_T \in \mathrm{GL}(r,2)$ one checks $(G_C^{-1} \oplus G_T)\,\pi_q\,(G_C \oplus G_T^{-1}) = \pi_{G_T q \circ G_C}$, and $\pi_{q+l} = \pi_q \circ (w, u \mapsto (w, u + l(w)))$ for affine l. All maps involved are affine, hence Clifford permutations, and conjugating a diagonal gate by a permutation gate gives a diagonal gate. The polar forms $B_k(x,y) = q_k(x+y) + q_k(x) + q_k(y)$ are alternating, and $q_k \mapsto B_k$ is a bijection from pure quadratic forms onto $\mathrm{Alt}(m)$, the space of alternating bilinear forms, under which G_C acts by congruence. Hence only the $\mathrm{GL}(m,2)$-orbit of $W = \mathrm{span}(B_1, \ldots, B_r)$ matters. If the B_k are dependent, a change by G_T makes some q_k affine, and that target becomes idle. If they have a common radical vector, a change by G_C moves it to e_m, and w_m becomes idle. We may therefore assume that W has dimension $t = |\mathsf{T}|$ and trivial common radical, with $m + t \le n \le 6$ and the remaining qubits idle.

Proposition 10. *Up to* $\mathrm{GL}(m,2)$*, the* t*-dimensional subspaces* $W \subseteq \mathrm{Alt}(m)$ *with trivial common radical and* $m + t \le 6$ *form exactly seven orbits, represented (controls* $0, \ldots, m-1$*, the* t *forms listed) by*

class	(m,t)	*representative* f	*orbit size*
TOF	$(2,1)$	x_0x_1	1
B	$(4,1)$	$x_0x_1 + x_2x_3$	28
C	$(3,2)$	$(x_0x_1,\ x_0x_2)$	7
TRI	$(3,3)$	$(x_0x_1,\ x_0x_2,\ x_1x_2)$	1
SEC	$(4,2)$	$(x_0x_1,\ x_2x_3)$	280
TAN	$(4,2)$	$(x_0x_1,\ x_0x_2 + x_1x_3)$	210
EXT	$(4,2)$	$(x_0x_1 + x_2x_3,\ x_0x_1 + x_0x_2 + x_1x_3)$	56

With idle qubits these give 7 *classes at* $n = 6$ *(TOF with three idle qubits, B and C with one, SEC, TAN, EXT, TRI),* 3 *at* $n = 5$ *(TOF with two idle qubits, B, C), TOF with at most one idle qubit at* $n \in \{3,4\}$*, and none at* $n \le 2$*.*

Proof. $\mathrm{Alt}(2)$ is one-dimensional. Every nonzero form in $\mathrm{Alt}(3)$ has rank two, so $t = 1$ is degenerate for $m = 3$; two independent forms in $\mathrm{Alt}(3)$ have distinct one-dimensional radicals, and $\mathrm{GL}(3,2)$ acts transitively on the seven planes of $\mathrm{Alt}(3) \cong \mathbb{F}_2^3$; $t = 3$ is all of $\mathrm{Alt}(3)$. For $m = 4$, $t = 1$, the form must have rank four, and there are $|\mathrm{GL}(4,2)|/|\mathrm{Sp}(4,2)| = 20160/720 = 28$ of them, forming one orbit. For $m = 4$, $t = 2$, identify the projective space of $\mathrm{Alt}(4)$ with $\mathrm{PG}(5,2)$; the rank-two forms are the 35 points of the Klein quadric. The tangent hyperplane at a singular point P meets the quadric in a cone with vertex P over a hyperbolic quadric of $\mathrm{PG}(3,2)$ with 9 points, so P lies on 9 totally singular lines and is collinear on the quadric with 18 other singular points. Of the 31 lines through P, therefore, 9 are totally singular, 16 join P to the non-collinear singular points (secant lines, with exactly two singular points) and 6 are tangent (one singular point). This gives $35 \cdot 9/3 = 105$ totally singular, $35 \cdot 16/2 = 280$ secant, $35 \cdot 6 = 210$ tangent and $651 - 595 = 56$ external lines. A line through a rank-four form is nondegenerate. A totally singular line consists of forms $e \wedge a$, $e \wedge b$, $e \wedge (a+b)$ with a common linear form e, and has a nonzero common radical. That each of the three nondegenerate line types is a single $\mathrm{GL}(4,2)$-orbit is verified by the orbit census of Section 9, which also confirms all orbit sizes. For $m = 5$, $t = 1$, alternating forms on $\mathbb{F}_2^5$ are degenerate, and every other (m,t) with $t \le \dim \mathrm{Alt}(m)$ has $m + t \ge 7$ (note $\mathrm{Alt}(1) = 0$ and $\dim \mathrm{Alt}(2) = 1$). Direct inspection places the representatives in the stated types. □

6 Admissible diagonal phases

Fix a representative $\pi = \pi_f$ from Proposition 10, padded with idle qubits to $n \in \{5, 6\}$, and let

$$A_\pi := \{\lambda \in \mathbb{Z}_8^{\mathbb{F}_2^n} : \ P_\pi D_\lambda \in \mathcal{C}_3(n)\}.$$

Lemma 11. *If $P_\pi D_\theta \in \mathcal{C}_3(n)$ for real θ, then $\theta - \theta(0) \in \frac{\pi}{4}\mathbb{Z}$. For $\lambda \in \mathbb{Z}_8^{\mathbb{F}_2^n}$, $\lambda \in A_\pi$ if and only if $\psi_i := (\partial_{e_i}\lambda) \circ \pi \in \mathrm{Cliff}_n$ for $i = 1, \ldots, n$. In particular A_π is a subgroup.*

Proof. Let $V = P_\pi D_\theta$. Then $VZ^aV^\dagger = P_\pi Z^a P_\pi \in \mathcal{C}_2(n)$ and, by (4) and $\pi^{-1} = \pi$, $VX^bV^\dagger = (P_\pi X^b P_\pi)\,\mathrm{diag}(e^{i(\partial_b\theta)\circ\pi})$ with $P_\pi X^b P_\pi \in \mathcal{C}_2(n)$. So $V \in \mathcal{C}_3(n)$ iff each $\mathrm{diag}(e^{i(\partial_{e_i}\theta)\circ\pi})$ is a Clifford gate (Lemma 3(i)). By Lemma 3(iii) the values of $\partial_b\theta$ differ by elements of $\frac{\pi}{2}\mathbb{Z}$; since $\partial_b\theta(0) = \theta(b) - \theta(0) = -\partial_b\theta(b)$, we get $\theta(b) - \theta(0) \in \frac{\pi}{4}\mathbb{Z}$. The second statement is the same criterion for $\theta = \frac{\pi}{4}\lambda$, and it is $\mathbb{Z}_8$-linear in λ. □

Let $H_\pi := \langle x_i,\ 2x_ix_j,\ 4x_ix_jx_k : \ i,j,k \notin \mathsf{T}\rangle$, the phase functions of T, controlled-S and doubly controlled-Z gates on the non-target qubits, and $\mathrm{Cliff}_n \circ \pi := \{\varphi \circ \pi : \varphi \in \mathrm{Cliff}_n\}$.

Theorem 12. *For each of the ten representatives ($n = 5$: TOF, B, C; $n = 6$: TOF, B, C, SEC, TAN, EXT, TRI),*

$$A_\pi = \mathrm{Cliff}_n + \mathrm{Cliff}_n \circ \pi + H_\pi,$$

and $\log_2 |A_\pi|$ equals $38, 38, 33$ and $56, 56, 47, 47, 47, 47, 43$ respectively.

Proof. Inclusion $\supseteq$. If $\lambda \in \mathrm{Cliff}_n$ then $P_\pi D_\lambda \in \mathcal{C}_3(n)$ by Lemma 3(i). If $\varphi \in \mathrm{Cliff}_n$ then $P_\pi D_{\varphi\circ\pi} = D_\varphi P_\pi \in \mathcal{C}_3(n)$. If $h \in H_\pi$, then D_h depends only on non-target bits, which π does not change, so D_h commutes with P_π and with every X_τ, $\tau \in \mathsf{T}$. Hence $V = P_\pi D_h$ fixes X_τ; for a non-target i, $VX_iV^\dagger = (P_\pi X_i P_\pi)\,\mathrm{diag}(\omega^{\partial_{e_i}h})$, and $\partial_{e_i}h \in \mathrm{Cliff}_n$ by (5) because the coefficients of h satisfy the conditions of Lemma 4. Since A_π is a group, the sum of the three subgroups lies in A_π.

Inclusion $\subseteq$ is computer-assisted. For each representative, the program of Section 9 computes A_π twice, from Lemma 11 and independently from the literal definition (1), certifies its order by the kernel certificate described there, and computes the order of the subgroup generated by $\mathrm{Cliff}_n \cup \mathrm{Cliff}_n \circ \pi \cup H_\pi$. The two orders agree and equal the stated values, so the inclusion is an equality. □

Remark 13. The orders can be read off by hand. For example, for TOF with k idle qubits, $\log_2 |\mathrm{Cliff}_n| = 3 + 2n + \binom{n}{2}$, H_π adds one bit for each of the $\binom{n-1}{1} + \binom{n-1}{2} + \binom{n-1}{3}$ monomials on the non-target qubits, and $\mathrm{Cliff}_n \circ \pi$ adds the single element $4x_0x_1x_2$, which comes from $2x_2 \circ \pi$; this gives $30 + 25 + 1 = 56$ at $n = 6$ and $23 + 14 + 1 = 38$ at $n = 5$. What the computation certifies is that no other admissible phase exists.

Corollary 14. *For $n \le 4$ and the layer TOF with at most one idle qubit, $A_\pi = \mathrm{Cliff}_n + \mathrm{Cliff}_n \circ \pi + H_\pi$ as well.*

Proof. Let ρ be the six-qubit representative of the same class, i.e. π padded with idle qubits. If $\lambda \in A_\pi$, then $\lambda'(x, z) := \lambda(x)$ lies in A_ρ, because $(P_\pi D_\lambda) \otimes I$ lies in $\mathcal{C}_3(6)$. By Theorem 12, $\lambda' = \varphi_1 + \varphi_2 \circ \rho + h$. Setting the added bits to 0 gives $\lambda = \varphi_1(\cdot, 0) + \varphi_2(\pi(\cdot), 0) + h(\cdot, 0)$, and restriction preserves the three coefficient patterns. The same argument reproves the case $n = 5$ from $n = 6$. □

7 Proof of the main theorem

Let $n \le 6$ and $U \in \mathcal{C}_3(n)$. By Sections 3–5, $U = C'_L P_\pi D_\theta C'_R$ with Clifford gates C'_L, C'_R, $\pi = \pi_f$ a representative of Proposition 10 (possibly the identity), and $P_\pi D_\theta \in \mathcal{C}_3(n)$.

If π is the identity, then $D_\theta \in \mathcal{C}_3(n)$, and Lemma 4 gives $D_\theta = e^{i\gamma} D_\varphi D_h$ with $\varphi \in \mathrm{Cliff}_n$ and $h \in H_n$; this is (3) with $\mathsf{T} = \varnothing$.

Otherwise, Lemma 11 gives $\theta = \gamma + \frac{\pi}{4}\lambda$ with $\lambda \in A_\pi$, and by Theorem 12 and Corollary 14 $\lambda = \varphi_1 + \varphi_2 \circ \pi + h$ with $\varphi_1, \varphi_2 \in \mathrm{Cliff}_n$, $h \in H_\pi$. Then $P_\pi D_\lambda = P_\pi D_{\varphi_2 \circ \pi} D_h D_{\varphi_1} = D_{\varphi_2} P_\pi D_h D_{\varphi_1}$. Because $P_{\pi_f} = \sum_{x_\mathsf{C}} |x_\mathsf{C}\rangle\langle x_\mathsf{C}| \otimes \prod_\tau X_\tau^{f_\tau(x_\mathsf{C})} \otimes I$, we have $H_\mathsf{T} P_{\pi_f} H_\mathsf{T} = \mathrm{diag}((-1)^{\sum_\tau x_\tau f_\tau(x_\mathsf{C})})$, and D_h acts on non-target qubits only, so it commutes with H_T. Hence $P_{\pi_f} D_h = H_\mathsf{T} E H_\mathsf{T}$ with E as in (3), and

$$U = (C'_L e^{i\gamma} D_{\varphi_2})\, P_{\pi_f} D_h\, (D_{\varphi_1} C'_R) = (C'_L e^{i\gamma} D_{\varphi_2} H_\mathsf{T})\, E\, (H_\mathsf{T} D_{\varphi_1} C'_R) \in \mathrm{SC}(n).$$

Equivalently, $P_{\pi_f} D_h$ maps each element of the Lagrangian $\langle Z_c \ (c \notin \mathsf{T}),\ X_\tau \ (\tau \in \mathsf{T})\rangle$ to itself under conjugation. This proves Theorem 1. □

Remark 15. The program also confirms semi-Cliffordness by a route that does not use Theorem 12. For $\lambda \in A_\pi$ and $\tau \in \mathsf{T}$ there is at most one $\alpha \in \mathbb{F}_2^\mathsf{T}$ (two would make $4(\alpha + \alpha') \cdot f$ affine, which is impossible because the f_τ have linearly independent polar forms) for which $V X_\tau Z^\alpha V^\dagger = X_\tau Z^\alpha \,\mathrm{diag}(\omega^{\psi_\tau + 4\alpha \cdot f})$ is a Pauli operator. The set of λ for which such an $\alpha = M_\tau$ exists for every τ, with M symmetric, is a subgroup, because ψ_τ is $\mathbb{Z}_8$-linear in λ and $4(\alpha+\alpha') \cdot f \equiv 4\alpha \cdot f + 4\alpha' \cdot f \pmod 8$; for such λ the space $\langle Z_c,\ X_\tau Z^{M_\tau}\rangle$ is a Lagrangian in S_V. The program checks that every generator of A_π lies in this subgroup.

8 Seven qubits

Lemma 16. *A permutation gate P_y lies in $\mathcal{C}_3(n)$ if and only if $y\tau_b y^{-1}$ is affine for every b and every coordinate of y^{-1} is a polynomial of degree at most two over $\mathbb{F}_2$.*

Proof. $P_y X^b P_y^\dagger = P_{y\tau_b y^{-1}}$ is Clifford iff $y\tau_b y^{-1}$ is affine (Lemma 3(ii)), and $P_y Z^a P_y^\dagger = \mathrm{diag}((-1)^{a \cdot y^{-1}(z)}) = D_{4g}$ with $g = a \cdot y^{-1}$. Since $4(u \oplus v) \equiv 4u + 4v \pmod 8$, the $\mathbb{Z}_8$-expansion of $4g$ is 4 times the algebraic normal form of g, which lies in Cliff_n iff g has degree at most two. □

Proposition 17. *The permutation of $\mathbb{F}_2^7$*

$$y(b) = (b_1,\ b_2,\ b_3,\ b_4 + b_1b_2,\ b_5 + b_1b_3,\ b_6 + b_2b_3,\ b_7 + b_1b_2b_3 + b_1b_6 + b_2b_5 + b_3b_4)$$

gives a gate $P_y \in \mathcal{C}_3(7) \setminus \mathrm{SC}(7)$.

Proof. The inverse is $b_7 = c_7 + c_1c_6 + c_2c_5 + c_3c_4$ with $b_4 = c_4 + c_1c_2$, $b_5 = c_5 + c_1c_3$, $b_6 = c_6 + c_2c_3$, so it has degree two. Writing $x = y^{-1}(c)$ one finds $y\tau_{e_i}y^{-1}(c) = c + e_i + \ell_i(c)$ with $\ell_1 = c_2e_4 + c_3e_5 + c_6e_7$, $\ell_2 = c_1e_4 + c_3e_6 + c_5e_7$, $\ell_3 = c_1e_5 + c_2e_6 + c_4e_7$, $\ell_4 = c_3e_7$, $\ell_5 = c_2e_7$, $\ell_6 = c_1e_7$ and $\ell_7 = 0$, all linear; so every $y\tau_b y^{-1}$, a composite of these, is affine, and $P_y \in \mathcal{C}_3(7)$ by Lemma 16. Moreover $P_y X^b Z^a P_y^\dagger = P_{y\tau_b y^{-1}} \,\mathrm{diag}((-1)^{a \cdot y^{-1}})$ is a Pauli operator if and only if $y\tau_b y^{-1}$ is a translation and $a \cdot y^{-1}$ is affine. Coordinates 4–7 of $y(x+b) + y(x)$ are constant in x only if $b \in \langle e_7\rangle$, and the quadratic monomials $c_1c_2, c_1c_3, c_2c_3, c_1c_6, c_2c_5, c_3c_4$ of y^{-1} are distinct, so $a \cdot y^{-1}$ is affine only if $a \in \langle e_1, e_2, e_3\rangle$. Hence $S_{P_y} = \langle X_7, Z_1, Z_2, Z_3\rangle$, which is isotropic, so $\dim S_{P_y} + \dim \mathrm{rad}\, S_{P_y} = 8 < 14$ and Lemma 5 shows $P_y \notin \mathrm{SC}(7)$. The program confirms both facts, the first directly from the definition (1). The underlying algebra is the free commutative algebra with vanishing squares on three generators, in which $b_1b_2b_3 \neq 0$, which is exactly the situation excluded for $n \le 6$ by Lemma 8(c). □

9 Verification

The file `check_c3_n5_n6.py` is standalone. It uses only the Python standard library, exact integer arithmetic and no network, and it takes no input. Its output is fixed in `expected_stdout.txt`. It checks:

1. the class list, by comparing the hand classification (entered as data) with a computed census of the $\mathrm{GL}(m,2)$-orbits of t-dimensional subspaces of $\mathrm{Alt}(m)$ for $m+t \le 6$; the census does not use the representatives or the expected sizes, and it also checks that the representatives meet the nondegenerate orbits bijectively at $n=5$ and $n=6$;

2. for each of the ten representatives, A_π derived by two independent routes (Lemma 11, and the literal definition (1) with phases carried as $\mathbb{Z}_8$-linear forms through every generator conjugation), each with a kernel certificate, mutual containment and equal orders; the literal route's Pauli test is first calibrated against an independently built list of Pauli operators; then the orders in Theorem 12 and their equality with the order of the subgroup generated by Cliff_n, $\mathrm{Cliff}_n \circ \pi$ and H_π;

3. that $P_\pi D_h$ fixes the Lagrangian of Section 7 for every generator h of H_π, and the Lagrangian check of the Remark in Section 7;

4. Lemma 4 for $n = 1, \dots, 6$, as an equality of subgroups;

5. three controls: a Toffoli gate times T on an idle qubit, which is correctly accepted with $\dim S_U = 8$; a Toffoli gate times T on its target, which is correctly rejected; and the seven-qubit gate of Proposition 17.

The kernel certificate works as follows. Let R be a $\mathbb{Z}_8$-matrix and V a matrix invertible modulo 2 such that $RV_j = 2^{v_j} w_j$ for $j < r$, with the w_j independent modulo 2, and $RV_j = 0$ for $j \ge r$. Then $\ker R = V\{y : 2^{v_j} y_j = 0\}$. Indeed, if $\sum_j 2^{v_j} y_j w_j = 0$ with some term nonzero, dividing by the least power of two among the nonzero terms and reducing modulo 2 contradicts independence. Hence $\log_2 |\ker R| = \sum_j v_j + 3(\text{number of columns} - r)$. The certificate is checked independently of the elimination that produced it.

The program prints 54 lines in about seven seconds on a laptop, and its output is byte-identical across isolated runs. Twenty single-point mutations of the program each make it exit with a nonzero status. They include one for every line that the internal reviews found no earlier mutation reached. The harness `mutate.py` and its report are included; the harness also checks statically that the two derivations of A_π share no function other than the $\mathbb{Z}_8$ linear algebra. Reproduction confirms only that the program computes what is described. The steps that the program relies on but does not check (Lemmas 3, 4, 5, 8 and 11, Proposition 6, the normalisation at the start of Section 5, Proposition 10 apart from the transitivity of $\mathrm{GL}(4,2)$ on each line type, and Corollary 14) are established by the written argument. The program's output uses a few internal labels: "Lemma A" is Lemma 5, the "Lemma-C certificate" is the kernel certificate above, "Step 3" is Lemma 11, "Theorem 1" is Theorem 12, "Step 6" is the Remark in Section 7, and the "M9 control" is the first control of item 5.

10 Discussion

The argument separates cleanly into a structural part and a finite computation. The structural part explains why seven is the threshold. The permutation part of a third-level gate carries a commutative algebra with vanishing squares, and such an algebra satisfies $N^3 = 0$ below dimension seven. That condition means a single Toffoli layer, whose admissible diagonal phases are only the obvious ones. He, Robitaille and Tan characterized third-level permutation gates through their

staircase form and showed that those on at most six qubits are semi-Clifford [5, Appendix A]. The present result extends semi-Cliffordness from permutations to all third-level gates on at most six qubits by controlling the diagonal factor. The only non-textbook published inputs are the theorem of Beigi and Shor [3] and the factorization of Zeng, Chen and Chuang [2]. The seven-qubit example is checked by hand and confirmed by the program; the inclusion $\subseteq$ of Theorem 12, on which the lower bound rests, is computer-assisted.

Two natural questions remain. One is whether the computation in Theorem 12 can be replaced by a uniform argument that also describes the admissible phases for multi-layer permutations when $n \geq 7$. The other is whether the same algebraic approach separates the generalized semi-Clifford property from the semi-Clifford property at higher levels, where de Silva and Lautsch report a five-qubit gate in $\mathcal{C}_5 \setminus \mathcal{C}_4$ that is not generalized semi-Clifford [8].

AI assistance and review status

AI tools were used to develop the argument, to write the verification program, to assist with exposition and to conduct internal adversarial reviews. The manuscript follows the evidence-tracking guidance of Scientific Agent Skills [10]. These activities are disclosed as AI assistance, not as external refereeing or formal proof verification.

Three adversarial reviews were performed by AI in isolated contexts. Each recomputed with its own code the admissible-phase groups it covered (the first covered $n = 5$ only); the second and third re-derived the classification; the last worked directly from definition (1), shared none of the program's routes, and re-checked the proof of Lemma 8 line by line. A further audit of this manuscript and its program recomputed every number in it independently. The last review also established that the argument needs only the refereed results of [3, 2], which is how Section 3 is written. **No independent human expert has examined this argument, and no novelty certification by a human is recorded.** Searches of arXiv through 26 September 2026 found no prior resolution of the five- and six-qubit cases. The problem record [9] was still listed as unsolved when this manuscript was prepared.

References

[1] D. Gottesman and I. L. Chuang, "Quantum teleportation is a universal computational primitive," arXiv:quant-ph/9908010; Nature **402**, 390–393 (1999).

[2] B. Zeng, X. Chen and I. L. Chuang, "Semi-Clifford operations, structure of $\mathcal{C}_k$ hierarchy, and gate complexity for fault-tolerant quantum computation," Phys. Rev. A **77**, 042313 (2008), arXiv:0712.2084. Proposition 2 and the abstract were consulted on 26 September 2026.

[3] S. Beigi and P. W. Shor, "$\mathcal{C}_3$, semi-Clifford and generalized semi-Clifford operations," Quantum Inf. Comput. **10**, 41–59 (2010), arXiv:0810.5108. Theorem 1.1 and Section 1.1 were consulted on 26 September 2026.

[4] J. T. Anderson and A. Connelly, "Affine equivalence in the Clifford hierarchy," arXiv:2507.14370v2 (2025). Preprint.

[5] Z. He, L. Robitaille and X. Tan, "Characterization of permutation gates in the third level of the Clifford hierarchy," arXiv:2510.04993v1 (2025). Corollary 3.9, Section 2.1 and Appendix A were consulted on 26 September 2026.

[6] A. Caranti, F. Dalla Volta and M. Sala, "Abelian regular subgroups of the affine group and radical rings," Publ. Math. Debrecen **69**, 297–308 (2006), arXiv:math/0510166.

[7] S. X. Cui, D. Gottesman and A. Krishna, "Diagonal gates in the Clifford hierarchy," Phys. Rev. A **95**, 012329 (2017), arXiv:1608.06596.

[8] N. de Silva and O. Lautsch, "The generalised semi-Clifford conjecture is false," arXiv:2609.11903 (2026), preliminary version; cited as summarised in [9].

[9] Quantum Information and Quantum Computation Open Problem Zoo, "Smallest qubit number for a non-semi-Clifford third-level gate," https://qiqc-op.com/problem/op_5a17b19b8adeb191/. Status re-checked 26 September 2026.

[10] T. Kassis, V. Agarwal, Y. He, D. Patel and A. M. Brueckner, "Scientific Agent Skills: A Library of Procedural Knowledge for Research Agents," arXiv:2609.00065v2 (2026). Preprint; Software repository. The scientific-writing guidance was consulted at commit `330c8e764435`.