arXiv:2302.07966v2 [quant-ph] 29 Mar 2024

The qudit Pauli group: non-commuting pairs, non-commuting sets, and structure theorems

Rahul Sarkar[1] and Theodore J. Yoder[2]

[1]Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA 94305
[2]IBM T.J. Watson Research Center, Yorktown Heights, NY

April 1, 2024

Qudits with local dimension $d > 2$ can have unique structure and uses that qubits ($d = 2$) cannot. Qudit Pauli operators provide a very useful basis of the space of qudit states and operators. We study the structure of the qudit Pauli group for any, including composite, d in several ways. To cover composite values of d, we work with modules over commutative rings, which generalize the notion of vector spaces over fields. For any specified set of commutation relations, we construct a set of qudit Paulis satisfying those relations. We also study the maximum size of sets of Paulis that mutually non-commute and sets that non-commute in pairs. Finally, we give methods to find near minimal generating sets of Pauli subgroups, calculate the sizes of Pauli subgroups, and find bases of logical operators for qudit stabilizer codes. Useful tools in this study are normal forms from linear algebra over commutative rings, including the Smith normal form, alternating Smith normal form, and Howell normal form of matrices. Possible applications of this work include the construction and analysis of qudit stabilizer codes, entanglement assisted codes, parafermion codes, and fermionic Hamiltonian simulation.

1 Introduction

Typical models for quantum computation assume that quantum information is encoded into qubits, which are $d = 2$ level systems analogous to classical bits. However, systems with $d > 2$ levels, called qudits, also have theoretical advantages, including information storage [1, 2, 3], computation [4, 5, 6], the fault-tolerant creation of magic states [7, 8], and simpler realizations of topological phases [9, 10]. Experimental implementations of qudits exist. For instance qutrits ($d = 3$) with entangling operations can be implemented in superconducting [11, 12] and photonic [13] systems. Larger, composite values of d exist as well across superconducting [14, 15], photonic [16], and molecular systems [17, 18]. We assume throughout that $d = p_0^{\alpha_0} p_1^{\alpha_1} \dots p_{m-1}^{\alpha_{m-1}}$ is the prime factorization of a positive integer $d \geq 2$ into unique primes p_i.

Central to the understanding of qubits is the Pauli group on n qubits, a subgroup of $\mathrm{U}(2^n)$ generated by products and tensor products of $X = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and $Z = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. The Pauli group has myriad uses in quantum theory, including facilitating the description of stabilizer states, error-correcting codes, and quantum errors [19]. The qudit Pauli group, or Heisenberg-Weyl group, plays a similar role in understanding qudits [20]. Here we take

$$X = \sum_{j=0}^{d-1} |j+1\rangle\langle j|, \quad Z = \sum_{j=0}^{d-1} \omega^j |j\rangle\langle j|, \tag{1}$$

Rahul Sarkar: rsarkar@stanford.edu, Rahul Sarkar was funded by the Stanford Exploration Project for the duration of this study.

Theodore J. Yoder: ted.yoder@ibm.com,

where $\omega = e^{2\pi i/d}$, as generators of the n-qudit Pauli group, a subgroup of $\mathrm{U}(d^n)$. All matrices $X^a Z^b$ for $a, b \in \mathbb{Z}_d$ are distinct. If we take the group commutator $[A, B] := ABA^{-1}B^{-1}$ of two single-qudit Paulis, we find $[X^a Z^b, X^{a'} Z^{b'}] = \omega^{a'b - b'a} I$, for identity matrix I. Thus, while qubit Paulis P, Q only either commute $[P, Q] = I$ or anticommute $[P, Q] = -I$, qudit Paulis can fail to commute in any one of $d-1$ distinct ways, $[P, Q] = \omega^c I$ for any $c = 1, \ldots, d-1$. This is one reason there is interesting structure in the qudit Pauli group that is absent for qubits.

In this paper we communicate several structural theorems about sets and groups of qudit Pauli operators. We think it is illustrative to present our results in comparison with the corresponding ideas for qubits, which might be familiar to some readers.

Non-commuting pairs. We say Paulis $s_0, \ldots, s_{k-1}$ and $t_0, \ldots, t_{k-1}$ are non-commuting pairs if all pairs (s_i, t_i) do not commute with each other, but do commute with all other s_j and t_j for $j \neq i$. For qubits (and in fact any prime d), the maximum number of non-commuting pairs on n qudits is simply $k = n$, for instance (X, Z) on one qubit. However, for composite d, we can have more. For example, $d = 6$ permits (X^2, Z^2) and (X^3, Z^3) as a collection of $k = 2$ non-commuting pairs on one qudit. In Section 3, we will show that the maximum number of non-commuting pairs is $k = mn$, where m is the number of unique primes in the factorization of d. We use this result to answer some further questions in Sections 5 and 6. The key tool used to obtain this result is Lemma 3.1, which also finds use in Section 5.

Non-commuting sets. A set of Paulis $\{q_0, q_1, \ldots, q_{h-1}\}$ is non-commuting if $[q_i, q_j] \neq I$ for every $i \neq j$. On a single qubit, an example largest non-commuting set is $\{Z, X, XZ\}$ and, more generally for prime d, a largest set has size $h = d + 1$. In contrast, on a single $d = 6$ qudit, the largest non-commuting sets have size $h = 12$, e.g.

$$\{Z, X, XZ, XZ^2, XZ^3, XZ^4, XZ^5, X^2Z, X^2Z^3, X^2Z^5, X^3Z, X^3Z^2\}. \tag{2}$$

We will show in Section 4 that $h = \Psi(d) \geq d+1$ is the maximum size non-commuting set on a single qudit, where Ψ is the Dedekind Psi function from number theory. The situation of non-commuting sets on more than one qudit is complicated, and we discuss some bounds on the size in that setting. We also note here that if one additionally requires the Paulis in the non-commuting set to all have the same group commutator value, the size of the maximum set is $2n + 1$, independent of d [21].

Qudits needed to create commutation patterns. Suppose $C \in \mathbb{Z}^{l \times l}$ has zero diagonal, $C_{ii} = 0$ for all i, and is anti-symmetric, $C = -C^T$. These two properties define what is known as an *alternating* matrix. What is the minimum number of qudits n such that there exists a list of n-qudit Paulis $\{q_0, q_1, \ldots, q_{l-1}\}$ with $[q_i, q_j] = \omega^{C_{ij}} I$? In Appendix B of [22] and in [23], it was shown that the number of qubits required is half the rank of C as a matrix over the field $\mathbb{F}_2$. For instance, if

$$C = \begin{pmatrix} 0 & 3 & 0 & 0 \\ -3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & -5 & 0 \end{pmatrix}, \tag{3}$$

then two qubits are necessary and sufficient via the set $\{X \otimes I, Z \otimes I, I \otimes X, I \otimes Z\}$. However, just one $d = 15$ qudit suffices: $\{X^3, Z^9, X^5, Z^5\}$. In Theorem 5.4 of Section 5, we show that in general, the number of qudits is half the minimal number of vectors generating the column-space of C as a matrix over $\mathbb{Z}_d$, which can differ from other definitions of matrix rank when d is composite. We also give an efficient algorithm to find a satisfying set of Paulis using a matrix decomposition called the alternating Smith normal form [24]. This approach effectively reduces the problem to the construction of collections of non-commuting pairs, solved in Section 3.

Minimal and Gram-Schmidt generating sets. Suppose $A^{\otimes 3}$ is shorthand for $A \otimes A \otimes A$, $S = \{X^{\otimes 3}, Z^{\otimes 3}, \omega(XZ)^{\otimes 3}\}$ is a set of qudit Paulis, and $G = \langle S \rangle$ is the group generated by S. What is the smallest generating set of G? We answer this question in general in Section 6.1 using the Smith normal form of the matrix over $\mathbb{Z}_d$ representing S. For $d = 2$, $\omega = -1$ and a smallest generating set for this example is $\{X^{\otimes 3}, Z^{\otimes 3}\}$. However, for $d = 6$, the last generator contributes a phase unobtainable from the first two generators alone and so any minimal generating set is larger, e.g. $\{X^{\otimes 3}, Z^{\otimes 3}, \omega I^{\otimes 3}\}$. A Gram-Schmidt generating set for G is one that can be written

as some number of non-commuting pairs (s_i, t_i) and additional elements u_i that commute with all elements of G. If G is the centralizer of a qudit stabilizer group, then the non-commuting pairs in Gram-Schmidt generating set provide a basis for the logical operators of the stabilizer code, as is done for qubits in [25]. For instance, $\{X^{\otimes 3}, Z^{\otimes 3}, \omega I^{\otimes 3}\}$ is also a Gram-Schmidt generating set for the example G with $d = 6$, where the first two elements are a non-commuting pair. Our construction of Gram-Schmidt generating sets in Section 6.2 guarantees the minimal number of non-commuting pairs, and again makes use of the alternating Smith normal form.

Subgroups generated by maximum non-commuting pairs. For qubits (and other qudits with prime d), maximum collections of non-commuting pairs generate the full Pauli group. This is clearly the case, for instance, for the maximum collection (X, Z) on a single prime-d qudit. However, for composite d this is no longer necessarily the case. For instance, for $d = 12$, (X^3, Z^6) and (X^4, Z^4) generate only $\langle \omega I, X, Z^2 \rangle$, a proper subgroup of the full single-qudit Pauli group, despite being a maximum size collection of non-commuting pairs. We show in Section 6.3 that this can only happen when d is not square-free. We also go further and find the size of groups generated by arbitrary generating sets as well.

1.1 Related work

In the remainder of this section, we discuss prior works and similar results in the existing literature, and mention applications and contributions of our current results.

One of the important uses of qubit based quantum computers is the simulation of Hamiltonians of quantum systems, such as atoms or molecules. The way this is achieved is by mapping the algebra of fermions or bosons, depending on whether one wants to simulate a fermionic or bosonic system, to the algebra of qubits. Popular encoding techniques to achieve such a transformation are the Jordan-Wigner transform [26], the Bravyi-Kitaev encoding [27], the Verstraete-Cirac mapping [28], Fenwick trees [29], and ternary trees [30], but other options also exist in the literature [31, 32, 33, 34, 35]. One desirable property during this mapping process is to use as few qubits as possible; in fact this has been the focus of the methods described in [31, 32, 33, 35]. Equivalently, this problem is the same as finding a set of Pauli operators using the minimum number of qubits, where the Paulis must satisfy a certain set of commutation relations — for example in the case of fermionic simulation, the Paulis that the fermionic modes are mapped to must anticommute with one another. This exact question for the qubit case has been answered in [23], and the result there generalizes to all qudits of prime dimension. Here, we cover the case of qudits of composite dimension.

The question of finding Paulis using a minimum number of qubits that satisfy a prescribed set of commutation relations also comes up in the context of entanglement assisted quantum error-correcting codes (EAQECC) [36, 37, 38], and quantum convolutional codes [39]. A typical EAQECC is described as $[[n, k, d; c]]$, where n qubits are used to encode k logical qubits with code distande d, and using c ebits (or extended qubits). The setup for EAQECC is as follows: we initially have a set of Paulis $\{p_0, p_1, \ldots, p_{k-1}\}$ on n qubits which do not necessarily commute, and then one seeks to create stabilizer codes out of them by using a set of Paulis $\{q_0, q_1, \ldots, q_{k-1}\}$ supported on c extended qubits, such that the set of Paulis $\{p_0 \otimes q_0, p_1 \otimes q_1, \ldots, p_{k-1} \otimes q_{k-1}\}$ on $n + c$ qubits form a valid set of stabilizer generators (meaning that they must mutually commute). Clearly, if one wants to be efficient with respect to the number of physical qubits used to encode a fixed number k of logical qubits, one would want to minimize the number c of extended qubits used in the process (we assume n is fixed). This leads to the task of finding these Paulis $\{q_0, q_1, \ldots, q_{k-1}\}$ such that $[p_i, p_j] = [q_i, q_j]$, for all $0 \leq i, j \leq k - 1$, while at the same time minimizing c. Now, one may use the same procedure to construct EAQECC over qudits, and this has been previously considered in [38, Remark 1], and the result there again applies to prime dimensional qudits only. In the case of quantum convolutional codes, an exactly analogous thing happens — there one seeks to find Paulis supported on a minimum number m of memory qubits that satisfy a memory commutativity matrix specified by the input-output commutation relations [39, Theorem 2]. In order to generalize these convolutional codes to the qudit case, one needs to be able to answer the same question for any qudit dimension d. Our results in Section 5 achieve exactly this.

Yet another instance where the problem of finding Paulis with specified commutation relations

comes up is in the setting of parafermion codes. Parafermion codes [21] are the higher dimensional generalization of Majorana codes [40, 41], just as qudit stabilizer codes generalize qubit stabilizer codes. A key part of the translation between Majorana codes and qubit codes is the construction of sets of qubit Paulis with certain commutation relations, matching those of the underlying Majorana operators [22]. We anticipate a similar translation between parafermion codes and qudit stabilizer codes to involve constructions of qudit Paulis with specified commutation relations that we provide in Section 5. It is to be noted that much of the results in Sections 3 and 4 set up the necessary machinery to be able to resolve this central question of Section 5.

In [42], Georghiu provides a construction of the standard form of a qudit stabilizer group S, i.e. an abelian subgroup of the qudit Pauli group. This is similar to Gottesman's standard form [19] for prime d-dimensional qudits and likewise makes use of Clifford transformations of the group. Our constructions of minimal and Gram-Schmidt generating sets apply to any Pauli subgroup and do not involve applying Cliffords. In the qubit case, such a Gram-Schmidt generating set is useful for improving simulations of stabilizer circuits [43]. We can also efficiently calculate the size of a qudit stabilizer group and thus the dimension of the stabilized codespace using the results in Section 6.3. Qudit stabilizer codes have also been studied previously in [44, 45] in the context of quantum codes with exotic local dimensions, qudit surface codes, and qudit hypermap codes, and some of our results in Section 6 could be applicable to those works also.

2 Preliminaries

In this section, we introduce some background material, some definitions and notation to be used throughout in the paper. We also review some necessary facts from ring theory and establish a few basic results that are helpful later on. Throughout the paper, entries of matrices and vectors (both row and column) will follow zero-based indexing.

2.1 Qudit Pauli group

Consider a collection of n qudits, each with dimension d. Thus each qudit corresponds to a complex d-dimensional Hilbert space $\mathbb{C}^d$ with a basis of orthonormal states: $|0\rangle, |1\rangle, \ldots, |d-1\rangle$. The Pauli operators on one qudit are the invertible operators X and Z from Eq. (1) which we will think of concretely as being given by their $d \times d$ matrix representations with respect to the chosen basis. Thus $X, Z \in \mathbb{C}^{d\times d}$, and notice that $X^d = Z^d = (-1)^{d-1}(XZ)^d = I$, where I is the $d \times d$ identity matrix. We will always use I to denote a context-appropriately-sized identity matrix and specify its shape only when there is a chance for confusion. No smaller number $0 < a < d$ results in X^a, Z^a, or $(XZ)^a$ even being proportional to I.

To define Pauli operators on n qudits, start with the disjoint union space $\mathcal{Z} := \bigsqcup_{n=1}^{\infty} \mathbb{Z}_d^{2n}$, and the map $P : \mathcal{Z} \to \bigsqcup_{n=1}^{\infty} \mathbb{C}^{d^n \times d^n}$ by

$$P(v) := X^{v_0} Z^{v_n} \otimes X^{v_1} Z^{v_{n+1}} \otimes \cdots \otimes X^{v_{n-1}} Z^{v_{2n-1}} \in \mathbb{C}^{d^n \times d^n}, \quad v \in \mathbb{Z}_d^{2n}. \tag{4}$$

Then the set of n-qudit Pauli operators is $\mathcal{P}_n := \{P(v) : v \in \mathbb{Z}_d^{2n}\} \subseteq \mathbb{C}^{d^n \times d^n}$. We call an n-qudit Pauli $P(v)$ X-type if $v_n = v_{n+1} = \cdots = v_{2n-1} = 0$, and Z-type if $v_0 = v_1 = \cdots = v_{n-1} = 0$. These are abbreviated as $X(u)$ and $Z(u)$ respectively, where $u \in \mathbb{Z}_d^n$ specifies just the nonzero entries of v.

The set $\mathcal{P}_n$ does not form a group. However, the set $\overline{\mathcal{P}}_n := \{\omega^j P(v) : v \in \mathbb{Z}_d^{2n}, j \in \mathbb{Z}_d\}$ does form a group with the operation of matrix multiplication, called the *Heisenberg-Weyl Pauli* group. Defining $K = \{\omega^j I : j = 0, 1, \ldots, d-1\}$, we may identify $\mathcal{P}_n$ in one-to-one fashion with the quotient group $\overline{\mathcal{P}}_n / K$ in the obvious way by ignoring phase factors.

Many of our results arise from interest in commutativity relations of elements of $\overline{\mathcal{P}}_n$. For this, we start by noticing that $XZ = \omega^{-1} ZX$. Defining the group commutator of two invertible matrices $[A, B] := ABA^{-1}B^{-1}$, we then have $[X, Z] = \omega^{-1} I$. This implies that for any two elements $\omega^j P(u)$

and $\omega^k P(v)$ of $\overline{\mathcal{P}}_n$

$$[\omega^j P(u), \omega^k P(v)] = [P(u), P(v)] = \omega^{u^T \Lambda v} I, \quad \Lambda = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}, \tag{5}$$

where Λ is written in four $n \times n$ blocks. The first equality of this equation suggests that it suffices to work with the set $\mathcal{P}_n$ for studying commutation relations, and this is largely what we will do outside of Section 6, where some more notation will be introduced just for the material in that section. To isolate the phase, we also define the notations

$$[\![P(u), P(v)]\!] := u^T \Lambda v \in \mathbb{Z}, \quad [\![P(u), P(v)]\!]_d := u^T \Lambda v \mod d \in \mathbb{Z}_d. \tag{6}$$

Two Pauli operators $P(u), P(v) \in \mathcal{P}_n$ are said to commute if and only if $[\![P(u), P(v)]\!]_d = 0$, because that implies $P(u)P(v) = P(v)P(u)$.

We now introduce some key definitions that underlies much of our results. First, in Section 3, we determine the maximum size of collections of non-commuting pairs.

Definition 2.1. A collection of k *non-commuting pairs* on n qudits consists of two ordered sets $S = \{s_0, s_1, \ldots, s_{k-1}\}$, and $T = \{t_0, t_1, \ldots, t_{k-1}\}$, where $S, T \subseteq \mathcal{P}_n$, for all $i, j \in \{0, 1, \ldots, k-1\}$, $[\![s_i, s_j]\!]_d = 0$ and $[\![t_i, t_j]\!]_d = 0$, and moreover $[\![s_i, t_j]\!]_d = 0$ if and only if $i \neq j$. A collection of k *non-commuting CSS pairs* is a collection of non-commuting pairs where every element of S is X-type and every element of T is Z-type.

Thus, in a collection of k non-commuting pairs, there are a total of $2k$ Paulis, and they all commute except for the k pairs (s_i, t_i). An easy consequence of this definition is that all elements of S and T are necessarily distinct from one another and not equal to I.

Later, in Section 5, we will also be concerned with what values the commutators of a collection of non-commuting pairs can have.

Definition 2.2. Suppose we are given n qudits, and a k-tuple of numbers $f := (f_0, \ldots, f_{k-1}) \in (\mathbb{Z}_d \setminus \{0\})^k$. We say that f is an *achievable non-commuting pair relation* on n qudits if there exist non-commuting pairs $S = \{s_0, s_1, \ldots, s_{k-1}\}$ and $T = \{t_0, t_1, \ldots, t_{k-1}\}$ of size k, such that $[\![s_i, t_i]\!]_d = f_i$, for all $i = 0, \ldots, k-1$. We say that S, T *achieves* f.

In Section 4, we count the elements of non-commuting sets.

Definition 2.3. A non-empty subset $S \subseteq \mathcal{P}_n$ is called a *non-commuting set* on n qudits if and only if $[\![p, q]\!]_d \neq 0$ whenever $p, q \in S$ are distinct.

It should be clear from this definition that a non-commuting subset $S \subseteq \mathcal{P}_n$ cannot contain I and all elements of S must be distinct.

2.2 Modules over commutative rings

Our next goal is to provide a short review of some basic results in linear algebra over the commutative ring $\mathbb{Z}_d$ with $d \geq 2$, for the reader who is unfamiliar with these notions. The main difficulties in working with $\mathbb{Z}_d$ and matrices over $\mathbb{Z}_d$ arise when d is composite, because then and only then is $\mathbb{Z}_d$ a ring and not a field. We introduce some ring-theoretic terminology here that will find use in our proofs, and some well-known canonical matrix decompositions such as the Smith normal form and the Howell normal form, which for certain applications generalize the utility of the singular value decomposition and reduced row echelon form, respectively, to certain classes of commutative rings. Separately, we dedicate a section (Appendix A) to the alternating Smith normal form, which is central to this paper and provides a symmetry-respecting Smith normal form for matrices with alternating symmetry, to be defined later. The related material is mostly borrowed from [46, Chapters II, III, X, XIII, XV], and [47, Chapters I-V, XIV,XV].

2.2.1 Basic definitions

Generically, we denote a ring by R, and assume, unless stated otherwise, it is non-zero, commutative, and contains a multiplicative identity 1 and additive identity 0. Largely, we work over the ring of integers $\mathbb{Z}$ or the ring of integers modulo d, $\mathbb{Z}_d$. If $a, b \in R$ such that $a = bc$ for some $c \in R$, we say b *divides* a, and we write $b \mid a$. An element $x \in R$ is called *invertible* (or *unit*) if it has a multiplicative inverse. The subset of units of a ring form a multiplicative group called the *group of invertible elements*. Ring R is called a *field* if every non-zero element is a unit. R is called an *integral domain* if $ab = 0$ for $a, b \in R$, implies either $a = 0$ or $b = 0$. All finite integral domains, such as $\mathbb{Z}_d$ for prime d, are fields by Wedderburn's little theorem [48], but this is not so for infinite rings, like $\mathbb{Z}$. For a ring R that is not an integral domain, an element $0 \neq x \in R$ is called a *zero divisor* if $xy = 0$ for some $R \ni y \neq 0$. The set of zero divisors of R is denoted $\mathcal{Z}(R)$.

An ideal J is an additive subgroup of R that is also closed under multiplication by R, i.e. $J = RJ := \{rj : j \in J, r \in R\}$. An ideal $J \subseteq R$ is called a *principal ideal* if it is generated by a single element, i.e. $J = xR := \{xy : y \in R\}$, for some $x \in R$. A commutative ring R is called a *principal ideal ring* (or PIR), if all its ideals are principal ideals. It is well-known that $\mathbb{Z}_d$ is a PIR [49, Chapter 2.2, Ex. 8]. An integral domain that is also a PIR is called a *principal ideal domain* (or PID). For example $\mathbb{Z}$ is a PID, as it can be shown that any ideal of $\mathbb{Z}$ is of the form $x\mathbb{Z}$ (hence a principal ideal), for some $x \in \mathbb{Z}$.

Given a ring R, a *left module* over R, or simply a left R-module M is an abelian group together with an operation $\cdot : R \times M \to M$, such that for all $a, b \in R$ and $x, y \in M$ we have (i) $(a+b) \cdot x = a \cdot x + b \cdot x$, (ii) $a \cdot (x+y) = a \cdot x + a \cdot y$, (iii) $(ab) \cdot x = a \cdot (b \cdot x)$, and (iv) $1 \cdot x = x$. One can analogously define a *right module* over R, but any right R-module can be viewed as a left R-module and vice-versa, and for us this distinction will not be needed. Thus when the ring R is understood, we will simply say "a module M". A ring R is also a R-module, with the $\cdot$ operation being the same as ring multiplication. The most important modules we will have to deal with are the $\mathbb{Z}_d$-module $\mathbb{Z}_d^k$ and the $\mathbb{Z}$-module $\mathbb{Z}^k$, i.e. the set of all k-tuples of elements of the ring $\mathbb{Z}_d$ and $\mathbb{Z}$ respectively, where the operation $\cdot$ is that of ring multiplication applied to each component of the tuple. Given a R-module M, a subset $S \subseteq M$ is called a *generating set* (over R) of M if any $y \in M$ can be expressed as $y = \sum_{i=1}^{k} a_i x_i$ for some finite k, $a_1, \ldots, a_k \in R$, and $x_1, \ldots, x_k \in S$. We say that M is *finitely generated* if it has a finite generating set. A non-empty subset $S \subseteq M$ is called *linearly independent* (over R) if and only if the following condition holds for every finite positive integer k, $a_1, \ldots, a_k \in R$, and distinct elements $x_1, \ldots, x_k \in S$:

$$\sum_{i=1}^{k} a_i \cdot x_i = 0 \implies a_i = 0 \text{ for every } i = 1, \ldots, k. \tag{7}$$

A linearly independent subset $S \subseteq M$ that is also a generating set of M, is called a *basis* of M. We say that a R-module M is a *free R-module* if M has a basis, or is the zero module. For example, it is easy to show that both $\mathbb{Z}_d^k$ and $\mathbb{Z}^k$ are finitely generated and free: in both cases the set $S = \{e_1, e_2, \ldots, e_k\}$ is a basis for these modules, where $(e_j)_i = 1$ if $i = j$, else 0.

2.2.2 Minimal generating sets

A commutative ring R with a multiplicative identity element is an *invariant basis number ring* (see [50, Chapter 4, Definition 2.8] and [51, Section 1A] for a discussion on invariant basis number rings). This means that any finitely generated free R-module M has a unique finite dimension given by the size of any (hence all) basis of M, which we will call the *rank* of M where, by convention, the zero module has zero rank. For a proof of this statement, the reader is referred to [47, Corollary 5.13], and the discussions thereafter. For example, the modules $\mathbb{Z}_d^k$ and $\mathbb{Z}^k$ discussed above both have rank k. Given a R-module M, a subset $M' \subseteq M$ is called a *submodule* if it is also a R-module. A key fact is that a submodule of a free module need not be free; for example, consider the submodule $\{0, 2\}$ of $\mathbb{Z}_4$ which does not have a basis. Thus, apriori, the rank of the submodule M' is not defined.

In order to be able to extend the notion of rank to submodules of a finitely generated free R-module, we need the notion of minimal generators. Given a finitely generated R-module M, a generating set $S \subseteq M$ is called a *minimal generating set* of M if there does not exist another generating set $S' \subseteq M$ such that $|S'| < |S|$. The quantity $|S|$ is called the *minimal number of generators* of M, and we will denote it as $\Theta(M)$. For example, for the submodule $\{0, 2\}$ of $\mathbb{Z}_4$, the generating set $\{2\}$ is a minimal generating set, and so $\Theta(\{0, 2\}) = 1$. If M is free, it is a well-known result that $\Theta(M)$ equals the rank of M, since R is commutative. So $\Theta(M)$ is indeed a generalization of rank to non-free modules.

Our main interest in the minimal number of generators is motivated by the need to define a quantity analogous to the column rank of a matrix with entries in a field, for the situation we encounter in this paper where the entries of the matrix belong to a commutative ring R (in particular $\mathbb{Z}$ or $\mathbb{Z}_d$). Suppose we have a matrix $C \in R^{k\times t}$, and we label each of its t columns as $c_0, c_1, \ldots, c_{t-1}$. Each column is an element of the R-module R^k, and together they generate a submodule of R^k, namely $M_C := \left\{\sum_{i=0}^{t-1} a_i c_i : a_i \in R \,\forall\, i\right\}$. Now suppose $A \in R^{k\times k}$ and $B \in R^{t\times t}$ are invertible matrices, and recall that a square matrix with entries in R is invertible if and only if the determinant of the matrix is a unit in R [47, Corollary 2.21]. Let $\bar{C} := ACB \in R^{k\times t}$, let its columns be $\bar{c}_0, \bar{c}_1, \ldots, \bar{c}_{t-1}$, and let $M_{\bar{C}}$ be the submodule of R^k generated by the columns of $\bar{C}$. The important property of the minimal number of generators that we will need is the following lemma, which is well-known (although we also provide a proof in Appendix B.1):

Lemma 2.1. *The minimal number of generators of the submodules M_C and $M_{\bar{C}}$ of R^k are equal.*

We also note an important special case in the next lemma, still assuming that R is a commutative ring with a multiplicative identity, where the minimal number of generators for a matrix with special structure is known. We thank Luc Guyot for an outline of this proof in [52]. See Appendix B.1 for the proof.

Lemma 2.2. *Let $C \in R^{k\times t}$ be a matrix such that each row and column has at most one non-zero element, and suppose one of those non-zero elements is divisible by all the others. Then the minimal number of generators of the R-module M_C, generated by the columns of C, is equal to the number of non-zero elements of the matrix C.*

2.2.3 Smith normal form

Let R be a commutative ring with multiplicative identity. Following [47, Chapter 15], R is called an *elementary divisor ring* if for every $k, \ell \geq 1$ and for every $C \in R^{k\times\ell}$, there exists invertible matrices $B \in R^{k\times k}$ and $\tilde{B} \in R^{\ell\times\ell}$, such that (i) $A := BC\tilde{B}$ is a diagonal matrix, and (ii) $a_0 \mid a_1 \mid \cdots \mid a_{r-1}$, where $r = \min\{k, \ell\}$, and $a_i := A_{ii}$ for every i. The diagonal matrix A is called the *Smith normal form* (SNF) of the matrix C, and we say that C admits a *diagonal reduction*. The non-zero elements of the diagonal matrix A are called the *invariant factors* of C.

We recall the following well-known result about the Smith normal form of a matrix with entries from a principal ideal ring R:

Theorem 2.3 ([47, Theorems 15.9, 15.24])**.** *If a commutative ring R is a principal ideal ring, then for all $k, \ell \geq 1$, every matrix $C \in R^{k\times\ell}$ admits a diagonal reduction $A := BC\tilde{B}$, where $B \in R^{k\times k}$ and $\tilde{B} \in R^{\ell\times\ell}$ are invertible matrices, and A is the Smith normal form of C. The Smith normal form A is unique up to multiplication of the diagonal elements of A by units of R.*

The following result about the SNF of a matrix and its relation to the minimal number of generators of the module generated by its columns, will be useful in Section 6, and follows as a direct consequence of Lemma 2.1 and Lemma 2.2:

Lemma 2.4. *Let R be a commutative principal ideal ring. Suppose $C \in R^{k\times\ell}$ has a Smith normal form $A = BC\tilde{B}$, for invertible matrices $B, \tilde{B}$ and A a diagonal matrix of appropriate shape. Let M_C be the submodule generated by the columns of C. Then $\Theta(M_C)$ equals the number of non-zero elements of A.*

Choosing integer n so that $k, \ell, r = O(n)$, finding the Smith normal form of a matrix takes $\tilde{O}(n^{\theta+1})$ time [53] where $\tilde{O}$ hides log factors and assuming $2 < \theta \leq 3$ is the exponent of matrix multiplication.

2.2.4 Howell normal form

We now introduce the Howell normal form of a matrix with entries in the ring $\mathbb{Z}_d$, which will be useful for us in Section 6. The material here is borrowed from [54, 55]. Although we only discuss the $\mathbb{Z}_d$ case here, the Howell normal form is also valid for a matrix over a principal ideal ring, and the reader is referred to [53, 56] for details. Also note that we present the Howell normal form in terms of column spans of a matrix, instead of row spans as done in prior literature (of course they are equivalent).

A matrix $H \in \mathbb{Z}_d^{k\times\ell}$ is said to be in *reduced column echelon* form if H satisfies the following properties:

(i) If H has r non-zero columns, then its first r columns are non-zero.

(ii) For $0 \leq i \leq r-1$, let j_i be the row index of the first non-zero entry of column i. Then $j_0 < j_1 < \cdots < j_{r-1}$.

(iii) For $0 \leq i \leq r-1$, the matrix entry $H_{j_i i}$ is a divisor of d over integers.

(iv) For each $0 \leq t < i \leq r-1$, we have that $0 \leq H_{j_i t} < H_{j_i i}$.

If the matrix H only satisfied properties (i) and (ii) above, then it can be made to satisfy properties (iii) and (iv) also by post-multiplying H on the right by an invertible matrix over $\mathbb{Z}_d$. We say that H is in *Howell normal form* (HNF), if in addition to properties (i)-(iv), the matrix H satisfies the following additional property:

(v) Let M_H be the submodule of $\mathbb{Z}_d^k$ generated by the columns of H. For each $0 \leq i \leq r-1$, consider the submodule $\overline{M}_i := \{v \in M_H : v_t = 0,\ \forall\ 0 \leq t < j_i\}$. Then $\overline{M}_i$ is generated by columns i to $r-1$ of H, for every $0 \leq i \leq r-1$.

Let us now state the main result about the Howell normal form. Given $A \in \mathbb{Z}_d^{k\times\ell}$, we construct $\overline{A} := [A\ \ 0]$, by adding k zero columns to A. Then there exists an invertible matrix $L \in \mathbb{Z}_d^{(\ell+k)\times(\ell+k)}$, such that $\overline{A}L = H$, where $H \in \mathbb{Z}_d^{k\times(k+\ell)}$ is a matrix in Howell normal form. Moreover, the matrix H is uniquely determined.

As part of the calculation of the Howell normal form, some algorithms (e.g. [53]) include computation of a matrix $G \in \mathbb{Z}_d^{(k+\ell)\times(k+\ell)}$, whose columns generate the kernel of H, $\mathrm{Ker}(H) = \{x \in \mathbb{Z}_d^\ell : Hx = 0\}$. So $HG = 0$. Moreover, the columns of LG provides a generating set for the kernel of A. Choosing integer n so that $k, \ell, r = O(n)$, finding both G and H takes $\tilde{O}(n^{\theta+1})$ time.

With the Howell normal form H and kernel G, we can test membership in the column space of A, or equivalently solve linear systems, $\overline{A}x = b$, where we are given $b \in \mathbb{Z}_d^k$ and want to determine whether $x \in \mathbb{Z}^{k+\ell}$ exists. Note that by invertibility of L, this is the same as finding a $y = L^{-1}x$ such that $Hy = b$. Also, if y is a solution then so is $y + Gv$ for any $v \in \mathbb{Z}_d^\ell$. Solving the inhomogenous equation $Hy = b$ is made easy by properties (i-v). In principle, one can observe, [53]

$$\underbrace{\left(\begin{array}{c|c} 1 & 0 \\ \hline b & H \end{array}\right)}_{H_b}\underbrace{\left(\begin{array}{c|c} 1 & 0 \\ \hline -y & I \end{array}\right)}_{L_b} = \left(\begin{array}{c|c} 1 & 0 \\ \hline b' & H \end{array}\right) \tag{8}$$

where the righthand-side is the Howell normal form of H_b and the invertible matrix L_b relating them necessarily has the indicated form. Now, $b' = 0$ if and only if $Hy = b$. In practice, one does not need to run the Howell normal form algorithm on H_b, but can instead achieve the same result by solving $Hy = b$ one entry at a time using the fact H is in reduced row echelon form and has the Howell property (v).

3 Maximum number of non-commuting pairs of Paulis

In this section, we give a precise count of the maximum number of non-commuting pairs that one can achieve with n qudits. This generalizes the well-known result in the qubit ($d = 2$) case, where this count is equal to the number of qubits [19] and prepares for our construction in Section 5 of Pauli sets satisfying given commutation relations using the minimum number of qudits.

It is easy to establish the lower bound nm on the maximum number of non-commuting pairs as illustrated by the following example.

Example 1. *Define the sets* $S = \{s_i = X^{d/p_i^{\alpha_i}} : i \in \{0, 1, \ldots, m-1\}\}$ *and* $T = \{t_i = Z^{d/p_i^{\alpha_i}} : i \in \{0, 1, \ldots, m-1\}\}$. *Then* S, T *is a collection of* m *non-commuting CSS pairs on a single qudit. On* n *qudits, this example can be applied individually to each qudit to get a collection of non-commuting CSS pairs of size* nm.

We now show that nm is also the upper bound. The key result that underlies the proof is the following number-theoretic obstruction:

Lemma 3.1 (Obstruction lemma)**.** *Suppose* $\ell = p^\gamma$, *for some prime number* p, *and positive integer* $\gamma \geq 1$. *For some* $k \geq 2$, *let* A *be a* $k \times k$ *matrix with entries in* $\mathbb{Z}$. *Then the following conditions cannot hold simultaneously*

(i) $\det(A) = 0$,

(ii) A *contains exactly* k *entries that are not divisible by* ℓ, *such that each row and column of* A *contains exactly one such element.*

Proof. For contradiction, assume that such a matrix A exists satisfying both conditions (i) and (ii). Let us denote the columns of A as $c_0, \ldots, c_{k-1}$. We claim that there exist integers $x_0, \ldots, x_{k-1}$, not all zeros, such that $\sum_{i=0}^{k-1} x_i c_i = 0$, which we prove at the end. Now pick any row j. This row contains an element A_{jq} which is not divisible by ℓ, while all other elements of this row are divisible by ℓ. From the condition $\sum_{i=0}^{k-1} x_i c_i = 0$, we thus we have

$$x_q A_{jq} = -\sum_{\substack{i=0 \\ i \neq q}}^{k-1} x_i A_{ji}. \tag{9}$$

The right hand side of Eq. (9) is divisible by ℓ, and A_{jq} is not divisible by ℓ, so we deduce that x_q is divisible by p. By repeating this argument for each row, and using condition (ii), we deduce that x_i is divisible by p for each $i = 0, \ldots, k-1$. Defining $x_i' = x_i/p$ for each i, we now obtain $\sum_{i=0}^{k-1} x_i' c_i = 0$. We can keep repeating the argument, giving an infinite sequence of integers

$$x_i, \frac{x_i}{p}, \frac{x_i}{p^2}, \frac{x_i}{p^3}, \ldots, \quad \text{for every } i = 0, \ldots, k-1. \tag{10}$$

But this implies $x_i = 0$ for every i, giving a contradiction.

We now prove the claim made in the previous paragraph. We may regard A as a matrix over the set of rational numbers $\mathbb{Q}$, and then by condition (i) we have that $\det(A) = 0$ over $\mathbb{Q}$. Thus the columns $c_0, \ldots, c_{k-1}$ are linearly dependent as column vectors over the field $\mathbb{Q}$, which implies there exists $y_0, \ldots, y_{k-1} \in \mathbb{Q}$, not all zeros, such that $\sum_{i=0}^{k-1} y_i c_i = 0$. Multiplying through by the least common denominator of the y_i immediately proves the claim. □

We can now prove the following theorem:

Theorem 3.2 (Largest size non-commuting pairs)**.** *The largest size of a collection of non-commuting pairs on* n *qudits is* nm.

Proof. Suppose by way of contradiction that we have a collection of $k > nm$ non-commuting pairs $S = \{s_0, \ldots, s_{k-1}\}$ and $T = \{t_0, \ldots, t_{k-1}\}$. Then for every j, $[\![s_j, t_j]\!]$ is not divisible by d. By the pigeonhole principle, there is some prime p (from the factorization of d) and a set $J \subseteq \{0, 1, \ldots, k-1\}$ of $n+1$ indices so that for each $j \in J$, $[\![s_j, t_j]\!]$ is not divisible by p^α, where α is the largest integer such that p^α divides d. Without loss of generality, one may assume that $J = \{0, 1, \ldots, n\}$.

Form the matrices $M_{ij}^{(1,1)} = [\![s_i, s_j]\!] = 0$, $M_{ij}^{(1,2)} = [\![s_i, t_j]\!]$, $M_{ij}^{(2,1)} = [\![t_i, s_j]\!]$, and $M_{ij}^{(2,2)} = [\![t_i, t_j]\!] = 0$ for $i, j \in J$. Write $M = \begin{pmatrix} M^{(1,1)} & M^{(1,2)} \\ M^{(2,1)} & M^{(2,2)} \end{pmatrix} \in \mathbb{Z}^{2(n+1)\times 2(n+1)}$. We note that every row and column of M contains exactly one element not divisible by p^α (i.e. the diagonal elements of $M^{(1,2)}$ and $M^{(2,1)}$).

However, we also claim $\det(M) = 0$. Thus, Lemma 3.1 gives our contradiction. To prove that the determinant is zero, we take vectors $u_0, \ldots, u_{2n+1} \in \mathbb{Z}_d^{2n}$, so that $s_i = P(u_i)$ and $t_i = P(u_{i+n+1})$, for every $i = 0, \ldots, n$. Then $M_{ij} = u_i^T \Lambda u_j \in \mathbb{Z}$, where Λ is defined in Eq. (5). We now lift M and the vectors u_i to the field of real numbers, i.e. we may treat $M \in \mathbb{R}^{2(n+1)\times 2(n+1)}$, and $u_i \in \mathbb{R}^{2n}$ for every i. Because there are $2(n+1)$ vectors u_i for $i = 0, \ldots, 2n+1$, it now follows that there are some real numbers $x_0, \ldots, x_{2n+1}$, such that $\sum_{i=0}^{2n+1} x_i u_i = 0$. From this it follows that the columns of M (and the rows) are also linearly dependent and $\det(M) = 0$ over the reals, and hence also over $\mathbb{Z}$. □

Note a consequence of the theorem above:

Corollary 3.3. *Let S, T be a collection of non-commuting pairs of size nm on n qudits. Let $G = \{r \in \mathcal{P}_n : [\![r, r']\!]_d = 0, \ \forall\ r' \in S \cup T\}$. Then G does not contain a non-commuting pair of Paulis.*

Proof. If G contained such a non-commuting pair (s, t), then the ordered sets $S' := S \cup \{s\}$ and $T' := T \cup \{t\}$ would be a non-commuting pair of size $nm+1$, which contradicts Theorem 3.2. □

4 Maximum size of a non-commuting set

It is well-known in the qubit case ($d = 2$), that the maximum size of an anticommuting set on n qubits is $2n + 1$ (see [57, Lemma 8], [58, Appendix G], [59, Theorem 1]). In Section 4.1, we generalize this result to a single qudit of arbitrary dimension d, and in Section 4.2 we provide bounds for the multi-qudit case. The case of constructing Pauli sets satisfying the commutation relations of parafermion operators [21] is a special case of the results in this section, which we specifically solve later in Corollary 5.5.

4.1 Non-commuting sets on a single qudit

In this section, we evaluate gcd of elements of $\mathbb{Z}$ (or $\mathbb{Z}_d$ considered as elements of $\mathbb{Z}$) over $\mathbb{Z}$ and define it to be positive to ensure that it is unique. Also whenever we write a/b for $a, b \in \mathbb{Z}_d$ (or $\mathbb{Z}$), we will always mean division over integers (in particular this must mean that b divides a over integers). Our goal in this section is to prove the following theorem:

Theorem 4.1 (Largest size non-commuting set on 1 qudit)**.** *Let $d = p_0^{\alpha_0} p_1^{\alpha_1} \ldots p_{m-1}^{\alpha_{m-1}}$ be the prime factorization of $d \geq 2$. Then the size of a largest non-commuting subset of $\mathcal{P}_1$ on one d-dimensional qudit is given by the Dedekind psi function $\Psi(d) := d \prod_{j=0}^{m-1}(1 + 1/p_j)$.*

The Dedekind psi function is bounded by $\Psi(d) \geq d + 1$, with equality when d is prime, and $\Psi(d) < e^\gamma d \log\log(d)$ for all $d > 30$ [60].

Recall from the paragraph below Definition 2.3 that it suffices to find the size of a largest non-commuting subset of $\mathcal{P}_1$. The basic strategy of the proof is now to reduce this problem to

finding a maximum clique[a] in a graph. To achieve this, we start by noticing that if $\binom{a}{b}, \binom{s}{t} \in \mathbb{Z}_d^2$, then for the Paulis $P(\binom{a}{b})$ and $P(\binom{s}{t})$ on 1 qudit represented by these vectors, Eq. (6) simplifies to

$$[\![P(\tbinom{a}{b}), P(\tbinom{s}{t})]\!]_d = -\det \begin{pmatrix} a & b \\ s & t \end{pmatrix} \mod d. \tag{11}$$

As a consequence, finding the largest set of non-commuting Paulis on 1 qudit is equivalent to finding a maximum clique C_m in the undirected graph $G = (V, E)$ with vertices $V = (\mathbb{Z}_d \times \mathbb{Z}_d) \setminus \{(0,0)\}$ and edge $((a,b),(s,t)) \in E$ if and only if $\det \binom{a\ b}{s\ t} \not\equiv 0 \mod d$. This commutation graph (or its complement) has been studied before in relation to projective geometry in e.g. [61, 62, 63] and mutually unbiased bases [64]. We will show in several steps below that $|C_m| = \Psi(d)$, thus proving Theorem 4.1.

The first step is to define a special subset of vertices $W \subseteq V$, where $(a,b) \in W$ if and only if $\gcd(a,b,d) = 1$. We prove the following Lemma in Appendix B.2.

Lemma 4.2. *The graph $G = (V, E)$ has the following properties.*

(i) If $C \subseteq V$ is a clique in G, then there is a clique $C' \subseteq W$ of the same size, $|C| = |C'|$.

(ii) If $v_0, v_2 \in V$, $v_1 \in W$, $(v_0, v_1) \notin E$, and $(v_1, v_2) \notin E$, then $(v_0, v_2) \notin E$.

(iii) If (a,b) and (s,t) are in W, then $((a,b),(s,t)) \notin E$ if and only if there exists $u \in \mathbb{Z}$ with $\gcd(u,d) = 1$ such that $(a,b) = (us, ut)$, the right hand side evaluated modulo d.

Lemma 4.2(i) implies there is a maximum clique that is just a subset of W. Part (ii) implies transitivity of an equivalence relation $u \sim v$ for $u, v \in W$, where u and v are said to be equivalent if $(u,v) \notin E$. Part (iii) says equivalence classes of W are all the same size, exactly the size of the group of units of $\mathbb{Z}_d$, which is $\phi(d) = d\prod_{j=0}^{m-1}(1 - 1/p_j)$, the Euler's totient function. Combining these facts, the size of a maximum clique of G is $|C_m| = |W|/\phi(d)$. Thus the remaining step is to count $|W|$.

In the remainder of the proof of Theorem 4.1, we will adopt the following notation: any product of the form $\prod_{p|r}$, where r is a positive integer, will mean that the product is taken over all unique primes p that divide r. To count $|W|$, we let $(a,b) \in W$ and allow $a \in \mathbb{Z}_d$ to be arbitrary. Now $b \in \mathbb{Z}_d$ is restricted by the choice of a. Namely, if $\gcd(a,d) > 1$, b cannot be 0 or share any prime factors with $\gcd(a,d)$. Each prime factor p of $\gcd(a,d)$ reduces the allowable set of b by a factor of $(p-1)/p = (1 - 1/p)$. We thus have the following equality over rationals:

$$\frac{1}{d}|W| = \sum_{a \in \mathbb{Z}_d} \prod_{p|\gcd(a,d)} \left(1 - \frac{1}{p}\right) =: H(d). \tag{12}$$

It turns out H is multiplicative, which we show in Appendix B.2.

Lemma 4.3. *If $d_0, d_1 \in \mathbb{Z}$ are relatively prime, then $H(d_0 d_1) = H(d_0)H(d_1)$.*

Therefore, it is easy to evaluate H for prime powers: $H(p^k) = p^k(1 - 1/p^2)$. Consequently, $H(d) = d\prod_{p|d}(1 - 1/p^2)$. We note for curiosity's sake that $|W| = dH(d) = J_2(d)$ is the second of Jordan's totient functions [65]. With $H(d)$ in hand, we now conclude the calculation:

$$|C_m| = |W|/\phi(d) = dH(d)/\phi(d) = d\prod_{p|d}(1 + 1/p) = \Psi(d). \tag{13}$$

This completes the proof of Theorem 4.1.

[a] In an undirected graph $G(V, E)$, a clique is a subset of vertices with the property that any two distinct vertices have an edge connecting them.

4.2 Bounds on the size of non-commuting sets on n qudits

Suppose we have a non-commuting set $S = \{s_i : i = 0, \ldots, |S| - 1\}$ on n qudits and another non-commuting set $S' = \{s'_i : i = 0, \ldots, |S'| - 1\}$ on n' qudits. Then,

$$S \circ S' := \{s_i \otimes I : i = 1, \ldots, |S| - 1\} \cup \{s_0 \otimes s'_i : i = 0, \ldots, |S'| - 1\} \tag{14}$$

is a non-commuting set of size $|S| + |S'| - 1$ on $n + n'$ qudits. We call this the Jordan-Wigner composition of sets S and S', in analogy to the Jordan-Wigner encoding of fermions into qubits [26]. The Jordan-Wigner composition gives a simple lower-bound on the size of non-commuting sets on n qudits.

Corollary 4.4. *The size of a maximum non-commuting set on n qudits is at least $(\Psi(d)-1)n+1$.*

Proof. The Jordan-Wigner composition of t non-commuting sets S_i, $i = 0, \ldots, t-1$, gives a non-commuting set of size $\sum_{i=0}^{t-1} |S_i| - (t-1)$. With n qudits, each supporting a maximum size non-commuting set S_i with size $|S_i| = \Psi(d)$ (Theorem 4.1), this gives a lower bound of $(\Psi(d)-1)n+1$ on the size of the maximum non-commuting set. □

If larger non-commuting sets are found in special cases, the lower bound can be improved. For instance, for $d = 3$, maximum non-commuting sets on $n = 1$ and $n = 2$ qutrits have sizes 4 and 7 (the case $n = 2$ was found via exhaustive computer search), respectively, matching the $dn+1$ lower bound that Corollary 4.4 implies. But for $n = 3$, we have a computer-verified, size-13, maximum non-commuting set, where the Paulis are $P(c)$ for each column c of the matrix

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 1 & 1 & 1 & 1 & 2 & 2 & 1 & 1 & 2 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 2 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 2 & 0 & 1 & 2 & 1 & 0 & 1 \\ 1 & 0 & 1 & 2 & 1 & 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 \end{pmatrix}. \tag{15}$$

The Jordan-Wigner composition therefore implies a lower-bound on non-commuting sets of qutrit Paulis of $12\lfloor n/3 \rfloor + 3(n \bmod 3) + 1$.

Similarly, for $d = 4$ and $n = 2$, a maximum non-commuting set has size 20. This implies the lower bound of $19\lfloor n/2 \rfloor + 5(n \bmod 2) + 1$ for the maximum size of non-commuting sets on n qudits with $d = 4$.

5 Achievable commutation relations

We established in Section 3 that the maximum number of non-commuting pairs that can exist on n qudits is nm. Thus, if $(f_0, \ldots, f_{k-1}) \in (\mathbb{Z}_d \setminus \{0\})^k$ is an achievable non-commuting pair relation on n qudits, we must have $1 \leq k \leq nm$. Our first goal in this section is to further characterize exactly what tuples $(f_0, \ldots, f_{k-1})$ are achievable non-commuting pair relations on n qudits. Then, we will use this characterization to construct Pauli multisets with given commutation relations.

5.1 Achievable tuples

Note that if $k \leq n$, then every $f = (f_0, \ldots, f_{k-1}) \in (\mathbb{Z}_d \setminus \{0\})^k$ is an achievable non-commuting pair relation on n qudits. On one qubit, $[\![Z^{f_i}, X]\!]_d = f_i$, and so we can use a separate qubit to achieve each of the k non-commuting pair relations. See also Lemma C.1(iii) for a more formal argument.

Our first result obtains a lower bound on the minimum number of qudits needed to achieve a given non-commuting pair relation $(f_0, \ldots, f_{k-1}) \in (\mathbb{Z}_d \setminus \{0\})^k$, that improves on the lower bound $\lceil k/m \rceil$ implied by Theorem 3.2.

Lemma 5.1. *Let $S = \{s_0, s_1, \ldots, s_{k-1}\}$ and $T = \{t_0, t_1, \ldots, t_{k-1}\}$ be a collection of non-commuting pairs of size k on n qudits, and for every i, let $f_i := [\![s_i, t_i]\!]$. Define the multiset $F_j := \{f_i : f_i \not\equiv 0 \pmod{p_j^{\alpha_j}},\ i \in \{0, \ldots, k-1\}\}$, where $d = p_0^{\alpha_0} p_1^{\alpha_1} \ldots p_{m-1}^{\alpha_{m-1}}$. Then $|F_j| \leq n$, for every $j = 0, \ldots, m-1$, and thus $n \geq \max_j |F_j|$.*

Proof. If $k \leq n$, then there is nothing to prove. So we assume $k > n$, and let j be arbitrary. For the sake of contradiction, suppose that $\ell := |F_j| > n$, and without loss of generality, one may then further assume that $F_j = \{f_0, f_1, \ldots, f_{\ell-1}\}$. Define $J := \{0, \ldots, \ell-1\}$, and then define the matrix $M \in \mathbb{Z}^{2\ell \times 2\ell}$, as in the proof of Theorem 3.2. Then by exactly the same argument there, and using $\ell > n$, we conclude that (i) $\det(M) = 0$, and (ii) every row and column of M has exactly one element not divisible by $p_j^{\alpha_j}$. But this is impossible by Lemma 3.1, which gives us the desired contradiction. □

Remark. *In Lemma 5.1, one can replace $f_i := [\![s_i, t_i]\!]$ by $f_i := [\![s_i, t_i]\!]_d$, and the conclusion still holds. This is because, for every i and j, we have $[\![s_i, t_i]\!] \equiv 0 \pmod{p_j^{\alpha_j}}$ if and only if $[\![s_i, t_i]\!]_d \equiv 0 \pmod{p_j^{\alpha_j}}$.*

Remarkably, $n = \max_j |F_j|$ is also a sufficient number of qubits to achieve the non-commuting pair relation $(f_0, \ldots, f_{k-1})$. We prove this below, but before that, we present an example of the construction.

Example 2. *Suppose $d = 30 = 2 \times 3 \times 5$ (so $m = 3$ factors), $n = 3$, and $f = (2, 5, 6, 11, 15)$. Then, $F_0 = \{5, 11, 15\}$ (elements of f not divisible by 2), $F_1 = \{2, 5, 11\}$ (elements not divisible by 3), and $F_2 = \{2, 6, 11\}$ (elements not divisible by 5). We fill a $|f| \times m$ matrix Q whose elements are from $\mathbb{Z}_{n+1}$ subject to the constraints (1) $Q_{ij} = 0$ if and only if $f_i \notin F_j$ and (2) all non-zero entries in each column are unique. Because $n = \max_j |F_j|$, constraint (2) is satisfiable. For instance,*

$$Q = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & 0 \\ 0 & 0 & 2 \\ 2 & 3 & 3 \\ 3 & 0 & 0 \end{pmatrix}. \tag{16}$$

Rows $0, \ldots, 4$ of Q are used to construct $p_0, \ldots, p_4$, X-type Paulis, and $q_0, \ldots, q_4$, Z-type Paulis, that form non-commuting pairs achieving f. To construct p_i, a non-zero entry Q_{ij} says that the prime factor labeled j should be excluded from the power of X on qubit Q_{ij}. We do the same for q_i, but with powers of Z and additional factors of $-\beta_i \in \mathbb{Z}_d$ in the exponents. For this example,

$$\begin{aligned}
p_0 &= X^{30/(3\times5)} \otimes X^{30} \otimes X^{30} = X^2 \otimes I \otimes I, & q_0 &= Z^{-2\beta_0} \otimes I \otimes I \\
p_1 &= X^{30/2} \otimes X^{30/3} \otimes X^{30} = X^{15} \otimes X^{10} \otimes I, & q_1 &= Z^{-15\beta_1} \otimes Z^{-10\beta_1} \otimes I \\
p_2 &= X^{30} \otimes X^{30/5} \otimes X^{30} = I \otimes X^6 \otimes I, & q_2 &= I \otimes Z^{-6\beta_2} \otimes I \\
p_3 &= X^{30} \otimes X^{30/2} \otimes X^{30/(3\times5)} = I \otimes X^{15} \otimes X^2, & q_3 &= I \otimes Z^{-15\beta_3} \otimes Z^{-2\beta_3} \\
p_4 &= X^{30} \otimes X^{30} \otimes X^{30/2} = I \otimes I \otimes X^{15}, & q_4 &= I \otimes I \otimes Z^{-15\beta_4}.
\end{aligned} \tag{17}$$

Note that $[\![p_i, q_j]\!]_d = 0$ for $i \neq j$ regardless of the values of the β_i. The β_i must be chosen so that $[\![p_i, q_i]\!]_d = f_i$. Namely, $\beta_0 = 8$, $\beta_1 = 5$, $\beta_2 = 1$, $\beta_3 = 29$, and $\beta_4 = 1$.

It was not a coincidence we could find values of β_i to make the desired non-commutation relations in the previous example. Lemma B.1 in the appendix gives the necessary statement in general. This also enables the following constructive proof that $n = \max_j |F_j|$ qubits are sufficient in general to achieve a non-commuting pair relation.

Theorem 5.2 (Minimum qudits achieving a non-commuting pair relation)**.** *Suppose that we are given $f := (f_0, \ldots, f_{k-1}) \in (\mathbb{Z}_d \setminus \{0\})^k$. For every $j = 0, \ldots, m-1$, define the multisets $F_j := \{f_i : f_i \not\equiv 0 \pmod{p_j^{\alpha_j}},\ i \in \{0, \ldots, k-1\}\}$. Then the minimum number of qudits needed for which f is an achievable non-commuting pair relation is $\max_j |F_j|$. The non-commuting pairs generating f can be chosen to be CSS.*

Proof. By Lemma 5.1 $n \geq \max_j |F_j|$ qudits are necessary to achieve f. To show that $\max_j |F_j|$ qudits are sufficient, consider a matrix $Q \in \mathbb{Z}_{n+1}^{k\times m}$, where $Q_{ij} \neq 0$ if and only if $f_i \in F_j$. Because $n = \max_j |F_j|$, by choosing non-zero entries from $\mathbb{Z}_{n+1} \setminus \{0\}$, we can arrange that within each column of Q, every non-zero entry is unique. We interpret these non-zero entries as qudit labels for qudits $1, 2, \ldots, n$.

For each row $i = 0, \ldots, k-1$, we will now show how to construct a pair of Paulis $X(u_i)$ and $Z(v_i)$ that are supported only on qudits indicated in that row and $[\![X(u_i), Z(v_i)]\!]_d = f_i$. Then the ordered sets $\{X(u_0), X(u_1), \ldots, X(u_{k-1})\}$ and $\{Z(v_0), Z(v_1), \ldots, Z(v_{k-1})\}$ are CSS non-commuting pairs generating f. Define the set of indices in row i where Q takes value h: $S_{ih} := \{j : Q_{ij} = h\}$ and $\overline{S}_{ih} := \{j : Q_{ij} \neq h\}$. By construction, $p_j^{\alpha_j}$ divides f_i for all $j \in S_{i0}$, and so $f_i = \gamma \prod_{j\in S_{i0}} p_j^{\alpha_j}$ for some integer γ.

Set $(u_i)_{h-1} = \prod_{j\in\overline{S}_{ih}} p_j^{\alpha_j}$ and $v_i = -\beta u_i$, where β is some integer (that depends on i) to be determined. So, we have

$$[\![X(u_i), Z(v_i)]\!] = \beta \sum_{h\in\mathbb{Z}_{n+1}\setminus\{0\}} \Big(\prod_{j\in\overline{S}_{ih}} p_j^{2\alpha_j}\Big) = \beta \prod_{j\in S_{i0}} p_j^{2\alpha_j} \sum_{h\in\mathbb{Z}_{n+1}\setminus\{0\}} \Big(\prod_{j\in\overline{S}_{ih}\setminus S_{i0}} p_j^{2\alpha_j}\Big). \tag{18}$$

Choose $\beta = \gamma\beta'$ for another integer β'. Apply Lemma B.1 with $d' = f_i$ and

$$d'' = \Big(\prod_{j\in S_{i0}} p_j^{\alpha_j}\Big) \sum_{h\in\mathbb{Z}_{n+1}\setminus\{0\}} \Big(\prod_{j\in\overline{S}_{ih}\setminus S_{i0}} p_j^{2\alpha_j}\Big), \tag{19}$$

where we note that for any $j \notin S_{i0}$ (i.e. exactly the set of indices for which $p_j^{\alpha_j}$ does not divide f_i), p_j does not divide d''. To elaborate, p_j divides all but one term in the sum in Eq. (19). Therefore, Lemma B.1 implies there is some integer β' so that $[\![X(u_i), Z(v_i)]\!] = \beta' d' d'' \equiv f_i \pmod d$.

Lastly, we show that $[\![X(u_i), Z(v_j)]\!]_d = 0$ for $i \neq j$.

$$[\![X(u_i), Z(v_j)]\!] = \beta \sum_{h\in\mathbb{Z}_{n+1}\setminus\{0\}} \prod_{l\in\overline{S}_{ih}} p_l^{2\alpha_l} \prod_{k\in\overline{S}_{jh}} p_k^{2\alpha_k}. \tag{20}$$

Because within each column of Q the entries are unique, $\overline{S}_{ih} \cup \overline{S}_{jh} = \{0, 1, \ldots, m-1\}$. Thus, d divides $[\![X(u_i), Z(v_j)]\!]$ and $[\![X(u_i), Z(v_j)]\!]_d = 0$. □

Achievable non-commuting pair relations f of maximum size, i.e. $|f| = nm$, can be characterized more directly. We do so in Appendix C.

5.2 Qudits needed to achieve a matrix of commutation relations

If R is a principal ideal ring, and $C \in R^{k\times k}$ is a matrix satisfying $C_{ii} = 0$ and $C_{ij} = -C_{ji}$, for all $i, j = 0, 1, \ldots, k-1$, then such a matrix is called an *alternating* matrix over R. Suppose we are given an alternating matrix $C \in \mathbb{Z}_d^{k\times k}$ and wish to find $P \in \mathbb{Z}_d^{k\times 2n}$ such that

$$P\Lambda P^T = C. \tag{21}$$

Here, rows of P can be interpreted as n-qudit Paulis possessing the commutation relations specified by C. The goal of this section is to answer the following question: what is the minimum number of qudits n for which such a P can be found?

The following lemma, proved constructively in appendix A, is useful in answering this question.

Lemma 5.3 (Alternating Smith Normal Form)**.** *Given an alternating matrix $A \in \mathbb{Z}_d^{k\times k}$, there are matrices $L, B \in \mathbb{Z}_d^{k\times k}$, where B is alternating with at most one non-zero entry per row and column and L is invertible, such that $A = LBL^T$. We may further arrange B so that it is non-zero only in the top-left $2r \times 2r$ block which has the form $\bigoplus_{i=1}^{r} \left(\begin{smallmatrix} 0 & \beta_i \\ -\beta_i & 0 \end{smallmatrix}\right)$ for integer $r = \Theta(M_A)/2$, where M_A is the $\mathbb{Z}_d$-submodule generated by the columns of A, and each $\beta_i \in \mathbb{Z}_d$ non-zero, satisfying $\beta_i \mid \beta_{i+1}$ for all $i < r$. Also, for all $i = 1, \ldots, r$, β_i is uniquely determined up to multiplication by a unit by the formula $\beta_i = d_{2i}/d_{2i-1} \bmod d$ (or, alternatively, the formula $\beta_i = d_{2i-1}/d_{2i-2} \bmod d$), where d_j is the greatest common divisor of all $j \times j$ minors of A (and $d_0 := 1$).*

Remark. *Note that in the formulas for β_i in the above lemma, both the minors and the greatest common divisor are first evaluated over integers, and then the division is also performed over integers. Finally the modulo d operation gives back an element of $\mathbb{Z}_d$. Similar to the remark following Lemma A.2, one should note that the smallest integer j for which $d_j/d_{j-1} \equiv 0 \pmod{d}$ is $2r+1$, and thus an odd integer (the fact that this integer is odd for any d is itself quite an interesting fact). In fact, if k is odd, then such an odd integer j must exist as $d_k = 0$. We can also easily deduce that $d_{j'} \neq 0$ for all $j' < j$. Furthermore, it follows from Lemma A.2 that for all $j' \geq j$ such that $d_{j'-1} \neq 0$, we also have $d_{j'}/d_{j'-1} \equiv 0 \pmod{d}$.*

Apply the Lemma to C, finding invertible L and alternating B such that $C = LBL^T$. Then defining $Q := L^{-1}P \in \mathbb{Z}_d^{k\times 2n}$ we have

$$(L^{-1}P)\Lambda(L^{-1}P)^T = Q\Lambda Q^T = B, \tag{22}$$

and let $r = \Theta(M_C)/2$ with M_C denoting the submodule generated by the columns of C. Now, since B is alternating with at most one non-zero entry per row and column, the Paulis represented by the first $2r$ rows of Q are simply non-commuting pairs (the last $k-2r$ rows can be chosen to be all 0s, representing identity Paulis). In Theorem 5.2, we concluded that the necessary and sufficient number of qudits needed to achieve a set of non-commuting relations $\{\beta_1, \beta_2, \ldots, \beta_r\}$ is $\max_j |F_j|$, where $F_j = \{\beta_i : \beta_i \not\equiv 0 \pmod{p_j^{\alpha_j}}, i \in \{1, 2, \ldots, r\}\}$. Now there exists some p_j so that $\beta_r \not\equiv 0 \pmod{p_j^{\alpha_j}}$, because otherwise we would have $\beta_r = 0$. Since $\beta_i | \beta_{i+1}$ for all $i < r$, if $\beta_r \not\equiv 0 \pmod{p_j^{\alpha_j}}$, then $\beta_i \not\equiv 0 \pmod{p_j^{\alpha_j}}$ for all i. Thus, $\max_j |F_j| = r$ qudits are necessary and sufficient to construct Q and also $P = LQ$.

The above establishes the following theorem:

Theorem 5.4 (Minimum qudits achieving commutation relations)**.** *If $C \in \mathbb{Z}_d^{k\times k}$ is an alternating matrix, and $Q \in \mathbb{Z}_d^{k\times 2n}$ satisfies $Q\Lambda Q^T = C$, then $n \geq \Theta(M_C)/2$, where M_C is the submodule generated by the columns of C. Moreover, there exists a matrix $P \in \mathbb{Z}_d^{k\times\Theta(M_C)}$, such that $C = P\Lambda P^T$. Rows of P indicate k $(\Theta(M_C)/2)$-qudit Paulis possessing the commutations relations specified by C.*

For some particular forms of alternating matrix $C \in \mathbb{Z}_d^{k\times k}$, one can compute $\Theta(M_C)$ analytically. We mentioned one such case already before in Lemma 2.2. Another case happens when all entries in the upper triangular part of C, except the diagonal, are equal. That is, suppose C is alternating with $C_{ij} = t \in \mathbb{Z}_d \setminus \{0\}$, for all $j > i$. Then applying Lemma 5.3 again gives matrices $L, B \in \mathbb{Z}_d^{k\times k}$, with L invertible and B alternating and of the form given by the lemma. Let us calculate the quantities $d_0, d_1, \ldots, d_k$ as defined in Lemma 5.3. First suppose $t = 1$. Then we have:

(i) $d_0 = 1$ by definition.

(ii) For every even j, the determinant (evaluated over integers) of the top-left $j \times j$ block of the matrix C is 1, which is easily verified by bringing C to upper-triangular form using row (or column) operations. In particular, if k is even, then $d_k = \det(C) = 1$.

(iii) If k is odd, $d_k = \det(C) = 0$ as C is an alternating matrix. Also by (ii) we get $d_{k-1} = 1$, as at least one $(k-1)\times(k-1)$ minor is 1.

Then by the chain of divisibilities condition mentioned in the remark following Lemma A.2, we conclude that for $t = 1$, we have $d_j = 1$ for all j if k is even, while if k is odd, then we have $d_k = 0$ and $d_j = 1$ for all $j < k$. Now for arbitrary $t \in \mathbb{Z}_d \setminus \{0\}$, we simply note that d_j equals t^j times the value of d_j for the case $t = 1$. Combining these facts, and using the formulas in Lemma 5.3, we obtain the following:

(i) If k is even, then $\beta_i = t$ for all $1 \leq i \leq k/2$. Thus $\Theta(M_C) = k$.

(ii) If k is odd, then $\beta_i = t$ for all $1 \leq i \leq \lfloor k/2 \rfloor$. Thus $\Theta(M_C) = k-1$.

From these observations, we immediately obtain the following corollary as a direct consequence of Theorem 5.4:

Corollary 5.5. *Let $t \in \mathbb{Z}_d \setminus \{0\}$. Then the largest size of a non-commuting set S on n-qudits, such that $[\![p, q]\!]_d = \pm t$, for every distinct $p, q \in S$, is $2n + 1$.*

A construction of such a non-commuting set of qudit Paulis can be found in [21]. To describe it, we let X_i (resp. Z_i) denote the single-qudit Pauli X (resp. Z) on the qudit i. For $j = 1, 2, \ldots, n$ let

$$\gamma_{2j-1} = Z_j \prod_{i=1}^{j-1} X_i, \quad \gamma_{2j} = \omega^{(d+1)/2} X_j Z_j \prod_{i=1}^{j-1} X_i. \tag{23}$$

Also, let $\gamma_0 = \prod_{j=1}^{n} \gamma_{2j-1}\gamma_{2j}^{\dagger}$. Then $[\![\gamma_i, \gamma_j]\!]_d = \omega$ for all $i < j$, and $S = \{\gamma_0, \gamma_1, \ldots, \gamma_{2n}\}$ forms the claimed non-commuting set of size $2n + 1$.

6 Some group theoretic results

In this section, we depart from the previous sections where we studied elements of $\mathcal{P}_n$, and instead we take up the study of the Heisenberg-Weyl Pauli group $\overline{\mathcal{P}}_n$ for a d-dimensional qudit, without ignoring the phases. We begin this section by establishing the notion of equivalent generating sets in Theorem 6.3. Using this, we give in Section 6.1 a characterization of a near minimal generating set of any subgroup of $\overline{\mathcal{P}}_n$ and also an algorithm to compute such a generating set. We also provide an (inefficient) algorithm to compute a minimal generating set of any subgroup of $\overline{\mathcal{P}}_n$ from a near minimal generating set. Next, in Section 6.2, we provide a way to compute a Gram-Schmidt generating set of any subgroup of $\overline{\mathcal{P}}_n$. This generalizes the well-known stabilizer-destabilizer decomposition [43] of the qubit Pauli group to the qudit case. Finally, in Section 6.3 we develop some results to compute the size of any subgroup of $\overline{\mathcal{P}}_n$. This generalizes similar results in [42], where only the special case of stabilizer subgroups was studied. Of particular note in this subsection is the square-free theorem (Theorem 6.15), which says that maximal sets of non-commuting pairs generate the qudit Pauli group when d is square-free.

Let us first introduce some notation that will help the discussion. All products in this section will be ordered, unless mentioned otherwise. Recall from Section 2.1 that any element $p \in \overline{\mathcal{P}}_n$ has the form $\omega^j P(v)$ for some $v \in \mathbb{Z}_d^{2n}$ and $j \in \mathbb{Z}_d$, where $\omega = e^{2\pi i/d}$, and for a given p, the corresponding values of j and v are uniquely determined. Thus we can equivalently represent p by the tuple (j, v), and the map $p \mapsto (j, v)$ sets up a bijection $\overline{\mathcal{P}}_n \to \mathbb{Z}_d \times \mathbb{Z}_d^{2n}$. For any $p \in \overline{\mathcal{P}}_n$, we define $p^0 := I$. Representing a Pauli as an element of $\mathbb{Z}_d \times \mathbb{Z}_d^{2n}$, we define the projection maps π_1 and π_2 onto the first and second factors respectively, i.e. $\pi_1((j, v)) = j$ and $\pi_2((j, v)) = v$, for a Pauli represented by the tuple (j, v). Similarly, for an ordered multiset $S := \{q_0, q_1, \ldots, q_{k-1}\}$, we use the notation $\pi_1(S) := \{\pi_1(q_0), \pi_1(q_1), \ldots, \pi_1(q_{k-1})\}$, and $\pi_2(S) := \{\pi_2(q_0), \pi_2(q_1), \ldots, \pi_2(q_{k-1})\}$. We may unambiguously associate $\pi_2(S)$ with a $2n \times k$ matrix with elements in $\mathbb{Z}_d$, where the j^{th} column is $\pi_2(q_j)$, and we will also denote the matrix by $\pi_2(S)$ when there is no chance for confusion.

Now suppose we have two Paulis $p, q \in \overline{\mathcal{P}}_n$ represented by (j, v) and (k, w) respectively. Then it is an easy exercise to check that pq is represented by $(\ell, v + w)$, where $\ell = j + k + v^T \left(\begin{smallmatrix} 0 & 0 \\ I & 0 \end{smallmatrix}\right) w$, evaluated modulo d, where I here is an $n \times n$ identity matrix. From this, one can also easily show that for $t \geq 1$, p^t is represented by (ℓ, tv), where $\ell = tj + \frac{t(t-1)}{2} v^T \left(\begin{smallmatrix} 0 & 0 \\ I & 0 \end{smallmatrix}\right) v$, evaluated modulo d, and thus computing both $p+q$ and p^t takes $O(n)$ operations (for t bounded). The following well-known lemma is easy to establish (see also [62, 64]):

Lemma 6.1. *Let $p = \omega^j P(v) \in \overline{\mathcal{P}}_n$. If d is odd, then $p^d = I$. If d is even, then $p^d = \pm I$, with $p^d = I$ if and only if $v^T \left(\begin{smallmatrix} 0 & 0 \\ I & 0 \end{smallmatrix}\right) v$ is even.*

Thus in the case of odd d, the maximum possible order of any Pauli is d, while in the case of even d, the maximum possible order of a Pauli is $2d$ (for example, on 1 qudit, the Pauli XZ has order $2d$ when d is even).

Next we would like to figure out a way to transform a generating set of a subgroup of $\overline{\mathcal{P}}_n$ to another generating set of the same group with the same number of generators. Before we present the main theorem on this, let us consider the case of one generator. Suppose $p \in \overline{\mathcal{P}}_n$ generates a group G. For some unit $s \in \mathbb{Z}_d$, let us define $q := p^s$. We can then show that q also generates G. To see this let $t \in \mathbb{Z}_d$ be the unit such that $st = 1$ in $\mathbb{Z}_d$, i.e. treating s, t as elements of $\mathbb{Z}$ we have $st = kd + 1$ for some non-negative integer k. Then $q^t = (p^s)^t = p^{st} = p^{kd+1} = p^{kd}p$. Now if d is odd, then $p^{kd} = I$, and if d is even, then $p^{kd} = \pm I$ by Lemma 6.1, and thus $q^t = \pm p$. If $q^t = p$, then q generates G. The case $q^t = -p$ is more interesting. This is precisely the case when $p^d = -I$ and d is even, and thus $q^{t(d+1)} = q^{td}q^t = p$, which again shows that q generates G. Thus we have proved that $\langle p \rangle = \langle q \rangle = G$. Theorem 6.3 generalizes this observation to multiple generators. We will also need Lemma 6.2, whose proof is easy and is left to the reader.

Notation: Recall that $[p, q]$ is the group commutator of $p, q \in \overline{\mathcal{P}}_n$. Let $S \subseteq \overline{\mathcal{P}}_n$ be a set or multiset. We will denote $I_S := \{p \in \langle S \rangle : \pi_2(p) = 0\}$ to be the set of all elements of $\langle S \rangle$ that are proportional to I. We will also denote $J_S := \{[p, p'] : p, p' \in S\} \cup \{p^d : p \in S\}$. Then I_S is a subgroup and moreover we have $\langle J_S \rangle \subseteq I_S \subseteq \langle S \rangle$.

Lemma 6.2. *Suppose $S := \{q_0, q_1, \ldots, q_{k-1}\} \subseteq \overline{\mathcal{P}}_n$ is a multiset. Then*

(i) $\langle J_S \rangle \subseteq \langle S \rangle$.

(ii) Let $T \subseteq \langle S \rangle$. Then $\langle J_T \rangle \subseteq \langle J_S \rangle$. Moreover, if $\langle T \rangle = \langle S \rangle$, then $\langle J_T \rangle = \langle J_S \rangle$.

(iii) If $S' \subseteq S$ is chosen such that $q \in S \setminus S'$ implies $\pi_2(q) = 0$, then $\langle J_S \rangle = \langle J_{S'} \rangle$.

(iv) Let $T := \{q'_0, q'_1, \ldots, q'_{k-1}\} \subseteq \overline{\mathcal{P}}_n$ be a different multiset such that $q'_j = q_j \omega^{\beta_j}$, for some $\beta_j \in \mathbb{Z}_d$ for each j. Then $\langle J_T \rangle = \langle J_S \rangle$.

Theorem 6.3 (Equivalent generating sets)**.** *Suppose $S := \{q_0, q_1, \ldots q_{k-1}\}$ is an ordered multiset of elements of $\overline{\mathcal{P}}_n$. Let $A \in \mathbb{Z}_d^{k \times k}$ be an invertible matrix, and consider the ordered multiset $T := \{q'_0, q'_1, \ldots, q'_{k-1}\} \subseteq \overline{\mathcal{P}}_n$, where we define $q'_i := \prod_{j=0}^{k-1} q_j^{A_{ij}}$. Then $\langle S \rangle = \langle T \rangle$.*

Proof. This follows a similar line as the one generator example above. Keeping track of the phases that is facilitated by Lemma 6.2. A full proof is given in Appendix B.4. □

Our next goal is to provide a complete characterization of the subgroup of phases I_S, given a generating set $S \subseteq \overline{\mathcal{P}}_n$. For instance, for any $p, q \in S$, both $[p, q]$ and p^d are proportional to identity and thus members of I_S. There may be additional products of generators that are proportional to identity as well, and these can be read off the kernel of $\pi_2(S)$. That these are all the elements of I_S is the content of the next lemma.

Recall that given a matrix $A \in \mathbb{Z}_d^{k \times \ell}$, the kernel of A is a submodule of $\mathbb{Z}_d^\ell$ defined as $\ker(A) := \{v \in \mathbb{Z}_d^\ell : Av = 0\}$.

Lemma 6.4. *Suppose $S := \{q_0, q_1, \ldots, q_{k-1}\} \subseteq \overline{\mathcal{P}}_n$ is an ordered multiset. Consider the matrix $\pi_2(S) \in \mathbb{Z}_d^{2n \times k}$, and let $\overline{K} \in \mathbb{Z}_d^{k \times \ell}$ be such that the columns of $\overline{K}$ is a generating set for $\ker(\pi_2(S))$. Then we have the following:*

(i) Let $v \in \ker(\pi_2(S))$, and define $q' := \prod_{j=0}^{k-1} q_j^{v_j}$. Then $\pi_2(q') = 0$.

(ii) Define the set $N_S := \left\{\prod_{j=0}^{k-1} q_j^{v_j} : v \in \ker(\pi_2(S))\right\}$. Then $I_S = \langle N_S, J_S \rangle$.

(iii) Define the multiset $\overline{K}_S := \left\{\prod_{j=0}^{k-1} q_j^{\overline{K}_{ji}} : i = 0, 1, \ldots, \ell - 1\right\}$. Then $I_S = \langle \overline{K}_S, J_S \rangle$.

Proof. See Appendix B.4. □

The group $K = \langle \omega I \rangle$ is isomorphic to the additive group on $\mathbb{Z}_d$. Subgroups of $\mathbb{Z}_d$ have the property that they can always be generated by one element (in other words, $\mathbb{Z}_d$ is a principal ideal ring). Thus, the following lemma says the same is true of subgroups of K.

Lemma 6.5. *Let* $S := \{\omega^{\mu_j} I : j = 0, 1, \ldots, t,\ \mu_j \in \mathbb{Z}_d\} \subset \overline{\mathcal{P}}_n$. *Let* $\mathcal{I}$ *be the ideal generated by* $\{\mu_0, \ldots, \mu_t\}$. *Then*

(i) $\langle S \rangle = \{\omega^{\mu} I : \mu \in \mathcal{I}\}$.

(ii) There exists $\mu \in \mathcal{I}$ *such that* $\mathcal{I}$ *is generated by* μ, *and* $\langle S \rangle = \langle \omega^{\mu} I \rangle$. *One can choose* $\mu = \gcd(\mu_0, \mu_1, \ldots, \mu_t, d)$, *where* gcd *is the greatest common divisor evaluated over integers.*

Algorithm 1 Return a generator for I_S given $S \subseteq \overline{\mathcal{P}}_n$

```
1:  procedure IDENTITY_GENERATOR(S := {q_0, q_1, ..., q_{k-1}}, d)
2:      μ ← d
3:      if d is even then                    ▷ Generate phases of I from d^th powers of generators
4:          for 0 ≤ j ≤ k − 1 do
5:              t ← π_1(q_j^d)
6:              if t = d/2 then
7:                  μ ← d/2;   Exit for loop
8:      if μ = 1 then
9:          return ωI
10:
11:     for 0 ≤ j < ℓ ≤ k − 1 do             ▷ Generate phases of I from group commutators
12:         t ← π_1([q_j, q_ℓ])
13:         μ ← gcd(t, μ)
14:         if μ = 1 then
15:             return ωI
16:
17:     Compute K̄ ∈ Z_d^{k×s} such that π_2(S)K̄ = 0, columns of K̄ span ker(π_2(S))
18:
19:     for 0 ≤ j ≤ s − 1 do                 ▷ Generate all other phases of I
20:         q ← ∏_{ℓ=0}^{k−1} q_ℓ^{K̄_{ℓj}}
21:         t ← π_1(q)
22:         μ ← gcd(t, μ)
23:         if μ = 1 then
24:             return ωI
25:
26:     return ω^μ I
```

Lemma 6.4 and Lemma 6.5 leads to a simple algorithm (Algorithm 1) to find a generator for I_S given a multiset $S \subseteq \overline{\mathcal{P}}_n$. The basic idea is to get a generating set for I_S and then use Lemma 6.5(ii). In the algorithm, $\gcd(a, b)$ refers to the evaluation of the greatest common divisor over integers, for integers $0 \leq a, b \leq d$, which takes $O(M(d) \log(\log(d)))$ operations, where $M(d)$ is the number of operations needed to multiply two integers no greater than d (so $M(d) = O(\log^2 d)$ for elementary-school arithmetic, for instance) [53]. In order to facilitate early termination, instead of evaluating all the generators and then evaluating the gcd using Lemma 6.5(ii) of the whole list, we use the property $\gcd(a, b, c) = \gcd(a, \gcd(b, c))$ recursively, to update the gcd everytime we have a new generator. If at any stage the gcd becomes 1, we can terminate the algorithm (Lines 8-9, Lines 14-15, Lines 23-24). In Lines 3-7, we use the fact from Lemma 6.1 that for odd d, any $p \in \overline{\mathcal{P}}_n$ satisfies $p^d = I$, and for even d satisfies $p^d = \pm I$. Lines 6-7 exploits this fact, and terminates checking the d^{th} powers of the remaining generators, if a generator $q_j \in S$ is detected such that $q_j^d = -I$. In Lines 11-15, notice that we exploit the fact that $\pi_1([q_j, q_\ell]) + \pi_1([q_\ell, q_j]) \equiv 0 \pmod{d}$, or equivalently $[q_j, q_\ell] = [q_\ell, q_j]^{-1}$; thus only the commutators $[q_j, q_\ell]$ need to be considered for $j < \ell$. In Line 17, the computation of the kernel matrix $\overline{K}$ can be carried out using the techniques in [53, Chapter 5]. In terms of computational cost, the computation of q_j^d and $[q_j, q_\ell]$ in Line 5 and Line 12 respectively involves $O(n)$ operations, while computation of q is Line 20 involves $O(kn)$ operations.

6.1 Minimal and near-minimal generating sets

We are now in a position to answer the following question: given a multiset $S \subseteq \overline{\mathcal{P}}_n$, can we find a non-empty generating set of $\langle S \rangle$ of the smallest size? Such a set is called a *minimal generating set* of $\langle S \rangle$. The "non-empty" condition is only relevant for the case $\langle S \rangle = \{I\}$. This subsection is dedicated to providing a *nearly complete* solution to this problem. We begin by stating a result from commutative algebra that will be needed below, whose proof can be found in [66] (we thank Jeremy Rickard for the outline of the proof):

Lemma 6.6. *Let $A, B \in \mathbb{Z}_d^{k \times \ell}$. Then the following conditions are equivalent.*

(i) The submodules of $\mathbb{Z}_d^k$ generated by the columns of A and B are the same.

(ii) There exists an invertible matrix $C \in \mathbb{Z}_d^{\ell \times \ell}$ such that $A = BC$.

Also note the following simple corollary of this lemma, which we prove for completeness in Appendix B.4.

Corollary 6.7. *Let $A \in \mathbb{Z}_d^{k \times p}$ and $B \in \mathbb{Z}_d^{k \times q}$. If the columns of A and B generate the same submodule of $\mathbb{Z}_d^k$, then the number of invariant factors of A and B are equal.*

Now for the rest of this subsection, suppose $S := \{s_0, s_1, \ldots, s_{k-1}\} \subseteq \overline{\mathcal{P}}_n$ is an ordered mutltiset. The case when $\pi_2(S)$ has zero invariant factors is special. In this case, we have $\langle S \rangle = I_S$, and then a minimal generating set of I_S can be obtained using Lemma 6.5(ii). Thus, through the remainder of this section we assume that $\pi_2(S)$ has at least one invariant factor. The following lemma is easy to establish, which gives us a lower and upper bound on the size of a minimal generating set of $\langle S \rangle$.

Lemma 6.8 (The size of a smallest generating set)**.** *For any ordered multiset $S \subseteq \overline{\mathcal{P}}_n$, if $\pi_2(S) \in \mathbb{Z}_d^{2n \times k}$ has r invariant factors in its Smith normal form, then the smallest generating set of $\langle S \rangle$ has either r or $r+1$ elements.*

Proof. This follows from Lemma B.4 in Appendix B.4. The crux is similar to the canonical generating set in [42], but finds a canonical generating set for an arbitrary group rather than just a qudit stabilizer group, and does not change the Pauli basis using Clifford operators. □

Based on this result, we make the following definition.

Definition 6.1. Given an ordered multiset $S := \{s_0, s_1, \ldots, s_{k-1}\} \subseteq \overline{\mathcal{P}}_n$, such that r is the number of invariant factors of $\pi_2(S)$, we call a subset $T \subseteq \langle S \rangle$ a *near-minimal generating set* of $\langle S \rangle$, if T is of the form $T := T' \cup \{p\}$ and satisfies (i) $\langle T \rangle = \langle S \rangle$, (ii) $\langle p \rangle = I_S$, and (iii) $|T'| = r$. Note that if $r = 0$, then T' is empty; so this definition also works for that case.

Lemma B.4 gives us a way to compute a near-minimal generating set of $\langle S \rangle$. At this point, it begs the question of whether a near-minimal generating set of $\langle S \rangle$ is also a minimal generating set of $\langle S \rangle$ or not. While we do not completely resolve this question here, we prove some partial results below in the remainder of this subsection. Let us first show that indeed there are cases where a near-minimal generating set is a minimal generating set. An easy example is the case when the number of invariant factors of $\pi_2(S)$ is zero. In this case, a near-minimal generating set of $\langle S \rangle$ has size one, and hence it is a minimal generating set. We give another example below.

Example 3. *Suppose $S = \{X, \omega I\} \subseteq \overline{\mathcal{P}}_n$, for any $d > 2$ and number of qudits n. Then one checks easily that $\langle S \rangle = \{\omega^j X^a : j, a \in \mathbb{Z}_d\}$. Thus $|\langle S \rangle| = d^2$. Computing the Smith normal form of $\pi_2(S)$ shows that the number of its invariant factors is one. Thus a near-minimal generating set of $\langle S \rangle$ has size two. Suppose for contradiction that there exists a minimal generating set T of $\langle S \rangle$ with $|T| = 1$. Since the maximum possible order of any element of $\overline{\mathcal{P}}_n$ is $2d$ (Lemma 6.1), this implies that $|\langle T \rangle| \leq 2d < d^2$. So T cannot generate $\langle S \rangle$. Thus all near-minimal generating sets of $\langle S \rangle$ are also minimal generating sets of $\langle S \rangle$ in this example.*

Let us now give an example where a near-minimal generating set is not a minimal generating set, i.e. there exists generating sets of $\langle S \rangle$ of exactly size $r \geq 1$, where r is the number of invariant factors of $\pi_2(S)$. One such example is when $r = k$ in Lemma 6.8 (then clearly S is itself minimal). Another one is given below in Example 4.

Definition 6.2 ([19])**.** A stabilizer group is a subgroup $G \subseteq \overline{\mathcal{P}}_n$ such that $\{\omega^j I : j \in \mathbb{Z}_d\} \cap G = \{I\}$, or, equivalently, $I_G = \{I\}$.

Example 4. *Assume that $\langle S \rangle$ is a stabilizer group and $\langle S \rangle \neq \{I\}$. In this case, if S is any generating set of $\langle S \rangle$, we know the size of a minimal generating set of $\langle S \rangle$ is exactly the number of invariant factors r of $\pi_2(S)$. This is because a near-minimal generating set $T' \cup \{p\}$ of G must satisfy $p = I$, and thus $\langle T' \cup \{p\} \rangle = \langle T' \rangle$.*

How far are we from computing a minimal generating set of $\langle S \rangle$, given we have a near-minimal generating set of $\langle S \rangle$? The next result gives a nice structure theorem to find minimal generating sets from near-minimal ones.

Theorem 6.9. *Given $S \subseteq \overline{\mathcal{P}}_n$, suppose that $T := T' \cup \{p\}$ is a near-minimal generating set of $\langle S \rangle$ with $\langle p \rangle = I_S$. Let the number of invariant factors of $\pi_2(S)$ be $r \geq 1$, and suppose $T' = \{q_0, q_1, \ldots, q_{r-1}\}$. Then the following conditions are equivalent.*

(i) T is not a minimal generating set of $\langle S \rangle$.

(ii) There exist integers $\gamma_0, \gamma_1, \ldots, \gamma_{r-1} \in \mathbb{Z}_d$, such that $T'' := \{p^{\gamma_0} q_0, p^{\gamma_1} q_1, \ldots, p^{\gamma_{r-1}} q_{r-1}\}$ is a minimal generating set of $\langle S \rangle$.

(iii) There exist integers $\gamma_0, \gamma_1, \ldots, \gamma_{r-1} \in \mathbb{Z}_d$, such that $p \in \langle T'' \rangle$, with T'' defined in part (ii).

Proof. Denote by M the submodule of $\mathbb{Z}_d^{2n}$ generated by the columns of $\pi_2(S)$. We first prove that (i) implies (ii). Suppose that $U := \{u_0, u_1, \ldots, u_{r-1}\}$ is a minimal generating set of $\langle S \rangle$. By Lemma B.4(ii), (v), we know that the columns of both $\pi_2(T')$ and $\pi_2(U)$ generate M. Thus by Lemma 6.6 we can conclude that there exists an invertible matrix $C \in \mathbb{Z}_d^{r \times r}$ such that $\pi_2(T') = \pi_2(U)C$. Now define a new set $T'' := \{t_0, t_1, \ldots, t_{r-1}\}$ such that $t_i = \prod_{j=0}^{r-1} u_j^{C_{ji}}$, for every i. Lemma 6.3 then implies that $\langle T'' \rangle = \langle U \rangle = \langle S \rangle$. It also follows from the equality $\pi_2(T') = \pi_2(U)C$ that $\pi_2(t_i) = \pi_2(q_i)$ for every i. Thus each t_i is equivalent to q_i up to some phase factor, and since $\langle p \rangle = I_S$, we immediately conclude that $t_i = p^{\gamma_i} q_i$ for some $\gamma_i \in \mathbb{Z}_d$, for every i.

Now assume (ii) is true. As $\langle T'' \rangle = \langle S \rangle$, this implies that $p \in \langle T'' \rangle$, proving (iii).

Finally assume that (iii) is true, and we want to prove (i). Clearly the definition of T'' in (iii) implies that $\langle T'' \rangle \subseteq \langle S \rangle$, because $p, q_i \in T$, and hence $p^{\gamma_i} q_i \in \langle S \rangle$ for every i. To prove the reverse containment, note that $p \in \langle T'' \rangle$ implies that $q_i \in \langle T'' \rangle$ for every i (since $p_i^{\gamma_i} q_i \in T''$ by definition). This shows that $T \subseteq \langle T'' \rangle$ and thus $\langle S \rangle \subseteq \langle T'' \rangle$. Thus T'' is a minimal generating set of $\langle S \rangle$ (since it has size r), and this proves (i). □

Algorithm 2 Given a near-minimal generating set $T = T' \cup \{p\}$, find a minimal generating set

```
procedure FIND_MINIMAL_GENERATING_SET(T' := {q_0, q_1, ..., q_{r-1}}, p, d)
    for (γ_0, γ_1, ..., γ_{r-1}) ∈ Z_d^r do
        T'' ← {p^{γ_0} q_0, p^{γ_1} q_1, ..., p^{γ_{r-1}} q_{r-1}}
        p̃ ← IDENTITY_GENERATOR(T'', d)
        if p ∈ ⟨p̃⟩ then                          ▷ Check membership of p in ⟨p̃⟩
            return T''
    return T' ∪ {p}
```

Condition (iii) of Theorem 6.9 can be used to obtain a simple (but inefficient) algorithm to test whether a near-minimal generating set of $\langle S \rangle$ is also a minimal generating set or not, and then output a minimal generating set of $\langle S \rangle$. This is given in Algorithm 2. We assume that $\pi_2(S)$ has

$r \geq 1$ invariant factors, so that a computed near-minimal generating set (using Lemma B.4) of $\langle S \rangle$ has size $r+1$. Suppose here that $r < |S|$, so that we are in the setting of Lemma B.4(v). Let $T = T' \cup \{p\}$ be one such near-minimal generating set with $\langle p \rangle = I_S$, and let $T' = \{q_0, q_1, \ldots, q_{r-1}\}$. The algorithm proceeds by looping over each r-tuple $(\gamma_0, \gamma_1, \ldots, \gamma_{r-1}) \in \mathbb{Z}_d^r$, and then constructing the set $T'' \subseteq \langle S \rangle$ of size r. If T'' is a minimal generating set of $\langle S \rangle$, then by Theorem 6.9 we must have $p \in I_{T''}$. To check this condition, in Lines 3-5 we use Algorithm 1 to compute $\tilde{p} := \omega^\beta I$ with the property that $\beta \in \mathbb{Z}_d$ is the smallest possible integer satisfying $\langle \tilde{p} \rangle = I_{T''}$. Then if $p = \omega^\delta I$ for some $\delta \in \mathbb{Z}_d$, it follows that $p \in I_{T''}$ if and only if β divides δ. If this condition check succeeds for any $(\gamma_0, \gamma_1, \ldots, \gamma_{r-1}) \in \mathbb{Z}_d^r$, then we have found a minimal generating set T'' of $\langle S \rangle$ of size r, and otherwise we conclude that $T = T' \cup \{p\}$ is a minimal generating set of $\langle S \rangle$ of size $r+1$.

Another thing to note about Algorithm 2 is that the computational work in Line 4 can be significantly reduced by precomputing and storing a few quantities. Note that due to Lemma 6.2(iv), the results of Lines 2-17 (of Algorithm 1) in the execution of IDENTITY_GENERATOR do not change irrespective of T'' (or equivalently irrespective of the choice of the r-tuple $(\gamma_0, \gamma_1, \ldots, \gamma_{r-1}) \in \mathbb{Z}_d^r$ in Line 2 of Algorithm 2) — thus one can compute and store $\overline{K}$ and the μ value (let us call this μ_0) up to this point (Line 17), with the choice $\gamma_0 = \gamma_1 = \cdots = \gamma_{r-1} = 0$ (i.e. $T'' = T'$). Subsequently, every time IDENTITY_GENERATOR gets called in Line 4 of Algorithm 2, we can initialize Algorithm 1 at Line 19, using the precomputed $\overline{K}$ and setting $\mu = \mu_0$. In fact, if $\mu_0 = 1$, then there is nothing to compute – we know that IDENTITY_GENERATOR will return $\tilde{p} = \omega I$, which also means that the membership check in Line 5 of Algorithm 2 will succeed. For the case $\mu_0 \neq 1$, there are also ways to speed up the execution of Lines 19-26 of Algorithm 1. The main bottleneck is Line 20 which has a complexity of $O(rn)$ operations. But this can be reduced to $O(r)$ using another precomputation step: note that we need to compute $q \leftarrow \prod_{\ell=0}^{r-1} (p^{\gamma_\ell} q_\ell)^{\overline{K}_{\ell j}}$ in Line 20, which can be simplified as $\prod_{\ell=0}^{r-1} (p^{\gamma_\ell} q_\ell)^{\overline{K}_{\ell j}} = \left(\prod_{\ell=0}^{r-1} q_\ell^{\overline{K}_{\ell j}} \right) p^{\sum_{\ell=0}^{r-1} \gamma_\ell \overline{K}_{\ell j}}$, and thus the quantity $\left(\prod_{\ell=0}^{r-1} q_\ell^{\overline{K}_{\ell j}} \right)$ can be precomputed for each $j = 0, 1, \ldots, s-1$ (here s is the number of columns of $\overline{K}$). Putting everything together, and ignoring the precomputation step, we see that Algorithm 2 has a run time complexity of $O(sd^r(r + M(d) \log\log(d)))$, where $M(d)$ is the cost of integer multiplication of non-negative integers less than d. The computational complexity of the precomputation step is upper bounded by the complexity of Algorithm 1. The exponential factor of d^r in the complexity of Algorithm 2 is undesirable, and coming up with a more efficient algorithm is left for future work.

We finish this subsection by stating a special case of Theorem 6.9 below, when d is prime. In this case, the theorem simplifies.

Lemma 6.10. *Given $S \subseteq \overline{\mathcal{P}}_n$, suppose that $T := T' \cup \{p\}$ is a near-minimal generating set of $\langle S \rangle$ with $\langle p \rangle = I_S$. Let d be prime. Then the following conditions are equivalent.*

(i) T is a minimal generating set of $\langle S \rangle$.

(ii) $\langle T' \rangle$ is a stabilizer subgroup of $\overline{\mathcal{P}}_n$, and $I_S \neq \{I\}$.

Proof. See Appendix B.4. □

6.2 A Gram-Schmidt generating set of a subgroup

The symplectic Gram-Schmidt procedure (described in [25] for qubits and easily extended to prime d) takes a generating set S for a Pauli subgroup $G = \langle S \rangle$ and returns another generating set $S' = S_1 \cup S_2 \cup U$, where S_1, S_2 is a collection of non-commuting pairs (in the qubit case, anti-commuting pairs), U is a subset of the center $\mathcal{Z}(G)$ of G,[b] and $G = \langle S' \rangle$. Note, in particular, that for such a generating set S_1, S_2, and U must be disjoint. Here we describe a procedure achieving the same result, a Gram-Schmidt generating set, for qudit Pauli groups.

[b] Recall that, for a group G, its center $\mathcal{Z}(G)$ is the subgroup of elements in G that commute with everything in G.

Lemma 6.11 (Gram-Schmidt generators)**.** *Let $S = \{s_0, s_1, \ldots, s_{k-1}\} \subseteq \overline{\mathcal{P}}_n$ and let $A_{ij} = [\![s_i, s_j]\!]_d$ define a matrix $A \in \mathbb{Z}_d^{k\times k}$. The group $G := \langle S\rangle$ has a Gram-Schmidt generating set $S_1 \cup S_2 \cup U$ where $|S_1| = |S_2| = \Theta(M_A)/2$, M_A is the submodule of $\mathbb{Z}_d^k$ generated by the columns of A, and moreover there does not exist a Gram-Schmidt generating set with smaller S_1 or S_2.*

Proof. Constructing a Gram-Schmidt generating set makes use of the alternating Smith normal form from Lemma 5.3. Denote by $Q \in \mathbb{Z}_d^{k\times 2n}$ the matrix whose rows correspond to each Pauli in S, ignoring the phase factors. Then $A = Q\Lambda Q^T \in \mathbb{Z}_d^{k\times k}$. Now Lemma 5.3 implies the existence of $L, B \in \mathbb{Z}_d^{k\times k}$ such that $A = LBL^T$, where L is invertible and B is of the form specified by the lemma, whose top-left block has the form $\bigoplus_{i=1}^{r} \left(\begin{smallmatrix} 0 & \beta_i \\ -\beta_i & 0 \end{smallmatrix}\right)$ for non-zero $\beta_1, \beta_2, \ldots, \beta_r$, and all other elements are zero. Here $r = \Theta(M_A)/2$. Let $S' = \{s'_0, \ldots, s'_{k-1}\}$, where $s'_i := \prod_{j=0}^{k-1} s_j^{L^{-1}_{ij}}$ with ordering of the product irrelevant to our result. Because the commutation relations of S are given by $A = Q\Lambda Q^T$, the commutation relations of S' are given by $(L^{-1}Q)\Lambda(L^{-1}Q)^T = B$. Therefore, in particular it is a Gram-Schmidt generating set $S' = S_1 \cup S_2 \cup U$ with $|S_1| = |S_2| = r$. Moreover, the invertibility of L gives $\langle S\rangle = \langle S'\rangle$ via Theorem 6.3. Thus, this procedure gives the required Gram-Schmidt generating set.

Now we prove that there is no Gram-Schmidt generating set containing fewer non-commuting pairs. Suppose for contradiction, there exists sets $T, H \subseteq G$ such that $\langle T, H\rangle = G$, $|T| < 2r$, and $H \subseteq \mathcal{Z}(G)$. Construct a multiset $E = T \cup J \cup H$, where J is a multiset of $2r - |T|$ phaseless identity Paulis. Let $P, P' \in \mathbb{Z}_d^{(2r+|H|+|U|)\times 2n}$ be matrices such that the rows of P (resp. P') represent the Paulis in the multiset $E \cup U$ (resp. $S' \cup H$), modulo the phase factors. We first note that since $\langle E, U\rangle = \langle S', H\rangle = G$, this implies that the submodules generated by the columns of P^T and P'^T are equal. Thus, by Lemma 6.6 there exists an invertible matrix V such that $P = VP'$. The number of non-zero entries in the ASNF of $B' = \left(\begin{smallmatrix} B & 0 \\ 0 & 0 \end{smallmatrix}\right) = P'\Lambda P'^T$ is $2r$, and by the uniqueness part of the ASNF in Lemma 5.3, this should equal the number of non-zero entries in the ASNF of $P\Lambda P^T = VB'V^T$. However, by the assumptions on the sets T and H,

$$C = P\Lambda P^T = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}, \tag{24}$$

where $D \in \mathbb{Z}_d^{|T|\times|T|}$, which in turn implies that C cannot have more than $|T| < 2r$ non-zero entries in its ASNF, obtaining our contradiction. □

With Lemma 6.11, we can make a Gram-Schmidt generating set $S_1 \cup S_2 \cup U$ with minimally sized S_1 and S_2. However, there is no such minimality guarantee on U. If d is prime, the only elements of $\mathcal{Z}(G)$ that can be generated by $\langle S_1, S_2\rangle$ are exactly those in $K = \{\omega^j I : j \in \mathbb{Z}_d\}$. If d is not prime, it is possible for non-trivial elements of $\mathcal{Z}(G)$ to be generated by $\langle S_1, S_2\rangle$, including elements of U. The ways this can happen is limited however.

Lemma 6.12. *Suppose $S = \{s_0, \ldots, s_{k-1}\}, T = \{t_0, \ldots t_{k-1}\}$ is a collection of non-commuting pairs, $f_i = [\![s_i, t_i]\!]$ for every i, and let $H = \langle S, T\rangle$. Then*

$$\mathcal{Z}(H) = \langle \omega^{f_i} I, s_i^{a_i}, t_i^{a_i} : i = 0, \ldots, k-1, a_i \in \mathbb{Z} \text{ s.t. } a_i f_i \equiv 0 \pmod d\rangle. \tag{25}$$

Proof. Let $A = \langle \omega^{f_i} I, s_i^{a_i}, t_i^{a_i} : i = 0, \ldots, k-1, a_i \in \mathbb{Z} \text{ s.t. } a_i f_i \equiv 0 \pmod d\rangle$ and note that $J_H \subseteq A$. We see that $\omega^{f_i} I = [s_i, t_i]$, $s_i^{a_i}$ and $t_i^{a_i}$ are all in $\mathcal{Z}(H)$ because they are products of generators of H and they commute with all generators of H. Thus, $A \subseteq \mathcal{Z}(H)$.

To show containment in the other direction, let $p \in \mathcal{Z}(H)$. Therefore, p can be written in terms of the generating set of H as

$$p = \omega^c \prod_{i=0}^{k-1} s_i^{a_i} t_i^{b_i} \tag{26}$$

for some $a_i, b_i \in \mathbb{Z}_d$ and $\omega^c I \in J_H \subseteq A$. Evaluate the commutator: $[\![p, t_j]\!] = [\![s_j^{a_j}, t_j]\!] = a_j [\![s_j, t_j]\!] = a_j f_j$. Since $p \in \mathcal{Z}(H)$ the value of this commutator modulo d must be 0, implying $a_j f_j \equiv 0 \pmod d$. This argument was independent of j, and an analogous argument works for commutators with s_j, implying $b_j f_j \equiv 0 \pmod d$. Therefore, $p \in A$, completing the proof. □

Let $U = \{u_0, \ldots, u_{k-1}\}$. One can use the Howell normal form and the approach outlined in Section 2.2.4 to check membership of $\pi_2(u_0)$ in $\mathcal{Z}(\langle S_1, S_2\rangle)$ and, if it is, check that the phase can be corrected by an element of $I_{\langle S_1,S_2\rangle}$. If that is also the case, then u_0 is redundant and can be removed from U. Otherwise, continue by checking whether u_1 is in $\mathcal{Z}(\langle S_1, S_2, u_0\rangle)$ and so forth.

6.3 Sizes of subgroups of $\overline{\mathcal{P}}_n$

In this section, we investigate the sizes of subgroups of $\overline{\mathcal{P}}_n$ given their generating sets. First, let's assume nothing about the generating set. Using the Smith normal form, we can prove the following.

Lemma 6.13 (Pauli subgroup size)**.** *Given an ordered multiset $S := \{q_0, q_1, \ldots, q_{k-1}\} \subseteq \overline{\mathcal{P}}_n$, suppose that the invariant factors of $\pi_2(S)$ are $d_0, d_1, \ldots, d_{r-1}$. Then $|\langle S\rangle| = |I_S| \prod_{j=0}^{r-1} |d_i\mathbb{Z}_d| = |I_S| \prod_{j=0}^{r-1} (d/\gcd(d_i, d))$.*

Proof. Firstly, note that since invariant factors are unique up to units in $\mathbb{Z}_d$, the sizes of the ideals $|d_i\mathbb{Z}_d|$ do not depend on this choice; so the expression for $|\langle S\rangle|$ is well-defined. Thus, let $\pi_2(S)$ have a Smith normal form $D \in \mathbb{Z}_d^{2n\times k}$, a diagonal matrix satisfying $\pi_2(S)P = QD$, where $P \in \mathbb{Z}_d^{k\times k}$ and $Q \in \mathbb{Z}_d^{2n\times 2n}$ are invertible, and the first r (and only) non-zero diagonal entries of D are $d_0, d_1, \ldots, d_{r-1}$. Let the first r columns of Q be $Q_0, Q_1, \ldots, Q_{r-1}$. Then the submodule of $\mathbb{Z}_d^{2n}$ generated by the columns of $\pi_2(S)$ equals that generated by the set $\{d_iQ_i : i = 0, 1, \ldots, r-1\}$; call this submodule M. It is then clear that $|\langle S\rangle| = |I_S|\,|\langle S\rangle/I_S| = |I_S|\,|M|$.

It follows from the structure theorem of finitely generated modules over a principal ideal ring [47, Chapter 15] that $|M| = \prod_{j=0}^{r-1} |d_j\mathbb{Z}_d|$. But it is possible to prove the same in an elementary fashion. Since Q is invertible, any non-empty subset of the columns of Q generate a free submodule. Thus $Q_0, Q_1, \ldots, Q_{r-1}$ form a basis, and it follows that $|M| = \prod_{j=0}^{r-1} |\{xd_jQ_j : x \in \mathbb{Z}_d\}| = \prod_{j=0}^{r-1} |\{xd_j : x \in \mathbb{Z}_d\}| = \prod_{j=0}^{r-1} (d/\gcd(d_j, d))$. □

In the special case that S is a qudit stabilizer group, i.e. an abelian subgroup of the n-qudit Pauli group, there is a codespace consisting of qudit states that are $+1$ eigenvectors of all the elements of S. This codespace has dimension $d^n/|\langle S\rangle|$ [42], which we can efficiently calculate using Lemma 6.13.

Next, consider subgroups of $\overline{\mathcal{P}}_n$ that are generated by non-commuting pairs of qudit Paulis. Such subgroups enjoy some surprising properties. For example, if $S = \{s_0, s_1, \ldots, s_{k-1}\}$ and $T = \{t_0, t_1, \ldots, t_{k-1}\}$ is a set of non-commuting pairs, then $S \cup T$ is an independent set of generators, in the sense that the group generated by removing any generator from $S \cup T$ leads to a proper subgroup of $\langle S, T\rangle$. To see this note that if $s_0 \in \langle S \setminus \{s_0\}, T\rangle$, then it must imply that $[\![s_0, t_0]\!]_d = 0$, which is a contradiction. We could have argued the same with any other generator of $S \cup T$.

A particularly interesting result is Theorem 6.15, where it is shown that a maximum collection of non-commuting pairs generate the entire qudit Pauli group $\overline{\mathcal{P}}_n$ when d is square free. The following lemma proves useful.

Lemma 6.14. *Let $p, q \in \overline{\mathcal{P}}_n$ and $G \subseteq \overline{\mathcal{P}}_n$. If every element in G commutes with both p and q and $[\![p, q]\!]_d = c$ is a unit in $\mathbb{Z}_d$, then $|\langle p, q, G\rangle/K| = d^2|\langle G\rangle/I_G|$ and $|\langle p, q, G\rangle| = d^3|\langle G\rangle/I_G|$.*

Proof. The second equality differs from the first in that it includes arbitrary phases, which one can create from products of p and q: $\{(pqp^\dagger q^\dagger)^j : j = 0, 1, \ldots, d-1\} = \{\omega^{jc}I : j = 0, 1, \ldots, d-1\} = K$, as c is a unit in $\mathbb{Z}_d$.

To prove the first equality, note $|\langle p\rangle/I_{\{p\}}| = |\langle q\rangle/I_{\{q\}}| = d$. If $|\langle p\rangle/I_{\{p\}}|$ were not d, then $p^a \in K$ for $0 < a < d$ and thus $0 = [\![p^a, q]\!]_d \equiv ac \pmod d$, contradicting the fact that c is a unit. Next, we see that for any $j \not\equiv 0 \pmod d$, p^j is not in $\langle q, G\rangle$ even up to a phase. If it were, then p^j would commute with q, but $[\![p^j, q]\!]_d = jc \bmod d$, which is not zero because c is a unit and $j \not\equiv 0 \pmod d$. Likewise, q^j is not in $\langle G\rangle$ up to a phase for any $j \not\equiv 0 \pmod d$. Thus, $|\langle p, q, G\rangle/K| = |\langle p\rangle/I_{\{p\}}|\,|\langle q, G\rangle/I_{\{q\}\cup G}| = |\langle p\rangle/I_{\{p\}}|\,|\langle q\rangle/I_{\{q\}}|\,|\langle G\rangle/I_G| = d^2|\langle G\rangle/I_G|$. □

Theorem 6.15 (The square-free theorem)**.** *Suppose $d = p_0 p_1 \dots p_{m-1}$ is square-free. If S, T is a collection of nm non-commuting pairs, then $\langle S, T\rangle = \overline{\mathcal{P}}_n$.*

Proof. A collection of nm non-commuting pairs is a maximum size such collection, characterized by Theorem C.2. This implies that, for the appropriate order of the set $S \cup T$, the commutator matrix C is block-diagonal with 2×2 blocks: $C = \bigoplus_{j=0}^{m-1} C_j = \bigoplus_{j=0}^{m-1} \bigoplus_{i=0}^{n-1} C_{ij}$, $C_{ij} = \left(\begin{smallmatrix} 0 & f_{ij} \\ -f_{ij} & 0 \end{smallmatrix}\right)$, and $\gcd(f_{ij}, d) = \prod_{h\neq j} p_h$ for all i, j.

Put C into alternating Smith normal form $C = LBL^T$ using Lemma 5.3. Due to the simple form of C, it is easy to evaluate d_{2i}, the greatest common divisors of all $2i \times 2i$ matrix minors of C:

$$d_{2i} \equiv \begin{cases} 1, & i \leq n \\ d^{2(i-n)}, & i > n \end{cases}. \tag{27}$$

Therefore, for appropriate invertible L, we have $B = \bigoplus_{i=1}^{n} \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$.

The Smith normal form implies the existence of a collection of non-commuting pairs, $S' = \{s'_0, \dots, s'_{n-1}\}$ and $T' = \{t'_0, \dots, t'_{n-1}\}$ where $[\![s'_i, t'_i]\!]_d = 1$, and $G \subseteq \overline{\mathcal{P}}_n$ that commutes with every element of S' and T'. Moreover $\langle S', T', G\rangle = \langle S, T\rangle$. More specifically, S', T', and G can be found from S and T by taking appropriate products specified by L.

By Lemma 6.14, $|\langle S', T'\rangle| = d^{2n+1}$. This is the size of $\overline{\mathcal{P}}_n$. Therefore, $G = \{I\}$ and $\overline{\mathcal{P}}_n = \langle S', T'\rangle = \langle S, T\rangle$. □

A collection of non-commuting pairs S, T that generates the full Heisenberg-Weyl Pauli group $\overline{\mathcal{P}}_n$ is a convenient generating set $B = S \cup T = \{s_0, t_0, \dots, s_{k-1}, t_{k-1}\}$ for the entire group. If we have $p \in \overline{\mathcal{P}}_n$, it has some decomposition in terms of the generators in B which can be rearranged to

$$p = \omega^c \prod_{i=0}^{k-1} s_i^{a_i} t_i^{b_i}, \tag{28}$$

where $\omega^c I \in J_B$ and $a_i, b_i \in \mathbb{Z}_d$. We can write $\omega^c I$ in terms of a product of the generators by using a modified form of Algorithm 1 that records which product of generators gives ωI. Also, using the fact that S, T are non-commuting pairs, we have $[\![p, t_i]\!] = a_i [\![s_i, t_i]\!]$, for every $i = 0, 1, \dots, k-1$. In particular, $a_i = [\![p, t_i]\!]/[\![s_i, t_i]\!] \mod d$ must be an integer. Likewise, we have $b_i = [\![s_i, p]\!]/[\![s_i, t_i]\!] \mod d$, for every i. This makes the decomposition of p into generators from B rather simple, just amounting to calculation of $[\![s_i, p]\!]$ and $[\![p, t_i]\!]$, and the inclusion of an overall phase.

To conclude this section, we note that for more general non-square-free d, we can get lower bounds on the group size generated by non-commuting pairs. We state this as a corollary of Lemma B.3, which we state and prove in Appendix B.4.

Corollary 6.16. *Suppose $S = \{s_0, \dots, s_{k-1}\}, T = \{t_0, \dots t_{k-1}\}$ is a collection of non-commuting pairs. For every i, let $f_i = [\![s_i, t_i]\!]_d$, and denote the order of f_i in $\mathbb{Z}_d$ as $a_i := d/\gcd(f_i, d)$. Then we have the lower bound $|\langle S, T\rangle / I_{S\cup T}| \geq \prod_{i=0}^{k-1} a_i^2$.*

7 Discussion

We have applied the theory of modules over commutative rings to study properties of the qudit Pauli group. One problem we have left open is the maximum size of a non-commuting set on $n > 1$ qudits. Though we have provided some lower bounds on the size, we are also lacking a non-trivial upper bound. A second problem that we left open is the efficient construction of minimal generating sets of qudit Pauli subgroups in all cases. The missing step to get from a near-minimal generating set to a minimal one involves determining whether appropriate phases can be added to each generator as in Theorem 6.9.

Further work on qudit group theory might involve applying techniques similar to those we developed here to selecting an element of the qudit Clifford group uniformly at random. If this

line of work parallels similar developments in the qubit case [67, 68], interesting structure of the qudit Clifford group could be discovered.

Acknowledgments

The authors would like to acknowledge Sergey Bravyi, Andrew Cross, Robert König, Seth Merkel, and John Smolin for motivating discussions. We are also extremely grateful to Prof. Brian Conrad for bringing to our attention the existence of the ASNF for bilinear alternating forms (Theorem A.1), which was key to this paper, and for clarifying several aspects of commutative algebra that arose in this work.

Appendices

A An Alternating Smith Normal Form

A.1 Bilinear alternating forms

For a commutative ring R with multiplicative identity, consider a R-module M. Let $g : M \times M \to R$ be a *bilinear form*. This means that for all $x, y, z \in M$ and $a \in R$, we have (i) $g(x + y, z) = g(x, z) + g(y, z)$, (ii) $g(x, y + z) = g(x, y) + g(x, z)$, and (iii) $g(a \cdot x, y) = g(x, a \cdot y) = ag(x, y)$. We say that g is *alternating* if $g(x, x) = 0$ for all $x \in M$, which implies $g(x, y) + g(y, x) = 0$ for all $x, y \in M$. A key fact about alternating bilinear forms is the following theorem:

Theorem A.1 ([46, Ch XV, Exercise 17])**.** *Let M be a finitely generated free module of rank k over a commutative ring R with multiplicative identity, and let $g : M \times M \to R$ be a bilinear alternating form. If R is a PIR, then there exists a basis $\{e_i : i = 0, 1, \ldots, k-1\}$ of M, and a non-negative integer r with $2r \leq k$, such that*

(i) $g(e_{2i}, e_{2i+1}) = -g(e_{2i+1}, e_{2i}) = a_i \neq 0$*, for $i = 0, 1, \ldots, r-1$, where $a_0, a_1, \ldots, a_{r-1} \in R$, while $g(e_i, e_j) = 0$ for all other pairs of indices i, j,*

(ii) $a_0 \mid a_1 \mid \cdots \mid a_{r-1}$.

Moreover, the ideals $\{ba_i : b \in R\}$ are uniquely determined, for every $i = 0, 1, \ldots, r-1$, irrespective of the choice of the basis.

Now suppose that R is a principal ideal ring, and $A \in R^{k \times k}$ is an alternating matrix over R (see beginning of Section 5.2 for definition of an alternating matrix). Then the module R^k is a finitely generated free R-module of rank k, as mentioned previously. The matrix A defines a bilinear alternating form $g : R^k \times R^k \to R$ as follows:

$$g(x, y) := x^T A y, \quad \text{for all } x, y \in R^k. \tag{29}$$

We can then apply Theorem A.1 to this setting to extract a basis $\{e_i : i = 0, 1, \ldots, k-1\}$ of R^k, a non-negative integer $r \leq \lfloor k/2 \rfloor$, and $a_0, a_1, \ldots, a_{r-1} \in R$ as in the theorem, such that $B \in R^{k \times k}$ defined by the equation $LBL^T = A$ is an alternating matrix, where the columns of $(L^{-1})^T \in R^{k \times k}$ are $e_0, e_1, \ldots, e_{k-1}$, and

$$B_{jj'} = \begin{cases} a_{j/2} & \text{if } (j, j') \in \{(2i, 2i+1) : i = 0, 1, \ldots, r-1\}, \\ -a_{j'/2} & \text{if } (j', j) \in \{(2i, 2i+1) : i = 0, 1, \ldots, r-1\}, \\ 0 & \text{otherwise.} \end{cases} \tag{30}$$

Furthermore, by applying [47, Corollary 5.16] we can conclude that L is an invertible matrix over R, and then by Lemma 2.1 and Lemma 2.2 we also know that $\Theta(M_C) = \Theta(M_A) = 2r$, where M_A and M_C are the submodules of R^k generated by the columns of A and C respectively. This decomposition $A = LBL^T$ is known as the *alternating Smith Normal form* (ASNF), for example see [24, Theorem 18] where it is done assuming R is a PID. In this paper, we are interested in the specific cases when R is either $\mathbb{Z}$ or $\mathbb{Z}_d$. In these cases, one even has an explicit formula for the non-zero entries of B in the ASNF, in terms of the minors of the matrix C. In the next subsection, Lemma A.2 summarizes all these facts when $R = \mathbb{Z}$, and provides a complete proof that has the added advantage of specifying an explicit algorithm to compute the ASNF. We also discuss how Lemma A.2 easily generalizes to the case $R = \mathbb{Z}_d$, leading to Lemma 5.3, which we already encountered before.

A.2 An algorithm for the ASNF

For the following lemma, for $a, b \in \mathbb{Z}$ we define the greatest common divisor of a and b to be the largest positive integer that divides both a and b over $\mathbb{Z}$, and denote it $\gcd(a, b)$. We also define $\gcd(0, 0) = 0$, and $\gcd(a) = a$ for any $a \in \mathbb{Z}$.

Lemma A.2. *Suppose $A \in \mathbb{Z}^{k\times k}$ is an alternating matrix. Then, there are matrices $L, B \in \mathbb{Z}^{k\times k}$, where B is alternating and has at most one non-zero entry per row and column, and L is invertible, such that $A = LBL^T$. Moreover, we may further arrange B so that it is non-zero only in the top-left $2r \times 2r$ block which has the form $\bigoplus_{i=1}^{r} \left(\begin{smallmatrix} 0 & \beta_i \\ -\beta_i & 0 \end{smallmatrix}\right)$ for integer $r = \Theta(M_A)/2$, where M_A is the $\mathbb{Z}$-submodule generated by the columns of A, and each $\beta_i \in \mathbb{Z}$ non-zero, satisfying $\beta_i \mid \beta_{i+1}$ for all $i < r$. Moreover, for all $i = 1, 2, \ldots, r$, we have $|\beta_i| = d_{2i}/d_{2i-1} = d_{2i-1}/d_{2i-2}$, $\beta_i^2 = d_{2i}/d_{2i-2}$, and for all $i > 2r$ we have $d_i = 0$, where d_j is the greatest common divisor of all $j \times j$ minors of A (with $d_0 := 1$), which implies that the β_i are unique up to choice of $\pm$ sign.*

Remark. *In the context of the lemma above, we have a chain of divisibilities $d_0 \mid d_1 \mid \cdots \mid d_k$. This is because for any $j \geq 0$, any $(j+1) \times (j+1)$ minor of A is divisible by d_j, and thus d_j must also divide d_{j+1}. This means that if there is any $j' \leq k$ such that $d_{j'} = 0$, then it automatically implies that $d_j = 0$ for all $j > j'$. Suppose such a j' exists and moreover assume that $d_j \neq 0$ for all $j < j'$ (for example if k is odd, then $d_k = \det(A) = 0$ by property of alternating matrices). In other words, suppose j' is the smallest integer such that all $j' \times j'$ minors of A are zero. Then we can see from the above lemma that $j' = 2r+1$, and hence must be odd. Thus we also conclude that $j' = \Theta(M_A) + 1$.*

This lemma appears in more or less general forms elsewhere. As alluded in the previous section, it is a consequence of Theorem A.1 from [46]. Part of this lemma is also Theorem 18 in [24] applied to the ring of integers. However, we go further and include the formulas for β_i. Finally, the analogous lemma for matrices over $\mathbb{F}_2$ appears as Lemma B.1 in [22]. After the proof, we discuss two variations on the theme: Lemma 5.3, which replaces the ring $\mathbb{Z}$ with $\mathbb{Z}_d$, and a generalization to matrices over any principal ideal ring (of which $\mathbb{Z}_d$ is just one).

Proof of Lemma A.2. The construction of B and L is iterative. We let $B_0 = A$. Entries of a matrix that are the only non-zero entry in both their row and column we call "pivotal" entries. Rows and columns containing those entries are also called pivotal. Suppose B_{i-1} is an antisymmetric matrix with zero diagonal and each of its first $i-1$ rows and columns are either pivotal or all zero. Because it is antisymmetric with zero diagonal, the total number of pivotal rows (counting also those outside the first $i-1$ rows, if any) is even and equals the number of pivotal columns. See Fig. 1 for a depiction of the structure of B_{i-1}.

Starting with B_{i-1}, we will construct a matrix $B_i = L_i B_{i-1} L_i^T$ with the same properties for one additional row and column, and L_i is invertible. The sets of pivotal rows and columns of B_i will contain the sets of pivotal rows and columns of B_{i-1}. For notational simplicity we denote $B_{i-1} = B'$ below.

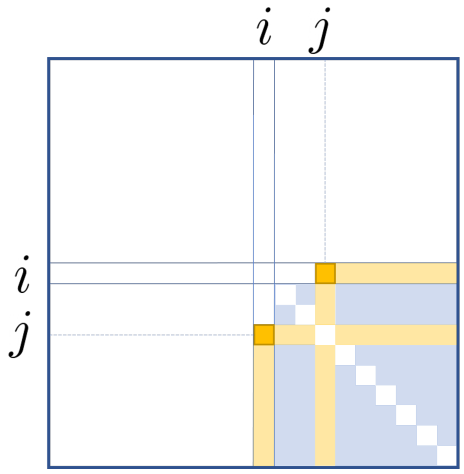

Figure 1: A schematic of the matrix B_{i-1} at the start of iteration i of the alternating Smith normal form construction. Non-shaded regions are guaranteed to be zero except for at most $2(i-1)$ existing pivotal entries. Dark yellow squares indicate the entries that iteration i is making into pivotal entries. Yellow-shaded regions contain entries that must be zeroed-out to complete iteration i.

If row i of B' is already pivotal or all zero, then we can just set $B_i = B'$ and L_i to the identity matrix. Otherwise, there exists a smallest j such that $B'_{ji} = -B'_{ij} \neq 0$. Note that $j > i$; otherwise, row i must already be pivotal.

We now alternate two subroutines until a termination condition is met. The first subroutine zeroes out entries in the column i other than B'_{ji}, and the second zeroes out entries in the column j other than B'_{ij}. Updates are made to B' throughout each subroutine, but it always remains antisymmetric with zero diagonal. Thus, the subroutines also simultaneously zero out entries in row i and row j, respectively.

The fundamental task in both subroutines is the same — we have two non-zero entries a and b in the same column, and wish to take linear combinations of their rows so that a is replaced by $\gcd(a,b) > 0$ and b is replaced by 0. By Bezout's identity, there exist integers x, y such that $ax + by = \gcd(a,b)$ (and these can be found by performing the extended version of Euclid's GCD algorithm). Let $z = a/\gcd(a,b)$ and $w = b/\gcd(a,b)$. Then,

$$\begin{pmatrix} x & y \\ -w & z \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \gcd(a,b) \\ 0 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} x & y \\ -w & z \end{pmatrix}^{-1} = \begin{pmatrix} z & -y \\ w & x \end{pmatrix}. \tag{31}$$

Let $K \in \mathbb{Z}^{k\times k}$ be the identity matrix except for the entries $K_{\alpha\alpha} = x$, $K_{\alpha\beta} = y$, $K_{\beta\alpha} = -w$, $K_{\beta\beta} = z$, where α is the row index of a and β the row index of b. Then $B' \leftarrow KB'K^T$ has replaced entry a with $\gcd(a,b)$ and b with 0, while also preserving the antisymmetry of B' (since we likewise took linear combinations of columns by acting via K^T on the right).

The first subroutine sequentially zeros out all entries $B'_{j'i} \neq 0$ for $j' \neq j$, leaving a non-zero entry at just position (j,i) in column i. The second subroutine sequentially zeros out all entries $B'_{i'j}$ for $i' \neq i$, leaving a non-zero entry at just position (i,j) in column j. Note, the second subroutine may cause entries $B'_{j'i}$ to become non-zero again, undoing some of the work of the first subroutine, and likewise the first subroutine may cause entries $B_{i'j}$ to become non-zero. For this reason we must alternate application of subroutines until rows i and j are both pivotal. This termination is guaranteed – the absolute values of entries at (i,j) and (j,i) are non-increasing because $\gcd(a,b)$ divides $|a|$. Moreover, if $\gcd(a,b) = |a|$ (and this holds for all entries b we are trying to zero out in this iteration), then we may choose $y = 0$ in Bezout's identity and the row containing a remains unchanged (in absolute value) – that is, the work of the previously applied subroutine is not undone, the rows i and j have been made pivotal, and we have constructed B_i and L_i (the latter a product of the appropriate sequence of matrices K above).

After k iterations, we have obtained invertible matrices $L_1, \ldots, L_k$ such that

$$(L_k L_{k-1} \ldots L_1) A (L_k L_{k-1} \ldots L_1)^T = B_k \tag{32}$$

where B_k is antisymmetric with zero diagonal and at most one non-zero entry per row and column. There is a permutation matrix P such that $B_{k+1} = P B_k P^T = \bigoplus_{i=1}^{r} \left(\begin{smallmatrix} 0 & \gamma_i \\ -\gamma_i & 0 \end{smallmatrix}\right)$ for some integers r and γ_i. We are still lacking the divisibility condition on entries of B, which we rectify next.

Consider the matrix of integers

$$T = \begin{pmatrix} 0 & a & 0 & 0 \\ -a & 0 & 0 & 0 \\ 0 & 0 & 0 & b \\ 0 & 0 & -b & 0 \end{pmatrix}. \tag{33}$$

Suppose we have Bezout's identity $ax + by = \gcd(a, b)$. We claim there exists invertible matrix H such that

$$HTH^T = \begin{pmatrix} 0 & \gcd(a,b) & 0 & 0 \\ -\gcd(a,b) & 0 & 0 & 0 \\ 0 & 0 & 0 & \operatorname{lcm}(a,b) \\ 0 & 0 & -\operatorname{lcm}(a,b) & 0 \end{pmatrix}. \tag{34}$$

This is exemplified by the following sequence of row and column operations:

$$T \to \begin{pmatrix} 0 & a & 0 & b \\ -a & 0 & 0 & 0 \\ 0 & 0 & 0 & b \\ -b & 0 & -b & 0 \end{pmatrix} \to \begin{pmatrix} 0 & \gcd(a,b) & 0 & 0 \\ -\gcd(a,b) & 0 & -by & 0 \\ 0 & by & 0 & \operatorname{lcm}(a,b) \\ 0 & 0 & -\operatorname{lcm}(a,b) & 0 \end{pmatrix} \to HTH^T. \tag{35}$$

The first step adds row 3 to row 1 (and likewise on columns). The second uses Eq. (31) to zero out entries $(1,4)$ and $(4,1)$. The third step zeros entries at $(2,3)$ and $(3,2)$ by subtracting $by/\gcd(a,b)$ times row 1 from row 3 (and likewise on columns).

Apply this procedure iteratively to pairs of 2×2 blocks of B_{k+1}. In iteration i, apply it sequentially $r - i$ times to block i paired with each of blocks $j = i+1, \ldots, r$. After iteration i, the non-zero entry of the i^{th} block divides all entries in blocks $i+1, \ldots, r$. After iteration $r-1$, we have an invertible H so that $B_{k+2} = H B_{k+1} H^T = \bigoplus_{i=1}^{r} \left(\begin{smallmatrix} 0 & \beta_i \\ -\beta_i & 0 \end{smallmatrix}\right)$ where $\beta_i \mid \beta_{i+1}$.

Set $L = (HPL_k L_{k-1} \ldots L_1)^{-1} = L_1^{-1} L_2^{-1} \ldots L_k^{-1} P^{-1} H^{-1}$ and $B = B_{k+2}$, so that $A = LBL^T$. Note that by combination of Lemmas 2.1 and 2.2, we deduce that $r = \Theta(M_A)/2$.

To show that $|\beta_i| = d_{2i}/d_{2i-1} = d_{2i-1}/d_{2i-2}$ and $\beta_i^2 = d_{2i}/d_{2i-2}$, we note that invertible row and column operations do not change the greatest common divisor of matrix minors (see Corollary 4.8 in [47] or the proof of Proposition 8.1 in [69]). Thus, since the formulas for β_i are easily verified using matrix minors of $B = \bigoplus_{i=1}^{r} \left(\begin{smallmatrix} 0 & \beta_i \\ -\beta_i & 0 \end{smallmatrix}\right)$, they hold also for A.

We summarize the construction in this proof in Algorithm 3. □

To start the commentary on Lemma A.2, let us apply it more specifically to the ring $\mathbb{Z}_d$ that is relevant for d-dimensional qudits. Suppose we have an alternating matrix $\bar{A} \in \mathbb{Z}_d^{k \times k}$. Treating $\bar{A}$ as an element of $\mathbb{Z}^{k \times k}$, we use Lemma A.2 to find matrices $\bar{L}, \bar{B} \in \mathbb{Z}^{k \times k}$ so that $\bar{A} = \bar{L}\bar{B}\bar{L}^T$. We note that invertibility of $\bar{L}$ implies that $\det(\bar{L}) = \pm 1$ over integers, since ± 1 are the only units in $\mathbb{Z}$. We may reduce this expression modulo d to obtain the equality $A = LBL^T$, where A, B, and L are obtained by reducing $\bar{A}, \bar{B}$, and $\bar{L}$ modulo d, respectively. Note that $\bar{A} = A$, and L is an invertible matrix in $\mathbb{Z}_d^{k \times k}$ since $\det(L) = \det(\bar{L}) \mod d = \pm 1$. The matrix B is alternating in $\mathbb{Z}_d^{k \times k}$ since $\bar{B}$ was alternating in $\mathbb{Z}^{k \times k}$. Moreover, any zero entry in $\bar{B}$ remains zero in B, while some non-zero entries of $\bar{B}$ may become zero in B, due to the modulo d operation. The divisibility and uniqueness parts of Lemma A.2 are also modified in accordance with the following observations on divisibility in $\mathbb{Z}_d$.

Algorithm 3 Given alternating matrix $A \in \mathbb{Z}^{k\times k}$, find $L, B \in \mathbb{Z}^{k\times k}$ such that $A = LBL^T$ and B, L have the properties stated in Lemma A.2.

procedure ALTERNATING_SMITH_NORMAL_FORM(A)
  $B \leftarrow A$
  $L \leftarrow \mathsf{IdentityMatrix}(k \times k)$
  for $1 \le i \le k$ **do**
    if Row i of B is all 0 **then**
      Continue
    $j \leftarrow$ the smallest index such that $B_{ij} \neq 0$
    while $S = \{h > j : B_{ih} \neq 0\}$ is not empty or $T = \{l > i : B_{lj} \neq 0\}$ is not empty **do**
      for $h \in S$ **do**
        $x, y, g = \mathsf{ExtendedGCD}(B_{ij}, B_{ih})$
        $K \leftarrow \mathsf{IdentityMatrix}(k \times k)$
        $K_{jj} \leftarrow x$, $K_{jh} \leftarrow y$, $K_{hj} \leftarrow -B_{ih}/g$, $K_{hh} \leftarrow B_{ij}/g$
        $B \leftarrow KBK^T$
        $L \leftarrow LK^{-1}$
      for $l \in T$ **do**
        $x, y, g = \mathsf{ExtendedGCD}(B_{ij}, B_{lj})$
        $K \leftarrow \mathsf{IdentityMatrix}(k \times k)$
        $K_{ii} \leftarrow x$, $K_{il} \leftarrow y$, $K_{li} \leftarrow -B_{lj}/g$, $K_{ll} \leftarrow B_{ij}/g$
        $B \leftarrow KBK^T$
        $L \leftarrow LK^{-1}$
    $P \leftarrow$ permutation matrix so that PBP^T is block diagonal with r non-zero 2×2 blocks
    $B = PBP^T$
    $L = LP^{-1}$
    for $1 \le i \le r - 1$ **do**
      for $i + 1 \le j \le r$ **do**
        $H \leftarrow$ the matrix enacting the transformation in Eq. (35) on blocks i and j
        $B = HBH^T$
        $L = LH^{-1}$
  Return B, L
  # Note: $\mathsf{ExtendedGCD}(a, b)$ finds $x, y \in \mathbb{Z}$ and $0 < g \in \mathbb{Z}$ such that $xa+yb = g = \gcd(a, b)$. Also, if $\gcd(a, b) = |a|$, $y = 0$.

A greatest common divisor $\gcd(a, b)$ for $a, b \in \mathbb{Z}_d$ is any element of $\mathbb{Z}_d$ that divides a and b and has the property that if $c \in \mathbb{Z}_d$ divides both a and b, then c divides $\gcd(a, b)$. The group of units U_d is the multiplicative group consisting of elements of $\mathbb{Z}_d$ that have multiplicative inverses in $\mathbb{Z}_d$. Elements of U_d are the integers less than d that are relatively prime to d. Multiplication by units does not affect divisibility, i.e. if $a \mid b$, then $ua \mid b$ and $a \mid ub$ for all $u \in U_d$. Therefore, the value of $\gcd(a, b)$ is not unique, and if g and h are both greatest common divisors of a and b, then there is a unit $u \in U_d$ such that $g = uh$, a fact that holds over commutative principal ideal rings [70, 71]. These observations lead to the ASNF over the ring $\mathbb{Z}_d$ as stated in Lemma 5.3 in the main text.

Finally, we note that Lemma A.2 can be generalized to matrices over commutative principal ideal rings. A commutative principal ideal ring R is defined by two properties

1. Noetherian property (e.g. Chapter 10 of [46]): all ideals are finitely generated, or, equivalently, any ascending chain of ideals $I_0 \subseteq I_1 \subseteq I_2 \subseteq \dots$ terminates with $I_n = I_m$ for all $m \geq n$.
2. Hermite property [70]: for any two elements $a, b \in R$, there exists $c \in R$ and an invertible matrix $T \in R^{2\times 2}$ such that $(a\ b)T = (c\ 0)$.

These properties generalize key steps in the proof of Lemma A.2. The Hermite property replaces Eq. (31), which is used to zero out matrix entries. Note that it also implies $(a\ b) = (c\ 0)T^{-1}$ and therefore $c \mid a$ and $c \mid b$. The Noetherian property is used to argue that the repeated application of the zeroing subroutines eventually terminates with a new pivotal row and column. Because $c \mid a$, the ideal generated by c contains the ideal generated by a. Since the Noetherian property means the chain of ideals must terminate, the algorithm must also terminate. These modifications to the proof of Lemma A.2 lead to a more general theorem.

Theorem A.3. *Let R be a commutative principal ideal ring, and suppose $A \in R^{k\times k}$ is an alternating matrix. Then, there are matrices $L, B \in R^{k\times k}$, where B is alternating and has at most one non-zero entry per row and column, and L is invertible, such that $A = LBL^T$. We may further arrange B so that it is non-zero only in the top-left $2r \times 2r$ block which has the form $\bigoplus_{i=1}^{r} \left(\begin{smallmatrix} 0 & \beta_i \\ -\beta_i & 0 \end{smallmatrix}\right)$ for integers $r = \Theta(M_A)/2$, where M_A is the R-submodule generated by the columns of A, and each $\beta_i \in R$ non-zero, satisfying $\beta_i \mid \beta_{i+1}$ for all $i < r$. Also, β_i is uniquely determined up to multiplication by a unit and satisfies the formula $\beta_i d_{2i-1} = d_{2i}$ (or, alternatively, the formula $\beta_i d_{2i-2} = d_{2i-1}$), where d_j is a greatest common divisor in R of all $j \times j$ minors of A (and $d_0 := 1$).*

B Technical lemmas and proofs

B.1 Proofs for Section 2

In this section, we give proofs for some of the preliminary lemmas involving modules over commutative rings.

Lemma 2.1. Suppose $C \in R^{k\times t}$ and $A \in R^{k\times k}$, $B \in R^{t\times t}$ are invertible matrices. Let $\bar{C} := ACB$. The minimal number of generators of the submodules M_C and $M_{\bar{C}}$ of R^k are equal.

Proof. Let $\hat{C} := CB$, let its columns be $\hat{c}_0, \hat{c}_1, \dots, \hat{c}_{t-1}$, and let $M_{\hat{C}}$ be the submodule of R^k generated by the columns of $\hat{C}$. For every i, each column $\hat{c}_i \in M_C$ is a linear combination of the columns of C, from which we conclude that $M_{\hat{C}} \subseteq M_C$. By invertibility of B, we also have $C = \hat{C}B^{-1}$, and by the same argument we conclude that $M_C \subseteq M_{\hat{C}}$. Thus $M_C = M_{\hat{C}}$, and so we have $\Theta(M_C) = \Theta(M_{\hat{C}})$. It remains to show that $\Theta(M_{\bar{C}}) = \Theta(M_{\hat{C}})$, where $\bar{C} = A\hat{C}$.

Let $D \in R^{k\times t'}$, with $t' \leq t$, be a matrix whose columns form a minimal generating set of $M_{\hat{C}}$. Then each $\hat{c}_i$ can be expressed as a linear combination of the columns of D; so we can write $\hat{C} = DE$, for some $E \in R^{t'\times t}$. Letting $\bar{A} := AD \in R^{k\times t'}$, we thus have $\bar{C} = \bar{A}E$. This shows that $\Theta(M_{\bar{C}}) \leq t = \Theta(M_{\hat{C}})$. Again by invertibility of A, we also have $\hat{C} = A^{-1}\bar{C}$, and now repeating the argument gives $\Theta(M_{\hat{C}}) \leq \Theta(M_{\bar{C}})$. □

Lemma 2.2. Let $C \in R^{k\times t}$ be a matrix such that each row and column has at most one non-zero element, and suppose one of those non-zero elements is divisible by all the others. Then the minimal number of generators of the R-module M_C, generated by the columns of C, is equal to the number of non-zero elements of the matrix C.

Proof. The lemma is clearly true if $C = 0$; so assume that C has at least one non-zero element. Without loss of generality, we may assume that C is a diagonal matrix so that $C_{00}, C_{11}, \ldots, C_{r-1,r-1}$ all divide $C_{rr} \neq 0$ (which implies $C_{ii} \neq 0$ for all $i = 0, 1, \ldots, r$), where $r+1$ is the number of non-zero elements of C. If C is not of this form, then one can permute the rows and columns of C to bring it to this form, and Lemma 2.1 ensures that the minimal number of generators remain the same. Notice that if $x \in M_C$, then for $i = 0, 1, \ldots, r$ we have $x_i = a_i C_{ii}$ for some $a_i \in R$, and $x_i = 0$ for all $i > r$.

Define the ideals $N_i := \{x \in R : xC_{ii} = 0\}$ for each $i = 0, 1, \ldots, r$, and note that by the divisibility condition we have $N_0, N_1, \ldots, N_{r-1} \subseteq N_r \neq R$. Let N be a maximal ideal of R containing N_r. Then R/N is a field [46, Chapter 2], and let $f : R \to R/N$ be the corresponding quotient map (which is a ring homomorphism) taking an element $y \in R$ to its coset in R/N. Treating $(R/N)^{r+1}$ as a R-module, we now define a R-module homomorphism $\kappa : M_C \to (R/N)^{r+1}$ as follows: if $x \in M_C$, then $(\kappa(x))_i = f(a_i)$ for every $i = 0, 1, \ldots, r$, where $x_i = a_i C_{ii}$ (one can check that κ is well-defined, i.e. it does not depend on the choice of a_i as $N_i \subseteq N$, and that it is a module homomorphism). The map κ is also clearly surjective. Thus if M_C has a generating set S of size $|S| < r+1$, then surjectivity of κ implies that a generating set for $(R/N)^{r+1}$ is $\{\kappa(x) : x \in S\}$. But then $\{\kappa(x) : x \in S\}$ is also a generating set of the R/N-module $(R/N)^{r+1}$, which is a vector space of dimension $r+1$ as R/N is a field, and thus it cannot have less than $r+1$ generators. □

B.2 Proofs for Section 4

Recall that in Section 4, we studied a graph G with a vertex for each single-qudit Pauli in $\mathcal{P}_1$ and edges between pairs that do not commute. A subset of vertices W contains those corresponding to Paulis $X^a Z^b$ with $\gcd(a, b, d) = 1$.

Lemma 4.2. The graph $G = (V, E)$ has the following properties.

(i) If $C \subseteq V$ is a clique in G, then there is a clique $C' \subseteq W$ of the same size, $|C| = |C'|$.

(ii) If $v_0, v_2 \in V$, $v_1 \in W$, $(v_0, v_1) \notin E$, and $(v_1, v_2) \notin E$, then $(v_0, v_2) \notin E$.

(iii) If (a, b) and (s, t) are in W, then $((a, b), (s, t)) \notin E$ if and only if there exists $u \in \mathbb{Z}$ with $\gcd(u, d) = 1$ such that $(a, b) = (us, ut)$, the right hand side evaluated modulo d.

Proof. To prove (i), we show that if $C \subseteq V$ is a clique, then for any $v = (a, b) \in C$, $C' = (C \setminus \{v\}) \cup \{v'\}$, where $v' = v/\gcd(a, b, d) \in W$, is also a clique. Suppose for contradiction, there is some $(s, t) \in C$ that is connected by an edge to v but not to v'. Then, for $g = \gcd(a, b, d)$, we have $\det \left(\begin{smallmatrix} a/g & b/g \\ s & t \end{smallmatrix}\right) \equiv 0 \pmod{d}$ which implies $\det \left(\begin{smallmatrix} a & b \\ s & t \end{smallmatrix}\right) \equiv 0 \pmod{d}$, contradicting that (s, t) is connected to v.

To prove (ii), let $v_i := (a_i, b_i)$ for $i = 0, 1, 2$. Note that by Bezout's identity, there are integers x, y, z such that $xa_1 + yb_1 + zd = \gcd(a_1, b_1, d) = 1$. Let

$$A := \begin{pmatrix} a_0 & b_0 & xa_0 + yb_0 + zd \\ a_1 & b_1 & 1 \\ a_2 & b_2 & xa_2 + yb_2 + zd \end{pmatrix} \in \mathbb{Z}^{3\times 3}. \tag{36}$$

Due to the last column of A being a linear combination of the first two and $(d, d, d)^T$, we have $\det(A) \equiv 0 \pmod{d}$. Also, for $c_i = xa_i + yb_i + zd$, the cofactor expansion gives $\det(A) = c_0 \det \left(\begin{smallmatrix} a_1 & b_1 \\ a_2 & b_2 \end{smallmatrix}\right) - \det \left(\begin{smallmatrix} a_0 & b_0 \\ a_2 & b_2 \end{smallmatrix}\right) + c_2 \det \left(\begin{smallmatrix} a_0 & b_0 \\ a_1 & b_1 \end{smallmatrix}\right) \equiv -\det \left(\begin{smallmatrix} a_0 & b_0 \\ a_2 & b_2 \end{smallmatrix}\right) \pmod{d}$. Thus we have $\det \left(\begin{smallmatrix} a_0 & b_0 \\ a_2 & b_2 \end{smallmatrix}\right) \equiv 0 \pmod{d}$, proving $(v_0, v_2) \notin E$.

The reverse direction of part (iii) is a simple calculation. The forward direction is more involved. Apply the Smith normal form (Theorem 2.3) to obtain $A := \left(\begin{smallmatrix} a & b \\ s & t \end{smallmatrix}\right) = S^{-1}BT$ for invertible matrices $S, T \in \mathbb{Z}_d^{2\times 2}$ and diagonal $B \in \mathbb{Z}_d^{2\times 2}$ with $B_{00} = \gcd(a, b, s, t) \bmod d = \gcd(a, b, s, t)$ and $B_{11} = (\det(A)/B_{00}) \bmod d$. The formulas for B_{00} and B_{11} follow from [72, Theorem 2.4], even though $\mathbb{Z}_d$ is not a unique factorization domain (as required to apply the theorem), but one may simply compute the Smith normal form of A over $\mathbb{Z}$, which is a unique factorization domain, and then reduce the resulting matrices modulo d, leading to the formulas for B_{00} and B_{11}. Since $((a, b), (s, t)) \notin E$, we know $\det(A) \equiv 0 \pmod d$, and because we also have $\gcd(a, b, d) = 1$, we get $B_{11} = 0$. Rearrange to obtain $SA = BT$ and note that the last row of BT is $(0, 0)$. This implies

$$S_{10}(a, b) + S_{11}(s, t) = (0, 0), \tag{37}$$

evaluated over $\mathbb{Z}_d$. This equation means $\gcd(S_{10}, d)$ divides $S_{11}s$ and $S_{11}t$ over integers. Because $\gcd(s, t, d) = 1$, it must then be that $\gcd(S_{10}, d)$ also divides S_{11} over integers. Therefore, $\gcd(S_{10}, d)$ divides $\det(S)$ over integers, and hence also over $\mathbb{Z}_d$. Since S is invertible, $\det(S)$ (modulo d) is a unit of the ring $\mathbb{Z}_d$, so its only divisors are other units. This means $\gcd(S_{10}, d) = 1$. Likewise, we can argue $\gcd(S_{11}, d) = 1$. So S_{10}, S_{11}, and $S_{11}(S_{10})^{-1}$ are all units in $\mathbb{Z}_d$. Multiply Eq. (37) by S_{10}^{-1} and rearrange to finish the proof. □

The other lemma we prove here involves the arithmetic function $H(d) = |W|/d$. We argued in Section 4 that it is equal to

$$H(d) = \sum_{a\in\mathbb{Z}_d} \prod_{p|\gcd(a,d)} \left(1 - \frac{1}{p}\right). \tag{38}$$

Lemma 4.3. If $d_0, d_1 \in \mathbb{Z}$ are relatively prime, then $H(d_0d_1) = H(d_0)H(d_1)$.

Proof. By Bezout's identity, there are integers x_0, x_1 such that $x_0d_0 + x_1d_1 = 1$. Note that x_0 is an inverse of d_0 modulo d_1, so $\gcd(x_0, d_1) = 1$. Likewise $\gcd(x_1, d_0) = 1$.

The Chinese remainder theorem implies there is an isomorphism between $\mathbb{Z}_{d_0d_1}$ and $\mathbb{Z}_{d_0} \times \mathbb{Z}_{d_1}$. Explicitly, it says that $a \equiv a_0 \pmod{d_0}$ and $a \equiv a_1 \pmod{d_1}$, if and only if $a \equiv a_0x_1d_1 + a_1x_0d_0 \pmod{d_0d_1}$. This means that $\gcd(a, d_0d_1) = \gcd(a_0x_1d_1 + a_1x_0d_0, d_0d_1)$. We also have $\gcd(x, d_0d_1) = \gcd(x, d_0)\gcd(x, d_1)$ for any integer x, as d_0, d_1 are relatively prime.

Putting these facts together, we complete the proof:

$$\begin{aligned} H(d_0d_1) &= \sum_{a\in\mathbb{Z}_{d_0d_1}} \prod_{p|\gcd(a,d_0d_1)} (1 - 1/p) && (39)\\ &= \sum_{a_0\in\mathbb{Z}_{d_0}} \sum_{a_1\in\mathbb{Z}_{d_1}} \prod_{p|\,\gcd(a_0x_1d_1+a_1x_0d_0,d_0d_1)} (1 - 1/p) && (40)\\ &= \left(\sum_{a_0\in\mathbb{Z}_{d_0}} \prod_{p|\,\gcd(a_0,d_0)} (1 - 1/p)\right)\left(\sum_{a_1\in\mathbb{Z}_{d_1}} \prod_{p|\gcd(a_1,d_1)} (1 - 1/p)\right) = H(d_0)H(d_1). && (41) \end{aligned}$$

□

B.3 Proofs for Section 5

Lemma B.1. *Let $d = p_0^{\alpha_0}p_1^{\alpha_1}\dots p_{m-1}^{\alpha_{m-1}}$ be the prime factorization of an integer $d \geq 2$, where $p_0, \dots, p_{m-1}$ are primes. Let $d', d'' \geq 2$ be two integers. Define $J := \{j \in \{0, \dots, m-1\} : d' \not\equiv 0 \pmod{p_j^{\alpha_j}}\}$, and suppose that $d'' \not\equiv 0 \pmod{p_j}$ for every $j \in J$. Then we have the set equality*

$$\{x \bmod d : x \in d'\mathbb{Z}\} = \{x \bmod d : x \in d'd''\mathbb{Z}\}.$$

Proof. If J is empty, then d' is divisible by d, so the statement is trivial. Thus assume $J \neq \emptyset$, so d' is not divisible by d. It is clear that $\{x \bmod d : x \in d'd''\mathbb{Z}\} \subseteq \{x \bmod d : x \in d'\mathbb{Z}\}$. Let

ℓ be the smallest non-negative integer such that $\ell d' \equiv 0 \pmod{d}$. Then ℓ must be of the form $\ell = \prod_{j\in J} p_j^{\beta_j}$, where $1 \leq \beta_j \leq \alpha_j$ for each j. Note that ℓ is an upper bound on the size of the set $\{x \bmod d : x \in d'\mathbb{Z}\}$. Now consider the multiset of elements $K := \{td'd'' \mod d : t = 0, \ldots, \ell-1\}$. We will show that all elements of K are distinct, which will prove the lemma as it implies that ℓ is a lower bound on the size of the set $\{x \bmod d : x \in d'd''\mathbb{Z}\}$.

Suppose this is not the case, and there exist distinct $t, t' \in \{0, \ldots, \ell-1\}$, with $t > t'$, such that $td'd'' \equiv t'd'd'' \pmod{d}$, or equivalently $(t-t')d'd'' \equiv 0 \pmod{d}$. By assumptions on J and d'', we then conclude that $(t-t')d' \equiv 0 \pmod{d}$. Since $0 < t - t' \leq \ell - 1$, this implies that ℓ is not the smallest non-negative integer satisfying $\ell d' \equiv 0 \pmod{d}$, which is a contradiction. □

The following corollary now follows immediately from the above lemma by setting $d' = l_J$ and $d'' = l_J^{r-1}$, for $r \geq 1$.

Corollary B.2. *Let $d = p_0^{\alpha_0} p_1^{\alpha_1} \ldots p_{m-1}^{\alpha_{m-1}}$ be the prime factorization of an integer $d \geq 2$, where $p_0, \ldots, p_{m-1}$ are primes. Let $J \subseteq \{0, \ldots, m-1\}$, and define $\ell_J := \prod_{j\in J} p_j^{\alpha_j}$. Then for every integer $r \geq 1$, we have $\{x \bmod d : x \in \ell_J^r \mathbb{Z}\} = \{x \bmod d : x \in \ell_J \mathbb{Z}\} = \{k\ell_J : 0 \leq k < d/\ell_J\}$.*

B.4 Proofs for Section 6

We start by giving a full proof of the following theorem on transforming generating sets of Pauli groups.

Theorem 6.3. (Equivalent generating sets) Suppose $S := \{q_0, q_1, \ldots q_{k-1}\}$ is an ordered multiset of elements of $\overline{\mathcal{P}}_n$. Let $A \in \mathbb{Z}_d^{k\times k}$ be an invertible matrix, and consider the ordered multiset $T := \{q'_0, q'_1, \ldots, q'_{k-1}\} \subseteq \overline{\mathcal{P}}_n$, where we define $q'_i := \prod_{j=0}^{k-1} q_j^{A_{ij}}$. Then $\langle S\rangle = \langle T\rangle$.

Proof. It is clear that $\langle T\rangle \subseteq \langle S\rangle$, so we only need to show that $\langle S\rangle \subseteq \langle T\rangle$. Since A is invertible, there exists a matrix $B \in \mathbb{Z}_d^{k\times k}$ such that $BA = I$. Now define the multiset $U := \{r_0, r_1, \ldots, r_{k-1}\} \subseteq \overline{\mathcal{P}}_n$, where $r_\ell := \prod_{j=0}^{k-1} (q'_j)^{B_{\ell j}}$ for each $\ell = 0, 1, \ldots, k-1$. Then clearly we have $\langle U\rangle \subseteq \langle T\rangle$, and we will show that $\langle S\rangle \subseteq \langle U\rangle$. Corresponding to the sets S and U, we also define the sets J_S and J_U, according to the notation above.

Now fix any ℓ, and note that $r_\ell = \prod_{i=0}^{k-1} \left(\prod_{j=0}^{k-1} q_j^{A_{ij}}\right)^{B_{\ell i}}$, and this product can be rearranged using group commutators as

$$r_\ell = \lambda_\ell \prod_{j=0}^{k-1} q_j^{\sum_{i=0}^{k-1} B_{\ell i} A_{ij}}, \quad \lambda_\ell \in \langle J_S\rangle. \tag{42}$$

Notice that the fact $BA = I$ over the ring $\mathbb{Z}_d$ implies that $\sum_{i=0}^{k-1} B_{\ell i} A_{ij} \equiv 1 \pmod{d}$ if $\ell = j$, and otherwise $\sum_{i=0}^{k-1} B_{\ell i} A_{ij} \equiv 0 \pmod{d}$. Thus using Eq. (42) and the definition of J_S, we can conclude that $r_\ell = \overline{\lambda}_\ell q_\ell$ for some $\overline{\lambda}_\ell \in \langle J_S\rangle$, for every ℓ. We can now use Lemma 6.2(i),(iv) to deduce that $\langle J_S\rangle = \langle J_U\rangle \subseteq \langle U\rangle$. It follows that for every ℓ, we have $\left(\overline{\lambda}_\ell\right)^{-1} \in \langle U\rangle$, and this implies $q_\ell \in \langle U\rangle$. We thus conclude that $\langle S\rangle \subseteq \langle U\rangle$, completing the proof. □

The next lemma characterizes, for generating set $S \subseteq \overline{\mathcal{P}}_n$, the group of elements proportional to identity that S generates.

Lemma 6.4. Suppose $S := \{q_0, q_1, \ldots, q_{k-1}\} \subseteq \overline{\mathcal{P}}_n$ is an ordered multiset. Consider the matrix $\pi_2(S) \in \mathbb{Z}_d^{2n\times k}$, and let $\overline{K} \in \mathbb{Z}_d^{k\times \ell}$ be such that the columns of $\overline{K}$ is a generating set for $\ker(\pi_2(S))$. Then we have the following:

(i) Let $v \in \ker(\pi_2(S))$, and define $q' := \prod_{j=0}^{k-1} q_j^{v_j}$. Then $\pi_2(q') = 0$.

(ii) Define the set $N_S := \left\{\prod_{j=0}^{k-1} q_j^{v_j} : v \in \ker(\pi_2(S))\right\}$. Then $I_S = \langle N_S, J_S\rangle$.

(iii) Define the multiset $\overline{K}_S := \left\{ \prod_{j=0}^{k-1} q_j^{\overline{K}_{ji}} : i = 0, 1, \ldots, \ell - 1 \right\}$. Then $I_S = \langle \overline{K}_S, J_S \rangle$.

Proof. For (i) note that since $v \in \ker(\pi_2(S))$, we have $\pi_2(q') = \sum_{j=0}^{k-1} v_j \pi_2(q_j) = \pi_2(S)v = 0$, where all expressions are evaluated modulo d.

For part (ii) we only need to prove that $I_S \subseteq \langle N_S, J_S \rangle$, as the reverse containment is obvious by part (i) and definition of J_S. Take any $q \in I_S$. Then one can write $q = \prod_{j=0}^{\ell} h_j$, where each $h_j \in S$. Using group commutators we can rearrange this product to obtain $q = \lambda \prod_{j=0}^{k-1} q_j^{r_j}$, for $\lambda \in \langle J_S \rangle$, and non-negative integers r_j for each j. Using Lemma 6.1, we can further simplify this expression by reducing the powers r_j modulo d, to obtain $q = \lambda\lambda' \prod_{j=0}^{k-1} q_j^{s_j}$, for some $\lambda' \in \langle J_S \rangle$, and each $s_j \in \mathbb{Z}_d$. Define $s := (s_0, s_1, \ldots, s_{k-1}) \in \mathbb{Z}_d^k$. Now we know that $\pi_2(q) = 0$. This then implies that $\sum_{j=0}^{k-1} s_j \pi_2(q_j) = 0$, evaluated over $\mathbb{Z}_d$, or equivalently $s \in \ker(\pi_2(S))$. We can then conclude that $\prod_{j=0}^{k-1} q_j^{s_j} \in N_S$, and thus $q \in \langle N_S, J_S \rangle$.

We now prove part (iii). Each column of $\overline{K}_S$ is an element of $\ker(\pi_2(S))$. Thus it is clear that $\overline{K}_S \subseteq N_S$, and so $\langle \overline{K}_S, J_S \rangle \subseteq \langle N_S, J_S \rangle$. We will now show that $N_S \subseteq \langle \overline{K}_S, J_S \rangle$, which will prove $\langle N_S, J_S \rangle = \langle \overline{K}_S, J_S \rangle$, and then by (ii) we will obtain $\langle \overline{K}_S, J_S \rangle = I_S$. For this, take any $v \in \ker(\pi_2(S))$. Since the columns of $\overline{K}$ is a generating set for $\ker(\pi_2(S))$, we have $v = \overline{K}w$, for some $w \in \mathbb{Z}_d^{\ell}$. Then using the group commutators one obtains for $\lambda, \lambda', \lambda'' \in \langle J_S \rangle$,

$$\begin{aligned} \prod_{j=0}^{k-1} q_j^{v_j} &= \prod_{j=0}^{k-1} q_j^{(\sum_{i=0}^{\ell-1} \overline{K}_{ji} w_i) \bmod d} = \lambda \prod_{j=0}^{k-1} q_j^{\sum_{i=0}^{\ell-1} \overline{K}_{ji} w_i} \\ &= \lambda\lambda' \prod_{i=0}^{\ell-1} \left(\prod_{j=0}^{k-1} q_j^{\overline{K}_{ji} w_i} \right) = \lambda\lambda'\lambda'' \prod_{i=0}^{\ell-1} \left(\prod_{j=0}^{k-1} q_j^{\overline{K}_{ji}} \right)^{w_i}. \end{aligned} \tag{43}$$

Now since $\prod_{j=0}^{k-1} q_j^{\overline{K}_{ji}} \in \langle \overline{K}_S \rangle$ for every i, we deduce that $\prod_{j=0}^{k-1} q_j^{v_j} \in \langle \overline{K}_S, J_S \rangle$, and since v is arbitrary, we conclude that $N_S \subseteq \langle \overline{K}_S, J_S \rangle$. □

Next, we prove the corollary of Lemma 6.6.

Corollary 6.7. Let $A \in \mathbb{Z}_d^{k \times p}$ and $B \in \mathbb{Z}_d^{k \times q}$. If the columns of A and B generate the same submodule of $\mathbb{Z}_d^k$, then the number of invariant factors of A and B are equal.

Proof. Without loss of generality, assume that $p \leq q$. Define $\overline{A} = \begin{pmatrix} A & 0 \end{pmatrix} \in \mathbb{Z}_d^{k \times q}$, where we have added $q - p$ zero columns to A. Then the columns of $\overline{A}$ and B still generate the same submodule of $\mathbb{Z}_d^k$. Thus by Lemma 6.6 we may conclude that there exists an invertible matrix $C \in \mathbb{Z}_d^{q \times q}$ such that $\overline{A} = BC$. Next, suppose that A has the Smith normal form $AP = QD$ where P, Q are invertible matrices, and D is a diagonal matrix (all matrices are over the ring $\mathbb{Z}_d$). Let D have r non-zero elements. Then we can note that $\overline{A} \left(\begin{smallmatrix} P & 0 \\ 0 & I \end{smallmatrix} \right) = Q \begin{pmatrix} D & 0 \end{pmatrix}$, or equivalently $BC \left(\begin{smallmatrix} P & 0 \\ 0 & I \end{smallmatrix} \right) = Q \begin{pmatrix} D & 0 \end{pmatrix}$, and since $\left(\begin{smallmatrix} P & 0 \\ 0 & I \end{smallmatrix} \right)$ is invertible, we can conclude by the uniqueness part of the Smith normal form (Theorem 2.3) that $\begin{pmatrix} D & 0 \end{pmatrix}$ is the Smith normal form of B, and hence B also has r invariant factors. □

There is a generalization of Lemma 6.14, when the commutator $[\![p, q]\!]_d = c$, as appearing in the lemma, is not a unit in $\mathbb{Z}_d$. In this case, instead of exact equality, we get lower bounds:

Lemma B.3. *Let $p, q \in \overline{\mathcal{P}}_n$, $G \subseteq \overline{\mathcal{P}}_n$, $[\![p, q]\!]_d = c \in \mathbb{Z}_d \setminus \{0\}$, and assume the following:*

(a) Each element in G commutes with both p and q.

(b) a, b are the smallest positive integers such that p^a and q^b are equivalent to I up to phase.

(c) a_1, b_1 are the smallest positive integers such that $[\![p^{a_1}, q]\!]_d = [\![p, q^{b_1}]\!]_d = 0$.

Then the following statements are true:

(i) $a_1 = b_1$, a_1 *divides both* a *and* b, *and both* a, b *divide* d, *where division is over integers. Moreover, the order of* c *in* $\mathbb{Z}_d$ *is* a_1.

(ii) We have the lower bounds $\left|\langle G, p\rangle / I_{G\cup\{p\}}\right| \geq a_1 \left|\langle G\rangle / I_G\right|$, $\left|\langle G, q\rangle / I_{G\cup\{q\}}\right| \geq a_1 \left|\langle G\rangle / I_G\right|$, *and* $\left|\langle G, p, q\rangle / I_{G\cup\{p,q\}}\right| \geq a_1^2 \left|\langle G\rangle / I_G\right|$.

Proof. We first prove (i). It is clear that $a_1 \leq a$: if not, since p^a is equivalent to I up to phase factors, we have $[\![p^a, q]\!]_d = 0$, which contradicts (c). Moreover, if a_1 does not divide a, i.e. $a = ta_1 + s$ for some integers $t \geq 0$ and $0 < s < a_1$, then we also obtain a contradiction to (c) as $0 = [\![p^a, q]\!]_d = [\![p^s, q]\!]_d$. Similarly, we can argue that b_1 divides b. Also it is easy to see that a divides d: if this is not true, then $d = ta + s$ for some integers $t \geq 0$ and $0 < s < a$, which implies $p^d = (p^a)^t p^s$, and since both p^d and p^a are equivalent to I up to phase factors, we conclude that the same is true for p^s, which contradicts (b). Similarly we may also conclude that b divides d.

Next, suppose that the order of c in $\mathbb{Z}_d$ is $0 < \gamma < a_1$. Thus $c\gamma \bmod d = 0$, which implies $[\![p^\gamma, q]\!]_d = c\gamma \bmod d = 0$, contradicting (c). Thus the order of c in $\mathbb{Z}_d$ is at least a_1. On the other hand, $ca_1 \bmod d = [\![p^{a_1}, q]\!]_d = 0$, implying the order of c is at most a_1. From this we conclude that the order of c is a_1. By repeating the same argument using the fact that b_1 be the smallest positive integer such that $[\![p, q^{b_1}]\!]_d = 0$, we also get that the order of c equals b_1, and thus $a_1 = b_1$.

Now we prove (ii). All quotient groups appearing in the following argument will be treated as subgroups of $\overline{\mathcal{P}}_n / K$. Define the sets $P = \{p^j : j = 0, 1, \ldots, a-1\}$, and $Q = \{q^j : j = 0, 1, \ldots, b-1\}$. By (b) we know that the elements in the set P (or Q) are all distinct, even up to phase factors. This implies that the quotient groups $\langle p\rangle / I_{\{p\}}$ and $\langle q\rangle / I_{\{q\}}$ are in one-to-one correspondence with the sets P and Q respectively, where the identification is made by taking equivalence up to phase factors. Next, we make two claims, which we prove at the end:

Claim A: Let $H \subseteq \overline{\mathcal{P}}_n$ is a subgroup such that p commutes with every element of H. Define $H_Q := \{q^j \in Q : \omega^t q^j \in H, \text{ for some } t \in \mathbb{Z}_d\}$. Then $H_Q \subseteq \{q^j : 0 \leq j \leq b-1, j \equiv 0 \pmod{a_1}\}$.

Claim B: Let $H \subseteq \overline{\mathcal{P}}_n$ is a subgroup such that q commutes with every element of H. Define $H_P := \{p^j \in P : \omega^t p^j \in H, \text{ for some } t \in \mathbb{Z}_d\}$. Then $H_P \subseteq \{p^j : 0 \leq j \leq a-1, j \equiv 0 \pmod{a_1}\}$.

Consider the group $\langle G\rangle$, each of whose elements commute with q (resp. p) by (a). Then it follows from Claim B and Claim A respectively, that

$$\begin{aligned} \left|\frac{\langle G\rangle}{I_G} \bigcap \frac{\langle p\rangle}{I_{\{p\}}}\right| &\leq |\{a_1 t \bmod a : t \in \mathbb{Z}_a\}| = a/a_1, \\ \left|\frac{\langle G\rangle}{I_G} \bigcap \frac{\langle q\rangle}{I_{\{q\}}}\right| &\leq |\{a_1 t \bmod b : t \in \mathbb{Z}_b\}| = b/a_1. \end{aligned} \tag{44}$$

This now implies

$$\begin{aligned} \left|\langle G, p\rangle / I_{G\cup\{p\}}\right| &= \frac{\left|\langle G\rangle / I_G\right| \left|\langle p\rangle / I_{\{p\}}\right|}{\left|\langle G\rangle / I_G \cap \langle p\rangle / I_{\{p\}}\right|} \geq \frac{\left|\langle G\rangle / I_G\right| a}{a/a_1} = a_1 \left|\langle G\rangle / I_G\right|, \\ \left|\langle G, q\rangle / I_{G\cup\{q\}}\right| &= \frac{\left|\langle G\rangle / I_G\right| \left|\langle q\rangle / I_{\{q\}}\right|}{\left|\langle G\rangle / I_G \cap \langle q\rangle / I_{\{q\}}\right|} \geq \frac{\left|\langle G\rangle / I_G\right| b}{b/a_1} = a_1 \left|\langle G\rangle / I_G\right|. \end{aligned} \tag{45}$$

For the bound on $\left|\langle G, p, q\rangle / I_{G\cup\{p,q\}}\right|$, consider the subgroup generated by $\overline{G} := G \cup \{p\}$. Then every element of $\langle \overline{G}\rangle$ commutes with p. Again by Claim A, we have that $\left|\langle \overline{G}\rangle / I_{\overline{G}} \cap \langle q\rangle / I_{\{q\}}\right| \leq b/a_1$, and then it follows that

$$\left|\langle G, p, q\rangle / I_{G\cup\{p,q\}}\right| = \frac{\left|\langle \overline{G}\rangle / I_{\overline{G}}\right| \left|\langle q\rangle / I_{\{q\}}\right|}{\left|\langle \overline{G}\rangle / I_{\overline{G}} \cap \langle q\rangle / I_{\{q\}}\right|} \geq a_1 \left|\langle \overline{G}\rangle / I_{\overline{G}}\right| \geq a_1^2 \left|\langle G\rangle / I_G\right|, \tag{46}$$

where we used Eq. (45) in the last inequality.

We now prove Claim A, and then proof of Claim B is exactly similar. For contradiction, assume that there exists $q^j \in H_Q$ such that $j \not\equiv 0 \pmod{a_1}$. Then $0 = [\![p, q^j]\!]_d = cj \bmod d$, as q^j

is equivalent to some element of H up to phase factors. But this implies that j must be a multiple of a_1, as a_1 is the order of c (specifically order implies that for any positive integer $0 < \gamma < a_1$, $c\gamma \bmod d \neq 0$). Thus we have a contradiction, and the claim is proved. □

Lemma 6.8 follows from parts (iii)-(v) of the next lemma. Several parts of this stronger lemma are used in subsequent proofs in Section 6.1.

Lemma B.4. *Let $\pi_2(S) \in \mathbb{Z}_d^{2n\times k}$ have the Smith normal form $\left(\begin{smallmatrix} D & 0 \\ 0 & 0 \end{smallmatrix}\right)$, so that $\pi_2(S)P = Q\left(\begin{smallmatrix} D & 0 \\ 0 & 0 \end{smallmatrix}\right)$, for $D \in \mathbb{Z}_d^{r\times r}$ a diagonal matrix with all diagonal entries non-zero, $r \geq 1$, and invertible matrices $P \in \mathbb{Z}_d^{k\times k}$, $Q \in \mathbb{Z}_d^{2n\times 2n}$. Let M denote the submodule of $\mathbb{Z}_d^{2n}$ generated by the columns of $\pi_2(S)$. Then we have the following:*

(i) $v \in \pi_2(\langle S\rangle)$ if and only if $v \in M$, where $\pi_2(\langle S\rangle)$ can be regarded as a multiset of size $|\langle S\rangle|$.

(ii) If $T \subseteq \langle S\rangle$ is a generating set of $\langle S\rangle$, then the submodule of $\mathbb{Z}_d^{2n}$ generated by the columns of $\pi_2(T)$ equals M.

(iii) If $T \subseteq \langle S\rangle$ is a generating set of $\langle S\rangle$, then T contains a subset T', with $|T'| \geq r$, such that $u \in T'$ implies $\pi_2(u) \neq 0$. Moreover, the number of invariant factors of $\pi_2(T)$ is r.

(iv) If $r = k$, then S is a generating set of $\langle S\rangle$ of the smallest size.

(v) If $r \leq k$, there exists a generating set $T = T' \cup \{p\}$ of $\langle S\rangle$ such that $\langle p\rangle = I_S$ and $|T'| = r$. Moreover, any such generating set has the following properties: (a) the columns of $\pi_2(T')$ generate the submodule M, (b) the matrix $\pi_2(T')$ has r invariant factors, and (c) for distinct elements $q, r \in T'$, $\pi_2(q)$ and $\pi_2(r)$ are distinct and non-zero.

Proof. For part (i), suppose that $v \in \pi_2(\langle S\rangle)$. This means that there exists $p \in \langle S\rangle$ such that $\pi_2(p) = v$. Now one can write $p = \prod_{j=0}^{\ell} h_j$, where each $h_j \in S$, which upon rearrangement of the order of the product using group commutators and simplifying the result using $s_j^d = \pm I$ for each j (by Lemma 6.1), gives $p = \lambda \prod_{j=0}^{k-1} s_j^{w_j}$, where $w := (w_0, w_1, \ldots, w_{k-1}) \in \mathbb{Z}_d^k$, and $\lambda \in \langle S\rangle$ with $\pi_2(\lambda) = 0$. We now conclude that $v = \pi_2(p) = \sum_{j=0}^{k-1} w_j \pi_2(s_j) = \pi_2(S)w \in M$. For the other direction, assume $v \in M$, which then implies that there exists $w \in \mathbb{Z}_d^k$, such that $v = \pi_2(S)w$. Now define $p := \prod_{j=0}^{k-1} s_j^{w_j}$. Then clearly $p \in \langle S\rangle$, and $\pi_2(p) = v$, proving that $v \in \pi_2(\langle S\rangle)$.

For part (ii), let the ordered multiset $T := \{q_0, q_1, \ldots, q_{\ell-1}\} \subseteq \langle S\rangle$ be a generating set of $\langle S\rangle$, and let M' be the module generated by the columns of $\pi_2(T) \in \mathbb{Z}_d^{2n\times \ell}$. Take $v \in M$. Then by part (i), $v \in \pi_2(\langle S\rangle)$. Since $\langle T\rangle = \langle S\rangle$ by assumption, we again get $v \in M'$ by part (i). Thus $M \subseteq M'$, and running the argument backwards gives $M' \subseteq M$.

Note that by Lemma 2.4, the minimal number of generators of M is $\Theta(M) = r$. Part (ii) implies that the columns of $\pi_2(T)$ generate the submodule M, and so by Corollary 6.7, $\pi_2(M)$ has r invariant factors. Also, if $|\{v \in T : \pi_2(v) \neq 0\}| < r$, then this leads to a contradiction as it implies that the minimal number of generators of M is less than r.

Part (iv) follows because S is a generating set for $\langle S\rangle$ of size k, and part (iii) implies that any generating set of $\langle S\rangle$ must have size k.

For part (v), if $r = k$, we can take $T' = S$ and any $p \in I_S$ such that $\langle p\rangle = I_S$. So assume $r < k$, and let us denote $\hat{Q} := Q\left(\begin{smallmatrix} D & 0 \\ 0 & 0 \end{smallmatrix}\right) = \left(\overline{Q} \;\; 0\right)$, where $\overline{Q} \in \mathbb{Z}_d^{2n\times r}$ with no non-zero columns and all columns distinct (by assumption on invertibility of Q, and uniqueness of Smith normal form). We define the ordered multiset $U := \{u_0, u_1, \ldots, u_{k-1}\} \subseteq \langle S\rangle$, with $u_i := \prod_{j=0}^{k-1} s_j^{P_{ji}}$. The equation $\pi_2(S)P = Q\left(\begin{smallmatrix} D & 0 \\ 0 & 0 \end{smallmatrix}\right)$ then implies $\pi_2(u_i)$ equals the i^{th} column of $\hat{Q}$ for every i, from which we deduce that for every $i = r, r+1, \ldots, k-1$, $u_i = \omega^{\mu_i} I \in I_S$ for some $\mu_i \in \mathbb{Z}_d$. Now Theorem 6.3 implies that $\langle S\rangle = \langle U\rangle$, as P is invertible. Also Lemma 6.5(ii) implies that there exists $\mu \in \mathbb{Z}_d$ such that $\langle \omega^\mu I\rangle = I_S$. Combining these facts we deduce that $\langle S\rangle = \langle u_0, u_1, \ldots, u_{r-1}, \omega^\mu I\rangle$. Thus we may choose $T' = \{u_0, u_1, \ldots, u_{r-1}\}$ and $p = \omega^\mu I$. We now prove properties (a)-(c) for any such generating set. Note that the columns of $\pi_2(T)$ and $\pi_2(T')$ both generate the same submodule.

Then property (a) follows by part (ii), while property (b) follows by part (iii) and Corollary 6.7. Property (c) follows because if $\pi_2(T')$ had two columns which are the same or if one column was zero, then the number of invariant factors of $\pi_2(T')$ (which equals the minimal number of generators for the submodule generated by the columns of T') would be less than r, which would contradict property (b). □

Finally, we present the proof of Lemma 6.10.

Lemma 6.10. Given $S \subseteq \overline{\mathcal{P}}_n$, suppose that $T := T' \cup \{p\}$ is a near-minimal generating set of $\langle S \rangle$ with $\langle p \rangle = I_S$. Let d be prime. Then the following conditions are equivalent.

(i) T is a minimal generating set of $\langle S \rangle$.

(ii) $\langle T' \rangle$ is a stabilizer subgroup of $\overline{\mathcal{P}}_n$, and $I_S \neq \{I\}$.

Proof. Let the number of invariant factors of $\pi_2(S)$ be $r \geq 1$, and suppose $T' = \{q_0, q_1, \ldots, q_{r-1}\}$. We first note a few facts about T'. Since d is prime, $\mathbb{Z}_d$ is a field. By Lemma B.4(v), we know that $\pi_2(T')$ has $r \geq 1$ invariant factors, and since $|T'| = r$, this implies that the kernel of the matrix $\pi_2(T')$ is trivial (as $\mathbb{Z}_d$ is a field). Now let $T'' := \{p^{\gamma_0} q_0, p^{\gamma_1} q_1, \ldots, p^{\gamma_{r-1}} q_{r-1}\}$, for some $(\gamma_0, \gamma_1, \ldots, \gamma_{r-1}) \in \mathbb{Z}_d^r$. Then by Lemma 6.2(iv), we know that $\langle J_{T'} \rangle = \langle J_{T''} \rangle$. Also since $\pi_2(T') = \pi_2(T'')$, we conclude that the kernel of $\pi_2(T'')$ is trivial. Combining these observations, and by using Lemma 6.4(iii) we deduce that $I_{T'} = I_{T''}$. Let us call this Fact (a). The other fact we need is that, since d is prime, $\langle \omega^j I \rangle = \{\omega^\ell I : \ell \in \mathbb{Z}_d\}$, for every $j \in \mathbb{Z}_d \setminus \{0\}$. Let us call this Fact (b). We now return to the proof.

First we prove that (i) implies (ii). So assume that T is a minimal generating set of $\langle S \rangle$. For contradiction, assume that $I_S = \{I\}$, which then implies $p = I$. Hence T cannot be a minimal generating set as $\langle T' \rangle = \langle T', p \rangle$. Thus $I_S \neq \{I\}$. Next for contradiction again assume that $\langle T' \rangle$ is not a stabilizer group. This means that there exists $\omega^j I \in \langle T' \rangle$, for some $j \in \mathbb{Z}_d \setminus \{0\}$, and then Fact (b) implies that $p \in \langle T' \rangle$. Thus again we conclude that T is not a minimal generating set of $\langle S \rangle$.

Next we prove that (ii) implies (i). So assume that T' is a stabilizer group and $I_S \neq \{I\}$, which implies that $I_{T'} = \{I\}$ and $p \neq I$, respectively. For contradiction, assume that T is not a minimal generating set of $\langle S \rangle$. Then by Theorem 6.9(iii), we can conclude that there exist integers $\gamma_0, \gamma_1, \ldots, \gamma_{r-1} \in \mathbb{Z}_d$, such that $p \in \langle T'' \rangle$, where $T'' := \{p^{\gamma_0} q_0, p^{\gamma_1} q_1, \ldots, p^{\gamma_{r-1}} q_{r-1}\}$. Now by Fact (a) we also have that $I_{T''} = I_{T'} = \{I\}$. Hence we can conclude that $p = I$, giving a contradiction. This proves the lemma. □

C Maximum collections of non-commuting pairs

While subsection 5.1 gave both necessary and sufficient conditions for a k-tuple $(f_0, \ldots, f_{k-1}) \in (\mathbb{Z}_d \setminus \{0\})^k$ to be an achievable non-commuting pair relation on n qudits, it is possible to give a more direct characterization of which k-tuples are achievable non-commuting pair relations for the case $k = nm$, i.e. when we have the maximum number of non-commuting pairs on n qudits.

To simplify the presentation, let us introduce the following notation.

Notation. Suppose $f := (f_0, \ldots, f_{k-1})$ is an achievable non-commuting pair relation on n qudits. Suppose $\beta := (\beta_0, \ldots, \beta_{k-1}) \in (\mathbb{Z}_d \setminus \{0\})^k$ be such that $\beta_i f_i \not\equiv 0 \pmod{d}$, for every $i = 0, \ldots, k-1$. For a fixed f, we will denote the set of all such β by $\boldsymbol{\beta}(f)$. If $\beta \in \boldsymbol{\beta}(f)$, we will denote $f_\beta := (f'_0, \ldots, f'_{k-1})$, where $f'_i = \beta_i f_i \mod d$, for every i. Let $\mathrm{Sym}(k)$ denote the permutation group on the set of indices $\{0, \ldots, k-1\}$. If $\sigma \in \mathrm{Sym}(k)$, we will denote $f_\sigma := (f_{\sigma(0)}, \ldots, f_{\sigma(k-1)})$.

Given $f_i \in \mathbb{Z}_d \setminus \{0\}$ as in the above notation, it is useful to explicitly write down for which values of $\beta_i \in \mathbb{Z}_d \setminus \{0\}$, we have $\beta_i f_i \not\equiv 0 \pmod{d}$. This is easy to calculate. We start by writing $f_i = h \gcd(f_i, d)$, the gcd evaluated over integers, so that we have $\gcd(h, d) = 1$. Now consider

the set $\delta_i := \{td/\gcd(f_i,d) : t \in \mathbb{Z}, 1 \le t < \gcd(f_i,d)\}$. Then we claim that $\beta_i f_i \not\equiv 0 \pmod d$ if and only if $\beta_i \notin \delta_i$. The forward direction is easy (we prove the contrapositive): if $\beta_i \in \delta_i$, then we have $\beta_i f_i = htd$ for some integer t, and so $\beta_i f_i \equiv 0 \pmod d$. For the converse direction, assume for contradiction that $\beta_i f_i \equiv 0 \pmod d$, and $\beta_i \in (\mathbb{Z}_d \setminus \{0\}) \setminus \delta_i$. This means that d divides $h\gcd(f_i,d)\beta_i$, and since $\gcd(h,d) = 1$ this implies that β_i is a multiple of $d/\gcd(f_i,d)$, which is a contradiction.

With the notation above, we state a helper lemma:

Lemma C.1. *If $f := (f_0, \ldots, f_{k-1})$ is an achievable non-commuting pair relation on n qudits, then*

(i) f_σ is an achievable non-commuting pair relation on n qudits, for every $\sigma \in Sym(k)$.

(ii) f_β is an achievable non-commuting pair relation on n qudits, for every $\beta \in \boldsymbol{\beta}(f)$.

(iii) $\bar{f} := (f_0, \ldots, f_{k-1}, f_k)$ is an achievable non-commuting pair relation on $n+1$ qudits, for every $f_k \in \mathbb{Z}_d \setminus \{0\}$.

Proof. Let $S = \{s_0, \ldots, s_{k-1}\}$ and $T = \{t_0, \ldots, t_{k-1}\}$ be a non-commuting pair that achieves f. For part (i), let $\sigma \in \mathrm{Sym}(k)$. Define the ordered sets $S_\sigma := \{s_{\sigma(0)}, \ldots, s_{\sigma(k-1)}\}$ and $T_\sigma := \{t_{\sigma(0)}, \ldots, t_{\sigma(k-1)}\}$. Then S_σ and T_σ achieves f_σ.

For part (ii), let $u_0, \ldots, u_{k-1} \in \mathbb{Z}_d^{2n}$ be such that $s_i = P(u_i)$, for every i, and let $\beta \in \boldsymbol{\beta}(f)$. Define the ordered set $\tilde{S} := \{P(\tilde{u}_0), \ldots, P(\tilde{u}_{k-1})\}$, where $\tilde{u}_i = \beta_i u_i \bmod d$. Then $\tilde{S}, T$ achieves $\tilde{f}$, because for all $i, j \in \{0, \ldots, k-1\}$, we have

$$[\![P(\tilde{u}_i), P(\tilde{u}_j)]\!]_d = [\![s_i, s_j]\!]_d = [\![t_i, t_j]\!]_d = 0, \qquad [\![P(\tilde{u}_i), t_j]\!]_d = \begin{cases} \beta_i f_i \bmod d & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases} \tag{47}$$

For part (iii), let $f_k \in \mathbb{Z}_d \setminus \{0\}$ be arbitrary. Now construct the ordered sets $\bar{S} := \{\bar{s}_0, \ldots, \bar{s}_k\} \subseteq \mathcal{P}_{n+1}$ and $\bar{T} := \{\bar{t}_0, \ldots, \bar{t}_k\} \subseteq \mathcal{P}_{n+1}$, where $\bar{s}_i = s_i \otimes I$, $\bar{t}_i = t_i \otimes I$ (here I is $d \times d$), for every $i = 0, \ldots, k-1$, and $\bar{s}_k = I \otimes X$, $\bar{t}_k = I \otimes Z^{f_k}$ (here I is $d^n \times d^n$). Then $\bar{S}, \bar{T}$ achieves $\bar{f}$. □

We can now provide the complete characterization of all achievable non-commuting pair relations on n qudits, when $k = nm$.

Theorem C.2. *Let $Sym(nm)$ denote the permutation group on the set of indices $\{0, 1, \ldots, nm-1\}$. Define $f := (f_0, \ldots, f_{nm-1}) \in (\mathbb{Z}_d \setminus \{0\})^{nm}$ given by*

$$f_{i+jn} = \frac{d}{p_j^{\alpha_j}}, \quad \text{for all } i = 0, 1, \ldots, n-1, \;\; j = 0, 1, \ldots, m-1, \tag{48}$$

and also the set

$$\mathcal{F} := \{(f_\beta)_\sigma \in (\mathbb{Z}_d \setminus \{0\})^{nm} : \beta \in \boldsymbol{\beta}(f),\ \sigma \in Sym(nm)\}. \tag{49}$$

Then the following hold:

(i) If $\beta \in \boldsymbol{\beta}(f)$, then for every i, j, we have $\beta_{i+jn} \in \mathbb{Z}_d \setminus \{0\} \setminus \{tp_j^{\alpha_j} : t \in \mathbb{Z}, 1 \le t < d/p_j^{\alpha_j}\}$, and $(f_\beta)_{i+jn} \in \{td/p_j^{\alpha_j} : t \in \mathbb{Z}, 1 \le t < p_j^{\alpha_j}\}$.

(ii) Each element of $\mathcal{F}$ is an achievable non-commuting pair relation on n qudits. Conversely, if any $f' \in (\mathbb{Z}_d \setminus \{0\})^{nm}$ is an achievable non-commuting pair relation on n qudits, then $f' \in \mathcal{F}$.

Proof. (i) The statement about β_{i+jn} follows from the argument in the paragraph above Lemma C.1. This now implies that $(f_\beta)_{i+jn} \in \{td/p_j^{\alpha_j} : t \in \mathbb{Z}, 1 \le t < p_j^{\alpha_j}\}$.

(ii) We first prove the forward direction. We choose the ordered sets $S := \{P(u_0), \ldots, P(u_{nm-1})\}$ and $T := \{P(v_0), \ldots, P(v_{nm-1})\}$, where the vectors $u_0, \ldots, u_{nm-1}, v_0, \ldots, v_{nm-1} \in \mathbb{Z}_d^{2n}$ have components defined by

$$(u_{i+jn})_\ell = \begin{cases} d/p_j^{\alpha_j} & \text{if } \ell = i \\ 0 & \text{if } \ell \neq i \end{cases}, \quad (v_{i+jn})_\ell = \begin{cases} -d/p_j^{\alpha_j} & \text{if } \ell = i+n \\ 0 & \text{if } \ell \neq i+n \end{cases}, \tag{50}$$

for all $i \in \{0, 1, \ldots, n-1\}$, $j \in \{0, 1, \ldots, m-1\}$, and $\ell \in \{0, \ldots, 2n-1\}$. Then one can easily check that S, T achieves the non-commuting pair relation $\bar{f} = (\bar{f}_0, \ldots, \bar{f}_{nm-1}) \in (\mathbb{Z}_d \setminus \{0\})^{nm}$, where $\bar{f}_i = (f_i)^2 \bmod d$, for every $i = 0, \ldots, nm-1$. Now by Corollary B.2, there exists $\beta_0, \beta_1, \ldots, \beta_{m-1} \in \mathbb{Z}_d \setminus \{0\}$, such that $\beta_j (d/p_j^{\alpha_j})^2 \bmod d = d/p_j^{\alpha_j}$, for every j. By Lemma C.1(ii), we then conclude that f is also an achievable non-commuting pair relation on n qudits. Finally, each element of $\mathcal{F}$ is an achievable non-commuting pair relation on n qudits by Lemma C.1(i),(ii).

For the converse, suppose that the ordered sets $S = \{s_0, \ldots, s_{nm-1}\}$ and $T = \{t_0, \ldots, t_{nm-1}\}$ are a collection of non-commuting pairs on n qudits that generate $f' = (f'_0, \ldots, f'_{nm-1}) \in (\mathbb{Z}_d \setminus \{0\})^{nm}$. For every $j = 0, \ldots, m-1$, define the multisets $F_j := \{f'_i : f'_i \not\equiv 0 \pmod{p_j^{\alpha_j}}, \ i \in \{0, \ldots, nm-1\}\}$, and then by Lemma 5.1 we know that $|F_j| \leq n$. We now claim that (i) $|F_j| = n$, for every j, and (ii) $F_\ell \cap F_j = \emptyset$, whenever $j \neq \ell$. To see this note that if any of these two conditions does not hold, then by a simple counting argument we conclude that there exists $f'_i \notin \cup_{j=1}^m F_j$. But then that would imply that f'_i is divisible by every $p_j^{\alpha_j}$, or equivalently f'_i is divisible by d. This gives a contradiction, which proves the claim. The claim also implies that each f'_i is an element of exactly one F_j. Now fix a F_j and choose any $f'_i \in F_j$. Then we must have $f'_i = k(d/p_j^{\alpha_j})$, for some $k \not\equiv 0 \pmod{p_j^{\alpha_j}}$. This immediately implies $f' \in \mathcal{F}$, as F_j is arbitrary. □

References

[1] Andrew D. Greentree, S. G. Schirmer, F. Green, Lloyd C. L. Hollenberg, A. R. Hamilton, and R. G. Clark. "Maximizing the Hilbert space for a finite number of distinguishable quantum states". Phys. Rev. Lett. **92**, 097901 (2004). doi: 10.1103/PhysRevLett.92.097901.

[2] Markus Grassl, Thomas Beth, and Martin Rötteler. "On optimal quantum codes". International Journal of Quantum Information **02**, 55–64 (2004). doi: 10.1142/S0219749904000079.

[3] Suhail Ahmad Rather, Adam Burchardt, Wojciech Bruzda, Grzegorz Rajchel-Mieldzioć, Arul Lakshminarayan, and Karol Życzkowski. "Thirty-six entangled officers of Euler: Quantum solution to a classically impossible problem". Phys. Rev. Lett. **128**, 080507 (2022). doi: 10.1103/PhysRevLett.128.080507.

[4] Michael A. Nielsen, Michael J. Bremner, Jennifer L. Dodd, Andrew M. Childs, and Christopher M. Dawson. "Universal simulation of Hamiltonian dynamics for quantum systems with finite-dimensional state spaces". Phys. Rev. A **66**, 022317 (2002). doi: 10.1103/PhysRevA.66.022317.

[5] Jonathan E. Moussa. "Quantum circuits for qubit fusion". Quantum Information & Computation **16**, 1113–1124 (2016). doi: 10.26421/QIC16.13-14-3.

[6] Alex Bocharov, Martin Roetteler, and Krysta M. Svore. "Factoring with qutrits: Shor's algorithm on ternary and metaplectic quantum architectures". Phys. Rev. A **96**, 012306 (2017). doi: 10.1103/PhysRevA.96.012306.

[7] Earl T. Campbell, Hussain Anwar, and Dan E. Browne. "Magic-state distillation in all prime dimensions using quantum Reed-Muller codes". Phys. Rev. X **2**, 041021 (2012). doi: 10.1103/PhysRevX.2.041021.

[8] Anirudh Krishna and Jean-Pierre Tillich. "Towards low overhead magic state distillation". Phys. Rev. Lett. **123**, 070507 (2019). doi: 10.1103/PhysRevLett.123.070507.

[9] Tyler D. Ellison, Yu-An Chen, Arpit Dua, Wilbur Shirley, Nathanan Tantivasadakarn, and Dominic J. Williamson. "Pauli stabilizer models of twisted quantum doubles". PRX Quantum **3**, 010353 (2022). doi: 10.1103/PRXQuantum.3.010353.

[10] Tyler D. Ellison, Yu-An Chen, Arpit Dua, Wilbur Shirley, Nathanan Tantivasadakarn, and Dominic J. Williamson. "Pauli topological subsystem codes from Abelian anyon theories". Quantum **7**, 1137 (2023). doi: 10.22331/q-2023-10-12-1137.

[11] Noah Goss, Alexis Morvan, Brian Marinelli, Bradley K. Mitchell, Long B. Nguyen, Ravi K. Naik, Larry Chen, Christian Jünger, John Mark Kreikebaum, David I. Santiago, Joel J. Wallman, and Irfan Siddiqi. "High-fidelity qutrit entangling gates for superconducting circuits". Nature Communications **13**, 7481 (2022). doi: 10.1038/s41467-022-34851-z.

[12] Kai Luo, Wenhui Huang, Ziyu Tao, Libo Zhang, Yuxuan Zhou, Ji Chu, Wuxin Liu, Biying Wang, Jiangyu Cui, Song Liu, Fei Yan, Man-Hong Yung, Yuanzhen Chen, Tongxing Yan, and Dapeng Yu. "Experimental realization of two qutrits gate with tunable coupling in superconducting circuits". Phys. Rev. Lett. **130**, 030603 (2023). doi: 10.1103/PhysRevLett.130.030603.

[13] Peter B R Nisbet-Jones, Jerome Dilley, Annemarie Holleczek, Oliver Barter, and Axel Kuhn. "Photonic qubits, qutrits and ququads accurately prepared and delivered on demand". New Journal of Physics **15**, 053007 (2013). doi: 10.1088/1367-2630/15/5/053007.

[14] Michael Kues, Christian Reimer, Piotr Roztocki, Luis Romero Cortés, Stefania Sciara, Benjamin Wetzel, Yanbing Zhang, Alfonso Cino, Sai T. Chu, Brent E. Little, David J. Moss, Lucia Caspani, José Azaña, and Roberto Morandotti. "On-chip generation of high-dimensional entangled quantum states and their coherent control". Nature **546**, 622–626 (2017). doi: 10.1038/nature22986.

[15] Laurin E. Fischer, Alessandro Chiesa, Francesco Tacchino, Daniel J. Egger, Stefano Carretta, and Ivano Tavernelli. "Universal qudit gate synthesis for transmons". PRX Quantum **4**, 030327 (2023). doi: 10.1103/PRXQuantum.4.030327.

[16] Shuang Wang, Zhen-Qiang Yin, H F Chau, Wei Chen, Chao Wang, Guang-Can Guo, and Zheng-Fu Han. "Proof-of-principle experimental realization of a qubit-like qudit-based quantum key distribution scheme". Quantum Science and Technology **3**, 025006 (2018). doi: 10.1088/2058-9565/aaace4.

[17] Eufemio Moreno-Pineda, Clément Godfrin, Franck Balestro, Wolfgang Wernsdorfer, and Mario Ruben. "Molecular spin qudits for quantum algorithms". Chem. Soc. Rev. **47**, 501–513 (2018). doi: 10.1039/C5CS00933B.

[18] Mario Chizzini, Luca Crippa, Luca Zaccardi, Emilio Macaluso, Stefano Carretta, Alessandro Chiesa, and Paolo Santini. "Quantum error correction with molecular spin qudits". Phys. Chem. Chem. Phys. **24**, 20030–20039 (2022). doi: 10.1039/D2CP01228F.

[19] Daniel Gottesman. "Stabilizer codes and quantum error correction". PhD thesis. California Institute of Technology. (1997). doi: 10.7907/rzr7-dt72.

[20] Daniel Gottesman. "Fault-tolerant quantum computation with higher-dimensional systems". In C.P. Williams, editor, Quantum Computing and Quantum Communications, QCQC 1998. Volume 1509, pages 302–313. Springer Berlin Heidelberg (1999). doi: 10.1007/3-540-49208-9_27.

[21] Utkan Güngördü, Rabindra Nepal, and Alexey A. Kovalev. "Parafermion stabilizer codes". Phys. Rev. A **90**, 042326 (2014). doi: 10.1103/PhysRevA.90.042326.

[22] Rahul Sarkar and Theodore J Yoder. "A graph-based formalism for surface codes and twists" (2021). doi: 10.48550/arXiv.2101.09349.

[23] Lane G. Gunderman. "Transforming collections of Pauli operators into equivalent collections of Pauli operators over minimal registers". Phys. Rev. A **107**, 062416 (2023). doi: 10.1103/PhysRevA.107.062416.

[24] Greg Kuperberg. "Kasteleyn cokernels.". The Electronic Journal of Combinatorics [electronic only] **9**, R29, 30p. (2002). doi: 10.37236/1645.

[25] Mark M. Wilde. "Logical operators of quantum codes". Phys. Rev. A **79**, 062322 (2009). doi: 10.1103/PhysRevA.79.062322.

[26] Pascual Jordan and Eugene Paul Wigner. "Über das Paulische Äquivalenzverbot". Zeitschrift für Physik **47**, 631–651 (1928). doi: 10.1007/BF01331938.

[27] Sergey B. Bravyi and Alexei Yu. Kitaev. "Fermionic quantum computation". Annals of Physics **298**, 210–226 (2002). doi: 10.1006/aphy.2002.6254.

[28] F Verstraete and J. Ignacio Cirac. "Mapping local Hamiltonians of fermions to local Hamiltonians of spins". Journal of Statistical Mechanics: Theory and Experiment **2005**, P09012 (2005). doi: 10.1088/1742-5468/2005/09/P09012.

[29] Vojtěch Havlíček, Matthias Troyer, and James D. Whitfield. "Operator locality in the quantum simulation of fermionic models". Phys. Rev. A **95**, 032332 (2017). doi: 10.1103/PhysRevA.95.032332.

[30] Zhang Jiang, Amir Kalev, Wojciech Mruczkiewicz, and Hartmut Neven. "Optimal fermion-to-qubit mapping via ternary trees with applications to reduced quantum states learning". Quantum **4**, 276 (2020). doi: 10.22331/q-2020-06-04-276.

[31] Sergey Bravyi, Jay M. Gambetta, Antonio Mezzacapo, and Kristan Temme. "Tapering off qubits to simulate fermionic Hamiltonians" (2017). doi: 10.48550/arXiv.1701.08213.

[32] Kanav Setia, Sergey Bravyi, Antonio Mezzacapo, and James D. Whitfield. "Superfast encodings for fermionic quantum simulation". Phys. Rev. Res. **1**, 033033 (2019). doi: 10.1103/PhysRevResearch.1.033033.

[33] Kanav Setia, Richard Chen, Julia E. Rice, Antonio Mezzacapo, Marco Pistoia, and James D. Whitfield. "Reducing qubit requirements for quantum simulations using molecular point group symmetries". Journal of Chemical Theory and Computation **16**, 6091–6097 (2020). doi: 10.1021/acs.jctc.0c00113.

[34] Jacob T. Seeley, Martin J. Richard, and Peter J. Love. "The Bravyi-Kitaev transformation for quantum computation of electronic structure". The Journal of Chemical Physics **137**, 224109 (2012). doi: 10.1063/1.4768229.

[35] Mark Steudtner and Stephanie Wehner. "Fermion-to-qubit mappings with varying resource requirements for quantum simulation". New Journal of Physics **20**, 063010 (2018). doi: 10.1088/1367-2630/aac54f.

[36] Todd Brun, Igor Devetak, and Min-Hsiu Hsieh. "Correcting quantum errors with entanglement". Science **314**, 436–439 (2006). doi: 10.1126/science.1131563.

[37] Min-Hsiu Hsieh. "Entanglement-assisted coding theory". PhD thesis. University of Southern California. (2008). url: https://www.proquest.com/dissertations-theses/entanglement-assisted-coding-theory/docview/304492442/se-2.

[38] Mark M. Wilde and Todd A. Brun. "Optimal entanglement formulas for entanglement-assisted quantum coding". Phys. Rev. A **77**, 064302 (2008). doi: 10.1103/PhysRevA.77.064302.

[39] Monireh Houshmand, Saied Hosseini-Khayat, and Mark M. Wilde. "Minimal-memory, noncatastrophic, polynomial-depth quantum convolutional encoders". IEEE Transactions on Information Theory **59**, 1198–1210 (2013). doi: 10.1109/TIT.2012.2220520.

[40] A Yu Kitaev. "Unpaired Majorana fermions in quantum wires". Physics-Uspekhi **44**, 131 (2001). doi: 10.1070/1063-7869/44/10S/S29.

[41] Sagar Vijay and Liang Fu. "Quantum error correction for complex and Majorana fermion qubits" (2017). doi: 10.48550/arXiv.1703.00459.

[42] Vlad Gheorghiu. "Standard form of qudit stabilizer groups". Physics Letters A **378**, 505–509 (2014). doi: 10.1016/j.physleta.2013.12.009.

[43] Scott Aaronson and Daniel Gottesman. "Improved simulation of stabilizer circuits". Phys. Rev. A **70**, 052328 (2004). doi: 10.1103/PhysRevA.70.052328.

[44] Lane G. Gunderman. "Stabilizer codes with exotic local-dimensions". Quantum **8**, 1249 (2024). doi: 10.22331/q-2024-02-12-1249.

[45] Zihan Lei. "Qudit surface codes and hypermap codes". Quantum Information Processing **22**, 297 (2023). doi: 10.1007/s11128-023-04060-8.

[46] Serge Lang. "Algebra". Volume 211 of Graduate Texts in Mathematics, pages xvi+914. Springer-Verlag, New York. (2002). Third edition. doi: 10.1007/978-1-4613-0041-0.

[47] William C. Brown. "Matrices over commutative rings". Volume 169 of Monographs and textbooks in pure and applied mathematics. Marcel Dekker, Inc., New York. (1993).

[48] T. J. Kaczynski. "Another proof of Wedderburn's theorem". The American Mathematical Monthly **71**, 652–653 (1964). doi: 10.2307/2312328.

[49] Robert B. Ash. "Basic abstract algebra: for graduate students and advanced undergraduates". Dover Publications Inc., New York. (2013).

[50] Thomas W. Hungerford. "Algebra". Volume 73 of Graduate Texts in Mathematics. Springer-Verlag, New York. (1974). First edition. doi: 10.1007/978-1-4612-6101-8.

[51] Tsit-Yuen Lam. "Lectures on modules and rings". Volume 189 of Graduate Texts in Mathematics. Springer-Verlag, New York. (1999). First edition. doi: 10.1007/978-1-4612-0525-8.

[52] Rahul Sarkar. "Size of minimal generating set of a module generated by columns of a diagonal matrix with extra structure". MathOverflow. url: https://mathoverflow.net/q/431397 (version: 2022-09-28).

[53] Arne Storjohann. "Algorithms for matrix canonical forms". PhD thesis. ETH Zurich. Zürich (2000). doi: 10.3929/ethz-a-004141007.

[54] John A. Howell. "Spans in the module $(\mathbb{Z}_m)^s$". Linear and Multilinear Algebra **19**, 67–77 (1986). doi: 10.1080/03081088608817705.

[55] Mark A. Webster, Benjamin J. Brown, and Stephen D. Bartlett. "The XP stabiliser formalism: a generalisation of the Pauli stabiliser formalism with arbitrary phases". Quantum **6**, 815 (2022). doi: 10.22331/q-2022-09-22-815.

[56] Claus Fieker and Tommy Hofmann. "Computing in quotients of rings of integers". LMS Journal of Computation and Mathematics **17**, 349–365 (2014). doi: 10.1112/S1461157014000291.

[57] Rahul Sarkar and Ewout van den Berg. "On sets of maximally commuting and anticommuting Pauli operators". Research in the Mathematical Sciences **8**, 14 (2021). doi: 10.1007/s40687-020-00244-1.

[58] Xavier Bonet-Monroig, Ryan Babbush, and Thomas E. O'Brien. "Nearly optimal measurement scheduling for partial tomography of quantum states". Phys. Rev. X **10**, 031064 (2020). doi: 10.1103/PhysRevX.10.031064.

[59] Pavel Hrubeš. "On families of anticommuting matrices". Linear Algebra and its Applications **493**, 494–507 (2016). doi: 10.1016/j.laa.2015.12.015.

[60] Patrick Solé and Michel Planat. "Extreme values of the Dedekind ψ function". Journal of Combinatorics and Number Theory **3**, 33–38 (2011). url: https://www.proquest.com/scholarly-journals/extreme-values-dedekind-psi-function/docview/1728715084/se-2.

[61] Michel Planat and Metod Saniga. "On the Pauli graphs on N-qudits". Quantum Information & Computation **8**, 127–146 (2008). doi: 10.26421/qic8.1-2-9.

[62] Michel Planat. "Pauli graphs when the Hilbert space dimension contains a square: Why the Dedekind psi function?". Journal of Physics A: Mathematical and Theoretical **44**, 045301 (2011). doi: 10.1088/1751-8113/44/4/045301.

[63] Hans Havlicek and Metod Saniga. "Projective ring line of a specific qudit". Journal of Physics A: Mathematical and Theoretical **40**, F943 (2007). doi: 10.1088/1751-8113/40/43/F03.

[64] Michel Planat and Anne-Céline Baboin. "Qudits of composite dimension, mutually unbiased bases and projective ring geometry". Journal of Physics A: Mathematical and Theoretical **40**, F1005 (2007). doi: 10.1088/1751-8113/40/46/F04.

[65] Leonard Eugene Dickson. "History of the theory of numbers". Volume 1. Carnegie Institution of Washington. (1919). doi: https://doi.org/10.5962/t.174869.

[66] Jeremy Rickard. "Condition for equality of modules generated by columns of matrices". MathOverflow. url: https://mathoverflow.net/q/437972 (version: 2023-01-06).

[67] Robert Koenig and John A. Smolin. "How to efficiently select an arbitrary Clifford group element". Journal of Mathematical Physics **55**, 122202 (2014). doi: 10.1063/1.4903507.

[68] Sergey Bravyi and Dmitri Maslov. "Hadamard-free circuits expose the structure of the Clifford group". IEEE Transactions on Information Theory **67**, 4546–4563 (2021). doi: 10.1109/TIT.2021.3081415.

[69] Alexander Miller and Victor Reiner. "Differential posets and Smith normal forms". Order **26**, 197–228 (2009). doi: 10.1007/s11083-009-9114-z.

[70] Irving Kaplansky. "Elementary divisors and modules". Transactions of the American Mathematical Society **66**, 464–491 (1949). doi: 10.2307/1990591.

[71] Dan D. Anderson, Michael Axtell, Sylvia J. Forman, and Joe Stickles. "When are associates unit multiples?". Rocky Mountain Journal of Mathematics **34**, 811–828 (2004). doi: 10.1216/rmjm/1181069828.

[72] Richard P. Stanley. "Smith normal form in combinatorics". Journal of Combinatorial Theory, Series A **144**, 476–495 (2016). doi: 10.1016/j.jcta.2016.06.013.