Reconquering Bell sampling on qudits: stabilizer learning and testing, quantum pseudorandomness bounds, and more

Jonathan Allcock*[1], Joao F. Doriguello†[2], Gábor Ivanyos‡[3], and Miklos Santha§[4,5]

[1]Tencent Quantum Laboratory, Hong Kong, China
[2]HUN-REN Alfréd Rényi Institute of Mathematics, Budapest, Hungary
[3]HUN-REN Institute for Computer Science and Control, Budapest, Hungary
[4]Centre for Quantum Technologies, National University of Singapore, Singapore
[5]CNRS, IRIF, Université Paris Cité

November 13, 2025

Abstract

Bell sampling is a simple yet powerful tool based on measuring two copies of a quantum state in the Bell basis, and has found applications in a plethora of problems related to stabiliser states and measures of magic. However, it was not known how to generalise the procedure from qubits to d-level systems – qudits – for all dimensions $d > 2$ in a useful way. Indeed, a prior work of the authors (arXiv'24) showed that the natural extension of Bell sampling to arbitrary dimensions fails to provide meaningful information about the quantum states being measured. In this paper, we overcome the difficulties encountered in previous works and develop a useful generalisation of Bell sampling to qudits of all dimensions $d \geqslant 2$. At the heart of our primitive is a new unitary, based on Lagrange's four-square theorem, that maps four copies of any stabiliser state $|\mathcal{S}\rangle$ to four copies of its complex conjugate $|\mathcal{S}^*\rangle$ (up to some Pauli operator), which may be of independent interest. We then demonstrate the utility of our new Bell sampling technique by lifting several known results from qubits to qudits for any $d \geqslant 2$ (which involves working with submodules instead of subspaces):

- Learning an unknown stabiliser state $|\mathcal{S}\rangle \in (\mathbb{C}^d)^{\otimes n}$ in $O(n^3)$ time with $O(n)$ samples;
- Solving the Hidden Stabiliser Group Problem (a stabiliser version of the State Hidden Subgroup Problem) in $\widetilde{O}(n^3/\varepsilon)$ time with $\widetilde{O}(n/\varepsilon)$ samples;
- Testing whether $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ has stabiliser size (a generalisation of stabiliser dimension for submodules) at least d^t or is ε-far from all such states in $\widetilde{O}(n^3/\varepsilon)$ time with $\widetilde{O}(n/\varepsilon)$ samples if $\varepsilon = O(d^{-2})$;
- Testing whether $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ is Haar-random or the output of a Clifford circuit augmented with less than $n/2$ single-qudit non-Clifford gates in $O(n^3)$ time using $O(n)$ samples. As a corollary, we show that Clifford circuits with at most $n/2$ single-qudit non-Clifford gates cannot prepare pseudorandom states, an exponential improvement over previous works;
- Testing whether $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ has stabiliser fidelity at least $1 - \varepsilon_1$ or at most $1 - \varepsilon_2$ with $O(d^2/\varepsilon_2)$ samples if $\varepsilon_1 = 0$ or $O(d^2/\varepsilon_2^2)$ samples if $\varepsilon_1 = O(d^{-2})$.

*jonallcock@tencent.com †doriguello@renyi.hu ‡gabor.ivanyos@sztaki.mta.hu §cqtms@nus.edu.sg

Contents

1 Introduction **3**
- 1.1 Bell sampling on qudits . . . 3
- 1.2 Applications of Bell sampling on qudits . . . 7
  - 1.2.1 Stabiliser state learning . . . 7
  - 1.2.2 Hidden Stabiliser Group Problem . . . 8
  - 1.2.3 Stabiliser size property testing . . . 8
  - 1.2.4 Testing doped Clifford circuits and pseudorandomness lower bounds . . . 9
  - 1.2.5 Stabiliser state testing . . . 10

2 Preliminaries **11**
- 2.1 Background on rings and modules . . . 12
- 2.2 Scalar product modules and symplectic modules . . . 13
- 2.3 Symplectic Fourier analysis . . . 14
- 2.4 Generalised Pauli group and Weyl operators . . . 15
- 2.5 Stabiliser groups . . . 17
- 2.6 Stabiliser size . . . 20
- 2.7 Stabiliser fidelity . . . 20

3 Bell sampling on qudits **23**
- 3.1 Further properties of distributions p_ψ and b_ψ . . . 29

4 Stabiliser state learning **31**

5 Hidden Stabiliser Group Problem **32**

6 Property testing stabiliser size **35**

7 Testing doped Clifford circuits and pseudorandomness bounds **37**
- 7.1 Anti-concentration of Haar-random states . . . 37
- 7.2 Stabiliser size of t-doped Clifford circuits . . . 38
- 7.3 Distinguishing t-doped Clifford circuits from Haar randomness . . . 39

8 Stabiliser state testing **40**
- 8.1 Tolerant stabiliser testing . . . 44

1 Introduction

Stabiliser states are central to quantum computation, and find applications in quantum error correction [Sho95, CS96, Got96, Got97], efficient classical simulations of quantum circuits [AG04, BSS16, BBC+19], low-rank recovery [KZG16], randomized benchmarking [KLR+08, MGE11, HWFW19], measurement-based quantum computation [RB00], tensor networks for holography codes [HNQ+16, NW20], and quantum learning algorithms [HKP20]. A major tool in exploring stabiliser states is *Bell sampling* [Mon17], which involves measuring two copies of a given n-qubit state $|\psi\rangle \in (\mathbb{C}^2)^{\otimes n}$ in the Bell basis. In spite of its simplicity, Bell sampling – along with a variant called Bell difference sampling [GNW21] – has proved to be a powerful primitive, and has found widespread applications [Mon17, GNW21, HG24, GIKL25, GIKL24b, GIKL24a].

The first application of Bell sampling was due to Montanaro [Mon17] to identify an unknown stabiliser state, which simplified or improved upon previous works on the same task [AG08, Röt09, ZPDF16]. Later, Gross, Nezami, and Walter [GNW21] formalised a slightly modified measurement primitive called *Bell difference sampling* consisting of performing Bell sampling twice and taking the difference of the two outcomes. Based on this, they developed an algorithm to decide whether a given unknown state is either a stabiliser state or is far from one. Bell (difference) sampling has since been employed to infer several properties associated with stabiliser states [MW16, AA24]. Hangleiter and Gullans [HG24] used Bell sampling to detect and learn circuit errors and to extract information such as the depth and number of T-gates in a circuit, while Haug, Lee, and Kim [HLK24] employed Bell sampling to measure stabiliser entropies. Around the same time, Grewal, Iyer, Kretschmer, and Liang [GIKL23, GIKL25, GIKL24b, GIKL24c, GIKL24a] applied Bell difference sampling to several tasks: obtaining the classical description of the output of a Clifford circuit augmented with a few non-Clifford single-qubit gates (called *doped Clifford circuits* [LOH24, OLLH24, LOLH24]); testing whether a state is Haar-random or the output of a doped Clifford circuit; approximating an arbitrary quantum state with a stabiliser state; and determining the stabiliser fidelity of a given state, which is its maximum overlap with any stabiliser state [BBC+19]. Finally, other works have focused on several measures of nonstabiliserness (or 'magic') [HLK24, HK23] and understanding the entanglement-magic interplay [GOL25].

While the aforementioned works focused on qubits, the scenario regarding *qudits* is still mostly unexplored, the main reason being that it is not known how to extend Bell (difference) sampling to qudits in a meaningful and useful way that retains most of the compelling properties from the qubit case. The limitations of Bell sampling on qudits was recently explored by the authors [ADIS24], who showed that a natural extension based on measuring two copies $|\psi\rangle^{\otimes 2}$ of a given n-qudit state $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ in the generalised Bell basis is doomed to fail: even in the simplest case when $|\psi\rangle$ is a stabiliser state, Bell difference sampling can return a completely random output. This is in stark contrast to the qubit case where the output is a generator of the stabiliser group. At a high level, this stems from the well-known fact that the transpose map $\psi \mapsto \bar{\psi} = \psi^\top$ is not completely positive, or from the fact that the involution $x \mapsto -x$ is non-trivial if $x \in \mathbb{Z}_d \triangleq \{0, 1, \ldots, d-1\}$ for $d > 2$.

In this work, we successfully extend Bell sampling to qudits of all dimensions d, such that the main and useful properties from the qubit case are preserved. Our generalisation is non-trivial yet simple, and at its heart it is a novel unitary that maps four copies $|\mathcal{S}\rangle^{\otimes 4}$ of a stabiliser state into their complex conjugate $|\mathcal{S}^*\rangle^{\otimes 4}$ with respect to the computational basis (up to some qudit Pauli operator), which we believe might be of independent interest. Given our new Bell sampling primitive on qudits, we then manage to lift several known results from qubits to qudits.

1.1 Bell sampling on qudits

To explain Bell (difference) sampling and our generalisation in more detail, let us introduce some notation and recall various concepts relating to the Pauli group on qudits. Let $\omega \triangleq \mathrm{e}^{\mathrm{i}2\pi/d}$

be the d-th root of unity and $\tau \triangleq (-1)^d e^{i\pi/d} = e^{i\pi(d^2+1)/d}$ such that $\tau^2 = \omega$. Let $D = d$ if d is odd and $D = 2d$ if d is even, so that D is the order of τ. The generalised Pauli operators are defined using the shift and clock operators X and Z, respectively, as

$$\mathsf{X}|q\rangle = |q+1\rangle, \quad \mathsf{Z}|q\rangle = \omega^q|q\rangle, \quad \forall q \in \mathbb{Z}_d.$$

A useful way of writing Pauli operators on n qudits is via the so-called *Weyl operators* defined as

$$\mathcal{W}_{\mathbf{x}} = \mathcal{W}_{\mathbf{v},\mathbf{w}} \triangleq \tau^{\langle \mathbf{v},\mathbf{w}\rangle}(\mathsf{X}^{v_1}\mathsf{Z}^{w_1}) \otimes \cdots \otimes (\mathsf{X}^{v_n}\mathsf{Z}^{w_n}) \qquad \forall \mathbf{x} = (\mathbf{v},\mathbf{w}) \in \mathbb{Z}_d^{2n},$$

where $\langle \mathbf{v},\mathbf{w}\rangle = \sum_{i=1}^n v_i w_i$ is the usual scalar product, considering $\mathbf{v},\mathbf{w}$ as embedded into $\mathbb{Z}^n$. One can think of Weyl operators as Pauli operators labeled by a $2n$-bit string. There is then a nice connection between operators in the Hilbert space $(\mathbb{C}^d)^{\otimes n}$ and strings in $\mathbb{Z}_d^{2n}$. The generalised Bell basis is defined as

$$|\mathcal{W}_{\mathbf{x}}\rangle \triangleq (\mathcal{W}_{\mathbf{x}} \otimes \mathbb{I})|\Phi^+\rangle, \qquad \text{where} \quad |\Phi^+\rangle \triangleq d^{-\frac{n}{2}} \sum_{\mathbf{q}\in\mathbb{Z}_d^n} |\mathbf{q},\mathbf{q}\rangle \quad \text{is a maximally entangled state.}$$

The n-qudit Pauli group $\mathscr{P}_d^n \triangleq \{\tau^s \mathcal{W}_{\mathbf{x}} : \mathbf{x} \in \mathbb{Z}_d^{2n}, s \in \mathbb{Z}\}$ is the group of all Weyl operators up to a power of τ, while the n-qudit Clifford group $\mathscr{C}_d^n \triangleq \{\mathcal{U} \in \mathbb{C}^{d^n \times d^n} : \mathcal{U}^\dagger\mathcal{U} = \mathbf{I}, \mathcal{U}\mathscr{P}_d^n\mathcal{U}^\dagger = \mathscr{P}_d^n\}$ is the group of $d^n \times d^n$ unitary operators that normalise the Pauli group. A stabilizer group $\mathcal{S} \subset \mathscr{P}_d^n$ is an abelian subgroup of the n-qudit Pauli group of size d^n.

Bell sampling on qubits [Mon17] measures two copies of an n-qubit state $|\psi\rangle = \sum_{\mathbf{q}\in\mathbb{F}_2^{2n}} \alpha_{\mathbf{q}}|\mathbf{q}\rangle$ in the Bell basis $|\mathcal{W}_{\mathbf{x}}\rangle$ to obtain a string $\mathbf{x} \in \mathbb{F}_2^{2n}$ from the distribution $2^{-n}|\langle\psi|\mathcal{W}_{\mathbf{x}}|\psi^*\rangle|^2$, where $|\psi^*\rangle = \sum_{\mathbf{q}\in\mathbb{F}_2^{2n}} \alpha_{\mathbf{q}}^*|\mathbf{q}\rangle$ is the complex conjugate of $|\psi\rangle$ with respect to the computational basis. Bell difference sampling [Mon17, GNW21] performs Bell sampling twice and computes the difference of the two outcome strings. Gross, Nezami, and Walter [GNW21] proved that Bell difference sampling on $|\psi\rangle^{\otimes 4}$ returns a string $\mathbf{x} \in \mathbb{F}_2^{2n}$ with probability

$$q_\psi(\mathbf{x}) \triangleq \sum_{\mathbf{y}\in\mathbb{F}_2^{2n}} p_\psi(\mathbf{y})p_\psi(\mathbf{x}-\mathbf{y}), \qquad \text{where} \quad p_\psi(\mathbf{x}) \triangleq 2^{-n}|\langle\psi|\mathcal{W}_{\mathbf{x}}|\psi\rangle|^2$$

is called the *characteristic distribution*. If $|\psi\rangle$ is a stabiliser state $|\mathcal{S}\rangle$ of some stabiliser group $\mathcal{S}$, then $p_{\mathcal{S}}(\mathbf{x}) = 2^{-n}$ for all strings $\mathbf{x} \in \mathbb{F}_2^{2n}$ such that $\omega^s\mathcal{W}_{\mathbf{x}} \in \mathcal{S}$ for some $s \in \mathbb{Z}_d$ and 0 otherwise. Therefore $q_{\mathcal{S}}(\mathbf{x}) = 2^{-n}$ if $\omega^s\mathcal{W}_{\mathbf{x}} \in \mathcal{S}$ and 0 otherwise. Bell difference sampling on copies of a stabiliser state thus returns an element of its stabiliser group, and that is where its power lies.

As mentioned, the generalisation of Bell (difference) sampling to d-level systems has proved challenging. Indeed, the authors [ADIS24] recently showed that the natural extension of the qubit scheme, based on measuring two copies of a given n-qudit state $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ in the generalised Bell basis, can return a uniformly random string, and thus may provide no information about the corresponding stabilizer group at all. More precisely such a procedure samples a string $\mathbf{x} \in \mathbb{Z}_d^{2n}$ with probability

$$q_\psi(\mathbf{x}) \triangleq \sum_{\mathbf{y}\in\mathbb{Z}_d^{2n}} p_\psi(\mathbf{y})p_\psi(J(\mathbf{x}-\mathbf{y})), \qquad \text{where} \quad p_\psi(\mathbf{x}) \triangleq d^{-n}|\langle\psi|\mathcal{W}_{\mathbf{x}}|\psi\rangle|^2$$

and $J : \mathbb{Z}_d^{2n} \to \mathbb{Z}_d^{2n}$ is the involution $J(\mathbf{x}) = J((\mathbf{v},\mathbf{w})) = (-\mathbf{v},\mathbf{w})$ for $\mathbf{x} = (\mathbf{v},\mathbf{w}) \in \mathbb{Z}_d^{2n}$ [App05, App09, GNW21]. Although the above expression is a clear generalisation of the qubit case, since the involution is trivial over $\mathbb{F}_2$, i.e., $J(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}_2^{2n}$, in the worst case, sampling from $q_{\mathcal{S}}$ when $|\psi\rangle = |\mathcal{S}\rangle$ is a stabiliser state returns a uniformly random string $\mathbf{x} \in \mathbb{Z}_d^{2n}$, in stark contrast to the qubit case where Bell difference sampling returns a generator of the stabiliser group. It was clear from the results of [ADIS24] that a useful generalisation of Bell difference

sampling, one which carried over important properties like sampling a generator of the stabiliser group given copies of its stabiliser state, is non-trivial.

If we are given the complex conjugate $|\psi^*\rangle$ of some state $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$, then Bell sampling on $|\psi\rangle|\psi^*\rangle$ returns a string $\mathbf{x} \in \mathbb{Z}_d^{2n}$ with probability $p_\psi(\mathbf{x}) = d^{-n}|\langle\psi|\mathcal{W}_\mathbf{x}|\psi\rangle|^2$. In other words, we get a sample from the characteristic distribution, which is exactly what we want since $p_\mathcal{S}(\mathbf{x}) = d^{-n}$ if $\omega^s\mathcal{W}_\mathbf{x} \in \mathcal{S}$ and 0 otherwise for a stabiliser state $|\mathcal{S}\rangle$. The crucial fact for qubits is that $|\mathcal{S}^*\rangle = \mathcal{W}_\mathbf{z}|\mathcal{S}\rangle$ for some Weyl operator $\mathcal{W}_\mathbf{z}$, so Bell sampling on two copies of $|\mathcal{S}\rangle \in (\mathbb{C}^2)^{\otimes n}$ is equivalent to sampling from a shifted version of the characteristic distribution $p_\mathcal{S}$. Bell difference sampling then removes this shift by $\mathbf{z}$. The same cannot be said for qudits in general, i.e., $|\mathcal{S}\rangle$ and $|\mathcal{S}^*\rangle$ are not necessarily related by a Pauli operator. If it were possible to transform $|\mathcal{S}\rangle$ into $|\mathcal{S}^*\rangle$, then Bell difference sample would work on qudits as well.

Our generalisation of Bell (difference) sampling is based exactly on this idea: transforming enough copies of a stabiliser state $|\mathcal{S}\rangle$ into its complex conjugate $|\mathcal{S}^*\rangle$. More specifically, we propose a unitary transformation that maps at most four copies of any stabiliser state, $|\mathcal{S}\rangle^{\otimes 4}$, into $|\mathcal{S}^*\rangle^{\otimes 4}$ up to a Pauli operator. Interestingly enough, our unitary is based on Lagrange's four-square theorem, which states that every positive integer can be written as the sum of four integer squares!

Result 1. *For every integer $d \geqslant 2$, let a_1, a_2, a_3, a_4 be non-negative integers such that $a_1^2 + a_2^2 + a_3^2 + a_4^2 = D-1$ coming from Lagrange's four-square theorem. Define the unitary $\mathcal{B}_\mathbf{R} \in ((\mathbb{C}^d)^{\otimes n})^{\otimes 4}$*

$$\mathcal{B}_\mathbf{R} : |\mathbf{Q}\rangle \mapsto |\mathbf{QR} \pmod d\rangle, \quad \textit{where} \quad \mathbf{Q} = [\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}_4] \in \mathbb{Z}_d^{n\times 4}$$

and $\mathbf{R} \in \mathbb{Z}^{4\times 4}$ is the matrix

$$\mathbf{R} = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ a_2 & -a_1 & a_4 & -a_3 \\ a_3 & -a_4 & -a_1 & a_2 \\ a_4 & a_3 & -a_2 & -a_1 \end{pmatrix}.$$

Given any stabiliser state $|\mathcal{S}\rangle \in (\mathbb{C}^d)^{\otimes n}$, then $\mathcal{B}_\mathbf{R}|\mathcal{S}\rangle^{\otimes 4} = \mathcal{P}(\mathcal{S})|\mathcal{S}^*\rangle^{\otimes 4}$ for some Pauli operator $\mathcal{P}(\mathcal{S})$ that depends on $|\mathcal{S}\rangle$.

The notation $|\mathbf{Q}\rangle$ is short for $|\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}_4\rangle$, while $|(\mathbf{QR})_1, (\mathbf{QR})_2, (\mathbf{QR})_3, (\mathbf{QR})_4\rangle$ is more compactly represented as $|\mathbf{QR}\rangle$, where $(\mathbf{QR})_i$ is the i-th column of $\mathbf{QR}$. More explicitly, the first n-qudit register $|\mathbf{q}_1\rangle$ is transformed into $|a_1\mathbf{q}_1 + a_2\mathbf{q}_2 + a_3\mathbf{q}_3 + a_4\mathbf{q}_4\rangle$ and similarly for the other three registers. The unitary $\mathcal{B}_\mathbf{R}$ thus mixes four copies of a state in a non-trivial way.

The crucial property behind the matrix $\mathbf{R}$ that guarantees the transformation of $|\mathcal{S}\rangle^{\otimes 4}$ into $|\mathcal{S}^*\rangle^{\otimes 4}$ is that $\mathbf{R}^\top\mathbf{R} = \mathbf{R}\mathbf{R}^\top = (D-1)\mathbf{I}$. It is well known that stabiliser states are quadratic phase states [HDDM05], meaning they are of the form $|\mathscr{A}|^{-\frac{1}{2}}\sum_{\mathbf{q}\in\mathscr{A}}\omega^{L(\mathbf{q})}\tau^{Q(\mathbf{q})}|\mathbf{q}\rangle$ where $\mathscr{A}$ is some affine submodule of $\mathbb{Z}_d^n$ and $L, Q : \mathbb{Z}^n \to \mathbb{Z}$ are linear and quadratic polynomials, respectively (see Section 2.5 for detailed expressions). The unitary $\mathcal{B}_\mathbf{R} : |\mathbf{Q}\rangle \mapsto |\mathbf{QR} \pmod d\rangle$ basically introduces, by a change of variables, the matrix $\mathbf{R}$ into the joint phases of $|\mathcal{S}\rangle^{\otimes 4}$, and the factor $\mathbf{R}^\top\mathbf{R} = (D-1)\mathbf{I}$ appears in the quadratic polynomial $Q(\mathbf{Q})$, mapping it to $(D-1)Q(\mathbf{Q})$ and so $\tau^{Q(\mathbf{Q})} \mapsto \tau^{-Q(\mathbf{Q})}$. The complex conjugate of the linear part can be obtained by some Pauli operator $\mathcal{P}(\mathcal{S})$. Together, $\mathcal{P}(\mathcal{S})\mathcal{B}_\mathbf{R}|\mathcal{S}\rangle^{\otimes 4} = |\mathcal{S}^*\rangle^{\otimes 4}$.

The number of required stabiliser copies is dictated by the number of quadratic terms needed to decompose $D-1$, so in special cases fewer than four copies can suffice. For example, in prime dimensions d there always exist $a_1, a_2 \in \mathbb{F}_d$ such that $a_1^2 + a_2^2 \equiv -1 \pmod d$, so only two copies of $|\mathcal{S}\rangle$ suffice and we employ $\mathbf{R} = \left(\begin{smallmatrix} a_1 & a_2 \\ a_2 & -a_1 \end{smallmatrix}\right)$. If moreover $d \equiv 1 \pmod 4$, -1 is a quadratic residue modulo d, meaning that only one copy of $|\mathcal{S}\rangle$ suffices. Furthermore, by picking integers $a_1, \ldots, a_4$ whose squares sum to any positive integer m (not necessarily $D-1$), the unitary $\mathcal{B}_\mathbf{R}$ can more generally be seen as transforming the quadratic part $Q(\mathbf{q})$ of any quantum phase state into $mQ(\mathbf{q})$, a fact that might be of independent interest.

Our generalised Bell sampling method then proceeds as follows: take 8 copies of a stabiliser state $|\mathcal{S}\rangle$, apply $\mathcal{B}_{\mathbf{R}}$ onto 4 out the 8 copies to obtain $|\mathcal{S}\rangle^{\otimes 4}|\mathcal{S}^*\rangle^{\otimes 4}$ up to some Pauli operators, and carry on with the standard procedure of measuring each pair $|\mathcal{S}\rangle|\mathcal{S}^*\rangle$ in the Bell basis. The result is four strings in $\mathbb{Z}_d^{2n}$. If we repeat this procedure again and take the difference between the outcomes, the four resulting strings $\mathbf{x}_1, \dots, \mathbf{x}_4 \in \mathbb{Z}_d^{2n}$ are samples from the characteristic distribution $p_{\mathcal{S}}(\mathbf{x}) = d^{-n}|\langle \mathcal{S}|\mathcal{W}_{\mathbf{x}}|\mathcal{S}\rangle|^2$ which, as mentioned above, is uniformly distributed over the strings corresponding to the stabiliser group $\mathcal{S}$. We thus successfully sample four generators of the stabiliser group through this procedure. We formalise this in the next definition.

Definition 1 (Skewed Bell difference sampling)**.** *Let $\mathcal{B}_{\mathbf{R}} \in ((\mathbb{C}^d)^{\otimes n})^{\otimes 4}$ be unitary from Result 1 and let $\mathcal{U}_{\mathbf{R}} \in ((\mathbb{C}^d)^{\otimes n})^{\otimes 8}$ be the unitary acting on $\bigotimes_{i\in[8]} |\mathbf{q}_i\rangle$ that applies $\mathcal{B}_{\mathbf{R}}$ to $|\mathbf{q}_2, \mathbf{q}_4, \mathbf{q}_6, \mathbf{q}_8\rangle$ and the identity to the remaining registers. The* skewed Bell difference sampling *on 16 quantum states in $(\mathbb{C}^d)^{\otimes n}$ is the projective measurement given by*

$$\mathcal{U}_{\mathbf{R}} \otimes \mathcal{U}_{\mathbf{R}} \left(\bigotimes_{i\in[4]} \Pi_{\mathbf{x}_i} \right) \mathcal{U}_{\mathbf{R}}^\dagger \otimes \mathcal{U}_{\mathbf{R}}^\dagger$$

$$\text{where} \qquad \Pi_{\mathbf{x}_i} = \sum_{\mathbf{y}\in\mathbb{Z}_d^{2n}} |\mathcal{W}_{\mathbf{y}}\rangle\langle\mathcal{W}_{\mathbf{y}}| \otimes |\mathcal{W}_{\mathbf{x}_i+\mathbf{y}}\rangle\langle\mathcal{W}_{\mathbf{x}_i+\mathbf{y}}|, \qquad \forall \mathbf{x}_1, \dots, \mathbf{x}_4 \in \mathbb{Z}_d^{2n}.$$

By modifying Bell difference sampling with the unitary $\mathcal{U}_{\mathbf{R}} \otimes \mathcal{U}_{\mathbf{R}}$, we can successfully sample generators of a stabiliser group $\mathcal{S}$ given copies of its stabiliser state $|\mathcal{S}\rangle$. More generally, we prove that our generalised Bell difference sampling has a useful interpretation if we are given copies of an arbitrary (not necessarily stabiliser) state $|\psi\rangle$.

Result 2. *Given $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$, skewed Bell difference sampling on $|\psi\rangle^{\otimes 16}$ corresponds to sampling from the distribution*

$$b_\psi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) = \sum_{\mathbf{Y}\in\mathbb{Z}_d^{2n\times 4}} \prod_{i=1}^{4} p_\psi(\mathbf{x}_i + \mathbf{Y}_i) p_\psi((\mathbf{YR})_i), \qquad \text{where} \quad p_\psi(\mathbf{x}) \triangleq d^{-n}|\langle\psi|\mathcal{W}_{\mathbf{x}}|\psi\rangle|^2$$

and $\mathbf{Y}_i$ and $(\mathbf{YR})_i$ denote the i-th column of the corresponding matrix. If $|\psi\rangle = |\mathcal{S}\rangle$ is a stabiliser state, then

$$b_{\mathcal{S}}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) = \begin{cases} d^{-4n} & \text{if } \omega^{s_1}\mathcal{W}_{\mathbf{x}_1}, \omega^{s_2}\mathcal{W}_{\mathbf{x}_2}, \omega^{s_3}\mathcal{W}_{\mathbf{x}_3}, \omega^{s_4}\mathcal{W}_{\mathbf{x}_4} \in \mathcal{S}, \\ 0 & \text{otherwise.} \end{cases}$$

The above result is a direct generalisation of the distribution $q_\psi(\mathbf{x}) = \sum_{\mathbf{y}\in\mathbb{F}_2^{2n}} p_\psi(\mathbf{x}+\mathbf{y})p_\psi(\mathbf{y})$ for qubits, the main difference being the presence of more factors involving the characteristic distribution and the matrix $\mathbf{R}$ which connects these factors in a non-trivial way. Notice that sampling four times from the distribution $q_\psi(\mathbf{x})$ on qubits can alternatively be written as $\sum_{\mathbf{Y}\in\mathbb{F}_2^{2n\times 4}} \prod_{i=1}^{4} p_\psi(\mathbf{x}_i + \mathbf{Y}_i)p_\psi(\mathbf{Y}_i)$.

Several properties of the distributions p_ψ and q_ψ, especially related to their Fourier transforms, were explored in [GNW21, GIKL25, GIKL24b, GIKL24c, GIKL24a]. We prove several equivalent properties of p_ψ and b_ψ for all dimensions $d \geqslant 2$ which, together with skewed Bell difference sampling itself, allows us to lift several known results from qubits to qudits of all dimensions, as explained next.

Related work. Recently, Hinsche, Eisert, and Carrasco [HEC25] introduced a general measurement primitive called *Fourier sampling*. They showed that, when applied to the stabiliser formalism, Fourier sampling reduces to the standard Bell (difference) sampling on qubits ($d = 2$) and yields new measurement primitives for dimensions $d > 2$ (their procedure is defined for prime dimensions but appears to be valid for all $d \geqslant 2$). More precisely, they introduced the POVM

defined by the projectors $\Pi_{\mathbf{x}}^{\mathrm{HEC}} \triangleq d^{-2n} \sum_{\mathbf{y} \in \mathbb{Z}_d^{2n}} \omega^{[\mathbf{y},\mathbf{x}]} \mathcal{W}_{\mathbf{y}}^{\otimes D}$ acting on $((\mathbb{C}^d)^{\otimes n})^{\otimes D}$ and showed that the probability distribution behind this POVM has interesting and desirable properties like returning generators of the stabiliser group $\mathcal{S}$ when measuring copies of its stabiliser state $|\mathcal{S}\rangle$. Compared to our results, the POVM of [HEC25] requires D copies of a state $|\psi\rangle$, while our new generalised Bell difference sampling subroutine only needs at most 8 copies of $|\psi\rangle$ at a time (4 copies for prime dimensions). The two measurement primitives yield different distributions, so one might be preferable over the other depending on the application.

1.2 Applications of Bell sampling on qudits

Bell (difference) sampling has been applied to a plethora of problems [Mon17, GNW21, HG24, HLK24, HK23, GIKL25, GIKL24b, GIKL24c, GIKL24a], almost all on qubits. By employing our generalised Bell (difference) sampling primitive, we manage to lift several of these results to qudits for all $d \geqslant 2$. These generalisations are often non-trivial since the distribution b_ψ is slightly more involved than the corresponding distribution q_ψ for qubits, given the presence of the matrix $\mathbf{R}$. Moreover, for general dimensions d one loses several properties pertaining to subspaces that are normally taken for granted, and submodules must be analysed instead. See Table 1 for a summary of our results and a comparison to previous works.

In what follows, we assume a computational model in which classical operations over the ring $\mathbb{Z}_d$, measurements in the computational basis, and single and two-qudit gates from a universal gate set [MS00, BB02, BOB05, BBO06] all require $O(1)$ time to be performed. The latter assumption is supported by the fact that there exists quantum hardware that naturally encodes information in qudits [MHZ+23].

1.2.1 Stabiliser state learning

Stabiliser states are among the classes of states that can be efficiently identified [AG08, Röt09, ZPDF16, Mon17], in contrast to arbitrary n-qubit states which, by Holevo's theorem [Hol73], require exponentially many copies (in n) to determine. Other examples of efficiently learnable states include matrix product states [CPF+10, LCLP10], non-interacting fermion states [AG23], and low-degree phase states [ABDY23].

The simplest and most efficient quantum algorithm for learning an unknown stabiliser state $|\mathcal{S}\rangle \in (\mathbb{C}^2)^{\otimes n}$ on qubits is due to Montanaro [Mon17], who employed Bell (difference) sampling to identity $|\mathcal{S}\rangle$ in $O(n^3)$ time using $O(n)$ copies of $|\mathcal{S}\rangle$ and making collective measurements across at most two copies of $|\mathcal{S}\rangle$ at a time. For qudits, we are only aware of the works [ADIS24, HEC25]. The former focuses on prime dimensions and proposes an algorithm based on techniques from hidden polynomial problems [DHIS14, IS17] to learn n-qudit stabiliser states $|\mathcal{S}\rangle \in (\mathbb{C}^d)^{\otimes n}$ in $O(n^3)$ time using $O(n)$ copies of $|\mathcal{S}\rangle$. On the other hand, the algorithm of [HEC25] applies to all dimensions and requires $O(nd)$ copies of $|\mathcal{S}\rangle$ and $O(n^3 + nd)$ time.

Using our generalised Bell difference sampling protocol, we propose a quantum algorithm that identifies an unknown stabiliser state $|\mathcal{S}\rangle \in (\mathbb{C}^d)^{\otimes n}$ for all dimensions $d \geqslant 2$.

Result 3. *For all $d \geqslant 2$, there is a quantum algorithm that identifies an unknown stabiliser state $|\mathcal{S}\rangle \in (\mathbb{C}^d)^{\otimes n}$ in $O(n^3)$ time using $O(n)$ copies of $|\mathcal{S}\rangle$. The algorithm makes collective operations across at most 8 copies of $|\mathcal{S}\rangle$ at a time.*

Result 3 not only extends the work of [ADIS24] to all dimensions beyond primes, but also improves the sample complexity of [HEC25] from $O(nd)$ down to $O(n)$. Our algorithm is based on the same underlying idea of Montanaro's algorithm: by collecting $O(n)$ samples from the stabiliser group $\mathcal{S}$, with very high probability a set of generators of these samples will also generate, and thus identify, $\mathcal{S}$. Obtaining a set of generators from all the samples can be done by casting the matrix of samples into its Smith normal form in $O(n^3)$ time [Sto96]. Finally, note that any algorithm for learning a stabiliser state requires $\Omega(n^2)$ time just to write the output.

1.2.2 Hidden Stabiliser Group Problem

Bouland, Giurgică-Tiron, and Wright [BGTW25] recently introduced the *State Hidden Subgroup Problem* (StateHSP) as the task of identifying a hidden symmetry subgroup that leaves a quantum state invariant. In a nutshell, one is given a group G with a unitary representation $R : G \to \mathcal{H}$ acting on the Hilbert space $\mathcal{H}$ and copies of a quantum state $|\psi\rangle$ such that $R(h)|\psi\rangle = |\psi\rangle$ for all elements h from a subgroup H of G, and $|\langle\psi|R(g)|\psi\rangle| < 1 - \varepsilon$ for all elements $g \notin H$. The problem is to identify H. StateHSP not only generalises the traditional Hidden Subgroup Problem, but was also used by [BGTW25] to give an efficient quantum algorithm for the hidden cut problem.

Inspired by [BGTW25], Hinsche, Eisert, and Carrasco [HEC25] defined the *Hidden Stabiliser Group Problem* where the hidden subgroup is now a stabiliser group. More precisely, one is given copies of a state $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ promised to satisfy $|\langle\psi|\mathcal{P}|\psi\rangle| < 1 - \varepsilon$ for any Pauli $\mathcal{P} \in \mathscr{P}_d^n$ that does not leave $|\psi\rangle$ invariant up a power of ω. The problem is then to identify such a set, or, equivalently, the set of Pauli operators such that $\mathcal{P}|\psi\rangle = \omega^s|\psi\rangle$ for some $s \in \mathbb{Z}_d$, which is called the *unsigned stabiliser group* of $|\psi\rangle$, $\mathrm{Weyl}(|\psi\rangle) \triangleq \{\mathbf{x} \in \mathbb{Z}_d^{2n} : \mathcal{W}_{\mathbf{x}}|\psi\rangle = \omega^s|\psi\rangle \text{ for some } s \in \mathbb{Z}_d\}$. Hinsche, Eisert, and Carrasco [HEC25], by employing their new Fourier sampling procedure, gave a quantum algorithm for the Hidden Stabiliser Group Problem that uses $O\big(\frac{nd\log d}{1-(1-\varepsilon)^d}\big) = O(n\log d\max\{d,\varepsilon^{-1}\})$ copies of $|\psi\rangle$ and runs in polynomial time.[1]

By employing our generalised Bell difference sampling procedure, we improve the results of [HEC25].

Result 4. *For all $d \geqslant 2$ with prime decomposition $d = \prod_{i=1}^{\ell} p_i^{k_i}$ and largest prime factor p_ℓ, there is a quantum algorithm that solves the Hidden Stabiliser Group Problem with high probability by identifying* $\mathrm{Weyl}(|\psi\rangle)$ *using $O\big(\frac{n}{\varepsilon}\min\{\sum_{i=1}^{\ell} k_i, \log p_\ell\}\big)$ copies of $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ in $O\big(\frac{n^3}{\varepsilon}\min\{\sum_{i=1}^{\ell} k_i, \log p_\ell\}\big)$ time. The quantum algorithm acts on 8 copies of $|\psi\rangle$ at a time.*

Result 4 improves upon the sample complexity $O(n\log d\max\{d,\varepsilon^{-1}\})$ of Hinsche, Eisert, and Carrasco when $\varepsilon = \Omega(d^{-1})$. Moreover, our algorithm only acts on at most 8 copies of $|\psi\rangle$ at a time, while [HEC25] requires D copies at a time, which can be quite demanding for large dimensions. Our algorithm is similar to the one from [HEC25] (and from Result 3 and Montanaro [Mon17] for that matter): we gather enough samples from the skewed Bell distribution b_ψ which we can show is supported on the orthogonal complement of $\mathrm{Weyl}(|\psi\rangle)$ (a rigorous definition is left to Section 2.2). Therefore, with high probability their span can be used to obtain $\mathrm{Weyl}(|\psi\rangle)$ (by putting the matrix of samples into its Smith normal form).

1.2.3 Stabiliser size property testing

Given an n-qudit state $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$, recall its *unsigned stabiliser group* $\mathrm{Weyl}(|\psi\rangle)$ as the set of Weyl operators that stabilise $|\psi\rangle$ up to some complex phase. For qubits (or qudits over prime dimensions), the dimension of $\mathrm{Weyl}(|\psi\rangle)$ is called the *stabiliser dimension*. In [GIKL25], Grewal, Iyer, Kretschmer, and Liang studied the problem of testing whether a given n-qubit state $|\psi\rangle \in (\mathbb{C}^2)^{\otimes n}$ either has stabiliser dimension at least t, or else has fidelity less than $1 - \varepsilon$ with all states with stabiliser dimension at least t, promised that one of these cases hold. By employing Bell difference sampling, they proposed a quantum algorithm that tests stabiliser dimension with high probability when $\varepsilon < \frac{3}{8}$ in $O\big(\frac{n^3}{\varepsilon}\big)$ time and using $O\big(\frac{n}{\varepsilon}\big)$ copies of $|\psi\rangle$.

The concept of dimension of a submodule is not well defined in general, with concepts like rank and length serving as possible generalisations. We find that working with such generalisations for $\mathrm{Weyl}(|\psi\rangle)$ over a general dimension d can be too cumbersome, and instead simply define the *stabiliser size* of $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ as the cardinality of $\mathrm{Weyl}(|\psi\rangle)$. We then study and

[1] In [HEC25, Theorem 6], their sample complexity is stated as $O(n\log d/(d\varepsilon))$. There is a missing factor of d since their POVM requires D copies of $|\psi\rangle$ and the approximation $(1-\varepsilon)^d \approx 1 - O(\varepsilon d)$ is valid only if $\varepsilon = O(d^{-1})$.

propose a quantum algorithm for the problem of testing whether a given state has stabiliser size at least d^t or is ε-far in fidelity from any state with stabiliser size at least d^t, promised that one of these two cases hold.

Result 5. *Let $d = \prod_{i=1}^{\ell} p_i^{k_i}$ be the prime decomposition of d and $\varepsilon = O(d^{-2})$. Let $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ be a state promised to either have stabiliser size at least d^t or be ε-far in fidelity from all such states. There is a quantum algorithm that distinguishes the two cases with high probability in $O\big(\frac{n^3}{\varepsilon}\sum_{i=1}^{\ell} k_i\big)$ time using $O\big(\frac{n}{\varepsilon}\sum_{i=1}^{\ell} k_i\big)$ copies of $|\psi\rangle$.*

Our quantum algorithm is similar to that of [GIKL25], although its analysis is more involved since we are dealing with submodules. The main idea is to gather $O\big(\frac{n}{\varepsilon}\sum_{i=1}^{\ell} k_i\big)$ samples from the distribution b_ψ through skewed Bell difference sampling and analyse the cardinality of the submodule generated by them. All the samples are within the orthogonal complement of $\mathrm{Weyl}(|\psi\rangle)$, which has size at most d^{2n-t} if we are in the case when $|\mathrm{Weyl}(|\psi\rangle)| \geqslant d^t$. If, on the other hand, $|\psi\rangle$ is ε-far away from states with stabiliser size at least d^t, we then prove that, with high probability, the cardinality of the submodule generated by the samples is greater than d^{2n-t}. Therefore, we need only to check whether the samples generate a submodule of size smaller or greater than d^{2n-t}, which can be done by putting the matrix of all samples into its Smith normal form in $O\big(\frac{n^3}{\varepsilon}\sum_{i=1}^{\ell} k_i\big)$ time. We note that, just like in [GIKL25], our result is valid for *all* choices of t in d^t. Increasing the range of $\varepsilon = O(d^{-2})$ is left as an open problem.

1.2.4 Testing doped Clifford circuits and pseudorandomness lower bounds

Pseudorandom states were introduced by Ji, Liu, and Song [JLS18] and have attracted a lot of attention due to their applicability in quantum cryptography and complexity theory [JLS18, Kre21, AQY22, MY22, HMY23, KQST23]. At a high level, a set of quantum states is pseudorandom if they mimic the Haar measure over n-qudit states and can be prepared by some efficient, $\mathrm{poly}(n)$-time, quantum algorithm (a formal definition is left to Section 7).

A few works have explored the resources required for constructing pseudorandom states, e.g., Ref. [ABF+24] explored possible constructions of pseudorandom states using limited entanglement, while Ref. [HBK25] defined and studied several other pseudorandom resources. Grewal, Iyer, Kretschmer, and Liang [GIKL23, GIKL24b], on the other hand, analysed pseudorandomness from the perspective of stabiliser complexity. More specifically, they proposed a quantum algorithm to test whether a quantum state is Haar-random or has non-negligible stabiliser fidelity. As a consequence, they proved that any Clifford+T circuit that uses $O(\log n)$ T-gates cannot generate a set of n-qubit pseudorandom quantum states [GIKL23]. This bound was later improved to $\frac{n}{2}$ non-Clifford single-qubit gates [GIKL24b] by employing Bell difference sampling to distinguish Haar-random states from doped Clifford circuit outputs. Here a t-doped Clifford circuit is a quantum circuit composed of Clifford gates and at most t non-Clifford single-qudit gates that starts in the state $|0\rangle^{\otimes n}$.

Regarding qudits, [ADIS24] generalised the work of [GIKL23] by employing the stabiliser testing POVM $\{\Pi_r^+, \mathbf{I} - \Pi_r^+\}$ of [GNW21] to efficiently distinguish t-doped Clifford circuits from Haar-random states when $t = O(\log_d n)$. As a consequence, they proved that t-doped Clifford circuits with $t = O(\log_d n)$ cannot generate pseudorandom quantum states. This bound was later re-obtained by Hinsche, Eisert, and Carrasco [HEC25] as a consequence of learning (not testing) the output state of a t-doped Clifford circuit for $t = O(\log_d n)$.

Here, by employing our skewed Bell difference sampling routine, we exponentially improve the pseudorandomness bounds on qudits of [ADIS24, HEC25] from $O(\log_d n)$ to $\frac{n}{2}$, thus matching the bound of [GIKL24b] on qubits.

Result 6. *Let $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ be a state promised to either be Haar-random or the output of a t-doped Clifford circuit for $t < \frac{n}{2}$. There is a quantum algorithm that distinguishes between the two cases with high probability in $O(n^3)$ time and using $O(n)$ copies of $|\psi\rangle$. As a consequence,*

any doped Clifford circuit in $(\mathbb{C}^d)^{\otimes n}$ *that uses less than* $\frac{n}{2}$ *non-Clifford single-qudit gates cannot produce an ensemble of pseudorandom quantum states.*

Our algorithm to distinguish Haar-random states from outputs of doped Clifford circuits is a repackaging of our algorithm for testing stabiliser size. By sampling $O(n)$ strings from the distribution b_ψ using skewed Bell difference sampling, we once again study the cardinality of their generated submodule. If $|\psi\rangle$ is Haar-random, we prove that the $O(n)$ samples generate the whole space $\mathbb{Z}_d^{2n}$ with very high probability, which cannot happen if $|\psi\rangle$ is the output of a t-doped Clifford circuit with $t < \frac{n}{2}$, since the samples coming from b_ψ are supported on a submodule or size at most d^{n+2t}.

Finally, we point out that our above result does not contradict [HIN+23], which proved that learning the output distribution of a Clifford circuit over qubits with a single T-gate is as hard as the problem of learning parities with noise. The discrepancy between results stems from the difference in access models: while we assume access to multiple copies of $|\psi\rangle$, [HIN+23] considers only algorithms that deal with computational-basis measurements on $|\psi\rangle$.

1.2.5 Stabiliser state testing

Stabiliser state testing is the problem of deciding whether a given quantum state $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ is either a stabiliser state or else is ε-far in fidelity from all stabiliser states. The measure of closeness of $|\psi\rangle$ from any stabiliser state is usually referred to as the *stabiliser fidelity* [BBC+19], i.e., $F_{\mathcal{S}}(|\psi\rangle) \triangleq \max_{|\mathcal{S}\rangle} |\langle \mathcal{S}|\psi\rangle|^2$. We are thus interested in testing whether a given state has stabiliser fidelity 1 or at most $1-\varepsilon$.

Stabiliser testing with a constant number of samples was first achieved by Gross, Nezami, and Walter [GNW21] for all dimensions $d \geqslant 2$. The authors proposed a simple testing algorithm for *qubits* which is based on obtaining $\mathbf{x} \in \mathbb{F}_2^{2n}$ via Bell difference sampling, measuring the corresponding Weyl operator $\mathcal{W}_{\mathbf{x}}$ twice on two independent copies of $|\psi\rangle$, and checking whether the outcomes are equal or not. Both outcomes will always be equal if $|\psi\rangle$ is a stabiliser state, but might differ with some probability is $|\psi\rangle$ has stabiliser fidelity at most $1-\varepsilon$. By repeating this procedure $O(1/\varepsilon)$ times, one can distinguish both cases with high probability. Their algorithm for higher dimensions $d > 2$, however, does not have a simple interpretation. Gross, Nezami, and Walter [GNW21] introduced the operator

$$\Pi_r^+ \triangleq \frac{\mathbf{I} + \mathcal{V}_r}{2} \qquad \text{where} \qquad \mathcal{V}_r \triangleq \frac{1}{d^n} \sum_{\mathbf{x} \in \mathbb{Z}_d^{2n}} (\mathcal{W}_{\mathbf{x}} \otimes \mathcal{W}_{\mathbf{x}}^\dagger)^{\otimes r} \quad \text{for } r \in \mathbb{N}$$

and showed that Π_r^+ is projector for $r \geqslant 2$ such that $\gcd(d, r) = 1$. The authors then proved that the binary POVM $\{\Pi_r^+, \mathbf{I} - \Pi_r^+\}$ has the desired property of always measuring Π_r^+ if and only if $|\psi\rangle$ is a stabiliser state. Therefore, by measuring the POVM $\{\Pi_r^+, \mathbf{I} - \Pi_r^+\}$ $O\big(\frac{d^2}{\varepsilon}\big)$ times, it is possible to distinguish between the two cases with high probability.

By using skewed Bell difference sampling, we propose an alternative stabiliser testing algorithm that, contrary to the algorithm of Gross, Nezami, and Walter, has a very simple and natural interpretation: similarly to the qubit case, one obtains a string $\mathbf{x} \in \mathbb{Z}_d^{2n}$ via skewed Bell difference sampling, measures the corresponding operators $\frac{1}{2}\mathcal{W}_{\mathbf{x}} \otimes \mathcal{W}_{\mathbf{x}}^\dagger + \frac{1}{2}\mathcal{W}_{\mathbf{x}}^\dagger \otimes \mathcal{W}_{\mathbf{x}}$ on two independent copies of $|\psi\rangle$, and checks whether the outcome is 1 or not. It is not hard to see that the outcome is always 1 if and only if $|\psi\rangle$ is a stabiliser state.

Result 7. *Let* $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ *be a state promised to have stabiliser fidelity either equal to* 1 *or at most* $1-\varepsilon$*. There is a quantum algorithm that distinguishes both cases with high probability in* $O\big(\frac{d^2}{\varepsilon}\big)$ *time using* $O\big(\frac{d^2}{\varepsilon}\big)$ *copies of* $|\psi\rangle$*.*

Although our testing algorithm has the same asymptotic time and sample complexities than [GNW21], the overall constant factors can be substantially better for small values of r

(see Section 8). Moreover, for some dimensions d, the parameter r can be substantially large, being $r = O(\log d)$ in the worst case for dimensions that are the product of all the smallest primes. The POVM $\{\Pi_r^+, \mathbf{I} - \Pi_r^+\}$ thus requires collective measurements across $2r$ copies of $|\psi\rangle$ (through phase estimation), while skewed Bell difference sampling only involves $O(1)$ copies.

We then consider the tolerant version of stabiliser state testing, where one has to distinguish states with stabiliser fidelity at least $1-\varepsilon_1$ from states with stabiliser fidelity at most $1-\varepsilon_2$, $\varepsilon_1 < \varepsilon_2$. Tolerant stabiliser testing, in contrast with the $\varepsilon_1 = 0$ case, is still poorly understood and has received a lot of attention recently, all focused on n-qubit states [GIKL24b, ABD24, HH25, IL24, BvDH25, AD25, MT25, CGYZ25]. Grewal, Iyer, Kretschmer, and Liang [GIKL24b] analysed the testing protocol of [GNW21] under the tolerant setting and proved that $O(1/\gamma^2)$ copies of $|\psi\rangle \in (\mathbb{C}^2)^{\otimes n}$ suffices to determine whether $F_{\mathcal{S}}(|\psi\rangle) \geqslant 1-\varepsilon_1$ or $F_{\mathcal{S}}(|\psi\rangle) \leqslant 1-\varepsilon_2$, where $\gamma \triangleq (1-\varepsilon_1)^6 - \frac{4-3\varepsilon_2}{4}$. Their algorithm thus works as long as $\varepsilon_1 \leqslant 1 - \left(1 - \frac{3\varepsilon_2}{4}\right)^{\frac{1}{6}} \leqslant 1 - \left(\frac{1}{2}\right)^{\frac{1}{3}}$. Follow-up works [ABD24, BvDH25, AD25, MT25, CGYZ25] extended the range of ε_1 to $1-\varepsilon_2 \leqslant (1-\varepsilon_1)^C$ for some universal constant $C > 1$ while still requiring $\mathrm{poly}((1-\varepsilon_1)^{-1})$ samples (although the current polynomial dependence can be quite bad, sometimes to the power of hundreds). Finally, Iyer and Liang [IL24] considered tolerant testing algorithms that accept mixed state inputs.

Here we initiate the study of tolerant testing algorithms for qudits of all dimensions $d \geqslant 2$. For such, we propose two different algorithms to decide whether $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ has stabiliser fidelity at least $1-\varepsilon_1$ or at most $1-\varepsilon_2$. Our first algorithm is a simple extension of the stabiliser testing protocol of [GNW21] using the POVM $\{\Pi_r^+, \mathbf{I}-\Pi_r^+\}$, while our second algorithm employs skewed Bell difference sampling in a similar way than our algorithm from Result 7. In both cases, repeated measurements of $\{\Pi_r^+, \mathbf{I}-\Pi_r^+\}$ or $\frac{1}{2}\mathcal{W}_{\mathbf{x}}\otimes\mathcal{W}_{\mathbf{x}}^\dagger + \frac{1}{2}\mathcal{W}_{\mathbf{x}}^\dagger\otimes\mathcal{W}_{\mathbf{x}}$ for $\mathbf{x} \sim b_\psi$ allow us to estimate whether $F_{\mathcal{S}}(|\psi\rangle)$ is greater than $1-\varepsilon_1$ or smaller than $1-\varepsilon_2$.

Result 8. *Let $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ and $r \geqslant 2$ such that $\gcd(d, r) = 1$. Let $\varepsilon_2 > \varepsilon_1 \geqslant 0$ such that $\gamma_r \triangleq (1-\varepsilon_1)^{2r} - 1 + \left(1 - (1 - \frac{1}{4d^2})^{r-1}\right)\varepsilon_2 > 0$. There is a quantum algorithm that decides whether $F_{\mathcal{S}}(|\psi\rangle) \geqslant 1-\varepsilon_1$ or $F_{\mathcal{S}}(|\psi\rangle) \leqslant 1-\varepsilon_2$ with high probability using $O\left(\frac{r}{\gamma_r^2}\right)$ copies of $|\psi\rangle$.*

Result 9. *Let $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ and $\varepsilon_2 > \varepsilon_1 \geqslant 0$ such that $\alpha \triangleq (1-\varepsilon_1)^{16}(1-2\varepsilon_1)^4 - \left(1 - \frac{\varepsilon_2}{2d^2}\left(1 - \frac{1}{8d^2}\right)\right)^4 > 0$. There is a quantum algorithm that decides whether $F_{\mathcal{S}}(|\psi\rangle) \geqslant 1-\varepsilon_1$ or $F_{\mathcal{S}}(|\psi\rangle) \leqslant 1-\varepsilon_2$ with high probability using $O\left(\frac{1}{\alpha^2}\right)$ copies of $|\psi\rangle$.*

Our POVM-based algorithm works as long as $(1-\varepsilon_1)^{2r} - 1 + \left(1 - (1 - \frac{1}{4d^2})^{r-1}\right)\varepsilon_2 > 0$, which implies that, for $\varepsilon_2 = 1$, $\varepsilon_1 \leqslant 1 - \left(1 - \frac{1}{4d^2}\right)^{(r-1)/2r} = O(d^{-2})$. Our Bell-sampling-based algorithm works as long as $(1-\varepsilon_1)^{16}(1-2\varepsilon_1)^4 - \left(1 - \frac{\varepsilon_2}{2d^2}\left(1 - \frac{1}{8d^2}\right)\right)^4 > 0$, which implies that, for $\varepsilon_2 = 1$, $(1-\varepsilon_1)^4(1-2\varepsilon_1) \geqslant 1 - \frac{1}{2d^2}\left(1 - \frac{1}{8d^2}\right) \implies \varepsilon_1 = O(d^{-2})$. The range of parameters $\varepsilon_1, \varepsilon_2$ and sample complexities of both algorithms are similar. As numerically shown in Section 8, however, overall the POVM-based algorithm presents better range of parameters, while the Bell-sampling-based one has a better sample complexity in general.

2 Preliminaries

Given $n \in \mathbb{N} \triangleq \{1, 2, \dots\}$, let $[n] \triangleq \{1, \dots, n\}$. Throughout this paper, let $d \in \mathbb{N}$ be $d \geqslant 2$. Let $\omega \triangleq \mathrm{e}^{\mathrm{i}2\pi/d}$ be the d-th root of unity and $\tau \triangleq (-1)^d \mathrm{e}^{\mathrm{i}\pi/d} = \mathrm{e}^{\mathrm{i}\pi(d^2+1)/d}$. Note that $\tau^2 = \omega$. Let D be the order of τ. Then $D = d$ if d is odd and $D = 2d$ is d is even. Given a group $\mathcal{G}$ and $g_1, \dots, g_k \in \mathcal{G}$, let $\mathrm{Z}(\mathcal{G}) \triangleq \{g \in \mathcal{G} : gg' = g'g, \forall g' \in \mathcal{G}\}$ be the center of $\mathcal{G}$ and let $\langle g_1, \dots, g_k\rangle$ be the subgroup of $\mathcal{G}$ generated by $g_1, \dots, g_k$.

For any state $|\psi\rangle = \sum_{\mathbf{w}\in\mathbb{Z}_d^n} a_{\mathbf{w}}|\mathbf{w}\rangle$ in $(\mathbb{C}^d)^{\otimes n}$, let $|\psi^*\rangle = \sum_{\mathbf{w}\in\mathbb{Z}_d^n} \overline{a_{\mathbf{w}}}|\mathbf{w}\rangle$ denote its complex conjugate with respect to the computational basis. We shall often write ψ to denote $|\psi\rangle\langle\psi|$. By $\mathbf{I}$ we mean the identity matrix or operator, and $\mathbf{1}[\cdot]$ is the indicator function such that $\mathbf{1}[\text{statement}]$ equals 1 if the statement is true and 0 if it is false.

Problem	Dimension d	Work	Sample Complexity	Time complexity	Remarks
Stabiliser learning	$d = 2$	[AG08]	$O(n^2)$	$O(n^3)$	Operations over 1 copy
		[ZPDF16]	$O(n)$	$O(n^3)$	For graph states Operations over 2 copies
		[Mon17]	$O(n)$	$O(n^3)$	Operations over 2 copies
	$d \geqslant 2$ prime	[ADIS24]	$O(n)$	$O(n^3)$	Operations over 3 copies
	$d \geqslant 2$	[HEC25]	$O(nd \log d)$	$O(n^3 \log d + nd \log d)$	Operations over D copies
		Thr. 57	$O(n)$	$O(n^3)$	Operations over 8 copies
Hidden Stabiliser Group	$d \geqslant 2$	[HEC25]	$O(n \max\{d, \varepsilon^{-1}\} \log d)$	$O\big((n + \frac{n^3}{d}) \max\{d, \frac{1}{\varepsilon}\} \log d\big)$	Operations over D copies
		Thr. 61	$O((n/\varepsilon) \min\{\sum_{i=1}^{\ell} k_i, \log p_\ell\})$	$O((n^3/\varepsilon) \min\{\sum_{i=1}^{\ell} k_i, \log p_\ell\})$	Operations over 8 copies
Stabiliser size testing	$d = 2$	[GIKL25]	$O(n/\varepsilon)$	$O(n^3/\varepsilon)$	$\varepsilon < 3/8$
	$d \geqslant 2$	Thr. 62	$O\big((n/\varepsilon) \sum_{i=1}^{\ell} k_i\big)$	$O\big((n^3/\varepsilon) \sum_{i=1}^{\ell} k_i\big)$	$\varepsilon = O(d^{-2})$
t-doped Clifford testing	$d = 2$	[GIKL23]	poly(n) if $t = O(\log n)$	poly(n) if $t = O(\log n)$	t-doped Cliffords for $t = O(\log n)$ cannot generate pseudorandom states; Operations over 2 copies
		[GIKL24b]	$O(n)$	$O(n^3)$	t-doped Cliffords for $t < n/2$ cannot generate pseudorandom states; Operations over 2 copies
	$d \geqslant 2$ prime	[ADIS24]	$O(d^{4t})$	$O(nd^{4t})$	t-doped Cliffords for $t = O(\log_d n)$ cannot generate pseudorandom states; Operations over 2 copies
	$d \geqslant 2$	Thr. 70	$O(n)$	$O(n^3)$	t-doped Cliffords for $t < n/2$ cannot generate pseudorandom states; Operations over 8 copies
Tolerant stabiliser testing	$d = 2$	[GIKL24b]	$O(1/\gamma^2)$ where $\gamma \triangleq (1-\varepsilon_1)^6 - \frac{4-3\varepsilon_2}{4}$	$O(n/\gamma^2)$	$\varepsilon_1 \leqslant 1 - (1 - 3\varepsilon_2/4)^{1/6}$
		[AD25]	$1/\operatorname{poly}(1-\varepsilon_1)$	$n/\operatorname{poly}(1-\varepsilon_1)$	$1 - \varepsilon_2 \leqslant (1-\varepsilon_1)^{O(1)}$
		[BvDH25]	$1/\operatorname{poly}(1-\varepsilon_1)$	$n/\operatorname{poly}(1-\varepsilon_1)$	$1 - \varepsilon_2 \leqslant (1-\varepsilon_1)^{O(1)}$
		[MT25]	$1/\operatorname{poly}(1-\varepsilon_1)$	$n/\operatorname{poly}(1-\varepsilon_1)$	$1 - \varepsilon_2 \leqslant (1-\varepsilon_1)^{O(1)}$
		[IL24]	$1/\operatorname{poly}(1-\varepsilon_1)$	$n/\operatorname{poly}(1-\varepsilon_1)$	$1 - \varepsilon_2 \leqslant (1-\varepsilon_1)^{O(1)}$ Accepts mixed state inputs
	$d \geqslant 2$	[GNW21]	$O(d^2/\varepsilon)$	$O(nd^2/\varepsilon)$	$\varepsilon_1 = 0$ and $\varepsilon_2 = \varepsilon$ Operations over $2r$ copies, $\gcd(d, r) = 1$
		Thr. 80	$O(d^2/\varepsilon)$	$O(nd^2/\varepsilon)$	$\varepsilon_1 = 0$ and $\varepsilon_2 = \varepsilon$ Operations over 8 copies
		Thr. 84	$O(r/\gamma_r^2)$ where $\gamma_r \triangleq (1-\varepsilon_1)^{2r} - 1 + \big(1 - (1 - \frac{1}{4d^2})^{r-1}\big)\varepsilon_2$	$O(nr/\gamma_r^2)$	$\varepsilon_1 = O(d^{-2})$ Operations over $2r$ copies, $\gcd(d, r) = 1$
		Thr. 85	$O(1/\alpha^2)$ where $\alpha \triangleq (1-\varepsilon_1)^{16}(1-2\varepsilon_1)^4 - \big(1 - \frac{\varepsilon_2}{2d^2}\big(1 - \frac{1}{8d^2}\big)\big)^4$	$O(n/\alpha^2)$	$\varepsilon_1 = O(d^{-2})$ Operations over 8 copies

Table 1: A comparison of our results with several known results in the literature for applications of Bell sampling. In the remarks column, *operations over m copies* indicates that quantum operations are performed on m-fold tensor products of the state under consideration. The prime decomposition of d is $d = \prod_{i=1}^{\ell} p_i^{k_i}$ with $p_1 < p_2 < \cdots < p_\ell$.

2.1 Background on rings and modules

In this paper, we shall be concerned with the ring $\mathbb{Z}_d \triangleq \mathbb{Z}/d\mathbb{Z} \cong \{0, 1, \ldots, d-1\}$ of integers modulo d. For completeness, we review some basic concepts of rings. A commutative ring R is a set equipped with two operations, addition $(+)$ and multiplication $(\cdot)$, such that R is an abelian group under addition, i.e., $+$ is associative and commutative, there is an additive identity 0, and every element has an additive inverse; R is a commutative monoid under multiplication, i.e., $\cdot$ is associative and commutative and there is a multiplicative identity $1 \neq 0$; and multiplication is distributive under addition.

Given a commutative ring R, an R-module M is an abelian group $(M, +)$ together with an operation $\cdot : R \times M \mapsto M$ (whose images we denote by rx) which satisfies the conditions $r(x+y) = rx + ry$, $(r+s)x = rx + sx$, $(rs)x = r(sx)$, and $1x = x$ for all $r, s \in R$ and $x, y \in M$. We will call the cardinality of a module M its *size*. A set $\mathscr{B} \subseteq M$ is called a generating set of M if every element of M can be expressed as a linear combination of elements in $\mathscr{B}$. If M has a

finite generating set, it is said to be finitely generated. A finite subset $\{x_{i_1}, \ldots, x_{i_k}\}$ is *linearly independent* over R if $\sum_{j=1}^{k} r_j x_{i_j} = 0 \implies r_1 = \cdots = r_k = 0$ for all $r_1, \ldots, r_k \in R$. A generating set $\mathscr{B}$ that is linearly independent over R is called a basis of M. An R-module with a basis is said to be free. The rank of a free module is the cardinality of any of its bases. The $\mathbb{Z}_d$-module $\mathbb{Z}_d^n$ is free and has rank n. However, submodules of $\mathbb{Z}_d^n$ are not necessarily free.

Given an R-module M, a non-empty set $N \subseteq M$ is called a *submodule* if N is closed under addition and multiplication. A submodule $N \subseteq M$ is called *maximal* if there are no other submodules in between N and M. A chain of submodules of M of the form $M_1 \subset M_2 \subset \cdots \subset M_\ell$ is said to have *length* ℓ, and the *length* of M is largest length of any of its chains. If an R-module M has finite length, then it is finitely generated.

Given submodules $N_1, N_2 \subseteq M$, let

$$N_1 + N_2 \triangleq \{x_1 + x_2 : x_1 \in N_1, x_2 \in N_2\} \qquad \text{and} \qquad N_1 \times N_2 \triangleq \{(x_1, x_2) : x_1 \in N_1, x_2 \in N_2\}.$$

If $M = \sum_{i=1}^{k} N_i$ and $N_j \cap \sum_{i \in [k] : i \neq j} N_i = 0$ for all $j \in [k]$, then $M = \bigoplus_{i=1}^{k} N_i$ is said to be the direct sum of $N_1, \ldots, N_k$.

Fact 2 (Fundamental theorem of finitely generated abelian groups)**.** *Given a module M with size $|M| = d$ where $d = \prod_{i=1}^{\ell} p_i^{k_i}$ is the prime decomposition of d, then $M \cong \bigoplus_{i=1}^{\ell} M_i$ where $|M_i| = p_i^{k_i}$ and M_i decomposes as $M_i \cong \bigoplus_{j=1}^{r_i} \mathbb{Z}/p_i^{\lambda_{ij}}\mathbb{Z}$ such that $1 \leqslant \lambda_{ij} \leqslant k_i$ and $\sum_{j=1}^{r_i} \lambda_{ij} = k_i$. If M is a submodule of a free module with rank n, then $r_i \leqslant n$ in the decomposition of M.*

2.2 Scalar product modules and symplectic modules

For any $L \in \mathbb{N}$, define the *scalar product* $\langle \cdot, \cdot \rangle_L : \mathbb{Z}^n \times \mathbb{Z}^n \to \mathbb{Z}$, $(\mathbf{v}, \mathbf{w}) \mapsto \sum_i v_i w_i \bmod L$ and the *symplectic product* $[\cdot, \cdot]_L : \mathbb{Z}^{2n} \times \mathbb{Z}^{2n} \to \mathbb{Z}$, $(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} (x_i y_{n+i} - x_{n+i} y_i) \bmod L$. These definitions extend to scalar and symplectic products acting on $\mathbb{Z}_d^n \times \mathbb{Z}_d^n$ and $\mathbb{Z}_d^{2n} \times \mathbb{Z}_d^{2n}$ by considering the natural embedding of $\mathbb{Z}_d$ into $\mathbb{Z}$. We consider the $\mathbb{Z}_d^n$ module equipped with the scalar product $\langle \cdot, \cdot \rangle \triangleq \langle \cdot, \cdot \rangle_d$ and the module $\mathbb{Z}_d^{2n}$ equipped with the *symplectic product* $[\cdot, \cdot] \triangleq [\cdot, \cdot]_d$. It should be clear from the context whether we are considering a scalar product or a symplectic module. The scalar product $\langle \cdot, \cdot \rangle$ and the symplectic product $[\cdot, \cdot]$ are both *non-degenerate*, meaning that the zero module element $\mathbf{0}$ is the only element orthogonal to every element in the module.

Definition 3 (Involution)**.** *The involution $J : \mathbb{Z}_d^{2n} \to \mathbb{Z}_d^{2n}$ on the symplectic module $\mathbb{Z}_d^{2n}$ is defined as $J(\mathbf{x}) = J((\mathbf{v}, \mathbf{w})) = (-\mathbf{v}, \mathbf{w})$ where $\mathbf{x} = (\mathbf{v}, \mathbf{w})$ and $\mathbf{v}, \mathbf{w} \in \mathbb{Z}_d^n$.*

Let $\mathbf{e}_i \in \mathbb{Z}_d^n$ be the module element whose i-th component is 1 and the remaining components are 0. Let vec be the vectorisation operator defined by $\operatorname{vec}(|\mathbf{v}\rangle\langle\mathbf{w}|) = |\mathbf{v}\rangle|\mathbf{w}\rangle$ for computational basis states $\mathbf{v}, \mathbf{w} \in \mathbb{Z}_d^n$. Note that vec preserves inner products, i.e., $\langle \operatorname{vec}(\mathbf{A}) | \operatorname{vec}(\mathbf{B}) \rangle = \operatorname{Tr}[\mathbf{A}^\dagger \mathbf{B}]$ for all $\mathbf{A}, \mathbf{B} \in \mathbb{Z}_d^{m \times n}$. Given a matrix $\mathbf{M} \in \mathbb{Z}_d^{m \times n}$, let

$$\begin{aligned}
\operatorname{col}(\mathbf{M}) &\triangleq \{\mathbf{M}\mathbf{v} \in \mathbb{Z}_d^m : \mathbf{v} \in \mathbb{Z}_d^n\} && \text{be its column submodule,} \\
\operatorname{row}(\mathbf{M}) &\triangleq \operatorname{col}(\mathbf{M}^\top) && \text{be its row submodule,} \\
\operatorname{null}(\mathbf{M}) &\triangleq \{\mathbf{v} \in \mathbb{Z}_d^n : \mathbf{M}\mathbf{v} = \mathbf{0}\} && \text{be its null submodule.}
\end{aligned}$$

Given a submodule $\mathscr{V} \subseteq \mathbb{Z}_d^n$ of the scalar product module $\mathbb{Z}_d^n$, its *orthogonal complement* is $\mathscr{V}^\perp \triangleq \{\mathbf{w} \in \mathbb{Z}_d^n : \langle \mathbf{v}, \mathbf{w} \rangle = 0, \forall \mathbf{v} \in \mathscr{V}\}$. Given a submodule $\mathscr{X} \subseteq \mathbb{Z}_d^{2n}$ of the symplectic module $\mathbb{Z}_d^{2n}$, its *symplectic complement* is $\mathscr{X}^\perp \triangleq \{\mathbf{y} \in \mathbb{Z}_d^{2n} : [\mathbf{x}, \mathbf{y}] = 0, \forall \mathbf{x} \in \mathscr{X}\}$. We shall use the following facts about orthogonal and symplectic complements throughout.

Fact 4. *Let $\mathscr{V}, \mathscr{W} \subseteq \mathbb{Z}_d^n$ and $\mathscr{X}, \mathscr{Y} \subseteq \mathbb{Z}_d^{2n}$ be submodules of the scalar product and symplectic modules $\mathbb{Z}_d^n$ and $\mathbb{Z}_d^{2n}$, respectively. Then*

a. $\mathscr{V}^\perp$ *is a submodule;*

b. $(\mathscr{V}^\perp)^\perp = \mathscr{V}$*;*

c. $|\mathscr{V}||\mathscr{V}^\perp| = d^n$*;*

d. $\mathscr{V} \subseteq \mathscr{W} \iff \mathscr{W}^\perp \subseteq \mathscr{V}^\perp$*;*

e. $(\mathscr{V} + \mathscr{W})^\perp = \mathscr{V}^\perp \cap \mathscr{W}^\perp$*;*

f. $\mathscr{X}^\perp$ *is a submodule;*

g. $(\mathscr{X}^\perp)^\perp = \mathscr{X}$*;*

h. $|\mathscr{X}||\mathscr{X}^\perp| = d^{2n}$*;*

i. $\mathscr{X} \subseteq \mathscr{Y} \iff \mathscr{Y}^\perp \subseteq \mathscr{X}^\perp$*;*

j. $(\mathscr{X} + \mathscr{Y})^\perp = \mathscr{X}^\perp \cap \mathscr{Y}^\perp$.

Lemma 5. *Given* $\mathbf{M} \in \mathbb{Z}_d^{m\times n}$*, then* $\mathrm{null}(\mathbf{M})^\perp = \mathrm{row}(\mathbf{M})$.

Proof. Let $\mathbf{x} \in \mathrm{row}(\mathbf{M})$. Then $\mathbf{x} = \mathbf{M}^\top\mathbf{y}$ for some $\mathbf{y} \in \mathbb{Z}_d^m$. For all $\mathbf{z} \in \mathrm{null}(\mathbf{M})$, $\langle \mathbf{z}, \mathbf{x}\rangle = \langle \mathbf{z}, \mathbf{M}^\top\mathbf{y}\rangle = \langle \mathbf{M}\mathbf{z}, \mathbf{y}\rangle = 0$, thus $\mathbf{x} \in \mathrm{null}(\mathbf{M})^\perp$ and so $\mathrm{row}(\mathbf{M}) \subseteq \mathrm{null}(\mathbf{M})^\perp$. On the other hand, let $\mathbf{x} \in \mathrm{row}(\mathbf{M})^\perp$. Then $\langle \mathbf{x}, \mathbf{M}^\top\mathbf{y}\rangle = 0$ for all $\mathbf{y} \in \mathbb{Z}_d^n$. Since $0 = \langle \mathbf{x}, \mathbf{M}^\top\mathbf{y}\rangle = \langle \mathbf{M}\mathbf{x}, \mathbf{y}\rangle$ and the scalar product $\langle \cdot, \cdot \rangle$ is non-degenerate, that means that $\mathbf{M}\mathbf{x} = \mathbf{0}$ and hence $\mathbf{x} \in \mathrm{null}(\mathbf{M})$, therefore $\mathrm{row}(\mathbf{M})^\perp \subseteq \mathrm{null}(\mathbf{M}) \iff \mathrm{null}(\mathbf{M})^\perp \subseteq \mathrm{row}(\mathbf{M})$. ∎

The concept of isotropic and Lagrangian submodules will be fundamental to our analysis.

Definition 6. *A submodule* $\mathscr{X} \subset \mathbb{Z}_d^{2n}$ *of a symplectic module* $\mathbb{Z}_d^{2n}$ *is* isotropic *if* $[\mathbf{x}, \mathbf{y}] = 0$ *for all* $\mathbf{x}, \mathbf{y} \in \mathscr{X}$*, and is* Lagrangian *if* $\mathscr{X}^\perp = \mathscr{X}$.

Fact 7. *Every Lagrangian submodule of* $\mathbb{Z}_d^{2n}$ *has size* d^n*. Every isotropic submodule of a symplectic module can be extended to a Lagrangian one.*

Fact 8. *For any isotropic submodule* $\mathscr{X} \subset \mathbb{Z}_d^{2n}$*,* $\mathscr{X} \subseteq \mathscr{X}^\perp$.

The following fact will be useful.

Lemma 9. *Let* $\mathscr{V} \subseteq \mathbb{Z}_d^n$ *and* $\mathscr{X} \subseteq \mathbb{Z}_d^{2n}$ *be submodules and* $\mathbf{w} \in \mathbb{Z}_d^n$ *and* $\mathbf{y} \in \mathbb{Z}_d^{2n}$*. Then*

$$\sum_{\mathbf{v}\in\mathscr{V}} \omega^{\langle \mathbf{v},\mathbf{w}\rangle} = \begin{cases} |\mathscr{V}| & \text{if } \mathbf{w} \in \mathscr{V}^\perp, \\ 0 & \text{if } \mathbf{w} \notin \mathscr{V}^\perp, \end{cases} \qquad \sum_{\mathbf{x}\in\mathscr{X}} \omega^{[\mathbf{x},\mathbf{y}]} = \begin{cases} |\mathscr{X}| & \text{if } \mathbf{y} \in \mathscr{X}^\perp, \\ 0 & \text{if } \mathbf{y} \notin \mathscr{X}^\perp. \end{cases}$$

2.3 Symplectic Fourier analysis

Boolean Fourier analysis [Wol08, O'D14] is a well-established field wherein a Boolean function is studied through its Fourier expansion using the characters $\omega^{\langle \mathbf{v},\cdot\rangle}$ defined with respect to the scalar product. A character of $\mathbb{Z}_d^{2n}$ is a group homomorphism $\chi : \mathbb{Z}_d^{2n} \to \mathbb{C}$ such that $\chi(\mathbf{x} + \mathbf{y}) = \chi(\mathbf{x})\chi(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{Z}_d^{2n}$. It is a standard result that every character of the symplectic module $\mathbb{Z}_d^{2n}$ can be written as $\omega^{[\mathbf{x},\cdot]}$ for some $\mathbf{x} \in \mathbb{Z}_d^{2n}$ (see [Gro06, Appendix IX.C]). In this paper, we shall work with *symplectic* Fourier analysis, which is similar to the usual Boolean Fourier analysis but the Fourier characters are instead defined with respect to the symplectic product as $\omega^{[\mathbf{x},\cdot]}$.

Definition 10. *The symplectic Fourier transform* $\widehat{f} : \mathbb{Z}_d^{2n} \to \mathbb{C}$ *of* $f : \mathbb{Z}_d^{2n} \to \mathbb{C}$ *is*

$$\widehat{f}(\mathbf{y}) \triangleq \frac{1}{d^{2n}} \sum_{\mathbf{x}\in\mathbb{Z}_d^{2n}} \omega^{[\mathbf{y},\mathbf{x}]} f(\mathbf{x}).$$

Using Lemma 9, one can expand any function $f : \mathbb{Z}_d^{2n} \to \mathbb{C}$ using its symplectic Fourier transform as

$$f(\mathbf{x}) = \sum_{\mathbf{y}\in\mathbb{Z}_d^{2n}} \omega^{[\mathbf{x},\mathbf{y}]} \widehat{f}(\mathbf{y}).$$

The transformation $f \mapsto \widehat{f}$ is unitary, meaning that (a generalised) Parseval's identity holds.

Lemma 11 (Parseval's identity)**.** *Given $f, g : \mathbb{Z}_d^{2n} \to \mathbb{C}$ and $\mathbf{t} \in \mathbb{Z}_d^{2n}$, then*

$$\frac{1}{d^{2n}} \sum_{\mathbf{x} \in \mathbb{Z}_d^{2n}} \omega^{[\mathbf{t},\mathbf{x}]} f(\mathbf{x}) g(\mathbf{x}) = \sum_{\mathbf{y} \in \mathbb{Z}_d^{2n}} \widehat{f}(\mathbf{y}) \widehat{g}(\mathbf{t} - \mathbf{y}).$$

Proof. $$\sum_{\mathbf{x} \in \mathbb{Z}_d^{2n}} \omega^{[\mathbf{t},\mathbf{x}]} f(\mathbf{x}) g(\mathbf{x}) = \sum_{\mathbf{x},\mathbf{y},\mathbf{z} \in \mathbb{Z}_d^{2n}} \omega^{[\mathbf{x},\mathbf{y}+\mathbf{z}-\mathbf{t}]} \widehat{f}(\mathbf{y}) \widehat{g}(\mathbf{z})$$
$$= d^{2n} \sum_{\mathbf{y},\mathbf{z} \in \mathbb{Z}_d^{2n}} \widehat{f}(\mathbf{y}) \widehat{g}(\mathbf{z}) \cdot \mathbf{1}[\mathbf{z} = \mathbf{t} - \mathbf{y}] = d^{2n} \sum_{\mathbf{y} \in \mathbb{Z}_d^{2n}} \widehat{f}(\mathbf{y}) \widehat{g}(\mathbf{t} - \mathbf{y}). \qquad \blacksquare$$

The usual Parseval's identity follows from the above lemma by taking $\mathbf{t} = \mathbf{0}$ and $g(\mathbf{x}) = \overline{f(\mathbf{x})}$. We also define the convolution operation.

Definition 12 (Convolution)**.** *Let $f, g : \mathbb{Z}_d^{2n} \to \mathbb{C}$. The convolution between f and g is the function $f * g : \mathbb{Z}_d^{2n} \to \mathbb{C}$ defined by*

$$(f * g)(\mathbf{x}) \triangleq \frac{1}{d^{2n}} \sum_{\mathbf{y} \in \mathbb{Z}_d^{2n}} f(\mathbf{y}) g(\mathbf{x} - \mathbf{y}) = \frac{1}{d^{2n}} \sum_{\mathbf{y} \in \mathbb{Z}_d^{2n}} f(\mathbf{x} - \mathbf{y}) g(\mathbf{y}).$$

Lemma 13. *Let $f, g : \mathbb{Z}_d^{2n} \to \mathbb{C}$. Then, for all $\mathbf{x} \in \mathbb{Z}_d^{2n}$, $\widehat{f * g}(\mathbf{x}) = \widehat{f}(\mathbf{x}) \widehat{g}(\mathbf{x})$.*

Proof. $$\widehat{f * g}(\mathbf{x}) = \frac{1}{d^{2n}} \sum_{\mathbf{y} \in \mathbb{Z}_d^{2n}} \omega^{[\mathbf{x},\mathbf{y}]} (f * g)(\mathbf{y}) = \frac{1}{d^{4n}} \sum_{\mathbf{y},\mathbf{z} \in \mathbb{Z}_d^{2n}} \omega^{[\mathbf{x},\mathbf{y}]} f(\mathbf{z}) g(\mathbf{y} - \mathbf{z})$$
$$= \frac{1}{d^{4n}} \sum_{\mathbf{z} \in \mathbb{Z}_d^{2n}} \omega^{[\mathbf{x},\mathbf{z}]} f(\mathbf{z}) \sum_{\mathbf{y} \in \mathbb{Z}_d^{2n}} \omega^{[\mathbf{x},\mathbf{y}-\mathbf{z}]} g(\mathbf{y} - \mathbf{z}) = \widehat{f}(\mathbf{x}) \widehat{g}(\mathbf{x}). \qquad \blacksquare$$

2.4 Generalised Pauli group and Weyl operators

Define the unitary shift and clock operators X and Z, respectively, as

$$\mathsf{X}|q\rangle \triangleq |q+1\rangle, \quad \mathsf{Z}|q\rangle \triangleq \omega^q |q\rangle, \quad \forall q \in \mathbb{Z}_d.$$

The Pauli group on 1 qudit is the group generated by the clock and shift operators and powers of τ, i.e., $\mathscr{P}_d = \langle \tau \mathbf{I}, \mathsf{X}, \mathsf{Z} \rangle$. The n-qudit Pauli group $\mathscr{P}_d^n$ over $\mathbb{Z}_d$ (also known as Weyl-Heisenberg group) is the group generated by tensor products of clock and shift operators, together with powers of τ, i.e.,

$$\mathscr{P}_d^n \triangleq \langle \tau \mathbf{I}, \mathsf{X}, \mathsf{Z} \rangle^{\otimes n}.$$

The corresponding n-qudit Clifford group $\mathscr{C}_d^n$ is defined as the normaliser of the Pauli group in the unitary group,

$$\mathscr{C}_d^n \triangleq \{\mathcal{U} \in \mathbb{C}^{d^n \times d^n} : \mathcal{U}^\dagger \mathcal{U} = \mathbf{I},\ \mathcal{U} \mathscr{P}_d^n \mathcal{U}^\dagger = \mathscr{P}_d^n\}.$$

A Pauli gate is any element from $\mathscr{P}_d^n$ and a Clifford gate is any element from $\mathscr{C}_d^n$. More generally, an n-qudit quantum gate is any unitary in $\mathbb{C}^{d^n \times d^n}$. We assume that measuring a qudit on the computational basis and single and two-qudit quantum gates from a universal gate set [MS00, BB02, BOB05, BBO06] all require constant time to be performed. An example of a universal gate set is the set of all single-qudit gates plus the controlled shift operator $\mathbf{I}_{d(d-1)} \otimes \mathsf{X}$ [BBO06].

Another way to characterise the Pauli group $\mathscr{P}_d^n$ is through the phase space picture of finite-dimensional quantum mechanics [Woo87, App05, Gro06, Bea13]. The main objects here are the *Weyl operators* (also known as the generalised Pauli operators) defined as

$$\mathcal{W}_{\mathbf{x}} = \mathcal{W}_{\mathbf{v},\mathbf{w}} = \tau^{\langle \mathbf{v},\mathbf{w} \rangle_D} (\mathsf{X}^{w_1} \mathsf{Z}^{v_1}) \otimes \cdots \otimes (\mathsf{X}^{w_n} \mathsf{Z}^{v_n}), \qquad \forall \mathbf{x} = (\mathbf{v}, \mathbf{w}) \in \mathbb{Z}_d^{2n}.$$

The action of $\mathcal{W}_{\mathbf{x}}$ on the Hilbert space $(\mathbb{C}^d)^{\otimes n}$ is

$$\mathcal{W}_{\mathbf{x}}|\mathbf{q}\rangle = \mathcal{W}_{\mathbf{v},\mathbf{w}}|\mathbf{q}\rangle = \tau^{\langle \mathbf{v},\mathbf{w}\rangle_D}\omega^{\langle \mathbf{q},\mathbf{v}\rangle}|(\mathbf{q}+\mathbf{w}) \pmod d\rangle, \qquad \forall \mathbf{q} \in \mathbb{Z}_d^n, \quad \forall \mathbf{x} = (\mathbf{v},\mathbf{w}) \in \mathbb{Z}_d^{2n}.$$

It is clear that every Weyl operator is an element of the Pauli group $\mathscr{P}_d^n$. Conversely, every Pauli operator is equal to some Weyl operator up to a phase that is a power of τ. Therefore,

$$\mathscr{P}_d^n = \{\tau^s \mathcal{W}_{\mathbf{x}} : \mathbf{x} \in \mathbb{Z}_d^{2n},\ s \in \mathbb{Z}_D\}.$$

The following properties hold for the Weyl operators.

Fact 14. *For any $\mathbf{x}, \mathbf{y} \in \mathbb{Z}_d^{2n}$,*

a. $\mathcal{W}_{\mathbf{x}}\mathcal{W}_{\mathbf{y}} = \tau^{[\mathbf{x},\mathbf{y}]_D}\mathcal{W}_{\mathbf{x}+\mathbf{y}} = \omega^{[\mathbf{x},\mathbf{y}]}\mathcal{W}_{\mathbf{y}}\mathcal{W}_{\mathbf{x}}$*;*

b. $\mathcal{W}_{\mathbf{x}}$ *and* $\mathcal{W}_{\mathbf{y}}$ *commute if and only if* $[\mathbf{x},\mathbf{y}] = 0$*.*

c. $\mathcal{W}_{\mathbf{x}}^m = \mathcal{W}_{m\mathbf{x}} \quad \forall m \in \mathbb{Z}_d$*. In particular,* $\mathcal{W}_{\mathbf{x}}^\dagger = \mathcal{W}_{-\mathbf{x}}$ *and* $\mathcal{W}_{\mathbf{x}}^d = \mathbf{I}$*;*

d. $\operatorname{Tr}[\mathcal{W}_{\mathbf{y}}^\dagger \mathcal{W}_{\mathbf{x}}] = \operatorname{Tr}[\mathcal{W}_{\mathbf{y}}\mathcal{W}_{\mathbf{x}}^\dagger] = d^n \cdot \mathbf{1}[\mathbf{x} \equiv \mathbf{y}]$*;*

e. The operators $\{d^{-\frac{n}{2}}\mathcal{W}_{\mathbf{x}} : \mathbf{x} \in \mathbb{Z}_d^{2n}\}$ form an orthonormal basis with respect to the Hilbert-Schmidt inner product $\langle \mathcal{A}, \mathcal{B}\rangle = \operatorname{Tr}[\mathcal{A}^\dagger\mathcal{B}]$.

To ease notation, scalar and symplectic products $\langle \cdot,\cdot\rangle_D$ and $[\cdot,\cdot]_D$ in powers of τ will simply be denoted as $\langle \cdot,\cdot\rangle$ and $[\cdot,\cdot]$.

Any operator $\mathcal{A}$ on $(\mathbb{C}^d)^{\otimes n}$ can be expanded as

$$\mathcal{A} = d^{-\frac{n}{2}} \sum_{\mathbf{x}\in\mathbb{Z}_d^{2n}} c_{\mathcal{A}}(\mathbf{x})\mathcal{W}_{\mathbf{x}} \quad \text{where } c_{\mathcal{A}} : \mathbb{Z}_d^{2n} \to \mathbb{C} \text{ given by } c_{\mathcal{A}}(\mathbf{x}) = d^{-\frac{n}{2}} \operatorname{Tr}[\mathcal{W}_{\mathbf{x}}^\dagger \mathcal{A}]$$

is the *characteristic function* of the operator $\mathcal{A}$. We note that $\operatorname{Tr}[\mathcal{A}^\dagger\mathcal{B}] = \sum_{\mathbf{x}\in\mathbb{Z}_d^{2n}} \overline{c_{\mathcal{A}}(\mathbf{x})}c_{\mathcal{B}}(\mathbf{x})$. In particular, for any quantum state $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$,

$$\psi \triangleq |\psi\rangle\langle\psi| = d^{-\frac{n}{2}} \sum_{\mathbf{x}\in\mathbb{Z}_d^{2n}} c_\psi(\mathbf{x})\mathcal{W}_{\mathbf{x}} \quad \text{where } c_\psi(\mathbf{x}) = d^{-\frac{n}{2}} \operatorname{Tr}[\mathcal{W}_{\mathbf{x}}^\dagger\psi] = d^{-\frac{n}{2}}\langle\psi|\mathcal{W}_{\mathbf{x}}^\dagger|\psi\rangle.$$

Note that $\overline{c_\psi(\mathbf{x})} = c_\psi(-\mathbf{x})$ since $\mathcal{W}_{\mathbf{x}}^\dagger = \mathcal{W}_{-\mathbf{x}}$.

From the characteristic function of $\psi = |\psi\rangle\langle\psi|$ we define its *characteristic distribution*

$$p_\psi : \mathbb{Z}_d^{2n} \to [0, d^{-n}], \quad p_\psi(\mathbf{x}) = |c_\psi(\mathbf{x})|^2 = d^{-n}|\langle\psi|\mathcal{W}_{\mathbf{x}}|\psi\rangle|^2 = d^{-n}\operatorname{Tr}[\psi\mathcal{W}_{\mathbf{x}}\psi\mathcal{W}_{\mathbf{x}}^\dagger].$$

Note that $p_\psi(-\mathbf{x}) = p_\psi(\mathbf{x})$ and that $p_\psi(\mathbf{x}) \leqslant d^{-n}$ since $|\langle\psi|\mathcal{W}_{\mathbf{x}}|\psi\rangle| \leqslant 1$ for all $\mathbf{x} \in \mathbb{Z}_d^{2n}$. To check that p_ψ is indeed a probability distribution, note that $\sum_{\mathbf{x}\in\mathbb{Z}_d^{2n}} p_\psi(\mathbf{x}) = \sum_{\mathbf{x}\in\mathbb{Z}_d^{2n}} \overline{c_\psi(\mathbf{x})}c_\psi(\mathbf{x}) = \operatorname{Tr}[|\psi\rangle\langle\psi|\psi\rangle\langle\psi|] = 1$. The next result, which is a slight generalisation of [GNW21, Eq. (3.5)] for $d > 2$, shows that p_ψ is invariant (up to renormalisation) under the symplectic Fourier transform.

Lemma 15. *For any $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ and $\mathbf{x} \in \mathbb{Z}_d^{2n}$, $\widehat{p_\psi}(\mathbf{x}) = d^{-n}p_\psi(\mathbf{x})$.*

Proof.

$$\begin{aligned}
\widehat{p_\psi}(\mathbf{x}) &= \frac{1}{d^{2n}} \sum_{\mathbf{y}\in\mathbb{Z}_d^{2n}} \omega^{[\mathbf{x},\mathbf{y}]}p_\psi(\mathbf{y}) = \frac{1}{d^{3n}} \sum_{\mathbf{y}\in\mathbb{Z}_d^{2n}} \omega^{[\mathbf{x},\mathbf{y}]}\langle\psi|\mathcal{W}_{\mathbf{y}}|\psi\rangle\langle\psi|\mathcal{W}_{\mathbf{y}}^\dagger|\psi\rangle \\
&= \frac{1}{d^{3n}} \sum_{\mathbf{y}\in\mathbb{Z}_d^{2n}} \langle\psi|\mathcal{W}_{\mathbf{y}}|\psi\rangle\langle\psi|\mathcal{W}_{\mathbf{x}}^\dagger\mathcal{W}_{\mathbf{y}}^\dagger\mathcal{W}_{\mathbf{x}}|\psi\rangle = \frac{1}{d^{2n}} \sum_{\mathbf{y}\in\mathbb{Z}_d^{2n}} \overline{c_\psi(\mathbf{y})}c_{\mathcal{W}_{\mathbf{x}}\psi\mathcal{W}_{\mathbf{x}}^\dagger}(\mathbf{y}) \\
&= \frac{1}{d^{2n}} \operatorname{Tr}[\psi\mathcal{W}_{\mathbf{x}}\psi\mathcal{W}_{\mathbf{x}}^\dagger] \qquad \text{(by Parseval's identity)} \\
&= d^{-n}p_\psi(\mathbf{x}).
\end{aligned}$$

■

Moreover, the next identity regarding the mass on a submodule $\mathscr{X} \subseteq \mathbb{Z}_d^{2n}$ under p_ψ is true.

Lemma 16. *Let $\mathscr{X} \subseteq \mathbb{Z}_d^{2n}$ be a submodule. Then*

$$\sum_{\mathbf{x}\in\mathscr{X}} p_\psi(\mathbf{x}) = \frac{|\mathscr{X}|}{d^n} \sum_{\mathbf{x}\in\mathscr{X}^\perp} p_\psi(\mathbf{x}).$$

Proof. Using Lemma 9 and 15,

$$\sum_{\mathbf{x}\in\mathscr{X}} p_\psi(\mathbf{x}) = \sum_{\mathbf{x}\in\mathscr{X}} \sum_{\mathbf{y}\in\mathbb{Z}_d^{2n}} \omega^{[\mathbf{x},\mathbf{y}]}\widehat{p_\psi}(\mathbf{y}) = \frac{1}{d^n} \sum_{\mathbf{x}\in\mathscr{X}} \sum_{\mathbf{y}\in\mathbb{Z}_d^{2n}} \omega^{[\mathbf{x},\mathbf{y}]} p_\psi(\mathbf{y}) = \frac{|\mathscr{X}|}{d^n} \sum_{\mathbf{y}\in\mathscr{X}^\perp} p_\psi(\mathbf{y}). \qquad \blacksquare$$

2.5 Stabiliser groups

We now review *stabiliser groups*, which are among the most important subgroups of $\mathscr{P}_d^n$. For completeness and the reader's convenience, we provide proofs for a number of important claims. We point the reader to [App05, Gro06, App09, Bea13, GNW21] for more information.

Definition 17 (Partial stabiliser group)**.** *A partial stabiliser group $\mathcal{X}$ is a subgroup of $\mathscr{P}_d^n$ which contains only the identity from the center of $\mathscr{P}_d^n$, i.e., $\mathcal{X} \cap \mathrm{Z}(\mathscr{P}_d^n) = \{\mathbf{I}\}$. A stabiliser group $\mathcal{S}$ is a partial stabiliser group which is also maximal.*

Lemma 18. *Any partial stabiliser group $\mathcal{X} \subset \mathscr{P}_d^n$ can be written as $\mathcal{X} = \{\omega^{s(\mathbf{x})}\mathcal{W}_\mathbf{x} : \mathbf{x} \in \mathscr{X}\}$, where $\mathscr{X} \subset \mathbb{Z}_d^{2n}$ is an isotropic submodule and $s : \mathbb{Z}_d^{2n} \to \mathbb{Z}_d$. If d is odd, then s is linear, i.e., $s(\mathbf{x}+\mathbf{y}) = s(\mathbf{x}) + s(\mathbf{y})$. If d is even, then $\omega^{s(\mathbf{x}+\mathbf{y})} = \omega^{s(\mathbf{x})+s(\mathbf{y})}\tau^{[\mathbf{x},\mathbf{y}]}$. If $\mathcal{X}$ is a stabiliser group, then $\mathscr{X}$ is also Lagrangian and $|\mathscr{X}| = d^n$.*

Proof. First notice that, for d odd, any operator $\tau^s\mathcal{W}_\mathbf{x}$ can be expressed as $\omega^{s'}\mathcal{W}_\mathbf{x}$ for $s' = 2^{-1}s$, where 2^{-1} is the multiplicative inverse of 2 mod d. For d even, no operator $\tau^s\mathcal{W}_\mathbf{x}$ for s odd can belong to $\mathcal{X}$ otherwise $-\mathbf{I} \in \mathcal{X}$, so all elements of $\mathcal{X}$ are of the form $\omega^s\mathcal{W}_\mathbf{x}$ for any d. There cannot be two operators $\omega^s\mathcal{W}_\mathbf{x}, \omega^{s'}\mathcal{W}_\mathbf{x} \in \mathcal{X}$ with $s \neq s'$ since $\mathcal{X} \cap \mathrm{Z}(\mathscr{P}_d^n) = \{\mathbf{I}\}$. Thus $\mathcal{X} = \{\omega^{s(\mathbf{x})}\mathcal{W}_\mathbf{x} : \mathbf{x} \in \mathscr{X}\}$ for some set $\mathscr{X} \subset \mathbb{Z}_d^{2n}$ and $s : \mathscr{X} \to \mathbb{Z}_d$. That $\mathscr{X}$ is a submodule follows from $\mathcal{X}$ being a group. It is isotropic because, given $\omega^{s(\mathbf{x})}\mathcal{W}_\mathbf{x}, \omega^{s(\mathbf{y})}\mathcal{W}_\mathbf{y} \in \mathcal{X}$, then $(\omega^{s(\mathbf{x})}\mathcal{W}_\mathbf{x})^\dagger(\omega^{s(\mathbf{y})}\mathcal{W}_\mathbf{y})^\dagger(\omega^{s(\mathbf{x})}\mathcal{W}_\mathbf{x})(\omega^{s(\mathbf{y})}\mathcal{W}_\mathbf{y}) = \omega^{[\mathbf{x},\mathbf{y}]}\cdot\mathbf{I}$ is an element of $\mathcal{X}$, which implies that $[\mathbf{x},\mathbf{y}] = 0$ since $\mathcal{X} \cap \mathrm{Z}(\mathscr{P}_d^n) = \{\mathbf{I}\}$. Now notice that $(\omega^{s(\mathbf{x})}\mathcal{W}_\mathbf{x})(\omega^{s(\mathbf{y})}\mathcal{W}_\mathbf{y}) = \tau^{[\mathbf{x},\mathbf{y}]}\omega^{s(\mathbf{x})+s(\mathbf{y})}\mathcal{W}_{\mathbf{x}+\mathbf{y}}$, where $[\mathbf{x},\mathbf{y}] = 0 \pmod d$ due to commutativity, therefore, by the uniqueness of phase of each Weyl operator, $s(\mathbf{x}) + s(\mathbf{y}) = s(\mathbf{x}+\mathbf{y})$ if d is odd, and $\omega^{s(\mathbf{x}+\mathbf{y})} = \omega^{s(\mathbf{x})+s(\mathbf{y})}\tau^{[\mathbf{x},\mathbf{y}]}$ is d is even. Moreover, $s(\mathbf{0}) = 0$ again due to $\mathcal{X} \cap \mathrm{Z}(\mathscr{P}_d^n) = \{\mathbf{I}\}$. Therefore, for d odd, $\omega^{s(\cdot)}$ is a character and we can write $s(\mathbf{x}) = [\mathbf{a},\mathbf{x}]$ for some $\mathbf{a} \in \mathbb{Z}_d^{2n}$. Finally, we prove that $\mathscr{X}$ is Lagrangian if $\mathcal{X}$ is maximal. Assume by contradiction that exists $\mathbf{y} \in \mathscr{X}^\perp$ such that $\mathbf{y} \notin \mathscr{X}$. This would mean that $\omega^{s(\mathbf{y})}\mathcal{W}_\mathbf{y} \notin \mathcal{X}$ while $\mathcal{W}_\mathbf{y}$ commutes with all elements of $\mathcal{X}$, a contradiction since $\mathcal{X}$ is maximal. Therefore $\mathbf{y} \in \mathscr{X}$ and $\mathscr{X}$ is Lagrangian. As a consequence, $|\mathscr{X}| = d^n$. $\blacksquare$

Corollary 19. *Any stabiliser group $\mathcal{S}$ of $\mathscr{P}_d^n$ is commutative (abelian) and has size $|\mathcal{S}| = d^n$.*

Given a generating set $\{(\mathbf{v}_i, \mathbf{w}_i)\}_{i\in[\ell]}$ for $\mathscr{M}$, let the matrices $\mathbf{V}, \mathbf{W} \in \mathbb{Z}_d^{n\times\ell}$ with columns $\mathbf{v}_1, \ldots, \mathbf{v}_\ell$ and $\mathbf{w}_1, \ldots, \mathbf{w}_\ell$, respectively. Then $\mathscr{M} = \mathrm{col}\left(\left[\begin{smallmatrix}\mathbf{V}\\ \mathbf{W}\end{smallmatrix}\right]\right)$. Note that $\mathbf{V}^\top\mathbf{W} = \mathbf{W}^\top\mathbf{V}$ since $\mathscr{M}$ is isotropic.

Together with the concept of stabiliser group is the concept of stabiliser state, which is the joint $+1$-eigenstate of all elements of a stabiliser group.

Definition 20 (Partial stabiliser state)**.** *A non-zero state $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ is* stabilised *by $\mathcal{P} \in \mathscr{P}_d^n$ if $\mathcal{P}|\psi\rangle = |\psi\rangle$. A non-zero state $|\Psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ is a* partial stabiliser state *of a partial stabiliser group $\mathcal{X}$ if $\mathcal{P}|\Psi\rangle = |\Psi\rangle$ for all $\mathcal{P} \in \mathcal{X}$. Let $\mathfrak{X}_d^n(K)$ denote the set of all partial stabiliser states in $(\mathbb{C}^d)^{\otimes n}$ of partial stabiliser groups $\mathcal{X}$ of size at least K.*

Definition 21 (Stabiliser state)**.** *A non-zero state* $|\Psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ *is a* stabiliser state *of a stabiliser group* $\mathcal{S}$ *if* $\mathcal{P}|\Psi\rangle = |\Psi\rangle$ *for all* $\mathcal{P} \in \mathcal{S}$*. Let* $\mathcal{S}_d^n$ *be the set of all stabiliser states in* $(\mathbb{C}^d)^{\otimes n}$*. Clearly* $\mathcal{S}_d^n = \mathcal{X}_d^n(d^n)$*.*

Consider the stabiliser group $\mathcal{S} = \{\omega^{s(\mathbf{x})}\mathcal{W}_\mathbf{x} : \mathbf{x} \in \mathscr{M}\}$ and let $V_\mathcal{S} \triangleq \{|\Psi\rangle \in (\mathbb{C}^d)^{\otimes n} : \omega^{s(\mathbf{x})}\mathcal{W}_\mathbf{x}|\Psi\rangle = |\Psi\rangle, \forall \mathbf{x} \in \mathscr{M}\}$ be the set of stabiliser states of $\mathcal{S}$. The next lemma constructs an projector onto $V_\mathcal{S}$ and shows that any stabiliser group has a unique stabiliser state.

Lemma 22. *Let* $\mathcal{X} = \{\omega^{s(\mathbf{x})}\mathcal{W}_\mathbf{x} : \mathbf{x} \in \mathscr{X}\}$ *be a partial stabiliser group and let* $V_\mathcal{X} \triangleq \{|\Psi\rangle \in (\mathbb{C}^d)^{\otimes n} : \omega^{s(\mathbf{x})}\mathcal{W}_\mathbf{x}|\Psi\rangle = |\Psi\rangle, \forall \mathbf{x} \in \mathscr{X}\}$*. Then the projector onto* $V_\mathcal{X}$ *is*

$$\Pi_\mathcal{X} = \frac{1}{|\mathscr{X}|}\sum_{\mathbf{x}\in\mathscr{X}} \omega^{s(\mathbf{x})}\mathcal{W}_\mathbf{x}.$$

Moreover, $\dim(V_\mathcal{X}) = d^n/|\mathscr{X}|$*. It then follows that any stabiliser group* $\mathcal{S}$ *has a unique stabiliser state, i.e.,* $\dim(V_\mathcal{S}) = 1$*.*

Proof. Consider the operators $\Pi_\mathbf{a} = |\mathscr{X}|^{-1}\sum_{\mathbf{x}\in\mathscr{X}} \omega^{s(\mathbf{x})+[\mathbf{a},\mathbf{x}]}\mathcal{W}_\mathbf{x}$ for $\mathbf{a} \in \mathbb{Z}_d^{2n}$. It is easy to see that $\Pi_\mathbf{a}$ is Hermitian by the change of index $\mathbf{x} \mapsto -\mathbf{x}$ since $s(-\mathbf{x}) = -s(\mathbf{x})$. It is also idempotent,

$$\Pi_\mathbf{a}^2 = \frac{1}{|\mathscr{X}|^2}\sum_{\mathbf{x},\mathbf{y}\in\mathscr{X}} \omega^{s(\mathbf{x})+s(\mathbf{y})+[\mathbf{a},\mathbf{x}+\mathbf{y}]}\mathcal{W}_\mathbf{x}\mathcal{W}_\mathbf{y} = \frac{1}{|\mathscr{X}|^2}\sum_{\mathbf{x},\mathbf{y}\in\mathscr{X}} \omega^{s(\mathbf{x}+\mathbf{y})+[\mathbf{a},\mathbf{x}+\mathbf{y}]}\mathcal{W}_{\mathbf{x}+\mathbf{y}} = \Pi_\mathbf{a},$$

where we used that $\omega^{s(\mathbf{x}+\mathbf{y})} = \omega^{s(\mathbf{x})+s(\mathbf{y})}\tau^{[\mathbf{x},\mathbf{y}]}$ (which is valid for both odd and even d). Therefore, $\Pi_\mathbf{a}$ (and so $\Pi_\mathcal{X}$) is a projector. Moreover, $\Pi_\mathcal{X}|\Psi\rangle = |\Psi\rangle$ for any $|\Psi\rangle \in V_\mathcal{X}$. On the other hand, given $\Pi_\mathcal{X}|\psi\rangle$ for any $|\psi\rangle$, $\omega^{s(\mathbf{x})}\mathcal{W}_\mathbf{x}(\Pi_\mathcal{X}|\psi\rangle) = \Pi_\mathcal{X}|\psi\rangle$ for all $\mathbf{x} \in \mathscr{X}$, and thus $\Pi_\mathcal{X}|\psi\rangle \in V_\mathcal{X}$. This proves that $\Pi_\mathcal{X}$ is the projector onto $V_\mathcal{X}$.

Due to the Pontryagin duality theorem (see [Rud90, Theorem 1.7.2]) and to $\mathscr{X}$ being a finite group, there are $|\mathscr{X}|$ characters of $\mathscr{X}$ (since the set of characters of $\mathscr{X}$ is isomorphic to $\mathscr{X}$). Each character gives rise to a distinct projector $\Pi_C = |\mathscr{X}|^{-1}\sum_{\mathbf{x}\in\mathscr{X}} \omega^{s(\mathbf{x})+[\mathbf{a}_C,\mathbf{x}]}\mathcal{W}_\mathbf{x}$ where $C \in \mathbb{Z}_d^{2n}/\mathscr{X}^\perp$ is a coset and $\mathbf{a}_C$ is any element from C. Two different projectors are orthogonal since they belong to different eigenvalues of at least one Weyl operator or, alternatively,

$$\Pi_C\Pi_{C'} = \frac{1}{|\mathscr{X}|^2}\sum_{\mathbf{x},\mathbf{y}\in\mathscr{X}} \omega^{s(\mathbf{x})+s(\mathbf{y})+[\mathbf{a}_C,\mathbf{x}]+[\mathbf{a}_{C'},\mathbf{y}]}\mathcal{W}_\mathbf{x}\mathcal{W}_\mathbf{y} = \frac{1}{|\mathscr{X}|^2}\sum_{\mathbf{x},\mathbf{z}\in\mathscr{X}} \omega^{s(\mathbf{z})+[\mathbf{a}_C,\mathbf{x}]+[\mathbf{a}_{C'},\mathbf{z}-\mathbf{x}]}\mathcal{W}_\mathbf{z} = 0,$$

using that $\omega^{s(\mathbf{x})+s(\mathbf{y})}\mathcal{W}_\mathbf{x}\mathcal{W}_\mathbf{y} = \omega^{s(\mathbf{x}+\mathbf{y})}\mathcal{W}_{\mathbf{x}+\mathbf{y}}$ and $\sum_{\mathbf{x}\in\mathscr{X}} \omega^{[\mathbf{a}_C-\mathbf{a}_{C'},\mathbf{x}]} = 0$ since $\mathbf{a}_C - \mathbf{a}_{C'} \neq \mathscr{X}^\perp$ (Lemma 9). This means that the rank of each projector must be a $|\mathscr{X}|$-fraction of the Hilbert space dimension d^n, i.e., $d^n/|\mathscr{X}|$. This implies that $\dim(V_\mathcal{X}) = \operatorname{rank}(\Pi_C) = d^n/|\mathscr{X}|$. ∎

Following [Gro06, GNW21], we can denote the unique stabiliser state $|\mathcal{S}\rangle$ of a stabiliser group $\mathcal{S} = \{\omega^{s(\mathbf{x})}\mathcal{W}_\mathbf{x} : \mathbf{x} \in \mathscr{M}\}$ as $|\mathscr{M}, s\rangle$, where $\mathscr{M} \subset \mathbb{Z}_d^{2n}$ and $s : \mathscr{M} \to \mathbb{Z}_d$ are the Lagrangian submodule and phase function defining $\mathcal{S}$, respectively. We note that we can obtain another function from s by replacing it with $s'(\mathbf{x}) = s(\mathbf{x}) + [\mathbf{z},\mathbf{x}]$ for some $\mathbf{z} \in \mathbb{Z}_d^{2n}$, in which case $|\mathscr{M}, s'\rangle = \mathcal{W}_\mathbf{z}|\mathscr{M}, s\rangle$. Even more precisely, any two strings $\mathbf{z}, \mathbf{z}'$ from a coset $C \in \mathbb{Z}_d^{2n}/\mathscr{M}$ yields the same stabiliser state $|\mathscr{M}, s + [\mathbf{z},\cdot]\rangle$ since then $\mathbf{z} - \mathbf{z}' \in \mathscr{M}$. Given $\mathscr{M}$ and s, we can write the stabiliser state $|\mathscr{M}, s + [\mathbf{z},\cdot]\rangle$ as $|\mathscr{M}, s, C\rangle$ where $\mathbf{z} \in C$ and $C \in \mathbb{Z}_d^{2n}/\mathscr{M}$. Conversely, given a Lagrangian subspace $\mathscr{M}$, any state that is a simultaneous eigenvector of $\{\mathcal{W}_\mathbf{x}\}_{\mathbf{x}\in\mathscr{M}}$ is a stabiliser state of some stabiliser group. The eigenvectors of $\{\mathcal{W}_\mathbf{x}\}_{\mathbf{x}\in\mathscr{M}}$ thus determine a stabiliser basis $\{|\mathscr{M}, s\rangle\}_s = \{|\mathscr{M}, s, C\rangle\}_{C\in\mathbb{Z}_d^{2n}/\mathscr{M}}$.

Consider a stabiliser group $\mathcal{S} = \{\omega^{s(\mathbf{x})}\mathcal{W}_\mathbf{x} : \mathbf{x} \in \mathscr{M}\}$. According to Lemma 22, the operator $d^{-n}\sum_{\mathbf{x}\in\mathscr{M}} \omega^{s(\mathbf{x})}\mathcal{W}_\mathbf{x}$ is the orthogonal projection onto the set $V_\mathcal{S}$ spanned by the stabiliser state $|\mathcal{S}\rangle$ of $\mathcal{S}$. Therefore,

$$|\mathcal{S}\rangle\langle\mathcal{S}| = \frac{1}{d^n}\sum_{\mathbf{x}\in\mathscr{M}} \omega^{s(\mathbf{x})}\mathcal{W}_\mathbf{x} \tag{1}$$

and the characteristic function and probability of $|\mathcal{S}\rangle\langle\mathcal{S}|$ are

$$c_{\mathcal{S}}(\mathbf{x}) = \begin{cases} d^{-n/2}\omega^{s(\mathbf{x})} & \text{if } \mathbf{x} \in \mathscr{M}, \\ 0 & \text{otherwise,} \end{cases} \qquad \text{and} \qquad p_{\mathcal{S}}(\mathbf{x}) = \begin{cases} d^{-n} & \text{if } \mathbf{x} \in \mathscr{M}, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Thus $p_{\mathcal{S}}$ is the uniform distribution on the subspace $\mathscr{M}$, a fact which will be of paramount importance when studying Bell (difference) sampling in the next section. While we have obtained an expression for $|\mathcal{S}\rangle\langle\mathcal{S}|$, in the following lemma, we obtain an expression for $|\mathcal{S}\rangle$. We note that a similar form for stabiliser states was given in [HDDM05, Theorem 1].

Lemma 23. *Let $\mathcal{S} = \langle\{\omega^{s_i}\mathcal{W}_{\mathbf{v}_i,\mathbf{w}_i}\}_{i\in[\ell]}\rangle$ be a stabiliser group. Let $\mathbf{s} = (s_1, \ldots, s_\ell) \in \mathbb{Z}_d^\ell$ and let $\mathbf{V}, \mathbf{W} \in \mathbb{Z}_d^{n\times\ell}$ be the matrices with column vectors $\mathbf{v}_1, \ldots, \mathbf{v}_\ell$ and $\mathbf{w}_1, \ldots, \mathbf{w}_\ell$, respectively. Then the unique stabiliser state of $\mathcal{S}$ is*

$$|\mathcal{S}\rangle = \frac{1}{\sqrt{d^\ell |\operatorname{null}(\mathbf{W})|}} \sum_{\mathbf{q}\in\mathbb{Z}_d^\ell} \omega^{\mathbf{q}^\top(\mathbf{s}+\mathbf{V}^\top\mathbf{u})} \cdot \tau^{\mathbf{q}^\top\mathbf{V}^\top\mathbf{W}\mathbf{q}} |\mathbf{W}\mathbf{q} + \mathbf{u}\rangle$$

for any $\mathbf{u} \in \mathbb{Z}_d^n$ such that $\mathbf{s} + \mathbf{V}^\top\mathbf{u} \in \operatorname{row}(\mathbf{W})$ (the multiplication in $\mathbf{q}^\top\mathbf{V}^\top\mathbf{W}\mathbf{q}$ is over $\mathbb{Z}_D$).

Proof. First note that the unormalised state $|\mathcal{S}_\mathbf{u}\rangle \triangleq |\mathcal{S}\rangle\langle\mathcal{S}|\mathbf{u}\rangle$, for any $\mathbf{u} \in \mathbb{Z}_d^n$, is either the stabiliser state $|\mathcal{S}\rangle$ of $\mathcal{S}$ or the zero state. Then

$$\begin{aligned} |\mathcal{S}_\mathbf{u}\rangle &= \frac{1}{d^n} \sum_{\mathbf{x}\in\mathscr{M}} \omega^{s(\mathbf{x})} \mathcal{W}_\mathbf{x} |\mathbf{u}\rangle && \text{(by Equation (1))} \\ &\propto \sum_{\mathbf{q}\in\mathbb{Z}_d^\ell} \omega^{s(q_1\mathbf{x}_1+\cdots+q_\ell\mathbf{x}_\ell)} \mathcal{W}_{q_1\mathbf{x}_1+\cdots+q_\ell\mathbf{x}_\ell} |\mathbf{u}\rangle \\ &= \sum_{\mathbf{q}\in\mathbb{Z}_d^\ell} \omega^{\sum_{i=1}^\ell q_i s(\mathbf{x}_i)} \mathcal{W}_{\mathbf{V}\mathbf{q},\mathbf{W}\mathbf{q}} |\mathbf{u}\rangle && \text{(since } [\mathbf{x}_i, \mathbf{x}_j] = 0\text{)} \\ &= \sum_{\mathbf{q}\in\mathbb{Z}_d^\ell} \omega^{\mathbf{q}^\top(\mathbf{s}+\mathbf{V}^\top\mathbf{u})} \cdot \tau^{\mathbf{q}^\top\mathbf{V}^\top\mathbf{W}\mathbf{q}} |\mathbf{W}\mathbf{q} + \mathbf{u}\rangle. \end{aligned}$$

To analyse the case when $|\mathcal{S}_\mathbf{u}\rangle$ is the stabiliser state, we compute $|\langle\mathcal{S}_\mathbf{u}|\mathcal{S}_\mathbf{u}\rangle|$ as follows:

$$\begin{aligned} |\langle\mathcal{S}_\mathbf{u}|\mathcal{S}_\mathbf{u}\rangle| &\propto \left| \sum_{\mathbf{q},\mathbf{q}'\in\mathbb{Z}_d^\ell} \omega^{(\mathbf{s}+\mathbf{V}^\top\mathbf{u})^\top(\mathbf{q}-\mathbf{q}')} \cdot \tau^{\mathbf{q}^\top\mathbf{V}^\top\mathbf{W}\mathbf{q}-\mathbf{q}'^\top\mathbf{V}^\top\mathbf{W}\mathbf{q}'} \langle\mathbf{u}+\mathbf{W}\mathbf{q}|\mathbf{u}+\mathbf{W}\mathbf{q}'\rangle \right| \\ &= \left| \sum_{\mathbf{q}\in\mathbb{Z}_d^\ell} \sum_{\mathbf{q}'\in\operatorname{null}(\mathbf{W})} \omega^{(\mathbf{s}+\mathbf{V}^\top\mathbf{u})^\top\mathbf{q}'} \cdot \tau^{\mathbf{q}^\top\mathbf{V}^\top\mathbf{W}\mathbf{q}-(\mathbf{q}+\mathbf{q}')^\top\mathbf{V}^\top\mathbf{W}(\mathbf{q}+\mathbf{q}')} \right| \\ &= \left| \sum_{\mathbf{q}\in\mathbb{Z}_d^\ell} \sum_{\mathbf{q}'\in\operatorname{null}(\mathbf{W})} \omega^{(\mathbf{s}+\mathbf{V}^\top\mathbf{u})^\top\mathbf{q}'} \right| && \text{(by } \mathbf{V}^\top\mathbf{W} = \mathbf{W}^\top\mathbf{V} \text{ and } \mathbf{W}\mathbf{q}' = \mathbf{0}\text{)} \\ &= d^\ell \left| \sum_{\mathbf{q}'\in\operatorname{null}(\mathbf{W})} \omega^{(\mathbf{s}+\mathbf{V}^\top\mathbf{u})^\top\mathbf{q}'} \right|. \end{aligned}$$

According to Lemma 5 and 9,

$$\sum_{\mathbf{q}'\in\operatorname{null}(\mathbf{W})} \omega^{(\mathbf{s}+\mathbf{V}^\top\mathbf{u})^\top\mathbf{q}'} = \begin{cases} |\operatorname{null}(\mathbf{W})| & \text{if } \mathbf{s} + \mathbf{V}^\top\mathbf{u} \in \operatorname{null}(\mathbf{W})^\perp = \operatorname{row}(\mathbf{W}), \\ 0 & \text{if } \mathbf{s} + \mathbf{V}^\top\mathbf{u} \notin \operatorname{null}(\mathbf{W})^\perp = \operatorname{row}(\mathbf{W}). \end{cases}$$

Thus $|\langle\mathcal{S}_\mathbf{u}|\mathcal{S}_\mathbf{u}\rangle| \propto d^\ell |\operatorname{null}(\mathbf{W})|$ if $\mathbf{s} + \mathbf{V}^\top\mathbf{u} \in \operatorname{row}(\mathbf{W})$, and $|\langle\mathcal{S}_\mathbf{u}|\mathcal{S}_\mathbf{u}\rangle| = 0$ otherwise. Finally, it is clear that there always is $\mathbf{u} \in \mathbb{Z}_d^n$ such that $\mathbf{s} + \mathbf{V}^\top\mathbf{u} \in \operatorname{row}(\mathbf{W})$, otherwise $|\mathcal{S}_\mathbf{u}\rangle = 0$ for all $\mathbf{u} \in \mathbb{Z}_d^n$ and $|\mathcal{S}\rangle\langle\mathcal{S}|$ would not be an orthogonal projection. ∎

2.6 Stabiliser size

Sometimes we are interested in the Pauli operators that stabilise a given pure state $|\psi\rangle$ up to a phase. This is captured by the *unsigned stabiliser group* of $|\psi\rangle$.

Definition 24 (Unsigned stabiliser group and stabiliser size)**.** *Given* $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$*, its* unsigned stabiliser group *is* $\mathrm{Weyl}(|\psi\rangle) \triangleq \{\mathbf{x} \in \mathbb{Z}_d^{2n} : |\langle\psi|\mathcal{W}_{\mathbf{x}}|\psi\rangle| = 1\}$ *and its* stabiliser size *is* $|\,\mathrm{Weyl}(|\psi\rangle)|$.

Remark 25. *Note that for all* $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$, $\mathrm{Weyl}(|\psi\rangle)$ *is a partial stabiliser group. On the other hand, for any partial stabiliser group* $\mathcal{X}$*, there always is* $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ *such that* $\mathrm{Weyl}(|\psi\rangle) = \mathcal{X}$.

Lemma 26. *For any non-zero* $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$, $\mathrm{Weyl}(|\psi\rangle)$ *is an isotropic submodule.*

Proof. Suppose there are $\mathbf{x}, \mathbf{y} \in \mathrm{Weyl}(|\psi\rangle)$ such that $[\mathbf{x}, \mathbf{y}] \neq 0 \pmod d$. On the one hand, $\mathcal{W}_{\mathbf{x}}\mathcal{W}_{\mathbf{y}}|\psi\rangle = \omega^{[\mathbf{x},\mathbf{y}]}\mathcal{W}_{\mathbf{y}}\mathcal{W}_{\mathbf{x}}|\psi\rangle = \omega^{[\mathbf{x},\mathbf{y}]+s(\mathbf{x})+s(\mathbf{y})}|\psi\rangle$, while on the other hand, $\mathcal{W}_{\mathbf{x}}\mathcal{W}_{\mathbf{y}}|\psi\rangle = \omega^{s(\mathbf{x})+s(\mathbf{y})}|\psi\rangle$, which is a contradiction. Moreover, given $\mathbf{x}, \mathbf{y} \in \mathrm{Weyl}(|\psi\rangle)$, then $\mathcal{W}_{\mathbf{x}+\mathbf{y}}|\psi\rangle = \tau^{[\mathbf{y},\mathbf{x}]}\mathcal{W}_{\mathbf{x}}\mathcal{W}_{\mathbf{y}}|\psi\rangle = \tau^{[\mathbf{y},\mathbf{x}]}\omega^{s(\mathbf{x})+s(\mathbf{y})}|\psi\rangle$, but since $[\mathbf{y}, \mathbf{x}] = 0 \pmod d$, either $\tau^{[\mathbf{y},\mathbf{x}]} = 1$ or $\tau^{[\mathbf{y},\mathbf{x}]} = \omega^{d/2}$, the latter for d even only. Thus $\mathbf{x} + \mathbf{y} \in \mathrm{Weyl}(|\psi\rangle)$ and $\mathrm{Weyl}(|\psi\rangle)$ is a submodule. ■

The stabiliser size is invariant under Clifford transformations as a corollary of the next fact.

Definition 27 (Symplectic matrix)**.** *A matrix* $\Gamma \in \mathbb{Z}_d^{2n\times 2n}$ *is called* symplectic *if it preserves the symplectic product, i.e.,* $[\Gamma\mathbf{x}, \Gamma\mathbf{y}] = [\mathbf{x}, \mathbf{y}]$ *for all* $\mathbf{x}, \mathbf{y} \in \mathbb{Z}_d^{2n}$*. Equivalently,* Γ *satisfies* $\Gamma^\top\Omega\Gamma = \Omega$*, where* $\Omega \triangleq \left(\begin{smallmatrix} \mathbf{0}_n & \mathbf{I}_n \\ -\mathbf{I}_n & \mathbf{0}_n \end{smallmatrix}\right)$*. Let* $\mathrm{Sp}(2n, d)$ *be the set of symplectic matrices over* $\mathbb{Z}_d^{2n}$.

Fact 28 ([DB13, Theorem 5])**.** *Any Clifford operator* $\mathcal{C} \in \mathscr{C}_d^n$ *is of the form* $\mathcal{C} = \mathcal{W}_{\mathbf{a}}\mu(\Gamma)$ *for some* $\mathbf{a} \in \mathbb{Z}_d^{2n}$ *and a unitary operator* $\mu(\Gamma) \in (\mathbb{C}^d)^{\otimes n}$ *such that* $\mu(\Gamma)\mathcal{W}_{\mathbf{x}}\mu(\Gamma)^\dagger = \mathcal{W}_{\Gamma\mathbf{x}}$ *for all* $\mathbf{x} \in \mathbb{Z}_d^{2n}$*, where* $\Gamma \in \mathrm{Sp}(2n, d)$.

Corollary 29. *For all* $\mathcal{C} \in \mathscr{C}_d^n$ *and* $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$, $|\,\mathrm{Weyl}(\mathcal{C}|\psi\rangle)| = |\,\mathrm{Weyl}(|\psi\rangle)|$.

Proof. Write $\mathcal{C} = \mathcal{W}_{\mathbf{a}}\mu(\Gamma)$ for some $\mathbf{a} \in \mathbb{Z}_d^{2n}$ and $\Gamma \in \mathrm{Sp}(2n, d)$. The result follows because $|\,\mathrm{Weyl}(\mathcal{C}|\psi\rangle)| = |\,\mathrm{Weyl}(\mu(\Gamma)|\psi\rangle)|$ and Γ is invertible ($\Gamma^{-1} = \Omega^{-1}\Gamma^\top\Omega$), so that $\mathrm{Weyl}(\mu(\Gamma)|\psi\rangle) = \{\Gamma\mathbf{x} : \mathbf{x} \in \mathrm{Weyl}(|\psi\rangle)\}$ and $\mathrm{Weyl}(|\psi\rangle) = \{\Gamma^{-1}\mathbf{x} : \mathbf{x} \in \mathrm{Weyl}(\mu(\Gamma)|\psi\rangle)\}$. Indeed, it is not hard to see that $\{\Gamma\mathbf{x} : \mathbf{x} \in \mathrm{Weyl}(|\psi\rangle)\} \subseteq \mathrm{Weyl}(\mu(\Gamma)|\psi\rangle)$ and $\{\Gamma^{-1}\mathbf{x} : \mathbf{x} \in \mathrm{Weyl}(\mu(\Gamma)|\psi\rangle)\} \subseteq \mathrm{Weyl}(|\psi\rangle)$. Assume by contradiction that there is $\mathbf{y} \in \mathrm{Weyl}(\mu(\Gamma)|\psi\rangle)$ such that $\mathbf{y} \notin \{\Gamma\mathbf{x} : \mathbf{x} \in \mathrm{Weyl}(|\psi\rangle)\}$. Then $\Gamma^{-1}\mathbf{y} \in \{\Gamma^{-1}\mathbf{x} : \mathbf{x} \in \mathrm{Weyl}(\mu(\Gamma)|\psi\rangle)\}$ but $\Gamma^{-1}\mathbf{y} \notin \mathrm{Weyl}(|\psi\rangle)$, a contradiction. Similarly for $\mathrm{Weyl}(|\psi\rangle) = \{\Gamma^{-1}\mathbf{x} : \mathbf{x} \in \mathrm{Weyl}(\mu(\Gamma)|\psi\rangle)\}$. ■

The next lemma covers another important property of unsigned stabiliser groups: the stabiliser size is at most d^n and is maximised for stabiliser states.

Lemma 30. *For any* $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$, $|\,\mathrm{Weyl}(|\psi\rangle)| \leqslant d^n$*, with equality if and only if* $|\psi\rangle$ *is a stabiliser state.*

Proof. A simple consequence of $\mathrm{Weyl}(|\psi\rangle)$ being isotropic and any isotropic set of elements having cardinality at most d^n, with equality if and only if the set is Lagrangian. ■

2.7 Stabiliser fidelity

Alongside its stabiliser size, another important measure of the "stabiliser complexity" of a quantum state is its *stabiliser fidelity* [BBC+19]. Here we propose a slight generalisation called "K-sized stabiliser fidelity". Recall the definitions of $\mathcal{X}_d^n(K)$ and $\mathcal{S}_d^n$ from Definition 20 and 21.

Definition 31 ((K-sized) Stabiliser fidelity)**.** *Given an* n*-qudit quantum state* $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$*, its* K-sized stabiliser fidelity *is* $F_K(|\psi\rangle) \triangleq \max_{|\Psi\rangle\in\mathcal{X}_d^n(K)} |\langle\Psi|\psi\rangle|^2$*, while its* stabiliser fidelity *is* $F_{\mathcal{S}}(|\psi\rangle) \triangleq \max_{|\mathcal{S}\rangle\in\mathcal{S}_d^n} |\langle\mathcal{S}|\psi\rangle|^2$.

It is possible to relate the K-sized stabiliser fidelity $F_K(|\psi\rangle)$ of a quantum state $|\psi\rangle$ to its characteristic distribution p_ψ as shown next.

Lemma 32. *Given $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$, let $\mathscr{X} \subset \mathbb{Z}_d^{2n}$ be the isotropic submodule characterising the partial stabiliser state that maximises $F_K(|\psi\rangle)$. Then*

$$\sum_{\mathbf{x}\in\mathscr{X}^\perp} p_\psi(\mathbf{x}) \leqslant F_K(|\psi\rangle) \leqslant \sqrt{\sum_{\mathbf{x}\in\mathscr{X}^\perp} p_\psi(\mathbf{x})}.$$

Proof. Let $|\Psi\rangle = \arg\max_{|\phi\rangle\in\mathfrak{X}_d^n(K)} |\langle\phi|\psi\rangle|^2$ and $\mathcal{X} = \{\omega^{s(\mathbf{x})}\mathcal{W}_\mathbf{x} : \mathbf{x} \in \mathscr{X}\}$ be its partial stabiliser group. Let $\Pi_\mathcal{X} = |\mathscr{X}|^{-1}\sum_{\mathbf{x}\in\mathscr{X}} \omega^{s(\mathbf{x})}\mathcal{W}_\mathbf{x}$ be the projection onto the set of partial stabiliser states of $\mathcal{X}$, i.e., $V_\mathcal{X} = \{|\Psi\rangle \in (\mathbb{C}^d)^{\otimes n} : \omega^{s(\mathbf{x})}\mathcal{W}_\mathbf{x}|\Psi\rangle = |\Psi\rangle, \forall \mathbf{x} \in \mathscr{X}\}$. For the upper bound,

$$\begin{aligned}
F_K(|\psi\rangle) &= \langle\psi|\Pi_\mathcal{X}|\psi\rangle \\
&= \frac{1}{|\mathscr{X}|}\sum_{\mathbf{x}\in\mathscr{X}} \omega^{s(\mathbf{x})}\langle\psi|\mathcal{W}_\mathbf{x}|\psi\rangle && \text{(by Lemma 22)} \\
&\leqslant \sqrt{\frac{1}{|\mathscr{X}|}\sum_{\mathbf{x}\in\mathscr{X}} |\langle\psi|\mathcal{W}_\mathbf{x}|\psi\rangle|^2} && \text{(by Cauchy-Schwarz)} \\
&= \sqrt{\frac{d^n}{|\mathscr{X}|}\sum_{\mathbf{x}\in\mathscr{X}} p_\psi(\mathbf{x})} && \text{(definition of } p_\psi(\mathbf{x})\text{)} \\
&= \sqrt{\sum_{\mathbf{x}\in\mathscr{X}^\perp} p_\psi(\mathbf{x})}. && \text{(by Lemma 16)}
\end{aligned}$$

For the lower bound, define the projections $\Pi_C = |\mathscr{X}|^{-1}\sum_{\mathbf{x}\in\mathscr{X}} \omega^{s(\mathbf{x})+[\mathbf{b},\mathbf{x}]}\mathcal{W}_\mathbf{x}$ from the proof of Lemma 22, where $C \in \mathbb{Z}_d^{2n}/\mathscr{X}^\perp$ is a coset and $\mathbf{b} \in \mathbb{Z}_d^{2n}$ is any element from C. Then

$$\begin{aligned}
F_K(|\psi\rangle) &= \max_{C\in\mathbb{Z}_d^{2n}/\mathscr{X}^\perp} \langle\psi|\Pi_C|\psi\rangle \\
&\geqslant \sum_{C\in\mathbb{Z}_d^{2n}/\mathscr{X}^\perp} \langle\psi|\Pi_C|\psi\rangle^2 && \text{(by } \textstyle\sum_{C\in\mathbb{Z}_d^{2n}/\mathscr{X}^\perp}\Pi_C = \mathbf{I}\text{)} \\
&= \frac{1}{|\mathscr{X}|^2}\sum_{C\in\mathbb{Z}_d^{2n}/\mathscr{X}^\perp}\sum_{\mathbf{x},\mathbf{y}\in\mathscr{X}} \omega^{s(\mathbf{x})+s(\mathbf{y})+[\mathbf{b}_C,\mathbf{x}+\mathbf{y}]}\langle\psi|\mathcal{W}_\mathbf{x}|\psi\rangle\langle\psi|\mathcal{W}_\mathbf{y}|\psi\rangle, && \text{(by Lemma 22)}
\end{aligned}$$

where $\mathbf{b}_C$ is any element of the coset C. By taking the average over all elements $\mathbf{b} \in C$, then

$$\begin{aligned}
F_K(|\psi\rangle) &\geqslant \frac{1}{|\mathscr{X}|^2|\mathscr{X}^\perp|}\sum_{C\in\mathbb{Z}_d^{2n}/\mathscr{X}^\perp}\sum_{\mathbf{b}\in C}\sum_{\mathbf{x},\mathbf{y}\in\mathscr{X}} \omega^{s(\mathbf{x})+s(\mathbf{y})+[\mathbf{b},\mathbf{x}+\mathbf{y}]}\langle\psi|\mathcal{W}_\mathbf{x}|\psi\rangle\langle\psi|\mathcal{W}_\mathbf{y}|\psi\rangle \\
&= \frac{d^{-2n}}{|\mathscr{X}|}\sum_{\mathbf{b}\in\mathbb{Z}_d^{2n}}\sum_{\mathbf{x},\mathbf{y}\in\mathscr{X}} \omega^{s(\mathbf{x}+\mathbf{y})+[\mathbf{b},\mathbf{x}+\mathbf{y}]}\langle\psi|\mathcal{W}_\mathbf{x}|\psi\rangle\langle\psi|\mathcal{W}_\mathbf{y}|\psi\rangle \\
&\qquad\qquad (\omega^{s(\mathbf{x}+\mathbf{y})} = \omega^{s(\mathbf{x})+s(\mathbf{y})} \text{ for } \mathbf{x},\mathbf{y} \in \mathscr{X}) \\
&= \frac{1}{|\mathscr{X}|}\sum_{\mathbf{x}\in\mathscr{X}} \langle\psi|\mathcal{W}_\mathbf{x}|\psi\rangle\langle\psi|\mathcal{W}_{-\mathbf{x}}|\psi\rangle && \text{(by Lemma 9 over } \mathbf{b} \text{ and } s(\mathbf{0}) = 0\text{)} \\
&= \frac{d^n}{|\mathscr{X}|}\sum_{\mathbf{x}\in\mathscr{M}} p_\psi(\mathbf{x}) && \text{(definition of } p_\psi(\mathbf{x})\text{)} \\
&= \sum_{\mathbf{x}\in\mathscr{X}^\perp} p_\psi(\mathbf{x}). && \text{(by Lemma 16)} \quad \blacksquare
\end{aligned}$$

Corollary 33. *Given a quantum state $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$, let $|\mathscr{M}, s\rangle = \arg\max_{|\mathcal{S}\rangle\in\mathcal{S}_d^n} |\langle\mathcal{S}|\psi\rangle|^2$ be the stabiliser state that maximises the stabiliser fidelity $F_{\mathcal{S}}(|\psi\rangle)$. Then*

$$\sum_{\mathbf{x}\in\mathscr{M}} p_\psi(\mathbf{x}) \leqslant F_{\mathcal{S}}(|\psi\rangle) \leqslant \sqrt{\sum_{\mathbf{x}\in\mathscr{M}} p_\psi(\mathbf{x})}.$$

We also prove that given an isotropic submodule $\mathscr{X}$ with large enough p_ψ-mass, it is possible to find a state $|\phi\rangle$ with large fidelity with $|\psi\rangle$ such that $\mathscr{X} \subseteq \mathrm{Weyl}(|\phi\rangle)$.

Lemma 34. *Let $\mathscr{X} \subset \mathbb{Z}_d^{2n}$ be an isotropic submodule with size d^n/K and such that*

$$\sum_{\mathbf{x}\in\mathscr{X}} p_\psi(\mathbf{x}) \geqslant \frac{1-\varepsilon}{K} \qquad \text{for } \varepsilon \in [0,1].$$

Then there is $|\phi\rangle \in (\mathbb{C}^d)^{\otimes n}$ with $\mathscr{X} \subseteq \mathrm{Weyl}(|\phi\rangle)$ such that the fidelity between $|\psi\rangle$ and $|\phi\rangle$ is at least $1-\varepsilon$, i.e., $|\langle\phi|\psi\rangle|^2 \geqslant 1-\varepsilon$.

Proof. For any suitable phase function $s : \mathbb{Z}_d^{2n} \to \mathbb{Z}_d$, consider the partial stabiliser group $\mathcal{X} = \{\omega^{s(\mathbf{x})}\mathcal{W}_\mathbf{x} : \mathbf{x} \in \mathscr{X}\}$. Define the projections $\Pi_C = |\mathscr{X}|^{-1}\sum_{\mathbf{x}\in\mathscr{X}} \omega^{s(\mathbf{x})+[\mathbf{b},\mathbf{x}]}\mathcal{W}_\mathbf{x}$ from the proof of Lemma 22, where $C \in \mathbb{Z}_d^{2n}/\mathscr{X}^\perp$ is a coset and $\mathbf{b} \in \mathbb{Z}_d^{2n}$ is any element from C. Let $V_{\mathcal{X}}(C) \triangleq \{|\Psi\rangle \in (\mathbb{C}^d)^{\otimes n} : \omega^{s(\mathbf{x})+[\mathbf{b},\mathbf{x}]}\mathcal{W}_\mathbf{x}|\Psi\rangle = |\Psi\rangle, \forall \mathbf{x} \in \mathscr{X}\}$ for any $\mathbf{b} \in C$ and let $|\phi\rangle$ be the state that maximises $\max_{C\in\mathbb{Z}_d^{2n}/\mathscr{X}^\perp} \max_{|\Psi\rangle\in V_{\mathcal{X}}(C)} |\langle\Psi|\psi\rangle|^2$. Clearly $\mathscr{X} \subseteq \mathrm{Weyl}(|\phi\rangle)$. By following the exact same steps as the proof Lemma 32, one can show that $\max_{C\in\mathbb{Z}_d^{2n}/\mathscr{X}^\perp} \max_{|\Psi\rangle\in V_{\mathcal{X}}(C)} |\langle\Psi|\psi\rangle|^2 = \max_{C\in\mathbb{Z}_d^{2n}/\mathscr{X}^\perp} \langle\psi|\Pi_C|\psi\rangle \geqslant \sum_{\mathbf{x}\in\mathscr{X}^\perp} p_\psi(\mathbf{x})$. Therefore, and also using Lemma 16,

$$|\langle\phi|\psi\rangle|^2 \geqslant \sum_{\mathbf{x}\in\mathscr{X}^\perp} p_\psi(\mathbf{x}) = \frac{d^n}{|\mathscr{X}|}\sum_{\mathbf{x}\in\mathscr{X}} p_\psi(\mathbf{x}) = K\sum_{\mathbf{x}\in\mathscr{X}} p_\psi(\mathbf{x}) \geqslant 1-\varepsilon. \qquad \blacksquare$$

We now bound the value of $p_\psi(\mathbf{x})$ in terms of $F_{\mathcal{S}}(|\psi\rangle)$.

Lemma 35. *Let $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ and $|\mathscr{M}, s\rangle = \arg\max_{|\mathcal{S}'\rangle\in\mathcal{S}_d^n} |\langle\mathcal{S}'|\psi\rangle|^2$ such that $F_{\mathcal{S}}(|\psi\rangle) \geqslant \frac{1}{2}$. If $\mathbf{x} \in \mathscr{M}$, then $d^n p_\psi(\mathbf{x}) \geqslant (2F_{\mathcal{S}}(|\psi\rangle) - 1)^2$.*

Proof. Let $\theta \triangleq F_{\mathcal{S}}(|\psi\rangle)$ for simplicity and write $|\psi\rangle = \sqrt{\theta}|\mathcal{S}\rangle + \sqrt{1-\theta}|\mathcal{S}^\perp\rangle$, where $|\mathcal{S}\rangle$ is the stabiliser state that maximises $F_{\mathcal{S}}(|\psi\rangle)$ and $|\mathcal{S}^\perp\rangle$ is the part orthogonal to $|\mathcal{S}\rangle$. For $\mathbf{x} \in \mathscr{M}$,

$$\begin{aligned}
\langle\psi|\mathcal{W}_\mathbf{x}|\psi\rangle &= \langle\psi|(\sqrt{\theta}\mathcal{W}_\mathbf{x}|\mathcal{S}\rangle + \sqrt{1-\theta}\mathcal{W}_\mathbf{x}|\mathcal{S}^\perp\rangle) \\
&= \sqrt{\theta}\omega^{s(\mathbf{x})}\langle\psi|\mathcal{S}\rangle + \sqrt{1-\theta}\langle\psi|\mathcal{W}_\mathbf{x}|\mathcal{S}^\perp\rangle && (\mathcal{W}_\mathbf{x}|\mathcal{S}\rangle = \omega^{s(\mathbf{x})}|\mathcal{S}\rangle) \\
&= \theta\omega^{s(\mathbf{x})} + \sqrt{1-\theta}\langle\psi|\mathcal{W}_\mathbf{x}|\mathcal{S}^\perp\rangle \\
&= \theta\omega^{s(\mathbf{x})} + \sqrt{1-\theta}(\sqrt{\theta}\langle\mathcal{S}|\mathcal{W}_\mathbf{x}|\mathcal{S}^\perp\rangle + \sqrt{1-\theta}\langle\mathcal{S}^\perp|\mathcal{W}_\mathbf{x}|\mathcal{S}^\perp\rangle) \\
&= \theta\omega^{s(\mathbf{x})} + (1-\theta)\langle\mathcal{S}^\perp|\mathcal{W}_\mathbf{x}|\mathcal{S}^\perp\rangle. && (\mathcal{W}_\mathbf{x}^\dagger|\mathcal{S}\rangle = \omega^{s(-\mathbf{x})}|\mathcal{S}\rangle \text{ and } \langle\mathcal{S}|\mathcal{S}^\perp\rangle = 0)
\end{aligned}$$

Then

$$\begin{aligned}
|\langle\psi|\mathcal{W}_\mathbf{x}|\psi\rangle|^2 &= \theta^2 + (1-\theta)^2|\langle\mathcal{S}^\perp|\omega^{-s(\mathbf{x})}\mathcal{W}_\mathbf{x}|\mathcal{S}^\perp\rangle|^2 + 2\theta(1-\theta)\,\mathrm{Re}[\langle\mathcal{S}|\omega^{-s(\mathbf{x})}\mathcal{W}_\mathbf{x}^\dagger|\mathcal{S}^\perp\rangle] \\
&\geqslant \theta^2 + (1-\theta)^2\,\mathrm{Re}[\langle\mathcal{S}|\omega^{-s(\mathbf{x})}\mathcal{W}_\mathbf{x}^\dagger|\mathcal{S}^\perp\rangle]^2 + 2\theta(1-\theta)\,\mathrm{Re}[\langle\mathcal{S}|\omega^{-s(\mathbf{x})}\mathcal{W}_\mathbf{x}^\dagger|\mathcal{S}^\perp\rangle] \\
&\geqslant \theta^2 + (1-\theta)^2 - 2\theta(1-\theta) && (\mathrm{Re}[\langle\mathcal{S}|\omega^{-s(\mathbf{x})}\mathcal{W}_\mathbf{x}^\dagger|\mathcal{S}^\perp\rangle] \in [-1,1]) \\
&= (2\theta-1)^2. && \blacksquare
\end{aligned}$$

3 Bell sampling on qudits

In this section, we generalise Bell sampling as proposed by [Mon17, GNW21] to qudits. To do so, we first introduce the generalised Bell states. Let $|\Phi^+\rangle \triangleq d^{-n/2} \sum_{\mathbf{q}\in\mathbb{Z}_d^n} |\mathbf{q}\rangle^{\otimes 2}$ be a maximally entangled state. The generalised Bell states over $\mathbb{Z}_d^n$ are defined as

$$|\mathcal{W}_{\mathbf{x}}\rangle \triangleq (\mathcal{W}_{\mathbf{x}} \otimes \mathbf{I})|\Phi^+\rangle = \frac{1}{\sqrt{d^n}} \sum_{\mathbf{q}\in\mathbb{Z}_d^n} \tau^{\langle \mathbf{v},\mathbf{w}\rangle} \cdot \omega^{\langle \mathbf{q},\mathbf{v}\rangle} |\mathbf{q}+\mathbf{w}\rangle|\mathbf{q}\rangle, \quad \forall \mathbf{x} = (\mathbf{v},\mathbf{w}) \in \mathbb{Z}_d^{2n}.$$

Notice that $|\mathcal{W}_{\mathbf{x}}\rangle = d^{-\frac{n}{2}} \operatorname{vec}(\mathcal{W}_{\mathbf{x}})$ since $\mathcal{W}_{\mathbf{x}} = \sum_{\mathbf{q}\in\mathbb{Z}_d^n} \tau^{\langle \mathbf{v},\mathbf{w}\rangle} \cdot \omega^{\langle \mathbf{q},\mathbf{v}\rangle} |\mathbf{q}+\mathbf{w}\rangle\langle\mathbf{q}|$. As previously mentioned, the re-scaled Weyl operators $\{d^{-\frac{n}{2}}\mathcal{W}_{\mathbf{x}} : \mathbf{x} \in \mathbb{Z}_d^{2n}\}$ form an orthonormal basis in the space of operators, therefore the states $\{|\mathcal{W}_{\mathbf{x}}\rangle : \mathbf{x} \in \mathbb{Z}_d^{2n}\}$ form an orthonormal basis in the Hilbert space $(\mathbb{C}^d)^{\otimes n} \otimes (\mathbb{C}^d)^{\otimes n}$. Indeed, $\langle \mathcal{W}_{\mathbf{x}}|\mathcal{W}_{\mathbf{y}}\rangle = d^{-n} \operatorname{Tr}[\mathcal{W}_{\mathbf{x}}^\dagger \mathcal{W}_{\mathbf{y}}] = \mathbf{1}[\mathbf{x} = \mathbf{y}]$. In addition, note that, for all $\mathbf{x} = (\mathbf{v},\mathbf{w}), \mathbf{y} = (\mathbf{v}',\mathbf{w}') \in \mathbb{Z}_d^{2n}$,

$$\begin{aligned}
(\mathcal{W}_{\mathbf{x}} \otimes \mathcal{W}_{\mathbf{y}})|\Phi^+\rangle &= \frac{1}{\sqrt{d^n}} \sum_{\mathbf{q}\in\mathbb{Z}_d^n} \tau^{\langle \mathbf{v},\mathbf{w}\rangle + \langle \mathbf{v}',\mathbf{w}'\rangle} \cdot \omega^{\langle \mathbf{q},\mathbf{v}+\mathbf{v}'\rangle} |\mathbf{q}+\mathbf{w}\rangle|\mathbf{q}+\mathbf{w}'\rangle \\
&= \frac{1}{\sqrt{d^n}} \sum_{\mathbf{q}\in\mathbb{Z}_d^n} \tau^{\langle \mathbf{v},\mathbf{w}\rangle + \langle \mathbf{v}',\mathbf{w}'\rangle} \cdot \omega^{\langle \mathbf{q}-\mathbf{w}',\mathbf{v}+\mathbf{v}'\rangle} |\mathbf{q}+\mathbf{w}-\mathbf{w}'\rangle|\mathbf{q}\rangle \\
&= \tau^{\langle \mathbf{v},\mathbf{w}\rangle - \langle \mathbf{v}',\mathbf{w}'\rangle - 2\langle \mathbf{v},\mathbf{w}'\rangle} \cdot \tau^{-\langle \mathbf{v}+\mathbf{v}',\mathbf{w}-\mathbf{w}'\rangle} (\mathcal{W}_{\mathbf{x}-J(\mathbf{y})} \otimes \mathbf{I})|\Phi^+\rangle \\
&= \tau^{[J(\mathbf{y}),\mathbf{x}]} (\mathcal{W}_{\mathbf{x}-J(\mathbf{y})} \otimes \mathbf{I})|\Phi^+\rangle. \qquad (3)
\end{aligned}$$

Thus $(\mathcal{W}_{\mathbf{x}} \otimes \mathcal{W}_{J(\mathbf{x})})|\Phi^+\rangle = |\Phi^+\rangle$. The following representation of $|\mathcal{W}_{\mathbf{x}}\rangle\langle\mathcal{W}_{\mathbf{x}}|$ will be useful.

Lemma 36. *For all* $\mathbf{x} \in \mathbb{Z}_d^{2n}$,

$$|\mathcal{W}_{\mathbf{x}}\rangle\langle\mathcal{W}_{\mathbf{x}}| = \frac{1}{d^{2n}} \sum_{\mathbf{y}\in\mathbb{Z}_d^{2n}} \omega^{[\mathbf{x},\mathbf{y}]} \mathcal{W}_{\mathbf{y}} \otimes \mathcal{W}_{J(\mathbf{y})}.$$

Proof. For all $\mathbf{z}, \mathbf{z}' \in \mathbb{Z}_d^{2n}$,

$$\begin{aligned}
\sum_{\mathbf{y}\in\mathbb{Z}_d^{2n}} \omega^{[\mathbf{x},\mathbf{y}]} \langle \mathcal{W}_{\mathbf{z}}|\mathcal{W}_{\mathbf{y}} \otimes \mathcal{W}_{J(\mathbf{y})}|\mathcal{W}_{\mathbf{z}'}\rangle &= \sum_{\mathbf{y}\in\mathbb{Z}_d^{2n}} \omega^{[\mathbf{x},\mathbf{y}]} \langle \Phi^+|(\mathcal{W}_{\mathbf{z}}^\dagger \otimes \mathbb{I})(\mathcal{W}_{\mathbf{y}} \otimes \mathcal{W}_{J(\mathbf{y})})(\mathcal{W}_{\mathbf{z}'} \otimes \mathbb{I})|\Phi^+\rangle \\
&= \sum_{\mathbf{y}\in\mathbb{Z}_d^{2n}} \omega^{[\mathbf{y},\mathbf{z}'-\mathbf{x}]} \langle \Phi^+|(\mathcal{W}_{\mathbf{z}}^\dagger \mathcal{W}_{\mathbf{z}'} \otimes \mathbb{I})(\mathcal{W}_{\mathbf{y}} \otimes \mathcal{W}_{J(\mathbf{y})})|\Phi^+\rangle \\
&= \sum_{\mathbf{y}\in\mathbb{Z}_d^{2n}} \omega^{[\mathbf{y},\mathbf{z}'-\mathbf{x}]} \langle \Phi^+|(\mathcal{W}_{\mathbf{z}}^\dagger \mathcal{W}_{\mathbf{z}'} \otimes \mathbb{I})|\Phi^+\rangle \\
&= \mathbf{1}[\mathbf{z} = \mathbf{z}'] \sum_{\mathbf{y}\in\mathbb{Z}_d^{2n}} \omega^{[\mathbf{y},\mathbf{z}'-\mathbf{x}]} \\
&= d^{2n} \mathbf{1}[\mathbf{z} = \mathbf{z}' = \mathbf{x}],
\end{aligned}$$

using Lemma 9 together with $(\mathbb{Z}_d^{2n})^{\perp\!\!\!\perp} = \{\mathbf{0}\}$ (the symplectic product is non-degenerate). ∎

Given a pure state of $2n$ qudits divided into systems $A_1, \ldots, A_n$ and $B_1, \ldots, B_n$, we call *Bell sampling* the operation of measuring each pair A_iB_i of qudits in the generalised Bell basis, which returns a vector in $\mathbb{Z}_d^{2n}$. A similar operation, called *Bell difference sampling*, performs Bell sampling twice and subtracts the results from each other.

Definition 37 (Bell sampling)**.** *Bell sampling is the projective measurement given by* $|\mathcal{W}_{\mathbf{x}}\rangle\langle\mathcal{W}_{\mathbf{x}}|$ *for* $\mathbf{x} \in \mathbb{Z}_d^{2n}$.

Definition 38 (Bell difference sampling)**.** *Bell difference sampling is defined as performing Bell sampling twice and subtracting the results from each other (modulo d). More precisely, it is the projective measurement given by*

$$\Pi_{\mathbf{x}} = \sum_{\mathbf{y}\in\mathbb{Z}_d^{2n}} |\mathcal{W}_{\mathbf{y}}\rangle\langle\mathcal{W}_{\mathbf{y}}| \otimes |\mathcal{W}_{\mathbf{x}+\mathbf{y}}\rangle\langle\mathcal{W}_{\mathbf{x}+\mathbf{y}}|, \quad \forall \mathbf{x} \in \mathbb{Z}_d^{2n}.$$

Performing Bell sampling on $|\psi_1\rangle|\psi_2\rangle$ is equivalent to sampling from a surprisingly simple probability distribution.

Lemma 39. *Given two pure states $|\psi_1\rangle, |\psi_2\rangle \in (\mathbb{C}^d)^{\otimes n}$, Bell sampling on $|\psi_1\rangle|\psi_2\rangle$ returns $\mathbf{x} \in \mathbb{Z}_d^{2n}$ with probability $d^{-n}|\langle\psi_1|\mathcal{W}_{\mathbf{x}}|\psi_2^*\rangle|^2$. In particular, if $|\psi_1\rangle = |\psi_2^*\rangle = |\psi\rangle$, Bell sampling on $|\psi\rangle|\psi^*\rangle$ corresponds to sampling from the distribution $p_\psi(\mathbf{x})$.*

Proof. We have $|\langle\mathcal{W}_{\mathbf{x}}|(|\psi_1\rangle|\psi_2\rangle)|^2 = d^{-n}|\mathrm{Tr}[\mathcal{W}_{\mathbf{x}}^\dagger|\psi_1\rangle\langle\psi_2^*|]|^2 = d^{-n}|\langle\psi_1|\mathcal{W}_{\mathbf{x}}|\psi_2^*\rangle|^2$, since $|\psi_1\rangle|\psi_2\rangle = \mathrm{vec}(|\psi_1\rangle\langle\psi_2^*|)$. ∎

As an immediate consequence of the above result, performing Bell sampling on $|\mathcal{S}\rangle|\mathcal{S}^*\rangle$ is equivalent to sampling from $p_{\mathcal{S}}(\mathbf{x})$, which according to Equation (2) is constant on $\mathscr{M}$. This means that Bell sampling on $|\mathcal{S}\rangle|\mathcal{S}^*\rangle$ returns an element of $\mathscr{M}$.

It is well known that the transformation $|\psi\rangle \mapsto |\psi^*\rangle$ cannot be implemented by a physical process since it is not completely positive. Therefore, there is little hope to sample from $p_\psi(\mathbf{x})$ using Bell sampling with only access to copies of $|\psi\rangle$. The situation is more subtle if we restrict ourselves to stabiliser states $|\mathcal{S}\rangle$. First observe that $|\mathcal{S}^*\rangle$ is itself a stabiliser state since $\mathcal{W}_{\mathbf{v},\mathbf{w}}^* = (-1)^{(d+1)\langle\mathbf{v},\mathbf{w}\rangle}\mathcal{W}_{J(\mathbf{v},\mathbf{w})}$. This means that if $|\mathcal{S}\rangle = |\mathscr{M}, s\rangle$, then $|\mathcal{S}^*\rangle = |J(\mathscr{M}), t\rangle$, where $\omega^{t(\mathbf{x})} = \omega^{-s(\mathbf{x})}(-1)^{(d+1)\langle\mathbf{v},\mathbf{w}\rangle}$.

For qubits ($d = 2$), the involution operation J is trivial, i.e., $J(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}_2^{2n}$. As a consequence, the stabiliser state $|\mathcal{S}^*\rangle$ is characterised by the same Lagrangian subspace as $|\mathcal{S}\rangle$, possibly only with a different phase. Therefore, $|\mathcal{S}^*\rangle = \mathcal{W}_{\mathbf{z}}|\mathcal{S}\rangle$ for some $\mathbf{z} \in \mathbb{F}_2^{2n}$, i.e., $|\mathcal{S}^*\rangle$ is related to $|\mathcal{S}\rangle$ via a Weyl operator $\mathcal{W}_{\mathbf{z}}$, which ultimately is the reason why a simple Bell sampling on $|\mathcal{S}\rangle^{\otimes 2}$ works. Allcock et al. [ADIS24] explored Bell (difference) sampling on $|\mathcal{S}\rangle^{\otimes 4}$ for any[2] dimension d and found that it returns a random string $\mathbf{x}$ in $\mathscr{M} + J(\mathscr{M})$ which, in the worst case, is the whole space $\mathbb{Z}_d^{2n}$. A new idea is thus needed in order to sample elements from $\mathscr{M}$.

We now show how one can uniformly sample from the Lagrangian submodule $\mathscr{M}$. At the core of our sampling procedure, we develop an algorithm that maps copies of a (possibly unknown) stabiliser state $|\mathcal{S}\rangle$ into its complex conjugate $|\mathcal{S}^*\rangle$. More precisely, we show that there is a unitary $\mathcal{B}_{\mathbf{R}} \in ((\mathbb{C}^d)^{\otimes n})^{\otimes 4}$ that transforms four copies of $|\mathcal{S}\rangle$ into four copies of $|\mathcal{S}^*\rangle$ up to some Weyl operator and phase depending on $\mathcal{S}$. Interestingly enough, the unitary $\mathcal{B}_{\mathbf{R}}$ is constructed using Lagrange's four-square theorem, which states that every positive integer can be written as the sum of four integer squares (hence why we need four copies of $|\mathcal{S}\rangle$). In the following, we denote the functional dependence of $\mathbf{v} \in \mathbb{Z}_d^n$ and $s \in \mathbb{Z}_d$ on a stabiliser group $\mathcal{S}$ as $\mathbf{v}(\mathcal{S})$ and $s(\mathcal{S})$.

Definition 40. *For integers $d \geqslant 2$, $k \geqslant 1$, and matrix $\mathbf{R} \in \mathbb{Z}^{k\times k}$, let the unitary $\mathcal{B}_{\mathbf{R}} \in ((\mathbb{C}^d)^{\otimes n})^{\otimes k}$*

$$\mathcal{B}_{\mathbf{R}} : |\mathbf{Q}\rangle \mapsto |\mathbf{QR} \pmod d\rangle, \quad \text{where} \quad \mathbf{Q} = [\mathbf{q}_1, \ldots, \mathbf{q}_k] \in \mathbb{Z}_d^{n\times k}.$$

Fact 41 (Lagrange's four-square theorem)**.** *For every $m \in \mathbb{N}$, there are $a_1, a_2, a_3, a_4 \in \mathbb{N} \cup \{0\}$ such that $m = a_1^2 + a_2^2 + a_3^2 + a_4^2$.*

Lemma 42. *For integer $d \geqslant 2$, let $a_1, a_2, a_3, a_4 \in \mathbb{N} \cup \{0\}$ such that $a_1^2 + a_2^2 + a_3^2 + a_4^2 = D - 1$. Then the matrix*

$$\mathbf{R} = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ a_2 & -a_1 & a_4 & -a_3 \\ a_3 & -a_4 & -a_1 & a_2 \\ a_4 & a_3 & -a_2 & -a_1 \end{pmatrix} \in \mathbb{Z}^{4\times 4}$$

[2] Allcock et al. [ADIS24] technically focused on prime dimensions, but their results can be generalised to any d.

satisfies $\mathbf{R}^\top\mathbf{R} = \mathbf{R}\mathbf{R}^\top = (D-1)\mathbf{I}$.

Theorem 43. *For every integer $d \geqslant 2$, consider the unitary $\mathcal{B}_\mathbf{R} \in ((\mathbb{C}^d)^{\otimes n})^{\otimes 4}$ from Definition 40 with any $\mathbf{R} \in \mathbb{Z}^{4\times 4}$ such that $\mathbf{R}^\top\mathbf{R} = \mathbf{R}\mathbf{R}^\top = (D-1)\mathbf{I}$. Then for any stabiliser group $\mathcal{S}$,*

$$\mathcal{B}_\mathbf{R}|\mathcal{S}\rangle^{\otimes 4} = \left(\bigotimes_{i=1}^{4} \mathcal{W}_{\mathbf{0},\mathbf{w}_i(\mathcal{S})}\mathcal{W}_{\mathbf{v}_i(\mathcal{S}),\mathbf{0}}\right)\omega^{s(\mathcal{S})}|\mathcal{S}^*\rangle^{\otimes 4}$$

for some $\mathbf{v}_1(\mathcal{S}), \dots, \mathbf{v}_4(\mathcal{S}), \mathbf{w}_1(\mathcal{S}), \dots, \mathbf{w}_4(\mathcal{S}) \in \mathbb{Z}_d^n$ and $s(\mathcal{S}) \in \mathbb{Z}_d$ depending on $\mathcal{S}$.

Proof. According to Lemma 23, every stabiliser state $|\mathcal{S}\rangle$ can be written as (up to normalisation)

$$|\mathcal{S}\rangle = \sum_{\mathbf{q}\in\mathbb{Z}_d^\ell} \omega^{\mathbf{p}^\top\mathbf{q}} \cdot \tau^{\mathbf{q}^\top\mathbf{M}\mathbf{q}}|\mathbf{W}\mathbf{q}+\mathbf{u}\rangle, \quad \text{where } \mathbf{p} = \mathbf{s} + \mathbf{V}^\top\mathbf{u} \text{ and } \mathbf{M} = \mathbf{V}^\top\mathbf{W}.$$

Write $\mathbf{q}^\top\mathbf{M}\mathbf{q} = \operatorname{Tr}[\mathbf{M}\mathbf{q}\mathbf{q}^\top]$ and $\mathbf{p}^\top\mathbf{q} = \operatorname{Tr}[\mathbf{q}\mathbf{p}^\top]$. Then

$$\begin{aligned}|\mathcal{S}\rangle^{\otimes 4} &= \sum_{\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3,\mathbf{q}_4\in\mathbb{Z}_d^\ell} \tau^{\operatorname{Tr}[\mathbf{M}\sum_{i=1}^4 \mathbf{q}_i\mathbf{q}_i^\top]} \cdot \omega^{\operatorname{Tr}[\sum_{i=1}^4 \mathbf{q}_i\mathbf{p}^\top]}|\mathbf{W}\mathbf{q}_1+\mathbf{u}, \cdots, \mathbf{W}\mathbf{q}_4+\mathbf{u}\rangle \\ &= \sum_{\mathbf{Q}\in\mathbb{Z}_d^{\ell\times 4}} \tau^{\operatorname{Tr}[\mathbf{M}\mathbf{Q}\mathbf{Q}^\top]} \cdot \omega^{\operatorname{Tr}[\mathbf{Q}\mathbf{P}^\top]}|\mathbf{W}\mathbf{Q}+\mathbf{U}\rangle,\end{aligned}$$

where $\mathbf{Q} = [\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}_4] \in \mathbb{Z}_d^{\ell\times 4}$ is a matrix with columns $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}_4$, and similarly for $\mathbf{U} = [\mathbf{u}, \mathbf{u}, \mathbf{u}, \mathbf{u}] \in \mathbb{Z}_d^{\ell\times 4}$ and $\mathbf{P} = [\mathbf{p}, \mathbf{p}, \mathbf{p}, \mathbf{p}] \in \mathbb{Z}_d^{\ell\times 4}$.

Consider now the unitary $\mathcal{B}_\mathbf{R} : |\mathbf{Q}\rangle \mapsto |\mathbf{Q}\mathbf{R} \pmod d\rangle$ from Definition 40, where $\mathbf{R} \in \mathbb{Z}^{4\times 4}$ is such that $\mathbf{R}^\top\mathbf{R} = \mathbf{R}\mathbf{R}^\top = (D-1)\mathbf{I}$. Notice that $\mathbf{Y} \equiv \mathbf{Q}\mathbf{R} \pmod d \implies \mathbf{Y} \equiv \mathbf{Q}\mathbf{R} + \mathbf{E} \pmod D$ for some $\mathbf{E} \in \mathbb{Z}_D^{\ell\times 4}$ such that $\mathbf{E} \equiv \mathbf{0} \pmod d$ and $2\mathbf{E} \equiv \mathbf{0} \pmod D$. We can see that

$$\begin{aligned}\mathcal{B}_\mathbf{R}|\mathcal{S}\rangle^{\otimes 4} &= \sum_{\mathbf{Q}\in\mathbb{Z}_d^{\ell\times 4}} \tau^{\operatorname{Tr}[\mathbf{M}\mathbf{Q}\mathbf{Q}^\top]} \cdot \omega^{\operatorname{Tr}[\mathbf{Q}\mathbf{P}^\top]}|\mathbf{W}\mathbf{Q}\mathbf{R}+\mathbf{U}\mathbf{R}\rangle \\ &= \sum_{\mathbf{Q}\in\mathbb{Z}_d^{\ell\times 4}} \tau^{\operatorname{Tr}[\mathbf{M}(-\mathbf{Q}\mathbf{R}^\top \pmod d)(-\mathbf{Q}\mathbf{R}^\top \pmod d)^\top]} \cdot \omega^{-\operatorname{Tr}[\mathbf{Q}\mathbf{R}^\top\mathbf{P}^\top]}|\mathbf{W}\mathbf{Q}+\mathbf{U}\mathbf{R}\rangle \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (\text{by } \mathbf{Q} \leftarrow -\mathbf{Q}\mathbf{R}^\top \pmod d) \\ &= \sum_{\mathbf{Q}\in\mathbb{Z}_d^{\ell\times 4}} \tau^{\operatorname{Tr}[\mathbf{M}(\mathbf{E}-\mathbf{Q}\mathbf{R}^\top)(\mathbf{E}-\mathbf{Q}\mathbf{R}^\top)^\top]} \cdot \omega^{-\operatorname{Tr}[\mathbf{Q}\mathbf{R}^\top\mathbf{P}^\top]}|\mathbf{W}\mathbf{Q}+\mathbf{U}\mathbf{R}\rangle \quad (\text{for some } \mathbf{E} \in \mathbb{Z}_D^{\ell\times 4}) \\ &= \sum_{\mathbf{Q}\in\mathbb{Z}_d^{\ell\times 4}} \tau^{\operatorname{Tr}[\mathbf{M}\mathbf{Q}\mathbf{R}^\top\mathbf{R}\mathbf{Q}^\top-\mathbf{M}\mathbf{E}\mathbf{R}\mathbf{Q}^\top-\mathbf{M}\mathbf{Q}\mathbf{R}^\top\mathbf{E}^\top+\mathbf{M}\mathbf{E}\mathbf{E}^\top]} \cdot \omega^{-\operatorname{Tr}[\mathbf{Q}\mathbf{R}^\top\mathbf{P}^\top]}|\mathbf{W}\mathbf{Q}+\mathbf{U}\mathbf{R}\rangle.\end{aligned}$$

Notice now that, if d is odd, then $2\mathbf{E} \equiv \mathbf{0} \pmod D \implies \mathbf{E} \equiv \mathbf{0} \pmod{2d}$. If d is even, then $\mathbf{E} \equiv \mathbf{0} \pmod d \implies \mathbf{E}^\top \equiv \mathbf{0} \pmod 2$, i.e., all the entries of $\mathbf{E}^\top$ are even, and so $\mathbf{E}\mathbf{E}^\top \equiv 2\mathbf{E}\mathbf{I}_{\ell\times 4} \equiv \mathbf{0} \pmod{2d}$. In both cases, $\mathbf{E}\mathbf{E}^\top \equiv \mathbf{0} \pmod D$. Regarding the term $\operatorname{Tr}[\mathbf{M}\mathbf{E}\mathbf{R}\mathbf{Q}^\top + \mathbf{M}\mathbf{Q}\mathbf{R}^\top\mathbf{E}^\top] = \operatorname{Tr}[(\mathbf{M}+\mathbf{M}^\top)\mathbf{E}\mathbf{R}\mathbf{Q}^\top]$ (using the trace cyclicality), we know that $\mathbf{M} = \mathbf{M}^\top$ is symmetrical, therefore $\operatorname{Tr}[(\mathbf{M}+\mathbf{M}^\top)\mathbf{E}\mathbf{R}\mathbf{Q}^\top] = \operatorname{Tr}[\mathbf{M}(2\mathbf{E})\mathbf{R}\mathbf{Q}^\top] = 0$, since $2\mathbf{E} \equiv \mathbf{0} \pmod D$. Hence, the expression $\mathcal{B}_\mathbf{R}|\mathcal{S}\rangle^{\otimes 4}$ simplifies to

$$\begin{aligned}\mathcal{B}_\mathbf{R}|\mathcal{S}\rangle^{\otimes 4} &= \sum_{\mathbf{Q}\in\mathbb{Z}_d^{\ell\times 4}} \tau^{\operatorname{Tr}[\mathbf{M}\mathbf{Q}\mathbf{R}^\top\mathbf{R}\mathbf{Q}^\top]} \cdot \omega^{-\operatorname{Tr}[\mathbf{Q}\mathbf{R}^\top\mathbf{P}^\top]}|\mathbf{W}\mathbf{Q}+\mathbf{U}\mathbf{R}\rangle \\ &= \sum_{\mathbf{Q}\in\mathbb{Z}_d^{\ell\times 4}} \tau^{-\operatorname{Tr}[\mathbf{M}\mathbf{Q}\mathbf{Q}^\top]} \cdot \omega^{-\operatorname{Tr}[\mathbf{Q}\mathbf{R}^\top\mathbf{P}^\top]}|\mathbf{W}\mathbf{Q}+\mathbf{U}\mathbf{R}\rangle. \qquad (\text{since } \mathbf{R}^\top\mathbf{R} = (D-1)\mathbf{I})\end{aligned}$$

The above expression is almost $|\mathcal{S}^*\rangle^{\otimes 4}$, the only issue being the remaining matrices $\mathbf{R}$. It thus only remains to show that the above expression can be transformed into $|\mathcal{S}^*\rangle^{\otimes 4}$ via some Weyl operator. For this, we use that, according to Lemma 23, $\mathbf{p} \in \mathrm{row}(\mathbf{W}) \implies \mathbf{p} = \mathbf{W}^\top \mathbf{t}$ for some $\mathbf{t} \in \mathbb{Z}_d^n$. Let $\mathbf{T} = [\mathbf{t}, \mathbf{t}, \mathbf{t}, \mathbf{t}] \in \mathbb{Z}_d^{\ell \times 4}$. In the following, let $(\mathbf{R} - \mathbf{I})_i$ be the i-th column of $\mathbf{R} - \mathbf{I} \in \mathbb{Z}_d^{4\times 4}$. We note that, by the definition of the Weyl operator $\mathcal{W}_{\mathbf{v},\mathbf{w}}$,

$$\begin{aligned}
\left(\bigotimes_{i=1}^{4} \mathcal{W}_{\mathbf{0},\mathbf{U}(\mathbf{R}-\mathbf{I})_i} \mathcal{W}_{\mathbf{T}(\mathbf{I}-\mathbf{R})_i,\mathbf{0}}\right) |\mathbf{WQ} + \mathbf{U}\rangle &= \omega^{\sum_{i=1}^{4} (\mathbf{WQ}_i + \mathbf{U}_i)^\top (\mathbf{T}(\mathbf{I}-\mathbf{R})_i)} |\mathbf{WQ} + \mathbf{U} + \mathbf{U}(\mathbf{R} - \mathbf{I})\rangle \\
&= \omega^{\mathrm{Tr}[(\mathbf{T}(\mathbf{I}-\mathbf{R})(\mathbf{WQ}+\mathbf{U})^\top]} |\mathbf{WQ} + \mathbf{UR}\rangle \\
&= \omega^{\mathrm{Tr}[\mathbf{TQ}^\top \mathbf{W}^\top - \mathbf{TRQ}^\top \mathbf{W}^\top + \mathbf{T}(\mathbf{I}-\mathbf{R})\mathbf{U}^\top]} |\mathbf{WQ} + \mathbf{UR}\rangle \\
&= \omega^{\mathrm{Tr}[\mathbf{QP}^\top - \mathbf{QR}^\top \mathbf{P}^\top + \mathbf{T}(\mathbf{I}-\mathbf{R})\mathbf{U}^\top]} |\mathbf{WQ} + \mathbf{UR}\rangle.
\end{aligned}$$

In summary,

$$\omega^{-\mathrm{Tr}[\mathbf{QR}^\top \mathbf{P}^\top]} |\mathbf{WQ} + \mathbf{UR}\rangle = \left(\bigotimes_{i=1}^{4} \mathcal{W}_{\mathbf{0},\mathbf{U}(\mathbf{R}-\mathbf{I})_i} \mathcal{W}_{\mathbf{T}(\mathbf{I}-\mathbf{R})_i,\mathbf{0}}\right) \omega^{\mathrm{Tr}[-\mathbf{QP}^\top + \mathbf{T}(\mathbf{R}-\mathbf{I})\mathbf{U}^\top]} |\mathbf{WQ} + \mathbf{U}\rangle,$$

which means that

$$\mathcal{B}_{\mathbf{R}} |\mathcal{S}\rangle^{\otimes 4} = \left(\bigotimes_{i=1}^{4} \mathcal{W}_{\mathbf{0},\mathbf{U}(\mathbf{R}-\mathbf{I})_i} \mathcal{W}_{\mathbf{T}(\mathbf{I}-\mathbf{R})_i,\mathbf{0}}\right) \omega^{\mathrm{Tr}[\mathbf{T}(\mathbf{R}-\mathbf{I})\mathbf{U}^\top]} |\mathcal{S}^*\rangle^{\otimes 4}. \qquad \blacksquare$$

Remark 44. *Since $\mathcal{B}_{\mathbf{R}}^\dagger = \mathcal{B}_{-\mathbf{R}^\top}$, then also $\mathcal{B}_{\mathbf{R}}^\dagger |\mathcal{S}\rangle^{\otimes 4} = \mathcal{P}(\mathcal{S}) |\mathcal{S}^*\rangle^{\otimes 4}$ for some Pauli $\mathcal{P}(\mathcal{S})$ depending on $\mathcal{S}$.*

From the above proof, it should be clear that the number of copies of $|\mathcal{S}\rangle$ required by the unitary $\mathcal{B}_{\mathbf{R}}$ is bounded by the minimum number of integer squares required to decompose $D-1$. As a consequence, specific dimensions d might require less copies of $|\mathcal{S}\rangle$. As examples, the sum-of-two-squares theorem states that $m \in \mathbb{N}$ can be written as the sum of two squares of integers, $m = a_1^2 + a_2^2$, if and only if its prime decomposition contains no factor p^k, where prime $p \equiv 3 \pmod 4$ and k is odd. More specifically for odd prime dimensions, it is well known that if $d \equiv 3 \pmod 4$, then there are $a_1, a_2 \in \mathbb{F}_d$ such that $a_1^2 + a_2^2 \equiv -1 \pmod d$. Also, $d \equiv 1 \pmod 4$ if and only if -1 is a quadratic residue modulo d, i.e., there is $a_1 \in \mathbb{F}_d$ such that $a_1^2 \equiv -1 \pmod d$. In this case, the unitary $\mathcal{B}_{\mathbf{R}}$ can be replaced by the identity. We summarise these remarks in the next result.

Corollary 45. *Let $|\mathcal{S}\rangle \in (\mathbb{C}^d)^{\otimes n}$ be any stabiliser state. Then:*

1. *If $p^k \nmid (D-1)$ for prime $p \equiv 3 \pmod 4$ and k odd, then there is a unitary that maps $|\mathcal{S}\rangle^{\otimes 2} \mapsto (\bigotimes_{i=1}^{2} \mathcal{W}_{\mathbf{x}_i}) \tau^s |\mathcal{S}^*\rangle^{\otimes 2}$ for all $|\mathcal{S}\rangle \in (\mathbb{C}^d)^{\otimes n}$, where $\mathbf{x}_i$ and s depend on $\mathcal{S}$;*
2. *If d is prime and $d \equiv 3 \pmod 4$, there is a unitary that maps $|\mathcal{S}\rangle^{\otimes 2} \mapsto (\bigotimes_{i=1}^{2} \mathcal{W}_{\mathbf{x}_i}) \omega^s |\mathcal{S}^*\rangle^{\otimes 2}$ for all $|\mathcal{S}\rangle \in (\mathbb{C}^d)^{\otimes n}$, where $\mathbf{x}_i$ and s depend on $\mathcal{S}$;*
3. *If d is prime and $d \equiv 1 \pmod 4$, then $|\mathcal{S}\rangle = \mathcal{W}_{\mathbf{x}} \omega^s |\mathcal{S}^*\rangle$ where $\mathbf{x}$ and s depend on $\mathcal{S}$.*

As previously discussed, for qubits ($d = 2$) both $|\mathcal{S}\rangle$ and $|\mathcal{S}^*\rangle$ are related by a Weyl operator, so there is no need to perform any transformation. If we employ Theorem 43 and Corollary 45 for $d = 2$, we paradoxically find that 3 copies of $|\mathcal{S}\rangle$ would be required to obtain any copy of $|\mathcal{S}^*\rangle$ (since $D - 1 = 3$ for $d = 2$). However, since $\tau^d = \omega = -1$ for $d = 2$ and the quadratic term $\mathbf{q}^\top \mathbf{Mq}$ becomes linear over $\mathbb{F}_2$, i.e., $(-1)^{\mathbf{q}^\top \mathbf{Mq}} = (-1)^{\sum_{i=1}^{n} M_{ii} q_i}$ due to $\mathbf{M}$ being symmetric and $q_i^2 = q_i \in \mathbb{F}_2$, one only needs to decompose $d - 1$, and not $D - 1 = 2d - 1$, into integer squares,

which is trivial as $d-1=1$. Therefore, the procedure explained in Theorem 43 also recovers the qubit case ($\mathcal{B}_2$ is the identity).

Finally, we mention that there are randomised algorithms for computing a single representation $m=a_1^2+a_2^2+a_3^2+a_4^2$ of a given $m\in\mathbb{N}$ in expected time $O(\log^2 m/\log\log m)$ [RS86, PT18].

Even though in Theorem 43 we phrased the action of $\mathcal{B}_{\mathbf{R}}$ onto $|\mathcal{S}\rangle^{\otimes 4}$, it is possible to describe the effect of $\mathcal{B}_{\mathbf{R}}$ onto tensor products of Weyl operators.

Theorem 46. *Let $\mathcal{B}_{\mathbf{R}}\in((\mathbb{C}^d)^{\otimes n})^{\otimes 4}$ be the unitary from Definition 40 with $\mathbf{R}\in\mathbb{Z}^{4\times 4}$ such that $\mathbf{R}^\top\mathbf{R}=\mathbf{R}\mathbf{R}^\top=(D-1)\mathbf{I}$. Then, for all $\mathbf{X}=[\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4]\in\mathbb{Z}_d^{2n\times 4}$,*

$$\mathcal{B}_{\mathbf{R}}\left(\bigotimes_{i\in[4]}\mathcal{W}_{\mathbf{x}_i}\right)\mathcal{B}_{\mathbf{R}}^\dagger=\bigotimes_{i\in[4]}\mathcal{W}_{J((\mathbf{X}\mathbf{R})_i)},$$

where $(\mathbf{X}\mathbf{R})_i$ is the i-th column of $\mathbf{X}\mathbf{R}$ and $J:\mathbb{Z}_d^{2n}\to\mathbb{Z}_d^{2n}$ is the involution.

Proof. Write $\mathbf{x}_i=(\mathbf{v}_i,\mathbf{w}_i)$ for $i\in[4]$ and let $\mathbf{V}=[\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3,\mathbf{v}_4],\mathbf{W}=[\mathbf{w}_1,\mathbf{w}_2,\mathbf{w}_3,\mathbf{w}_4]\in\mathbb{Z}_d^{n\times 4}$. For any $\mathbf{Q}\in\mathbb{Z}_d^{n\times 4}$,

$$\begin{aligned}
\mathcal{B}_{\mathbf{R}}\left(\bigotimes_{i\in[4]}\mathcal{W}_{\mathbf{x}_i}\right)\mathcal{B}_{\mathbf{R}}^\dagger|\mathbf{Q}\rangle &= \mathcal{B}_{\mathbf{R}}\left(\bigotimes_{i\in[4]}\mathcal{W}_{\mathbf{x}_i}\right)|-\mathbf{Q}\mathbf{R}^\top\rangle\\
&= \tau^{\sum_{i=1}^4\langle\mathbf{v}_i,\mathbf{w}_i\rangle}\cdot\omega^{-\sum_{i=1}^4\langle\mathbf{v}_i,(\mathbf{Q}\mathbf{R}^\top)_i\rangle}\mathcal{B}_{\mathbf{R}}|\mathbf{W}-\mathbf{Q}\mathbf{R}^\top\rangle\\
&= \tau^{\sum_{i=1}^4\langle\mathbf{v}_i,\mathbf{w}_i\rangle}\cdot\omega^{-\sum_{i=1}^4\langle\mathbf{v}_i,(\mathbf{Q}\mathbf{R}^\top)_i\rangle}|\mathbf{W}\mathbf{R}+\mathbf{Q}\rangle \qquad (\mathbf{R}^\top\mathbf{R}=(D-1)\mathbf{I})\\
&= \tau^{\mathrm{Tr}[\mathbf{W}\mathbf{V}^\top]}\cdot\omega^{-\mathrm{Tr}[\mathbf{Q}\mathbf{R}^\top\mathbf{V}^\top]}|\mathbf{W}\mathbf{R}+\mathbf{Q}\rangle\\
&= \tau^{-\mathrm{Tr}[\mathbf{W}\mathbf{R}\mathbf{R}^\top\mathbf{V}^\top]}\cdot\omega^{-\mathrm{Tr}[\mathbf{Q}\mathbf{R}^\top\mathbf{V}^\top]}|\mathbf{W}\mathbf{R}+\mathbf{Q}\rangle \qquad (\mathbf{R}^\top\mathbf{R}=(D-1)\mathbf{I})\\
&= \tau^{\mathrm{Tr}[\mathbf{W}\mathbf{R}(-\mathbf{V}\mathbf{R})^\top]}\cdot\omega^{\mathrm{Tr}[\mathbf{Q}(-\mathbf{V}\mathbf{R})^\top]}|\mathbf{W}\mathbf{R}+\mathbf{Q}\rangle\\
&= \bigotimes_{i\in[4]}\mathcal{W}_{J((\mathbf{X}\mathbf{R})_i)}|\mathbf{Q}\rangle. \qquad \blacksquare
\end{aligned}$$

We define a modified Bell sampling procedure.

Definition 47 (U-skewed Bell sampling)**.** *Let $U\in((\mathbb{C}^d)^{\otimes n})^{\otimes 2k}$ be a unitary. The U-skewed Bell sampling on $2k$ states in $(\mathbb{C}^d)^{\otimes n}$ is the projective measurement*

$$U^\dagger\left(\bigotimes_{i\in[k]}|\mathcal{W}_{\mathbf{x}_i}\rangle\langle\mathcal{W}_{\mathbf{x}_i}|\right)U \qquad \forall\mathbf{x}_1,\dots,\mathbf{x}_k\in\mathbb{Z}_d^{2n}.$$

Definition 48 (U-skewed Bell difference sampling)**.** *Let $U\in((\mathbb{C}^d)^{\otimes n})^{\otimes 4k}$ be a unitary. The U-skewed Bell difference sampling on $4k$ states in $(\mathbb{C}^d)^{\otimes n}$ is the projective measurement*

$$U^\dagger\left(\bigotimes_{i\in[k]}\Pi_{\mathbf{x}_i}\right)U \quad \text{where} \quad \Pi_{\mathbf{x}_i}=\sum_{\mathbf{y}\in\mathbb{Z}_d^{2n}}|\mathcal{W}_{\mathbf{y}}\rangle\langle\mathcal{W}_{\mathbf{y}}|\otimes|\mathcal{W}_{\mathbf{x}_i+\mathbf{y}}\rangle\langle\mathcal{W}_{\mathbf{x}_i+\mathbf{y}}|, \quad \forall\mathbf{x}_1,\dots,\mathbf{x}_k\in\mathbb{Z}_d^{2n}.$$

Performing a skewed Bell sampling on copies of a stabiliser state $|\mathcal{S}\rangle=|\mathscr{M},s\rangle$ returns elements from $\mathscr{M}+\mathbf{z}$ for some $\mathbf{z}\in\mathbb{Z}_d^{2n}$ depending on $\mathcal{S}$.

Theorem 49. *Let $|\mathcal{S}\rangle\in(\mathbb{C}^d)^{\otimes n}$ be a stabiliser state. Let $\mathcal{B}_{\mathbf{R}}\in((\mathbb{C}^d)^{\otimes n})^{\otimes 4}$ be the unitary from Definition 40 with $\mathbf{R}\in\mathbb{Z}^{4\times 4}$ such that $\mathbf{R}^\top\mathbf{R}=\mathbf{R}\mathbf{R}^\top=(D-1)\mathbf{I}$. Let $\mathcal{U}_{\mathbf{R}}\in((\mathbb{C}^d)^{\otimes n})^{\otimes 8}$ be the unitary acting on $\bigotimes_{i\in[8]}|\mathbf{q}_i\rangle$ that applies $\mathcal{B}_{\mathbf{R}}$ onto $|\mathbf{q}_2,\mathbf{q}_4,\mathbf{q}_6,\mathbf{q}_8\rangle$ and the identity onto the remaining registers. Then $\mathcal{U}_{\mathbf{R}}$-skewed Bell sampling on $|\mathcal{S}\rangle^{\otimes 8}$ corresponds to sampling from the distribution $\prod_{i=1}^4 p_{\mathcal{S}}(\mathbf{x}_i+\mathbf{z}_i(\mathcal{S}))$ for some $\mathbf{z}_1(\mathcal{S}),\dots,\mathbf{z}_4(\mathcal{S})\in\mathbb{Z}_d^{2n}$.*

Proof. The probability distribution is simply

$$\begin{aligned}
\left|\left(\bigotimes_{i=1}^{4}\langle\mathcal{W}_{\mathbf{x}_i}|\right)\mathcal{U}_{\mathbf{R}}^{\dagger}|\mathcal{S}\rangle^{\otimes 8}\right|^2 &= \left|\left(\bigotimes_{i=1}^{4}\langle\mathcal{W}_{\mathbf{x}_i}|\right)\bigotimes_{i=1}^{4}((\mathbf{I}\otimes\mathcal{W}_{\mathbf{z}_i})|\mathcal{S}\rangle|\mathcal{S}^*\rangle)\right|^2 && \text{(by Theorem 43)}\\
&= \prod_{i=1}^{4}|\langle\mathcal{W}_{\mathbf{x}_i+J(\mathbf{z}_i)}|(|\mathcal{S}\rangle|\mathcal{S}^*\rangle)|^2 && \text{(by Equation (3) and } \mathcal{W}_{\mathbf{y}}^{\dagger}=\mathcal{W}_{-\mathbf{y}})\\
&= \prod_{i=1}^{4}p_{\mathcal{S}}(\mathbf{x}_i+J(\mathbf{z}_i)) && \text{(by Lemma 39)}
\end{aligned}$$

for some $\mathbf{z}_1,\dots,\mathbf{z}_4\in\mathbb{Z}_d^{2n}$ that depend on $\mathcal{S}$. Simply relabel $\mathbf{z}_1,\dots,\mathbf{z}_4$ to $J(\mathbf{z}_1),\dots,J(\mathbf{z}_4)$. ∎

We now explore skewed Bell difference sampling on $|\psi\rangle^{\otimes 16}$ for any state $|\psi\rangle\in(\mathbb{C}^d)^{\otimes n}$ and not only stabiliser states.

Theorem 50. *Let $|\psi\rangle\in(\mathbb{C}^d)^{\otimes n}$ be a pure state. Let $\mathcal{B}_{\mathbf{R}}\in((\mathbb{C}^d)^{\otimes n})^{\otimes 4}$ be the unitary from Definition 40 with $\mathbf{R}\in\mathbb{Z}^{4\times 4}$ such that $\mathbf{R}^{\top}\mathbf{R}=\mathbf{R}\mathbf{R}^{\top}=(D-1)\mathbf{I}$. Let $\mathcal{U}_{\mathbf{R}}\in((\mathbb{C}^d)^{\otimes n})^{\otimes 8}$ be the unitary acting on $\bigotimes_{i\in[8]}|\mathbf{q}_i\rangle$ that applies $\mathcal{B}_{\mathbf{R}}$ onto $|\mathbf{q}_2,\mathbf{q}_4,\mathbf{q}_6,\mathbf{q}_8\rangle$ and the identity onto the remaining registers. Then $(\mathcal{U}_{\mathbf{R}}\otimes\mathcal{U}_{\mathbf{R}})$-skewed Bell difference sampling on $|\psi\rangle^{\otimes 16}$ corresponds to sampling from the distribution*

$$\begin{aligned}
b_{\psi}(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4) &\triangleq \operatorname{Tr}\left[(\mathcal{U}_{\mathbf{R}}\otimes\mathcal{U}_{\mathbf{R}})\left(\bigotimes_{i\in[4]}\Pi_{\mathbf{x}_i}\right)(\mathcal{U}_{\mathbf{R}}^{\dagger}\otimes\mathcal{U}_{\mathbf{R}}^{\dagger})\psi^{\otimes 16}\right]\\
&= \sum_{\mathbf{Y}\in\mathbb{Z}_d^{2n\times 4}}\prod_{i=1}^{4}p_{\psi}(\mathbf{x}_i+\mathbf{Y}_i)p_{\psi}((\mathbf{Y}\mathbf{R})_i).
\end{aligned}$$

If $|\psi\rangle=|\mathcal{S}\rangle=|\mathscr{M},s\rangle$ is a stabiliser state, then

$$b_{\mathcal{S}}(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4)=\begin{cases}d^{-4n} & \text{if } \mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4\in\mathscr{M},\\ 0 & \text{otherwise.}\end{cases}$$

We call the distribution $b_{\psi}(\mathbf{x})$ the $\mathbf{R}$-skewed (or simply skewed) Bell distribution of $|\psi\rangle$.

Proof. First note that

$$\begin{aligned}
&d^{16n}(\mathcal{U}_{\mathbf{R}}\otimes\mathcal{U}_{\mathbf{R}})\left(\bigotimes_{i\in[4]}\Pi_{\mathbf{x}_i}\right)(\mathcal{U}_{\mathbf{R}}^{\dagger}\otimes\mathcal{U}_{\mathbf{R}}^{\dagger})\\
&= \sum_{\mathbf{y}_1,\dots,\mathbf{y}_4\in\mathbb{Z}_d^{2n}}\mathcal{U}_{\mathbf{R}}\left(\bigotimes_{i\in[4]}\sum_{\mathbf{z}_i'\in\mathbb{Z}_d^{2n}}\omega^{[\mathbf{y}_i,\mathbf{z}_i]}\mathcal{W}_{\mathbf{z}_i'}\otimes\mathcal{W}_{J(\mathbf{z}_i')}\right)\mathcal{U}_{\mathbf{R}}^{\dagger}\otimes\mathcal{U}_{\mathbf{R}}\left(\bigotimes_{i\in[4]}\sum_{\mathbf{z}_i\in\mathbb{Z}_d^{2n}}\omega^{[\mathbf{x}_i+\mathbf{y}_i,\mathbf{z}_i]}\mathcal{W}_{\mathbf{z}_i'}\otimes\mathcal{W}_{J(\mathbf{z}_i)}\right)\mathcal{U}_{\mathbf{R}}^{\dagger}\\
&\qquad\qquad\text{(by Lemma 36)}\\
&= \sum_{\mathbf{y}_1,\dots,\mathbf{y}_4\in\mathbb{Z}_d^{2n}}\sum_{\mathbf{Z},\mathbf{Z}'\in\mathbb{Z}_d^{2n\times 4}}\omega^{\sum_{i=1}^{4}[\mathbf{y}_i,\mathbf{Z}_i'+\mathbf{Z}_i]}\left(\bigotimes_{i\in[4]}\mathcal{W}_{\mathbf{Z}_i'}\otimes\mathcal{W}_{(\mathbf{Z}'\mathbf{R})_i}\right)\otimes\left(\bigotimes_{i\in[4]}\omega^{[\mathbf{x}_i,\mathbf{Z}_i]}\mathcal{W}_{\mathbf{Z}_i}\otimes\mathcal{W}_{(\mathbf{Z}\mathbf{R})_i}\right)\\
&\qquad\qquad\text{(by Theorem 46)}\\
&= d^{8n}\sum_{\mathbf{Z},\mathbf{Z}'\in\mathbb{Z}_d^{2n\times 4}}\mathbf{1}[\mathbf{Z}=-\mathbf{Z}']\left(\bigotimes_{i\in[4]}\mathcal{W}_{\mathbf{Z}_i'}\otimes\mathcal{W}_{(\mathbf{Z}'\mathbf{R})_i}\right)\otimes\left(\bigotimes_{i\in[4]}\omega^{[\mathbf{x}_i,\mathbf{Z}_i]}\mathcal{W}_{\mathbf{Z}_i}\otimes\mathcal{W}_{(\mathbf{Z}\mathbf{R})_i}\right)\\
&\qquad\qquad\text{(by Lemma 9 over } \mathbf{y}_i)\\
&= d^{8n}\sum_{\mathbf{Z}\in\mathbb{Z}_d^{2n\times 4}}\left(\bigotimes_{i\in[4]}\mathcal{W}_{\mathbf{Z}_i}^{\dagger}\otimes\mathcal{W}_{(\mathbf{Z}\mathbf{R})_i}^{\dagger}\right)\otimes\left(\bigotimes_{i\in[4]}\omega^{[\mathbf{x}_i,\mathbf{Z}_i]}\mathcal{W}_{\mathbf{Z}_i}\otimes\mathcal{W}_{(\mathbf{Z}\mathbf{R})_i}\right). \qquad \text{(due to } \mathcal{W}_{-\mathbf{z}}=\mathcal{W}_{\mathbf{z}}^{\dagger})
\end{aligned}$$

Therefore,

$$b_\psi(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4) = \frac{1}{d^{8n}} \sum_{\mathbf{Z}\in\mathbb{Z}_d^{2n\times 4}} \mathrm{Tr}\left[\left(\bigotimes_{i\in[4]} \mathcal{W}_{\mathbf{Z}_i}^\dagger \psi \otimes \mathcal{W}_{(\mathbf{ZR})_i}^\dagger \psi\right) \otimes \left(\bigotimes_{i\in[4]} \omega^{[\mathbf{x}_i,\mathbf{Z}_i]} \mathcal{W}_{\mathbf{Z}_i} \psi \otimes \mathcal{W}_{(\mathbf{ZR})_i} \psi\right)\right]$$

$$= \sum_{\mathbf{Z}\in\mathbb{Z}_d^{2n\times 4}} \prod_{i=1}^4 \omega^{[\mathbf{x}_i,\mathbf{Z}_i]} p_\psi(\mathbf{Z}_i) p_\psi((\mathbf{ZR})_i). \tag{4}$$

The above expression can be further simplified to

$$b_\psi(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4) = \sum_{\mathbf{Z},\mathbf{Y},\mathbf{Y}'\in\mathbb{Z}_d^{2n\times 4}} \prod_{i=1}^4 \omega^{[\mathbf{x}_i,\mathbf{Z}_i]} \omega^{[\mathbf{Z}_i,\mathbf{Y}_i]+[(\mathbf{ZR})_i,\mathbf{Y}'_i]} \widehat{p_\psi}(\mathbf{Y}_i)\widehat{p_\psi}(\mathbf{Y}'_i)$$

(Fourier expansion)

$$= \frac{1}{d^{8n}} \sum_{\mathbf{Z},\mathbf{Y},\mathbf{Y}'\in\mathbb{Z}_d^{2n\times 4}} \prod_{i=1}^4 \omega^{[\mathbf{x}_i,\mathbf{Z}_i]} \omega^{[\mathbf{Z}_i,\mathbf{Y}_i]+[(\mathbf{ZR})_i,(\mathbf{Y}'\mathbf{R})_i]} p_\psi(\mathbf{Y}_i) p_\psi((\mathbf{Y}'\mathbf{R})_i)$$

(by Lemma 15 and $\mathbf{Y}' \leftarrow \mathbf{Y}'\mathbf{R}$)

$$= \frac{1}{d^{8n}} \sum_{\mathbf{Z},\mathbf{Y},\mathbf{Y}'\in\mathbb{Z}_d^{2n\times 4}} \prod_{i=1}^4 \omega^{[\mathbf{x}_i,\mathbf{Z}_i]} \omega^{[\mathbf{Z}_i,\mathbf{Y}_i]-[\mathbf{Z}_i,\mathbf{Y}'_i]} p_\psi(\mathbf{Y}_i) p_\psi((\mathbf{Y}'\mathbf{R})_i)$$

(due to $\mathbf{R}^\top\mathbf{R} = (D-1)\mathbf{I}$)

$$= \sum_{\mathbf{Y}\in\mathbb{Z}_d^{2n\times 4}} \prod_{i=1}^4 p_\psi(\mathbf{x}_i+\mathbf{Y}_i) p_\psi((\mathbf{YR})_i).$$

(by Lemma 9 over $\mathbf{Z}_i$)

If $|\psi\rangle = |\mathcal{S}\rangle = |\mathscr{M}, s\rangle$ is a stabiliser state, then we know that $p_\mathcal{S}(\mathbf{y}) = d^{-n}$ if $\mathbf{y} \in \mathscr{M}$ and $p_\mathcal{S}(\mathbf{y}) = 0$ if $\mathbf{y} \notin \mathscr{M}$. This means that $\prod_{i=1}^4 p_\mathcal{S}(\mathbf{x}_i+\mathbf{Y}_i) p_\mathcal{S}((\mathbf{YR})_i) \neq 0$ if and only if $\mathbf{x}_i+\mathbf{Y}_i \in \mathscr{M}$ and $(\mathbf{YR})_i \in \mathscr{M}$ for $i \in [4]$. Since $\mathscr{M}$ is a submodule, this implies that $\mathbf{x}_i, \mathbf{Y}_i \in \mathscr{M}$ for $i \in [4]$ if and only if $\prod_{i=1}^4 p_\mathcal{S}(\mathbf{x}_i+\mathbf{Y}_i) p_\mathcal{S}((\mathbf{YR})_i) \neq 0$. Therefore,

$$b_\mathcal{S}(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4) = \begin{cases} d^{-4n} & \text{if } \mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4 \in \mathscr{M}, \\ 0 & \text{otherwise.} \end{cases}$$

■

Theorem 50 is the much sought-after generalisation of Bell sampling: a properly skewed Bell difference sampling can successfully sample generators from a stabiliser group if sufficient copies of its stabiliser state are given. We compare the above result with the qubit case ($d = 2$) $q_\psi(\mathbf{x}) = 2^{2n}(p_\psi * p_\psi)(\mathbf{x}) = \sum_{\mathbf{y}\in\mathbb{F}_2^{2n}} p_\psi(\mathbf{x}+\mathbf{y}) p_\psi(\mathbf{y})$ from Gross, Nezami, and Walter [GNW21]. Due to the action of the matrix $\mathbf{R}$ in order to map $|\mathcal{S}\rangle^{\otimes 4} \mapsto |\mathcal{S}^*\rangle^{\otimes 4}$, all characteristic distributions inside b_ψ become non-trivially correlated. From now on, whenever we mention the matrix $\mathbf{R}$, we refer to the 4×4 integer matrix from Lemma 42. Finally, we shall succinctly write $b_\psi(\mathbf{X})$ for $\mathbf{X} = [\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4] \in \mathbb{Z}_d^{2n\times 4}$.

3.1 Further properties of distributions p_ψ and b_ψ

The distributions p_ψ and b_ψ are quite rich and we now explore them in more details. From Equation (4) and Lemma 9, we obtain the following simple corollaries. The first regards the mass on a submodule $\mathscr{X} \subseteq \mathbb{Z}_d^{2n}$ under b_ψ.

Corollary 51. *Let $\mathscr{X}_1, \mathscr{X}_2, \mathscr{X}_3, \mathscr{X}_4 \subseteq \mathbb{Z}_d^{2n}$ be submodules and $\mathbf{X} = [\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4] \in \bigotimes_{i=1}^4 \mathscr{X}_i$ be the matrix with columns $\mathbf{x}_i \in \mathscr{X}_i$, $i \in [4]$. Then*

$$\frac{1}{\prod_{i=1}^4 |\mathscr{X}_i|} \sum_{\mathbf{X}\in\bigotimes_{i=1}^4 \mathscr{X}_i} b_\psi(\mathbf{X}) = \sum_{\mathbf{X}\in\bigotimes_{i=1}^4 \mathscr{X}_i^\perp} \prod_{i=1}^4 p_\psi(\mathbf{X}_i) p_\psi((\mathbf{XR})_i).$$

The second gives an expression for the marginal distribution of b_ψ.

Corollary 52. *For all* $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$,

$$\sum_{\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4 \in \mathbb{Z}_d^{2n}} b_\psi(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4) = d^{3n} \sum_{\mathbf{y}\in\mathbb{Z}_d^{2n}} \omega^{[\mathbf{x}_1,\mathbf{y}]} p_\psi(\mathbf{y}) \prod_{i=1}^{4} p_\psi(a_i\mathbf{y}).$$

Corollary 51 has several important applications. The first one tells us that the mass of b_ψ on a submodule can be upper bounded by the p_ψ-mass.

Lemma 53. *For any submodule* $\mathscr{X} \subseteq \mathbb{Z}_d^{2n}$, $\sum_{\mathbf{X}\in\mathscr{X}^{\otimes 4}} b_\psi(\mathbf{X}) \leqslant \left(\sum_{\mathbf{x}\in\mathscr{X}} p_\psi(\mathbf{x})\right)^4$.

Proof.

$$\begin{aligned}
\sum_{\mathbf{X}\in\mathscr{X}^{\otimes 4}} b_\psi(\mathbf{X}) &= |\mathscr{X}|^4 \sum_{\mathbf{X}\in(\mathscr{X}^\perp)^{\otimes 4}} \prod_{i=1}^{4} p_\psi(\mathbf{X}_i) p_\psi((\mathbf{XR})_i) && \text{(by Corollary 51)} \\
&\leqslant \frac{|\mathscr{X}|^4}{d^{4n}} \sum_{\mathbf{X}\in(\mathscr{X}^\perp)^{\otimes 4}} \prod_{i=1}^{4} p_\psi(\mathbf{X}_i) && (p_\psi((\mathbf{XR})_i) \leqslant d^{-n}) \\
&= \left(\sum_{\mathbf{x}\in\mathscr{X}} p_\psi(\mathbf{x})\right)^4. && \text{(by Lemma 16)} \quad \blacksquare
\end{aligned}$$

We further obtain that p_ψ and b_ψ are supported on $\mathrm{Weyl}(|\psi\rangle)^\perp$, which is a generalisation of [GIKL24b, Lemma 4.3 & Corollary 4.4].

Lemma 54. *Let* $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ *and its characteristic and skewed Bell distributions* p_ψ *and* b_ψ. *The support of* p_ψ *and* b_ψ *are contained in* $\mathrm{Weyl}(|\psi\rangle)^\perp$ *and* $(\mathrm{Weyl}(|\psi\rangle)^\perp)^{\otimes 4}$, *respectively.*

Proof. We show that the mass of p_ψ on $\mathrm{Weyl}(|\psi\rangle)^\perp$ equals 1.

$$\begin{aligned}
\sum_{\mathbf{x}\in\mathrm{Weyl}(|\psi\rangle)^\perp} p_\psi(\mathbf{x}) &= \frac{|\mathrm{Weyl}(|\psi\rangle)^\perp|}{d^n} \sum_{\mathbf{y}\in\mathrm{Weyl}(|\psi\rangle)} p_\psi(\mathbf{y}) && \text{(by Lemma 16)} \\
&= \frac{|\mathrm{Weyl}(|\psi\rangle)^\perp|}{d^n} \frac{|\mathrm{Weyl}(|\psi\rangle)|}{d^n} = 1. && (p_\psi(\mathbf{x}) = d^{-n} \text{ iff } \mathbf{x} \in \mathrm{Weyl}(|\psi\rangle))
\end{aligned}$$

Now, since $b_\psi(\mathbf{X}) = \sum_{\mathbf{Y}\in\mathbb{Z}_d^{2n\times 4}} \prod_{i=1}^4 p_\psi(\mathbf{X}_i+\mathbf{Y}_i) p_\psi((\mathbf{YR})_i)$ and p_ψ is supported on $\mathrm{Weyl}(|\psi\rangle)^\perp$, which is a submodule, we must have that $b_\psi(\mathbf{X}) \neq 0 \iff \mathbf{X}_i+\mathbf{Y}_i, (\mathbf{YR})_i \in \mathrm{Weyl}(|\psi\rangle)^\perp$, and thus $\mathbf{X} \in (\mathrm{Weyl}(|\psi\rangle)^\perp)^{\otimes 4}$. $\blacksquare$

We have previously seen that the characteristic function of a stabiliser state $|\mathcal{S}\rangle$ is uniformly distributed on its associated Lagrangian submodule $\mathscr{M}$, i.e., $p_{\mathcal{S}}(\mathbf{x}) = d^{-n}$ if $\mathbf{x} \in \mathscr{M}$ and $p_{\mathcal{S}}(\mathbf{x}) = 0$ otherwise (Equation (2)). Such fact can be recovered from the above Lemma 54, since $\mathrm{Weyl}(|\mathcal{S}\rangle)^\perp = \mathscr{M}^\perp = \mathscr{M}$ and $p_{\mathcal{S}}(\mathbf{x}) = d^{-n}$ if and only if $\mathbf{x} \in \mathrm{Weyl}(|\mathcal{S}\rangle) = \mathscr{M}$.

We now prove that a submodule whose symplectic complement has high b_ψ-mass must be isotropic. For such, we shall need the following fact.

Fact 55 ([GNW21, Lemma 3.10])**.** *Let* $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ *and* $\mathbf{x},\mathbf{y} \in \mathbb{Z}_d^{2n}$ *such that* $p_\psi(\mathbf{x}), p_\psi(\mathbf{y}) \geqslant d^{-n}\left(1-\frac{1}{4d^2}\right)$. *Then* $[\mathbf{x},\mathbf{y}] = 0$.

Lemma 56. *Let* $\mathscr{X} \subseteq \mathbb{Z}_d^{2n}$ *be a submodule such that*

$$\sum_{\mathbf{X}\in(\mathscr{X}^\perp)^{\otimes 4}} b_\psi(\mathbf{X}) > \left(1 - \frac{1}{2d^2}\left(1-\frac{1}{p_1}\right)\left(1-\frac{1}{8d^2}\right)\right)^4,$$

where p_1 *is the smallest prime that divides* d. *Then* $\mathscr{X}$ *is isotropic.*

Proof. Consider the set $\mathscr{A} \triangleq \{\mathbf{x} \in \mathscr{X} : d^n p_\psi(\mathbf{x}) \geqslant 1 - \frac{1}{4d^2}\}$. According to Fact 55, $[\mathbf{x}, \mathbf{y}] = 0$ for all $\mathbf{x}, \mathbf{y} \in \mathscr{A}$. Therefore, the submodule $\langle \mathscr{A} \rangle$ generated by the span of all elements in $\mathscr{A}$ is isotropic. We now show that actually $\langle \mathscr{A} \rangle = \mathscr{X}$. For such, we assume by contradiction that $\langle \mathscr{A} \rangle \subset \mathscr{X}$ is a proper submodule of $\mathscr{X}$. We then have that $|\mathscr{A}| \leqslant \frac{|\mathscr{X}|}{p_1}$. Hence

$$
\begin{aligned}
\sum_{\mathbf{X} \in (\mathscr{X}^\perp)^{\otimes 4}} b_\psi(\mathbf{X}) &= |\mathscr{X}^\perp|^4 \sum_{\mathbf{X} \in \mathscr{X}^{\otimes 4}} \prod_{i=1}^4 p_\psi(\mathbf{X}_i) p_\psi((\mathbf{X}\mathbf{R})_i) && \text{(by Corollary 51)} \\
&\leqslant |\mathscr{X}^\perp|^4 \sum_{\mathbf{X} \in \mathscr{X}^{\otimes 4}} \prod_{i=1}^4 p_\psi(\mathbf{X}_i)^2 && \text{(by Cauchy-Schwarz)} \\
&= |\mathscr{X}^\perp|^4 \left(\sum_{\mathbf{x} \in \mathscr{A}} p_\psi(\mathbf{x})^2 + \sum_{\mathbf{x} \in \mathscr{X} \setminus \mathscr{A}} p_\psi(\mathbf{x})^2 \right)^4 \\
&\leqslant |\mathscr{X}^\perp|^4 \left(\frac{|\mathscr{A}|}{d^{2n}} + \frac{|\mathscr{X}| - |\mathscr{A}|}{d^{2n}} \left(1 - \frac{1}{4d^2}\right)^2 \right)^4 \\
&\leqslant \frac{|\mathscr{X}|^4 |\mathscr{X}^\perp|^4}{d^{8n}} \left(\frac{1}{p_1} + \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{4d^2}\right)^2 \right)^4 && (|\mathscr{A}| \leqslant \tfrac{|\mathscr{X}|}{p_1}) \\
&= \left(\frac{1}{p_1} + \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{4d^2}\right)^2 \right)^4, && (|\mathscr{X}||\mathscr{X}^\perp| = d^{2n})
\end{aligned}
$$

which contradicts the assumption of the lemma. Hence $\langle \mathscr{A} \rangle = \mathscr{X}$ and so $\mathscr{X}$ is isotropic. ∎

4 Stabiliser state learning

Using our unitary $\mathcal{B}_\mathbf{R}$ from Theorem 43 plus Bell difference sampling (Theorem 49), it is simple to learn an unknown stabiliser group $\mathcal{S}$ given enough copies of $|\mathcal{S}\rangle$, as shown in Algorithm 1.

Algorithm 1: Learning stabiliser states $|\mathcal{S}\rangle$

Input: $8m + 2n + 8$ copies of a stabiliser state $|\mathcal{S}\rangle \in (\mathbb{C}^d)^{\otimes n}$, $m = \lceil \frac{n}{2} + \frac{1}{4}\log_{p_1}(\ell/\delta)) \rceil$ and $d = \prod_{i=1}^\ell p_i^{k_i}$ is the prime decomposition of d with smallest prime factor p_1.

Output: a succinct description of the stabiliser group $\mathcal{S}$.

for $i \leftarrow -1$ **to** $m - 1$ **do**
  Perform Bell sampling on $(\mathbf{I} \otimes \mathcal{B}_\mathbf{R}^\dagger)|\mathcal{S}\rangle^{\otimes 8}$ to obtain $\mathbf{x}_{4i+1}, \mathbf{x}_{4i+2}, \mathbf{x}_{4i+3}, \mathbf{x}_{4i+4} \in \mathbb{Z}_d^{2n}$
  if $i \neq -1$ **then**
    $\mathbf{x}_{4i+j} \leftarrow \mathbf{x}_{4i+j} - \mathbf{x}_{-4+j}$ for $j = 1, 2, 3, 4$
Determine a generating set $\mathscr{B}$ for $\langle \{\mathbf{x}_i\}_{i \in [4m]} \rangle$ by putting the matrix $[\mathbf{x}_1, \ldots, \mathbf{x}_{4m}] \in \mathbb{Z}_d^{2n \times 4m}$ into its Smith normal form
For each $\mathbf{x} \in \mathscr{B}$, measure a copy of $|\mathcal{S}\rangle$ in the eigenbasis of $\mathcal{W}_\mathbf{x}$ to determine $s(\mathbf{x}) \in \mathbb{Z}_d$ such that $\mathcal{W}_\mathbf{x}|\mathcal{S}\rangle = \omega^{-s(\mathbf{x})}|\mathcal{S}\rangle$
return $\mathcal{S} = \langle \{\omega^{s(\mathbf{x})}\mathcal{W}_\mathbf{x}\}_{\mathbf{x} \in \mathscr{B}} \rangle$

Theorem 57. *Let $\delta \in (0, 1)$ and $\mathcal{S}$ an unknown stabiliser group with stabiliser state $|\mathcal{S}\rangle \in (\mathbb{C}^d)^{\otimes n}$. Let $d = \prod_{i=1}^\ell p_i^{k_i}$ be the prime factorisation of d with smallest prime factor p_1. Algorithm 1 identifies $\mathcal{S}$ with probability $1 - \delta$ using $6n + 2\log_{p_1} \frac{\ell}{\delta} + O(1)$ copies of $|\mathcal{S}\rangle$ in $O(n^3 + n^2 \log_{p_1} \frac{\ell}{\delta})$ time.*

Proof. Write $\mathcal{S} = \{\omega^{s(\mathbf{x})}\mathcal{W}_{\mathbf{x}} : \mathbf{x} \in \mathscr{M}\}$. According to Theorem 49, Bell sampling on $(\mathbf{I}\otimes\mathcal{B}_{\mathbf{R}}^\dagger)|\mathcal{S}\rangle^{\otimes 8}$, where $\mathcal{B}_{\mathbf{R}}$ is the unitary of Definition 40, returns $(\mathbf{x}_{4i+1}, \mathbf{x}_{4i+2}, \mathbf{x}_{4i+3}, \mathbf{x}_{4i+4}) \in (\mathbb{Z}_d^{2n})^{\otimes 4}$ with probability $\prod_{j=1}^4 p_{\mathcal{S}}(\mathbf{x}_{4i+j}+\mathbf{z}_j(\mathcal{S}))$ for some $\mathbf{z}_1(\mathcal{S}), \dots, \mathbf{z}_4(\mathcal{S}) \in \mathbb{Z}_d^{2n}$. Thus, perform Bell sampling $(\mathbf{I}\otimes\mathcal{B}_{\mathbf{R}}^\dagger)|\mathcal{S}\rangle^{\otimes 8}$ a number of $m+1$ times to obtain $(\mathbf{x}'_{4i+1}, \mathbf{x}'_{4i+2}, \mathbf{x}'_{4i+3}, \mathbf{x}'_{4i+4})_{i=-1}^{m-1}$. By removing $(\mathbf{x}'_{-3}, \mathbf{x}'_{-2}, \mathbf{x}'_{-1}, \mathbf{x}'_0)$ from all the other elements, we end up with $4m$ elements $\mathbf{x}_1, \dots, \mathbf{x}_{4m} \in \mathscr{M}$. We then compute a generating set $\mathscr{B}$ for $\langle\{\mathbf{x}_i\}_{i\in[4m]}\rangle$ (which has size at most $2n$) by putting the matrix $[\mathbf{x}_1, \dots, \mathbf{x}_{4m}] \in \mathbb{Z}_d^{2n\times 4m}$ into its Smith normal formal, which takes $O(mn^2)$ time [Sto96]. Finally, for each element $\mathbf{x} \in \mathscr{B}$, measure a copy of $|\mathcal{S}\rangle$ in the eigenbasis of $\mathcal{W}_{\mathbf{x}}$ to determine $s(\mathbf{x}) \in \mathbb{Z}_d$ such that $\mathcal{W}_{\mathbf{x}}|\mathcal{S}\rangle = \omega^{-s(\mathbf{x})}|\mathcal{S}\rangle$, which requires $O(n^2)$ time. The algorithm fails if the $4m$ samples $\mathbf{x}_1, \dots, \mathbf{x}_{4m}$ are contained in a maximal proper submodule of $\mathscr{M}$. Let then $\mathscr{X} \subset \mathscr{M}$ be a maximal proper submodule such that $\mathscr{M}/\mathscr{X} \cong \mathbb{Z}/p_i\mathbb{Z}$. Since $p_{\mathcal{S}}$ is uniform over $\mathscr{M}$, the probability that all $4m$ samples fall into $\mathscr{X}$ is $(|\mathscr{X}|/|\mathscr{M}|)^{4m} = 1/p_i^{4m}$. There are at most $\frac{p_i^{2n}-1}{p_i-1}$ submodules such that $\mathscr{M}/\mathscr{X} \cong \mathbb{Z}/p_i\mathbb{Z}$ (Lemma 58). Hence, by a union bound over all proper submodules $\mathscr{M}/\mathscr{X} \cong \mathbb{Z}/p_i\mathbb{Z}$ and over all $i \in [\ell]$, the probability that all $4m$ samples are contained in a proper submodule of $\mathscr{M}$ is, already taking $m = \lceil \frac{n}{2} + \frac{1}{4}\log_{p_1}(\ell/\delta)\rceil$,

$$\sum_{i=1}^{\ell} \frac{p_i^{2n}-1}{p_i-1} p_i^{-4m} \leqslant \sum_{i=1}^{\ell} p_i^{2n-4m} \leqslant \sum_{i=1}^{\ell} p_i^{-\log_{p_1}(\ell/\delta)} \leqslant \sum_{i=1}^{\ell} p_1^{-\log_{p_1}(\ell/\delta)} = \delta. \qquad \blacksquare$$

Lemma 58. *Let $\mathscr{M} \subseteq \mathbb{Z}_d^{2n}$ be a module of size d^n. There are at most $\frac{p^{2n}-1}{p-1}$ maximal proper submodules $\mathscr{X} \subset \mathscr{M}$ such that $\mathscr{M}/\mathscr{X} \cong \mathbb{Z}/p\mathbb{Z}$ for any prime $p|d$.*

Proof. Let $d = \prod_{i=1}^{\ell} p_i^{k_i}$ be the prime decomposition of d. By Fact 2, $\mathscr{M} \cong \bigoplus_{i=1}^{\ell} \mathscr{M}_i$ where $\mathscr{M}_i$ has size $|\mathscr{M}_i| = p_i^{nk_i}$ and decomposes as $\mathscr{M}_i \cong \bigoplus_{j=1}^{r_i} \mathbb{Z}/p_i^{\lambda_{ij}}\mathbb{Z}$ with $\lambda_{ij} \geqslant 1$ and $\sum_{j=1}^{r_i} \lambda_{ij} = nk_i$. Note that $r_i \leqslant 2n$ for all $i \in [\ell]$ since $\mathscr{M}$ has at most $2n$ generators. We can similarly write $\mathscr{X} \cong \bigoplus_{i=1}^{\ell} \mathscr{X}_i$ and by the condition $\mathscr{M}/\mathscr{X} \cong \mathbb{Z}/p_a\mathbb{Z}$ for $p_a|d$, then $\mathscr{X}_i = \mathscr{M}_i$ for $i \neq a$ and $\mathscr{M}_a/\mathscr{X}_a \cong \mathbb{Z}/p_a\mathbb{Z}$. We therefore just need to count maximal proper $\mathbb{Z}/p_a^{k_a}\mathbb{Z}$-submodules of $\mathscr{M}_a$.

We claim that there is a one-to-one correspondence between maximal proper submodules $\mathscr{X}_a \subset \mathscr{M}_a$ and kernels of surjective homomorphisms $\phi : \mathscr{M}_a \to \mathbb{Z}/p_a\mathbb{Z}$. Indeed, by isomorphism theorems, $\mathscr{M}_a/\ker(\phi) \cong \mathbb{Z}_p$, meaning that $\ker(\phi)$ must be a maximal proper submodule, while given a maximal proper submodule $\mathscr{X}_a$, the natural projection $\pi : \mathscr{M}_a \to \mathscr{M}_a/\mathscr{X}_a$ defined by $\pi(x) = x + \mathscr{X}_a$ for all $x \in \mathscr{M}_a$ is a surjective homomorphism with $\ker(\pi) = \mathscr{X}_a$. We now count the number of surjective homomorphisms $\phi : \mathscr{X}_a \to \mathbb{Z}/p_a\mathbb{Z}$. A homomorphism is determined by the images of all r_a generators of $\mathscr{M}_a$, which can be mapped to any element in $\mathbb{Z}/p_a\mathbb{Z}$, totaling $p_a^{r_a} - 1$ possible surjective homomorphisms, where we excluded the zero homomorphism since it is not surjective. On the other hand, different homomorphisms can have the same kernel. More specifically, two surjective homomorphisms ϕ and φ have the same kernel if and only if $\phi = \lambda\varphi$ for $\lambda \in \mathbb{Z}_p\backslash\{0\}$. Since there are $p_a - 1$ possible values of λ, there are $\frac{p_a^{r_a}-1}{p_a-1}$ possible kernels of surjective homomorphisms $\phi : \mathscr{X}_a \to \mathbb{Z}/p_a\mathbb{Z}$, and thus of maximal proper submodules. $\blacksquare$

5 Hidden Stabiliser Group Problem

The problem of identifying an unknown stabiliser state $|\mathcal{S}\rangle \in (\mathbb{C}^d)^{\otimes n}$ from the previous section can be framed as an instance of the so-called *State Hidden Subgroup Problem* (StateHSP) introduced by Bouland, Giurgică-Tiron, and Wright [BGTW25]. Hinsche, Eisert, and Carrasco [HEC25] narrowed down StateHSP to the case when the hidden group is a partial stabiliser group and defined the *Hidden Stabiliser Group Problem* as follows.[3]

[3] [HEC25] defines the Hidden Stabiliser Group Problem by supposing that $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ has a non-trivial partial stabiliser group $\mathcal{X} \subseteq \mathscr{P}_d^n$ such that $\mathcal{P}|\psi\rangle = |\psi\rangle$ for all $\mathcal{P} \in \mathcal{X}$ and $|\langle\psi|\mathcal{P}|\psi\rangle| \leqslant 1-\varepsilon$ for all $\mathcal{P} \in \mathscr{P}_d^n\backslash\mathcal{X}$. We believe that this is not an appropriate definition since for any $\mathcal{P} \in \mathcal{X}$ and $s \in \mathbb{Z}_d\backslash\{0\}$, $\omega^s\mathcal{P} \notin \mathcal{X}$ and still $|\langle\psi|\omega^s\mathcal{P}|\psi\rangle| = 1$.

Definition 59 (Hidden Stabiliser Group Problem)**.** *Given $\varepsilon > 0$, let $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ such that $|\langle\psi|\mathcal{W}_{\mathbf{x}}|\psi\rangle| < 1-\varepsilon$ for all $\mathbf{x} \notin \mathrm{Weyl}(|\psi\rangle)$. The* hidden stabiliser group problem *asks to identify $\mathrm{Weyl}(|\psi\rangle)$ given access to copies of $|\psi\rangle$.*

Hinsche, Eisert, and Carrasco [HEC25] proposed a quantum algorithm to solve the Hidden Stabiliser Group Problem that uses $O\big(\frac{nd\log d}{1-(1-\varepsilon)^d}\big)$ copies of $|\psi\rangle$ and runs in polynomial time.[4] Their algorithm uses the POVM given by $\Pi_{\mathbf{x}}^{\mathrm{HEC}} = d^{-2n}\sum_{\mathbf{y}\in\mathbb{Z}_d^{2n}} \omega^{[\mathbf{y},\mathbf{x}]}\mathcal{W}_{\mathbf{y}}^{\otimes D}$ to sample elements from, and thus learn, $\mathrm{Weyl}(|\psi\rangle)^{\perp\!\!\!\perp}$. A similar idea can be employed for all dimensions $d \geqslant 2$ but now using our generalised Bell difference sampling subroutine: by sampling enough elements from the distribution b_ψ, their span will generate $\mathrm{Weyl}(|\psi\rangle)^{\perp\!\!\!\perp}$ with high probability, from which $\mathrm{Weyl}(|\psi\rangle)$ can be obtained. In order to analyse the probability of successfully generating the whole submodule $\mathrm{Weyl}(|\psi\rangle)^{\perp\!\!\!\perp}$, we shall need the following lemma, which generalises [GIKL25, Lemma 2.3] and [CGYZ25, Lemma 4.21].

Lemma 60. *Let $\delta \in (0,1)$, $\varepsilon > 0$, and $\mathcal{D}$ be a distribution over $\mathbb{Z}_d^n$. Let $d = \prod_{i=1}^{\ell} p_i^{k_i}$ be the prime decomposition of d. Let $\mathscr{X} \subseteq \mathbb{Z}_d^n$ be the submodule spanned by m i.i.d. samples from $\mathcal{D}$. If $m \geqslant \frac{2}{\varepsilon}\big(\ln\frac{1}{\delta} + n\sum_{i=1}^{\ell} k_i\big)$, then $\mathcal{D}(\mathscr{X}) \triangleq \sum_{\mathbf{x}\in\mathscr{X}}\mathcal{D}(\mathbf{x}) \geqslant 1-\varepsilon$ with probability at least $1-\delta$.*

Proof. For $0 \leqslant i \leqslant m$, define $\mathscr{X}_i \triangleq \langle \mathbf{x}_1, \ldots, \mathbf{x}_i\rangle$, with $\mathscr{X}_0 \triangleq \{0^n\}$. Define the indicator random variable Z_i as

$$Z_i = \begin{cases} 1 & \text{if } \mathbf{x}_i \in \mathbb{Z}_d^n \backslash \mathscr{X}_{i-1} \text{ or } \mathcal{D}(\mathscr{X}_{i-1}) \geqslant 1-\varepsilon, \\ 0 & \text{otherwise.} \end{cases}$$

The value $Z_i = 1$ indicates that either the new sample $\mathbf{x}_i$ has increased the span of $\mathscr{X}_i$ or that $\mathscr{X}_{i-1}$ already accounts for a $(1-\varepsilon)$-fraction of the mass of $\mathcal{D}$. The following facts are true:

1. For any $\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}$, $\Pr[Z_i = 1 | \mathbf{x}_1, \ldots, \mathbf{x}_{i-1}] \geqslant \varepsilon$. Indeed, if $\mathcal{D}(\mathscr{X}_{i-1}) > 1-\varepsilon$, then already $Z_i = 1$, while if $\mathcal{D}(\mathscr{X}_{i-1}) \leqslant 1-\varepsilon$, then the probability that $Z_i = 1$ is $1 - \mathcal{D}(\mathscr{X}_{i-1}) \geqslant \varepsilon$.

2. As a consequence of the above fact, $\mathbb{E}[Z_i] \geqslant \varepsilon$.

3. Whenever $\sum_{i=1}^{m} Z_i \geqslant n\sum_{i=1}^{\ell} k_i$, we have $\mathcal{D}(\mathscr{X}_m) \geqslant 1-\varepsilon$. To see that, assume by contradiction that $\mathcal{D}(\mathscr{X}_m) < 1-\varepsilon$, in which case $\mathcal{D}(\mathscr{X}_i) < 1-\varepsilon$ for all $i \leqslant m$, and so $Z_i = 1$ implies that $\mathbf{x}_i \in \mathbb{Z}_d^n \backslash \mathscr{X}_{i-1}$. There are at most $n\sum_{i=1}^{\ell} k_i$ such i, which is the length of the module $\mathbb{Z}_d^n$, therefore $\mathbf{x}_1, \ldots, \mathbf{x}_m$ must span the whole space $\mathbb{Z}_d^n$, which contradicts the assumption $\mathcal{D}(\mathscr{X}_m) < 1-\varepsilon$.

Building on the above facts, we now prove the statement via a Chernoff bound, the only caveat being the fact that the Z_i's not independent. In order to deal with that, consider the new random variables Z_i' (for $i \in [m]$) as m i.i.d. samples from the Bernoulli distribution $\Pr[Z_i' = 1] = \varepsilon$. Then $\Pr[Z_i' = 1 | Z_1', \ldots, Z_{i-1}'] = \varepsilon \leqslant \Pr[Z_i = 1 | Z_1, \ldots, Z_{i-1}]$ from the first fact, and thus

[4] We note that the factor $\log d$ can be improved by using our Lemma 60, which is sharper than [HEC25, Lemma 1].

$\Pr[\sum_{i=1}^m < n\sum_{i=1}^\ell k_i] \leqslant \Pr[\sum_{i=1}^m Z_i' < n\sum_{i=1}^\ell k_i]$. Let $\theta \triangleq 1 - \frac{n}{m\varepsilon}\sum_{i=1}^\ell k_i$. Then

$$\begin{aligned}
\Pr[\mathcal{D}(\mathscr{X}_m) \geqslant 1-\varepsilon] &\geqslant \Pr\left[\sum_{i=1}^m Z_i \geqslant n\sum_{i=1}^\ell k_i\right] && \text{(by third fact)}\\
&\geqslant 1 - \Pr\left[\sum_{i=1}^m Z_i' < n\sum_{i=1}^\ell k_i\right] && \text{(by first fact)}\\
&= 1 - \Pr\left[\sum_{i=1}^m Z_i' < (1-\theta)m\varepsilon\right]\\
&\geqslant 1 - \exp\left(-\frac{\theta^2 m\varepsilon}{2}\right) && \text{(Chernoff bound)}\\
&\geqslant 1 - \exp\left(-\frac{m\varepsilon}{2} + n\sum_{i=1}^\ell k_i\right) \geqslant 1-\delta. && \blacksquare
\end{aligned}$$

Algorithm 2: Hidden Stabiliser Group Problem

Input: $8m+8$ copies of $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$, $m \triangleq \min\left\{\left\lceil \frac{4\ln\frac{1}{\delta} + 32n\sum_{i=1}^\ell k_i}{\varepsilon}, \frac{\ln\frac{\ell}{\delta} + 2n\ln p_\ell}{2\varepsilon}\right\rceil\right.$ and $d = \prod_{i=1}^\ell p_i^{k_i}$ is the prime decomposition of d and p_ℓ its largest prime factor.
Promise: $|\langle\psi|\mathcal{W}_{\mathbf{x}}|\psi\rangle| < 1-\varepsilon$ for all $\mathbf{x} \notin \mathrm{Weyl}(|\psi\rangle)$.
Output: a succinct description of $\mathrm{Weyl}(|\psi\rangle)$.

for $i \leftarrow -1$ **to** $m-1$ **do**
  Perform Bell sampling on $(\mathbf{I}\otimes\mathcal{B}_{\mathbf{R}}^\dagger)|\mathcal{S}\rangle^{\otimes 8}$ to obtain $\mathbf{x}_{4i+1}, \mathbf{x}_{4i+2}, \mathbf{x}_{4i+3}, \mathbf{x}_{4i+4} \in \mathbb{Z}_d^{2n}$
  if $i \neq -1$ **then**
    $\mathbf{x}_{4i+j} \leftarrow \mathbf{x}_{4i+j} - \mathbf{x}_{-4+j}$ for $j = 1,2,3,4$
Determine $\mathscr{W} = \langle\{\mathbf{x}_i\}_{i\in[4m]}\rangle^{\perp\!\!\!\perp}$ by putting the matrix $[\mathbf{x}_1, \ldots, \mathbf{x}_{4m}] \in \mathbb{Z}_d^{2n\times 4m}$ into its Smith normal form
return $\mathscr{W}$

Theorem 61. *Let $\varepsilon > 0$ and $\delta \in (0,1)$. Let $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ be an unknown state such that $|\langle\psi|\mathcal{W}_{\mathbf{x}}|\psi\rangle| < 1-\varepsilon$ for all $\mathbf{x} \notin \mathrm{Weyl}(|\psi\rangle)$. Let $d = \prod_{i=1}^\ell p_i^{k_i}$ be the prime factorisation of d with largest prime factor p_ℓ. Algorithm 2 identifies $\mathrm{Weyl}(|\psi\rangle)$ with probability $1-\delta$ in $O\big(\frac{n^2}{\varepsilon}\min\big\{n\sum_{i=1}^\ell k_i + \log\frac{1}{\delta}, n\log p_\ell + \log\frac{\ell}{\delta}\big\}\big)$ time using $O\big(\frac{1}{\varepsilon}\min\big\{n\sum_{i=1}^\ell k_i + \log\frac{1}{\delta}, n\log p_\ell + \log\frac{\ell}{\delta}\big\}\big)$ copies of $|\psi\rangle$ and acting on at most 8 copies of $|\psi\rangle$ at a time.*

Proof. By performing skewed Bell difference sampling m times (which requires $8m+8$ copies of $|\psi\rangle$ since one can subtract the first 4-tuple sample from all the subsequent ones), we obtain $\mathbf{x}_1, \ldots, \mathbf{x}_{4m} \in \mathrm{Weyl}(|\psi\rangle)^{\perp\!\!\!\perp}$. The time complexity is simply $O(mn^2)$. Algorithm 2 fails if all $4m$ samples are contained in a maximal proper submodule of $\mathrm{Weyl}(|\psi\rangle)^{\perp\!\!\!\perp}$.

We provide two error analyses. For the first one, consider a submodule $\mathscr{X}^{\perp\!\!\!\perp} \subset \mathrm{Weyl}(|\psi\rangle)^{\perp\!\!\!\perp}$ ($\mathrm{Weyl}(|\psi\rangle) \subset \mathscr{X}$) such that $\mathrm{Weyl}(|\psi\rangle)^{\perp\!\!\!\perp}/\mathscr{X}^{\perp\!\!\!\perp} \cong \mathbb{Z}/p_i\mathbb{Z}$. The probability that all samples

$\mathbf{x}_1, \dots, \mathbf{x}_{4m}$ are in $\mathscr{X}^\perp$ is $\left(\sum_{\mathbf{X} \in (\mathscr{X}^\perp)^{\otimes 4}} b_\psi(\mathbf{X})\right)^m$. We can bound

$$\begin{aligned}
\sum_{\mathbf{X} \in (\mathscr{X}^\perp)^{\otimes 4}} b_\psi(\mathbf{X}) &= |\mathscr{X}^\perp|^4 \sum_{\mathbf{X} \in \mathscr{X}^{\otimes 4}} \prod_{i=1}^4 p_\psi(\mathbf{X}_i) p_\psi((\mathbf{X}\mathbf{R})_i) && \text{(by Corollary 51)} \\
&\leqslant |\mathscr{X}^\perp|^4 \sqrt{\sum_{\mathbf{X} \in \mathscr{X}^{\otimes 4}} \prod_{i=1}^4 p_\psi(\mathbf{X}_i)^2} \sqrt{\sum_{\mathbf{X} \in \mathscr{X}^{\otimes 4}} \prod_{i=1}^4 p_\psi((\mathbf{X}\mathbf{R})_i)^2} && \text{(Cauchy-Schwarz)} \\
&= |\mathscr{X}^\perp|^4 \left(\sum_{\mathbf{x} \in \mathscr{X}} p_\psi(\mathbf{x})^2\right)^4 && (\mathbf{X} \leftarrow \mathbf{X}\mathbf{R}^\top) \\
&= |\mathscr{X}^\perp|^4 \left(\sum_{\mathbf{x} \in \operatorname{Weyl}(|\psi\rangle)} p_\psi(\mathbf{x})^2 + \sum_{\mathbf{x} \in \mathscr{X} \setminus \operatorname{Weyl}(|\psi\rangle)} p_\psi(\mathbf{x})^2\right)^4 \\
&< |\mathscr{X}^\perp|^4 \left(d^{-2n} |\operatorname{Weyl}(|\psi\rangle)| + d^{-2n}(1-\varepsilon)^4 (|\mathscr{X}| - |\operatorname{Weyl}(|\psi\rangle)|)\right)^4 \\
&\leqslant \left(\frac{1}{p_i} + (1-\varepsilon)^4 \left(1 - \frac{1}{p_i}\right)\right)^4. && (|\operatorname{Weyl}(|\psi\rangle)| \leqslant \tfrac{|\mathscr{X}|}{p_i}) \\
&\leqslant (1 - \varepsilon/2)^4. && (5)
\end{aligned}$$

Hence, by a union bound over all $\leqslant \frac{p_i^{2n}-1}{p_i-1}$ submodules $\mathscr{X}^\perp$ such that $\operatorname{Weyl}(|\psi\rangle)^\perp / \mathscr{X}^\perp \cong \mathbb{Z}/p_i\mathbb{Z}$ (Lemma 58) and all $i \in [\ell]$, the failure probability is, already taking $m \geqslant \frac{2n \ln p_\ell + \ln(\ell/\delta)}{2\varepsilon}$, at most

$$\sum_{i=1}^{\ell} \frac{p_i^{2n}-1}{p_i - 1} (1-\varepsilon/2)^{4m} \leqslant \sum_{i=1}^{\ell} \mathrm{e}^{2n \ln p_i - 2\varepsilon m} \leqslant \sum_{i=1}^{\ell} \mathrm{e}^{-\ln(\ell/\delta)} = \delta.$$

For our second error analysis, we consider a slightly weaker algorithm wherein one samples $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(m)} \sim b_\psi$, computes the four separate spans $\mathscr{W}_j \triangleq \langle \{\mathbf{X}_j^{(i)}\}_{i \in [m]} \rangle$ for $j \in [4]$, and considers the largest one. In practice there is no need to separate the four samples coming from one Bell difference sampling, but since Lemma 60 requires i.i.d. samples and we cannot guarantee their independence, we opted to separate the samples. We know that all $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(m)} \sim b_\psi$ belong to $(\operatorname{Weyl}(|\psi\rangle)^\perp)^{\otimes 4}$. By collecting $m \geqslant \frac{2}{\varepsilon}(2 \ln \frac{1}{\delta} + 16n \sum_{i=1}^{\ell} k_i)$ samples $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(m)}$ (b_ψ is over $(\mathbb{Z}_d^{2n})^{\otimes 4}$), then according to Lemma 60 the b_ψ-mass of $\langle \mathbf{X}^{(1)}, \dots, \mathbf{X}^{(m)} \rangle$ is at least $1 - \varepsilon/2$. On the other hand, by a very similar calculation to Equation (5), the b_ψ-mass of $\bigotimes_{i=1}^4 \mathscr{X}_i^\perp$ for any proper submodules $\mathscr{X}_1^\perp, \dots, \mathscr{X}_4^\perp \subset \operatorname{Weyl}(|\psi\rangle)^\perp$ is at most $(1-\varepsilon/2)^4$. This means that the largest span out of $\mathscr{W}_1, \mathscr{W}_2, \mathscr{W}_3, \mathscr{W}_4$ must equal $\operatorname{Weyl}(|\psi\rangle)^\perp$ (notice that we do not need to consider proper submodules of $(\operatorname{Weyl}(|\psi\rangle)^\perp)^{\otimes 4}$ but only of $\operatorname{Weyl}(|\psi\rangle)^\perp$ since we pick the largest span out of $\mathscr{W}_1, \mathscr{W}_2, \mathscr{W}_3, \mathscr{W}_4$). ∎

The sample complexity of Algorithm 2 is basically $O\left(\frac{n}{\varepsilon} \min\{\log p_\ell, \sum_{i=1}^{\ell} k_i\}\right)$, or simply $O\left(\frac{n}{\varepsilon}\right)$ for prime dimensions. The quantum algorithm of [HEC25], on the other hand, has sample complexity $O\left(\frac{nd \log d}{1-(1-\varepsilon)^d}\right) = O(n \log d \max\{d, \varepsilon^{-1}\})$ (since it is $O\left(\frac{n \log d}{\varepsilon}\right)$ for $\varepsilon = O(d^{-1})$ and $O(nd \log d)$ for $\varepsilon = \Omega(d^{-1})$). Our quantum algorithm has thus a better sample (and time) complexity and acts on at most 8 copies of $|\psi\rangle$ at a time (4 copies for prime dimensions and 2 if further $d \equiv 1 \pmod 4$), while $\{\Pi_{\mathbf{x}}^{\mathrm{HEC}}\}_{\mathbf{x}}$ requires D copies of $|\psi\rangle$ at a time, which can be impractical for large d.

6 Property testing stabiliser size

In this section, we focus on testing whether a given state $|\psi\rangle$ has stabiliser size at least d^t or is ε-far from any state with stabiliser size at least d^t, promised that one of these cases

holds. As we shall see, this section serves as a warm-up to Section 7. The results presented here generalise [GIKL25] to qudits, who proposed a property testing algorithm for stabiliser dimension. Since the concept of dimension is not well defined for general d, the right concept to focus on is that of stabiliser size.

The idea of Algorithm 3 is quite simple: sample several values from the skewed Bell distribution b_ψ and compute the size of their span. If it is at most d^{2n-t}, then we output the case where $|\psi\rangle$ has stabiliser size at least d^t. If, on the other hand, the span has size greater than d^{2n-t}, then we output the case where $|\psi\rangle$ is ε-far from any state with stabiliser size at least d^t. Implicit in the algorithm is learning the submodule $\mathrm{Weyl}(|\psi\rangle)$, which was the task behind the Hidden Stabiliser Group Problem from Section 5. As a consequence, Algorithms 2 and 3 are quite similar. The correctness of Algorithm 3 is proven in the next theorem.

Algorithm 3: Property testing stabiliser size

Input: $8m+8$ copies of $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$, where $m \triangleq \left\lceil \frac{2}{3\varepsilon}\left(\ln\frac{1}{\delta} + 8n\sum_{i=1}^{\ell} k_i\right)\right\rceil$.
Promise: $|\psi\rangle$ has stabiliser size at least d^t or is ε-far in fidelity from all such states.
Output: 1 if $|\psi\rangle$ has stabiliser size at least d^t and 0 otherwise, with probability $\geqslant 1-\delta$.

for $i \leftarrow -1$ **to** $m-1$ **do**
  Perform Bell sampling on $(\mathbf{I}\otimes\mathcal{B}_{\mathbf{R}}^\dagger)|\psi\rangle^{\otimes 8}$ to obtain $\mathbf{x}_{4i+1}, \mathbf{x}_{4i+2}, \mathbf{x}_{4i+3}, \mathbf{x}_{4i+4} \in \mathbb{Z}_d^{2n}$
  if $i \neq -1$ **then**
    $\mathbf{x}_{4i+j} \leftarrow \mathbf{x}_{4i+j} - \mathbf{x}_{-4+j}$ for $j = 1,2,3,4$
Compute the size K of $\langle\{\mathbf{x}_i\}_{i\in[4m]}\rangle$ by putting the matrix $[\mathbf{x}_1,\ldots,\mathbf{x}_{4m}] \in \mathbb{Z}_d^{2n\times 4m}$ into its Smith normal form
return 1 if $K \leqslant d^{2n-t}$ and 0 otherwise

Theorem 62. *Let $d = \prod_{i=1}^{\ell} p_i^{k_i}$ be the prime decomposition of d and p_1 the smallest prime to divide d. Let $\delta \in (0,1)$ and $0 < \varepsilon \leqslant \frac{1}{6d^2}\left(1-\frac{1}{p_1}\right)\left(1-\frac{1}{8d^2}\right)$. Let $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ be a state promised to either have stabiliser size at least d^t or be ε-far in fidelity from all such states. Algorithm 3 distinguishes the two cases with probability at least $1-\delta$, uses $8\left\lceil \frac{2}{3\varepsilon}\left(\ln\frac{1}{\delta} + 8n\sum_{i=1}^{\ell} k_i\right)\right\rceil + 8$ copies of $|\psi\rangle$, and has runtime $O\left(\frac{n^2}{\varepsilon}\left(\log\frac{1}{\delta} + n\sum_{i=1}^{\ell} k_i\right)\right)$.*

Proof. Assume first that $|\psi\rangle$ has stabiliser size at least d^t. According to Lemma 54, the distribution b_ψ is supported on $(\mathrm{Weyl}(|\psi\rangle)^{\perp\!\!\!\perp})^{\otimes 4}$. Therefore, $\mathscr{X}^{\perp\!\!\!\perp} \triangleq \langle\{\mathbf{x}_i\}_{i\in[4m]}\rangle$ is supported on $\mathrm{Weyl}(|\psi\rangle)^{\perp\!\!\!\perp}$, which has size at most d^{2n-t}, and so Algorithm 3 accepts with probability 1.

Now assume that $|\psi\rangle$ is ε-far in fidelity from all states with stabiliser size at least d^t. Since $m \geqslant \frac{1}{3\varepsilon}\left(2\ln\frac{1}{\delta} + 16n\sum_{i=1}^{\ell} k_i\right)$ (recall that b_ψ is over $(\mathbb{Z}_d^{2n})^{\otimes 4}$), Lemma 53 and 60 imply that, with probability at least $1-\delta$,

$$\sum_{\mathbf{x}\in\mathscr{X}^{\perp\!\!\!\perp}} p_\psi(\mathbf{x}) \geqslant \left(\sum_{\mathbf{X}\in(\mathscr{X}^{\perp\!\!\!\perp})^{\otimes 4}} b_\psi(\mathbf{X})\right)^{\frac{1}{4}} \geqslant (1-3\varepsilon)^{\frac{1}{4}} \geqslant 1-\varepsilon,$$

where we used that $\varepsilon < \frac{1}{2}$ in the last step. Under this event, Lemma 16 leads to

$$\sum_{\mathbf{x}\in\mathscr{X}} p_\psi(\mathbf{x}) = \frac{|\mathscr{X}|}{d^n} \sum_{\mathbf{x}\in\mathscr{X}^{\perp\!\!\!\perp}} p_\psi(\mathbf{x}) \geqslant \frac{1-\varepsilon}{K}. \tag{6}$$

On the other hand, since

$$\sum_{\mathbf{X}\in(\mathscr{X}^{\perp\!\!\!\perp})^{\otimes 4}} b_\psi(\mathbf{X}) \geqslant 1-3\varepsilon \geqslant 1 - \frac{1}{2d^2}\left(1-\frac{1}{p_1}\right)\left(1-\frac{1}{8d^2}\right),$$

Lemma 56 implies that $\mathscr{X}$ is isotropic, which together with Equation (6) allows us to use Lemma 34 to argue that there is a state with stabiliser size at least d^{2n}/K and fidelity at least $1-\varepsilon$ with $|\psi\rangle$. By assumption, $|\psi\rangle$ is ε-far from states with stabiliser size at least d^t, therefore $d^{2n}/K < d^t \implies K > d^{2n-t}$ and thus Algorithm 3 rejects with probability at least $1-\delta$.

Regarding the complexities, Algorithm 3 uses $8m+8$ copies of $|\psi\rangle$ and requires $O(mn^2)$ time to compute the Smith normal form of $\mathscr{X}^{\perp}$. ■

7 Testing doped Clifford circuits and pseudorandomness bounds

In this section, we prove that the output of any Clifford circuit augmented with a few non-Clifford single-qudit gates can be efficiently distinguished from a Haar-random quantum state, and therefore Clifford circuits require several non-Clifford single-qudit gates in order to generate pseudorandom quantum states (Definition 71). These results generalise [GIKL23, GIKL24b] from qubits to qudits. Moreover, as we shall see, our final lower bound is exponentially better compared to the prior work of [ADIS24], who also focused on qudits. The ideas behind our proof are similar to the ones from [GIKL24b]: a Haar-random state has vanishing stabiliser size with very high probability, while the output of a Clifford circuit with just a few non-Clifford single-qudit gates has a high stabiliser size. Before presenting our algorithm, though, we prove a few auxiliary results regarding Haar-random states.

7.1 Anti-concentration of Haar-random states

In this section, we show that the stabiliser size of a Haar random state $|\psi\rangle$ is 1 with high probability, or, equivalently, b_ψ is supported on the entire space $\mathbb{Z}_d^{2n\times 4}$ with high probability. To do so, we will make use of Lévy's lemma [MS86, Led01].

Fact 63 (Lévy's lemma)**.** *Let $f : \mathbb{S}^d \to \mathbb{R}$ be a function defined on the d-dimensional hypersphere $\mathbb{S}^d$. Assume f is K-Lipschitz, meaning that $|f(\psi) - f(\phi)| \leqslant K\|\psi - \phi\|$. Then, for every $\epsilon > 0$,*

$$\mathop{\mathbb{P}}_{\psi\sim\mu_{\text{Haar}}}[|f(\psi) - \mathbb{E}[f]| \geqslant \epsilon] \leqslant 2\exp\left(-\frac{(d+1)\epsilon^2}{9\pi^3 K^2}\right).$$

We will apply Lévy's lemma to functions of the form $\langle\psi|\mathcal{W}_{\mathbf{x}}|\psi\rangle$, which are 2-Lipschitz as shown next. The next result is basically [GIKL23, Lemma 20].

Lemma 64. *For any Weyl operator $\mathcal{W}_{\mathbf{x}} \in \mathscr{P}_d^n$, the function $f_{\mathbf{x}} : \mathbb{S}^{d^n} \to \mathbb{R}$ defined as $f_{\mathbf{x}}(|\psi\rangle) = \langle\psi|\mathcal{W}_{\mathbf{x}}|\psi\rangle$ is 2-Lipschitz.*

Proof.
$$\begin{aligned}
|\langle\psi|\mathcal{W}_{\mathbf{x}}|\psi\rangle - \langle\phi|\mathcal{W}_{\mathbf{x}}|\phi\rangle| &= |\langle\psi|\mathcal{W}_{\mathbf{x}}|\psi\rangle - \langle\phi|\mathcal{W}_{\mathbf{x}}|\psi\rangle + \langle\phi|\mathcal{W}_{\mathbf{x}}|\psi\rangle - \langle\phi|\mathcal{W}_{\mathbf{x}}|\phi\rangle| \\
&\leqslant |(\langle\psi| - \langle\phi|)\mathcal{W}_{\mathbf{x}}|\psi\rangle| + |\langle\phi|\mathcal{W}_{\mathbf{x}}(|\psi\rangle - |\phi\rangle)| \\
&\leqslant \|\mathcal{W}_{\mathbf{x}}|\psi\rangle\|\||\psi\rangle - |\phi\rangle\| + \|\mathcal{W}_{\mathbf{x}}|\phi\rangle\|\||\psi\rangle - |\phi\rangle\| \\
&= 2\||\psi\rangle - |\phi\rangle\|.
\end{aligned}$$
■

Putting the above two results together, we show that all Weyl operators $\mathcal{W}_{\mathbf{x}}$ for $\mathbf{x} \neq \mathbf{0}$ have very low expectation under Haar random states.

Lemma 65. *For any $\epsilon > 0$,*

$$\mathop{\mathbb{P}}_{|\psi\rangle\sim\mu_{\text{Haar}}}\left[\exists \mathbf{x} \in \mathbb{Z}_d^{2n}\backslash\{\mathbf{0}\} : |\langle\psi|\mathcal{W}_{\mathbf{x}}|\psi\rangle| \geqslant \epsilon\right] \leqslant 2d^{2n}\exp\left(-\frac{d^n\epsilon^2}{36\pi^3}\right).$$

Proof. Consider the function $f_{\mathbf{x}}(|\psi\rangle) = \langle\psi|\mathcal{W}_{\mathbf{x}}|\psi\rangle$ for $\mathbf{x} \in \mathbb{Z}_d^{2n}\backslash\{\mathbf{0}\}$, which is 2-Lipschitz according to Lemma 64. Moreover, $\mathbb{E}[f_{\mathbf{x}}] = 0$ over the Haar measure because d^{-1}-fraction of the eigenvalues of $\mathcal{W}_{\mathbf{x}}$ are ω^s, $s \in \mathbb{Z}_d$. The result thus follows from Lévy's lemma and a union bound over all d^{2n} possible Weyl operators. ■

As a consequence of the above result, b_ψ is supported on the entire space $\mathbb{Z}_d^{2n\times 4}$ with high probability if $|\psi\rangle$ is Haar random, or more precisely, that b_ψ has a bounded fraction of mass on maximal proper submodules of $\mathbb{Z}_d^{2n\times 4}$, which generalises [GIKL24b, Lemma 4.7].

Lemma 66. *Let $\epsilon > 0$ and let $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ be a Haar random n-qudit state. Let $d = \prod_{i=1}^{\ell} p_i^{k_i}$ be the prime decomposition of d. Then, for all maximal proper submodules $\mathscr{X} \subset \mathbb{Z}_d^{2n}$, $\sum_{\mathbf{X}\in\mathscr{X}^4} b_\psi(\mathbf{X}) \leqslant \left(\frac{1}{p_j} + \epsilon^4\right)^4$ with probability at least $1 - 2d^{2n}\exp\left(-\frac{d^n\epsilon^2}{36\pi^3}\right)$, where the prime $p_j|d$ is such that $\mathbb{Z}_d^{2n}/\mathscr{X} \cong \mathbb{Z}/p_j\mathbb{Z}$.*

Proof. Consider any maximal proper submodule $\mathscr{X} \subset \mathbb{Z}_d^{2n}$. Then $\mathscr{X}^\perp \cong \mathbb{Z}/p_j\mathbb{Z}$ must be simple for some $p_j|d$. By Corollary 51,

$$\sum_{\mathbf{X}\in\mathscr{X}^4} b_\psi(\mathbf{X}) = \frac{d^{8n}}{p_j^4} \sum_{\mathbf{X}\in(\mathscr{X}^\perp)^4} \prod_{i=1}^{4} p_\psi(\mathbf{X}_i) p_\psi((\mathbf{X}\mathbf{R})_i) = \frac{1}{p_j^4} \sum_{\mathbf{X}\in(\mathscr{X}^\perp)^4} \prod_{i=1}^{4} |\langle\psi|\mathcal{W}_{\mathbf{X}_i}|\psi\rangle\langle\psi|\mathcal{W}_{(\mathbf{X}\mathbf{R})_i}|\psi\rangle|^2.$$

Note that, in the above expression, there are $(p_j-1)^4\binom{4}{0}$ terms of the form $\prod_{i=1}^{8} |\langle\psi|\mathcal{W}_{\mathbf{x}_i}|\psi\rangle|^2$, $(p_j-1)^3\binom{4}{1}$ terms of the form $\prod_{i=1}^{6} |\langle\psi|\mathcal{W}_{\mathbf{x}_i}|\psi\rangle|^2$, $(p_j-1)^2\binom{4}{2}$ terms of the form $\prod_{i=1}^{4} |\langle\psi|\mathcal{W}_{\mathbf{x}_i}|\psi\rangle|^2$, and $(p_j-1)\binom{4}{3}$ terms of the form $\prod_{i=1}^{2} |\langle\psi|\mathcal{W}_{\mathbf{x}_i}|\psi\rangle|^2$, all with $\mathbf{x}_1, \ldots, \mathbf{x}_8 \neq \mathbf{0}$. Finally, there is 1 term $\langle\psi|\mathcal{W}_{\mathbf{0}}|\psi\rangle^8 = 1$. Conditioning on the event that $|\langle\psi|\mathcal{W}_{\mathbf{x}}|\psi\rangle| \leqslant \epsilon$ for all $\mathbf{x} \in \mathbb{Z}_d^{2n}\backslash\{\mathbf{0}\}$, which happens with probability at least $1 - 2d^{2n}\exp\left(-\frac{d^n\epsilon^2}{36\pi^3}\right)$ (Lemma 65), the mass of b_ψ over $\mathscr{X}^4$ is

$$\sum_{\mathbf{X}\in\mathscr{X}^4} b_\psi(\mathbf{X}) \leqslant \frac{1}{p_j^4} \sum_{i=0}^{4} \binom{4}{i} (p_j-1)^i \epsilon^{16i} = \frac{(1+\epsilon^4(p_j-1))^4}{p_j^4} \leqslant \left(\frac{1}{p_j} + \epsilon^4\right)^4.$$

The above is valid for any maximal proper submodule $\mathscr{X} \subset \mathbb{Z}_d^{2n}$, which concludes the proof. ∎

7.2 Stabiliser size of t-doped Clifford circuits

Following [LOH24, OLLH24, LOLH24], we first define t-doped Clifford circuits over qudits. Recall that a Clifford gate is any element belonging to the Clifford group $\mathscr{C}_d^n$.

Definition 67 (t-doped Clifford circuits)**.** *A t-doped Clifford circuit is a quantum circuit composed of Clifford gates and at most t non-Clifford single-qudit gates that starts in the state $|0\rangle^{\otimes n}$.*

The main technical lemma from this section is the fact that every non-Clifford gate in a Clifford circuit decreases the stabiliser length by at most 1.

Lemma 68. *Let $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ be the output state of a t-doped Clifford circuit. Then the stabiliser size of $|\psi\rangle$ is at least d^{n-2t}, i.e., $|\operatorname{Weyl}(|\psi\rangle)| \geqslant d^{n-2t}$.*

Proof. The proof is by induction on t. For $t = 0$, the output state is a stabiliser state, which has stabiliser size d^n. Assume then the induction hypothesis for $t-1$. Write $|\psi\rangle = \mathcal{C}\mathcal{U}|\phi\rangle$, where $|\phi\rangle$ is the output of a $(t-1)$-doped Clifford circuit, $\mathcal{U}$ is a single-qudit gate, and $\mathcal{C}$ is a Clifford circuit. Since the stabiliser size is unchanged by Clifford circuits (Corollary 29), we just need to show that the stabiliser size of $\mathcal{U}|\phi\rangle$ is at least d^{n-2t}. For such, consider $\operatorname{Weyl}(|\phi\rangle)$ with $|\operatorname{Weyl}(|\phi\rangle)| \geqslant d^{n-2(t-1)}$ by the induction assumption. Note that, for any $\mathbf{x} \in \operatorname{Weyl}(|\phi\rangle)$, if $\mathcal{W}_{\mathbf{x}}$ commutes with $\mathcal{U}$, then

$$\langle\phi|\mathcal{U}^\dagger\mathcal{W}_{\mathbf{x}}\mathcal{U}|\phi\rangle = \langle\phi|\mathcal{W}_{\mathbf{x}}|\phi\rangle \in \{1, \omega, \ldots, \omega^{d-1}\}.$$

Consider the set of commuting elements with $\mathcal{U}$, $\mathscr{E} \triangleq \{\mathbf{x} \in \operatorname{Weyl}(|\phi\rangle) : \mathcal{U}^\dagger\mathcal{W}_{\mathbf{x}}\mathcal{U} = \mathcal{W}_{\mathbf{x}}\}$. Then the stabiliser size of $\mathcal{U}|\phi\rangle$ is at least the size of $\mathscr{E}$, but $|\mathscr{E}| \geqslant |\operatorname{Weyl}(|\phi\rangle)|/d^2$, because $\mathscr{E}$ contains all elements $\mathbf{x} \in \operatorname{Weyl}(|\phi\rangle)$ for which $\mathcal{W}_{\mathbf{x}}$ restricts to the identity on the qudit to which $\mathcal{U}$ applies, i.e., $x_i = (v_i, w_i) = (0,0) \in \mathbb{Z}_d^2$ if $\mathcal{U}$ acts on the i-th qudit. Therefore, the stabiliser size of $\mathcal{U}|\phi\rangle$ is at least d^{n-2t}. ∎

In Lemma 54 we showed that the skewed Bell distribution b_ψ is supported on $(\mathrm{Weyl}(|\psi\rangle)^{\perp\!\!\!\perp})^4$. Together with the above lemma, we conclude that if $|\psi\rangle$ is the output of a t-doped Clifford circuit, then b_ψ is supported on a submodule of size at most d^{4n+8t}.

Corollary 69. *If $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ is the output of a t-doped Clifford circuit, then the support of the skewed Bell distribution b_ψ is a cartesian product of four submodules, each of size at most d^{n+2t}.*

Proof. According to Lemma 54, the support of b_ψ is contained in $(\mathrm{Weyl}(|\psi\rangle)^{\perp\!\!\!\perp})^4$, each of which with size $d^{2n}/|\mathrm{Weyl}(|\psi\rangle)| \leqslant d^{n+2t}$, where the last inequality follows from Lemma 68. ∎

7.3 Distinguishing t-doped Clifford circuits from Haar randomness

In hold of all the results from the previous sections, we now prove that, given enough copies of $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$, it is possible to distinguish whether (i) $|\psi\rangle$ is a Haar random state (ii) or $|\psi\rangle$ is the output of a $(n/2-1)$-doped Clifford circuit, under the promise that one case is the correct one. The idea of the algorithm is very simple: after performing skewed Bell difference sampling a number of $O(n)$ times, we check whether their span generates the whole space $\mathbb{Z}_d^{2n}$ or not. If $|\psi\rangle$ is Haar-random, then the sampled elements generate $\mathbb{Z}_d^{2n}$ with high probability, which can never happen if $|\psi\rangle$ is the output of a t-doped Clifford circuit for $t < n/2$, since b_ψ is supported on a submodule of size at most $d^{4(n+2t)}$.

Algorithm 4: Distinguishing $(n/2-1)$-doped Clifford circuits from Haar-random

Input: $8m+8$ copies of $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$, where $m \triangleq \left\lceil \frac{n}{2}\frac{\ln p_1}{\ln(2p_1/3)} + \frac{\ln(\ell/\delta)}{4\ln(2p_1/3)} \right\rceil$ and $d = \prod_{i=1}^{\ell} p_i^{k_i}$ is the prime decomposition of d with smallest prime factor.

Promise: $|\psi\rangle$ is Haar-random or the output of a $(n/2-1)$-doped Clifford circuit

Output: 0 if $|\psi\rangle$ is Haar-random and 1 otherwise, with probability at least $1-\delta$.

for $i \leftarrow -1$ **to** $m-1$ **do**
  Perform Bell sampling on $(\mathbf{I}\otimes\mathcal{B}_{\mathbf{R}}^\dagger)|\psi\rangle^{\otimes 8}$ to obtain $\mathbf{x}_{4i+1}, \mathbf{x}_{4i+2}, \mathbf{x}_{4i+3}, \mathbf{x}_{4i+4} \in \mathbb{Z}_d^{2n}$
  if $i \neq -1$ **then**
    $\mathbf{x}_{4i+j} \leftarrow \mathbf{x}_{4i+j} - \mathbf{x}_{-4+j}$ for $j = 1,2,3,4$
Compute the size K of $\langle\{\mathbf{x}_i\}_{i\in[4m]}\rangle$ by putting the matrix $[\mathbf{x}_1,\ldots,\mathbf{x}_{4m}] \in \mathbb{Z}_d^{2n\times 4m}$ into its Smith normal form
return 0 if $K = d^{2n}$ and 1 otherwise

Theorem 70. *Let $\delta \in (0,1)$ and $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ a state promised to be either Haar-random or the output of a t-doped Clifford circuit for $t < \frac{n}{2}$. Let $d = \prod_{i=1}^{\ell} p_i^{k_i}$ be the prime decomposition of d with smallest and largest prime factors p_1 and p_ℓ, respectively. Algorithm 4 distinguishes the two cases with probability at least $1 - \delta - 2d^{2n}\exp\left(-\frac{d^n}{36\pi^3\sqrt{2p_\ell}}\right)$ using $O(n + \log_{p_1}\frac{\ell}{\delta})$ copies of $|\psi\rangle$ and in $O(n^3 + n^2\log_{p_1}\frac{\ell}{\delta})$ time.*

Proof. If $|\psi\rangle$ is the output of a t-doped Clifford circuit with $t < n/2$, then $\langle\{\mathbf{x}_i\}_{i\in[4m]}\rangle$ always has size less than d^{2n} according to Corollary 45, where $\mathbf{x}_1,\ldots,\mathbf{x}_{4m} \in \mathrm{Weyl}(|\psi\rangle)^{\perp\!\!\!\perp}$ are skewed Bell difference samples. This means that Algorithm 4 can only fail when $|\psi\rangle$ is Haar random, which happens if $\mathbf{x}_1,\ldots,\mathbf{x}_{4m} \in \mathrm{Weyl}(|\psi\rangle)^{\perp\!\!\!\perp}$ all fall in the same maximal proper submodule $\mathscr{X} \subset \mathbb{Z}_d^{2n}$. Assume that $\mathbb{Z}_d^{2n}/\mathscr{X} \cong \mathbb{Z}/p_i\mathbb{Z}$ for some prime $p_i|d$. The probability that all samples $\mathbf{x}_1,\ldots,\mathbf{x}_{4m}$ are in $\mathscr{X}$ is $\left(\sum_{\mathbf{X}\in\mathscr{X}^4} b_\psi(\mathbf{X})\right)^m$. According to Lemma 66, with probability at least $1 - 2d^{2n}\exp\left(-\frac{d^n\epsilon^2}{36\pi^3}\right)$, we have that $\left(\sum_{\mathbf{X}\in\mathscr{X}^4} b_\psi(\mathbf{X})\right)^m \leqslant \left(\frac{1}{p_i}+\epsilon^4\right)^{4m}$. Take $\epsilon^4 = \frac{1}{2p_\ell}$ so that $\frac{1}{p_i}+\epsilon^4 \leqslant \frac{3}{2p_i}$. There are at most $\frac{p_i^{2n}-1}{p_i-1}$ submodules such that $\mathbb{Z}_d^{2n}/\mathscr{X} \cong \mathbb{Z}/p_i\mathbb{Z}$ according to

Lemma 58. A union bound over all such submodules and $i \in [\ell]$ leads to

$$\sum_{i=1}^{\ell} \frac{p_i^{2n}-1}{p_i-1}\left(\frac{1}{p_i}+\epsilon^4\right)^{4m} \leqslant \sum_{i=1}^{\ell} p_i^{2n}\left(\frac{3}{2p_i}\right)^{4m} = \sum_{i=1}^{\ell} \mathrm{e}^{2n\ln p_i - 4m\ln(2p_i/3)} \leqslant \sum_{i=1}^{\ell} \mathrm{e}^{-\ln(\ell/\delta)} = \delta$$

taking $m \geqslant \frac{n}{2}\frac{\ln p_1}{\ln(2p_1/3)} + \frac{\ln(\ell/\delta)}{4\ln(2p_1/3)}$. The total failure probability also needs to take into account the Haar measure failure probability $2d^{2n}\exp\left(-\frac{d^n}{36\pi^3\sqrt{2p_\ell}}\right)$. ■

The above results generalises [GIKL24b, Lemma 4.8 & Theorem 4.9] to qudits and also exponentially improves the range of t compared to [ADIS24], who provided a $\mathrm{poly}(n)$-sample-and-time-complexity algorithm for distinguishing Haar random states in $(\mathbb{C}^d)^{\otimes n}$ from t-doped Clifford circuits with $t = O(\log n/\log d)$.

As an immediate corollary of the above result, no t-doped Clifford circuit with $t < n/2$ can produce pseudorandom quantum states, otherwise they would be distinguishable from Haar random states in $\mathrm{poly}(n)$ time.

Definition 71 ([JLS18, Definition 2])**.** *Given a security parameter κ, a keyed family of n-qudit quantum states $\{|\phi_k\rangle \in (\mathbb{C}^d)^{\otimes n}\}_{k\in\{0,1\}^\kappa}$ is pseudorandom if (i) there is a polynomial-time quantum algorithm that generates $|\phi_k\rangle$ on input k, and (ii) for any $\mathrm{poly}(\kappa)$-time quantum adversary $\mathcal{A}$,*

$$\left|\mathop{\mathbb{P}}_{k\sim\{0,1\}^\kappa}[\mathcal{A}(|\phi_k\rangle^{\otimes \mathrm{poly}(\kappa)}) = 1] - \mathop{\mathbb{P}}_{|\psi\rangle\sim\mu_{\mathrm{Haar}}}[\mathcal{A}(|\psi\rangle^{\otimes \mathrm{poly}(\kappa)}) = 1]\right| = \mathrm{negl}(\kappa),$$

where μ_{Haar} is the n-qudit Haar measure and $\mathrm{negl}(\kappa)$ is an arbitrary negligible function of κ, i.e., $\mathrm{negl}(\kappa) = o(\kappa^{-c})$ $\forall c > 0$.

Corollary 72. *Any doped Clifford circuit that uses at most $n/2-1$ non-Clifford single-qudit gates cannot produce an ensemble of pseudorandom quantum states in $(\mathbb{C}^d)^{\otimes n}$.*

8 Stabiliser state testing

Gross, Nezami, and Walter [GNW21] developed a stabiliser testing algorithm for qudits (for all $d \geqslant 2$) which distinguishes between a stabiliser state and a state with stabiliser fidelity at most $1-\varepsilon$ using $O(1/\varepsilon)$ copies, which we review. For qubits, their algorithm is quite simple: perform Bell difference sampling to obtain $\mathbf{x} \in \mathbb{F}_2^{2n}$, measure the Weyl operator $\mathcal{W}_\mathbf{x}$ twice, and check whether the outcomes are equal or not. For large enough ε, their algorithm thus employs only 6 copies of $|\psi\rangle$, which is known to be almost optimal [Dam18]. Their algorithm for higher dimensions, however, does not have a simple interpretation. The authors introduced the operator

$$\mathcal{V}_r \triangleq \frac{1}{d^n}\sum_{\mathbf{x}\in\mathbb{Z}_d^{2n}} (\mathcal{W}_\mathbf{x}\otimes\mathcal{W}_\mathbf{x}^\dagger)^{\otimes r} \qquad \text{for } r \in \mathbb{N}$$

and showed that it possesses all the required properties for stabiliser testing. We are interested in exploring such operator in more details, so we review some of its properties.

Lemma 73 ([GNW21])**.** *Let $\mathcal{V}_r \triangleq d^{-n}\sum_{\mathbf{x}\in\mathbb{Z}_d^{2n}}(\mathcal{W}_\mathbf{x}\otimes\mathcal{W}_\mathbf{x}^\dagger)^{\otimes r}$. For $r \geqslant 2$ such that $\gcd(d,r) = 1$, $\mathcal{V}_r$ is a Hermitian unitary. As a consequence, the operator $\Pi_r^+ = \frac{1}{2}(\mathbf{I}+\mathcal{V}_r)$ is a projector and the binary POVM $\{\Pi_r^+, \mathbf{I}-\Pi_r^+\}$ can be implemented in $O(nr)$ time.*

Proof. It is easy to see that the operator $\mathcal{V}_r$ is Hermitian by a change of index. Moreover, $\mathcal{V}_r$ is unitary, since

$$\begin{aligned}\mathcal{V}_r^2 &= \frac{1}{d^{2n}}\sum_{\mathbf{x},\mathbf{y}\in\mathbb{Z}_d^{2n}}(\mathcal{W}_\mathbf{x}\mathcal{W}_\mathbf{y}\otimes\mathcal{W}_\mathbf{x}^\dagger\mathcal{W}_\mathbf{y}^\dagger)^{\otimes r} = \frac{1}{d^{2n}}\sum_{\mathbf{x},\mathbf{y}\in\mathbb{Z}_d^{2n}}\omega^{r[\mathbf{x},\mathbf{y}]}(\mathcal{W}_{\mathbf{x}+\mathbf{y}}\otimes\mathcal{W}_{\mathbf{x}+\mathbf{y}}^\dagger)^{\otimes r}\\ &= \frac{1}{d^{2n}}\sum_{\mathbf{x},\mathbf{z}\in\mathbb{Z}_d^{2n}}\omega^{r[\mathbf{x},\mathbf{z}]}(\mathcal{W}_\mathbf{z}\otimes\mathcal{W}_\mathbf{z}^\dagger)^{\otimes r} = \frac{1}{d^{2n}}\sum_{\mathbf{x},\mathbf{z}\in\mathbb{Z}_d^{2n}}\omega^{[\mathbf{x},\mathbf{z}]}(\mathcal{W}_\mathbf{z}\otimes\mathcal{W}_\mathbf{z}^\dagger)^{\otimes r} = \mathbf{I},\end{aligned}$$

where we used that $\gcd(d,r) = 1$ to remove r from $\omega^{r[\mathbf{x},\mathbf{z}]}$. Therefore, $\Pi_r^+ = \frac{1}{2}(\mathbf{I} + \mathcal{V}_r)$ is a projector, since $(\Pi_r^+)^2 = \frac{1}{4}(\mathbf{I} + 2\mathcal{V}_r + \mathcal{V}_r^2) = \Pi_r^+$.

It is possible to implement the binary POVM $\{\Pi_r^+, \mathbf{I} - \Pi_r^+\}$ by employing quantum phase estimation with $O(1)$ ancillae, $O(1)$ controlled versions of $\mathcal{V}_r$, and $O(1)$ auxiliary operations (we do not require a high accuracy in determining the eigenvalues of $\mathcal{V}_r$ since they are simply ± 1). And since we can write $\mathcal{V}_r = \left(d^{-1}\sum_{\mathbf{x}\in\mathbb{Z}_d^2}(\mathcal{W}_\mathbf{x}\otimes\mathcal{W}_\mathbf{x}^\dagger)^{\otimes r}\right)^{\otimes n}$, then a controlled version of $\mathcal{V}_r$ is equal to a composition of n controlled versions of $d^{-1}\sum_{\mathbf{x}\in\mathbb{Z}_d^2}(\mathcal{W}_\mathbf{x}\otimes\mathcal{W}_\mathbf{x}^\dagger)^{\otimes r}$, each of which acts on $2r$ qudits. Therefore, the POVM $\{\Pi_r^+, \mathbf{I} - \Pi_r^+\}$ can be implemented in time $O(nr)$. ■

Define the binary POVM $\{\Pi_r^+, \mathbf{I} - \Pi_r^+\}$ and let $G_r(|\psi\rangle)$ be the expected value of the random variable Z which equals 1 if we measure Π_r^+ and -1 if we measure $\mathbf{I} - \Pi_r^+$, i.e., $\mathbb{P}[Z = 1] = \operatorname{Tr}[\psi^{\otimes 2r}\Pi_r^+]$ and $\mathbb{P}[Z = -1] = 1 - \operatorname{Tr}[\psi^{\otimes 2r}\Pi_r^+]$. Then

$$G_r(|\psi\rangle) \triangleq 2\operatorname{Tr}[\psi^{\otimes 2r}\Pi_r^+] - 1 = \operatorname{Tr}[\psi^{\otimes 2r}\mathcal{V}_r] = \frac{1}{d^n}\sum_{\mathbf{x}\in\mathbb{Z}_d^{2n}}|\operatorname{Tr}[\psi\mathcal{W}_\mathbf{x}]|^{2r} = d^{(r-1)n}\sum_{\mathbf{x}\in\mathbb{Z}_d^{2n}}p_\psi(\mathbf{x})^r. \quad (7)$$

It is possible to relate $G_r(|\psi\rangle)$ with the stabiliser fidelity $F_\mathcal{S}(|\psi\rangle)$ using Fact 55.

Lemma 74. *Let $r \in \mathbb{N}$, $r \geqslant 2$, such that $\gcd(d,r) = 1$ and $C_{d,r} \triangleq \frac{1}{2}\left(1 - (1 - \frac{1}{4d^2})^{r-1}\right)$. Then*

$$F_\mathcal{S}(|\psi\rangle) \geqslant 1 + \frac{G_r(|\psi\rangle) - 1}{2C_{d,r}} \iff G_r(|\psi\rangle) \leqslant 1 - 2C_{d,r}(1 - F_\mathcal{S}(|\psi\rangle)) \qquad \forall |\psi\rangle \in (\mathbb{C}^d)^{\otimes n}.$$

Proof. Consider the set $\mathscr{X}_0 = \{\mathbf{x} \in \mathbb{Z}_d^{2n} : d^n p_\psi(\mathbf{x}) \geqslant 1 - \frac{1}{4d^2}\}$, which according to Fact 55 is isotropic. It can thus be extended a Lagrangian (maximal) submodule. Then

$$\begin{aligned}
F_\mathcal{S}(|\psi\rangle) &\geqslant \sum_{\mathbf{x}\in\mathscr{X}_0} p_\psi(\mathbf{x}) && \text{(by Corollary 33)}\\
&= 1 - \sum_{\mathbf{x}\in\mathbb{Z}_d^{2n}} p_\psi(\mathbf{x})\mathbf{1}\left[d^n p_\psi(\mathbf{x}) < 1 - \frac{1}{4d^2}\right]\\
&\geqslant 1 - \frac{\sum_{\mathbf{x}\in\mathbb{Z}_d^{2n}} p_\psi(\mathbf{x})(1 - d^{(r-1)n}p_\psi(\mathbf{x})^{r-1})}{1 - (1 - \frac{1}{4d^2})^{r-1}} && \text{($(r-1)$-th moment Markov's inequality)}\\
&= 1 + \frac{G_r(|\psi\rangle) - 1}{1 - (1 - \frac{1}{4d^2})^{r-1}}.
\end{aligned}$$

■

Using the above theorem, we review the stabiliser testing algorithm from [GNW21]: measure the POVM $\{\Pi_r^+, \mathbf{I} - \Pi_r^+\}$ several times and estimate the quantity $G_r(|\psi\rangle)$. If $G_r(|\psi\rangle) < 1$, then $|\psi\rangle$ is far from a stabiliser state.

Algorithm 5: Stabiliser testing algorithm from [GNW21]

Input: $2rm$ copies of $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ where $m \triangleq \lceil \frac{1}{C_{d,r}\varepsilon} \ln \frac{1}{\delta} \rceil$, $C_{d,r} \triangleq \frac{1}{2}\left(1 - (1 - \frac{1}{4d^2})^{r-1}\right)$, and $r \geqslant 2$ is an integer such that $\gcd(d,r) = 1$.

Promise: either $F_\mathcal{S}(|\psi\rangle) = 1$ or $F_\mathcal{S}(|\psi\rangle) \leqslant 1 - \varepsilon$.

Output: 1 if $F_\mathcal{S}(|\psi\rangle) = 1$ and 0 if $F_\mathcal{S}(|\psi\rangle) \leqslant 1 - \varepsilon$.

1 **for** $i = 1$ **to** m **do**

2 Measure the POVM $\{\Pi_r^+, \mathbf{I} - \Pi_r^+\}$ on $|\psi\rangle^{\otimes 2r}$ and obtain the outcome $Z_i \in \{-1, 1\}$

3 **return** 1 if $Z_i = 1$ for all $i \in [m]$ and 0 otherwise

Theorem 75 ([GNW21, Corollary 3.13]). *Let $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$, $\varepsilon > 0$, $\delta \in (0,1)$, and $r \geqslant 2$ an integer such that $\gcd(d,r) = 1$. Algorithm 5 always accepts if $F_{\mathcal{S}}(|\psi\rangle) = 1$ and rejects if $F_{\mathcal{S}}(|\psi\rangle) \leqslant 1-\varepsilon$ with probability $\geqslant 1-\delta$. It uses $2r\lceil \frac{1}{C_{d,r}\varepsilon} \ln \frac{1}{\delta} \rceil$ copies of $|\psi\rangle$ where $C_{d,r} \triangleq \frac{1}{2}\big(1-(1-\frac{1}{4d^2})^{r-1}\big)$.*

Proof. If $F_{\mathcal{S}}(|\psi\rangle) = 1$, then measuring the POVM $\{\Pi_r^+, \mathbf{I}-\Pi_r^+\}$ always returns 1 and Algorithm 5 always succeeds (it is perfectly complete). If $F_{\mathcal{S}}(|\psi\rangle) \leqslant 1-\varepsilon$, then according to Lemma 74, the expected outcome $G_r(|\psi\rangle)$ of measuring the POVM $\{\Pi_r^+, \mathbf{I}-\Pi_r^+\}$ is $G_r(|\psi\rangle) \leqslant 1-2C_{d,r}\varepsilon$. Then failure probability equals the probability that all $m = \lceil \frac{1}{C_{d,r}\varepsilon} \ln \frac{1}{\delta} \rceil$ outcomes equal 1, i.e.,

$$\left(\frac{1+G_r(|\psi\rangle)}{2}\right)^m \leqslant (1-C_{d,r}\varepsilon)^m \leqslant \mathrm{e}^{-mC_{d,r}\varepsilon} \leqslant \delta.$$

Finally, each measurement requires $2r$ copies of $|\psi\rangle$. ∎

Remark 76. *For $r = O(d^2)$, $C_{d,r} = \Theta(\frac{r}{d^2})$ and Algorithm 5 requires $O(\frac{d^2}{\varepsilon} \ln \frac{1}{\delta})$ copies of $|\psi\rangle$.*

We now develop an alternative for the above testing procedure: obtain $\mathbf{X} \in \mathbb{Z}_d^{2n\times 4}$ through skewed Bell difference sampling and measure $\frac{1}{2}\bigotimes_{i=1}^4(\mathcal{W}_{\mathbf{X}_i} \otimes \mathcal{W}_{\mathbf{X}_i}^\dagger) + \frac{1}{2}\bigotimes_{i=1}^4(\mathcal{W}_{\mathbf{X}_i}^\dagger \otimes \mathcal{W}_{\mathbf{X}_i})$ on $|\psi\rangle^{\otimes 8}$. We now study the average outcome of such procedure, i.e.,

$$\begin{aligned} A(|\psi\rangle) &\triangleq \frac{1}{2}\mathop{\mathbb{E}}_{\mathbf{X}\sim b_\psi}\left[\bigotimes_{i=1}^4 \langle\psi|^{\otimes 2}\mathcal{W}_{\mathbf{X}_i} \otimes \mathcal{W}_{\mathbf{X}_i}^\dagger |\psi\rangle^{\otimes 2} + \bigotimes_{i=1}^4 \langle\psi|^{\otimes 2}\mathcal{W}_{\mathbf{X}_i}^\dagger \otimes \mathcal{W}_{\mathbf{X}_i} |\psi\rangle^{\otimes 2}\right] \\ &= \mathop{\mathbb{E}}_{\mathbf{X}\sim b_\psi}\left[\prod_{i=1}^4 |\langle\psi|\mathcal{W}_{\mathbf{X}_i}|\psi\rangle|^2\right] = d^{4n}\mathop{\mathbb{E}}_{\mathbf{X}\sim b_\psi}\left[\prod_{i=1}^4 p_\psi(\mathbf{X}_i)\right]. \end{aligned}$$

Remark 77. *Note that $\frac{1}{2}\bigotimes_{i=1}^4(\mathcal{W}_{\mathbf{X}_i} \otimes \mathcal{W}_{\mathbf{X}_i}^\dagger) + \frac{1}{2}\bigotimes_{i=1}^4(\mathcal{W}_{\mathbf{X}_i}^\dagger \otimes \mathcal{W}_{\mathbf{X}_i})$ can be performed transversally and takes $O(n)$ time.*

Lemma 78. *For any $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$,*

$$A(|\psi\rangle) = d^{8n}\sum_{\mathbf{X}\in\mathbb{Z}_d^{2n\times 4}}\prod_{i=1}^4 p_\psi(\mathbf{X}_i)^2 p_\psi((\mathbf{X}\mathbf{R})_i).$$

Proof.

$$\begin{aligned} A(|\psi\rangle) &= d^{4n}\sum_{\mathbf{X},\mathbf{Y}\in\mathbb{Z}_d^{2n\times 4}}\prod_{i=1}^4 p_\psi(\mathbf{X}_i)p_\psi(\mathbf{X}_i+\mathbf{Y}_i)p_\psi((\mathbf{Y}\mathbf{R})_i) && \text{(by definition of } b_\psi) \\ &= \frac{1}{d^{4n}}\sum_{\mathbf{X},\mathbf{Y},\mathbf{Z},\mathbf{Z}'\in\mathbb{Z}_d^{2n\times 4}}\prod_{i=1}^4 p_\psi(\mathbf{X}_i)\omega^{[\mathbf{X}_i+\mathbf{Y}_i,\mathbf{Z}_i]}p_\psi(\mathbf{Z}_i)\omega^{[(\mathbf{Y}\mathbf{R})_i,(\mathbf{Z}'\mathbf{R})_i]}p_\psi((\mathbf{Z}'\mathbf{R})_i) \\ & && \text{(Fourier expansion and Lemma 15)} \\ &= \frac{1}{d^{4n}}\sum_{\mathbf{X},\mathbf{Y},\mathbf{Z},\mathbf{Z}'\in\mathbb{Z}_d^{2n\times 4}}\prod_{i=1}^4 p_\psi(\mathbf{X}_i)\omega^{[\mathbf{X}_i+\mathbf{Y}_i,\mathbf{Z}_i]}p_\psi(\mathbf{Z}_i)\omega^{-[\mathbf{Y}_i,\mathbf{Z}_i']}p_\psi((\mathbf{Z}'\mathbf{R})_i) \\ & && (\mathbf{R}\mathbf{R}^\top = (D-1)\mathbf{I}) \\ &= d^{4n}\sum_{\mathbf{X},\mathbf{Z}\in\mathbb{Z}_d^{2n\times 4}}\prod_{i=1}^4 p_\psi(\mathbf{X}_i)\omega^{[\mathbf{X}_i,\mathbf{Z}_i]}p_\psi(\mathbf{Z}_i)p_\psi((\mathbf{Z}\mathbf{R})_i) && \text{(by Lemma 9 over } \mathbf{Y}) \\ &= d^{8n}\sum_{\mathbf{Z}\in\mathbb{Z}_d^{2n\times 4}}\prod_{i=1}^4 p_\psi(\mathbf{Z}_i)^2 p_\psi((\mathbf{Z}\mathbf{R})_i). && \text{(Fourier expansion and Lemma 15)} \end{aligned}$$

∎

Similarly to Lemma 74, we can relate A_ψ with the stabiliser fidelity $F_{\mathcal{S}}(|\psi\rangle)$.

Lemma 79. *Let* $C_{d,3} \triangleq \frac{1}{2}\big(1-(1-\frac{1}{4d^2})^2\big)$. *For all* $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$,

$$A(|\psi\rangle) \leqslant G_3(|\psi\rangle)^4 \leqslant \big(1-2C_{d,3}(1-F_{\mathcal{S}}(|\psi\rangle))\big)^4.$$

Proof.

$$\begin{aligned}
A(|\psi\rangle) &= d^{8n} \sum_{\mathbf{X}\in\mathbb{Z}_d^{2n\times 4}} \prod_{i=1}^{4} p_\psi(\mathbf{X}_i)^2 p_\psi((\mathbf{XR})_i) && \text{(by Lemma 78)}\\
&\leqslant d^{8n} \sum_{\mathbf{X}\in\mathbb{Z}_d^{2n\times 4}} \prod_{i=1}^{4} p_\psi(\mathbf{X}_i)^3 \\
&= \left(d^{2n} \sum_{\mathbf{x}\in\mathbb{Z}_d^{2n}} p_\psi(\mathbf{x})^3\right)^4 \\
&\leqslant \big(1-2C_{d,3}(1-F_{\mathcal{S}}(|\psi\rangle))\big)^4. && \text{(by Equation (7) and Lemma 74)} \quad \blacksquare
\end{aligned}$$

By measuring $\frac{1}{2}\bigotimes_{i=1}^4(\mathcal{W}_{\mathbf{X}_i}\otimes\mathcal{W}_{\mathbf{X}_i}^\dagger)+\frac{1}{2}\bigotimes_{i=1}^4(\mathcal{W}_{\mathbf{X}_i}^\dagger\otimes\mathcal{W}_{\mathbf{X}_i})$ on $|\psi\rangle^{\otimes 8}$ with $\mathbf{X}\sim b_\psi$, we develop a different algorithm for testing stabiliser states.

Algorithm 6: Alternative stabiliser testing algorithm

Input: $16m+8$ copies of $|\psi\rangle\in(\mathbb{C}^d)^{\otimes n}$ where $m \triangleq \lceil\big(2+\frac{1}{4C_{d,3}\varepsilon}\big)\ln\frac{1}{\delta}\rceil$ and $C_{d,3} \triangleq \frac{1}{2}\big(1-(1-\frac{1}{4d^2})^2\big)$.

Promise: either $F_{\mathcal{S}}(|\psi\rangle)=1$ or $F_{\mathcal{S}}(|\psi\rangle)\leqslant 1-\varepsilon$.

Output: 1 if $F_{\mathcal{S}}(|\psi\rangle)=1$ and 0 if $F_{\mathcal{S}}(|\psi\rangle)\leqslant 1-\varepsilon$.

for $i\leftarrow 0$ **to** m **do**
  Perform Bell sampling on $(\mathbf{I}\otimes\mathcal{B}_{\mathbf{R}}^\dagger)|\psi\rangle^{\otimes 8}$ to obtain $\mathbf{X}^{(i)}\in\mathbb{Z}_d^{2n\times 4}$
  if $i\neq 0$ **then**
    $\mathbf{X}^{(i)}\leftarrow\mathbf{X}^{(i)}-\mathbf{X}^{(0)}$
    Measure $\frac{1}{2}\bigotimes_{j=1}^4(\mathcal{W}_{\mathbf{X}_j^{(i)}}\otimes\mathcal{W}_{\mathbf{X}_j^{(i)}}^\dagger)+\frac{1}{2}\bigotimes_{j=1}^4(\mathcal{W}_{\mathbf{X}_j^{(i)}}^\dagger\otimes\mathcal{W}_{\mathbf{X}_j^{(i)}})$ on $|\psi\rangle^{\otimes 8}$ to obtain $Z_i\in[-1,1]$
return 1 if $Z_i=1$ for all $i\in[m]$ and 0 otherwise

Theorem 80. *Let* $|\psi\rangle\in(\mathbb{C}^d)^{\otimes n}$, $\varepsilon>0$, *and* $\delta\in(0,1)$. Algorithm 6 *always accepts if* $F_{\mathcal{S}}(|\psi\rangle)=1$ *and rejects if* $F_{\mathcal{S}}(|\psi\rangle)\leqslant 1-\varepsilon$ *with probability at least* $1-\delta$. *It uses* $16\lceil\big(2+\frac{1}{4C_{d,3}\varepsilon}\big)\ln\frac{1}{\delta}\rceil+8$ *copies of* $|\psi\rangle$, *where* $C_{d,3}\triangleq\frac{1}{2}\big(1-(1-\frac{1}{4d^2})^2\big)$.

Proof. Let $A(|\psi\rangle)\in[-1,1]$ be the average outcome of measuring $\frac{1}{2}\bigotimes_{i=1}^4(\mathcal{W}_{\mathbf{X}_i}\otimes\mathcal{W}_{\mathbf{X}_i}^\dagger)+\frac{1}{2}\bigotimes_{i=1}^4(\mathcal{W}_{\mathbf{X}_i}^\dagger\otimes\mathcal{W}_{\mathbf{X}_i})$ on $|\psi\rangle^{\otimes 8}$ with $\mathbf{X}\sim b_\psi$. If $F_{\mathcal{S}}(|\psi\rangle)=1$, then $A(|\psi\rangle)=1$ and Algorithm 6 always succeeds (it is perfectly complete). If $F_{\mathcal{S}}(|\psi\rangle)\leqslant 1-\varepsilon$, then according to Lemma 79, $A(|\psi\rangle)\leqslant(1-2C_{d,3}\varepsilon)^4$. Note that we can, without loss of generality, shift the measurement outcomes by 1. That means that the new measurement outcomes satisfy $Z_i\in[0,2]$, while their expected value becomes $1+A(|\psi\rangle)\leqslant 1+(1-2C_{d,3}\varepsilon)^4$. By Markov's inequality,

$$\Pr[Z_i\geqslant 2]\leqslant\frac{1+A(|\psi\rangle)}{2}\leqslant\frac{1+(1-2C_{d,3}\varepsilon)^4}{2}.$$

The failure probability of Algorithm 6 equals the probability that all $m = \lceil \left(2 + \frac{1}{4C_{d,3}\varepsilon}\right) \ln \frac{1}{\delta} \rceil$ outcomes Z_i equal 2, i.e.,

$$\left(\frac{1 + (1 - 2C_{d,3}\varepsilon)^4}{2}\right)^m \leqslant \left(\frac{1 + \frac{1}{1+8C_{d,3}\varepsilon}}{2}\right)^m = \left(\frac{1 + 4C_{d,3}\varepsilon}{1 + 8C_{d,3}\varepsilon}\right)^m \leqslant \exp\left(-m\frac{4C_{d,3}\varepsilon}{1 + 8C_{d,3}\varepsilon}\right) \leqslant \delta.$$

Finally, we need $16m + 8 = 16\lceil \left(2 + \frac{1}{4C_{d,3}\varepsilon}\right) \ln \frac{1}{\delta} \rceil + 8$ copies of $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$. ■

Remark 81. *For small enough ε or large enough d, Algorithm 6 uses $\approx \frac{4}{C_{d,3}\varepsilon} \ln \frac{1}{\delta}$ copies of $|\psi\rangle$, which is less than Algorithm 5 for $r = 3$, i.e., $\approx \frac{6}{C_{d,3}\varepsilon} \ln \frac{1}{\delta}$.*

8.1 Tolerant stabiliser testing

Having worked with the "vanilla" testing problem of whether $F_{\mathcal{S}}(|\psi\rangle) = 1$ or $F_{\mathcal{S}}(|\psi\rangle) \leqslant 1 - \varepsilon$, we now focus on its tolerant version where one wants to test whether $F_{\mathcal{S}}(|\psi\rangle) \geqslant 1 - \varepsilon_1$ or $F_{\mathcal{S}}(|\psi\rangle) \leqslant 1 - \varepsilon_2$ for $\varepsilon_1 < \varepsilon_2$. A tolerant stabiliser testing procedure for qubits was proposed by Grewal et al. [GIKL24b] which requires $O(1/\gamma^2)$ copies of $|\psi\rangle$, where $\gamma \triangleq (1 - \epsilon_1)^6 - \frac{4-3\varepsilon_2}{4}$. Their algorithm thus works as long as $(1 - \varepsilon_1)^6 > \frac{4-3\varepsilon_2}{4}$. In order to prove their result, Grewal et al. [GIKL24b] employed the inequalities $F_{\mathcal{S}}(|\psi\rangle) \geqslant \frac{4G_3(|\psi\rangle)-1}{3}$ (a shaper version of Lemma 74 exclusive for $d = 2$) and $F_{\mathcal{S}}(|\psi\rangle) \leqslant G_3(|\psi\rangle)^6$, the latter coming from a simple combination of Hölder's inequality and Corollary 33 (see Lemma 82 below).

We propose two tolerant stabiliser testing algorithms, one based on the POVM of [GNW21] and one based on skewed Bell (difference) sampling, similarly to the previous section. For such, we must be able to bound the expected quantities $G_r(|\psi\rangle)$ and $A(|\psi\rangle)$ in terms of $F_{\mathcal{S}}(|\psi\rangle)$. Upper bounds were already proven in Lemma 74 and 79. We next provide corresponding lower bounds.

Lemma 82. *Let $r \in \mathbb{N}$, $r \geqslant 2$, such that $\gcd(d, r) = 1$. Then*

$$G_r(|\psi\rangle) \geqslant F_{\mathcal{S}}(|\psi\rangle)^{2r} \qquad \forall |\psi\rangle \in (\mathbb{C}^d)^{\otimes n}.$$

Proof.

$$\begin{aligned}
G_r(|\psi\rangle) &= d^{(r-1)n} \sum_{\mathbf{x} \in \mathbb{Z}_d^{2n}} p_\psi(\mathbf{x})^r && \text{(by Equation (7))} \\
&\geqslant d^{(r-1)n} \sum_{\mathbf{x} \in \mathscr{M}} p_\psi(\mathbf{x})^r \\
&\geqslant \left(\sum_{\mathbf{x} \in \mathscr{M}} p_\psi(\mathbf{x})\right)^r && \text{(Hölder's inequality)} \\
&\geqslant F_{\mathcal{S}}(|\psi\rangle)^{2r}. && \text{(by Corollary 33)}
\end{aligned}$$ ■

Lemma 83. *For any $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ with $F_{\mathcal{S}}(|\psi\rangle) \geqslant \frac{1}{2}$,*

$$A(|\psi\rangle) \geqslant (2F_{\mathcal{S}}(|\psi\rangle) - 1)^4 F_{\mathcal{S}}(|\psi\rangle)^{16}.$$

Proof.

$$\begin{aligned}
A(|\psi\rangle) &= d^{8n} \sum_{\mathbf{X}\in\mathbb{Z}_d^{2n\times 4}} \prod_{i=1}^{4} p_\psi(\mathbf{X}_i)^2 p_\psi((\mathbf{XR})_i) && \text{(by Lemma 78)}\\
&\geqslant d^{8n} \sum_{\mathbf{X}\in\mathscr{M}^4} \prod_{i=1}^{4} p_\psi(\mathbf{X}_i)^2 p_\psi((\mathbf{XR})_i)\\
&\geqslant (2F_{\mathcal{S}}(|\psi\rangle)-1)^4 \left(d^n \sum_{\mathbf{x}\in\mathscr{M}} p_\psi(\mathbf{x})^2\right)^4 && \text{(by Lemma 35)}\\
&\geqslant (2F_{\mathcal{S}}(|\psi\rangle)-1)^4 \left(\sum_{\mathbf{x}\in\mathscr{M}} p_\psi(\mathbf{x})\right)^8 && \text{(Hölder's inequality)}\\
&\geqslant (2F_{\mathcal{S}}(|\psi\rangle)-1)^4 F_{\mathcal{S}}(|\psi\rangle)^{16}. && \text{(by Corollary 33)} \quad \blacksquare
\end{aligned}$$

Our first tolerant stabiliser testing algorithm (Algorithm 7) employs the POVM $\{\Pi_r^+, \mathbf{I}-\Pi_r^+\}$ to estimate the expectation value $G_r(|\psi\rangle)$. If the estimate of $G_r(|\psi\rangle)$ is large enough, then $F_{\mathcal{S}}(|\psi\rangle) \geqslant 1-\varepsilon_1$ with high probability, otherwise $F_{\mathcal{S}}(|\psi\rangle) \leqslant 1-\varepsilon_2$. Due to Lemma 74 and 82,

$$F_{\mathcal{S}}(|\psi\rangle)^{2r} \leqslant G_r(|\psi\rangle) \leqslant 1 - 2C_{d,r}(1-F_{\mathcal{S}}(|\psi\rangle)).$$

Therefore, the cases $F_{\mathcal{S}}(|\psi\rangle) \geqslant 1-\varepsilon_1$ and $F_{\mathcal{S}}(|\psi\rangle) \leqslant 1-\varepsilon_2$ can be distinguished as long as

$$\gamma_r \triangleq (1-\varepsilon_1)^{2r} - 1 + 2C_{d,r}\varepsilon_2 \geqslant \frac{1}{\mathrm{poly}(n)}.$$

Algorithm 7: Tolerant stabiliser testing algorithm based on [GNW21]

Input: $2rm$ copies of $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ where $m \triangleq \lceil \frac{8}{\gamma_r^2} \ln \frac{2}{\delta} \rceil$, $r \geqslant 2$ is an integer such that $\gcd(d,r) = 1$, and $\gamma_r \triangleq (1-\varepsilon_1)^{2r} - 1 + \left(1 - (1-\frac{1}{4d^2})^{r-1}\right)\varepsilon_2$.

Promise: either $F_{\mathcal{S}}(|\psi\rangle) \geqslant 1-\varepsilon_1$ or $F_{\mathcal{S}}(|\psi\rangle) \leqslant 1-\varepsilon_2$.

Output: 1 if $F_{\mathcal{S}}(|\psi\rangle) \geqslant 1-\varepsilon_1$ and 0 if $F_{\mathcal{S}}(|\psi\rangle) \leqslant 1-\varepsilon_2$.

1 **for** $i = 1$ **to** m **do**

2 Measure the POVM $\{\Pi_r^+, \mathbf{I}-\Pi_r^+\}$ on $|\psi\rangle^{\otimes 2r}$ and obtain the outcome $Z_i \in \{-1,1\}$

3 **return** 1 if $\frac{1}{m}\sum_{i=1}^m Z_i > (1-\varepsilon_1)^{2r} - \frac{\gamma_r}{2}$ and 0 otherwise

Theorem 84. *Let $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$, $\delta \in (0,1)$, and $r \geqslant 2$ an integer such that $\gcd(d,r) = 1$. Let $\varepsilon_2 > \varepsilon_1 \geqslant 0$ such that $\gamma_r \triangleq (1-\varepsilon_1)^{2r} - 1 + \left(1 - (1-\frac{1}{4d^2})^{r-1}\right)\varepsilon_2 > 0$. Algorithm 7 accepts if $F_{\mathcal{S}}(|\psi\rangle) \geqslant 1-\varepsilon_1$ and rejects if $F_{\mathcal{S}}(|\psi\rangle) \leqslant 1-\varepsilon_2$ with probability at least $1-\delta$. It uses $2r\lceil \frac{8}{\gamma_r^2} \ln \frac{2}{\delta} \rceil$ copies of $|\psi\rangle$.*

Proof. Consider $m \triangleq \lceil \frac{8}{\gamma_r^2} \ln \frac{2}{\delta} \rceil$ i.i.d. copies $Z_1, \ldots, Z_m \in \{-1,1\}$ of the random variable defined by measuring the POVM $\{\Pi_r^+, \mathbf{I}-\Pi_r^+\}$ on $|\psi\rangle^{\otimes 2r}$ and the quantity $\widetilde{G}_r \triangleq \frac{1}{m}\sum_{i=1}^m Z_i$. Then $\mathbb{E}[\widetilde{G}_r] = G_r(|\psi\rangle)$. Algorithm 7 fails if $|\widetilde{G}_r - G_r(|\psi\rangle)| \geqslant \frac{\gamma_r}{2}$. According to Hoeffding's inequality,

$$\Pr\left[|\widetilde{G}_r - G_r(|\psi\rangle)| \geqslant \frac{\gamma_r}{2}\right] \leqslant 2\mathrm{e}^{-m\gamma_r^2/8} \leqslant \delta.$$

The number of employed copies of $|\psi\rangle$ is $2r\lceil \frac{8}{\gamma_r^2} \ln \frac{2}{\delta} \rceil$. $\blacksquare$

Our second tolerant stabiliser testing algorithm (Algorithm 8) employs skewed Bell difference sampling to estimate the expected value $A(|\psi\rangle)$ instead. Once again, If the estimate of $A(|\psi\rangle)$ is

large enough, then $F_{\mathcal{S}}(|\psi\rangle) \geqslant 1-\varepsilon_1$ with high probability, otherwise $F_{\mathcal{S}}(|\psi\rangle) \leqslant 1-\varepsilon_2$. According to Lemma 79 and 83,

$$(2F_{\mathcal{S}}(|\psi\rangle)-1)^4 F_{\mathcal{S}}(|\psi\rangle)^{16} \leqslant A(|\psi\rangle) \leqslant \big(1-2C_{d,3}(1-F_{\mathcal{S}}(|\psi\rangle))\big)^4.$$

Therefore, the cases $F_{\mathcal{S}}(|\psi\rangle) \geqslant 1-\varepsilon_1$ and $F_{\mathcal{S}}(|\psi\rangle) \leqslant 1-\varepsilon_2$ can be distinguished as long as

$$\alpha \triangleq (1-\varepsilon_1)^{16}(1-2\varepsilon_1)^4 - (1-2C_{d,3}\varepsilon_2)^4 \geqslant \frac{1}{\mathrm{poly}(n)}.$$

Algorithm 8: Alternative tolerant stabiliser testing algorithm

Input: $16m+8$ copies of $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$ where $m \triangleq \lceil \frac{8}{\alpha^2} \ln \frac{2}{\delta} \rceil$ and
$\alpha \triangleq (1-\varepsilon_1)^{16}(1-2\varepsilon_1)^4 - \big(1-\frac{\varepsilon_2}{2d^2}\big(1-\frac{1}{8d^2}\big)\big)^4$.
Promise: either $F_{\mathcal{S}}(|\psi\rangle) = 1$ or $F_{\mathcal{S}}(|\psi\rangle) \leqslant 1-\varepsilon$.
Output: 1 if $F_{\mathcal{S}}(|\psi\rangle) = 1$ and 0 if $F_{\mathcal{S}}(|\psi\rangle) \leqslant 1-\varepsilon$.

for $i \leftarrow 0$ **to** m **do**
  Perform Bell sampling on $(\mathbf{I} \otimes \mathcal{B}_{\mathbf{R}}^\dagger)|\psi\rangle^{\otimes 8}$ to obtain $\mathbf{X}^{(i)} \in \mathbb{Z}_d^{2n\times 4}$
  if $i \neq 0$ **then**
    $\mathbf{X}^{(i)} \leftarrow \mathbf{X}^{(i)} - \mathbf{X}^{(0)}$
    Measure $\frac{1}{2}\bigotimes_{j=1}^4 (\mathcal{W}_{\mathbf{X}_j^{(i)}} \otimes \mathcal{W}^\dagger_{\mathbf{X}_j^{(i)}}) + \frac{1}{2}\bigotimes_{j=1}^4 (\mathcal{W}^\dagger_{\mathbf{X}_j^{(i)}} \otimes \mathcal{W}_{\mathbf{X}_j^{(i)}})$ on $|\psi\rangle^{\otimes 8}$ to obtain $Z_i \in [-1,1]$
return 1 if $\frac{1}{m}\sum_{i=1}^m Z_i > (1-\varepsilon_1)^{16}(1-2\varepsilon_1)^4 - \frac{\alpha}{2}$ and 0 otherwise

Theorem 85. *Let $|\psi\rangle \in (\mathbb{C}^d)^{\otimes n}$, $\delta \in (0,1)$, and $\varepsilon_2 > \varepsilon_1 \geqslant 0$ such that $\alpha \triangleq (1-\varepsilon_1)^{16}(1-2\varepsilon_1)^4 - \big(1-\frac{\varepsilon_2}{2d^2}\big(1-\frac{1}{8d^2}\big)\big)^4 > 0$. Algorithm 8 accepts if $F_{\mathcal{S}}(|\psi\rangle) \geqslant 1-\varepsilon_1$ and rejects if $F_{\mathcal{S}}(|\psi\rangle) \leqslant 1-\varepsilon_2$ with probability at least $1-\delta$. It uses $16\lceil \frac{8}{\alpha^2} \ln \frac{2}{\delta} \rceil + 8$ copies of $|\psi\rangle$.*

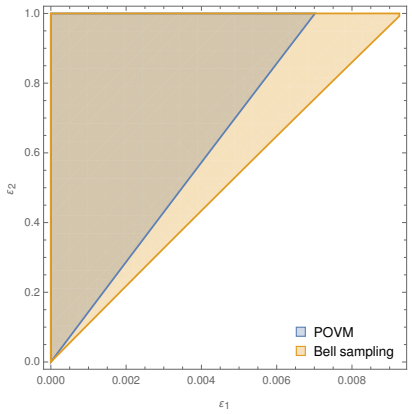
Proof. Consider $m \triangleq \lceil \frac{8}{\alpha^2} \ln \frac{2}{\delta} \rceil$ i.i.d. copies $Z_1, \ldots, Z_m \in [-1,1]$ of the random variable defined by measuring $\frac{1}{2}\bigotimes_{j=1}^4 (\mathcal{W}_{\mathbf{X}_j} \otimes \mathcal{W}^\dagger_{\mathbf{X}_j}) + \frac{1}{2}\bigotimes_{j=1}^4 (\mathcal{W}^\dagger_{\mathbf{X}_j} \otimes \mathcal{W}_{\mathbf{X}_j})$ on $|\psi\rangle^{\otimes 8}$ where $\mathbf{X} \sim b_\psi$ and $\widetilde{A} \triangleq \frac{1}{m}\sum_{i=1}^m Z_i$. Then $\mathbb{E}[\widetilde{A}] = A(|\psi\rangle)$. Algorithm 8 fails if $|\widetilde{A} - A(|\psi\rangle)| \geqslant \frac{\alpha}{2}$. According to Hoeffding's inequality,

$$\Pr\left[|\widetilde{A} - A(|\psi\rangle)| \geqslant \frac{\alpha}{2}\right] \leqslant 2\mathrm{e}^{-m\alpha^2/8} \leqslant \delta.$$

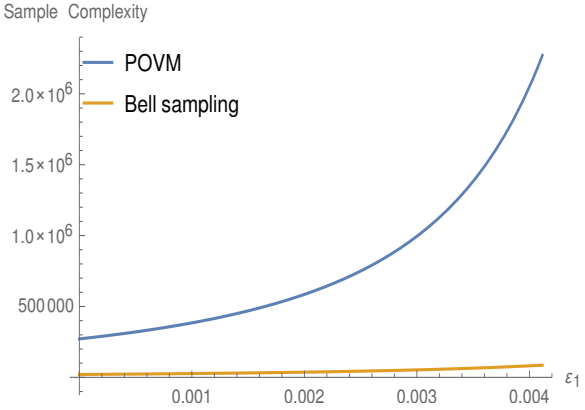
Instead of using 16 copies of $|\psi\rangle$ for each skewed Bell difference sampling, the outcome of one skewed Bell sampling can be subtracted from all the other, reducing the number of copies to 8 per sample. Another 8 copies are required to measure $\frac{1}{2}\bigotimes_{j=1}^4 (\mathcal{W}_{\mathbf{X}_j} \otimes \mathcal{W}^\dagger_{\mathbf{X}_j}) + \frac{1}{2}\bigotimes_{j=1}^4 (\mathcal{W}^\dagger_{\mathbf{X}_j} \otimes \mathcal{W}_{\mathbf{X}_j})$. The number of employed copies of $|\psi\rangle$ is thus $16\lceil \frac{8}{\alpha^2} \ln \frac{2}{\delta} \rceil + 8$. ∎

We compare the range of parameters $\varepsilon_2 > \varepsilon_1 > 0$ for which Algorithms 7 and 8 work in Figure 1, i.e., the values of $\varepsilon_1, \varepsilon_2$ for which $\gamma_r > 0$ and $\alpha > 0$. We also compare their sample complexities. Both algorithms have a similar range of parameters for small values of r, while Algorithm 7 can reach a larger range of parameters for high values of r at the expense of performing joint measurements involving $2r$ qudits. Algorithm 8, in comparison, only requires measurements across 8 qudits. Moreover, the Bell-sampling-based Algorithm 8 has a much better sample complexity for most values of ε_1.

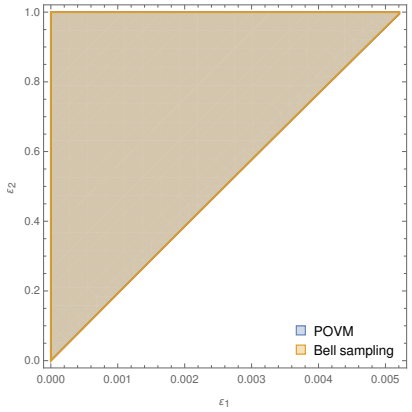
Acknowledgments. We thank David Gross for very useful discussions and ideas that led to the development of the unitary $\mathcal{B}_{\mathbf{R}}$. This research is funded by ERC grant No. 810115-DYNASNET. This work was done in part while JFD and MS were visiting the Simons Institute for the Theory of Computing.



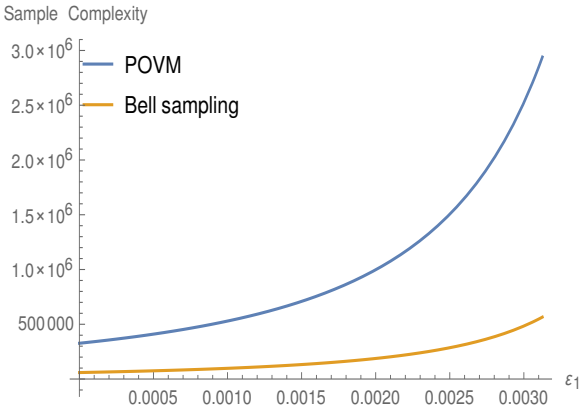
(a) Range comparison for dimension $d = 3$. The POVM-based algorithm uses $r = 2$.



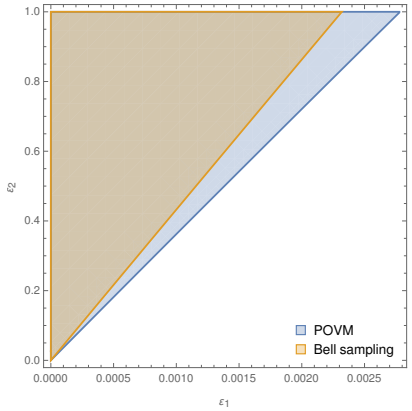
(b) Sample complexity comparison for $d = 3$. The POVM-based algorithm uses $r = 2$.



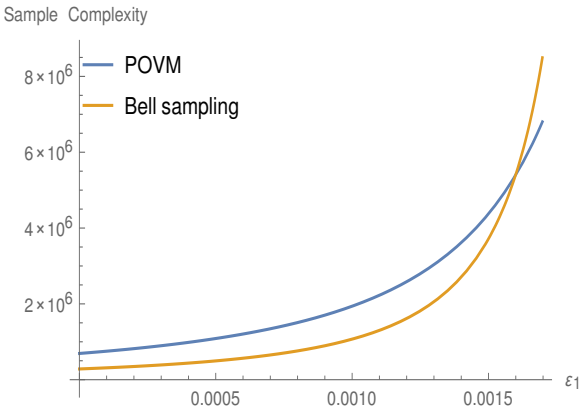
(c) Range comparison for dimension $d = 4$. The POVM-based algorithm uses $r = 3$.



(d) Sample complexity comparison for $d = 4$. The POVM-based algorithm uses $r = 3$.



(e) Range comparison for dimension $d = 6$. The POVM-based algorithm uses $r = 5$.



(f) Sample complexity comparison for $d = 6$. The POVM-based algorithm uses $r = 5$.

Figure 1: Comparison of sample complexities and range of parameters $\varepsilon_1, \varepsilon_2$ between Algorithm 7 (POVM-based) and Algorithm 8 (Bell-sampling-based). For the sample complexity, $\varepsilon_2 = 0.9$ and $\delta = 0.01$.

References

[AA24] Anurag Anshu and Srinivasan Arunachalam. A survey on the complexity of learning quantum states. *Nature Reviews Physics*, 6(1):59–69, Jan 2024. 3

[ABD24] Srinivasan Arunachalam, Sergey Bravyi, and Arkopal Dutt. A note on polynomial-time tolerant testing stabilizer states. *arXiv preprint arXiv:2410.22220*, 2024. 11

[ABDY23] Srinivasan Arunachalam, Sergey Bravyi, Arkopal Dutt, and Theodore J. Yoder. Optimal algorithms for learning quantum phase states. In Omar Fawzi and Michael Walter, editors, *18th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2023)*, volume 266 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 3:1–3:24, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. 7

[ABF+24] Scott Aaronson, Adam Bouland, Bill Fefferman, Soumik Ghosh, Umesh Vazirani, Chenyi Zhang, and Zixin Zhou. Quantum pseudoentanglement. In Venkatesan Guruswami, editor, *15th Innovations in Theoretical Computer Science Conference (ITCS 2024)*, volume 287 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 2:1–2:21, Dagstuhl, Germany, 2024. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. 9

[AD25] Srinivasan Arunachalam and Arkopal Dutt. Polynomial-time tolerant testing stabilizer states. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, STOC '25, page 1234–1241, New York, NY, USA, 2025. Association for Computing Machinery. 11, 12

[ADIS24] Jonathan Allcock, Joao F Doriguello, Gábor Ivanyos, and Miklos Santha. Beyond Bell sampling: stabilizer state learning and quantum pseudorandomness lower bounds on qudits. *arXiv preprint arXiv:2405.06357*, 2024. 3, 4, 7, 9, 12, 24, 37, 40

[AG04] Scott Aaronson and Daniel Gottesman. Improved simulation of stabilizer circuits. *Phys. Rev. A*, 70:052328, Nov 2004. 3

[AG08] Scott Aaronson and Daniel Gottesman. Identifying stabilizer states. http://pirsa.org/08080052/, 2008. 3, 7, 12

[AG23] Scott Aaronson and Sabee Grewal. Efficient tomography of non-interacting-fermion states. In Omar Fawzi and Michael Walter, editors, *18th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2023)*, volume 266 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 12:1–12:18, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. 7

[App05] D. M. Appleby. Symmetric informationally complete–positive operator valued measures and the extended Clifford group. *Journal of Mathematical Physics*, 46(5):052107, 04 2005. 4, 15, 17

[App09] D. M. Appleby. Properties of the extended Clifford group with applications to SIC-POVMs and MUBs. *arXiv preprint arXiv:0909.5233*, 2009. 4, 17

[AQY22] Prabhanjan Ananth, Luowen Qian, and Henry Yuen. Cryptography from pseudorandom quantum states. In Yevgeniy Dodis and Thomas Shrimpton, editors, *Advances in Cryptology – CRYPTO 2022*, pages 208–236, Cham, 2022. Springer Nature Switzerland. 9

[BB02] Jean-Luc Brylinski and Ranee Brylinski. Universal quantum gates. In *Mathematics of quantum computation*, pages 117–134. Chapman and Hall/CRC, 2002. 7, 15

[BBC+19] Sergey Bravyi, Dan Browne, Padraic Calpin, Earl Campbell, David Gosset, and Mark Howard. Simulation of quantum circuits by low-rank stabilizer decompositions. *Quantum*, 3:181, September 2019. 3, 10, 20

[BBO06] G. K. Brennen, S. S. Bullock, and D. P. O'Leary. Efficient circuits for exact-universal computationwith qudits. *Quantum Info. Comput.*, 6(4):436–454, jul 2006. 7, 15

[Bea13] Niel de Beaudrap. A linearized stabilizer formalism for systems of finite dimension. *Quant. Inf. Comput.*, 13:0073–0115, 2013. 15, 17

[BGTW25] Adam Bouland, Tudor Giurgică-Tiron, and John Wright. The state hidden subgroup problem and an efficient algorithm for locating unentanglement. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, STOC '25, page 463–470, New York, NY, USA, 2025. Association for Computing Machinery. 8, 32

[BOB05] Gavin K. Brennen, Dianne P. O'Leary, and Stephen S. Bullock. Criteria for exact qudit universality. *Phys. Rev. A*, 71:052318, May 2005. 7, 15

[BSS16] Sergey Bravyi, Graeme Smith, and John A. Smolin. Trading classical and quantum computational resources. *Phys. Rev. X*, 6:021043, Jun 2016. 3

[BvDH25] Zongbo Bao, Philippe van Dordrecht, and Jonas Helsen. Tolerant testing of stabilizer states with a polynomial gap via a generalized uncertainty relation. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, STOC '25, page 1254–1262, New York, NY, USA, 2025. Association for Computing Machinery. 11, 12

[CGYZ25] Sitan Chen, Weiyuan Gong, Qi Ye, and Zhihan Zhang. Stabilizer bootstrapping: A recipe for efficient agnostic tomography and magic estimation. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, STOC '25, page 429–438, New York, NY, USA, 2025. Association for Computing Machinery. 11, 33

[CPF+10] Marcus Cramer, Martin B. Plenio, Steven T. Flammia, Rolando Somma, David Gross, Stephen D. Bartlett, Olivier Landon-Cardinal, David Poulin, and Yi-Kai Liu. Efficient quantum state tomography. *Nature Communications*, 1(1):149, Dec 2010. 7

[CS96] A. R. Calderbank and Peter W. Shor. Good quantum error-correcting codes exist. *Phys. Rev. A*, 54:1098–1105, Aug 1996. 3

[Dam18] Raja Oktovin Parhasian Damanik. Optimality in stabilizer testing. *Report, August*, 2018. 40

[DB13] Niel De Beaudrap. A linearized stabilizer formalism for systems of finite dimension. *Quantum Info. Comput.*, 13(1–2):73–115, January 2013. 20

[DHIS14] Thomas Decker, Peter Høyer, Gábor Ivanyos, and Miklos Santha. Polynomial time quantum algorithms for certain bivariate hidden polynomial problems. *Quant. Inf. Comput.*, 14:790–806, 2014. 7

[GIKL23] Sabee Grewal, Vishnu Iyer, William Kretschmer, and Daniel Liang. Low-stabilizer-complexity quantum states are not pseudorandom. In Yael Tauman Kalai, editor, *14th Innovations in Theoretical Computer Science Conference (ITCS 2023)*, volume 251 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 64:1–64:20, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. 3, 9, 12, 37

[GIKL24a] Sabee Grewal, Vishnu Iyer, William Kretschmer, and Daniel Liang. Agnostic tomography of stabilizer product states. *arXiv preprint arXiv:2404.03813*, 2024. 3, 6, 7

[GIKL24b] Sabee Grewal, Vishnu Iyer, William Kretschmer, and Daniel Liang. Improved stabilizer estimation via Bell difference sampling. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, STOC 2024, page 1352–1363, New York, NY, USA, 2024. Association for Computing Machinery. 3, 6, 7, 9, 11, 12, 30, 37, 38, 40, 44

[GIKL24c] Sabee Grewal, Vishnu Iyer, William Kretschmer, and Daniel Liang. Pseudoentanglement ain't cheap. *arXiv preprint arXiv:2404.00126*, 2024. 3, 6, 7

[GIKL25] Sabee Grewal, Vishnu Iyer, William Kretschmer, and Daniel Liang. Efficient Learning of Quantum States Prepared With Few Non-Clifford Gates. *Quantum*, 9:1907, November 2025. 3, 6, 7, 8, 9, 12, 33, 36

[GNW21] David Gross, Sepehr Nezami, and Michael Walter. Schur–Weyl duality for the Clifford group with applications: Property testing, a robust Hudson theorem, and de Finetti representations. *Communications in Mathematical Physics*, 385(3):1325–1393, Aug 2021. 3, 4, 6, 7, 9, 10, 11, 12, 16, 17, 18, 23, 29, 30, 40, 41, 42, 44, 45

[GOL25] Andi Gu, Salvatore F.E. Oliviero, and Lorenzo Leone. Magic-induced computational separation in entanglement theory. *PRX Quantum*, 6:020324, May 2025. 3

[Got96] Daniel Gottesman. Class of quantum error-correcting codes saturating the quantum hamming bound. *Phys. Rev. A*, 54:1862–1868, Sep 1996. 3

[Got97] Daniel Gottesman. *Stabilizer codes and quantum error correction*. California Institute of Technology, 1997. 3

[Gro06] D. Gross. Hudson's theorem for finite-dimensional quantum systems. *Journal of Mathematical Physics*, 47(12):122107, 12 2006. 14, 15, 17, 18

[HBK25] Tobias Haug, Kishor Bharti, and Dax Enshan Koh. Pseudorandom unitaries are neither real nor sparse nor noise-robust. *Quantum*, 9:1759, June 2025. 9

[HDDM05] Erik Hostens, Jeroen Dehaene, and Bart De Moor. Stabilizer states and Clifford operations for systems of arbitrary dimensions and modular arithmetic. *Phys. Rev. A*, 71:042315, Apr 2005. 5, 19

[HEC25] Marcel Hinsche, Jens Eisert, and Jose Carrasco. The abelian state hidden subgroup problem: Learning stabilizer groups and beyond. *arXiv preprint arXiv:2505.15770*, 2025. 6, 7, 8, 9, 12, 32, 33, 35

[HG24] Dominik Hangleiter and Michael J. Gullans. Bell sampling from quantum circuits. *Phys. Rev. Lett.*, 133:020601, Jul 2024. 3, 7

[HH25] Marcel Hinsche and Jonas Helsen. Single-copy stabilizer testing. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, STOC '25, page 439–450, New York, NY, USA, 2025. Association for Computing Machinery. 11

[HIN+23] M. Hinsche, M. Ioannou, A. Nietner, J. Haferkamp, Y. Quek, D. Hangleiter, J.-P. Seifert, J. Eisert, and R. Sweke. One T gate makes distribution learning hard. *Phys. Rev. Lett.*, 130:240602, Jun 2023. 10

[HK23] Tobias Haug and M.S. Kim. Scalable measures of magic resource for quantum computers. *PRX Quantum*, 4:010301, Jan 2023. 3, 7

[HKP20] Hsin-Yuan Huang, Richard Kueng, and John Preskill. Predicting many properties of a quantum system from very few measurements. *Nature Physics*, 16(10):1050–1057, Oct 2020. 3

[HLK24] Tobias Haug, Soovin Lee, and M. S. Kim. Efficient quantum algorithms for stabilizer entropies. *Phys. Rev. Lett.*, 132:240602, Jun 2024. 3, 7

[HMY23] Minki Hhan, Tomoyuki Morimae, and Takashi Yamakawa. From the hardness of detecting superpositions to cryptography: Quantum public key encryption and commitments. In Carmit Hazay and Martijn Stam, editors, *Advances in Cryptology – EUROCRYPT 2023*, pages 639–667, Cham, 2023. Springer Nature Switzerland. 9

[HNQ+16] Patrick Hayden, Sepehr Nezami, Xiao-Liang Qi, Nathaniel Thomas, Michael Walter, and Zhao Yang. Holographic duality from random tensor networks. *Journal of High Energy Physics*, 2016(11):9, Nov 2016. 3

[Hol73] Alexander Semenovich Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problemy Peredachi Informatsii*, 9(3):3–11, 1973. 7

[HWFW19] Jonas Helsen, Joel J. Wallman, Steven T. Flammia, and Stephanie Wehner. Multiqubit randomized benchmarking using few samples. *Phys. Rev. A*, 100:032304, Sep 2019. 3

[IL24] Vishnu Iyer and Daniel Liang. Tolerant testing of stabilizer states with mixed state inputs. *arXiv preprint arXiv:2411.08765*, 2024. 11, 12

[IS17] Gábor Ivanyos and Miklos Santha. Solving systems of diagonal polynomial equations over finite fields. *Theoretical Computer Science*, 657:73–85, 2017. Frontiers of Algorithmics. 7

[JLS18] Zhengfeng Ji, Yi-Kai Liu, and Fang Song. Pseudorandom quantum states. In Hovav Shacham and Alexandra Boldyreva, editors, *Advances in Cryptology – CRYPTO 2018*, pages 126–152, Cham, 2018. Springer International Publishing. 9, 40

[KLR+08] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland. Randomized benchmarking of quantum gates. *Phys. Rev. A*, 77:012307, Jan 2008. 3

[KQST23] William Kretschmer, Luowen Qian, Makrand Sinha, and Avishay Tal. Quantum cryptography in algorithmica. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, STOC 2023, page 1589–1602, New York, NY, USA, 2023. Association for Computing Machinery. 9

[Kre21] William Kretschmer. Quantum Pseudorandomness and Classical Complexity. In Min-Hsiu Hsieh, editor, *16th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2021)*, volume 197 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 2:1–2:20, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. 9

[KZG16] Richard Kueng, Huangjun Zhu, and David Gross. Low rank matrix recovery from clifford orbits. *arXiv preprint arXiv:1610.08070*, 2016. 3

[LCLP10] Olivier Landon-Cardinal, Yi-Kai Liu, and David Poulin. Efficient direct tomography for matrix product states. *arXiv preprint arXiv:1002.4632*, 2010. 7

[Led01] Michel Ledoux. *The Concentration of Measure Phenomenon*. Mathematical surveys and monographs. American Mathematical Society, 2001. 37

[LOH24] Lorenzo Leone, Salvatore F. E. Oliviero, and Alioscia Hamma. Learning t-doped stabilizer states. *Quantum*, 8:1361, May 2024. 3, 38

[LOLH24] Lorenzo Leone, Salvatore F. E. Oliviero, Seth Lloyd, and Alioscia Hamma. Learning efficient decoders for quasichaotic quantum scramblers. *Phys. Rev. A*, 109:022429, Feb 2024. 3, 38

[MGE11] Easwar Magesan, J. M. Gambetta, and Joseph Emerson. Scalable and robust randomized benchmarking of quantum processes. *Phys. Rev. Lett.*, 106:180504, May 2011. 3

[MHZ+23] Michael Meth, Jan F Haase, Jinglei Zhang, Claire Edmunds, Lukas Postler, Alex Steiner, Andrew J Jena, Luca Dellantonio, Rainer Blatt, Peter Zoller, et al. Simulating 2D lattice gauge theories on a qudit quantum computer. *arXiv preprint arXiv:2310.12110*, 2023. 7

[Mon17] Ashley Montanaro. Learning stabilizer states by Bell sampling. *arXiv preprint arXiv:1707.04012*, 2017. 3, 4, 7, 8, 12, 23

[MS86] Vitali D. Milman and Gideon Schechtman. *Asymptotic theory of finite dimensional normed spaces: Isoperimetric inequalities in Riemannian manifolds*, volume 1200. Springer Science & Business Media, 1986. 37

[MS00] Ashok Muthukrishnan and C. R. Stroud. Multivalued logic gates for quantum computation. *Phys. Rev. A*, 62:052309, Oct 2000. 7, 15

[MT25] Saeed Mehraban and Mehrdad Tahmasbi. Improved bounds for testing low stabilizer complexity states. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, STOC '25, page 1222–1233, New York, NY, USA, 2025. Association for Computing Machinery. 11, 12

[MW16] Ashley Montanaro and Ronald de Wolf. *A Survey of Quantum Property Testing*. Number 7 in Graduate Surveys. Theory of Computing Library, 2016. 3

[MY22] Tomoyuki Morimae and Takashi Yamakawa. Quantum commitments and signatures without one-way functions. In Yevgeniy Dodis and Thomas Shrimpton, editors, *Advances in Cryptology – CRYPTO 2022*, pages 269–295, Cham, 2022. Springer Nature Switzerland. 9

[NW20] Sepehr Nezami and Michael Walter. Multipartite entanglement in stabilizer tensor networks. *Phys. Rev. Lett.*, 125:241602, Dec 2020. 3

[O'D14] Ryan O'Donnell. *Analysis of Boolean functions*. Cambridge University Press, 2014. 14

[OLLH24] Salvatore F. E. Oliviero, Lorenzo Leone, Seth Lloyd, and Alioscia Hamma. Unscrambling quantum information with Clifford decoders. *Phys. Rev. Lett.*, 132:080402, Feb 2024. 3, 38

[PT18] Paul Pollack and Enrique Treviño. Finding the four squares in Lagrange's theorem. *Integers*, 18(A15):7–17, 2018. 27

[RB00] Robert Raussendorf and Hans J. Briegel. Quantum computing via measurements only. *arXiv preprint quant-ph/0010033*, 2000. 3

[Röt09] Martin Rötteler. Quantum algorithms to solve the hidden shift problem for quadratics and for functions of large Gowers norm. In Rastislav Královič and Damian Niwiński, editors, *Mathematical Foundations of Computer Science 2009*, pages 663–674, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg. 3, 7

[RS86] Michael O. Rabin and Jeffery O. Shallit. Randomized algorithms in number theory. *Communications on Pure and Applied Mathematics*, 39(S1):S239–S256, 1986. 27

[Rud90] Walter Rudin. *Fourier analysis on groups*. Wiley-Interscience, New York, 1990. 18

[Sho95] Peter W. Shor. Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A*, 52:R2493–R2496, Oct 1995. 3

[Sto96] Arne Storjohann. Near optimal algorithms for computing Smith normal forms of integer matrices. In *Proceedings of the 1996 International Symposium on Symbolic and Algebraic Computation*, ISSAC '96, page 267–274, New York, NY, USA, 1996. Association for Computing Machinery. 7, 32

[Wol08] Ronald de Wolf. A brief introduction to Fourier analysis on the Boolean cube. *Theory of Computing*, pages 1–20, 2008. 14

[Woo87] William K Wootters. A Wigner-function formulation of finite-state quantum mechanics. *Annals of Physics*, 176(1):1–21, 1987. 15

[ZPDF16] Liming Zhao, Carlos A. Pérez-Delgado, and Joseph F. Fitzsimons. Fast graph operations in quantum computation. *Phys. Rev. A*, 93:032314, Mar 2016. 3, 7, 12