Schur-Weyl duality for the Clifford group with applications: Property testing, a robust Hudson theorem, and de Finetti representations

David Gross[*]   Sepehr Nezami[†]   Michael Walter[‡,†]

Abstract

Schur-Weyl duality is a ubiquitous tool in quantum information. At its heart is the statement that the space of operators that commute with the t-fold tensor powers $U^{\otimes t}$ of all unitaries $U \in U(d)$ is spanned by the permutations of the t tensor factors. In this work, we describe a similar duality theory for tensor powers of *Clifford unitaries*. The Clifford group is a central object in many subfields of quantum information, most prominently in the theory of fault-tolerance. The duality theory has a simple and clean description in terms of finite geometries. We demonstrate its effectiveness in several applications:

- We resolve an open problem in *quantum property testing* by showing that "stabilizerness" is efficiently testable: There is a protocol that, given access to six copies of an unknown state, can determine whether it is a stabilizer state, or whether it is far away from the set of stabilizer states. We give a related membership test for the Clifford group.

- We find that tensor powers of stabilizer states have an increased symmetry group. Conversely, we provide corresponding *de Finetti theorems*, showing that the reductions of arbitrary states with this symmetry are well-approximated by mixtures of stabilizer tensor powers (in some cases, exponentially well).

- We show that the distance of a pure state to the set of stabilizers can be lower-bounded in terms of the sum-negativity of its Wigner function. This gives a new quantitative meaning to the sum-negativity (and the related *mana*) – a measure relevant to fault-tolerant quantum computation. The result constitutes a *robust* generalization of the *discrete Hudson theorem*.

- We show that complex projective designs of arbitrary order can be obtained from a finite number (independent of the number of qudits) of Clifford orbits. To prove this result, we give explicit formulas for arbitrary moments of random stabilizer states.

[*]Institute for Theoretical Physics, University of Cologne, and Centre for Engineered Quantum Systems, School of Physics, The University of Sydney, `david.gross@thp.uni-koeln.de`. [†]Stanford Institute for Theoretical Physics, Stanford University, and Kavli Institute for Theoretical Physics, `nezami@stanford.edu`. [‡]Korteweg-de Vries Institute for Mathematics, Institute of Theoretical Physics, and Institute for Logic, Language and Computation, QuSoft, University of Amsterdam, `m.walter@uva.nl`.

Contents

1 Introduction

1.1 Background

Schur-Weyl duality. To motivate the symmetry this work is based on, we start by considering two types of problems that have frequently appeared in quantum information theory. First, assume that we have access to t copies $\rho^{\otimes t}$ of an unknown quantum state ρ on $\mathbb{C}^d$, and that we are interested in some property of ρ's eigenvalues (for example its entropy). Clearly, then, the problem has a $U^{\otimes t}$-symmetry in the sense that the inputs $\rho^{\otimes t}$ and

$$U^{\otimes t}(\rho^{\otimes t})U^{\dagger \otimes t}$$

represent equivalent properties. It thus makes sense to design a procedure that shares the $U^{\otimes t}$-symmetry, and indeed the resulting procedure has been shown to be optimal for estimating the eigenvalues [KW01, HM02, CM06, CHM07, OW15]. Moreover, consider *quantum state tomography*, the task of estimating the entire quantum state ρ. Essentially optimal estimators can be constructed by first estimating the eigenvalues and then the eigenbasis [OW16, OW17, HHJ$^+$17], crucially using the structure of $U^{\otimes t}$ in each step. There are many further problems in quantum information where this symmetry can be exploited – for example in quantum Shannon theory, where optimal rates mostly depend only on the eigenvalues of the quantum state [HM02, Har05].

Second, studying the properties of a Haar-random state vector $|\psi\rangle$ has proven to be extremely fruitful [HLW06, Has08]. Instead of working with the full distribution, it is often sufficient to exploit information about the statistical *moments* of the random matrix $|\psi\rangle\langle\psi|$. The t-th moment is described by the expected value of the t-th tensor power of the random matrix:

$$M_t = \mathbb{E}_{\psi \text{ Haar}}[(|\psi\rangle\langle\psi|)^{\otimes t}]. \tag{1.1}$$

Again, M_t is invariant under conjugation by $U^{\otimes t}$, $U \in U(\mathbb{C}^d)$.

The importance of Schur-Weyl duality in quantum information stems from the fact that it allows one to characterize the set of $U^{\otimes t}$-invariant operators on $(\mathbb{C}^d)^{\otimes t}$. Indeed, it implies that any such operator can be expressed as the linear combination of matrices r_π, $\pi \in S_t$ that act by permuting the tensor factors:

$$r_\pi(|\psi_1\rangle \otimes \cdots \otimes |\psi_t\rangle) = |\psi_{\pi_1}\rangle \otimes \cdots \otimes |\psi_{\pi_t}\rangle. \tag{1.2}$$

Clifford group and stabilizer states. Arguably, the subgroup of the full unitary group that is most important to quantum information is the Clifford group. The Clifford group and the closely related concept of stabilizer states and stabilizer codes feature centrally in fault-tolerant quantum computing, quantum coding in general, randomized benchmarking, measurement-based quantum computing, and many other subfields of quantum information.

To introduce the Clifford group, we first recall the definition of the set of *Pauli operators*. For a *qudit* (d-dimensional system), they are defined by their action on a some basis $\{|q\rangle\}_{q=0}^{d-1}$ via

$$X|q\rangle = |q+1\rangle, \quad Z|q\rangle = e^{2\pi i q/d}|q\rangle.$$

For n qudits, the Pauli group is defined as the finite group generated by the Pauli operators on each qudit. The *Clifford group* now is the natural symmetry group of the Pauli group. That is, a unitary U is Clifford if, for any Pauli operator P, $UPU^\dagger$ is again in the Pauli group. Ignoring overall phases,

the Clifford group is a finite group, which is intimately connected to the *metaplectic representation* of the discrete *symplectic group* (see, e.g., [Gro06]). Closely related to the Clifford group is the set of *stabilizer states*. These are the states that can be obtained by acting on a basis vector $|0 \ldots 0\rangle$ by arbitrary Clifford unitaries.

As before, there are many natural problems that are invariant under $U^{\otimes t}$, for U a Clifford unitary. Two examples we will discuss are: (1) Given access to $\psi^{\otimes t}$, decide whether ψ is a stabilizer state; (2) What are the t-th moments of a random stabilizer state ψ?

Randomized constructions. Another motivation arises from randomized constructions. Unitaries and states drawn from the Haar measure appear in many situations, including in quantum cryptography, coding, and data hiding [HLW06]. While randomized constructions are often near-optimal and frequently out-perform all known deterministic constructions, they have the drawback that generic quantum states cannot be efficiently prepared.

This contrasts with random Clifford unitaries and random stabilizer states, both of which can be efficiently realized (they require at most $O(n^2)$ gates to implement in a quantum circuit) [AG04]. They have therefore repeatedly been suggested as "drop-in replacements" for their Haar-measure analogues. Examples include randomized benchmarking [MGE11, HWFW17], low-rank recovery [KZG16], and tensor networks in the context of holography [HNQ$^+$16, NW16]. All these applications require information about the moments (in the sense of Eq. (1.1)) of random stabilizer states, which they all obtain from representation-theoretic data. To date, this representation theory and the associated stabilizer moments are understood only up to order $t = 4$ [ZKGG16, HWW16, NW16]. This contrasts with the Haar-random case, where Schur-Weyl duality gives this information for arbitrary orders t. Making analogous techniques available for the Clifford case was one important motivation for this work. Higher moments will generally lead to tighter performance bounds in randomized constructions, and are strictly required for some applications, like the *stabilizer testing problem* resolved here.

1.2 Schur-Weyl duality for the Clifford group

We start with an explicit description of the commutant of tensor powers of Clifford unitaries. While such a description has not yet appeared in the quantum information literature, we emphasize that some of the key results can already be deduced from work by Nebe, Rains, Sloane and colleagues on invariants of self-dual codes (see the excellent monograph [NRS06]). Also, in representation theory, there is a separate stream of closely related work regarding the structure of the oscillator representation and attempts to develop a Howe duality theory over finite fields, which is still an open problem (see, e.g., [How73, GH16] and references therein). We discovered the approach presented below independently, starting from our results in [NW16, App. C] for third tensor powers. Our proofs differ fundamentally from the preceding works in that they rely on the phase space formalism of finite-dimensional quantum mechanics, which offers additional insight.

To construct the commutant, start with the permutations r_π on $(\mathbb{C}^d)^{\otimes t}$ of Eq. (1.2). We assume for now that $\mathbb{C}^d$ is the Hilbert space of a single qudit with "computational basis" $\{|x\rangle\}_{x \in \mathbb{Z}_d}$ labeled by elements in $\mathbb{Z}_d = \mathbb{Z}/d\mathbb{Z}$ (this is anyway required for defining the Pauli and the Clifford group). Basis elements $|\mathbf{x}\rangle = |x_1\rangle \otimes \cdots \otimes |x_t\rangle$ of $(\mathbb{C}^d)^{\otimes t}$ are then labeled by vectors $\mathbf{x} \in \mathbb{Z}_d^t$. In this language:

$$r_\pi = \sum_{\mathbf{y} \in \mathbb{Z}_d^t} |\pi(\mathbf{y})\rangle \langle \mathbf{y}| = \sum_{(\mathbf{x},\mathbf{y}) \in T_\pi} |\mathbf{x}\rangle \langle \mathbf{y}|, \tag{1.3}$$

where $T_\pi = \{(\pi(\mathbf{y}), \mathbf{y}) : \mathbf{y} \in \mathbb{Z}_d^t\}$ and π permutes the components of $\mathbf{y}$. Because the Clifford group is a subgroup of the unitaries, the commutant is in general strictly larger. We thus have to add further operators to the r_π's in order to find a complete set.

The central message of this section is that, surprisingly, a minor modification of (1.3) suffices! Indeed, for any subspace T of $\mathbb{Z}_d^t \oplus \mathbb{Z}_d^t$ define

$$r(T) = \sum_{(\mathbf{x},\mathbf{y}) \in T} |\mathbf{x}\rangle \langle \mathbf{y}|.$$

We also consider the n-fold tensor power $R(T) := r(T)^{\otimes n}$, which is an operator on $((\mathbb{C}^d)^{\otimes t})^{\otimes n} \cong (\mathbb{C}^d)^{\otimes tn} \cong ((\mathbb{C}^d)^{\otimes n})^{\otimes t}$.

We now single out subspaces that satisfy certain geometric properties. Reflecting a well-known difference between even and odd dimensions in the stabilizer formalism, we define $D = d$ if d is odd, and $D = 2d$ if d is even.

Definition 4.1 ($\Sigma_{t,t}$). *Consider the quadratic form* $q \colon \mathbb{Z}_d^{2t} \to \mathbb{Z}_D$ *defined by* $q(\mathbf{x}, \mathbf{y}) := \mathbf{x} \cdot \mathbf{x} - \mathbf{y} \cdot \mathbf{y}$.[1] *We denote by* $\Sigma_{t,t}(d)$ *the set of subspaces* $T \subseteq \mathbb{Z}_d^{2t}$ *satisfying the following properties:*

1. *T is* totally q-isotropic*: i.e.,* $\mathbf{x} \cdot \mathbf{x} = \mathbf{y} \cdot \mathbf{y} \pmod{D}$ *for all* $(\mathbf{x}, \mathbf{y}) \in T$.

2. *T has dimension* t *(the maximal possible dimension).*

3. *T is* stochastic*:* $\mathbf{1}_{2t} = (1, \ldots, 1) \in T$.

We will summarize the first two conditions by saying that T is Lagrangian*. Thus, we will call $\Sigma_{t,t}(d)$ the set of* stochastic Lagrangian subspaces*.*

Our first main result is the following theorem, which states that the operators $R(T)$ obtained from these subspaces are a basis of the commutant:

Theorem 4.3 (Commutant of Clifford tensor powers)**.** *Let d be a prime and $n \geqslant t - 1$. Then the operators $R(T) = r(T)^{\otimes n}$ for $T \in \Sigma_{t,t}(d)$ are $\prod_{k=0}^{t-2}(d^k + 1)$ many linearly independent operators that span the commutant of the t-th tensor power action of the Clifford group for n qudits.*

Proof sketch. We use the phase space formalism of finite-dimensional quantum mechanics developed in [Woo87, App05, Gro06, GE08, DB13]. In particular, Clifford unitaries have a simple description on phase space: they act by affine symplectic transformations.

We use this structure to give a concise proof that the operators $R(T)$ commute with $U^{\otimes t}$ for any Clifford unitary. The linear independence is not hard, so it remains to argue that the number of subspaces equals the dimension of the commutant. We show this by a careful counting argument. We first compute the number of stochastic Lagrangian subspaces. Employing the fundamental Witt's theorem, we find recursive relations for the dimension of commutant of the Clifford group. We solve this recursion using Gaussian binomial identities (the result is a generalization of [Zhu15, (8)–(10)]) and find that the cardinalities match, concluding the proof. $\square$

There is a rich structure associated with the objects appearing in this theorem: It is easy to see that the spaces $T_\pi = \{(\pi(\mathbf{y}), \mathbf{y})\}$ that give rise to the commutant of $U(d)$ appear as special cases above. For general d and t, not all $R(T)$'s are invertible. In particular, for some T's, $R(T)$ is proportional to

---

[1]Note that for $x \in \mathbb{Z}_d$, x^2 is well-defined modulo D.

the projection onto a stabilizer code. This way, one can e.g., recover the code that has been used to describe the irreps contained in the 4th tensor power of the Clifford group in [ZKGG16]. The set of invertible $R(T)$'s are associated with spaces T of the form $(A\mathbf{y}, \mathbf{y})$, for A that are elements of a certain "stochastic orthogonal" group $O_t(d)$. This group is of interest to the formulation of modular Howe duality [GH16], and underlies several of our applications below.

Remarkably, the size of the commutant *stabilizes* as soon as $n \geqslant t - 1$. That is, just like the symmetric group in Schur-Weyl duality, the set that parametrizes the commutant of the Clifford tensor powers is independent of the number n of qudits. The fact that the operators $R(T) = r(T)^{\otimes n}$ are themselves tensor powers facilitates possible physical implementations. This, once more, generalizes a property of the symmetric group in Schur-Weyl duality.

To find novel applications of this theory, it is helpful to identify a set of non-trivial T's that afford an intuitive interpretation. Several of our multi-qubit results presented below are based on spaces with elements $(\bar{\pi}\mathbf{y}, \mathbf{y})$, where $\bar{\pi}$ is what we refer to as an *anti-permutation*. An anti-permutation is simply the binary complement of a permutation matrix. Formally, $\bar{\pi} = \mathbf{1}_t \mathbf{1}_t^\top - \pi$, where $\mathbf{1}_t = (1, \ldots, 1)$ contains t ones, and $\pi \in S_t$. Its operator representation is particularly straightforward. The n-qubit anti-identity, e.g., acts by

$$R(\bar{\mathbb{1}}) = 2^{-n} \left(I^{\otimes t} + X^{\otimes t} + Y^{\otimes t} + Z^{\otimes t} \right)^{\otimes n}, \tag{1.4}$$

which greatly facilitates the analysis (cf. Eq. (3.13) and Definition 4.29).

1.3 Quantum property testing: stabilizer testing

The theory of *quantum property testing* asks which properties of a "black box" many-body quantum system can be learned efficiently—in particular without having to resort to costly full tomography [BFNR03, BFNR08, MdW13]. A prototypical example of a testable property is *purity*. Indeed, given access to two copies $\rho \otimes \rho$ of an unknown quantum state ρ, the so-called *swap test* provides for a simple protocol that accepts with certainty if $\rho = |\psi\rangle \langle\psi|$ is pure, and rejects with probability $\Theta(1/\varepsilon^2)$ if ρ is ε-far away from the set of pure states in trace distance. The test is *perfectly complete* in the sense that it has a type-I error rate of zero (pure states are accepted with certainty); it requires a number of copies (two) that is independent of the dimension. It is also *transversal* in the sense that if ρ acts on n qubits, all operations are required to be coherent only across the two copies, and factorize w.r.t. the n qubits.

An open problem in this theory was whether *stabilizerness* and *Cliffordness* are testable properties of, respectively, states and unitaries [MdW13]. Both properties are clearly Clifford-invariant—so by the arguments presented in the introduction, it makes sense to search for tests in the commutant of the Clifford group. It is known that 2nd and 3rd moments of random stabilizer states are identical to the moments of Haar-random states [Zhu15, KG15, Web16]. This implies that three copies of a state are not sufficient to test for stabilizerness, and the results of [ZKGG16] can be used to show that four copies are also insufficient for a dimension-independent theory.

Prior work. Prior to our results, the best known algorithms for stabilizer testing required a number of copies that scaled linearly with n, the number of qubits. Indeed, these algorithms proceeded by attempting to *identify* the stabilizer state, which necessarily requires $\Omega(n)$ copies by the Holevo bound [AG08, Mon17, ZPDF16, KR08]. However, the existence of tests that require only a constant number of copies has been an important open question [MdW13]. We note that the stabilizer testing

Input: Six copies of an unknown multi-qubit quantum state ($\psi^{\otimes 6}$).

1. Perform *Bell difference sampling*: That is, Bell sample twice (on two independent copies of $\psi^{\otimes 2}$), with outcomes **x**, **y**, and set $\mathbf{a} = \mathbf{x} - \mathbf{y}$.

2. Measure the Weyl operator $W_\mathbf{a}$ twice (on two independent copies of ψ). Accept iff the outcomes agree.

Algorithm 1: Algorithm for testing whether an unknown multi-qubit state is a stabilizer state.

problem asks whether a given state is *any* stabilizer state – which is distinct from the problem of verifying whether it equals some *fixed* stabilizer state [HM15].

Our results. We show that for n qudits $O(1)$ copies suffice to give an *efficient*, *perfectly complete*, *dimension-independent*, and *transversal* test. For example, for qubits ($d = 2$) our test requires only 6 copies of the state to achieve a power independent of n (Algorithm 1). It requires coherent operations on only two qubits at a time, which means in particular that it can be implemented given a source that creates two copies of a fixed state at a time (Fig. 2).

First, we consider the problem for qubits. Here our protocol affords an intuitive description using a new primitive which we call *Bell difference sampling*. Then we proceed to the general case and discuss the connection to the commutant of the Clifford group described in Section 1.2.

Qubits: Bell difference sampling

We start with an intuitive motivation of the test. Let $|\psi\rangle \langle\psi| = 2^{-n/2} \sum_\mathbf{a} c_\mathbf{a} W_\mathbf{a}$ be its expansion w.r.t. the Weyl operators $W_\mathbf{a}$ (which are just the Pauli operators labeled in the usual way by bitstrings $\mathbf{a} \in \mathbb{Z}_2^{2n}$, cf. Section 2). Now measure two copies of ψ in the *Bell basis* $|W_\mathbf{x}\rangle$ defined by applying the Weyl operators to the maximally entangled state, i.e., $|W_\mathbf{x}\rangle = (W_\mathbf{x} \otimes I)|\Phi^+\rangle$ where $|\Phi^+\rangle = 2^{-n/2} \sum_\mathbf{q} |\mathbf{q}, \mathbf{q}\rangle$. If ψ is real in the computational basis then it is not hard to see that the measurement outcome is distributed according to the probability distribution $p_\psi(\mathbf{a}) = |c_\mathbf{a}|^2$. This is known as Bell sampling [Mon17, ZPDF16]. Now stabilizer states are distinguished by the fact that they are eigenvectors of all Weyl operators $W_\mathbf{a}$ for which $|c_\mathbf{a}|^2 \neq 0$ (these are its *stabilizer group*). This suggests using Bell sampling to obtain some $\mathbf{a}$, then measuring $W_\mathbf{a}$ twice on two fresh copies, and accepting ψ as a stabilizer if the same eigenvalue is obtained twice.

While we show that this works for real state vectors, Bell sampling unfortunately does not extend to complex state vectors. To overcome this challenge, we introduce a new primitive:

Definition 3.1 (Bell difference sampling). *We define* Bell difference sampling *as performing Bell sampling twice and subtracting (adding) the results from each other (modulo two). In other words, it is the projective measurement on four copies of a state, $\psi^{\otimes 4} \in ((\mathbb{C}^2)^{\otimes n})^{\otimes 4}$, given by*

$$\Pi_\mathbf{a} = \sum_\mathbf{x} |W_\mathbf{x}\rangle \langle W_\mathbf{x}| \otimes |W_{\mathbf{x}+\mathbf{a}}\rangle \langle W_{\mathbf{x}+\mathbf{a}}| .$$

For stabilizer states (whether real or complex) it is easy to see that Bell difference sampling will always sample an element **a** corresponding to a Weyl operator $W_\mathbf{a}$ in its stabilizer group. What
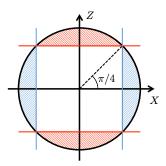
Figure 1: The X-Z-plane of the Bloch sphere. The area shaded in red indicates the projection of those states ρ for which $|\text{tr}\, Z\rho| > \sin\frac{\pi}{4} = \frac{1}{\sqrt{2}}$. Likewise, the blue area correspond to the states with $|\text{tr}\, X\rho| > \frac{1}{\sqrt{2}}$. As the two areas do not intersect, these two conditions cannot be simultaneously satisfied. This is a manifestation of the uncertainty principle.

is rather less obvious is that, even for arbitrary quantum states, Bell difference sampling still has a useful interpretation. The following theorem shows that this is indeed the case: it amounts to sampling from the probability distribution p_ψ twice and taking the difference.

Theorem 3.2 (Bell difference sampling)**.** *Let ψ be an arbitrary pure state of n qubits. Then:*

$$\text{tr}\left[\Pi_{\mathbf{a}}\psi^{\otimes 4}\right] = \sum_{\mathbf{x}} p_\psi(\mathbf{x})p_\psi(\mathbf{x}+\mathbf{a}).$$

If ψ is a stabilizer state, say $|S\rangle\langle S|$, then this is equal to $p_S(\mathbf{a})$ from Eq. (3.3).

Using Bell difference sampling as a primitive, we obtain the natural Algorithm 1.

Theorem 3.3 (Stabilizer testing for qubits)**.** *Let ψ be a pure state of n qubits. If ψ is a stabilizer state then Algorithm 1 accepts with certainty, $p_{accept} = 1$. On the other hand, if $\max_S |\langle S|\psi\rangle|^2 \leqslant 1 - \varepsilon^2$ then $p_{accept} \leqslant 1 - \varepsilon^2/4$.*

Proof sketch. We want to show that if the success probability, p_{accept}, is close to one then ψ has high overlap with a stabilizer state. The proof proceeds in two steps. First, we analyze the success probability and show that if $p_{accept} \approx 1$ then $p_\psi(\mathbf{a})$ is typically close to its maximum possible value 2^{-n}. Next, we use Markov's inequality to find a large set of $\mathbf{a}$ where $p_\psi(\mathbf{a}) > \frac{1}{2}2^{-n}$. Using a version of uncertainty principle (see Fig. 1), we show that the corresponding Weyl operators $W_{\mathbf{a}}$ necessarily commute, and therefore form a stabilizer subgroup. This finally means that our initial state must have a large overlap with a corresponding stabilizer state. □

Theorem 3.3 solves the stabilizer testing conjecture for qubits. It also implies a number of interesting corollaries. E.g., it directly follows that one can also test Cliffordness of a unitary efficiently, without given black-box access to the inverse as in [Low09, Wan11]; this resolves another open problem from [MdW13]. From a structural point of view, it shows that the Clifford group is the solution, within $U(2^n)$, of a set of polynomial equations of order 6. Our result is optimal in the sense that there exist no perfectly complete tests for fewer than six copies that achieve statistical power independent of the number of qubits (see Section 5 and [Dam18]).
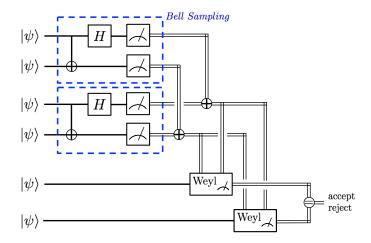
Figure 2: Quantum circuit implementing Algorithm 1 for qubit stabilizer testing. Inside the blue blocks: The quantum gates denote the controlled-NOT and the Hadamard gate, respectively; the measurements are in the n-qubit computational basis. Outside the blue blocks: Double lines represent classical information. The "$\oplus$"-operation is addition modulo two. The boxes labeled "Weyl" perform a two-outcome measurement with respect to the eigenspaces of $W_{\mathbf{a}}$, where $\mathbf{a}$ is determined by classical inputs. For n qubits, the circuit is fully transversal in the sense that all operations are required to be coherent only across two copies, and factorize with respect to the n qubits.

Qudits

A careful analysis of the measurement of Algorithm 1 shows that it is equivalent to a projective measurement of the form $\Pi_{\text{accept}} = \frac{1}{2}\left(I + V\right)$, where V is the following Hermitian unitary operator:

$$V = 2^{-n} \sum_{\mathbf{x}} W_{\mathbf{x}}^{\otimes 6}. \tag{1.5}$$

It can be readily seen that the operator Eq. (1.5) commutes with tensor powers of Clifford unitaries.

In fact, as discussed earlier, it is natural to approach the stabilizer testing problem by measuring operators in the commutant of the Clifford group. Since the stabilizer states are a single orbit under the Clifford group, any such measurement by design will have the same level of significance on all stabilizer states.

Equation (1.5) and corresponding measurement have a clear generalization to arbitrary qudits. Let $d \geqslant 2$ and consider the operators

$$\Pi_{s,\text{accept}} = \frac{1}{2}(I + V_s) \quad \text{where} \quad V_s = d^{-n} \sum_{\mathbf{x}} (W_{\mathbf{x}} \otimes W_{\mathbf{x}}^{\dagger})^{\otimes s}. \tag{1.6}$$

One can see that if $(d, s) = 1$, V_s is a Hermitian unitary and so $\Pi_{s,\text{accept}}$ is a projector. We now state our general stabilizer testing result:

Theorem 3.11 (Stabilizer testing for qudits). *Let $d \geqslant 2$ and choose $s \geqslant 2$ such that $(d, s) = 1$. Let ψ be a pure state of n qudits and denote by $p_{accept} = \text{tr}[\psi^{\otimes 2s}\Pi_{s,accept}]$ the probability that the POVM element $\Pi_{s,accept}$*

7

accepts given 2s *copies of* ψ. *If* ψ *is a stabilizer state then it accepts with certainty,* $p_{accept} = 1$. *On the other hand, if* $\max_S |\langle S|\psi\rangle|^2 \leqslant 1 - \varepsilon^2$ *then* $p_{accept} \leqslant 1 - C_{d,s}\varepsilon^2$, *where* $C_{d,s} = (1 - (1 - 1/4d^2)^{s-1})/2$.

The proof proceeds similarly to the one of Theorem 3.3. Again, an uncertainty relation for Weyl operators plays an important role. We record it since it may be of independent interest:

Lemma 3.10 (Uncertainty relation). *Let* $\delta = 1/2d$ *and* ψ *a pure state such that* $|\mathrm{tr}[\psi W_{\mathbf{x}}]|^2 > 1 - \delta^2$ *and* $|\mathrm{tr}[\psi W_{\mathbf{y}}]|^2 > 1 - \delta^2$. *Then* $W_{\mathbf{x}}$ *and* $W_{\mathbf{y}}$ *must commute.*

We also study the *minimal* number of copies required to distinguish stabilizer states from non-stabilizer states in such a way that the power of the statistical test does not decrease with the number of qubits. Since the stabilizer states share the same second moments with uniformly random states (see Section 1.6 below for more detail), one can see that any such test requires at least three copies. Our next result shows that this is sufficient at least when $d \equiv 1, 5 \pmod 6$. For this, consider the POVM element

$$\Pi_{\mathrm{accept}} = \frac{1}{2}(I + V) \quad \text{where} \quad V := d^{-n}\sum_{\mathbf{x}} A_{\mathbf{x}}^{\otimes 3}.$$

Theorem 8.6 (Stabilizer testing from three copies). *Let* $d \equiv 1, 5 \pmod 6$ *and* ψ *a pure state of* n *qudits. Denote by* $p_{accept} = \mathrm{tr}[\psi^{\otimes 3}\Pi_{accept}]$ *the probability that the POVM element* Π_{accept} *accepts given three copies of* ψ. *If* ψ *is a stabilizer state then it accepts with certainty,* $p_{accept} = 1$. *On the other hand, if* $\max_S |\langle S|\psi\rangle|^2 \leqslant 1 - \varepsilon^2$ *then* $p_{accept} \leqslant 1 - \varepsilon^2/16d^2$.

The operators $A_{\mathbf{x}}$ are known as phase-space point operators [Gro06], which are defined by a (symplectic) Fourier transform of the Weyl operator basis $W_{\mathbf{a}}$ (with respect to the index $\mathbf{a}$). Again, the test corresponds to a particular element of the commutant, and to establish Theorem 8.6 we also need another uncertainty relation, this time for phase-space point operators.

Lemma 8.2. *Let* d *be an odd integer and* ψ *a pure state of* n *qudits. Suppose that* $\mathrm{tr}[\psi A_{\mathbf{x}}], \mathrm{tr}[\psi A_{\mathbf{y}}], \mathrm{tr}[\psi A_{\mathbf{z}}] > \sqrt{1 - 1/2d^2}$. *Then* $[\mathbf{z} - \mathbf{x}, \mathbf{y} - \mathbf{x}] = 0$, *i.e.,* $W_{\mathbf{z}-\mathbf{x}}$ *and* $W_{\mathbf{y}-\mathbf{x}}$ *must commute.*

Lastly, we derive an explicit prescription for the minimal test that is perfectly complete, i.e., detects all stabilizer states with certainty. Here we use the full power of the algebraic theory. We assume that d is a prime.

Definition 4.11 (O_t). *Consider the quadratic form* $q: \mathbb{Z}_d^t \to \mathbb{Z}_D$ *defined by* $q(\mathbf{x}) := \mathbf{x} \cdot \mathbf{x}$.[2] *We define* $O_t(d)$ *as the group of* $t \times t$-*matrices* O *with entries in* $\mathbb{Z}_d$ *that satisfy the following properties:*

1. O *is a* q-*isometry: i.e.,* $O\mathbf{x} \cdot O\mathbf{x} = \mathbf{x} \cdot \mathbf{x} \pmod D$ *for all* $\mathbf{x} \in \mathbb{Z}_d^t$.

2. O *is stochastic:* $O\mathbf{1}_t = \mathbf{1}_t \pmod d$.

We will refer to $O_t(d)$ *as the* stochastic orthogonal group; *its elements will be called* stochastic isometries.

Equivalently, $O_t(d)$ is the group of $t \times t$-matrices O that are orthogonal in the ordinary sense (i.e., $O^\mathsf{T} O = I \bmod d$) and such that the sum of elements in each row is equal to 1 (mod D). See Remark 4.12 for more details.

---

[2] Recall that for $x \in \mathbb{Z}_d$, x^2 is well-defined modulo D.

Note that the subspace $T_O := \{(Oy, y) : y \in \mathbb{Z}_d^t\}$ is a stochastic Lagrangian subspace in $\Sigma_{t,t}(d)$ (as defined above in Definition 4.1), and so we obtain a corresponding operator in the commutant, which we abbreviate by $R(O) = R(T_O)$. It is easy to see that the operators $R(O)$ define a representation of the group $O_t(d)$, so

$$\Pi_t^{\min} := \frac{1}{|O_t(d)|} \sum_{O \in O_t(d)} R(O)$$

is the projector onto the invariant subspace for this action. Remarkably, not only do the $R(O)$ stabilize all stabilizer tensor powers $|S\rangle^{\otimes t}$ (Eq. (4.13)), but $\Pi_t^{\min}$ is in fact the minimal perfectly complete test for stabilizer states:

Theorem 5.6 (Minimal stabilizer test with perfect completeness)**.** *Let d be a prime and $n, t \geqslant 1$. Then the projector $\Pi_t^{\min}$ is the orthogonal projector onto* $\mathrm{span}\{|S\rangle^{\otimes t} : |S\rangle\langle S| \in \mathrm{Stab}(n,d)\}$.

Are there any other tensor power states in the support of $\Pi_t^{\min}$? For every $d \geqslant 2$, we have proved above there exists some $t \geqslant 3$ such that stabilizer testing is possible using t copies. Since the accepting POVM element is in each case the projector onto the invariant subspace of an element in $O_t(d)$ (e.g., the anti-identity for $d = 2$ and $t = 6$), it follows that in this case the only tensor power states contained in the support of $\Pi_t^{\min}$ are tensor powers of stabilizer states!

1.4 De Finetti theorems for stabilizer symmetries

Quantum de Finetti theorems provide versatile tools for the study of correlations in quantum states with permutation symmetry. They have found many important applications, from quantifying the monogamy of entanglement to proving security for quantum key distribution protocols, where de Finetti theorems allow to reduce general attacks to collective attacks [Ren05]. By now, several variants and generalizations are known [Stø69, HM76, RW89, Pet90, CFS02, KR05, DOS07, CKMR07, Ren07, NOP09, KM09, BCY11, BH13, BH17, BCHW16]. Generally speaking, de Finetti theorems state that when ρ is a quantum state on $(\mathbb{C}^\ell)^{\otimes t}$ that commutes with all permutations (i.e., $[r_\pi, \rho] = 0$ for all $\pi \in S_t$, where r_π are the permutation operators defined in Eq. (1.2)) then its reduced density operators $\rho_{1\ldots s} = \mathrm{tr}_{s+1\ldots t}[\rho]$ are well-approximated by convex mixtures of i.i.d. states in some suitable sense if $s \ll t$. E.g., for any such ρ there exists a probability measure $d\mu$ on the space of mixed states on $\mathbb{C}^\ell$ such that [CKMR07],

$$\frac{1}{2} \left\| \rho_{1\ldots s} - \int d\mu(\sigma) \sigma^{\otimes s} \right\|_1 \leqslant 2\frac{\ell^2 s}{t}. \tag{7.1}$$

Using the techniques developed for stabilizer testing, we prove two new versions of the quantum de Finetti theorem *adapted to the symmetries inherent in stabilizer states*. A key insight from the preceding section was that any stabilizer tensor power $|S\rangle^{\otimes t}$ is stabilized not only by the permutations, but by the larger group $O_t(d)$. This group contains includes in general many more elements, for example, the anti-identity (1.4) for the case of qubits. Our de Finetti theorems for stabilizer states show that if we consider *arbitrary* states ρ on $((\mathbb{C}^d)^{\otimes n})^{\otimes t}$ that show symmetries of this kind, then the conclusions of the de Finetti theorem can be strengthened. In this case, the reduced states can be well-approximated by convex mixtures of tensor powers of stabilizer states in $\mathrm{Stab}(n,d)$ (rather than of general pure states in $(\mathbb{C}^d)^{\otimes n}$).

Our first de Finetti theorem shows that the enlarged symmetry provided by the stochastic orthogonal group ensures that the approximation is *exponentially* good in the number of traced-out

subsystems. This is remarkable, since the ordinary permutation symmetry-based de Finetti theorem achieves exponential convergence only if the form of allowed states is relaxed to include "almost product states" [Ren07] or "high weight vectors" (as opposed to *highest* weight vectors) [KM09]. Such a relaxation is, in fact, already necessary for classical distributions [DF80]. In detail:

Theorem 7.6 (Exponential stabilizer de Finetti theorem). *Let* d *be a prime and* ρ *a quantum state on* $((\mathbb{C}^d)^{\otimes n})^{\otimes t}$ *that commutes with the action of* $O_t(d) \supseteq S_t$. *Let* $1 \leqslant s \leqslant t$. *Then there exists a probability distribution* p *on the (finite) set of mixed stabilizer states[3] of* n *qudits, such that*

$$\frac{1}{2}\left\|\rho_{1\ldots s} - \sum_{\sigma_S} p(\sigma_S)\sigma_S^{\otimes s}\right\|_1 \leqslant 2d^{\frac{1}{2}(2n+2)^2}d^{-\frac{1}{2}(t-s)}.$$

Our theorem can be understood as a stabilizer version of the Gaussian Finetti theorems established in [LC09, Lev16]. The latter have been successfully used to establish security of continuous-variable quantum key distribution (QKD) protocols which admit the required symmetries [LGPRC13, Lev17]. Since the input states of entanglement-based QKD schemes [Eke91], are usually taken to be powers of stabilizer states, they show the enlarged symmetry identified here—a fact that seems to have been overlooked so far. It is this natural to study applications of our de Finetti theorems to QKD security proofs—we will report results on this elsewhere.

We can also ask to which extent the conclusions of Theorem 7.6 hold if we only slightly enlarge the symmetry group. The following theorem shows that if we consider quantum states that commute with permutations as well as the anti-identity (but not necessarily other elements of $O_t(d)$) then we still get an approximation by mixtures of stabilizer tensor powers – but now with a polynomially rather than exponentially small error:

Theorem 7.7 (Stabilizer de Finetti theorem for the anti-identity). *Let* ρ *be a quantum state on* $((\mathbb{C}^2)^{\otimes n})^{\otimes t}$ *that commutes with all permutations as well as with the action of the anti-identity* (1.4) *on some (and hence every) subsystem consisting of six* n*-qubit blocks. Let* $s < t$ *be a multiple of six. Then there exists a probability distribution* p *on the (finite) set of mixed stabilizer states of* n *qubits, such that*

$$\frac{1}{2}\left\|\rho_{1\ldots s} - \sum_{\sigma_S} p(\sigma_S)\sigma_S^{\otimes s}\right\|_1 \leqslant 6\sqrt{2}\cdot 2^n\sqrt{\frac{s}{t}}.$$

While Theorem 7.7 is stated here only for qubits, we believe that a similar result can be established in any prime dimension.

1.5 Robust Hudson theorem

Similar techniques can also be applied to pure states with a small amount of negativity in their phase space representation. More precisely, recall that for odd d the *Wigner function* of a quantum state ψ is defined by $w_\psi(\mathbf{x}) = d^{-n}\operatorname{tr}[A_{\mathbf{x}}\psi]$, where the operators $A_{\mathbf{x}}$ are the phase-space point operators mentioned above. The Wigner function is a quasi-probability distribution, i.e., $\sum_{\mathbf{x}} w_\psi(\mathbf{x}) = 1$, but it can be negative. This negativity plays an important role – e.g., it is an obstruction to efficient classical simulability [VFGE12, ME12] and witnesses the onset of contextuality [HWVE14].

---

[3]A mixed stabilizer state is a maximally mixed state on a stabilizer code (see Section 2.4).

In fact, pure stabilizer states are characterized by having a nonnegative Wigner function—this is the discrete Hudson theorem [Gro06]. Our next result shows that this characterization is robust, and that the robustness is independent of the system size (number of qudits). The relevant quantity is the *Wigner* or *sum-negativity* $\mathrm{sn}(\psi) = \sum_{w_\psi(\mathbf{x})<0}|w_\psi(\mathbf{x})|$, i.e., the absolute sum of negative entries of the Wigner function.

Theorem 8.4 (Robust finite-dimensional Hudson theorem). *Let* d *be odd and* ψ *a pure quantum state of* n *qudits. Then there exist a stabilizer state* $|S\rangle$ *such that* $|\langle S|\psi\rangle|^2 \geqslant 1 - 9d^2\,\mathrm{sn}(\psi)$.

Our theorem gives a new quantitative meaning to the sum-negativity, and thereby to the related *mana*, a monotone from the resource theory of stabilizer states [VMGE14] that has attracted increasing attention in the theory of fault-tolerant quantum computation.

1.6 Random stabilizers, higher moments and designs

As mentioned to above, randomized constructions based on the Haar measure are often near-optimal, yet have the drawback that generic quantum states cannot be efficiently prepared. In contrast, random stabilizer states can be efficiently implemented, and early on, it had been discovered that they reproduce the same second moments as the Haar measure! More recently, there had been significant progress on the third and fourth moments [ZKGG16, HWW16, NW16], opening up several many applications where random Clifford unitaries and stabilizer states have successfully replaced the Haar measure [MGE11, HWFW17, KZG16, HNQ$^+$16, NW16]. To go beyond, however, a general understanding of the statistical properties of random stabilizer states is required.

The theory presented in this paper implies general formulas for the t-th moments of stabilizer states. For qudits,

$$\mathbb{E}_{S\ \text{stabilizer}}\left[|S\rangle\langle S|^{\otimes t}\right] = \frac{1}{Z_{n,d,t}}\sum_{T\in\Sigma_{t,t}(d)} R(T), \qquad (5.3)$$

where T ranges precisely over the maximal isotropic stochastic subspaces from Theorem 4.3!

Recall that a *(complex projective)* t-*design* is an ensemble of states $\{p_i, |\psi_i\rangle\}$ such that the average of any polynomial of degree t identically matches the average of the same polynomial with respect to the Haar measure. In other words, a t-design satisfies

$$\mathbb{E}_{i\sim p}\left[(|\psi_i\rangle\langle\psi_i|)^{\otimes t}\right] = \mathbb{E}_{\psi\ \text{Haar}}\left[(|\psi\rangle\langle\psi|)^{\otimes t}\right] = \frac{1}{N'_{d,t}}\sum_{\pi\in S_t} r_\pi, \qquad (1.7)$$

where the right-hand side is the familiar formula for the maximally mixed state on the symmetric subspace in terms of the symmetrizer—an easy consequence of Schur-Weyl duality. When the stabilizer states form a t-design (t $\leqslant$ 3 for qubits, t $\leqslant$ 2 otherwise), Eq. (5.3) reduces to Eq. (1.7). Equation (5.3) unifies and generalizes all previously known results [Zhu15, KG15, Web16, ZKGG16, HWW16, NW16].

Importantly, however, our formula allows us to compute an arbitrary t-th moment even when the stabilizer states deviate significantly from being a t-design. In fact, we demonstrate the power of the formula by using it to establish that following remarkable fact: Even when the stabilizer states (a single Clifford orbit) fail to be a t-design, we can obtain t-designs by taking a finitely many Clifford orbits with appropriately chosen weights:

Theorem 6.2. *Let* d *be a prime and* $n \geqslant t - 1$. *Then there exists an ensemble* $\{p_i, \Psi_i\}_{i=1}^{M_{t,d}}$ *of fiducial states in* $(\mathbb{C}^d)^{\otimes n}$ *such that:*

$$\mathbb{E}_{i \sim p} \mathbb{E}_{U \ Clifford}\left[\left(U |\Psi_i\rangle \langle \Psi_i| U^\dagger\right)^{\otimes t}\right] = \mathbb{E}_{\Psi \ Haar}\left[|\Psi\rangle \langle \Psi|^{\otimes t}\right]$$

That is, the corresponding ensemble of Clifford orbits is a complex projective t-design. Importantly, the number of fiducial states does *not* depend on the number of qudits n.

2 Preliminaries

2.1 Pauli and Clifford group

Let $d \geqslant 2$ be an arbitrary integer. We first consider a single qudit with computational basis vectors $|q\rangle$, where $q \in \{0, \ldots, d-1\}$ or $q \in \mathbb{Z}_d = \mathbb{Z}/d\mathbb{Z}$. We define unitary shift and boost operators

$$X |q\rangle = |q + 1\rangle, \quad Z |q\rangle = \omega^q |q\rangle,$$

where $\omega = e^{2\pi i q/d}$.

The algebra of shift and boost operators differs slightly depending on whether d is even or odd. For uniform treatment, one introduces $\tau = (-1)^d e^{i\pi/d} = e^{i\pi(d^2+1)/d}$. Note that $\tau^2 = \omega$. Let D denote the order of τ. Then $D = 2d$ if d is even, but $D = d$ if d is odd (indeed, in this case $\tau = \omega^{2^{-1}}$, where 2^{-1} denotes the multiplicative inverse of 2 mod d). Then $Y := \tau X^\dagger Z^\dagger$ is such that $XYZ = \tau I$, generalizing the commutation relation of the usual Pauli operators for qubits (where $\tau = i$). For a single qudit, the Pauli group is generated by X, Y, Z or, equivalently, by $\tau I, X, Z$.

For n qudits, the Hilbert space is $\mathcal{H}_n = (\mathbb{C}^d)^{\otimes n}$, with computational basis vectors $|\mathbf{q}\rangle = |q_1, \ldots, q_n\rangle$, and the *Pauli group* $\mathcal{P}_n$ is the group generated by the tensor product of I, X, Y, Z acting on each of the n qudits.

The Clifford group $\mathrm{Cliff}(n, d)$ is defined as normalizer of the Pauli group in the unitary group, modulo phases. That is, it consists of all unitary operators U that $U\mathcal{P}_n U^\dagger \subseteq \mathcal{P}_n$, up to phases. For qubits, the Clifford group is generated by the phase gate $P = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$, the Hadamard gate $H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, and the controlled-NOT gate.

2.2 Weyl operators and characteristic function

At this point it is useful to recall the phase space picture of finite-dimensional quantum mechanics developed in [Woo87, App05, Gro06, GE08, DB13], which is analogous to the phase space formalism for continuous-variable systems used, e.g., in quantum optics [Sch11]. For $\mathbf{x} = (\mathbf{p}, \mathbf{q}) \in \mathbb{Z}^{2n}$, define the *Weyl operator*

$$W_\mathbf{x} = W_{\mathbf{p},\mathbf{q}} = \tau^{-\mathbf{p}\cdot\mathbf{q}}(Z^{p_1} X^{q_1}) \otimes \cdots \otimes (Z^{p_n} X^{q_n}). \tag{2.1}$$

Clearly, each Weyl operator is an element of the Pauli group. Conversely, each element of the Pauli group is equal to a Weyl operator up to a phase that is a power of τ. It is not hard to see that the Weyl operators themselves only depend on $\mathbf{x}$ modulo D (which we recall is $2d$ or d, depending on whether d is even or odd). Indeed,

$$W_{\mathbf{x}+d\mathbf{z}} = (-1)^{(d+1)[\mathbf{x},\mathbf{z}]} W_\mathbf{x}, \tag{2.2}$$

where we have introduced the $\mathbb{Z}$-valued *symplectic form* on $\mathbb{Z}^{2n}$

$$[\mathbf{x}, \mathbf{y}] = [(\mathbf{p}, \mathbf{q}), (\mathbf{p}', \mathbf{q}')] = \mathbf{p} \cdot \mathbf{q}' - \mathbf{q} \cdot \mathbf{p}'. \tag{2.3}$$

We will often use the symplectic form in situations where $\mathbf{x}, \mathbf{y}$ are elements of $\mathbb{Z}_d^{2n}$ or $\mathbb{Z}_D^{2n}$, and interpret $[\mathbf{x}, \mathbf{y}]$ accordingly. For example,

$$W_\mathbf{x} W_\mathbf{y} = \tau^{[\mathbf{x},\mathbf{y}]} W_{\mathbf{x}+\mathbf{y}} \tag{2.4}$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{Z}_D^{2n}$. This implies that in particular

$$W_\mathbf{x} W_\mathbf{y} = \omega^{[\mathbf{x},\mathbf{y}]} W_\mathbf{y} W_\mathbf{x}. \tag{2.5}$$

Thus the commutation relations between Weyl operators only depend on $\mathbf{x}, \mathbf{y}$ mod d. In this sense, $\mathcal{V}_n = \{0, \ldots, d-1\}^{2n}$ is the natural classical phase space associated with the Hilbert space $\mathcal{H}_n = (\mathbb{C}^d)^{\otimes n}$. We will often write $W_\mathbf{x}$ for $\mathbf{x} \in \mathbb{Z}_d^{2n}$, identifying $\mathbb{Z}_d^{2n} \cong \mathcal{V}_n$ in the standard way.

Note that $\mathrm{tr}[W_\mathbf{x}] \neq 0$ if and only if $W_\mathbf{x}$ is a scalar multiple of the identity (necessarily $\pm I$), that is, if and only if $\mathbf{x} \equiv \mathbf{0} \pmod{d}$. Together with Eq. (2.4), it follows that the re-scaled Weyl operators $\{d^{-n/2} W_\mathbf{x}\}$ for $\mathbf{x} \in V_n$ form an orthonormal basis with respect to the Hilbert-Schmidt inner product $\langle A, B \rangle = \mathrm{tr}[A^\dagger B]$. In particular, any operator B on $\mathcal{H}_n$ can be expanded in the form $B = d^{-n/2} \sum_{\mathbf{x} \in \mathcal{V}_n} c_B(\mathbf{x}) W_\mathbf{x}$. The expansion coefficients $c_B(\mathbf{x})$ together define the *characteristic function* $c_B : \mathcal{V}_n \to \mathbb{C}$ of the operator B,

$$c_B(\mathbf{x}) = d^{-n/2} \mathrm{tr}[W_\mathbf{x}^\dagger B], \tag{2.6}$$

and we have Parseval's identity

$$\mathrm{tr}[A^\dagger B] = \sum_{\mathbf{x} \in \mathcal{V}_n} \overline{c}_A(\mathbf{x}) c_B(\mathbf{x}). \tag{2.7}$$

By definition, if U is a Clifford unitary then, for every $\mathbf{x} \in \mathcal{V}_n$, $U W_\mathbf{x} U^\dagger$ is proportional to a Weyl operator $W_{\mathbf{x}'}$, where we can take $\mathbf{x}' \in \mathcal{V}_n$ in view of Eq. (2.2). Since conjugation preserves the commutation relations, this action has substantially more structure. In particular, the mapping $\mathbf{x} \mapsto \mathbf{x}'$ is implemented by an element of the symplectic group $\mathrm{Sp}(2n, d)$, i.e., a linear transformation of $\mathbb{Z}_d^{2n}$ that preserves the symplectic form (2.3). The following facts are well-known in the literature (e.g. [App05, Gro06, DB13, Zhu15]).

Lemma 2.1. *For any prime* d *and any* $n \in \mathbb{N}$, *the following holds:*

1. *For each* $U \in \mathrm{Cliff}(n, d)$, *there is a* $\Gamma \in \mathrm{Sp}(2n, d)$ *and a function* $f \colon \mathbb{Z}_d^{2n} \to \mathbb{Z}_d$ *such that*

$$U W_\mathbf{x} U^\dagger = \omega^{f(\mathbf{x})} W_{\Gamma\mathbf{x}} \qquad \forall \mathbf{x} \in \mathbb{Z}_d^{2n}. \tag{2.8}$$

2. *Conversely, for each* $\Gamma \in \mathrm{Sp}(2n, d)$, *there is a* $U \in \mathrm{Cliff}(n, d)$ *and a phase function* $f \colon \mathbb{Z}_d^{2n} \to \mathbb{Z}_d$ *such that Eq.* (2.8) *holds. If* d *is odd, one can choose* U *such that* $f \equiv 0$.

3. *The quotient of the Clifford group by Weyl operators and phases is isomorphic to* $\mathrm{Sp}(2n, d)$.

Below, we will frequently assume that a correspondence $\Gamma \mapsto U_\Gamma$ has been fixed.

2.3 Wigner function and phase space point operators

It is also useful to consider the symplectic Fourier transform, which for any function $f\colon \mathcal{V}_n \to \mathbb{C}$ is defined as

$$\hat{f}(\mathbf{x}) = d^{-n} \sum_{\mathbf{y}} \omega^{-[\mathbf{x},\mathbf{y}]} f(\mathbf{y}). \tag{2.9}$$

The transformation $f \mapsto \hat{f}$ is unitary, i.e., we have Parseval's identity: $\sum_{\mathbf{x}} \overline{\hat{f}(\mathbf{x})}\hat{g}(\mathbf{x}) = \sum_{\mathbf{y}} \overline{f(\mathbf{y})}g(\mathbf{y})$.

The Fourier transform of the characteristic function is (up to normalization) known as the *Wigner function* [Woo87] $w_B\colon \mathcal{V}_n \to \mathbb{C}$, defined by

$$w_B(\mathbf{x}) = d^{-n/2}\hat{c}_B(\mathbf{x}) = d^{-3n/2} \sum_{\mathbf{y}} \omega^{-[\mathbf{x},\mathbf{y}]} c_B(\mathbf{y}) = d^{-2n} \sum_{\mathbf{y}} \omega^{-[\mathbf{x},\mathbf{y}]} \operatorname{tr}[W_{\mathbf{y}}^{\dagger} B] = d^{-n} \operatorname{tr}[A_{\mathbf{x}} B], \tag{2.10}$$

where we have introduced the *phase-space point operators*

$$A_{\mathbf{x}} = d^{-n} \sum_{\mathbf{y}} \omega^{-[\mathbf{x},\mathbf{y}]} W_{\mathbf{y}}^{\dagger}. \tag{2.11}$$

The operators $\{A_{\mathbf{x}}\}$ form an orthogonal basis of the space of all operators, $\operatorname{tr}[A_{\mathbf{x}}^{\dagger} A_{\mathbf{y}}] = d^n \delta_{\mathbf{x},\mathbf{y}}$, so the Wigner function can be seen as the set of coefficients of an operator as expanded in this basis, $B = \sum_{\mathbf{x}} w_B(\mathbf{x}) A_{\mathbf{x}}^{\dagger}$. Moreover, the Wigner function of a quantum state is a quasiprobability distribution in the sense that $\sum_{\mathbf{x}} w_\rho(\mathbf{x}) = 1$.

For odd d the Wigner function is particularly well-behaved. For one, the phase-space point operators are Hermitian (this is also true for qubits) and they square to the identity (so the eigenvalues are ± 1 and in particular $\|A_{\mathbf{x}}\| = 1$). This means that the Wigner function of a quantum state is real and $-d^{-n} \leqslant w_\psi(\mathbf{x}) \leqslant d^{-n}$. The phase-space point operators satisfy the following important identity:

$$A_{\mathbf{x}} A_{\mathbf{y}} A_{\mathbf{z}} = \omega^{2[\mathbf{z}-\mathbf{x},\mathbf{y}-\mathbf{x}]} A_{\mathbf{x}-\mathbf{y}+\mathbf{z}} \tag{2.12}$$

Moreover, (only) for odd d does the Wigner function transforms covariantly with respect to the Clifford group. Here, the Clifford operators can (up to overall phase) be parametrized by an affine symplectic transformation, i.e., by a symplectic matrix $\Gamma \in \operatorname{Sp}(2n, d)$ and a vector $\mathbf{b} \in \mathbb{Z}_d^{2n}$. Then $U = W_{\mathbf{v}}\,\mu_\Gamma$ is in $\operatorname{Cliff}(n, d)$, where μ_Γ is the so-called *metaplectic* representation of the symplectic group (see, e.g., [Gro06]), and the conjugation action of U on phase-space point operators is given by

$$UA_{\mathbf{x}}U^{\dagger} = A_{\Gamma\mathbf{x}+\mathbf{b}}. \tag{2.13}$$

In particular, the Weyl operators induce a translation in phase space.

2.4 Stabilizer groups, codes, and states

We now give uniform account of the stabilizer formalism [Got97, Got99] for qudits. Stabilizer states are commonly defined in terms of the Pauli group in the following way: Consider a subgroup of the Pauli group $S \subseteq \mathcal{P}_n$ that does not contain any (nontrivial) multiple of the identity operator. Then the operator

$$P_S = \frac{1}{|S|} \sum_{P \in S} P \tag{2.14}$$

14

is an orthogonal projection onto a subspace $V_S \subseteq \mathcal{H}_n$ of dimension $d^n/|S|$. We say that V_S is the *stabilizer code* associated with the *stabilizer group* S. If $|S| = d^n$ then this code is spanned by a single state, called a *(pure) stabilizer state* and denoted by $|S\rangle\langle S|$. It is given precisely by Eq. (2.14). In other words, a stabilizer state $|S\rangle$ is the unique $+1$ eigenvector (up to scalars) of all the Pauli operators in S,

$$P|S\rangle = |S\rangle \quad (\forall P \in S).$$

In the following we will mostly be talking about stabilizer groups that determine a pure state. We denote the (finite) set of pure stabilizer states in $(\mathbb{C}^d)^{\otimes n}$ by $\mathrm{Stab}(n, d)$.

In order to connect the stabilizer formalism to the phase space picture, we observe that the stabilizer group can be written in the form

$$S = \{\omega^{f(\mathbf{x})} W_{\mathbf{x}} : \mathbf{x} \in M\}, \tag{2.15}$$

for some subset $M \subseteq \mathcal{V}_n$ and some function $f \colon M \to \mathbb{Z}_d$. The two pieces of data determine the stabilizer state uniquely. Indeed, $|S\rangle$ can be characterized by demanding that

$$W_{\mathbf{x}}|S\rangle = \omega^{-f(\mathbf{x})}|S\rangle \quad (\forall P \in S).$$

Moreover, it is not hard to verify that M is closed under addition (because S is a group) and that $[\mathbf{x}, \mathbf{y}] = 0$ for any two elements $\mathbf{x}, \mathbf{y} \in M$. Thus, M is a totally isotropic submodule of the phase space $\mathcal{V}_n$ (which itself can be thought of as a $\mathbb{Z}_d$-module). For simplicity, we will usually say *subspace* instead of submodule, although the latter terminology is more appropriate for non-prime d. Moreover, $|M| = d^n$, which is the maximal possible cardinality of any such subspace – one often says that M is a Lagrangian subspace and it holds that $M = M^\perp$, where $M^\perp = \{\mathbf{y} \in \mathcal{V}_n | [\mathbf{x}, \mathbf{y}] = 0 \ \forall \mathbf{x} \in M\}$. See, e.g., [Gro06, GW13] for further detail on this symplectic point of view.

Conversely, suppose that M is a Lagrangian subspace of $\mathcal{V}_n$. Then there always exist functions f such that $\{\omega^{f(\mathbf{x})} W_{\mathbf{x}}\}_{\mathbf{x} \in M}$ is a stabilizer group; we will denote the corresponding stabilizer states by $|M, f\rangle$. Any other such function f can be obtained by replacing f by $g = f + \delta$, where $\delta \colon M \to \mathbb{Z}_d$ is a $\mathbb{Z}_d$-linear function. We can always write $\delta(\mathbf{x}) = [\mathbf{z}, \mathbf{x}]$; then $|M, g\rangle = W_{\mathbf{z}}|M, f\rangle$. In this way, M parametrizes an orthonormal basis of $\mathcal{H}_n$ worth of stabilizer states. In particular, any state that is a simultaneous eigenvector of the $\{W_{\mathbf{x}}\}_{\mathbf{x} \in M}$ is necessarily a stabilizer state. It is not hard to verify that the quantum channel that implements the projective measurement in this *stabilizer basis* $\{|M, f\rangle\}_f$ is given by

$$\Lambda_M[\rho] = \sum_f |M, f\rangle\langle M, f|\rho|M, f\rangle\langle M, f| = d^{-n} \sum_{\mathbf{x} \in M} W_{\mathbf{x}} \rho W_{\mathbf{x}}^\dagger. \tag{2.16}$$

The fact that any stabilizer state can be parametrized as $|S\rangle = |M, f\rangle$ will be of fundamental importance to our investigations. As a first consequence, we note that Eqs. (2.14) and (2.15) imply that $|S\rangle\langle S| = d^{-n} \sum_{\mathbf{x} \in M} \omega^{f(\mathbf{x})} W_{\mathbf{x}}$. This shows that the characteristic function is given by

$$c_S(\mathbf{x}) = \begin{cases} d^{-n/2}\omega^{f(\mathbf{x})} & \text{if } \mathbf{x} \in M, \\ 0 & \text{otherwise,} \end{cases} \tag{2.17}$$

i.e., it is supported precisely on the set M.

For odd d the phase is a linear function, so it can be written as $f(\mathbf{x}) = [\mathbf{a}, \mathbf{x}]$ for some suitable vector $\mathbf{a}$ (e.g., [Gro06, App. C]). This means that the Wigner functions of stabilizer states have the

following form [GW13]:

$$w_S(\mathbf{x}) = d^{-3n/2} \sum_{\mathbf{y} \in M} \omega^{-[\mathbf{x},\mathbf{y}]} d^{-n/2} \omega^{[\mathbf{a},\mathbf{y}]} = \begin{cases} d^{-n} & \text{if } \mathbf{x} \in \mathbf{a} + M, \\ 0 & \text{otherwise} \end{cases} \tag{2.18}$$

(using that $M = M^{\perp}$ for a pure stabilizer state). In particular, the Wigner function is non-negative. The finite-dimensional Hudson theorem asserts that, for pure states, the converse is also true [Gro06]. In Section 8 we will prove a robust version of this result.

3 Testing stabilizer states

Given two copies of an unknown pure state $\psi = |\psi\rangle\langle\psi|$ on $\mathcal{H}_n$, it is easy to verify using phase estimation whether $|\psi\rangle$ is an eigenvector of a given Weyl operator $W_{\mathbf{x}}$. In particular, if $W_{\mathbf{x}}$ is Hermitian then we simply measure twice and compare the result. The probability of obtaining the same outcome is

$$\text{tr}\left[\psi^{\otimes 2} \frac{I + W_{\mathbf{x}} \otimes W_{\mathbf{x}}^{\dagger}}{2}\right] = \frac{1}{2}\left(1 + d^n |c_{\psi}(\mathbf{x})|^2\right), \tag{3.1}$$

where we recall that c_{ψ} denotes the characteristic function defined in Eq. (2.6).

To turn this idea into an algorithm for testing whether ψ is a stabilizer state we need a way of generating good candidate Weyl operators. For this we note that, since ψ is a *pure* quantum state,

$$p_{\psi}(\mathbf{x}) = |c_{\psi}(\mathbf{x})|^2 = d^{-n} |\langle\psi|W_{\mathbf{x}}|\psi\rangle|^2 = d^{-n} \text{tr}[\psi W_{\mathbf{x}} \psi W_{\mathbf{x}}^{\dagger}] \tag{3.2}$$

is a probability distribution on the phase space $\mathcal{V}_n$. This follows directly from Eq. (2.7). We call p_{ψ} the *characteristic distribution* of ψ.

Now, if $|\psi\rangle = |S\rangle = |M, f\rangle$ is a stabilizer state then Eq. (2.17) implies that p_{ψ} is simply the uniform distribution on the subset $M \subseteq \mathcal{V}_n$:

$$p_S(\mathbf{x}) = \begin{cases} d^{-n} & \text{if } \mathbf{x} \in M, \\ 0 & \text{otherwise.} \end{cases} \tag{3.3}$$

Note that p_S is maximally sparse in the case of pure stabilizer states, since it always holds true that $0 \leqslant p_{\psi}(\mathbf{x}) \leqslant d^{-n}$. Therefore, if we sample from the characteristic distribution of a stabilizer state then Eq. (3.3) shows that we would with certainty obtain the label of a Weyl operator for which $|\psi\rangle$ is an eigenvector.

Importantly, the converse of this statement is also true: Suppose that $|\psi\rangle$ is an eigenvector of all Weyl operators $W_{\mathbf{x}}$ for $\mathbf{x}$ in the support of the characteristic distribution (i.e., $p_{\psi}(\mathbf{x}) > 0$). Since $p_{\psi}(\mathbf{x}) \leqslant d^{-n}$, the support of p_{ψ} contains at least d^n points. Thus if $|\psi\rangle$ is an eigenvector of all these Weyl operators then the support must be exactly of cardinality d^n and so $|\psi\rangle$ is a stabilizer state. This suggests the following algorithm:

1. Sample from the characteristic distribution of ψ. Denote the result $\mathbf{x}$.

2. Measure the corresponding Weyl operator $W_{\mathbf{x}}$ twice and accept if the result is the same.

By the preceding discussion, this test will accept if and only if the state is a stabilizer state. But how do we go about sampling from the characteristic distribution?

3.1 Qubit stabilizer testing and Bell difference sampling

When the wave function $|\psi\rangle$ is real in the computational basis then sampling from the characteristic distribution can be achieved by *Bell sampling*, introduced for qubits in [Mon17] (cf. [ZPDF16]). Bell sampling amounts to performing a basis measurement in the basis obtained by applying the Weyl operators to a fixed maximally entangled state, $|W_{\mathbf{x}}\rangle = (W_{\mathbf{x}} \otimes I) |\Phi^+\rangle$. Since the Weyl operators are orthogonal, $|W_{\mathbf{x}}\rangle$ is an orthonormal basis of the doubled Hilbert space $\mathcal{H}_n \otimes \mathcal{H}_n$. Using the transpose trick,

$$\left| \langle W_{\mathbf{x}} | \psi^{\otimes 2} \rangle \right|^2 = d^{-n} |\langle \psi | W_{\mathbf{x}} | \bar{\psi} \rangle|^2 \tag{3.4}$$

In case the wave function is real, Eq. (3.4) is exactly equal to $p_\psi(\mathbf{x})$; Bell sampling therefore allows us to implement step (1) above given two copies of the unknown state ψ.

In general, however, the transformation $\psi \mapsto \bar{\psi} = \psi^\mathsf{T}$ *cannot* be implemented by a physical process, since the transpose map is well-known not to be completely positive. Thus we need a new idea to treat the general case where the wave function can be complex.

We start with the observation that if ψ is a stabilizer state then so is $\bar{\psi}$. Indeed, $\overline{W_{\mathbf{p},\mathbf{q}}} = (-1)^{(d+1)(\mathbf{p} \cdot \mathbf{q})} W_{J(\mathbf{p},\mathbf{q})}$, where J is the involution [App05]

$$J \colon \mathcal{V}_n \to \mathcal{V}_n, \quad (p, q) \mapsto ((-p) \bmod d, q)$$

on phase space (note that the phase is trivial when d is odd and so always well-defined mod d). On the other hand, $\overline{\omega^{f(\mathbf{x})}} = \omega^{-f(\mathbf{x})}$. It follows that if $|\psi\rangle = |M, f\rangle$ then $|\bar{\psi}\rangle = |J(M), g\rangle$, where $\omega^{g(\mathbf{x})} = \omega^{-f(\mathbf{x})}(-1)^{(d+1)(\mathbf{p} \cdot \mathbf{q})}$ (again, this is well-defined for any d).

For qubits ($d = 2$), the involution J is trivial. This means that if ψ is a stabilizer state then ψ and $\bar{\psi}$ are characterized by the same subspace M, but possibly different phases. We saw above that (only) in this case there exists a Weyl operator $W_{\mathbf{z}}$ such that $|\bar{\psi}\rangle = W_{\mathbf{z}} |\psi\rangle$. As a consequence, if we perform Bell sampling on $|\psi\rangle \otimes |\psi\rangle$ then, from Eq. (3.4),

$$\left| \langle W_{\mathbf{x}} | \psi^{\otimes 2} \rangle \right|^2 = d^{-n} |\langle \psi | W_{\mathbf{x}+\mathbf{z}} | \psi \rangle|^2 = p_\psi(\mathbf{x} + \mathbf{z}).$$

Of course, $\mathbf{z}$ is an unknown vector that depends on the stabilizer state ψ. But since $\mathbf{z}$ depends *only* on the stabilizer state ψ, it is clear that we may Bell sample *twice* and take the difference of the result in order to obtain a uniform sample $\mathbf{a}$ from the subspace M. Formally:

Definition 3.1 (Bell difference sampling). *We define* Bell difference sampling *as performing Bell sampling twice and subtracting (adding) the results from each other (modulo two). In other words, it is the projective measurement on four copies of a state, $\psi^{\otimes 4} \in ((\mathbb{C}^2)^{\otimes n})^{\otimes 4}$, given by*

$$\Pi_{\mathbf{a}} = \sum_{\mathbf{x}} |W_{\mathbf{x}}\rangle \langle W_{\mathbf{x}}| \otimes |W_{\mathbf{x}+\mathbf{a}}\rangle \langle W_{\mathbf{x}+\mathbf{a}}|.$$

It is not obvious that Bell difference sampling should be meaningful for non-stabilizer quantum states ψ. The following theorem shows that it has a natural interpretation for general states:

Theorem 3.2 (Bell difference sampling). *Let ψ be an arbitrary pure state of n qubits. Then:*

$$\mathrm{tr}\left[\Pi_{\mathbf{a}} \psi^{\otimes 4} \right] = \sum_{\mathbf{x}} p_\psi(\mathbf{x}) p_\psi(\mathbf{x} + \mathbf{a}).$$

If ψ is a stabilizer state, say $|S\rangle \langle S|$, then this is equal to $p_S(\mathbf{a})$ from Eq. (3.3).

17

The proof of Theorem 3.2 uses the symplectic Fourier transform defined in Eq. (2.9). Remarkably, the characteristic *distribution* of any pure state is left invariant by the Fourier transform:

$$\hat{p}_\psi(\mathbf{a}) = 2^{-n} \sum_{\mathbf{x}} (-1)^{[\mathbf{a},\mathbf{x}]} c_\psi(\mathbf{x}) c_\psi(\mathbf{x}) = 2^{-n} \sum_{\mathbf{x}} c_\psi(\mathbf{x}) c_{W_\mathbf{a} \psi W_\mathbf{a}}(\mathbf{x}) \tag{3.5}$$
$$= 2^{-n} \operatorname{tr}[\psi W_\mathbf{a} \psi W_\mathbf{a}] = p_\psi(\mathbf{a}),$$

where the third step is Eq. (2.7) (note that for qubits the characteristic function is real).

We now give the proof of Theorem 3.2:

Proof of Theorem 3.2. We start with the observation that $\Pi_\mathbf{a} = (I \otimes I \otimes I \otimes W_\mathbf{a}) \Pi_\mathbf{0} (I \otimes I \otimes I \otimes W_\mathbf{a})$. On the other hand, it is easy to verify that

$$\Pi_\mathbf{0} = \frac{1}{2^{2n}} \sum_{\mathbf{x}} W_\mathbf{x}^{\otimes 4} \tag{3.6}$$

(i.e., it is the projection onto a stabilizer code of dimension 2^{2n}, which played an important role in [ZKGG16], and Bell difference sampling achieves precisely the syndrome measurement for this code). It follows that

$$\Pi_\mathbf{a} = \frac{1}{2^{2n}} \sum_{\mathbf{x}} (-1)^{[\mathbf{a},\mathbf{x}]} W_\mathbf{x}^{\otimes 4} \tag{3.7}$$

and so

$$\operatorname{tr}\left[\Pi_\mathbf{a} \psi^{\otimes 4}\right] = \frac{1}{2^{2n}} \sum_{\mathbf{x}} (-1)^{[\mathbf{a},\mathbf{x}]} \operatorname{tr}\left[W_\mathbf{x}^{\otimes 4} \psi^{\otimes 4}\right] = \sum_{\mathbf{x}} (-1)^{[\mathbf{a},\mathbf{x}]} p_\psi(\mathbf{x}) p_\psi(\mathbf{x})$$
$$= \sum_{\mathbf{x}} \hat{p}_\psi(\mathbf{x}) \hat{p}_\psi(\mathbf{x} + \mathbf{a}) = \sum_{\mathbf{x}} p_\psi(\mathbf{x}) p_\psi(\mathbf{x} + \mathbf{a});$$

the third equality is the unitarity of the Fourier transform, which also maps modulations to translations, and in the last step we used Eq. (3.5), namely that the characteristic distribution of a pure state is left invariant by the Fourier transform. □

Theorem 3.2 motivates Algorithm 1 as a natural algorithm for testing whether a multi-qubit state is a stabilizer state. The following theorem shows that stabilizer states are the only states that are accepted with certainty, and it quantifies this observation in a dimension-independent way:

Theorem 3.3 (Stabilizer testing for qubits). *Let ψ be a pure state of n qubits. If ψ is a stabilizer state then Algorithm 1 accepts with certainty, $p_{accept} = 1$. On the other hand, if $\max_S |\langle S|\psi\rangle|^2 \leqslant 1 - \varepsilon^2$ then $p_{accept} \leqslant 1 - \varepsilon^2/4$.*

The converse bound of Theorem 3.3 can be stated equivalently as

$$\max_S |\langle S|\psi\rangle|^2 \geqslant 4 p_{accept} - 3. \tag{3.8}$$

Proof. According to Theorem 3.2, step 1 of the algorithm samples elements $\mathbf{a}$ with probability $q(\mathbf{a}) = \sum_{\mathbf{x}} p_\psi(\mathbf{x}) p_\psi(\mathbf{x} + \mathbf{a})$.

Let us first discuss the case that ψ is a stabilizer state, say $|\psi\rangle = |M, f\rangle$. Since $p_\psi(\mathbf{x})$ is the uniform distribution over M, which is a *subspace*, it holds that $q(\mathbf{a}) = p_\psi(\mathbf{a})$, since, for $x \in M$, $x + a \in M$ if

and only if $a \in M$. But this means that $\mathbf{a} \in M$ with certainty. Thus, $|\psi\rangle$ is an eigenvector of the corresponding Weyl operator $W_{\mathbf{a}}$ and step 2 of the test always accepts.

We now consider the case that ψ is a general pure state. Our goal will be to show that if Algorithm 1 succeeds with high probability then there must exist a stabilizer state with high overlap with ψ. According to Eq. (3.1), the probability of acceptance is given by

$$p_{\text{accept}} = \frac{1}{2} \sum_{\mathbf{a}} q(\mathbf{a}) \left(1 + 2^n p_\psi(\mathbf{a}) \right)$$

where we recall that $q(\mathbf{a}) = \sum_{\mathbf{x}} p_\psi(\mathbf{x}) p_\psi(\mathbf{x} + \mathbf{a})$. Thus, by the Cauchy-Schwarz inequality,

$$
\begin{aligned}
p_{\text{accept}} &= \frac{1}{2} \sum_{\mathbf{x}} p_\psi(\mathbf{x}) \left(1 + 2^n \sum_{\mathbf{a}} p_\psi(\mathbf{x}+\mathbf{a}) p_\psi(\mathbf{a}) \right) \leqslant \frac{1}{2} \sum_{\mathbf{x}} p_\psi(\mathbf{x}) \left(1 + 2^n \sum_{\mathbf{a}} p_\psi(\mathbf{a})^2 \right) \\
&= \frac{1}{2} \left(1 + 2^n \sum_{\mathbf{a}} p_\psi(\mathbf{a})^2 \right) = \frac{1}{2} \sum_{\mathbf{a}} p_\psi(\mathbf{a}) \left(1 + 2^n p_\psi(\mathbf{a}) \right),
\end{aligned}
\tag{3.9}
$$

where we have also used the fact that p_ψ is a probability distribution. Intuitively, this bound shows that if our test accepts with high probability then $p_\psi(\mathbf{a}) \approx 2^{-n}$ with high probability. Indeed, let us consider

$$M_0 := \{ \mathbf{a} \in \mathcal{V}_n : 2^n p_\psi(\mathbf{a}) > 1/2 \}.$$

Then Markov's inequality (which can be applied since it is always true that $p_\psi \leqslant 2^{-n}$) asserts that

$$\sum_{\mathbf{a} \in M_0} p_\psi(\mathbf{a}) \geqslant 1 - 2 \sum_{\mathbf{a}} p_\psi(\mathbf{a}) \left(1 - 2^n p_\psi(\mathbf{a}) \right) = 1 - 4 \left(1 - p_{\text{accept}} \right). \tag{3.10}$$

The choice of threshold $1/2$ in the definition of M_0 ensures that the Weyl operators corresponding to any two points $\mathbf{a}, \mathbf{b} \in M_0$ commute. To see, we use that any pair of *anti*commuting $W_{\mathbf{a}}, W_{\mathbf{b}}$ can by a base change be mapped onto the Pauli operators X, Z; it can then verified on the Bloch sphere that there exists no qubit state ρ such that both $\text{tr}[\rho X]^2 > 1/2$ and $\text{tr}[\rho Z]^2 > 1/2$ (see Fig. 1 for a graphical proof).

Let us now extend the set M_0 to some maximal set M such that the corresponding Weyl operators commute. Then M is automatically a Lagrangian subspace, of dimension n.[4] As discussed in Section 2.4, it determines a whole basis of stabilizer states, $\{|M, f\rangle\}_f$. Thus:

$$
\begin{aligned}
\max_S |\langle S|\psi\rangle|^2 &\geqslant \max_f \langle M, f|\psi|M, f\rangle \geqslant \sum_f \langle M, f|\psi|M, f\rangle^2 = \text{tr}\left[\Lambda_M[\psi]^2 \right] \\
&= 2^{-2n} \sum_{\mathbf{x}, \mathbf{y} \in M} \text{tr}\left[\psi W_{\mathbf{x}}^\dagger W_{\mathbf{y}} \psi (W_{\mathbf{x}}^\dagger W_{\mathbf{y}})^\dagger \right] = 2^{-n} \sum_{\mathbf{x} \in M} \text{tr}\left[\psi W_{\mathbf{x}} \psi W_{\mathbf{x}}^\dagger \right] \\
&= \sum_{\mathbf{x} \in M} p_\psi(\mathbf{x}) \geqslant \sum_{\mathbf{x} \in M_0} p_\psi(\mathbf{x}) \geqslant 1 - 4 \left(1 - p_{\text{accept}} \right)
\end{aligned}
$$

where we used Eq. (2.16) for the measurement Λ_M in the stabilizer basis; the last bound is Eq. (3.10). In particular, if $\max_S |\langle S|\psi\rangle|^2 \leqslant 1 - \varepsilon^2$ then $p_{\text{accept}} \leqslant 1 - \varepsilon^2/4$. $\qquad\square$

---

[4]It is natural to ask whether the subspace M is uniquely determined by M_0. This is the case when, e.g., $p_{\text{accept}} > 7/8$. Indeed, in this case, Eq. (3.10) implies that $2^{-n}|M_0| \geqslant \sum_{\mathbf{a} \in M_0} p_\psi(\mathbf{a}) > 1 - 4(1 - 7/8) = 1/2$, so $|M_0| > 2^{n-1}$. It follows that M_0 spans an n-dimensional subspace which is necessarily contained in, and hence equal to, M.

19

Our theorem has the following consequence for quantum property testing, resolving an open question first raised by Montanaro and de Wolf [MdW13, Question 7].

Corollary 3.4. *Let ψ be a pure state of n qubits and let $\varepsilon > 0$. Then there exists a quantum algorithm that, given $O(1/\varepsilon^2)$ copies of ψ, accepts any stabilizer state (it is* perfectly complete*), while it rejects states such that $\max_S |\langle S|\psi\rangle|^2 \leqslant 1 - \varepsilon^2$ with probability at least 2/3.*

Before our result, the best known algorithms required a number of copies that scaled linearly with n, the number of qubits. Indeed, these algorithms proceeded by attempting to *identify* the stabilizer state, which requires $\Omega(n)$ copies by the Holevo bound [AG08, Mon17, ZPDF16]. Moreover, our algorithm is manifestly efficient (see the circuit in Fig. 2).

Remark 3.5. *For multi-qubit states ψ that are* real *in the computational basis, we can replace step 1 of the algorithm by a single Bell sampling, which in this case directly samples from the characteristic distribution p_ψ (see Eq. (3.4)). The resulting algorithm operators on* four *copies of ψ and achieves the same guarantees as Theorem 3.2.*

Remark 3.6. *The scaling in Theorem 3.2 is optimal. Indeed, it is known that distinguishing any fixed pair of states $|\psi\rangle$, $|\phi\rangle$ with $|\langle\psi|\phi\rangle|^2 = 1 - \varepsilon^2$ requires $\Omega(1/\varepsilon^2)$ copies [MdW13]. In particular, this lower bound holds if we choose $|\psi\rangle$ to be a stabilizer state and $|\phi\rangle$ a state that is ε-far away from being a stabilizer state, in which case our Algorithm 1 is applicable.*

Remark 3.7 (Clifford testing)**.** *It follows from Theorem 3.3 that we can also test whether a given unitary U is in the Clifford group or not (without given access to $U^\dagger$). This resolves another open question in the survey of Montanaro and de Wolf [MdW13, Question 9].*

Indeed, given black-box access to U alone we can create the Choi state $|U\rangle := (U \otimes I)|\Phi^+\rangle$, which is a stabilizer state if and only if U is a Clifford unitary. Moreover, the "average case" distance measure used in the literature for quantum property testing of unitaries is precisely equal to trace distance between the corresponding Choi states [MdW13, Section 5.1.1]. Thus, by first creating the Choi state and then running our Algorithm 1 we can efficiently test whether a given unitary U is a Clifford unitary.

It is instructive to write down the accepting POVM element for Algorithm 1. From Eqs. (3.1) and (3.7), we find that it is given by

$$\Pi_{\text{accept}} = \sum_{\mathbf{a}} \Pi_{\mathbf{a}} \otimes \frac{I + W_{\mathbf{a}} \otimes W_{\mathbf{a}}^\dagger}{2} = \frac{1}{2}(I + U), \tag{3.11}$$

where we have introduced the unitary

$$U = \frac{1}{2^{2n}} \sum_{\mathbf{x},\mathbf{a}\in\mathcal{V}_n} (-1)^{[\mathbf{a},\mathbf{x}]} W_{\mathbf{x}}^{\otimes 4} \otimes W_{\mathbf{a}}^{\otimes 2} = \underbrace{\left(\frac{1}{4} \sum_{\mathbf{x},\mathbf{a}\in\mathcal{V}_1} (-1)^{[\mathbf{a},\mathbf{x}]} W_{\mathbf{x}}^{\otimes 4} \otimes W_{\mathbf{a}}^{\otimes 2}\right)}_{=:u}^{\otimes n}.$$

It is easy to verify that $U = u^{\otimes n}$ is a Clifford unitary acting on the space $\mathcal{H}_n^{\otimes 6} \cong \mathcal{H}_{6n}$ of $6n$ qubits.

For any pure state ψ, $\psi^{\otimes n}$ is in the symmetric subspace, and so invariant under left and right-multiplication by permutations. In particular, we obtain a test of the same goodness as

Theorem 3.3 if we replace U by $V = U(I^{\otimes 4} \otimes \mathbb{F})$, where $\mathbb{F} = R((12))$ denotes the operator that swaps (or flips) two blocks of n qubits. Since $\mathbb{F} = 2^{-n} \sum_{\mathbf{b}} W_{\mathbf{b}}^{\otimes 2}$, we obtain the formula

$$
\begin{aligned}
V &= 2^{-3n} \sum_{\mathbf{x},\mathbf{a},\mathbf{b}} (-1)^{[\mathbf{a},\mathbf{x}]} W_{\mathbf{x}}^{\otimes 4} \otimes (W_{\mathbf{a}} W_{\mathbf{b}})^{\otimes 2} = 2^{-3n} \sum_{\mathbf{x},\mathbf{a},\mathbf{b}} (-1)^{[\mathbf{a},\mathbf{x}+\mathbf{b}]} W_{\mathbf{x}}^{\otimes 4} \otimes W_{\mathbf{a}+\mathbf{b} \bmod 2}^{\otimes 2} \\
&= 2^{-3n} \sum_{\mathbf{x},\mathbf{a},\mathbf{b}} (-1)^{[\mathbf{a},\mathbf{x}+\mathbf{b}]} W_{\mathbf{x}}^{\otimes 4} \otimes W_{\mathbf{b}}^{\otimes 2} = 2^{-n} \sum_{\mathbf{x} \in \mathcal{V}_n} W_{\mathbf{x}}^{\otimes 6} = \underbrace{\left(\frac{1}{2} \sum_{\mathbf{x} \in \mathcal{V}_1} W_{\mathbf{x}}^{\otimes 6} \right)}_{=:\nu}^{\otimes n}.
\end{aligned} \tag{3.12}
$$

Thus, we recognize that the unitary $V = \nu^{\otimes n}$ is precisely the action of the *anti-identity* (1.4) described in the introduction (for $t = 6$):

$$
V = R(\bar{\mathbb{1}}) = 2^{-n} \left(I^{\otimes 6} + X^{\otimes 6} + Y^{\otimes 6} + Z^{\otimes 6} \right)^{\otimes n} \tag{3.13}
$$

See also Remark 3.9. We discuss anti-permutations in more detail in Definition 4.29.

Equation (3.12) allows us to express the acceptance probability of Algorithm 1 in an interesting way:

$$
\begin{aligned}
p_{\text{accept}} &= \operatorname{tr}\left[\psi^{\otimes 6} \Pi_{\text{accept}} \right] = \frac{1}{2} \left(1 + \operatorname{tr}\left[\psi^{\otimes 6} U \right] \right) = \frac{1}{2} \left(1 + \operatorname{tr}\left[\psi^{\otimes 6} V \right] \right) \\
&= \frac{1}{2} \left(1 + 2^{-n} \sum_{\mathbf{x}} \operatorname{tr}\left[\psi^{\otimes 6} W_{\mathbf{x}}^{\otimes 6} \right] \right) = \frac{1}{2} \left(1 + 2^{2n} \sum_{\mathbf{x}} c_{\psi}(\mathbf{x})^6 \right) \\
&= \frac{1}{2} \left(1 + 2^{2n} \| c_{\psi} \|_{\ell_6}^6 \right) = \frac{1}{2} \left(1 + 2^{2n} \| p_{\psi} \|_{\ell_3}^3 \right) \\
&= \frac{1}{2} \left(1 + 2^{2n} \sum_{\mathbf{x}} p_{\psi}(\mathbf{x})^3 \right) = \sum_{\mathbf{x}} p_{\psi}(\mathbf{x}) \frac{1}{2} \left(1 + 2^{2n} p_{\psi}(\mathbf{x})^2 \right).
\end{aligned} \tag{3.14}
$$

It is intuitive that the ℓ_p-norms should appear, since stabilizer states can be characterized by having a maximally peaked characteristic function and distribution (Eqs. (2.17) and (3.3)).

In fact, the result of this calculation is plainly a strengthening of Eq. (3.9), since $2^n p_{\psi}(\mathbf{x}) \leqslant 1$. If we follow the rest of the proof of Theorem 3.3 then we obtain $p_{\text{accept}} \leqslant 1 - 3\varepsilon^2/8$, a slight improvement. More importantly, though, this argument completely avoids the analysis of Bell difference sampling in Theorem 3.2. This leads us towards an approach for testing general qudit stabilizer states.

3.2 Qudit stabilizer testing

While Bell sampling can only be used for qubit systems, Eq. (3.12) has a clear generalization to arbitrary qudits. Let $d \geqslant 2$ and consider the operator

$$
V_s = d^{-n} \sum_{\mathbf{x}} (W_{\mathbf{x}} \otimes W_{\mathbf{x}}^{\dagger})^{\otimes s}. \tag{3.15}
$$

(For qubits, the Weyl operators are Hermitian and so V_3 is precisely Eq. (3.12).) Suppose we choose s such that V_s is a Hermitian unitary (we will momentarily see that this can always be done). Then

$$
\Pi_{s,\text{accept}} = \frac{1}{2}(I + V_s)
$$

is a projection. If we think of it as the accepting element of a binary POVM then

$$p_{\text{accept}} = \text{tr}[\psi^{\otimes 2s}\Pi_{s,\text{accept}}] = \frac{1}{2}\left(1 + \text{tr}[\psi^{\otimes s}V_s]\right) = \frac{1}{2}\left(1 + d^{-n}\sum_{\mathbf{x}}|\text{tr}[\psi W_{\mathbf{x}}]|^{2s}\right)$$

$$= \frac{1}{2}\left(1 + d^{(s-1)n}\sum_{\mathbf{x}} p_\psi^s(\mathbf{x})\right) = \sum_{\mathbf{x}} p_\psi(\mathbf{x})\frac{1}{2}\left(1 + d^{(s-1)n}p_\psi^{s-1}(\mathbf{x})\right), \tag{3.16}$$

which generalizes Eq. (3.14).

When is V_s Hermitian and unitary? It is always Hermitian, since $W_{\mathbf{x}} \otimes W_{\mathbf{x}}^\dagger$ only depends on $\mathbf{x}$ modulo d. For unitarity we use Eq. (2.4) and calculate

$$V_s^2 = d^{-2n}\sum_{\mathbf{x},\mathbf{y}}(W_{\mathbf{x}}W_{\mathbf{y}} \otimes W_{\mathbf{x}}^\dagger W_{\mathbf{y}}^\dagger)^{\otimes s} = d^{-2n}\sum_{\mathbf{x},\mathbf{y}}\omega^{s[\mathbf{x},\mathbf{y}]}(W_{\mathbf{x}+\mathbf{y}} \otimes W_{-(\mathbf{x}+\mathbf{y})})^{\otimes s}$$

$$= d^{-2n}\sum_{\mathbf{x},\mathbf{y}}\omega^{s[\mathbf{x},\mathbf{y}]}(W_{\mathbf{x}+\mathbf{y} \bmod d} \otimes W_{-(\mathbf{x}+\mathbf{y} \bmod d)})^{\otimes s} = d^{-2n}\sum_{\mathbf{z}}\left(\sum_{\mathbf{x}}\omega^{s[\mathbf{x},\mathbf{z}]}\right)(W_{\mathbf{z}} \otimes W_{-\mathbf{z}})^{\otimes s}.$$

If s is invertible modulo d then $\omega^{s[-,\mathbf{z}]}$ is a nontrivial character for all $\mathbf{z}$, and so the inner sum simplifies to $d^{2n}\delta_{\mathbf{z},0}$. It follows that $V_s^2 = I$, as desired. We summarize:

Lemma 3.8. *Let* $d \geqslant 2$ *and* s *an integer that is invertible modulo* d *(i.e.,* $(s,d) = 1$*). Then* V_s *is a Hermitian unitary.*

Remark 3.9 (Qubits)**.** *For qubits, the operator* V_s *is a Hermitian unitary if and only if s is odd. E.g., for* $s = 1$ *it is the unitary swap operator* $\mathbb{F}$ *and for* $s = 3$ *it is precisely Eq. (3.12) (the anti-identity), while for* $s = 2$ *it is* not *unitary but in fact* proportional *to one of the POVM elements from Bell difference sampling. Indeed,* $V_2 = 2^n\Pi_0$ *where* Π_0 *is the projection from Eq. (3.6). Thus* $\|V_2\| = 2^n$ *and so we* cannot *interpret the associated* Π_2 *as a POVM element. This already partly explains why we had to resort to six copies to test stabilizerness.*

The second ingredient used to establish Theorem 3.3 was an uncertainty principle for Weyl operators. The following lemma supplies this for general d:

Lemma 3.10 (Uncertainty relation)**.** *Let* $\delta = 1/2d$ *and* ψ *a pure state such that* $|\text{tr}[\psi W_{\mathbf{x}}]|^2 > 1 - \delta^2$ *and* $|\text{tr}[\psi W_{\mathbf{y}}]|^2 > 1 - \delta^2$*. Then* $W_{\mathbf{x}}$ *and* $W_{\mathbf{y}}$ *must commute.*

Proof. Note that

$$\|W_{\mathbf{x}}|\psi\rangle - |\psi\rangle\langle\psi|W_{\mathbf{x}}|\psi\rangle\| < \delta$$

and likewise for $W_{\mathbf{y}}$. By the triangle inequality,

$$\|W_{\mathbf{x}}W_{\mathbf{y}}|\psi\rangle - |\psi\rangle\langle\psi|W_{\mathbf{x}}|\psi\rangle\langle\psi|W_{\mathbf{y}}|\psi\rangle\|$$
$$\leqslant \|W_{\mathbf{x}}W_{\mathbf{y}}|\psi\rangle - W_{\mathbf{x}}|\psi\rangle\langle\psi|W_{\mathbf{y}}|\psi\rangle\| + \|W_{\mathbf{x}}|\psi\rangle\langle\psi|W_{\mathbf{y}}|\psi\rangle - |\psi\rangle\langle\psi|W_{\mathbf{x}}|\psi\rangle\langle\psi|W_{\mathbf{y}}|\psi\rangle\|$$
$$\leqslant \|W_{\mathbf{x}}\|\|W_{\mathbf{y}}|\psi\rangle - |\psi\rangle\langle\psi|W_{\mathbf{y}}|\psi\rangle\| + \|W_{\mathbf{x}}|\psi\rangle - |\psi\rangle\langle\psi|W_{\mathbf{x}}|\psi\rangle\|\langle\psi|W_{\mathbf{y}}|\psi\rangle$$
$$\leqslant \|W_{\mathbf{y}}|\psi\rangle - |\psi\rangle\langle\psi|W_{\mathbf{y}}|\psi\rangle\| + \|W_{\mathbf{x}}|\psi\rangle - |\psi\rangle\langle\psi|W_{\mathbf{x}}|\psi\rangle\| < 2\delta,$$

but also

$$\|W_{\mathbf{x}}W_{\mathbf{y}}|\psi\rangle - \omega^{[\mathbf{x},\mathbf{y}]}|\psi\rangle\langle\psi|W_{\mathbf{x}}|\psi\rangle\langle\psi|W_{\mathbf{y}}|\psi\rangle\|$$

$$= \|\omega^{[x,y]}W_y W_x |\psi\rangle - \omega^{[x,y]} |\psi\rangle \langle\psi|W_y|\psi\rangle \langle\psi|W_x|\psi\rangle\| < 2\delta.$$

If we combine this with another triangle inequality, we obtain that

$$|1 - \omega^{[x,y]}| = \|\omega^{[x,y]} |\psi\rangle - |\psi\rangle\| < \frac{4\delta}{\langle\psi|W_x|\psi\rangle \langle\psi|W_y|\psi\rangle} < \frac{4\delta}{1-\delta^2}.$$

Now suppose that W_x and W_y do not commute. Then $[x,y] \neq 0$ and so

$$|1 - \omega^{[x,y]}| \geqslant |1 - \omega| = 2\sin(\pi/d) \geqslant \frac{4}{d}.$$

Thus, $4/d < 4\delta/(1-\delta^2)$, which plainly contradicts our choice of δ. This is the desired contradiction and we conclude that W_x and W_y commute. $\square$

We now show that stabilizer testing can be done in arbitrary local dimension:

Theorem 3.11 (Stabilizer testing for qudits). *Let $d \geqslant 2$ and choose $s \geqslant 2$ such that $(d,s) = 1$. Let ψ be a pure state of n qudits and denote by $p_{accept} = \mathrm{tr}[\psi^{\otimes 2s}\Pi_{s,accept}]$ the probability that the POVM element $\Pi_{s,accept}$ accepts given $2s$ copies of ψ. If ψ is a stabilizer state then it accepts with certainty, $p_{accept} = 1$. On the other hand, if $\max_S|\langle S|\psi\rangle|^2 \leqslant 1 - \varepsilon^2$ then $p_{accept} \leqslant 1 - C_{d,s}\varepsilon^2$, where $C_{d,s} = (1 - (1 - 1/4d^2)^{s-1})/2$.*

Proof. If ψ is a stabilizer state, say $|\psi\rangle = |M, f\rangle$, then $p_\psi(x)$ is the uniform distribution on M, which has d^n elements. In view of Eq. (3.16),

$$p_{accept} = \sum_x p_\psi(x)\frac{1}{2}\left(1 + d^{(s-1)n}p_\psi^{s-1}(x)\right) = 1,$$

so the test accepts with certainty.

Now suppose that ψ is a general state. Define

$$M_0 \coloneqq \{x \in \mathcal{V}_n : d^n p_\psi(x) > 1 - 1/4d^2\}.$$

By Lemma 3.10, the Weyl operators W_x for $x \in M_0$ all commute. We can thus extend M_0 to a maximal set M with this property. As in the proof of Theorem 3.3, we can bound

$$\max_S|\langle S|\psi\rangle|^2 \geqslant \sum_{x \in M_0} p_\psi(x).$$

But this probability can be bounded as before using the Markov inequality (but now for a $(s-1)$st moment):

$$\sum_{x \in M_0} p_\psi(x) = 1 - \sum_{d^n p_\psi(x) \leqslant 1 - 1/4d^2} p_\psi(x) \geqslant 1 - \frac{\sum p_\psi(x)\left(1 - d^{(s-1)n}p_\psi^{s-1}(x)\right)}{1 - (1 - 1/4d^2)^{s-1}}$$

$$= 1 - \frac{2}{1 - (1 - 1/4d^2)^{s-1}}\left(1 - p_{accept}\right).$$

The last equality is Eq. (3.16). This yields the desired bound. $\square$

Remark 3.12. *It is clear that* $s = d + 1$ *is always a valid choice in Theorem 3.11. This leads to* $C_{d,s} \approx 1/8d$ *for large* d, *but the resulting test involves gates that act on* $2d + 2$ *qudits at a time. However, this choice of* s *is in general rather pessimistic. E.g., if* d *is odd then we may always choose* $s = 2$, *meaning that our test acts on four copies at a time.*

Corollary 3.13. *Let* $d \geqslant 2$ *and fix* s *as in Theorem 3.11. Let* ψ *be a pure state of* n *qudits and let* $\varepsilon > 0$. *Then there exists an quantum algorithm that, given* $O(1/C_{d,s}\varepsilon^2)$ *copies of* ψ, *accepts any stabilizer state (it is perfectly complete), while it rejects states such that* $\max_S |\langle S|\psi\rangle|^2 \leqslant 1 - \varepsilon^2$ *with probability at least 2/3.*

It is clear that the POVM measurement $\{\Pi_{s,\text{accept}}, I - \Pi_{s,\text{accept}}\}$ can be implemented efficiently. Using phase estimation, it suffices to argue that the controlled version of V_s can be implemented efficiently. But $V_s = v_s^{\otimes n}$, so its controlled version is equal to a composition of n controlled versions of v_s, each of which acts only on a constant (with respect to n) number of qudits. It follows that our stabilizer test for qudits is efficient.

It is instructive to compute the action of the unitary $V_s = v_s^{\otimes n}$ more explicitly: Let $|\mathbf{x}\rangle = |\mathbf{x}_1, \ldots, \mathbf{x}_{2s}\rangle$ denote a computational basis vector of $\mathcal{H}_n^{\otimes 2s}$. Then, using Eq. (2.1),

$$V_s |\mathbf{x}\rangle = d^{-n} \sum_{\mathbf{a} \in \mathcal{V}_n} (W_{\mathbf{a}} \otimes W_{\mathbf{a}}^\dagger)^{\otimes s} |\mathbf{x}\rangle = d^{-n} \sum_{\mathbf{p},\mathbf{q} \in \mathbb{Z}_d^n} \omega^{\mathbf{p} \cdot (s\mathbf{q} + \mathbf{x}_1 - \mathbf{x}_2 + \cdots - \mathbf{x}_{2s})} |\mathbf{x}_1 + \mathbf{q}, \mathbf{x}_2 - \mathbf{q}, \ldots\rangle$$

$$= d^{-n} \sum_{\mathbf{p},\mathbf{q} \in \mathbb{Z}_d^n} \omega^{s\mathbf{p} \cdot (\mathbf{q} + \bar{\mathbf{x}}_{\text{odd}} - \bar{\mathbf{x}}_{\text{even}})} |\mathbf{x}_1 + \mathbf{q}, \mathbf{x}_2 - \mathbf{q}, \ldots\rangle = |\mathbf{x}_1 - \bar{\mathbf{x}}_{\text{odd}} + \bar{\mathbf{x}}_{\text{even}}, \mathbf{x}_2 + \bar{\mathbf{x}}_{\text{odd}} - \bar{\mathbf{x}}_{\text{even}}, \ldots\rangle,$$

where $\bar{\mathbf{x}}_{\text{even}} = s^{-1} \sum_{k \text{ even}} \mathbf{x}_k$ and $\bar{\mathbf{x}}_{\text{odd}}$ is defined analogously. If we re-order the tensor factors so that the odd systems come first, followed by the even ones, we find that a basis vector $|\mathbf{x}_{\text{odd}}, \mathbf{x}_{\text{even}}\rangle$ is mapped to $|\mathbf{x}_{\text{odd}} - \bar{\mathbf{x}}_{\text{odd}} + \bar{\mathbf{x}}_{\text{even}}, \mathbf{x}_{\text{even}} + \bar{\mathbf{x}}_{\text{odd}} - \bar{\mathbf{x}}_{\text{even}}\rangle$. Thus, V_s is a unitary that permutes the computational basis vectors by "swapping the mean" of the even and the odd sites of the $2s$ many blocks of n qudits.

Here is one last reformulation that will be useful to connect to our algebraic results. Let $\mathbf{p}_{2s} = (-1, 1, \ldots, -1, 1) \in \mathbb{Z}_d^{2s}$ denote the 'parity vector' that is ± 1 on even/odd sites, and consider the following $2s \times 2s$ matrix with entries in $\mathbb{Z}_d$:

$$\tilde{\mathbb{1}} = \mathbb{1} - s^{-1} \mathbf{p}_{2s} \mathbf{p}_{2s}^\mathsf{T} \tag{3.17}$$

Then we can write the action of V_s as

$$V_s |\mathbf{x}\rangle = |\tilde{\mathbb{1}}(\mathbf{x}_1, \ldots, \mathbf{x}_{2s})\rangle = |(\tilde{\mathbb{1}} \otimes I_n)\mathbf{x}\rangle. \tag{3.18}$$

It is easy to verify that $\tilde{\mathbb{1}}$ is a stochastic isometry (cf. Eq. (4.36) in Section 4.3). For qubits and $s = 3$, $\tilde{\mathbb{1}}$ is just the matrix obtained by taking the 6×6 identity matrix and inverting each bit (the 'anti-identity'). This gives a pleasant and insightful interpretation of Eq. (3.12), as we will see in Section 5.2. Interestingly, the anti-identity has previously appeared in the classification of Clifford gates in [GS16] (their T_6).

4 Algebraic theory of Clifford tensor powers

In this section, we present a general framework for studying the algebraic structure of stabilizer states and Clifford operators. We start by describing the commutant of the tensor powers of the

Clifford group, where we obtain results similar in flavor to the Schur-Weyl duality between the unitary group and the symmetric group. Next, we apply this machinery to compute arbitrary moments of qudit stabilizer states, and we describe how to construct t-designs of arbitrary order from weighted Clifford orbits. Lastly, we return to the stabilizer testing problem and explain how our solution from Section 3 can be understood more systematically and generalized. In particular, we find an optimal projection that characterizes the tensor powers of stabilizer states precisely.

Throughout this section we assume that d is prime.

4.1 Commutant of Clifford tensor powers

Schur-Weyl duality in its most fundamental form asserts that any operator on $(\mathbb{C}^D)^{\otimes t}$ that commutes with $U^{\otimes t}$ for all unitaries $U \in U(D)$ is necessarily a linear combination of permutation operators. Using the double commutant theorem, this implies at once that $(\mathbb{C}^d)^{\otimes t} = \bigoplus_\lambda V_{U(D),\lambda} \otimes V_{S_t,\lambda}$, where the $V_{U(D),\lambda}$ and $V_{S_t,\lambda}$ are pairwise inequivalent irreducible representations of the unitary group $U(D)$ and of the symmetric group S_t, respectively.

The main result of this section is that the commutant of the tensor powers of the Clifford group can be completely described in terms of a natural generalization of permutation operators (see Theorem 4.3 below). Mathematically, this generalization involves Lagrangian subspaces of a space equipped with a quadratic form. Since stabilizer states can be described in terms of Lagrangian subspaces with respect to a symplectic form (Section 2), this is reminiscent of Howe's classical duality between symplectic and orthogonal group actions.

To describe the result more precisely, let T denote a subspace of $\mathbb{Z}_d^t \oplus \mathbb{Z}_d^t$. We define a corresponding operator

$$r(T) = \sum_{(\mathbf{x},\mathbf{y})\in T} |\mathbf{x}\rangle \langle \mathbf{y}|$$

on $(\mathbb{C}^d)^{\otimes t}$, where $|\mathbf{x}\rangle = |x_1, x_2, \ldots, x_t\rangle \in (\mathbb{C}^d)^{\otimes t}$ denotes the computational basis vector associated with some $\mathbf{x} \in \mathbb{Z}_d^t$. We also consider the n-fold tensor power

$$R(T) := r(T)^{\otimes n},$$

which is an operator on $((\mathbb{C}^d)^{\otimes t})^{\otimes n} \cong (\mathbb{C}^d)^{\otimes tn} \cong ((\mathbb{C}^d)^{\otimes n})^{\otimes t}$. Both $r(T)$ and $R(T)$ are represented by real matrices in the computational basis.

Definition 4.1 ($\Sigma_{t,t}$). *Consider the quadratic form* $\mathfrak{q}: \mathbb{Z}_d^{2t} \to \mathbb{Z}_D$ *defined by* $\mathfrak{q}(\mathbf{x}, \mathbf{y}) := \mathbf{x} \cdot \mathbf{x} - \mathbf{y} \cdot \mathbf{y}$.[5] *We denote by* $\Sigma_{t,t}(d)$ *the set of subspaces* $T \subseteq \mathbb{Z}_d^{2t}$ *satisfying the following properties:*

1. T *is* totally $\mathfrak{q}$-isotropic*: i.e.,* $\mathbf{x} \cdot \mathbf{x} = \mathbf{y} \cdot \mathbf{y} \pmod{D}$ *for all* $(\mathbf{x}, \mathbf{y}) \in T$.

2. T *has dimension* t *(the maximal possible dimension).*

3. T *is* stochastic*:* $\mathbf{1}_{2t} = (1, \ldots, 1) \in T$.

We will summarize the first two conditions by saying that T *is* Lagrangian*. Thus, we will call* $\Sigma_{t,t}(d)$ *the set of* stochastic Lagrangian subspaces.

---

[5]Note that for $x \in \mathbb{Z}_d$, x^2 is well-defined modulo D.

See [NW16, App. C] for a complete list of the subspaces $\Sigma_{t,t}(d)$ for $t = 3$, and Section 4.3 for examples.

In Lemma 4.5, we will show that the operators $R(T)$ are indeed in the commutant of $\mathrm{Cliff}(n, d)^{\otimes t}$. The proof is straightforward and elucidates the role of the three conditions in Definition 4.1 as well as the difference between even and odd d.

Remark 4.2. *Recall that a subspace T is called* totally isotropic *with respect to a quadratic form $\mathfrak{q}$ if $\mathfrak{q}(\mathbf{v}) = 0$ for every $\mathbf{v} \in T$. This explain our terminology in Definition 4.1.*

We can also consider the $\mathbb{Z}_d$-valued bilinear form $\mathfrak{b}((\mathbf{x}, \mathbf{y}), (\mathbf{x}', \mathbf{y}')) := \mathbf{x} \cdot \mathbf{x}' - \mathbf{y} \cdot \mathbf{y}' \in \mathbb{Z}_d$. By a straightforward calculation,

$$\mathfrak{q}(\mathbf{v} + \mathbf{w}) = \mathfrak{q}(\mathbf{v}) + \mathfrak{q}(\mathbf{w}) + 2\mathfrak{b}(\mathbf{v}, \mathbf{w}) \pmod{D} \tag{4.1}$$

for all $\mathbf{v}, \mathbf{w} \in \mathbb{Z}_d^{2t}$. Thus, $\mathfrak{q}$ is a $\mathbb{Z}_D$-valued quadratic form associated to the $\mathbb{Z}_d$-bilinear form $\mathfrak{b}$ in the sense of [Woo93]. Note that if T is totally isotropic with respect to $\mathfrak{q}$ then Eq. (4.1) shows that T is self-orthogonal, i.e., $T \subseteq T^\perp$, where

$$T^\perp := \{\mathbf{v} \in \mathbb{Z}_d^{2t} \ : \ \mathfrak{b}(\mathbf{v}, \mathbf{w}) = 0 \quad \forall \mathbf{w} \in T\}.$$

If d is odd then $\mathfrak{q}(\mathbf{v}) = \mathfrak{b}(\mathbf{v}, \mathbf{v})$, so any self-orthogonal subspace is automatically totally isotropic with respect to $\mathfrak{q}$.

If $d = 2$ then Eq. (4.1) implies that, for a self-orthogonal subspace, the set of isotropic vectors forms a subspace – so we can check total isotropicity on a basis. Moreover, for $d = 2$, if T is Lagrangian then it is automatically stochastic; indeed, $\mathfrak{b}(\mathbf{v}, \mathbf{1}_{2t}) = \mathfrak{q}(\mathbf{v}) \pmod{2}$, so $\mathbf{1}_{2t}$ is contained in any maximal totally isotropic subspace.

Our goal of this section it to prove the following theorem:

Theorem 4.3 (Commutant of Clifford tensor powers). *Let d be a prime and $n \geqslant t - 1$. Then the operators $R(T) = r(T)^{\otimes n}$ for $T \in \Sigma_{t,t}(d)$ are $\prod_{k=0}^{t-2}(d^k + 1)$ many linearly independent operators that span the commutant of the t-th tensor power action of the Clifford group for n qudits.*

It is instructive to discuss a few key features of Theorem 4.3. First, we know that the permutation group on t elements, S_t, is in the commutant of the Clifford group $\mathrm{Cliff}(n, d)$, because it is even in the commutant of the larger unitary group $U(d^n)$. Indeed, let $\pi \cdot \mathbf{y} = (y_{\pi^{-1}(1)}, \ldots, y_{\pi^{-1}(t)})$ denote the permutation action of S_t on $\mathbb{Z}_d^t$. The one can see that, for any permutation $\pi \in S_t$, the subspace $T_\pi = \{(\pi \cdot \mathbf{y}, \mathbf{y}) : \mathbf{y} \in \mathbb{Z}_d^t\}$ is Lagrangian and stochastic. The corresponding operator $R(T_\pi) = r(T_\pi)^{\otimes n}$ agrees precisely with the usual permutation action of S_t on $((\mathbb{C}^d)^{\otimes n})^{\otimes t}$. Accordingly, we may identify S_t with a subset of $\Sigma_{t,t}(d)$. We will see below in Definition 4.11 that the set of subspaces T for which $R(T)$ is invertible forms a (in general, proper) subgroup that is (in general, strictly) larger than S_t.

Remarkably, Theorem 4.3 shows that the size of the commutant *stabilizes* as soon as $n \geqslant t - 1$. That is, just like for the symmetric group in Schur-Weyl duality of D, the set $\Sigma_{t,t}(d)$ that parametrizes the commutant of the Clifford tensor powers is independent of n, the number of qudits, provided that $n \geqslant t - 1$. This stabilization, along with the fact that the operators $R(T) = r(T)^{\otimes n}$ are tensor powers, are highly useful properties in applications (e.g., [NW16] and Sections 5 and 5.2 below).

Remark 4.4. *We believe that the results of Nebe et al [NRS06] show that the operators $R(T)$ span the commutant of $\mathrm{Cliff}(n, d)^{\otimes t}$ for any value of n. But we caution that if $n < t - 1$ then the $R(T)$ are in general no longer linearly independent (e.g., [Zhu15, eqs. (9) and (10)]).*

Theorem 4.3 will be established by combining a number of intermediate results of independent interest. We first show that the operators $R(T)$ are indeed in the commutant of $\mathrm{Cliff}(n,d)^{\otimes t}$.

Lemma 4.5. *For every* $T \in \Sigma_{t,t}(d)$ *and for every* $U \in \mathrm{Cliff}(n,d)$, *we have that* $[R(T), U^{\otimes t}] = [r(T)^{\otimes n}, U^{\otimes t}] = 0$.

Proof. Up to global phases, the Clifford group is generated by the following three operators, which are allowed to act on arbitrary qudits or pairs of qudits [Got99, Far14, NBD⁺02]: The Fourier transform (also known as the Hadamard gate for $d = 2$),

$$H = \frac{1}{\sqrt{d}} \sum_{a,b \in \mathbb{Z}_d} \omega^{ab} |a\rangle \langle b|,$$

the phase gate, which is defined as

$$P = \sum_{a \in \mathbb{Z}_2} i^{a^2} |a\rangle \langle a| \text{ for } d = 2, \qquad P = \sum_{a \in \mathbb{Z}_d} \omega^{2^{-1}a(a-1)} |a\rangle \langle a| \text{ for } d \neq 2,$$

(here we use that for $d = 2$, a^2 is well-defined modulo four, while for odd d, 2 has a multiplicative inverse, denoted 2^{-1}), and the controlled addition (also known as the CNOT gate for $d = 2$)

$$\mathrm{CADD} = \sum_{a,b \in \mathbb{Z}_d} |a, a+b\rangle \langle a, b|.$$

To establish the lemma we will prove the claim for each generator (cf. [NRS06]).

The Fourier transform H is a one-qudit gate, so it suffices to show that $[H^{\otimes t}, r(T)] = 0$ for every $T \in \Sigma_{t,t}(d)$. Indeed:

$$H^{\otimes t} r(T) H^{\dagger, \otimes t} = d^{-t} \sum_{\mathbf{a,b} \in \mathbb{Z}_d^t} \sum_{(\mathbf{x,y}) \in T} \omega^{\mathbf{a \cdot x} - \mathbf{b \cdot y}} |\mathbf{a}\rangle \langle \mathbf{b}| = d^{-t} \sum_{\mathbf{a,b} \in \mathbb{Z}_d^t} \sum_{(\mathbf{x,y}) \in T} \omega^{\flat((\mathbf{a,b}),(\mathbf{x,y}))} |\mathbf{a}\rangle \langle \mathbf{b}|$$

$$= \sum_{(\mathbf{a,b}) \in T^\perp} |\mathbf{a}\rangle \langle \mathbf{b}| = r(T).$$

In the second step and third steps, we used the notation $\flat$ and $T^\perp$ from Remark 4.2, respectively, as well as that $\dim T = t$. The last step holds since $T = T^\perp$, as T is a Lagrangian subspace.

Next, we consider the phase gate, which is likewise a single-qudit gate. For $d = 2$, we have that

$$P^{\otimes t} r(T) P^{\dagger, \otimes t} = \sum_{(\mathbf{x,y}) \in T} i^{\mathbf{x \cdot x} - \mathbf{y \cdot y}} |\mathbf{x}\rangle \langle \mathbf{y}| = r(T)$$

since T is totally isotropic. For odd d, we instead compute

$$P^{\otimes t} r(T) P^{\dagger, \otimes t} = \sum_{(\mathbf{x,y}) \in T} \omega^{2^{-1} \sum_j x_j(x_j-1) - y_j(y_j-1)} |\mathbf{x}\rangle \langle \mathbf{y}| = \sum_{\mathbf{v} = (\mathbf{x,y}) \in T} \omega^{2^{-1} \flat(\mathbf{v}, \mathbf{v} - \mathbf{1}_{2t})} |\mathbf{x}\rangle \langle \mathbf{y}| = r(T),$$

since T is totally isotropic and stochastic (so $\mathbf{w} = \mathbf{v} - \mathbf{1}_{2t} \in T$ and $\flat(\mathbf{v}, \mathbf{w}) = 0$ for every $\mathbf{v} \in T$).

Lastly, we consider the controlled addition gate, which is a two-qudit gate:

$$\mathrm{CADD}^{\otimes t} r(T)^{\otimes 2} \mathrm{CADD}^{\dagger, \otimes t} = \sum_{(\mathbf{x,y}) \in T} \sum_{(\mathbf{x',y'}) \in T} \mathrm{CADD}^{\otimes t} |\mathbf{x, x'}\rangle \langle \mathbf{y, y'}| \mathrm{CADD}^{\dagger, \otimes t}$$

27

$$= \sum_{(\mathbf{x},\mathbf{y})\in T} \sum_{(\mathbf{x}',\mathbf{y}')\in T} |\mathbf{x}, \mathbf{x}+\mathbf{x}'\rangle \langle \mathbf{y}, \mathbf{y}+\mathbf{y}'| = \sum_{(\mathbf{x},\mathbf{y})\in T} \sum_{(\mathbf{x}',\mathbf{y}')\in T} |\mathbf{x}, \mathbf{x}'\rangle \langle \mathbf{y}, \mathbf{y}'|$$

where we only used that T is a subspace. $\qquad\square$

We now show that the operators $R(T)$ are linearly independent as soon as $n \geqslant t - 1$. For this, we introduce the following useful notation:

Definition 4.6 (Vectorization). *The* vectorization *operator* vec *is defined by its action in the computational basis via*

$$\mathrm{vec}(|\mathbf{x}\rangle \langle \mathbf{y}|) = |\mathbf{x}\rangle \otimes |\mathbf{y}\rangle = |\mathbf{x}, \mathbf{y}\rangle \,.$$

Lemma 4.7. *If* $n \geqslant t - 1$ *then operators* $R(T)$ *are linearly independent.*

Proof. For each $T \in \Sigma_{t,t}(d)$, consider the vectorization of $r(T)$, which we denote by $|T\rangle := \mathrm{vec}\,(r(T)) = \sum_{\mathbf{v}\in T} |\mathbf{v}\rangle \in (\mathbb{C}^d)^{\otimes 2t}$. Note that $\langle \mathbf{v}|T\rangle = \delta_{\mathbf{v}\in T}$. Clearly, $\mathrm{vec}\,(R(T)) = \mathrm{vec}\,(r(T))^{\otimes n} = |T\rangle^{\otimes n}$. Therefore, we want to show that the vectors $|T\rangle^{\otimes n}$ are linearly independent as soon as $n \geqslant t - 1$. But each T is t-dimensional and contains the vector $\mathbf{1}_{2t}$. Extend it by $\mathbf{v}_1, \ldots, \mathbf{v}_{t-1}$ to a basis of T. Then, if T' is another subspace:

$$\langle \mathbf{v}_1|\ldots\langle \mathbf{v}_{t-1}|\langle 0|^{\otimes n-(t-1)} |T'\rangle^{\otimes n} = \langle \mathbf{v}_1|\ldots\langle \mathbf{v}_{t-1}||T'\rangle^{\otimes(t-1)} = \delta_{\mathbf{v}_1,\ldots,\mathbf{v}_{t-1}\in T'} = \delta_{T,T'}$$

This concludes the proof. $\qquad\square$

So far, we have accomplished the task of finding a large set of linearly independent operators in the commutant of $\mathrm{Cliff}(n, d)^{\otimes t}$, one for each element of $\Sigma_{t,t}(d)$. In the remainder of this section we will compute the dimension of the commutant as well as the cardinality of $\Sigma_{t,t}(d)$, and show that the two numbers agree precisely. We will use the *Gaussian binomial coefficients*, which are defined by

$$\binom{n}{k}_d = \frac{[n]_d[n-1]_d \cdots [n-k+1]_d}{[k]_d[k-1]_d \cdots [1]_d}, \quad \text{where } [k]_d = \sum_{i=0}^{k-1} d^i,$$

It is well-known that $\binom{n}{k}_d$ equals the number of k-dimensional subspaces in $\mathbb{Z}_d^n$. The Gaussian binomial coefficients satisfy the following analogs of Pascal's rule,

$$\binom{n}{k}_d = d^k \binom{n-1}{k}_d + \binom{n-1}{k-1}_d, \tag{4.2}$$

and of the binomial formula,

$$\sum_{k=0}^{n} d^{k(k-1)/2} \binom{n}{k}_d t^k = \prod_{k=0}^{n-1} \left(d^k t + 1\right). \tag{4.3}$$

We now compute the dimension of the commutant. This has previously been done for $t \leqslant 4$ by Zhu [Zhu15] and before that for $d = 2, n = 1$ by van den Nest et al [vdNDdM05].

We start with the following result from [Zhu15], which reduces the dimension computation to a counting problem. Zhu arrived at this result by computing the *frame potential* of the Clifford group – essentially, the norm squared of the character of the representation $U \mapsto U^{\otimes k}$. In contrast,

we will follow the approach by van den Nest et al, who considered the action of the Clifford average (also known as the *twirl operation* or *Reynolds operator*) on the basis of Weyl operators, correcting a glitch in [vdNDdM05] along the way.[6]

Lemma 4.8 ([Zhu15, vdNDdM05]). *The dimension of the commutant of* $\mathrm{Cliff}(n,d)^{\otimes t}$ *is equal to the number of orbits for the diagonal action of the symplectic group* $\mathrm{Sp}(2n,d)$ *on* $t-1$ *copies of the phase space* $\mathbb{Z}_d^{2n}$, *i.e., for the action*

$$\Gamma \cdot (\mathbf{x}_1, \ldots, \mathbf{x}_{t-1}) = (\Gamma\mathbf{x}_1, \ldots, \Gamma\mathbf{x}_{t-1}), \tag{4.4}$$

where $\Gamma \in \mathrm{Sp}(2n,d)$ *and* $(\mathbf{x}_1, \ldots, \mathbf{x}_{t-1}) \in (\mathbb{Z}_d^{2n})^{t-1}$ *and .*

Proof. We will show that the dimensoin of the commutant is equal to the number of orbits for the diagonal action of $\mathrm{Sp}(2n,d)$ on

$$\mathcal{W}_t := \{(\mathbf{x}_1, \ldots, \mathbf{x}_t) : \sum_{i=1}^{t} \mathbf{x}_i = 0\},$$

which is plainly an equivalent statement.

We start by noting that the Weyl operators $W_\mathbf{x}$ for $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_t) \in (\mathbb{Z}_d^{2n})^t$ form a basis of the space of operators on $((\mathbb{C}^d)^{\otimes n})^{\otimes t}$. We can thus obtain a generating set of the commutant by averaging each Weyl operator $W_\mathbf{x}$ with respect to the tensor power action of the Clifford group. According to Lemma 2.1, we can for each symplectic matrix Γ fix a Clifford unitary U_Γ such that the set of $\{U_\Gamma W_\mathbf{b}\}$ equals the Clifford group, up to global phases. Let us denote by f_Γ the phase function corresponding to U_Γ, as in Eq. (2.8). Thus, the average of the Weyl operator $W_\mathbf{x}$ is, up to overall normalization, given by

$$
\begin{aligned}
\Lambda_{\mathrm{Cliff}}(W_\mathbf{x}) &:= d^{-2n} \sum_{\Gamma \in \mathrm{Sp}(2n,d)} \sum_{\mathbf{b} \in \mathbb{Z}_d^{2n}} (U_\Gamma W_\mathbf{b})^{\otimes t} W_\mathbf{x} (U_\Gamma W_\mathbf{b})^{\dagger, \otimes t} \\
&= d^{-2n} \sum_{\Gamma \in \mathrm{Sp}(2n,d)} \sum_{\mathbf{b} \in \mathbb{Z}_d^{2n}} \omega^{[\mathbf{b}, \mathbf{x}_1 + \cdots + \mathbf{x}_t]} U_\Gamma^{\otimes t} W_\mathbf{x} U_\Gamma^{\dagger, \otimes t} \\
&= \delta_{\mathbf{x} \in \mathcal{W}_t} \sum_{\Gamma \in \mathrm{Sp}(2n,d)} \omega^{f_\Gamma(\mathbf{x})} W_{\Gamma\mathbf{x}},
\end{aligned}
$$

where $f_\Gamma(\mathbf{x}) = \sum_{i=1}^{t} f_\Gamma(\mathbf{x}_i)$ and $\Gamma\mathbf{x} := (\Gamma\mathbf{x}_1, \ldots, \Gamma\mathbf{x}_t)$.

When d is odd, the phase function f can be chosen to vanish (Lemma 2.1). Thus, the averaged operator is equal to the sum of Weyl operators over the $\mathrm{Sp}(2n,d)$-orbit of $\mathbf{x}$, provided $\mathbf{x} \in \mathcal{W}_t$, and zero otherwise. Since distinct orbits are disjoint, it is clear that we obtain a basis of the commutant by averaging one Weyl operator for each orbit of the diagonal action of $\mathrm{Sp}(2n,d)$ on $\mathcal{W}_t$.

Now consider the case where $d = 2$. To each $\mathbf{x} \in \mathcal{W}_t$, associate the phase $\phi_\mathbf{x}$ (a power of τ) such that

$$W_{\mathbf{x}_1} \cdots W_{\mathbf{x}_t} = \phi_\mathbf{x} I.$$

---

[6] In Ref. [vdNDdM05], the relative phases $\omega^{f_\Gamma(\mathbf{x})}$ that appear in our Eq. (4.5) are all taken to be trivial (for qubits). The origin seems to lie in their Section II, where it is stated—in their language—that $\alpha_1 \alpha_2 \alpha_3 = 1$. But this holds only if their π is cyclic. Clifford operations inducing non-cyclic permutations do, however, exist. We thank Huangjun Zhu for identifying the root of the apparent contradiction.

Then, for each $\Gamma \in \mathrm{Sp}(2n, d)$,

$$\phi_{\mathbf{x}} I = U_{\Gamma} W_{\mathbf{x}_1} \cdots W_{\mathbf{x}_t} U_{\Gamma}^{\dagger} = (U_{\Gamma} W_{\mathbf{x}_1} U_{\Gamma}^{\dagger}) \cdots (U_{\Gamma} W_{\mathbf{x}_t} U_{\Gamma}^{\dagger}) = \omega^{f_{\Gamma}(\mathbf{x})} W_{\Gamma \mathbf{x}_1} \cdots W_{\Gamma \mathbf{x}_t} = \omega^{f_{\Gamma}(\mathbf{x})} \phi_{\Gamma \mathbf{x}}.$$

It follows that the phase function $f_{\Gamma}(\mathbf{x})$ depends only on $\mathbf{x}$ and $\Gamma \mathbf{x}$ (rather than directly on Γ) and is given explicitly by the quotient

$$\omega^{f_{\Gamma}(\mathbf{x})} = \frac{\phi_{\mathbf{x}}}{\phi_{\Gamma \mathbf{x}}}. \tag{4.5}$$

Thus, for $\mathbf{x} \in \mathcal{W}_t$,

$$\Lambda_{\mathrm{Cliff}}(W_{\mathbf{x}}) = \sum_{\Gamma \in \mathrm{Sp}(2n,d)} \frac{\phi_{\mathbf{x}}}{\phi_{\Gamma \mathbf{x}}} W_{\Gamma \mathbf{x}} = \phi_{\mathbf{x}} \sum_{\Gamma \in \mathrm{Sp}(2n,d)} \frac{W_{\Gamma \mathbf{x}}}{\phi_{\Gamma \mathbf{x}}}.$$

In particular, if $\mathbf{y}$ is in the same $\mathrm{Sp}(2n, d)$-orbit as $\mathbf{x}$ then $\Lambda_{\mathrm{Cliff}}(W_{\mathbf{y}}) = \frac{\phi_{\mathbf{y}}}{\phi_{\mathbf{x}}} \Lambda_{\mathrm{Cliff}}(W_{\mathbf{x}})$, i.e., the two averaged operators only differ by a global phase. Thus, also for $d = 2$ we obtain a basis of the commutant by averaging one Weyl operator for each orbit of the diagonal action of $\mathrm{Sp}(2n, d)$ on $\mathcal{W}_t$. $\qquad \square$

We now derive an explicit formula for the dimension of the commutant.

Theorem 4.9 (Dimension of commutant). *Let $n \geqslant t - 1$. Then the dimension of the commutant of* $\mathrm{Cliff}(n, d)^{\otimes t}$ *is equal to* $\prod_{k=0}^{t-2} (d^k + 1)$.

Proof. To count the number of orbits of the action (4.4), we will associate to any orbit O an invariant, the *dimension*, defined by $\dim(O) = \dim \mathrm{span}\{\mathbf{x}_1, \ldots, \mathbf{x}_{t-1}\}$, where $(\mathbf{x}_1, \ldots, \mathbf{x}_{t-1})$ is any point in the orbit. We write Ω_t for the set of all orbits and Ω_t^{ℓ} for the set of orbits with dimension ℓ. We will establish and solve the following recursion relation:

$$|\Omega_t^{\ell}| = |\Omega_{t-1}^{\ell}| d^{\ell} + |\Omega_{t-1}^{\ell-1}| d^{\ell-1} \tag{4.6}$$

To see why this is true, suppose $(\mathbf{x}_1, \ldots, \mathbf{x}_{t-1}) \in \Omega_t^{\ell}$. Then there are two cases:

1. $\mathbf{x}_{t-1} \in \mathrm{span}\{\mathbf{x}_1, \ldots, \mathbf{x}_{t-2}\}$: Then the orbit through $(\mathbf{x}_1, \ldots, \mathbf{x}_{t-2})$ is in Ω_{t-1}^{ℓ}, and there are d^{ℓ} ways to choose $\mathbf{x}_{t-1} \in \mathrm{span}\{\mathbf{x}_1, \ldots, \mathbf{x}_{t-2}\}$. Together, this contributes $|\Omega_{t-1}^{\ell}| d^{\ell}$ many orbits to Ω_t^{ℓ}.

2. $\mathbf{x}_{t-1} \notin \mathrm{span}\{\mathbf{x}_1, \ldots, \mathbf{x}_{t-2}\}$: Then the orbit through $(\mathbf{x}_1, \ldots, \mathbf{x}_{t-2})$ is in $\Omega_{t-1}^{\ell-1}$, and we have to count the number of ways that we can add a new vector $\mathbf{x}_{t-1}$ to $\mathrm{span}\{\mathbf{x}_1, \ldots, \mathbf{x}_{t-2}\}$ such that we get different orbits. By Witt's theorem, which also holds for alternating forms in characteristic two [Wil09], the only invariants are the inner products between $\mathbf{x}_{t-1}$ and a basis of $\mathrm{span}\{\mathbf{x}_1, \ldots, \mathbf{x}_{t-2}\}$. By assumption, the latter space has dimension $\ell - 1 \leqslant t - 2 < n$, so we have $d^{\ell-1}$ options for $\mathbf{x}_{t-1}$. Together, this contributes $|\Omega_{t-1}^{\ell-1}| d^{\ell-1}$ many orbits to Ω_t^{ℓ}.

We have thus established the recursion relation (4.6). Since $\Omega_{2,0} = \{\{0\}\}$ and $\Omega_{2,1} = \{\mathbf{x}_1 \neq 0\}$, we find the initial conditions $|\Omega_{2,0}| = |\Omega_{2,1}| = 1$. The solution to the recursion relation is

$$|\Omega_t^{\ell}| = d^{\ell(\ell-1)/2} \binom{t-1}{\ell}_d,$$

as can be verified by using Pascal's rule (4.2). Using the binomial formula (4.3), we conclude that

$$|\Omega_t| = \sum_{\ell=0}^{t-1} |\Omega_t^\ell| = \sum_{\ell=0}^{t-1} d^{\ell(\ell-1)/2} \binom{t-1}{\ell}_d = \prod_{k=0}^{t-2} (d^k + 1). \tag{4.7}$$

This establishes the desired formula for the dimension of the commutant. □

Next, we count the number of stochastic Lagrangian subspaces.

Theorem 4.10 (Cardinality of $\Sigma_{t,t}$). *We have* $|\Sigma_{t,t}(d)| = \prod_{k=0}^{t-2}(d^k + 1)$.

Proof. Let $\Sigma_{t,t}^\ell(d)$ denote the set of subspaces $T \in \Sigma_{t,t}(d)$ such that $\dim(T \cap \Delta) = t - \ell$. We will show that

$$\left|\Sigma_{t,t}^\ell(d)\right| = d^{\ell(\ell-1)/2} \binom{t-1}{\ell}_d, \tag{4.8}$$

which implies the claim by the same calculation as in Eq. (4.7). To start, consider a subspace $T \in \Sigma_{t,t}^\ell(d)$ and consider

$$T_\Delta := T \cap \Delta = \{(\mathbf{x}, \mathbf{x}) : \mathbf{x} \in X\},$$

with X a $(t - \ell)$-dimensional subspace that is uniquely determined by T. Since T is stochastic, we know that $\mathbf{1}_t \in X$. Fix a basis $\mathbf{x}_1, \ldots, \mathbf{x}_{t-\ell}$ of X and extend it by vectors $\mathbf{z}_1, \ldots, \mathbf{z}_\ell$ to a basis of $\mathbb{Z}_d^t$. Denote the dual basis with respect to the ordinary dot product by $\hat{\mathbf{x}}_1, \ldots, \hat{\mathbf{x}}_{t-\ell}, \hat{\mathbf{z}}_1, \ldots, \hat{\mathbf{z}}_\ell$. Now, any vector in $\mathbb{Z}_d^{2t}$, so particularly in T, can be written uniquely in the form $(\mathbf{a} + \mathbf{b}, \mathbf{b})$. The subspace of vectors where $\mathbf{b}$ is a linear combination of $\mathbf{z}_1, \ldots, \mathbf{z}_\ell$ forms a complement of $T_\Delta \subseteq T$, which we shall denote by T_N. Since $T_N \cap \Delta = \{0\}$, we know that $\mathbf{a} \neq 0$ for any nonzero vector in T_N. The condition that T is self-orthogonal implies that $\mathbf{a} \cdot \mathbf{x}_i = 0$ for all $i = 1, \ldots, t - \ell$, so that $\mathbf{a}$ is a linear combination of $\hat{\mathbf{z}}_1, \ldots, \hat{\mathbf{z}}_\ell$. Since also $\dim T_N = \ell$, this implies that T_N has a *unique* basis of the form $(\hat{\mathbf{z}}_1 + \mathbf{w}_1, \mathbf{w}_1), \ldots, (\hat{\mathbf{z}}_\ell + \mathbf{w}_\ell, \mathbf{w}_\ell)$, where each $\mathbf{w}_i$ is of the form $\mathbf{w}_i = \sum_{j=1}^\ell A_{ij} \mathbf{z}_j$. We still need to implement the condition that T_N is self-orthogonal. In terms of the matrix $A = (A_{ij})$, this means that

$$0 = (\hat{\mathbf{z}}_i + \mathbf{w}_i) \cdot (\hat{\mathbf{z}}_j + \mathbf{w}_j) - \mathbf{w}_i \cdot \mathbf{w}_j = \hat{\mathbf{z}}_i \cdot \hat{\mathbf{z}}_j + A_{ij} + A_{ji} \pmod{d} \tag{4.9}$$

for any i, j. This means that the lower triangular part of A is uniquely determined by the upper triangular part.

For $d \neq 2$, (4.9) furthermore implies that the diagonal entries of A are fixed, so there are in total $d^{\ell(\ell-1)/2}$ many options for A. We have thus implemented all conditions for T to be a subspace in $\Sigma_{t,t}^\ell(d)$ since, according to Remark 4.2, for $d \neq 2$, any self-orthogonal T is automatically totally isotropic. The set of $(t - \ell)$-dimensional subspaces of $\mathbb{Z}_d^t$ that contain $\mathbf{1}_t$ are in bijection with the $(t - \ell - 1)$-dimensional subspaces in $\mathbb{Z}_d^t/\mathbb{Z}_d\mathbf{1}_t$, hence there are $\binom{t-1}{t-\ell-1}_d = \binom{t-1}{\ell}_d$ many choices for X. Together, we obtain (4.8).

For $d = 2$, (4.9) gives no constraint about the diagonal entries of A. Instead, it asserts that $\hat{\mathbf{z}}_i \cdot \hat{\mathbf{z}}_i = 0$ or, equivalently, that $\hat{\mathbf{z}}_i \cdot \mathbf{1}_t = 0$ for $i = 1, \ldots, \ell$, which is automatically satisfied since $\mathbf{1}_t \in X$. We will now show that there is a unique choice for the diagonal entries of A such that T is totally isotropic with respect to the $\mathbb{Z}_4$-valued quadratic form q. By the discussion in Remark 4.2,

31

since T is self-orthogonal, it suffices to consider T_Δ and its complement T_N separately. But the vectors in T_Δ are automatically isotropic, while for T_N total isotropy amounts to the condition that

$$0 = (\hat{\mathbf{z}}_i + \mathbf{w}_i) \cdot (\hat{\mathbf{z}}_i + \mathbf{w}_i) - \mathbf{w}_i \cdot \mathbf{w}_i = \hat{\mathbf{z}}_i \cdot \hat{\mathbf{z}}_i + 2A_{ii} \pmod 4,$$

which fixes the A_{ii} uniquely. We thus obtain (4.8) by the same counting as above. $\qquad\square$

We finally obtain Theorem 4.3 as a consequence of the preceding results.

Proof of Theorem 4.3. By combining Lemma 4.5 and Theorems 4.9 and 4.10, we see that the operators $R(T)$ form a basis of the commutant of $\mathrm{Cliff}(n,d)^{\otimes t}$ on $((\mathbb{C}^d)^{\otimes n})^{\otimes t}$. $\qquad\square$

It is interesting to note that all elements $R(T)$ of our basis of the commutant of $\mathrm{Cliff}(n,d)^{\otimes s}$ have the property that $\langle S^{\otimes t}|R(T)|S^{\otimes t}\rangle = 1$ for every stabilizer state $|S\rangle$. Indeed, if $T \in \Sigma_{t,t}(d)$ and $|S\rangle = U|0\rangle^{\otimes n}$ for some Clifford unitary U, then

$$\langle S^{\otimes t}|R(T)|S^{\otimes t}\rangle = \langle S^{\otimes t}|R(T)U^{\otimes t}|0^{\otimes tn}\rangle = \langle S^{\otimes t}|U^{\otimes t}R(T)|0^{\otimes tn}\rangle = \langle 0^{\otimes tn}|R(T)|0^{\otimes tn}\rangle = 1, \quad (4.10)$$

where we used that $\mathbf{0} \in T$ (see also Eq. (4.13) below).

4.2 Structure of the commutant

Theorem 4.3 is in the spirit of Schur-Weyl duality in that it establishes a natural basis of the commutant of the tensor power action of the Clifford group (a subgroup of the unitary group), generalizing the permutation operators. Yet, in contrast to the permutation group, $\Sigma_{t,t}(d)$ is *not* in general a group and the operators $R(T)$ for $T \in \Sigma_{t,t}(d)$ are not always invertible. In this section we show that $\Sigma_{t,t}(d)$ has a rich algebraic structure.

We first observe that there is a maximal subset of $\Sigma_{t,t}(d)$ that carries a group structure such that the $R(T)$ form a (unitary) representation. The following definition and lemma identify these elements:

Definition 4.11 (O_t). *Consider the quadratic form* $q\colon \mathbb{Z}_d^t \to \mathbb{Z}_D$ *defined by* $q(\mathbf{x}) := \mathbf{x} \cdot \mathbf{x}$.[7] *We define* $O_t(d)$ *as the group of* $t \times t$*-matrices* O *with entries in* $\mathbb{Z}_d$ *that satisfy the following properties:*

1. O *is a* q*-isometry: i.e.,* $O\mathbf{x} \cdot O\mathbf{x} = \mathbf{x} \cdot \mathbf{x} \pmod D$ *for all* $\mathbf{x} \in \mathbb{Z}_d^t$.

2. O *is stochastic:* $O\mathbf{1}_t = \mathbf{1}_t \pmod d$.

We will refer to $O_t(d)$ *as the* stochastic orthogonal group*; its elements will be called* stochastic isometries.

To see that $O_t(d)$ forms a group we only need to observe that $O^{-1} = O^{\mathsf{T}}$ is again in $O_t(d)$. The following remark is completely analogous to Remark 4.2.

Remark 4.12. *Recall that a linear map is an* isometry *with respect to a quadratic form* q *if* $q(O\mathbf{x}) = q(\mathbf{x})$ *for all* $\mathbf{x} \in \mathbb{Z}_d^t$. *This justifies our terminology in Definition 4.11. As before, we note that* q *is a* $\mathbb{Z}_D$*-valued quadratic form associated to the* $\mathbb{Z}_d$*-bilinear form* $\mathbf{x} \cdot \mathbf{y}$ *in the sense of [Woo93], namely,*

$$q(\mathbf{x} + \mathbf{y}) = q(\mathbf{x}) + q(\mathbf{y}) + 2\mathbf{x} \cdot \mathbf{y} \pmod D. \tag{4.11}$$

---

[7]Recall that for $x \in \mathbb{Z}_d$, x^2 is well-defined modulo D.

In particular, any $O \in O_t(d)$ *is an orthogonal matrix in the ordinary sense that* $O^\mathsf{T}O = I$ (mod d)*, i.e.,* $O\mathbf{x} \cdot O\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ *for all* $\mathbf{x}, \mathbf{y} \in \mathbb{Z}_d^t$. *If* d *is odd then* $q(\mathbf{x}) = \mathbf{x} \cdot \mathbf{x}$, *so any orthogonal matrix is automatically a* q-*isometry.*

If d = 2 *then Eq.* (4.11) *implies that an orthogonal matrix* O *is a* q-*isometry provided that* $q(\mathbf{x}) = 1$ (mod 4) *for every column of* O *or, equivalently, for every row of* O *(since* $O^\mathsf{T} = O^{-1}$). *In particular, any* q-*isometry is automatically stochastic.*

The significance of Definition 4.11 is the following observation.

Lemma 4.13. *For every* $O \in O_t(d)$, *the subspace*

$$T_O := \{(O\mathbf{x}, \mathbf{x}) \; : \; \mathbf{x} \in \mathbb{Z}_d^t\}$$

is an element of $\Sigma_{t,t}(d)$ *and the operators*

$$r(O) := r(T_O) = \sum_{\mathbf{x}} |O\mathbf{x}\rangle \langle \mathbf{x}|, \quad R(O) := r(O)^{\otimes n} = R(T_O) \tag{4.12}$$

are unitary. Conversely, if $R(T)$ *is invertible then* $T = T_O$ *for some* $O \in O_t(d)$. *Moreover, the operators* $R(O)$ *define a unitary representation of* $O_t(d)$ *on* $(\mathbb{C}^d)^{\otimes tn}$.

Proof. Only the converse needs justification. Note that in order for $R(T)$ to be invertible, both subspaces $\{\mathbf{x} : (\mathbf{x}, \mathbf{y}) \in T\}$ and $\{\mathbf{y} : (\mathbf{x}, \mathbf{y}) \in T\}$ of $\mathbb{Z}_d^t$ should be t-dimensional (corresponding to $r(T)$ having full row and column rank). The claim now follows easily. □

We will often regard $O_t(d)$ as a subset of $\Sigma_{t,t}(d)$ via the assignment $O \mapsto T_O$. Note that any permutation matrix satisfies the conditions of Definition 4.11, so we can consider S_t as a subgroup of $O_t(d)$ for every value of d, and hence as a subset of $\Sigma_{t,t}(d)$.

Remark 4.14. *The Clifford group is a* t-*design (for* $n \geqslant t - 1$*) if and only if* $S_t = O_t(d) = \Sigma_{t,t}(d)$, *i.e., if and only if*

$$t! = |S_t| = |\Sigma_{t,t}(d)| = \prod_{k=0}^{t-2} (d^k + 1).$$

This identity always holds up to t = 2, *and up to* t = 3 *precisely in the case of qubits (*d = 2*). Thus the multiqubit Clifford group is a 3-design (but not a 4-design), while in higher dimensions the Clifford group is only a 2-design (but not a 3-design), reproducing prior beautiful results [Zhu15, Web16].*

For $O \in O_t(d)$, Eq. (4.10) implies that

$$R(O) |S\rangle^{\otimes t} = |S\rangle^{\otimes t} \tag{4.13}$$

for every stabilizer state $|S\rangle$. That is, stochastic isometries stabilize any stabilizer tensor power. We will return to discussing the implications of this important fact in Section 5 below.

Next, we note that the group $O_t(d)$ naturally acts on the elements of $\Sigma_{t,t}(d)$ from left and right, suggesting that it is the natural symmetry group of $\Sigma_{t,t}(d)$.

Definition 4.15 (Left and right action on subspaces). *Consider a subspace* $T \in \Sigma_{t,t}(d)$ *and a matrix* $O \in O_t(d)$. *We define the* left action *of* O *on* T *as follows:*

$$OT = \{(O\mathbf{x}, \mathbf{y}) : (\mathbf{x}, \mathbf{y}) \in T\}. \tag{4.14}$$

Similarly, the right action *of* O *on* T *is defined as:*

$$TO = \{(\mathbf{x}, O^T\mathbf{y}) : (\mathbf{x}, \mathbf{y}) \in T\}. \tag{4.15}$$

It is easy to check that $OT, TO \in \Sigma_{t,t}(d)$.

Note that this action is consistent with the composition of the operators $R(T)$ and $R(O)$: For all $T \in \Sigma_{t,t}(d)$ and $O, O' \in O_t(d)$ we have that

$$R(O)R(T)R(O') = R(OTO').$$

We can therefore decompose $\Sigma_{t,t}(d)$ into a disjoint union of *double cosets* with respect to the left and right action:

$$\Sigma_{t,t}(d) = O_t(d)T_1O_t(d) \cup \cdots \cup O_t(d)T_kO_t(d), \tag{4.16}$$

where $T_1, \ldots, T_k$ are choices of subspaces in $\Sigma_{t,t}(d)$ that represent the different cosets. We note that $O_t(d)$ is always one of the doubly cosets in Eq. (4.16), corresponding to, e.g., the choice $T_1 = \Delta$.

We will now derive a complete classification of the double cosets. We start with the central definition. Recall the quadratic form $q: \mathbb{Z}_d^t \to \mathbb{Z}_D$, $q(\mathbf{x}) := \mathbf{x} \cdot \mathbf{x}$ from Definition 4.11.

Definition 4.16 (Defect subspaces). *A* defect subspace *is a subspace* $N \subseteq \mathbb{Z}_d^t$ *with the following properties:*

1. N *is* totally q-isotropic: *i.e.,* $q(\mathbf{x}) = 0 \pmod{D}$ *for all* $\mathbf{x} \in N$.

2. N *is* co-stochastic: $\mathbf{1}_t \in N^\perp$, *i.e.,* $\mathbf{x} \cdot \mathbf{1}_t = 0 \pmod{d}$ *for every* $\mathbf{x} \in N$.

The quotient $N^\perp/N$ *inherits a* $\mathbb{Z}_D$*-valued quadratic form, which we also denote by* $q([\mathbf{y}]) := q(\mathbf{y})$.

Given two defect subspaces N *and* M, *we write* $\mathrm{Iso}(N, M)$ *for the set of* defect isomorphisms, *by which we mean invertible linear maps* $J: N^\perp/N \to M^\perp/M$ *with the following two properties:*

1. J *is a* q-isometry: *i.e.,* $q(J[\mathbf{y}]) = q([\mathbf{y}])$ *for all* $[\mathbf{y}] \in N^\perp/N$.

2. J *is* stochastic: $J[\mathbf{1}_t] = [\mathbf{1}_t]$.

The inverse of J *is again a map in* $\mathrm{Iso}(M, N)$.

This definition is central as it allows us to construct elements in $\Sigma_{t,t}(d)$. If N, M are defect subspaces and $J \in \mathrm{Iso}(M, N)$, then

$$\begin{aligned}
T &= \{(\mathbf{x} + \mathbf{z}, \mathbf{y} + \mathbf{w}) : [\mathbf{y}] \in M^\perp/M, \, [\mathbf{x}] = J[\mathbf{y}], \, \mathbf{z} \in N, \, \mathbf{w} \in M\} \\
&= \{(\mathbf{x}, \mathbf{y}) : \mathbf{y} \in M^\perp, \, \mathbf{x} \in J[\mathbf{y}]\}
\end{aligned} \tag{4.17}$$

is an element in $\Sigma_{t,t}(d)$. Note that, necessarily, $\dim N = \dim M$ (since J is invertible) and $\mathbf{1}_t \in N$ if and only if $\mathbf{1}_t \in M$ (since J is also stochastic). We now show that *all* elements of $\Sigma_{t,t}(d)$ can be obtained in this way.

Proposition 4.17. *Let* $T \in \Sigma_{t,t}(d)$.

1. *The subspaces* $T_{LD} := \{\mathbf{x} : (\mathbf{x}, 0) \in T\}$ *and* $T_{RD} := \{\mathbf{y} : (0, \mathbf{y}) \in T\}$ *are defect subspaces. We call them the* left *and* right defect subspaces *of* T, *respectively.*

2. $\dim T_{LD} = \dim T_{RD}$ *and* $\mathbf{1}_t \in T_{LD}$ *if and only if* $\mathbf{1}_t \in T_{RD}$.

3. $T_{LD}^{\perp} = T_L := \{\mathbf{x} : (\mathbf{x}, \mathbf{y}) \in T\}$ *and* $T_{RD}^{\perp} = T_R := \{\mathbf{y} : (\mathbf{x}, \mathbf{y}) \in T\}$,

4. *For every* $\mathbf{y} \in T_{RD}^{\perp}$, *choose some* $\mathbf{x}(\mathbf{y})$ *such that* $(\mathbf{x}(\mathbf{y}), \mathbf{y}) \in T$. *Then* $T_J : [\mathbf{y}] \mapsto [\mathbf{x}(\mathbf{y})]$ *is a well-defined defect isomorphism, i.e., an element in* $\mathrm{Iso}(T_{RD}, T_{LD})$.

5. *The data* (T_{LD}, T_{RD}, T_J) *is uniquely determined by* T.

6. T *is of the form* (4.17), *with* $T_{LD} = N$, $T_{RD} = M$, *and* $T_J = J$.

Proof. The first claim is clear from Definition 4.1. Since T is stochastic, so $(\mathbf{1}_t, 0) \in T$ if and only if $(0, \mathbf{1}_t) = \mathbf{1}_{2t} - (\mathbf{1}_t, 0) \in T$, which proves half of the second claim. Next, consider the maps

$$\pi_L : T \to \mathbb{Z}_d^t, \ (\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x} \quad \text{and} \quad \pi_R : T \to \mathbb{Z}_d^t, \ (\mathbf{x}, \mathbf{y}) \mapsto \mathbf{y}.$$

Then $T_{LD} \cong \ker \pi_R$ and $T_{RD} \cong \ker \pi_L$, while $T_L = \mathrm{ran}\,\pi_L$ and $T_R = \mathrm{ran}\,\pi_R$. Note that $T_L \subseteq T_{LD}^{\perp}$ and $T_R \subseteq T_{RD}^{\perp}$, since T is totally isotropic. Using the rank-nullity theorem,

$$\dim T_{LD}^{\perp} = \dim T - \dim T_{LD} = \dim T/\ker \pi_R = \dim \mathrm{ran}\,\pi_R = \dim T_R \leqslant \dim T_{RD}^{\perp},$$
$$\dim T_{RD}^{\perp} = \dim T - \dim T_{RD} = \dim T/\ker \pi_L = \dim \mathrm{ran}\,\pi_L = \dim T_L \leqslant \dim T_{LD}^{\perp}.$$

Adding the two inequalities we see that they must both be equalities, hence $\dim T_{LD} = \dim T_{RD}$ as well as $T_L = T_{LD}^{\perp}$, $T_R = T_{RD}^{\perp}$. This establishes the second and third claim.

For the fourth claim, first recall from above that $T_{LD}^{\perp} = T_L$ and $T_{RD}^{\perp} = T_R$, which means that for any $\mathbf{y} \in T_{RD}^{\perp}$ there exists some $\mathbf{x} \in T_{LD}^{\perp}$ such that $(\mathbf{x}, \mathbf{y}) \in T$. Next, suppose that $(\mathbf{x}, \mathbf{y}), (\mathbf{x}', \mathbf{y}') \in T$ such that $\mathbf{y} - \mathbf{y}' \in T_{RD}$. Then $(0, \mathbf{y} - \mathbf{y}') \in T$, so

$$(\mathbf{x} - \mathbf{x}', 0) = (\mathbf{x}, \mathbf{y}) - (0, \mathbf{y} - \mathbf{y}') - (\mathbf{x}', \mathbf{y}') \in T,$$

which means that $\mathbf{x} - \mathbf{x}' \in T_{LD}$. As a consequence, $[\mathbf{y}] \mapsto [\mathbf{x}(\mathbf{y})]$ is well-defined as a map from $T_{RD}^{\perp}/T_{RD}$ to $T_{LD}^{\perp}/T_{LD}$. Using Definition 4.1, it is not hard to see that it defines an element of $\mathrm{Iso}(T_{RD}, T_{LD})$.

The fifth claim is clear by construction of T_{LD} and T_{RD} and from the fact that $[\mathbf{y}] \mapsto [\mathbf{x}(\mathbf{y})]$ is well-defined. And the last claim can be seen to hold since the right-hand side of (4.17) is clearly a subset of T for our choice of N, M, and J, but also of dimension t. $\square$

Thus we can cleanly decompose a subspace T into the two defect subspaces T_{LD} and T_{RD} as well as the defect isomorphism $T_J : [\mathbf{y}] \mapsto [\mathbf{x}(\mathbf{y})]$. When T corresponds to a stochastic isometry $O \in O_t(d)$, i.e., $T = T_O = \{(O\mathbf{y}, \mathbf{y})\}$, then both defect subspaces are trivial and the defect isomorphism T_J can be identified with O itself.

According to Proposition 4.17, for every $T \in \Sigma_{t,t}(d)$, the left and right defect subspaces T_{LD} and T_{RD} necessarily have the same dimension and $\mathbf{1}_t \in T_{LD}$ if and only if $\mathbf{1}_t \in T_{RD}$. Note that if $T \in \Sigma_{t,t}(d)$ and $O, O' \in O_t(d)$, then $T' = OTO'$ has defect subspaces

$$T'_{LD} = OT_{LD} \quad \text{and} \quad T'_{RD} = (O')^{\mathsf{T}} T_{RD} \tag{4.18}$$

and defect isomorphism

$$T'_J \colon [\mathbf{y}] \mapsto [\mathbf{O}\mathbf{x}(\mathbf{O}'\mathbf{y})]. \tag{4.19}$$

Which elements $T' \in \Sigma_{t,t}(d)$ can be obtained in this way? Clearly, the left-right action by $O_t(d)$ preserves the common dimension of the defect subspaces and whether the all-ones vector is contained. We will now show that these are the only two invariants.

Lemma 4.18. *Let* $N, M \subseteq \mathbb{F}_d^t$ *be two defect subspaces with* $\dim N = \dim M$ *and* $\mathbf{1}_t \in N$ *if and only if* $\mathbf{1}_t \in M$. *Then there exists* $O \in O_t(d)$ *such that* $ON = M$.

Proof. Let $\tilde{N} \coloneqq N + \mathbb{Z}_d \mathbf{1}_t$ and $\tilde{M} \coloneqq M + \mathbb{Z}_d \mathbf{1}_t$. The assumption implies that $\dim \tilde{N} = \dim \tilde{M}$. Choose any linear isomorphism $\tilde{O} \colon \tilde{N} \to \tilde{M}$ such that $\tilde{O}\mathbf{1}_t = \mathbf{1}_t$. Since both N and M are totally isotropic and co-stochastic, $\tilde{O}$ is an isometry with respect to the symmetric bilinear form $\mathbf{x} \cdot \mathbf{y}$.

If $d > 2$, we can directly apply the usual version of Witt's lemma for symmetric bilinear forms of odd characteristic [Wil09] to see that $\tilde{O}$ extends to an isometry map O which by construction is also stochastic, i.e., $O \in O_t(d)$.

For $d = 2$, we would like to appeal to the version of Witt's lemma from [Woo93] for the $\mathbb{Z}_4$-valued quadratic form $q(\mathbf{x}) = \mathbf{x} \cdot \mathbf{x} \pmod 4$ from Remark 4.12. Here we need to verify two conditions: (i) $\tilde{O}$ should be an isometry with respect to q, i.e., $q(\tilde{O}\mathbf{x}) = q(\mathbf{x})$ for every $\mathbf{x} \in \tilde{N}$. Since we already know that $\tilde{O}$ is orthogonal, it suffices to check this condition on a generating set. By construction, $\tilde{O}\mathbf{1}_t = \mathbf{1}_t$, so the condition is clearly true for $\mathbf{x} = \mathbf{1}_t$. On the other hand, both defect subspaces are totally isotropic, which means that $q(\tilde{O}\mathbf{x}) = 0 = q(\mathbf{x}) \pmod 4$ for every $\mathbf{x} \in N$. Thus, the first condition is satisfied. (ii) We also need to check is that $\tilde{N} \cap I^\perp = \tilde{M} \cap I^\perp$, where $I \coloneqq \{\mathbf{y} \in \mathbb{Z}_2^t : \mathbf{y} \cdot \mathbf{y} = 0 \pmod 2\}$. But $I = \mathbf{1}_t^\perp$ and hence $I^\perp = \mathbb{Z}_2 \mathbf{1}_t$. By construction, $\mathbf{1}_t \in \tilde{N}$ and $\mathbf{1}_t \in \tilde{M}$, so the second condition is also satisfied. We conclude that $\tilde{O}$ extends to an isometry O with respect to the quadratic form q, which implies that $O \in O_t(2)$ (Remark 4.12). $\square$

Corollary 4.19. *Let* $N, M \subseteq \mathbb{F}_d^t$ *be two defect subspaces with* $\dim N = \dim M$ *and* $\mathbf{1}_t \in N$ *if and only if* $\mathbf{1}_t \in M$. *Then there exists* $T \in \Sigma_{t,t}(d)$ *such that* $T_{LD} = N$ *and* $T_{RD} = M$.

Proof. Take $O \in O_t(d)$ as in Lemma 4.18. Then, $O^\mathsf{T} M = N$, $O^\mathsf{T} M^\perp = N^\perp$, and hence $J \colon [\mathbf{y}] \mapsto [O^\mathsf{T}\mathbf{y}]$ is a defect isomorphism. Then the subspace (4.17) has the desired defect spaces. $\square$

Lemma 4.20. *Let* $J \colon N^\perp/N \to M^\perp/M$ *be a defect isomorphism. Then there exists an* $O \in O_t(d)$ *inducing* J, *i.e.,* $ON = M$ *and* $[O\mathbf{x}] = J[\mathbf{x}]$ *for every* $[\mathbf{x}] \in N^\perp$.

Proof (Sketch). Note that the existence of J implies that $\dim M = \dim N$ as well as $\mathbf{1}_t \in M$ if and only if $\mathbf{1}_t \in N$. This means that we can choose a linear isomorphism $\tilde{J} \colon N^\perp \to M^\perp$ that fixes $\mathbf{1}_t$, sends N to M and which restricts to J. As in the proof of Lemma 4.18, we can use the appropriate version of Witt's lemma to obtain the existence of an extension $O \in O_t(d)$. $\square$

Corollary 4.21 (Equivalence of double cosets). *Let* $T, T' \in \Sigma_{t,t}(d)$. *Then,* $T' \in O_t(d)TO_t(d)$ *if and only if* $\dim T_{LD} = \dim T'_{LD}$ *and* $\mathbf{1}_t \in T_{LD}$ *iff* $\mathbf{1}_t \in T'_{RD}$. *In particular,* $\Sigma_{t,t}(d)$ *consists of no more than* t *double cosets.*

Proof. It is clear that the two conditions are necessary. We will now argue that they are sufficient. First, use Lemma 4.18 to find O and O' such that $OT'_{LD} = T_{LD}$ and $O'T_{RD} = T'_{RD}$. Then $T'' \coloneqq OT'O'$ is such that $T''_{LD} = T_{LD}$ and $T''_{RD} = T_{RD}$ (Eq. (4.18)). Next, use Lemma 4.20 to obtain some O'' that

induces the defect isomorphism $T_J(T_J'')^{-1} \colon T_{LD}^\perp/T_{LD} \to T_{LD}^\perp/T_{LD}$. Then $O''T'' = T$, concluding the proof. For the last remark, note that the dimension of a totally isotropic subspace is never larger than $t/2$. If the dimension is zero, then it cannot contain $\mathbf{1}_t$, while if the dimension is $t/2$ then it is Lagrangian, hence must contain $\mathbf{1}_t$. Hence the number of possible dimensions is at most $1 + (t/2 - 1)2 + 1 = t$. $\qquad\square$

Remark 4.22. *We can also restrict to either the left or the right action. In this case, the resulting cosets are classified by the right and left defect subspace, respectively, which is an arbitrary defect subspace in the sense of Definition 4.16.*

Next, we give an explicit description of the operators $R(T) = r(T)^{\otimes n}$ in terms of the defect subspaces and the defect isomorphism. If N is a defect subspace, define the *coset states*

$$|N, [\mathbf{x}]\rangle := |N|^{-1/2} \sum_{\mathbf{z} \in N} |\mathbf{x} + \mathbf{z}\rangle,$$

which form an orthonormal family for $[\mathbf{x}] \in N^\perp/N$.

Lemma 4.23. *Let $T \in \Sigma_{t,t}(d)$, with defect subspaces T_{LD}, T_{RD} and defect isomorphism $T_J \colon T_{LD}^\perp/T_{LD} \to T_{RD}^\perp/T_{RD}$. Then:*

$$\frac{1}{|T_{LD}|} r(T) = \sum_{[\mathbf{y}] \in T_{RD}^\perp/T_{RD}} |T_{LD}, T_J[\mathbf{y}]\rangle \langle T_{RD}, [\mathbf{y}]|. \tag{4.20}$$

Thus, $r(T)$ is proportional to a partial isometry, and $\operatorname{rank} r(T) = |T_{LD}^\perp/T_{LD}| = |T_{RD}^\perp/T_{RD}|$.

Proof. We obtain (4.20) directly from Proposition 4.17 and (4.17). Since the coset states form two orthonormal families and T_J is a bijection, the formula for the rank follows at once. $\qquad\square$

Now consider the case when the left and right defect subspaces coincide and the defect isomorphism is trivial. That is,

$$\begin{aligned} T &= \{(\mathbf{x} + \mathbf{z}, \mathbf{x} + \mathbf{w}) \ : \ [\mathbf{x}] \in N^\perp/N, \ \mathbf{z}, \mathbf{w} \in N\} \\ &= \{(\mathbf{x}, \mathbf{y}) \ : \ \mathbf{y} \in N^\perp, \mathbf{x} \in [\mathbf{y}]\} = \{(\mathbf{x}, \mathbf{y}) \ : \ \mathbf{x} \in N^\perp, \mathbf{y} \in [\mathbf{x}]\}, \end{aligned} \tag{4.21}$$

where $N := T_{LD} = T_{RD}$ is an arbitrary defect subspace. In view of Corollary 4.21, any double coset contains a subspace of this form. In this case, $r(T)$ and $R(T)$ are related to a well-known family of codes in quantum information theory. To state the result, define the Weyl operators of, respectively, shift and multiply type:

$$Z_{\mathbf{p}} = W_{\mathbf{p},0} \qquad \text{and} \qquad X_{\mathbf{q}} = W_{0,\mathbf{q}}.$$

Given any totally isotropic subspace $N \subseteq \mathbb{Z}_d^t$, the set

$$\mathrm{CSS}(N) := \{Z_{\mathbf{p}} X_{\mathbf{q}} : \mathbf{q}, \mathbf{p} \in N\}$$

forms a stabilizer group of cardinality $|N|^2$ (since N is self-orthogonal, the Weyl operators commute). Such codes are a simple variant of *Calderbank-Shor-Sloane (CSS) codes* [Ste96b, CS96, Ste96a]. The projection onto the code space can be written as

$$P_{\mathrm{CSS}(N)} = \frac{1}{|N|^2} \sum_{\mathbf{q}, \mathbf{p} \in N} Z_{\mathbf{p}} X_{\mathbf{q}} \tag{4.22}$$

By taking the trace of Eq. (4.22), one finds the dimension of the code is given by $d^{t-2\dim N} = |N^\perp/N|$. One can readily confirm that the coset states $|N, [\mathbf{z}]\rangle$ for $[\mathbf{z}] \in N^\perp/N$ form an orthonormal basis, so

$$P_{CSS(N)} = \sum_{[\mathbf{x}] \in N^\perp/N} |N, [\mathbf{x}]\rangle \langle N, [\mathbf{x}]|, \tag{4.23}$$

In particular, all this applies in the situation of Eq. (4.21). It follows that $r(T)$ and $R(T)$ are proportional to orthogonal projections onto CSS codes associated with the defect subspaces:

Theorem 4.24 (CSS codes). *Suppose that T is of the form* (4.21), *i.e., its left and right defect subspaces coincide and that the defect isomorphism is trivial. Let $N := T_{LD} = T_{RD}$. Then,*

$$r(T) = |N| P_{CSS(N)} = d^{\dim N} P_{CSS(N)}.$$

Conversely, if $T \in \Sigma_{t,t}(d)$ is such that $r(T)$ is an orthogonal projection, then T is of the form (4.21).

Proof. The formula for $r(T)$ follows directly by comparing Eqs. (4.20) and (4.23).

Conversely, suppose that $r(T)$ is an orthogonal projection. We see from Eq. (4.20) that the range of $r(T)$ is spanned by the $|T_{LD}, [\mathbf{x}]\rangle$, so we must have

$$|T_{LD}, [\mathbf{x}]\rangle = r(T) |T_{LD}, [\mathbf{x}]\rangle = \sum_{[\mathbf{y}] \in T_{RD}^\perp/T_{RD}} |T_{LD}, T_J[\mathbf{y}]\rangle \langle T_{RD}, [\mathbf{y}] | T_{LD}, [\mathbf{x}]\rangle.$$

Since the coset states $|T_{LD}, [\mathbf{x}]\rangle$ form a basis, it follows that

$$\langle T_{RD}, [\mathbf{y}] | T_{LD}, [\mathbf{x}]\rangle = \delta_{[\mathbf{x}], T_J[\mathbf{y}]}$$

for all $\mathbf{x} \in T_{LD}$ and $\mathbf{y} \in T_{RD}$. When $[\mathbf{x}] = T_J[\mathbf{y}]$, then $\langle T_{RD}, [\mathbf{y}] | T_{LD}, [\mathbf{x}]\rangle = 1$, which implies that $T_{LD} = T_{RD}$ (the inner product is at most $|T_{LD} \cap T_{RD}|/|T_{LD}|$ in absolute value). Denoting the common defect subspace by N, it follows that $\delta_{[\mathbf{x}],[\mathbf{y}]} = \langle N, [\mathbf{y}]|N, [\mathbf{x}]\rangle = \delta_{[\mathbf{x}], T_J[\mathbf{y}]}$, so T_J is trivial. $\square$

Finally, we can equip the set of subspaces $\Sigma_{t,t}(d)$ with a *semigroup structure*, denoted by $\circ$, such that the assignment $T \mapsto R(T)$ becomes a representation, i.e.,

$$R(T_1)R(T_2) = |N_1 \cap N_2|^n R(T_1 \circ T_2) = d^{n \dim(N_1 \cap N_2)} R(T_1 \circ T_2). \tag{4.24}$$

First, if T_1 or T_2 are associated to a stochastic isometry in $O_t(d)$, then we can simply define $T_1 \circ T_2$ as in Definition 4.15. In particular, the diagonal subspace is the identity element.

Next, consider the case that T_1 and T_2 are of the form (4.21), associated to defect subspaces N_1 and N_2. Then we may define $T_1 \circ T_2$ as the Lagrangian stochastic subspace T with

$$\begin{aligned} T_{LD} &= (N_1 + N_2) \cap N_1^\perp = N_1^\perp \cap N_2 + N_1, \\ T_{RD} &= (N_1 + N_2) \cap N_2^\perp = N_1 \cap N_2^\perp + N_2, \\ T_J &: T_{RD}^\perp/T_{RD} \to T_{LD}^\perp/T_{LD}, \quad [\mathbf{y}] \mapsto [\mathbf{x}] \end{aligned} \tag{4.25}$$

where $\mathbf{x}$ is such that $\mathbf{x} - \mathbf{y} \in N_1 + N_2$.

Lemma 4.25. *The data in* (4.25) *defines a subspace in $T \in \Sigma_{t,t}(d)$ such that Eq.* (4.24) *holds.*

Proof. We first verify that T_{LD} is a defect subspace. Thus, let $\mathbf{n}_1 + \mathbf{n}_2 \in N_1^\perp$, where $\mathbf{n}_1 \in N_1$ and $\mathbf{n}_2 \in N_2$. Then,

$$q(\mathbf{n}_1 + \mathbf{n}_2) = q(\mathbf{n}_1) + q(\mathbf{n}_2) + 2\mathbf{n}_1 \cdot \mathbf{n}_2 = 2\mathbf{n}_1 \cdot \mathbf{n}_2 = 0 \pmod{D},$$

The second step holds since N_1 and N_2 are defect subspaces, so $q(\mathbf{n}_1) = q(\mathbf{n}_2) = 0 \pmod{D}$, and the third step holds since $\mathbf{n}_2 \in N_1 + N_1^\perp = N_1^\perp$, so $\mathbf{n}_1 \cdot \mathbf{n}_2 = 0 \pmod{d}$. Moreover, $\mathbf{1}_t \in N_1^\perp \cap N_2^\perp \subseteq T_{LD}^\perp$, so T_{LD} is also co-stochastic. Similarly, one can check that T_{RD} is defect subspace.

Next, we verify that T_J is a well-defined defect space isomorphism. Note that

$$T_{LD}^\perp = (N_1 + N_2^\perp) \cap N_1^\perp = N_1^\perp \cap N_2^\perp + N_1,$$
$$T_{RD}^\perp = (N_1^\perp + N_2) \cap N_2^\perp = N_1^\perp \cap N_2^\perp + N_2.$$

which shows that for every $\mathbf{y} \in T_{RD}^\perp$ there exists $\mathbf{x} \in T_{LD}^\perp$ such that $\mathbf{x} - \mathbf{y} \in N_1 + N_2$. The same holds vice versa, so the map

$$T_{RD}^\perp \to T_{LD}^\perp / T_{LD}, \quad \mathbf{y} \mapsto [\mathbf{x}] \tag{4.26}$$

is surjective provided it is well-defined. Assume that $\mathbf{x}, \mathbf{x}' \in T_{LD}^\perp$ are two vectors such that $\mathbf{x} - \mathbf{y}, \mathbf{x}' - \mathbf{y} \in N_1 + N_2$. Then, $\mathbf{x} - \mathbf{x}' \in T_{LD}^\perp \cap (N_1 + N_2) = T_{LD}$, which shows that (4.26) is indeed well-defined. Note that its kernel is given by $T_{RD}^\perp \cap (N_1 + N_2) = T_{RD}$. Thus, the induced map, which is precisely T_J from (4.25), is a well-defined invertible linear map. We still need to verify that T_J is an isometry and stochastic. The latter is clear, since $\mathbf{1}_t - \mathbf{1}_t = 0 \in N_1 + N_2$. For the former, consider $[\mathbf{x}] \in T_{LD}^\perp / T_{LD}$ and $[\mathbf{y}] \in T_{RD}^\perp / T_{RD}$ such that $\mathbf{y} - \mathbf{x} = \mathbf{n}_1 + \mathbf{n}_2$, where $\mathbf{n}_1 \in N_1$ and $\mathbf{n}_2 \in N_2$. Without loss of generality, we may assume that $\mathbf{x}, \mathbf{y} \in N_1^\perp \cap N_2^\perp$, so in particular $\mathbf{x} - \mathbf{y} \perp \mathbf{y}$ and $\mathbf{n}_2 \in N_1^\perp$. Since N_1 and N_2 are totally isotropic, it follows that $q(\mathbf{x} - \mathbf{y}) = 2\mathbf{n}_1 \cdot \mathbf{n}_2 = 0 \pmod{D}$ and $q(\mathbf{x}) = q(\mathbf{y}) + q(\mathbf{x} - \mathbf{y}) + 2\mathbf{y} \cdot (\mathbf{x} - \mathbf{y}) = 0 \pmod{D}$.

Finally, we will establish that Eq. (4.24) holds with $T_1 \circ T_2 = T$. It sufices to prove the claim for $n = 1$:

$$\begin{aligned}
r(T_1)r(T_2) &= \sum_{\mathbf{x} \in N_1^\perp} \sum_{\mathbf{y} \in N_2^\perp} \big|[\mathbf{x}] \cap [\mathbf{y}]\big| \, |\mathbf{x}\rangle \langle \mathbf{y}| = |N_1 \cap N_2| \sum_{\mathbf{x} \in N_1^\perp} \sum_{\mathbf{y} \in N_2^\perp} \delta_{\mathbf{y} - \mathbf{x} \in N_1 + N_2} |\mathbf{x}\rangle \langle \mathbf{y}| \\
&= |N_1 \cap N_2| \sum_{\mathbf{y} \in T_{RD}^\perp} \sum_{\mathbf{x} \in T_{LD}^\perp} \delta_{\mathbf{y} - \mathbf{x} \in N_1 + N_2} |\mathbf{x}\rangle \langle \mathbf{y}| = |N_1 \cap N_2| \sum_{\mathbf{y} \in T_{RD}^\perp} \sum_{\mathbf{x} \in T_J[\mathbf{y}]} |\mathbf{x}\rangle \langle \mathbf{y}| \\
&= |N_1 \cap N_2| \, r(T).
\end{aligned}$$

In the third step, we used that for $\mathbf{x} \in N_1^\perp$ and $\mathbf{y} \in N_2^\perp$, the condition that $\mathbf{y} - \mathbf{x} \in N_1 + N_2$ implies that $\mathbf{x} \in T_{LD}^\perp$ and $\mathbf{y} \in T_{RD}^\perp$, and in the fourth step we used the definition of T_J. $\qquad\square$

Finally, if T_1 and T_2 are arbitrary subspaces in $\Sigma_{t,t}(d)$ then we can always left and right multiply T_1 and T_2 by suitable stochastic isometries, thereby reducing to the preceding two cases (cf. Eq. (4.18)). The semigroup structure is highly useful for calculations (see Section 4.3 below and [Dam18]). We believe that the projections exhibited by Theorem 4.24 and the semigroup structure of Eq. (4.24) will be instrumental in understanding the fine-grained decomposition of $\mathcal{H}_n^{\otimes t}$ into irreducible representations of $\mathrm{Cliff}(n, d)$ and $O_t(d)$, generalizing the results discussed below. First results in this direction will be reported in [MMG19], with a full analysis being a direction for future work.

4.3 Examples

It is instructive to compute the commutant for small values of t. One can verify that every subspace in $\Sigma_{2,2}(d)$, as well as in $\Sigma_{3,3}(2)$ corresponds to a permutation. That is, in this case, $\Sigma_{t,t}(d) = O_t(d) = S_t$. This is consistent with the fact that the Clifford group is always a unitary 2-design, and even a 3-design in the case of qubits [Zhu15].

For certain larger values of d and t, it is still true that $\Sigma_{t,t}(d) = O_t(d)$, e.g., for $t = 3$ and $d \equiv 2$ (mod 3) [NW16]. In this case, the double commutant theorem implies that we have a proper duality akin to Schur-Weyl duality:

$$((\mathbb{C}^d)^{\otimes n})^{\otimes t} = \bigoplus_\lambda V_{\text{Cliff}(n,d),\lambda} \otimes V_{O_t(d),\lambda}, \tag{4.27}$$

where the $V_{\text{Cliff}(n,d),\lambda}$ and $V_{O_t(d),\lambda}$ are pairwise inequivalent irreducible representations of $\text{Cliff}(n, d)$ and of $O_t(d)$, respectively. It would be interesting to identify these representations further. In fact, it would be more appropriate to call Eq. (4.27) a form of a *Howe duality*, of which an example are well-known dualities between metaplectic and orthogonal groups.

In general, however, $O_t(d)$ is a proper subset of $\Sigma_{t,t}(d)$, and it is an open problem to obtain a complete duality theory in positive characteristic [How73, GH16]. We now discuss some explicit examples.

Example 4.26 ($d = 3$, $t = 3$). *In this case, $O_3(3) = S_3$, and we have that [NW16]*

$$\Sigma_{3,3}(3) = S_3 \cup S_3 \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{array} \right] S_3 \tag{4.28}$$

where we identify the matrix with its row space, a Lagrangian stochastic subspace T. The double coset of T contains only two elements, T and (12)T. In total, $\Sigma_{3,3}(3)$ contains $6 + 2 = 8$ elements, which is in agreement with Theorem 4.10.

Next, we note that T corresponds to a CSS code as in Eq. (4.21) and Theorem 4.24, with defect subspace N spanned by the all-ones vectors $\mathbf{1}_3$. Thus, $R(T) = 3^n P$, where

$$P := P_{\text{CSS}}(N) = p^{\otimes n}, \quad p := \sum_{x=0}^{2} \left(\frac{1}{\sqrt{3}} \sum_{y=0}^{2} |x+y, y-x, y\rangle \right) \left(\frac{1}{\sqrt{3}} \sum_{z=0}^{2} \langle x+z, z-x, z| \right) \tag{4.29}$$

is a projector of rank 3^n (Eq. (4.23)).

It is now straightforward to derive the decomposition of $((\mathbb{C}^3)^{\otimes n})^{\otimes 3}$ into irreducible representations of the Clifford group (for $n \geqslant 2$). We start with Schur-Weyl duality, which asserts that

$$((\mathbb{C}^3)^{\otimes n})^{\otimes 3} = \bigoplus_{\lambda \vdash 3} V_{U(3^n),\lambda} \otimes V_{S_3,\lambda} \tag{4.30}$$

$$= \text{Sym}^3((\mathbb{C}^3)^{\otimes n}) \oplus U_{(2,1)}^{3^n} \otimes V_{U(3^n),(2,1)} \oplus \text{Alt}^3((\mathbb{C}^3)^{\otimes n}), \tag{4.31}$$

where λ runs over all partitions of 3. By Eq. (4.28), the commutant is generated by S_3 and the projection P. Since P commutes with all permutations, it follows that

$$P_+ := \Pi_3^{sym} P \Pi_3^{sym} = \frac{3^{-n}}{2} \left(R(T) + R((12)T) \right),$$

40

$$P_- := \Pi_3^{alt} P \Pi_3^{alt} = \frac{3^{-n}}{2} \left(R(T) + R((12)T) \right)$$

are orthogonal projections onto subrepresentations of the Clifford group. We can compute their dimensions readily by using the formula $\mathrm{tr}[R(S)] = d^{n \dim(S \cap \Delta)}$:

$$\dim W_\pm = \mathrm{tr}[P_\pm] = \frac{3^n \pm 1}{2} \tag{4.32}$$

Thus we can decompose the symmetric and anti-symmetric subspaces further into four subrepresentations:

$$\mathrm{Sym}^3((\mathbb{C}^3)^{\otimes n}) \cong W_+ \oplus W_+^\perp,$$
$$\mathrm{Alt}^3((\mathbb{C}^3)^{\otimes n}) \cong W_- \oplus W_-^\perp.$$

Since the commutant has dimension $|\Sigma_{3,3}(3)| = 8$, these four representations along with $V_{U(3^n),(2,1)}$ (which appears twice in (4.30)) are necessarily irreducible and pairwise inequivalent. We have thus fully decomposed $((\mathbb{C}^3)^{\otimes n})^{\otimes 3}$ into irreducible representations of $\mathrm{Cliff}(n,3) \times S_3$.

Next, we discuss some multi-qubit examples.

Example 4.27 ($d = 2$, $t = 4$). *As before, we find that $\mathcal{O}_4(2) = S_4$. In addition to the $4! = 24$ permutation subspaces, there exist 6 more Lagrangian subspaces in $\Sigma_{4,4}(2)$ – making a total of 30, which is known to be the dimension of the commutant of the multi-qubit Clifford group for $n \geqslant 3$ [Zhu15, (10)]. We can decompose $\Sigma_{4,4}(2)$ into two double cosets in a form that is completely analogous to Eq. (4.28):*

$$\Sigma_{4,4}(2) = S_4 \cup S_4 \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{array} \right] S_5 \tag{4.33}$$

The given matrix is the generator matrix of a Lagrangian subspace which we denote by T_4. Similarly to above, the operator $R(T_4)$ is proportional to a projector onto a CSS code, with defect subspace spanned by the all-ones vector $\mathbf{1}_4$. This projector is given by Eq. (3.6), and it can be used to decompose $((\mathbb{C}^2)^{\otimes n})^{\otimes 4}$ into irreducible representations of the Clifford group, as explained in [ZKGG16].

Example 4.28 ($d = 2$, $t = 5$). *Likewise, for $t = 5$, it is not hard to see that (cf. [Dam18])*

$$\Sigma_{5,5}(2) = S_5 \cup S_5 \left[\begin{array}{ccccc|ccccc} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right] S_5. \tag{4.34}$$

The displayed matrix corresponds to a subspace T_5 of the form Eq. (4.21), with defect subspace spanned by the vector $(1,1,1,1,0)$, and the operator $R(T_5)$ is proportional to a projector onto a CSS code. Indeed, we have

$$R(T_5) = R(T_4) \otimes I_2^{\otimes n}.$$

We now discuss some interesting elements in the groups $O_t(d)$. For qubits, we have the class of anti-permutations introduced previously in Eq. (1.4).

Definition 4.29 (Anti-permutation). *Let $\pi \in S_t$. We define the* anti-permutation *$\bar{\pi}$ as the binary complement of the corresponding permutation matrix. Formally, it is the $t \times t$-matrix*

$$\bar{\pi} = \mathbf{1}_t \mathbf{1}_t^\mathsf{T} - \pi$$

with entries in $\mathbb{F}_2$, where we identify π with the corresponding permutation matrix.

Lemma 4.30. *Let $\pi \in S_t$. If $t \equiv 2 \pmod 4$ then $\bar{\pi} \in O_t(d)$.*

Proof. By Remark 4.12, it suffices to check that $\bar{\pi}$ is orthogonal and that $q(\mathbf{x}) = 1$ for each column. The latter holds since each column of $\bar{\pi}$ contains $t - 1 \equiv 1 \pmod 4$ ones. For the former,

$$\bar{\pi}^\mathsf{T} \bar{\pi} = (\mathbf{1}_t \mathbf{1}_t^\mathsf{T} - \pi^\mathsf{T})(\mathbf{1}_t \mathbf{1}_t^\mathsf{T} - \pi) = (t - 2)\mathbf{1}_t \mathbf{1}_t^\mathsf{T} + I \equiv I \pmod 2,$$

where we used that ordinary permutation matrices are orthogonal and stochastic, as well as that t is even. $\qquad\square$

Remark 4.31. *More generally, the entrywise binary complement maps any $A \in O_t(2)$ to an element $\bar{A} \in O_t(2)$ provided that the rows of A each have Hamming weight w such that $t \equiv 2w \pmod 4$.*

For $t \geqslant 6$, the anti-permutations are distinct from the permutations, so in particular $O_t(2) \supsetneq S_t$. (For $t = 2$, the two sets coincide.)

Example 4.32 ($d = 2$, $t = 6$). *The* anti-identity *$\bar{\mathbb{1}} \in \mathcal{O}_6(2)$ and the corresponding subspace $T_{\bar{\mathbb{1}}} \subseteq \mathbb{Z}_2^6 \oplus \mathbb{Z}_2^6$ are given by (cf. Eqs. (1.4) and (3.13)).*

$$\bar{\mathbb{1}} = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}, \quad T_{\bar{\mathbb{1}}} = \left[\begin{array}{cccccc|cccccc} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right].$$

The anti-permutations admit several possible generalizations to odd primes d. One class of generalizations is given as follows. For $\pi \in S_t$ with $d \nmid t$, define

$$\bar{\pi} = 2t^{-1} \mathbf{1}_t \mathbf{1}_t^\mathsf{T} - \pi, \tag{4.35}$$

where t^{-1} denotes the multiplicative inverse of t in $\mathbb{F}_d$. It is easy to verify that $\bar{\pi} \in O_t(d)$. Moreover, $\bar{\pi}$ is the only nontrivial linear combination of $\mathbf{1}_t \mathbf{1}_t^\mathsf{T}$ and π with this property.

Another class of generalizations is given by the formula in Eq. (3.17). Let $\mathbf{p} \in \mathbb{F}_d^t$ be vector with entries in $\{\pm 1\}$ that is 'balanced', i.e., $\mathbf{p} \cdot \mathbf{1}_t = 0$ (this requires that t is even). If $d \nmid t$ and $\pi \in S_t$ is a permutation that stabilizes $\mathbf{p}$ up to a sign, i.e., $\pi \mathbf{p} = \pm \mathbf{p}$, then

$$\tilde{\pi} = \pi \mp 2t^{-1} \mathbf{p} \mathbf{p}^\mathsf{T} \tag{4.36}$$

is an element in $O_t(d)$. In particular, this yields a large family of 'anti-identities' for odd d.

Another non-trivial example of a stochastic isometry can be constructed from the adjacency matrix A of the edge-vertex graph of the icosahedron. The icosahedron has 12 vertices, so A is a 12×12 binary matrix. Any two vertices share either zero or two neighbors, which implies that A

is orthogonal. Moreover, each vertex has 5 neighbors, which implies $q(\mathbf{x}) = 1$ for each column $\mathbf{x}$ of A. By Remark 4.12, it follows that $A \in O_{12}(2)$. The space $T_{\bar{A}}$ generated by the element-wise complement of A is just the extended Golay code $\mathcal{G}_{24}$. The latter plays an important role in the invariant theory of the Clifford group as detailed in [NRS06]. Note, however, that unlike A and T_A, it is *not* the case that $\bar{A} \in O_{12}(2)$ or $T_{\bar{A}} \in \Sigma_{12,12}(2)$. Working out the precise connection between T_A, the extended Golay code, and their respective roles in the representation theory of the Clifford group is an interesting problem we leave open. Likewise, we leave open the question of whether $R(A)$ can be given a physical interpretation, as was the case for the anti-identity.

5 Statistical properties of stabilizer states

In this section we discuss the statistical properties of the stabilizer states. We use the techniques that we developed in the last section to prove an explicit formula for the t-th moment of random stabilizer states, which vastly generalizes previous results in the quantum information literature [Zhu15, KG15, Web16, ZKGG16, HWW16]. Throughout this section, d is assumed to be a prime.

5.1 Moments of random stabilizer states

We start by studying the operator-valued t-*th moment* of the uniform distribution over all stabilizer states in $(\mathbb{C}^d)^{\otimes n}$:

$$\mathbb{E}\left[|S\rangle\langle S|^{\otimes t}\right] := \frac{1}{|\mathrm{Stab}(n,d)|} \sum_{|S\rangle\langle S|\in\mathrm{Stab}(n,d)} |S\rangle\langle S|^{\otimes t} \tag{5.1}$$

Clearly this operator can be used to calculate the average value of any polynomial of degree t in the coefficients of the wavefunction of a random stabilizer state.

Note that the operator $\mathbb{E}[|S\rangle\langle S|^{\otimes t}]$ is invariant under conjugation by Clifford operators. This is because the set of stabilizer states is a single orbit of the Clifford group. Thus $\mathbb{E}[|S\rangle\langle S|^{\otimes t}]$ is in the commutant of the Clifford group and, assuming $n \geqslant t-1$, can be written in terms of the basis $R(T)$ from Theorem 4.3,

$$\mathbb{E}\left[|S\rangle\langle S|^{\otimes t}\right] = \sum_{T\in\Sigma_{t,t}(d)} \gamma_T\, R(T), \tag{5.2}$$

for certain coefficients $\gamma_T \in \mathbb{C}$. In this section, we will show that these coefficients are all equal and establish an explicit formula for the t-th moment of a random stabilizer state which holds for all values of t and n. We start with some useful lemmas.

Remark 5.1 (Sum of traces of the R(T)). *Recall that in order to establish Theorem 4.10, we determined the cardinality of the set $\Sigma_{t,t}^\ell(d)$, whose elements are the subspaces $T \in \Sigma_{t,t}(d)$ with $\dim(T \cap \Delta) = t - \ell$. The significance of the parameter ℓ is that*

$$\mathrm{tr}\, R(T) = (\mathrm{tr}\, r(T))^n = d^{\ell n}$$

for every subspace $T \in \Sigma_{t,t}^\ell(d)$. Thus, we can also compute the sum of the traces of all $R(T)$ by using Eq. (4.8):

$$\sum_{T\in\Sigma_{t,t}(d)} \mathrm{tr}\, R(T) = \sum_{\ell=1}^{t} \left|\Sigma_{t,t}^\ell(d)\right| d^{\ell n} = \sum_{\ell=1}^{t} \binom{t-1}{\ell-1}_d d^{\frac{(t-\ell)(t-\ell-1)}{2}} d^{\ell n} = d^n \prod_{k=0}^{t-2} \left(d^k + d^n\right)$$

by a well-known analog of the binomial identity for Gaussian binomial coefficients. For $n = 0$, we recover the cardinality of $\Sigma_{t,t}(d)$, in agreement with Theorem 4.10.

Next, we prove a formula that relates moments of stabilizer states for different numbers of qudits.

Lemma 5.2. *Let $N \geqslant n > 0$. Then:*

$$\left(I^{\otimes nt} \otimes \langle 0|^{\otimes(N-n)t}\right) \mathbb{E}_{\text{Stab}(N,d)} \left[|S\rangle \langle S|^{\otimes t}\right] \left(I^{\otimes nt} \otimes |0\rangle^{\otimes(N-n)t}\right) \; \propto \; \mathbb{E}_{\text{Stab}(n,d)} \left[|S\rangle \langle S|^{\otimes t}\right],$$

and both operators are nonzero.

Proof. If $|S\rangle \in \text{Stab}(N, d)$ is a stabilizer state then the partial projection $\left(I \otimes \langle 0|^{\otimes(N-n)}\right) |S\rangle$ is either zero or proportional to a stabilizer state in $\text{Stab}(n, d)$ (see, e.g., [HNQ$^+$16, App. G]). Thus, there exist coefficients $\alpha_{S'}$ for $S' \in \text{Stab}(n, d)$ such that

$$\left(I^{\otimes nt} \otimes \langle 0|^{\otimes(N-n)t}\right) \mathbb{E}_{\text{Stab}(N,d)} \left[|S\rangle \langle S|^{\otimes t}\right] \left(I^{\otimes nt} \otimes |0\rangle^{\otimes(N-n)t}\right) = \sum_{S' \in \text{Stab}(n,d)} \alpha_{S'} |S'\rangle \langle S'|^{\otimes t}$$

It is clear that the left-hand side operator is nonzero and invariant under conjugation by $U^{\otimes t}$ for any Clifford unitary $U \in \text{Cliff}(n, d)$. Thus, we can replace the right-hand side by its Clifford average, which is plainly proportional to $\mathbb{E}_{\text{Stab}(n,d)} \left[|S\rangle \langle S|^{\otimes t}\right]$. $\qquad \square$

Theorem 5.3 (t-th moment). *Let $n, t \geqslant 1$. Then the t-th moment of a random stabilizer state in $(\mathbb{C}^d)^{\otimes n}$ is given by the formula*

$$\mathbb{E}_{\text{Stab}(n,d)} \left[|S\rangle \langle S|^{\otimes t}\right] = \frac{1}{Z_{n,d,t}} \sum_{T \in \Sigma_{t,t}(d)} R(T), \tag{5.3}$$

where $Z_{n,d,t} = d^n \prod_{k=0}^{t-2}(d^k + d^n)$.

Proof. It suffices to argue that the left-hand side and right-hand side are proportional, since the formula for the proportionality constant $Z_{n,d,t}$ follows immediately from Remark 5.1 and comparing traces. Fixing d and t, we will proceed in two steps. First, we will argue that there exist $\alpha_n \in \mathbb{C}$ and $\beta_T \in \mathbb{C}$ for each $T \in \Sigma_{t,t}(d)$ such that

$$\mathbb{E}_{\text{Stab}(n,d)} \left[|S\rangle \langle S|^{\otimes t}\right] = \alpha_n \sum_{T \in \Sigma_{t,t}(d)} \beta_T \, r(T)^{\otimes n}. \tag{5.4}$$

We will show this first for $n \geqslant t - 1$ and then for all n. Afterwards, we will find that the β_T are necessarily equal, which as just discussed implies the claim.

Let us first assume that $n \geqslant t - 1$. By Theorem 4.3, there exist coefficients $\gamma_{n,T} \in \mathbb{C}$ such that

$$\mathbb{E}_{\text{Stab}(n,d)} \left[|S\rangle \langle S|^{\otimes t}\right] = \sum_{T \in \Sigma_{t,t}(d)} \gamma_{n,T} \, r(T)^{\otimes n}, \tag{5.5}$$

since the left-hand side commutes with arbitrary t-th tensor powers of Clifford unitaries. It follows that, for every $N \geqslant n$,

$$\left(I^{\otimes nt} \otimes \langle 0|^{\otimes(N-n)t}\right) \mathbb{E}_{\text{Stab}(N,d)} \left[|S\rangle \langle S|^{\otimes t}\right] \left(I^{\otimes nt} \otimes |0\rangle^{\otimes(N-n)t}\right)$$
$$= \sum_{T \in \Sigma_{t,t}(d)} \gamma_{N,T} \, r(T)^{\otimes n} \langle 0^{\otimes t}|r(T)|0^{\otimes t}\rangle^{N-n} = \sum_{T \in \Sigma_{t,t}(d)} \gamma_{N,T} \, r(T)^{\otimes n}, \tag{5.6}$$

44

since $\mathbf{0} \in T$ for every subspace T. From Lemma 5.2 we know that Eqs. (5.5) and (5.6) are proportional and nonzero. Since the operators $r(T)^{\otimes n}$ are also linearly independent for $n \geqslant t - 1$, it follows that there exist α_n and β_T such that $\gamma_{T,n} = \alpha_n \beta_T$ for all $n \geqslant t - 1$ (e.g., we can choose $\beta_T := \gamma_{T,t-1}$). Thus, we have established Eq. (5.4) for $n \geqslant t - 1$. To extend its validity to all values of n, we observe that Eq. (5.6) holds also when $N \geqslant t - 1 > n$. Together with Lemma 5.2, we find that, indeed,

$$\mathbb{E}_{\mathrm{Stab}(n,d)} \left[|S\rangle \langle S|^{\otimes t} \right] \propto \sum_{T \in \Sigma_{t,t}(d)} \gamma_{N,T}\, r(T)^{\otimes n} \propto \sum_{T \in \Sigma_{t,t}(d)} \beta_T\, r(T)^{\otimes n},$$

which shows that there exist constants α_n and β_T such that Eq. (5.4) holds for all values of n.

We will now argue that, in this case, the β_T are necessarily all equal. For this, we compute the expectation value of an operator $R(T)^\dagger = r(T)^{\otimes n,\dagger}$. On the one hand side, by Eq. (4.10),

$$\mathrm{tr}\left[R(T)^\dagger \mathbb{E}_{\mathrm{Stab}(n,d)} \left[|S\rangle \langle S|^{\otimes t} \right] \right] = \mathbb{E}_{\mathrm{Stab}(n,d)} \langle S^{\otimes t}| R(T)^\dagger |S^{\otimes t}\rangle = 1.$$

On the other hand,

$$\mathrm{tr}\left[R(T)^\dagger \mathbb{E}_{\mathrm{Stab}(n,d)} \left[|S\rangle \langle S|^{\otimes t} \right] \right] = \alpha_n \sum_{T' \in \Sigma_{t,t}(d)} \beta_{T'} \left(\mathrm{tr}\, r(T)^\dagger r(T') \right)^n$$

$$= \alpha_n \sum_{T' \in \Sigma_{t,t}(d)} \beta_{T'}\, d^{n \dim(T \cap T')} = \alpha_n d^{nt} \left(\beta_T + O(d^{-n}) \right)$$

in the limit of large n. Thus, all β_T must be equal, and the statement of the theorem follows. $\qquad\square$

Remark 5.4. *Theorem 5.3 is reminiscent of the following average formula for the tensor power of a Haar-random pure state ψ in $\mathbb{C}^D$, which follows from Schur's lemma and Schur-Weyl duality:*

$$\mathbb{E}_{\psi\, Haar} \left[|\psi\rangle \langle\psi|^{\otimes t} \right] = \frac{1}{\prod_{k=0}^{t-1}(k+D)} \sum_{\pi \in S_t} R(\pi). \tag{5.7}$$

Remark 5.5. *When d is an odd prime, $\Sigma_{t,t}(d) = S_t$ for $t \leqslant 2$, but not for $t \geqslant 3$. Thus Eqs. (5.3) and (5.7) match for $t \leqslant 2$ and deviate for $t \geqslant 3$. Since the operators $R(T)$ are linearly independent for sufficiently large n, this shows that stabilizer states in odd prime dimension are 2-designs, but not 3-designs or higher (provided $n \geqslant 2$). Similarly, for $d = 2$, $\Sigma_{t,t}(2) = S_t$ for $t \leqslant 3$, but not or $t \geqslant 4$, which shows that multiqubit stabilizer states form 3-designs, but not 4-designs or higher [KG15] (provided $n \geqslant 3$).*

Remarkably, the theory developed in this section allows us to design complex projective t-designs for any order t from the Clifford group orbits of a finite number of fiducial states. We explain this in Section 6 below.

5.2 Minimal projections for stabilizer testing

We now return to the problem of stabilizer testing; we revisit our solution from Section 3 and characterize minimal stabilizer tests with perfect completeness.

In Section 3, we found that perfectly complete stabilizer tests were in any local dimension d given by the following accepting POVM element on $t = 2s$ copies of $(\mathbb{C}^d)^{\otimes n}$,

$$\Pi_{s,\mathrm{accept}} = \frac{1}{2} (I + V_s), \tag{5.8}$$

where V_s is the Hermitian unitary defined in Eq. (3.15) and $(d, s) = 1$. This means that V_s is a unitary operator with the property that $2s$-th tensor powers of every pure stabilizer state $|S\rangle$ are contained in its $+1$ eigenspace:

$$V_s |S\rangle^{\otimes 2s} = |S\rangle^{\otimes 2s} \tag{5.9}$$

for any pure stabilizer states $|S\rangle$. Our soundness result implies that, conversely, these are the only tensor power states with this property.

Note that V_s is an operator in the commutant of the Clifford action. This is immediate by comparing Eqs. (3.18) and (4.12), which also shows that V_s is precisely the operator $R(\tilde{\mathbb{1}})$ associated with the 'anti-identity' $\tilde{\mathbb{1}}$ defined in Eq. (3.17)!

In fact, *any* $R(O)$ stabilizes the t-th tensor powers of stabilizer states: For all $O \in O_t(d)$,

$$R(O) |S\rangle^{\otimes t} = |S\rangle^{\otimes t}. \tag{5.10}$$

We proved this in Eq. (4.13). As we just saw, V_s is such an operator, so Eq. (5.10) generalizes Eq. (5.9).

Note that, since $\tilde{\mathbb{1}}$ squares to the identity, it generates a subgroup of $O_t(d)$ that contains two elements: $\{\mathbb{1}, \tilde{\mathbb{1}}\}$. We can thus interpret the projector (5.8) as the projector onto the invariant subspace for the action of this subgroup. This suggest that we look more generally at the invariant subspaces associated with subgroups of $O_t(d)$. Larger subgroups corresponds to projectors onto smaller invariant subspaces. In particular, the minimal projector corresponds to the full group $O_t(d)$, i.e.,

$$\Pi_t^{\min} := \frac{1}{|O_t(d)|} \sum_{O \in O_t(d)} R(O). \tag{5.11}$$

By Eq. (5.10), $\Pi_t^{\min}$ accepts all stabilizer tensor powers. Remarkably, it is the *minimal* projector with this property, as follows from the following theorem.

Theorem 5.6 (Minimal stabilizer test with perfect completeness). *Let d be a prime and $n, t \geqslant 1$. Then the projector $\Pi_t^{\min}$ is the orthogonal projector onto* $\mathrm{span}\{|S\rangle^{\otimes t} : |S\rangle \langle S| \in \mathrm{Stab}(n, d)\}$.

Proof. Note that the t-th moment $\rho := \mathbb{E}\left[|S\rangle \langle S|^{\otimes t}\right]$ defined in Eq. (5.1) is a density operator that is exactly supported on the span of the stabilizer tensor powers. By the preceding discussion, it remains to prove that the support of $\Pi_t^{\min}$ is contained in the support of ρ. We start with Eq. (5.3):

$$\mathbb{E}_{\mathrm{Stab}(n,d)}\left[|S\rangle \langle S|^{\otimes t}\right] = \frac{1}{Z_{n,d,t}} \sum_{T \in \Sigma_{t,t}(d)} R(T)$$

Recall from Eq. (4.16) that we can decompose $\Sigma_{t,t}(d)$ into k double cosets,

$$\Sigma_{t,t}(d) = O_t(d) T_1 O_t(d) \cup \cdots \cup O_t(d) T_k O_t(d).$$

One of the double cosets is just $O_t(d)$, say the first, corresponding to $T_1 = \Delta$ and $R(T_1) = I$. As a consequence of Corollary 4.21, we can choose each representative T_i to be of the form (4.21). Then, Theorem 4.24 shows that $R(T_i)$ is proportional to an orthogonal projection (by a positive proportionality constant) so in particular $R(T_i) \geqslant 0$. On the other hand, we can compute the sum over each double coset by

$$\sum_{T \in O_t(d) T_i O_t(d)} R(T) = \frac{|O_t(d) T_i O_t(d)|}{|O_t(d)| \times |O_t(d)|} \sum_{O, O' \in O_t(d)} R(O) R(T_i) R(O') = c_i \Pi_t^{\min} R(T_i) \Pi_t^{\min}$$

46

where $c_i := |O_t(d)T_i O_t(d)| > 0$. Together, we obtain that

$$\rho = \frac{1}{Z_{n,d,t}} \sum_{T \in \Sigma_{t,t}(d)} R(T) = \frac{1}{Z_{n,d,t}} \sum_{i=1}^{k} c_i \, \Pi_t^{\min} R(T_i) \Pi_t^{\min}$$

$$= \frac{c_1}{Z_{n,d,t}} \left(\Pi_t^{\min} + \sum_{i=2}^{k} \frac{c_i}{c_1} \Pi_t^{\min} R(T_i) \Pi_t^{\min} \right) \geqslant \frac{c_1}{Z_{n,d,t}} \Pi_t^{\min},$$

which shows that the support of ρ indeed contains the support of $\Pi_t^{\min}$. $\qquad\square$

Note that there is no condition on t in Theorem 5.6. Indeed, while the theorem identifies the projector onto the span of stabilizer tensor powers precisely, it makes no assertion about whether this subspace contains other tensor power states than stabilizer states or not. It therefore complements our results on stabilizer testing, from which we can read off values of t such that the projector Π_t and hence $\Pi_t^{\min} \leqslant \Pi_t$ contains only stabilizer tensor powers.

It is also interesting to ask about the minimal number of copies necessary for there to exist a perfectly complete stabilizer test that is dimension-independent. For $d = 2$, it is possible to show that $t = 4, 5$ copies of a random stabilizer state become on average indistinguishable from a Haar-random pure state as $n \to \infty$. This can be done by an explicit calculation of the 4th and 5th moments using our Theorem 5.3 and Eqs. (4.33) and (4.34) (for $t = 4$, this has been carried out in [Dam18]). Thus, $t = 6$ copies as in our Theorem 3.3 are indeed optimal for multiqubit stabilizer testing. For odd d (prime or not), we know from Theorem 3.11 that $t = 4$ copies always suffice. For $d \equiv 1, 5 \pmod 6$ it follows from Theorem 8.6 below that even $t = 3$ copies suffice (and are optimal). For $d \equiv 3 \pmod 6$, we leave the question of minimal t open.

6 Construction of designs

Next we describe a construction of projective t-designs for arbitrary t based on weighted Clifford orbits. As in Sections 4 and 5, we assume that d is prime.

First, we derive expressions for the average tensor powers of the Clifford orbits of arbitrary states. For any pure state $|\Psi\rangle$, the average $\mathbb{E}_{U \in \mathrm{Cliff}(n,d)}[(U|\Psi\rangle \langle\Psi| U^\dagger)^{\otimes t}]$ commutes with $\mathrm{Cliff}(n,d)^{\otimes t}$. By Theorem 4.3, and assuming that $n \geqslant t - 1$, it can therefore be expressed as

$$\mathbb{E}_{U \in \mathrm{Cliff}(n,d)} \left[\left(U|\Psi\rangle \langle\Psi| U^\dagger \right)^{\otimes t} \right] = \sum_{T \in \Sigma_{t,t}(d)} \alpha_T' R(T) \tag{6.1}$$

for some $\alpha_T' \in \mathbb{C}$. Not all of the α_T' are independent. This is because each $(U|\Psi\rangle \langle\Psi| U^\dagger)^{\otimes t}$ is invariant under the action of $S_t \times S_t$ (acting from the left and from right), and also under taking the conjugate transpose. This motivates the following definition:

Definition 6.1 (Equivalence relation $\sim_S$). *We define an* equivalence relation $\sim_S$ *on* $\Sigma_{t,t}(d)$ *in the following way:* $T \sim_S T'$ *if and only if there exist* $\pi, \pi' \in S_t$ *such that* $T' = \pi T \pi'$ *or* $T' = \pi T^t \pi'$, *where the transposed subspace* T^t *is defined by* $T^t = \{(\mathbf{y}, \mathbf{x}) : (\mathbf{x}, \mathbf{y}) \in T\}$.

We correspondingly decompose $\Sigma_{t,t}(d)$ into equivalence classes:

$$\Sigma_{t,t}(d) = \bigcup_{i=1}^{M_{t,d}} \mathcal{F}_{t,i}(d)$$

For convenience, we choose $\mathcal{F}_{t,1}(d)$ to be the set of subspaces corresponding to the permutation group S_t (these form a single equivalence class). We also define

$$\mathcal{R}_i := \sum_{T \in \mathcal{F}_{t,i}(d)} R(T).$$

We note that the operators $\mathcal{R}_i$ are Hermitian and linearly independent.

Since the $R(T)$ are linearly independent and $R(T)^\dagger = R(T^t)$, it follows that the coefficients α'_T in Eq. (6.1) must be the same for the elements of each equivalence class. Thus,

$$\mathbb{E}_{U \in \mathrm{Cliff}(n,d)}\left[\left(U |\Psi\rangle \langle\Psi| U^\dagger\right)^{\otimes t}\right] = \sum_{i=1}^{M_{t,d}} \alpha_i \mathcal{R}_i \tag{6.2}$$

for some coefficients α_i. Note that $\alpha_i \in \mathbb{R}$ because the $\mathcal{R}_i$ are Hermitian.

Theorem 6.2. *Let d be a prime and $n \geqslant t - 1$. Then there exists an ensemble $\{p_i, \Psi_i\}_{i=1}^{M_{t,d}}$ of fiducial states in $(\mathbb{C}^d)^{\otimes n}$ such that:*

$$\mathbb{E}_{i \sim p} \mathbb{E}_{U\ Clifford}\left[\left(U |\Psi_i\rangle \langle\Psi_i| U^\dagger\right)^{\otimes t}\right] = \mathbb{E}_{\Psi\ Haar}\left[|\Psi\rangle \langle\Psi|^{\otimes t}\right]$$

Importantly, the number $M_{t,d}$ of Clifford orbit is independent of n, the number of qudits.

Proof. We start the proof by taking an arbitrary finite t-design given by an ensemble $\{p_j, \Psi_j\}_{j=1}^K$. Such designs exist (see for example the early work [SZ84]), but K can be very large. If we replace each Ψ_j by a random element in its Clifford orbit then the resulting ensemble still forms a projective t-design. This means that

$$\mathbb{E}_{j \sim p} \mathbb{E}_{U \in \mathrm{Cliff}(n,d)}\left[(U |\Psi_j\rangle \langle\Psi_j| U^\dagger)^{\otimes t}\right] \propto \mathcal{R}_1.$$

Thus, if we define $\alpha_i^{(j)}$ as the coefficient of $\mathcal{R}_i$ in the Clifford average of the fiducial state Ψ_j,

$$\mathbb{E}_{U \in \mathrm{Cliff}(n,d)}\left[\left(U |\Psi_j\rangle \langle\Psi_j| U^\dagger\right)^{\otimes t}\right] = \sum_{i=1}^{M_{t,d}} \alpha_i^{(j)} \mathcal{R}_i,$$

then

$$\mathbb{E}_{j \sim p}\left[\alpha_i^{(j)}\right] = 0 \quad \text{for } i = 2, \ldots, M_{t,d}. \tag{6.3}$$

Conversely, if $\{p_j\}$ is an arbitrary probability distribution that satisfies Eq. (6.3) then the ensemble obtained by first choosing a random fiducial state according to this distribution and then a random state in its Clifford orbit is a projective t-design. We will now explain how to modify the probabilities p_j step by step, setting more and more probabilities to zero while ensuring that Eq. (6.3) continues to hold – until all but $M_{t,d}$ of them are zero. Without loss of generality, assume that $p_1 > 0$.

48

Each step proceeds as follows: Suppose that there exist indices $2 \leqslant j_1 < j_2 < \cdots < j_{M_{t,d}} \leqslant K$ such that $p_{j_m} > 0$ for all $m = 1, \ldots, M_{t,d}$. (If no such indices exist then we are done.) Consider the linear system

$$\sum_{m=1}^{M_{t,d}} q_m \alpha_i^{(j_m)} = 0 \quad \text{for } i = 2, \ldots, M_{t,d}$$

in the indeterminates $\{q_m\}_{m=1}^{M_{t,d}}$. This system is real, homogeneous, and underconstrained, so there always exists a nontrivial real solution $q \in \mathbb{R}^{M_{t,d}}$. We can also assume that some component of q is positive (otherwise replace q by $-q$). Now consider $p_{j_m} - x q_m$ for $x \in \mathbb{R}$. At $x = 0$, all p_{j_m} are strictly positive. At some critical $x = x_c$, one of the values $p_{j_m} - x_c q_m$ becomes zero, while all other ones are still non-negative. Thus, if we modify the probabilities $p_{j,m}$ by the rule

$$p_{j_m} \mapsto p_{j_m} - x_c q_m \quad \text{for } m = 1, \ldots, M_{t,d}$$

then it still holds true that $\sum_{j=1}^{M_{t,d}} p_j \alpha_i^{(j)} = 0$, but there is now at least one additional zero among the $p_{j,m}$. This continues to hold if we further normalize the $\{p_j\}$ to be a probability distribution, i.e.,

$$p_j \mapsto \frac{p_j}{\sum_{j'} p_{j'}} \quad \text{for } j = 1, \ldots, K,$$

which is always possible since $p_1 > 0$. Thus, we obtain a probability distribution $\{p_j\}$ with strictly smaller support satisfying Eq. (6.3) and $p_1 > 0$.

We can repeat this process until there are at most $M_{t,d} - 1$ nonzero probabilities among the $\{p_j\}_{j=2}^K$. By including p_1, we arrive at an ensemble of at most $M_{t,d}$ fiducial vectors. The corresponding probabilities satisfy Eq. (6.3), which is necessary and sufficient for the ensemble of Clifford orbits to be a design. This completes the proof. $\qquad\square$

Remark 6.3. *A simple upper bound for $M_{t,d}$ is $|\Sigma_{t,t}(d)| = \prod_{k=0}^{t-2}(d^k + 1)$ from Theorem 4.10. However, in general this is a rather pessimistic estimate. For example, consider $d = t = 3$. Then, $|\Sigma_{t,t}(d)| = 8$, while there are just $M_{t,d} = 2$ equivalence classes, as follows from Eq. (4.28). One of them is the set of permutations S_3, with 6 elements, and the other one has 2 elements.*

For $d = 3$ and $t = 4$, $|\Sigma_{t,t}(d)| = 80$, while $M_{t,d} = 3$. Again, one of the equivalence classes is the permutation group with $4! = 24$ elements. The second equivalence class is the class of the anti-permutations *as defined in Eq. (4.35), which is represented by the row space of the matrix*

$$\left[\begin{array}{cccc|cccc} 1 & 2 & 2 & 2 & 1 & 0 & 0 & 0 \\ 2 & 1 & 2 & 2 & 0 & 1 & 0 & 0 \\ 2 & 2 & 1 & 2 & 0 & 0 & 1 & 0 \\ 2 & 2 & 2 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

(the Lagrangian subspace corresponding to the qutrit anti-identity $\bar{\mathbb{1}}$). This equivalence class again has 24 elements. The last equivalence class can be represented by

$$\left[\begin{array}{cccc|cccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{array}\right],$$

and it has 32 elements.

Remark 6.4. *The criterion used in the proof of Theorem 6.2 can also be used to determine fiducial states that generate a projective t-design. For example, the Clifford orbit through a single fiducial state Ψ forms a projective t-design if and only if the coefficients α_i in Eq. (6.2) vanish for $i \neq 1$.*

Let us illustrate this strategy by showing that, for any $n \geqslant 2$, there exists a qutrit state $|\psi\rangle \in \mathbb{C}^3$ such that the Clifford orbit of $\Psi = \psi^{\otimes n}$ forms a projective 3-design. We note that this state cannot be a stabilizer state, since we know that the ensemble of qutrit stabilizer states does not form a 3-design (Remark 5.5)! Instead of with $\mathcal{R}_1$ and $\mathcal{R}_2$, we will work with their multiples $\Pi_3^{sym} \propto \mathcal{R}_1$ and $P_+ \coloneqq \Pi_3^{sym} P \Pi_3^{sym} \propto \mathcal{R}_2$, where P is the projector defined in Eq. (4.29). Now consider the third moment

$$\rho_3 \coloneqq \mathbb{E}_{U \in \mathrm{Cliff}(n,3)} \left[(U |\Psi\rangle \langle\Psi| U^\dagger)^{\otimes 3} \right],$$

and expand it as

$$\rho_3 = \alpha(\psi) \, \Pi_3^{sym} + \beta(\psi) \, P_+$$

for coefficients $\alpha(\psi), \beta(\psi) \in \mathbb{R}$ which depend on the choice of fiducial state. We wish to argue that for every n there exists a single-qutrit state $|\psi\rangle$ such that $\beta(\psi) = 0$. For this, we note that the coefficients can be computed as follows:

$$1 = \mathrm{tr}[\rho_3] = \alpha(\psi) \, \mathrm{tr}[\Pi_3^{sym}] + \beta(\psi) \, \mathrm{tr}[P_+]$$
$$\langle\psi^{\otimes 3}|r(T)|\psi^{\otimes 3}\rangle^n = \mathrm{tr}[R(T)\rho_3] = 3^n \alpha(\psi) \, \mathrm{tr}[P_+] + 3^n \beta(\psi) \, \mathrm{tr}[P_+].$$

It follows that $\beta(\psi) = 0$ if and only if

$$\langle\psi^{\otimes 3}|r(T)|\psi^{\otimes 3}\rangle^n = 3^n \frac{\mathrm{tr}[P_+]}{\mathrm{tr}[\Pi_3^{sym}]} = \frac{3}{3^n + 2},$$

where we used Eq. (4.32). Thus, the Clifford orbit through $\psi^{\otimes n}$ forms a projective 3-design if and only if

$$\langle\psi^{\otimes 3}|r(T)|\psi^{\otimes 3}\rangle = \left(\frac{3}{3^n + 2} \right)^{1/n} \in \left[\frac{1}{3}, \frac{3}{5} \right] \tag{6.4}$$

But the the left-hand side is equal to one if ψ is a stabilizer state, e.g., $|\psi\rangle = |0\rangle$ (Eq. (4.10)), while is vanishes for, e.g., the non-stabilizer state $|\psi\rangle = \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle)$. By continuity it follows that there always exists a single-qutrit state $|\psi\rangle$ satisfying Eq. (6.4). It is easy to find such an $|\psi\rangle$ explicitly, e.g., by considering the one-parameter family of states $|\psi(\theta)\rangle = \cos(\theta) |0\rangle - \sin(\theta) |1\rangle$ and solving Eq. (6.4) for $\theta \in [0, \frac{\pi}{2}]$.

7 De Finetti theorems for stabilizer symmetries

In this section we establish a direct connection between our results on stabilizer testing and the celebrated *quantum de Finetti theorems*, which play an important role in characterizing entanglement and correlations in quantum states with permutation symmetry (cf. discussion in Section 1.4).

We first recall the finite quantum de Finetti theorem from [CKMR07]. Let ρ be a quantum state on $(\mathbb{C}^\ell)^{\otimes t}$ that commutes with all permutations (i.e., $[r_\pi, \rho] = 0$ for all $\pi \in S_t$). Then there exists a probability measure μ on the space of mixed states on $\mathbb{C}^\ell$ such that

$$\frac{1}{2} \left\| \rho_{1 \dots s} - \int d\mu(\sigma) \sigma^{\otimes s} \right\|_1 \leqslant 2\ell^2 \frac{s}{t}. \tag{7.1}$$

Since any quantum state that commutes with permutations admits a purification on the symmetric subspace, Eq. (7.1) follows directly from a similar result for the symmetric subspace, namely, that for every $|\Psi\rangle \in \mathrm{Sym}^t(\mathbb{C}^\ell)$ there exists a probability measure μ on *pure states* on $\mathbb{C}^\ell$ such that

$$\frac{1}{2}\left\|\Psi_{1\dots s} - \int d\mu(\phi)\phi^{\otimes s}\right\|_1 \leqslant 2\ell\frac{s}{t}. \tag{7.2}$$

In this section, we prove de Finetti theorems adapted to stabilizer states. The key idea is to extend the permutation symmetry to invariance under a larger group:

1. the stochastic orthogonal group $O_t(d)$ (for qudits in any prime dimension d), or

2. the group generated by the permutations and the anti-identity (3.13) (for qubits).

These symmetries are natural since they are carried by the tensor powers of any stabilizer state, as we proved in Eq. (4.13).

In both cases, our theorems show that the reduced density matrices are close to convex combinations of *tensor powers of stabilizer states*. In the first case, we find that the reduced state is in fact *exponentially* (in the number of traced out systems) close to a state of this form, which is a much stronger guarantee than provided by the finite de Finetti theorems of Eqs. (7.1) and (7.2) (cf. [Ren07, KM09]). In the second case, we obtain power law convergence but the symmetry requirements are drastically reduced. We establish our results first for pure states (Sections 7.1 and 7.2) and then extend them by a standard purification argument to mixed states (Section 7.3).

7.1 Exponential stabilizer de Finetti theorem

Let d be an arbitrary prime. We start with the observation that, for any two distinct stabilizer states,

$$|\langle S|S'\rangle|^2 \leqslant \frac{1}{d} \tag{7.3}$$

(this can be seen from, e.g., Eq. (2.17)). It follows that, for fixed d and n, the stabilizer tensor powers $|S\rangle^{\otimes t}$ approach orthonormality as $t \to \infty$. The following lemma makes this precise.

Lemma 7.1. *Let* d *be a prime and* $n, t \geqslant 1$. *Consider the Gram matrix* $G_{S,S'} = \langle S|S'\rangle^t$, *where* $S, S' \in \mathrm{Stab}(n, d)$. *If*

$$\varepsilon := d^{\frac{1}{2}((n+2)^2 - t)} < \frac{1}{2}$$

then the following holds:

1. *The Gram matrix is ε-close to the identity matrix in operator norm:* $\|G - I\|_\infty \leqslant \varepsilon$. *In particular, the stabilizer tensor powers* $|S\rangle^{\otimes t}$ *are linearly independent.*

2. *The nonzero eigenvalues of* $Q := \sum_S |S\rangle^{\otimes t} \langle S|^{\otimes t}$ *and its pseudoinverse* Q^+ *lie in the interval* $1 \pm 2\varepsilon$.

3. *The vectors* $(Q^+)^{1/2} |S\rangle^{\otimes t}$ *for* $S \in \mathrm{Stab}(n, d)$ *are orthonormal.*

Proof. 1. The first claim follows directly from the element-wise bound (7.3):

$$\|G - I\|_\infty \leq \left(\max_{S \neq S'}|\langle S|S'\rangle^k|\right)|\mathrm{Stab}(n,d)| \leq d^{-t/2}\,d^{(n+2)^2/2} = \varepsilon,$$

where we used the bound

$$|\mathrm{Stab}(n,d)| = d^n \prod_{i=1}^n (d^i + 1) \leq d^{(n+2)^2/2}. \tag{7.4}$$

The cardinality of the set of stabilizer states has been computed in [AG04, Prop. 2] for $d = 2$ and in [Gro06, Cor. 21] for odd d. Since $\varepsilon < 1$, the statement about the Gram matrix implies that the stabilizer tensor powers are linearly independent.

2. Now define

$$H = \sum_{S \in \mathrm{Stab}(n,d)} |S\rangle^{\otimes t}\,\langle e_S|,$$

where $|e_S\rangle$ denotes an orthonormal basis labeled by the set of stabilizer states $\mathrm{Stab}(n,d)$. Then,

$$G = H^\dagger H \qquad \text{and} \qquad Q = HH^\dagger,$$

and thus the nonzero eigenvalues of G and Q are both identical (to the squared singular values of H). By part 1, the eigenvalues of G lie in the interval $1 \pm \varepsilon$, hence the same is true for the nonzero eigenvalues of Q. Since we assumed that $\varepsilon < 1/2$, it follows that the nonzero eigenvalues of the pseudoinverse Q^+ are in the interval $1 \pm 2\varepsilon$. This establishes the second claim.

3. By the first claim, the stabilizer tensor powers are linearly independent. On the other hand,

$$|S\rangle^{\otimes t} = QQ^+|S\rangle^{\otimes t} = \sum_{S'}|S'\rangle^{\otimes t}\,\langle S'|^{\otimes t}\,Q^+|S\rangle^{\otimes t}.$$

Thus, the linear independence implies that the vectors $(Q^+)^{1/2}|S\rangle^{\otimes t}$ are orthonormal. $\qquad\square$

Theorem 7.2 (Pure-state exponential stabilizer de Finetti theorem). *Let d be a prime and $|\Psi\rangle \in \mathrm{Sym}^t((\mathbb{C}^d)^{\otimes n})$ a pure quantum state that is left invariant by the action of $O_t(d)$. Let $1 \leq s \leq t$. Then there exists a probability distribution p on $\mathrm{Stab}(n,d)$, the set of pure stabilizer states of n qudits, such that*

$$\frac{1}{2}\left\|\Psi_{1\dots s} - \sum_S p(S)|S\rangle^{\otimes s}\,\langle S|^{\otimes s}\right\|_1 \leq 2d^{\frac{1}{2}(n+2)^2}d^{-\frac{1}{2}(t-s)}.$$

Proof. By assumption, $\Pi_t^{\min}|\Psi\rangle = |\Psi\rangle$, where $\Pi_t^{\min}$ is the minimal projector from Eq. (5.11). Theorem 5.6 shows that $|\Psi\rangle$ is contained in the span of stabilizer tensor powers, i.e.,

$$|\Psi\rangle = \sum_{S \in \mathrm{Stab}(n,d)} \alpha_S|S\rangle^{\otimes t}$$

for certain coefficients $\alpha_S \in \mathbb{C}$. We now use the third and second claim of Lemma 7.1 to see that

$$\sum_S |\alpha_S|^2 = \|(Q^+)^{1/2}|\Psi\rangle\|^2 \in [1 - 2\varepsilon, 1 + 2\varepsilon] \tag{7.5}$$

where $\varepsilon := d^{\frac{1}{2}((n+2)^2-t)}$. Here we have assumed that $\varepsilon < 1/2$, for otherwise the statement of the theorem is vacuous. We now compute the partial trace over all but s subsystems:

$$\Psi_{1\ldots s} = \sum_S |\alpha_S|^2 |S\rangle^{\otimes s} \langle S|^{\otimes s} + \sum_{S \neq S'} \alpha_S \overline{\alpha}_{S'} |S\rangle^{\otimes s} \langle S'|^{\otimes s} \langle S'|S\rangle^{t-s}.$$

The norm of the cross terms is small:

$$\left\| \sum_{S \neq S'} \alpha_S \overline{\alpha}_{S'} |S\rangle^{\otimes s} \langle S'|^{\otimes s} \langle S'|S\rangle^{t-s} \right\|_1 \leqslant \sum_{S \neq S'} |\alpha_S| |\alpha_{S'}| d^{-(t-s)/2} \leqslant \left(\sum_S |\alpha_S| \right)^2 d^{-(t-s)/2}$$

$$\leqslant \left(\sum_S |\alpha_S|^2 \right) d^{(n+2)^2/2} d^{-(t-s)/2} \leqslant (1 + 2\varepsilon) d^{(n+2)^2/2} d^{-(t-s)/2} \leqslant 2 d^{(n+2)^2/2} d^{-(t-s)/2};$$

the first inequality uses Eq. (7.3), the third inequality is Eq. (7.4), the fourth inequality is the upper bound in Eq. (7.5), and the last step uses that $\varepsilon < 1/2$. Finally, define $p(S) := |\alpha_S|^2 / \sum_{S'} |\alpha_{S'}|^2$ (the denominator is positive by the lower bound in Eq. (7.5) and $\varepsilon < 1/2$). Then:

$$\left\| \Psi_{1\ldots s} - \sum_S p(S) |S\rangle^{\otimes s} \langle S|^{\otimes s} \right\|_1 \leqslant \left\| \Psi_{1\ldots s} - \sum_S |\alpha_S|^2 |S\rangle^{\otimes s} \langle S|^{\otimes s} \right\|_1 + \sum_S \left| |\alpha_S|^2 - p(S) \right|$$

$$\leqslant 2 d^{(n+2)^2/2} d^{-(t-s)/2} + \left| 1 - \sum_{S'} |\alpha_{S'}|^2 \right| \leqslant 2 d^{(n+2)^2/2} d^{-(t-s)/2} + 2\varepsilon$$

$$\leqslant 4 d^{(n+2)^2/2} d^{-(t-s)/2}. \qquad \square$$

7.2 Stabilizer de Finetti theorem for the anti-identity

We now prove a stabilizer de Finetti theorem with reduced symmetry requirements. For concreteness, we restrict to the multi-qubit case ($d = 2$) and to tensor powers that are multiples of six. Neither restriction is essential. The following theorem shows that the reduced states of an arbitrary permutation-symmetric quantum state that is invariant under the anti-identity operator $V = R(\bar{\mathbb{1}})$ from Eq. (3.13), but not necessarily under other stochastic isometries, are well-approximated by convex mixtures of tensor powers of stabilizer states.

Theorem 7.3 (Pure-state stabilizer de Finetti theorem for the anti-identity). *Let $|\Psi\rangle \in \mathrm{Sym}^t((\mathbb{C}^2)^{\otimes n})$ be a quantum state that is left invariant by the action of the anti-identity (3.13) on some (and hence every) subsystem consisting of six n-qubit blocks. Let $s < t$ be a multiple of six. Then there exists a probability distribution p on $\mathrm{Stab}(n, 2)$, the set of pure stabilizer states of n qubits, such that*

$$\frac{1}{2} \left\| \Psi_{1\ldots s} - \sum_S p(S) |S\rangle^{\otimes s} \langle S|^{\otimes s} \right\|_1 \leqslant 6\sqrt{2^{n+1}} \sqrt{\frac{s}{t}}.$$

Proof. By the ordinary finite quantum de Finetti theorem (7.2), there exists a probability measure $d\mu(\phi)$ on the set of pure states such that

$$\frac{1}{2} \left\| \Psi_{1\ldots s} - \int d\mu(\phi) \phi^{\otimes s} \right\|_1 \leqslant 2^{n+1} \frac{s}{t}. \tag{7.6}$$

53

Let $\Pi_n = (I + R(\bar{\mathbb{1}}))/2$ denote the projector onto the $+1$-eigenspace of the 6×6-anti identity for n qubits. By assumption, $(\Pi_n^{\otimes(s/6)} \otimes I^{\otimes(t-s)}) |\Psi\rangle = |\Psi\rangle$, and hence $\mathrm{tr}[\Psi_{1\ldots s}\Pi_n^{\otimes s/6}] = 1$. Since the trace distance satisfies $\frac{1}{2}\|\rho - \sigma\|_1 = \max_{0 \leqslant Q \leqslant I} \mathrm{tr}[(\rho - \sigma)Q]$, it follows that

$$\int d\mu(\phi) \left(1 - \mathrm{tr}\left[\Pi_n^{\otimes(s/6)}\phi^{\otimes s}\right]\right) \leqslant 2^{n+1}\frac{s}{t}.$$

Now recall from Eq. (3.11) that the accepting POVM element for qubit stabilizer testing is given by $\Pi_{\mathrm{accept}} = \frac{1}{2}(I + U)$, where $U = V(I^{\otimes 4} \otimes \mathbb{F})$, where $\mathbb{F} = R((12))$ is the operator that swaps two blocks of n qubits (see discussion above Eq. (3.12)). Since tensor powers of pure states are permutation-symmetric,

$$\int d\mu(\phi) \left(1 - \mathrm{tr}\left[\Pi_{\mathrm{accept}}\phi^{\otimes 6}\right]^{s/6}\right) \leqslant 2^{n+1}\frac{s}{t} \tag{7.7}$$

According to Theorem 3.3 and Eq. (3.8) and using Lemma 7.4 below, for each pure state ϕ there exists a pure stabilizer state S_ϕ such that

$$|\langle S_\phi|\phi\rangle|^{2(s/6)} \geqslant 4\,\mathrm{tr}\left[\Pi_{\mathrm{accept}}\phi^{\otimes 6}\right]^{s/6} - 3.$$

Using the estimate $1 - p^6 \leqslant 6(1-p)$, which holds for all $p \in [0,1]$, we obtain

$$1 - |\langle S_\phi|\phi\rangle|^{2s} \leqslant 6\left(1 - |\langle S_\phi|\phi\rangle|^{2(s/6)}\right) \leqslant 24\left(1 - \mathrm{tr}\left[\Pi_{\mathrm{accept}}\phi^{\otimes 6}\right]^{s/6}\right).$$

Combining this estimate with Eq. (7.7), we get

$$\int d\mu(\phi) \left(1 - |\langle S_\phi|\phi\rangle|^{2s}\right) \leqslant 24 \cdot 2^{n+1}\frac{s}{t}.$$

It follows that replacing each pure state ϕ by the nearby stabilizer state S_ϕ incurs only a small error:

$$\frac{1}{2}\left\|\int d\mu(\phi)\phi^{\otimes s} - \int d\mu(\phi)S_\phi^{\otimes s}\right\|_1 \leqslant \int d\mu(\phi)\frac{1}{2}\left\|\phi^{\otimes s} - S_\phi^{\otimes s}\right\|_1 \leqslant \int d\mu(\phi)\sqrt{1 - |\langle\phi|S_\phi\rangle|^{2s}}$$
$$\leqslant \sqrt{\int d\mu(\phi)\left(1 - |\langle\phi|S_\phi\rangle|^{2s}\right)} \leqslant \sqrt{24 \cdot 2^{n+1}\frac{s}{t}},$$

where we have used the triangle inequality, the relation between the trace distance and the fidelity between pure states, and the concavity of the square root. Together with Eq. (7.6), we obtain

$$\frac{1}{2}\left\|\Psi_{1\ldots s} - \int d\mu(\phi)S_\phi^{\otimes s}\right\|_1 \leqslant 2^{n+1}\frac{s}{t} + \sqrt{24 \cdot 2^{n+1}\frac{s}{t}} \leqslant 6\sqrt{2^{n+1}}\sqrt{\frac{s}{t}}$$

where we have assumed without loss of generality that $2^{n+1}\frac{s}{t} \leqslant 1$ (otherwise, the right-hand side is larger than one so the resulting bound is trivially true). $\qquad\square$

Lemma 7.4. *The following bound holds for all $k \geqslant 1$ and $p \in [0,1]$ such that the right-hand side is non-negative:*

$$(4p - 3)^k \geqslant 4p^k - 3.$$

Proof. We will prove the inequality for all $k \geqslant 1$ and $p \in [\frac{3}{4}, 1]$. For this, note that the two expressions coincide for $p = 1$ and that the derivative of their difference is negative for all $p \in [\frac{3}{4}, 1]$. Indeed,

$$\partial_p\left((4p-3)^k - (4p^k - 3)\right) \leqslant 0 \quad \Leftrightarrow \quad (4p-3)^{k-1} \leqslant p^{k-1}.$$

In the interval that we are considering, $0 \leqslant 4p - 3 \leqslant p$, so the right-hand side condition holds. $\qquad\square$

7.3 Extension to mixed states

In this section, we extend Theorems 7.2 and 7.3 to the case of mixed density matrices. This is done using a standard purification argument as used to derive the ordinary quantum de Finetti theorem for mixed states from the version for pure states (i.e., Eq. (7.1) from Eq. (7.2)). For the next lemma, recall the vectorization operation from Definition 4.6.

Lemma 7.5 (Purification and symmetries). *Let ρ be positive semi-definite, $|\Psi\rangle = \text{vec}(\rho^{1/2})$ its standard purification, and O a unitary with real matrix elements in the computational basis. Then the following conditions are equivalent:*

1. $(O \otimes O)|\Psi\rangle = |\Psi\rangle$.

2. $[\rho, O] = 0$.

Proof. We observe that

$$(O \otimes O)|\Psi\rangle = \left(O \otimes (O^{-1})^\mathsf{T}\right)|\Psi\rangle = \text{vec}(O\rho^{1/2}O^{-1}).$$

Thus, condition 1 is equivalent to $O\rho^{1/2}O^{-1} = \rho^{1/2}$. It follows that condition 1 implies condition 2 by squaring. Conversely, assuming condition 2,

$$\rho = O\rho O^{-1} = \left(O\rho^{1/2}O^{-1}\right)\left(O\rho^{1/2}O^{-1}\right).$$

Hence $O\rho^{1/2}O^{-1}$ is a positive semi-definite square root of ρ. Since such square roots are unique, this implies that $O\rho^{-1/2}O^{-1} = \rho^{1/2}$ and hence condition 1. $\qquad\square$

Clearly, the operators $R(O)$ for $O \in O_t(d)$ have real matrix elements in the computational basis. In fact, they are given by permutation matrices in the computational basis, as can be seen from the formula given in Eq. (4.12). Thus they satisfy the conditions of Lemma 7.5. We use this now to extend our de Finetti theorems to mixed states.

Theorem 7.6 (Exponential stabilizer de Finetti theorem). *Let d be a prime and ρ a quantum state on $((\mathbb{C}^d)^{\otimes n})^{\otimes t}$ that commutes with the action of $O_t(d) \supseteq S_t$. Let $1 \leqslant s \leqslant t$. Then there exists a probability distribution p on the (finite) set of mixed stabilizer states[8] of n qudits, such that*

$$\frac{1}{2}\left\|\rho_{1\ldots s} - \sum_{\sigma_S} p(\sigma_S)\sigma_S^{\otimes s}\right\|_1 \leqslant 2d^{\frac{1}{2}(2n+2)^2}d^{-\frac{1}{2}(t-s)}.$$

Proof. Using Lemma 7.5, we can find a purification $|\Psi\rangle \in ((\mathbb{C}^d)^{\otimes 2n})^{\otimes t} \cong (\mathbb{C}^d)^{\otimes n \otimes t} \otimes (\mathbb{C}^d)^{\otimes n \otimes t}$ of ρ, which is invariant under the action of $O \in O_t(d)$. (In particular, $|\Psi\rangle$ is an elemenrt of the symmetric subspace.) Here we crucially use that the operators $R(O)$ for $2n$ qudits are just the second tensor powers of the corresponding operators for n qudits, as is clear from Eq. (4.12). Thus we can apply Theorem 7.2 to the state $|\Psi\rangle$. Since the local Hlibert space now contains $2n$ qudits, we obtain that there exists a probability distribution p over pure stabilizer states on $(\mathbb{C}^d)^{\otimes 2n}$ such that

$$\frac{1}{2}\left\|\Psi_{1\ldots s} - \sum_S p(S)|S\rangle\langle S|^{\otimes s}\right\|_1 \leqslant 2d^{\frac{1}{2}(2n+2)^2}d^{-\frac{1}{2}(t-s)}.$$

Taking the partial trace over the purifying systems does not increase the trace distance. Since reduced density matrices of pure stabilizer states are mixed stabilizer states, we obtain the result. $\quad\square$

---

[8] A mixed stabilizer state is a maximally mixed state on a stabilizer code (see Section 2.4).

The very same argument yields the following version of Theorem 7.3 for mixed states:

Theorem 7.7 (Stabilizer de Finetti theorem for the anti-identity). *Let ρ be a quantum state on $((\mathbb{C}^2)^{\otimes n})^{\otimes t}$ that commutes with all permutations as well as with the action of the anti-identity (1.4) on some (and hence every) subsystem consisting of six n-qubit blocks. Let $s < t$ be a multiple of six. Then there exists a probability distribution p on the (finite) set of mixed stabilizer states of n qubits, such that*

$$\frac{1}{2}\left\|\rho_{1\ldots s} - \sum_{\sigma_S} p(\sigma_S)\sigma_S^{\otimes s}\right\|_1 \leqslant 6\sqrt{2}\cdot 2^n\sqrt{\frac{s}{t}}.$$

8 Robust Hudson theorem

The methods developed in Section 3 also allow us to prove a robust version of the finite-dimensional Hudson theorem. Recall that from Eq. (2.18) that, for odd d, the Wigner function of a pure stabilizer state is necessarily nonnegative. Hudson theorem states that, for pure states, this condition is also sufficient, i.e., the Wigner function of a pure quantum state is non-negative if and only if the state is a stabilizer state [Gro06]. We will show in Theorem 8.4 that if the *Wigner* or *sum-negativity*

$$\mathrm{sn}(\psi) := \sum_{\mathbf{x}:w_\psi(\mathbf{x})<0} |w_\psi(\mathbf{x})| = \frac{1}{2}\left(\sum_{\mathbf{x}}|w_\psi(\mathbf{x})| - 1\right).$$

is small then the state is close to a stabilizer state.

The Wigner negativity is immediately related to the *mana* $\mathcal{M}(\psi) = \log(2\,\mathrm{sn}(\psi)+1)$, a monotone that plays an important role in the resource theory of stabilizer computation [Got97]. Throughout this section we assume that d is odd, so that the Wigner function is well-behaved (cf. Section 2.3).

8.1 Exact Hudson theorem

We first give a new and succinct proof of the finite-dimensional Hudson theorem. For pure states, $1 = \mathrm{tr}\,\psi^2 = \sum_{\mathbf{x}} d^n w_\psi(\mathbf{x})^2$. Thus we can define a probability distribution based on the Wigner function,

$$q_\psi(\mathbf{x}) = d^n w_\psi(\mathbf{x})^2,$$

similar to the p_ψ distribution that we defined in Eq. (3.2) via the characteristic function. Note that $0 \leqslant q_\psi(\mathbf{x}) \leqslant d^{-n}$, since $|w_\psi(\mathbf{x})| \leqslant d^{-n}$.

We now consider the sum of the absolute value of the Wigner function,

$$\|\psi\|_W := \sum_{\mathbf{x}}|w_\psi(\mathbf{x})| = d^{-n/2}\sum_{\mathbf{x}} q_\psi(\mathbf{x})^{1/2}.$$

It holds that $\|\psi\|_W \geqslant \sum_{\mathbf{x}} w_\psi(\mathbf{x}) = 1$, with equality if and only if $w_\psi(\mathbf{x}) \geqslant 0$ for all $\mathbf{x}$. By the Hölder inequality (with $p_1 = p_2 = p_3 = 3$, so $\sum_k 1/p_k = 1$):

$$1 = \sum_{\mathbf{x}} q_\psi(\mathbf{x}) = \sum_{\mathbf{x}} q_\psi(\mathbf{x})^{1/6}q_\psi(\mathbf{x})^{1/6}q_\psi(\mathbf{x})^{2/3} \leqslant \left(\sum_{\mathbf{x}} q_\psi(\mathbf{x})^{1/2}\right)^{2/3}\left(\sum_{\mathbf{x}} q_\psi(\mathbf{x})^2\right)^{1/3}.$$

Thus we obtain the following fundamental bound:

$$\sum_{\mathbf{x}} q_\psi(\mathbf{x})^2 \geqslant \frac{1}{d^n \|\psi\|_W^2}. \tag{8.1}$$

Crucially, we can interpret the left-hand side as the average of the function q_ψ with respect to the same probability distribution, $\mathbb{E}_{\mathbf{x} \sim q_\psi} q_\psi(\mathbf{x})$. Now suppose that the Wigner function is nowhere negative, so that the bound simplifies to $\sum_{\mathbf{x}} q_\psi(\mathbf{x})^2 \geqslant d^{-n}$. But $q_\psi(\mathbf{x}) \leqslant d^{-n}$ for all $\mathbf{x}$, so we conclude that the function q_ψ must be equal to d^{-n} on its support. In other words, $q_\psi(\mathbf{x})$ is the uniform distribution on a subset of cardinality d^n. This gives a rather direct proof of the finite-dimensional Hudson theorem:

Theorem 8.1 (Finite-dimensional Hudson theorem, [Gro06]). *Let d be an odd integer and ψ a pure quantum state of n qudits. Then the Wigner function of ψ is everywhere nonnegative, $w_\psi(\mathbf{x}) \geqslant 0$, if and only if ψ is a stabilizer state.*

Proof. In view of Eq. (2.18) we only need to show that if $w_\psi(\mathbf{x}) \geqslant 0$ for all $\mathbf{x}$ then ψ is a stabilizer state. By the preceding discussion, we know that $w_\psi(\mathbf{x}) = d^{-n} \mathbb{1}_T(\mathbf{x})$, where $\mathbb{1}_T$ denotes the indicator function of some subset $T \subseteq \mathcal{V}_n$ of cardinality d^n. In other words, $\langle \psi | A_\mathbf{x} | \psi \rangle = \mathbb{1}_T(\mathbf{x})$ and so $A_\mathbf{x} | \psi \rangle = | \psi \rangle$ for all $\mathbf{x} \in T$.

It remains to show that T is of the form $T = \mathbf{a} + M$, where M is a maximal isotropic subspace. For this, consider any three points $\mathbf{x}, \mathbf{y}, \mathbf{z} \in T$ and use Eq. (2.12), which asserts that $A_\mathbf{x} A_\mathbf{y} A_\mathbf{z} = \omega^{2[\mathbf{z}-\mathbf{x}, \mathbf{y}-\mathbf{x}]} A_{\mathbf{x}-\mathbf{y}+\mathbf{z}}$. Because $| \psi \rangle$ is an eigenvector of $A_\mathbf{x}, A_\mathbf{y}, A_\mathbf{z}$, with eigenvalue $+1$, we obtain

$$1 = \langle \psi | A_\mathbf{x} A_\mathbf{y} A_\mathbf{z} | \psi \rangle = \omega^{2[\mathbf{z}-\mathbf{x}, \mathbf{y}-\mathbf{x}]} \langle \psi | A_{\mathbf{x}-\mathbf{y}+\mathbf{z}} | \psi \rangle.$$

But $A_{\mathbf{x}-\mathbf{y}+\mathbf{z}}$ is Hermitian, so this is impossible unless $[\mathbf{z} - \mathbf{x}, \mathbf{y} - \mathbf{x}] = 0$. Therefore, T is the translate of an totally isotropic set M of cardinality d^n. Since the maximal size of a totally isotropic subspace is also d^n [Gro06, App. C], T is necessarily a maximal isotropic subspace. We conclude that ψ is a stabilizer state. $\square$

8.2 Robust Hudson theorem

To obtain a robust version of the Hudson theorem, we will, similarly as in our approach to stabilizer testing, combine Eq. (8.1) with an uncertainty relation that generalizes the proof of Theorem 8.1.

Lemma 8.2. *Let d be an odd integer and ψ a pure state of n qudits. Suppose that $\mathrm{tr}[\psi A_\mathbf{x}], \mathrm{tr}[\psi A_\mathbf{y}], \mathrm{tr}[\psi A_\mathbf{z}] > \sqrt{1 - 1/2d^2}$. Then $[\mathbf{z} - \mathbf{x}, \mathbf{y} - \mathbf{x}] = 0$, i.e., $W_{\mathbf{z}-\mathbf{x}}$ and $W_{\mathbf{y}-\mathbf{x}}$ must commute.*

Proof. Note that the assumption implies that

$$\| A_\mathbf{x} | \psi \rangle - | \psi \rangle \| < \sqrt{2 \left(1 - \sqrt{1 - \frac{1}{2d^2}} \right)} \leqslant \frac{1}{d},$$

and likewise for $\mathbf{y}$ and $\mathbf{z}$. Thus we obtain the following inequalities:

$$\| A_\mathbf{x} | \psi \rangle - | \psi \rangle \| < \frac{1}{d}, \quad \| A_\mathbf{y} | \psi \rangle - | \psi \rangle \| < \frac{1}{d}, \quad \| A_\mathbf{z} | \psi \rangle - | \psi \rangle \| < \frac{1}{d}.$$

As a consequence of the triangle inequality, and using $\|A_{\mathbf{z}}\| \leqslant 1$, along with Eq. (2.12), we obtain

$$\|A_{\mathbf{x}-\mathbf{y}+\mathbf{z}}\,|\psi\rangle - \omega^{-2[\mathbf{z}-\mathbf{x},\mathbf{y}-\mathbf{x}]}\,|\psi\rangle\| = \|A_{\mathbf{x}}A_{\mathbf{y}}A_{\mathbf{z}}\,|\psi\rangle - |\psi\rangle\| < \frac{3}{d}.$$

We can simply expand this relation and see that it is equivalent to

$$1 - \frac{9}{2d^2} < \langle\psi|\,A_{\mathbf{x}-\mathbf{y}+\mathbf{z}}\,|\psi\rangle \cos(2[\mathbf{z}-\mathbf{x},\mathbf{y}-\mathbf{x}]\frac{2\pi}{d}).$$

If $[\mathbf{z}-\mathbf{x},\mathbf{y}-\mathbf{x}] \neq 0$, then

$$1 - \frac{9}{2d^2} < \langle\psi|\,A_{\mathbf{x}-\mathbf{y}+\mathbf{z}}\,|\psi\rangle \cos(2[\mathbf{z}-\mathbf{x},\mathbf{y}-\mathbf{x}]\frac{2\pi}{d})$$
$$\leqslant -\cos(\frac{d-1}{2}\cdot\frac{2\pi}{d}) = \cos(\frac{\pi}{d}).$$

But, one can see that exactly for $d \geqslant 3$, $1 - \frac{9}{2d^2} \geqslant \cos(\frac{\pi}{d})$, which is contradiction. This shows that $[\mathbf{z}-\mathbf{x},\mathbf{y}-\mathbf{x}] = 0$. $\qquad\square$

Corollary 8.3. *Let d be an odd integer and ψ a pure state of n qudits. Then $T := \{\mathbf{x} \in \mathcal{V}_n : w_\psi(\mathbf{x}) > d^{-n}\sqrt{1 - 1/2d^2}\}$ is a subset of an affine totally isotropic subspace.*

We now prove the main result of this section:

Theorem 8.4 (Robust finite-dimensional Hudson theorem)**.** *Let d be odd and ψ a pure quantum state of n qudits. Then there exist a stabilizer state $|S\rangle$ such that $|\langle S|\psi\rangle|^2 \geqslant 1 - 9d^2\,\mathrm{sn}(\psi)$.*

Proof. Suppose that $\mathrm{sn}(\psi) \leqslant \varepsilon$. Then $\|\psi\|_W \leqslant 1 + 2\varepsilon$ and we find from Eq. (8.1) that

$$\sum_{\mathbf{x}} q_\psi(\mathbf{x})^2 \geqslant \frac{1}{d^n(1+2\varepsilon)^2},$$

i.e.,

$$\sum_{\mathbf{x}} q_\psi(\mathbf{x})\left(d^{-n} - q_\psi(\mathbf{x})\right) \leqslant d^{-n}\left(1 - \frac{1}{(1+2\varepsilon)^2}\right) \leqslant 4\varepsilon d^{-n}. \qquad (8.2)$$

We would like to show that the probability of the set T from Corollary 8.3 with respect to the probability distribution q_ψ is close to one. First, though, let us consider

$$\tilde{T} = \left\{\mathbf{x} : |w_\psi(\mathbf{x})| > d^{-n}\sqrt{1 - 1/2d^2}\right\} = \left\{\mathbf{x} : q_\psi(\mathbf{x}) > d^{-n}\left(1 - 1/2d^2\right)\right\}$$

which is defined just like T but for the absolute value of the Wigner function! Then, using Markov's inequality and Eq. (8.2), we have

$$\sum_{\mathbf{x}\in\tilde{T}} q_\psi(\mathbf{x}) \geqslant 1 - \frac{\sum_{\mathbf{x}} q_\psi(\mathbf{x})\left(d^{-n} - q_\psi(\mathbf{x})\right)}{d^{-n}\cdot 1/2d^2} \geqslant 1 - 8d^2\varepsilon.$$

But then T is likewise a high-probability subset:

$$\sum_{\mathbf{x}\in T} q_\psi(\mathbf{x}) \geqslant \sum_{\mathbf{x}\in\tilde{T}} q_\psi(\mathbf{x}) - \sum_{w_\psi(\mathbf{x})<0} q_\psi(\mathbf{x}) \geqslant \sum_{\mathbf{x}\in\tilde{T}} q_\psi(\mathbf{x}) - \sum_{w_\psi(\mathbf{x})<0} |w_\psi(\mathbf{x})| \geqslant 1 - 8d^2\varepsilon - \varepsilon,$$

where we used $q_\psi(\mathbf{x}) = d^n|w_\psi(\mathbf{x})|^2 \leqslant |w_\psi(\mathbf{x})|$.

As a result of Corollary 8.3, T is a subset of some affine totally isotropic subspace $\mathbf{a} + M$. If $|S\rangle$ denotes the corresponding stabilizer state then

$$|\langle\psi|S\rangle|^2 = d^n \sum_{\mathbf{x}} w_\psi(\mathbf{x})w_S(\mathbf{x}) = \sum_{\mathbf{x}\in\mathbf{a}+M} w_\psi(\mathbf{x}) = \sum_{\mathbf{x}\in T} w_\psi(\mathbf{x}) + \sum_{\mathbf{x}\in(\mathbf{a}+M)\setminus T} w_\psi(\mathbf{x}) \geqslant \sum_{\mathbf{x}\in T} w_\psi(\mathbf{x}) - \varepsilon$$

$$\geqslant \sum_{\mathbf{x}\in T} q_\psi(\mathbf{x}) - \varepsilon \geqslant 1 - (8d^2 + 2)\varepsilon > 1 - 9d^2\varepsilon.$$

In the fifth step we used that, for $\mathbf{x} \in T$, $w_\psi(\mathbf{x}) = |w_\psi(\mathbf{x})| \geqslant d^n|w_\psi(\mathbf{x})|^2 = q_\psi(\mathbf{x})$. $\qquad\square$

8.3 Stabilizer testing revisited: minimal number of copies

We will now revisit stabilizer testing from the perspective of the Wigner function and show that for $d \equiv 1, 5 \pmod 6$ it is in fact possible to perform stabilizer testing with just three copies of the state. This is clearly optimal, since the set of stabilizer states forms a projective 2-design.

We start with the phase space point operators $A_\mathbf{x}$ from (2.11), Consider the operator

$$V := d^{-n} \sum_{\mathbf{x}} A_\mathbf{x}^{\otimes 3} = d^{-2n} \sum_{\mathbf{y}_1+\mathbf{y}_2+\mathbf{y}_3=0} W_{\mathbf{y}_1} \otimes W_{\mathbf{y}_2} \otimes W_{\mathbf{y}_3}. \tag{8.3}$$

We remark that it is clear from Eq. (2.13) that V is an element of the commutant. Moreover, we have the following analog of Lemma 3.8:

Lemma 8.5. *For $d \equiv 1, 5 \pmod 6$, the operator V defined in Eq. (8.3) is a Hermitian unitary.*

Proof. Since the operators $A_\mathbf{x}$ are Hermitian, V is Hermitian as well. Thus it remains to prove that V is unitary. For this we compute:

$$V^2 = d^{-4n} \sum_{\mathbf{y}_1+\mathbf{y}_2+\mathbf{y}_3=0,\, \mathbf{z}_1+\mathbf{z}_2+\mathbf{z}_3=0} W_{\mathbf{y}_1}W_{\mathbf{z}_1} \otimes W_{\mathbf{y}_2}W_{\mathbf{z}_2} \otimes W_{\mathbf{y}_3}W_{\mathbf{z}_3}$$

$$= d^{-4n} \sum_{\mathbf{a}_1+\mathbf{a}_2+\mathbf{a}_3=0,\, \mathbf{b}_1+\mathbf{b}_2+\mathbf{b}_3=0} W_{\mathbf{a}_1} \otimes W_{\mathbf{a}_2} \otimes W_{\mathbf{a}_3} \omega^{\frac{1}{8}[\mathbf{a}_1+\mathbf{b}_1,\mathbf{a}_1-\mathbf{b}_1]+\frac{1}{8}[\mathbf{a}_2+\mathbf{b}_2,\mathbf{a}_2-\mathbf{b}_2]+\frac{1}{8}[\mathbf{a}_3+\mathbf{b}_3,\mathbf{a}_3-\mathbf{b}_3]}$$

$$= d^{-4n} \sum_{\mathbf{a}_1+\mathbf{a}_2+\mathbf{a}_3=0,\, \mathbf{b}_1,\mathbf{b}_2} W_{\mathbf{a}_1} \otimes W_{\mathbf{a}_2} \otimes W_{\mathbf{a}_3} \omega^{-\frac{1}{4}[\mathbf{a}_1-\mathbf{a}_3,\mathbf{b}_1]-\frac{1}{4}[\mathbf{a}_2-\mathbf{a}_3,\mathbf{b}_2]}$$

$$= \sum_{\mathbf{a}_1+\mathbf{a}_2+\mathbf{a}_3=0} \delta_{\mathbf{a}_1,\mathbf{a}_3}\delta_{\mathbf{a}_2,\mathbf{a}_3} W_{\mathbf{a}_1} \otimes W_{\mathbf{a}_2} \otimes W_{\mathbf{a}_3} = I,$$

where, for the second equality, we used the change of variables $\mathbf{a}_i = \mathbf{y}_i + \mathbf{z}_i$ and $\mathbf{b}_i = \mathbf{y}_i - \mathbf{z}_i$. $\qquad\square$

We now consider the binary POVM measurement with accepting projector

$$\Pi_{\text{accept}} = \frac{1}{2}(I + V).$$

Theorem 8.6 (Stabilizer testing from three copies). *Let $d \equiv 1, 5 \pmod 6$ and ψ a pure state of n qudits. Denote by $p_{accept} = \text{tr}[\psi^{\otimes 3}\Pi_{accept}]$ the probability that the POVM element Π_{accept} accepts given three copies of ψ. If ψ is a stabilizer state then it accepts with certainty, $p_{accept} = 1$. On the other hand, if $\max_S|\langle S|\psi\rangle|^2 \leqslant 1 - \varepsilon^2$ then $p_{accept} \leqslant 1 - \varepsilon^2/16d^2$.*

Proof. We first note that

$$p_{accept} = tr\left[\Pi_{accept}\psi^{\otimes 3}\right] = \frac{1}{2}\left(1 + d^{-n}\sum_x tr\left[A_x^{\otimes 3}\psi^{\otimes 3}\right]\right) = \frac{1}{2}\left(1 + d^{2n}\sum_x w_\psi^3(x)\right),$$

where w_ψ denotes the Wigner function defined in Eq. (2.10). It is clear from Eq. (2.18) that if ψ is a stabilizer state then $p_{accept} = 1$.

Now assume that ψ is an arbitrary pure state. Since $q(x) = d^n w_\psi^2(x)$ is a probability distribution, we can rewrite the above as

$$\sum_x q_\psi(x)\left(1 - d^n w_\psi(x)\right) = 2\left(1 - p_{accept}\right).$$

Moreover, $q(x) \leqslant d^{-n}$ for all x, so we can use Markov's probability for the set

$$T = \left\{x \in \mathcal{V}_n : w_\psi(x) > d^{-n}\sqrt{1 - 1/2d^2}\right\}$$

to see that

$$\sum_{x \in T} q_\psi(x) \geqslant 1 - \frac{2\left(1 - p_{accept}\right)}{1 - \sqrt{1 - 1/2d^2}} \geqslant 1 - 8d^2\left(1 - p_{accept}\right). \tag{8.4}$$

We now argue similarly as in the proof of the robust Hudson theorem (Theorem 8.4). From Corollary 8.3 below we know that there exists an affine Lagrangian subspace $a + M$ that contains T. Let $|S\rangle$ denote the corresponding stabilizer state. Then,

$$|\langle\psi|S\rangle|^2 = d^n\sum_x w_\psi(x)w_S(x) = \sum_{x \in a+M} w_\psi(x) = \sum_{x \in T} w_\psi(x) + \sum_{x \in (a+M)\setminus T} w_\psi(x)$$

For $x \in T$, $w_\psi(x) \geqslant 0$ and so $w_\psi(x) \geqslant q_\psi(x)$. Thus the first sum can be lower-bounded by using Eq. (8.4). For the second sum, we note that Eq. (8.4) also implies that $1 - 8d^2\left(1 - p_{accept}\right) \leqslant d^{-n}|T|$, since $q_\psi(x) \leqslant d^{-n}$, and so

$$\sum_{x \in (a+M)\setminus T} w_\psi(x) \geqslant -d^{-n}|(a + M)\setminus T| = -d^{-n}\left(d^n - |T|\right) \geqslant -8d^2\left(1 - p_{accept}\right).$$

Together, we obtain that $|\langle\psi|S\rangle|^2 \geqslant 1 - 16d^2\left(1 - p_{accept}\right)$, or $p_{accept} \leqslant 1 - \varepsilon^2/16d^2$, which is what wanted to show. $\qquad\square$

Acknowledgments

References

[AG04]   Scott Aaronson and Daniel Gottesman. Improved simulation of stabilizer circuits. *Phys. Rev. A*, 70(5):052328, 2004. `arXiv:quant-ph/0406196, doi:10.1103/PhysRevA.70.052328`.

[AG08]   Scott Aaronson and Daniel Gottesman. Identifying stabilizer states. 2008. URL: http://pirsa.org/08080052/.

[App05]   D. M. Appleby. SIC-POVMs and the Extended Clifford Group. *J. Math. Phys.*, 46:052107, 2005. `arXiv:quant-ph/0412001, doi:10.1063/1.1896384`.

[BCHW16]   Fernando GSL Brandao, Matthias Christandl, Aram W Harrow, and Michael Walter. The mathematics of entanglement. 2016. `arXiv:1604.01790`.

[BCY11]   Fernando GSL Brandão, Matthias Christandl, and Jon Yard. A quasipolynomial-time algorithm for the quantum separability problem. In *Proc. STOC*, pages 343–352. ACM, 2011. `arXiv:1011.2751, doi:10.1145/1993636.1993683`.

[BFNR03]   Harry Buhrman, Lance Fortnow, Ilan Newman, and Hein Röhrig. Quantum property testing. In *Proc. SODA*, pages 480–488, 2003.

[BFNR08]   Harry Buhrman, Lance Fortnow, Ilan Newman, and Hein Röhrig. Quantum property testing. *SIAM J. Comput.*, 5(37):1387–1400, 2008. `arXiv:quant-ph/0201117, doi:10.1137/S0097539704442416`.

[BH13]   Fernando GSL Brandao and Aram W Harrow. Quantum de Finetti theorems under local measurements with applications. In *Proc. STOC*, pages 861–870, 2013.

[BH17]   Fernando GSL Brandao and Aram W Harrow. Quantum de Finetti theorems under local measurements with applications. *Commun. Math. Phys.*, 353(2):469–506, 2017. `arXiv:1210.6367, doi:10.1007/s00220-017-2880-3`.

[CFS02]   Carlton M Caves, Christopher A Fuchs, and Rüdiger Schack. Unknown quantum states: the quantum de Finetti representation. *Journal of Mathematical Physics*, 43(9):4537–4559, 2002. `arXiv:quant-ph/0104088, doi:10.1063/1.1494475`.

[CHM07]   Matthias Christandl, Aram W Harrow, and Graeme Mitchison. Nonzero Kronecker coefficients and what they tell us about spectra. *Commun. Math. Phys.*, 270(3):575–585, 2007. `arXiv:quant-ph/0511029, doi:10.1007/s00220-006-0157-3`.

[CKMR07]   Matthias Christandl, Robert König, Graeme Mitchison, and Renato Renner. One-and-a-half quantum de Finetti theorems. *Commun. Math. Phys.*, 273(2), 2007. `arXiv:quant-ph/0602130, doi:10.1007/s00220-007-0189-3`.

[CM06]   Matthias Christandl and Graeme Mitchison. The spectra of quantum states and the Kronecker coefficients of the symmetric group. *Commun. Math. Phys.*, 261(3):789–797, 2006. `arXiv:quant-ph/0409016, doi:10.1007/s00220-005-1435-1`.

[CS96]   A Robert Calderbank and Peter W Shor. Good quantum error-correcting codes exist. *Phys. Rev. A*, 54(2):1098, 1996. `arXiv:quant-ph/9512032, doi:10.1103/PhysRevA.54.1098`.

[Dam18]   Raja Damanik. Optimality in stabilizer testing. 2018. URL: https://eprints.illc.uva.nl/1622/1/MoL-2018-09.text.pdf.

[DB13]   Niel De Beaudrap. A linearized stabilizer formalism for systems of finite dimension. *Quant. Inf. Comp.*, 13:73– 115, 2013. `arXiv:1102.3354`.

[DF80]   Persi Diaconis and David Freedman. Finite exchangeable sequences. *Ann. Prob.*, pages 745–764, 1980. `doi:10.1214/aop/1176994663`.

[DOS07]   Christian D'Cruz, Tobias J Osborne, and Rüdiger Schack. Finite de Finetti theorem for infinite-dimensional systems. *Phys. Rev. Lett.*, 98(16):160406, 2007. `arXiv:quant-ph/0606139, doi:10.1103/PhysRevLett.98.160406`.

[Eke91]   Artur K Ekert. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.*, 67(6):661, 1991. `doi:10.1103/PhysRevLett.67.661`.

[Far14]   JM Farinholt. An ideal characterization of the Clifford operators. *J. Phys. A*, 47(30):305303, 2014.

[GE08]   David Gross and Jens Eisert. Quantum Margulis expanders. *Quant. Inf. Comp.*, 8:722, 2008. `arXiv:0710.0651`.

[GH16]   Shamgar Gurevich and Roger Howe. Small representations of finite classical groups. 2016. `arXiv:1609.01276`.

[Got97]   Daniel Gottesman. *Stabilizer codes and quantum error correction*. PhD thesis, 1997. `arXiv:quant-ph/9705052`.

[Got99]   Daniel Gottesman. The Heisenberg representation of quantum computers. In *Proc. Group22*, pages 32–43, 1999. `arXiv:quant-ph/9807006`.

[Gro06]   David Gross. Hudson's theorem for finite-dimensional quantum systems. *J. Math. Phys.*, 47(12):122107, 2006. `arXiv:quant-ph/0602001, doi:10.1063/1.2393152`.

[GS16]   Daniel Grier and Luke Schaeffer. The classification of stabilizer operations over qubits. 2016. `arXiv:1603.03999`.

[GW13]   David Gross and Michael Walter. Stabilizer information inequalities from phase space distributions. *J. Math. Phys.*, 54(8):082201, 2013. `arXiv:1302.6990, doi:10.1063/1.4818950`.

[Har05]   Aram W Harrow. *Applications of coherent classical communication and the Schur transform to quantum information theory*. PhD thesis, 2005. `arXiv:quant-ph/0512255`.

[Has08]   Matthew B Hastings. A counterexample to additivity of minimum output entropy. *Nat. Phys.*, 5, 2008. `arXiv:0809.3972, doi:10.1038/nphys1224`.

[HHJ+17]   Jeongwan Haah, Aram W Harrow, Zhengfeng Ji, Xiaodi Wu, and Nengkun Yu. Sample-optimal tomography of quantum states. *IEEE Trans. Inf. Theo.*, 63(9):5628–5641, 2017. `arXiv:1508.01797`, `doi:10.1109/TIT.2017.2719044`.

[HLW06]   Patrick Hayden, Debbie W Leung, and Andreas Winter. Aspects of generic entanglement. *Commun. Math. Phys.*, 265(1):95–117, 2006. `arXiv:quant-ph/0407049`, `doi:10.1007/s00220-006-1535-6`.

[HM76]   Robin L Hudson and Graham R Moody. Locally normal symmetric states and an analogue of de Finetti's theorem. *Z. Wahrsch. verw. Geb.*, 33(4):343–351, 1976. `doi:10.1007/BF00534784`.

[HM02]   Masahito Hayashi and Keiji Matsumoto. Quantum universal variable-length source coding. *Phys. Rev. A*, 66(2):022311, 2002. `arXiv:quant-ph/0202001`, `doi:10.1103/PhysRevA.66.022311`.

[HM15]   Masahito Hayashi and Tomoyuki Morimae. Verifiable measurement-only blind quantum computing with stabilizer testing. *Phys. Rev. Lett.*, 115(22):220502, 2015. `arXiv:1505.07535`, `doi:10.1103/PhysRevLett.115.220502`.

[HNQ+16]   Patrick Hayden, Sepehr Nezami, Xiao-Liang Qi, Nathaniel Thomas, Michael Walter, and Zhao Yang. Holographic duality from random tensor networks. *J. High Energy Phys.*, 2016(11):9, 2016. `arXiv:1601.01694`, `doi:10.1007/JHEP11(2016)009`.

[How73]   Roger Howe. Invariant theory and duality for classical groups over finite fields with applications to their singular representation theory. Preprint, 1973.

[HWFW17]   Jonas Helsen, Joel J Wallman, Steven T Flammia, and Stephanie Wehner. Multi-qubit randomized benchmarking using few samples. 2017. `arXiv:1701.04299`.

[HWVE14]   Mark Howard, Joel J Wallman, Victor Veitch, and Joseph Emerson. Contextuality supplies the magic for quantum computation. *Nature*, 510(7505):351–355, 2014. `arXiv:1401.4174`, `doi:10.1038/nature13460`.

[HWW16]   Jonas Helsen, Joel J Wallman, and Stephanie Wehner. Representations of the multi-qubit Clifford group. 2016. `arXiv:1609.08188`.

[KG15]   Richard Kueng and David Gross. Qubit stabilizer states are complex projective 3-designs. 2015. `arXiv:1510.02767`.

[KM09]   Robert Koenig and Graeme Mitchison. A most compendious and facile quantum de Finetti theorem. *J. Math. Phys.*, 50(1):012105, 2009. `arXiv:quant-ph/0703210`, `doi:10.1063/1.3049751`.

[KR05]   Robert König and Renato Renner. A de Finetti representation for finite symmetric quantum states. *J. Math. Phys.*, 46(12):122108, 2005. `arXiv:quant-ph/0410229`, `doi:10.1063/1.2146188`.

[KR08]   Hari Krovi and Martin Rötteler. An efficient quantum algorithm for the hidden subgroup problem over Weyl-Heisenberg groups. In *Mathematical Methods in Computer Science*, pages 70–88. Springer, 2008.

[KW01]   Michael Keyl and Reinhard F Werner. Estimating the spectrum of a density operator. *Phys. Rev. A*, 64(5):052311, 2001. `arXiv:quant-ph/0102027`, `doi:10.1103/PhysRevA.64.052311`.

[KZG16]   Richard Kueng, Huangjun Zhu, and David Gross. Low rank matrix recovery from Clifford orbits. 2016. `arXiv:1610.08070`.

[LC09]   Anthony Leverrier and Nicolas J Cerf. Quantum de Finetti theorem in phase-space representation. *Phys. Rev. A*, 80(1):010102, 2009. `arXiv:0904.4862`, `doi:10.1103/PhysRevA.80.010102`.

[Lev16]   Anthony Leverrier. SU(p, q) coherent states and a Gaussian de Finetti theorem. 2016. `arXiv:1612.05080`.

[Lev17]   Anthony Leverrier. Security of continuous-variable quantum key distribution via a Gaussian de Finetti reduction. *Phys. Rev. Lett.*, 118(20):200501, 2017. `arXiv:1701.03393`, `doi:10.1103/PhysRevLett.118.200501`.

[LGPRC13] Anthony Leverrier, Raúl García-Patrón, Renato Renner, and Nicolas J Cerf. Security of continuous-variable quantum key distribution against general attacks. *Phys. Rev. Lett.*, 110(3):030502, 2013. `arXiv:1208.4920`, `doi:10.1103/PhysRevLett.110.030502`.

[Low09]   Richard A Low. Learning and testing algorithms for the Clifford group. *Phys. Rev. A*, 80(5):052314, 2009. `arXiv:0907.2833`, `doi:10.1103/PhysRevA.80.052314`.

[MdW13]   Ashley Montanaro and Ronald de Wolf. A survey of quantum property testing. 2013. `arXiv:1310.2035`.

[ME12]   Andrea Mari and Jens Eisert. Positive Wigner functions render classical simulation of quantum computation efficient. *Phys. Rev. Lett.*, 109(23):230503, 2012. `arXiv:1208.3660`, `doi:10.1103/PhysRevLett.109.230503`.

[MGE11]   Easwar Magesan, Jay M Gambetta, and Joseph Emerson. Scalable and robust randomized benchmarking of quantum processes. *Phys. Rev. Lett.*, 106(18):180504, 2011. `arXiv:1009.3639`, `doi:10.1103/PhysRevLett.106.180504`.

[MMG19]   Felipe Montealegre-Mora and David Gross. Rank-deficient representations in Howe duality over finite fields arise from quantum codes. 2019. In preparation.

[Mon17]   Ashley Montanaro. Learning stabilizer states by Bell sampling. 2017. `arXiv:1707.04012`.

[NBD$^+$02] Michael A Nielsen, Michael J Bremner, Jennifer L Dodd, Andrew M Childs, and Christopher M Dawson. Universal simulation of hamiltonian dynamics for quantum systems with finite-dimensional state spaces. *Physical Review A*, 66(2):022317, 2002.

[NOP09]   Miguel Navascues, Masaki Owari, and Martin B Plenio. Power of symmetric extensions for entanglement detection. *Phys. Rev. A*, 80(5):052306, 2009. `arXiv:0906.2731`, `doi:10.1103/PhysRevA.80.052306`.

[NRS06]   Gabriele Nebe, Eric M Rains, and Neil James Alexander Sloane. *Self-dual codes and invariant theory*, volume 17. Springer, 2006.

[NW16]   Sepehr Nezami and Michael Walter. Multipartite entanglement in stabilizer tensor networks. 2016. `arXiv:1608.02595`.

[OW15]   Ryan O'Donnell and John Wright. Quantum spectrum testing. In *Proc. STOC*, pages 529–538. ACM, 2015. `arXiv:1501.05028`.

[OW16]   Ryan O'Donnell and John Wright. Efficient quantum tomography. In *Proc. STOC*, pages 899–912. ACM, 2016. `arXiv:1508.01907`.

[OW17]   Ryan O'Donnell and John Wright. Efficient quantum tomography II. In *Proc. STOC*, pages 962–974. ACM, 2017. `arXiv:1612.00034`.

[Pet90]   Dénes Petz. A de Finetti-type theorem with m-dependent states. *Probability theory and related fields*, 85(1):65–72, 1990. `doi:10.1007/BF01377629`.

[Ren05]   Renato Renner. *Security of quantum key distribution*. PhD thesis, 2005. `arXiv:quant-ph/0512258`.

[Ren07]   Renato Renner. Symmetry of large physical systems implies independence of subsystems. *Nat. Phys.*, 3(9):645–649, 2007. `arXiv:quant-ph/0703069, doi:10.1038/nphys684`.

[RW89]   GA Raggio and RF Werner. Quantum statistical mechanics of general mean field systems. *Helv. Phys. Acta*, 62(8):980–1003, 1989. `doi:10.5169/seals-116175`.

[Sch11]   Wolfgang P Schleich. *Quantum optics in phase space*. John Wiley & Sons, 2011.

[Ste96a] Andrew Steane. Multiple-particle interference and quantum error correction. In *Proc. Royal Soc. A*, volume 452, pages 2551–2577. The Royal Society, 1996. `arXiv:quant-ph/9601029, doi:10.1098/rspa.1996.0136`.

[Ste96b] Andrew M Steane. Error correcting codes in quantum theory. *Phys. Rev. Lett.*, 77(5):793, 1996. `doi:10.1103/PhysRevLett.77.793`.

[Stø69]   Erling Størmer. Symmetric states of infinite tensor products of C*-algebras. *J. Func. Anal.*, 3(1):48–68, 1969. `doi:10.1016/0022-1236(69)90050-0`.

[SZ84]   PD Seymour and Thomas Zaslavsky. Averaging sets: a generalization of mean values and spherical designs. *Adv. Math.*, 52(3):213–240, 1984. `doi:10.1016/0001-8708(84)90022-7`.

[vdNDdM05] Maarten van den Nest, Jeroen Dehaene, and Bart de Moor. Invariants of the local clifford group. *Physical Review A*, 71(2):022310, 2005.

[VFGE12] Victor Veitch, Christopher Ferrie, David Gross, and Joseph Emerson. Negative quasi-probability as a resource for quantum computation. *New J. Phys.*, 14(11):113011, 2012. `arXiv:1201.1256, doi:10.1088/1367-2630/14/11/113011`.

[VMGE14]   Victor Veitch, SA Hamed Mousavian, Daniel Gottesman, and Joseph Emerson. The resource theory of stabilizer quantum computation. *New J. Phys.*, 16(1):013009, 2014. `arXiv:1307.7171, doi:10.1088/1367-2630/16/1/013009`.

[Wan11]   Guoming Wang. Property testing of unitary operators. *Phys. Rev. A*, 84(5):052328, 2011. `arXiv:1110.1133, doi:10.1103/PhysRevA.84.052328`.

[Web16]   Zak Webb. The Clifford group forms a unitary 3-design. *Quant. Inf. Comp.*, 26:1379–1400, 2016. `arXiv:1510.02769`.

[Wil09]   Robert Wilson. *The finite simple groups*, volume 251. Springer Science & Business Media, 2009.

[Woo87]   William K Wootters. A Wigner-function formulation of finite-state quantum mechanics. *Ann. Phys. (N.Y.)*, 176(1):1–21, 1987. `doi:10.1016/0003-4916(87)90176-X`.

[Woo93]   Jay A Wood. Witt's extension theorem for mod four valued quadratic forms. *Trans. Amer. Math. Soc.*, 336(1):445–461, 1993.

[Zhu15]   Huangjun Zhu. Multiqubit Clifford groups are unitary 3-designs. 2015. `arXiv:1510.02619`.

[ZKGG16]   Huangjun Zhu, Richard Kueng, Markus Grassl, and David Gross. The Clifford group fails gracefully to be a unitary 4-design. 2016. `arXiv:1609.08172`.

[ZPDF16]   Liming Zhao, Carlos A Pérez-Delgado, and Joseph F Fitzsimons. Fast graph operations in quantum computation. *Phys. Rev. A*, 93(3):032314, 2016. `arXiv:1510.03742, doi:10.1103/PhysRevA.93.032314`.