Stabilizer and numerical representation theory:

Dualities, algorithms and applications

Inaugural-Dissertation zur Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. nat.)

in Theoretische Physik

der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln
vorgelegt von

Felipe Montealegre Mora

aus San José, Costa Rica

December 16, 2021

Advisor, first reviewer and examiner	Prof. Dr. David Gross
Second reviewer and examiner	Prof. Dr. Frank Vallentin
Chair of the thesis advisory committee	Prof. Dr. Johannes Berg

Abstract

Symmetry arguments and representation theoretical methods are widely used in theoretical physics. Quantum information theory is no exception to this: Here representation theory has been used, for example, to guide the development of a variety of quantum computational protocols. In this thesis, I present a series of contributions of this sort—representation theory applied to quantum information. These contributions are divided into two parts.

Part I deals with the representation theory arising from the *stabilizer formalism*. Having originated within the context of quantum error correction, the stabilizer formalism quickly found more applications and today it is a vital tool in many areas of quantum information theory. Two important objects that appear in this formalism are the *Clifford group* and the *oscillator representation* of the symplectic group. In particular, tensor powers of the Clifford group and of the oscillator representation have recently come to the spotlight due to their applications in, for instance, quantum device characterization and the simulation of quantum computing. My contributions in this first part, Chapters 2 and 3, are in the context of these representations.

In Chapter 2, I study these representations through the lens of Howe duality. Here, I generalize the recently developed theory of the η *correspondence* to provide a full decomposition for tensor power representations (TPRs). This theory has been previously used to understand a certain "maximal rank" sector of oscillator TPRs. In other words, the η correspondence has been used to partially decompose oscillator tensor powers. I show that not only can this formalism be used to fully decompose oscillator TPRs, but that it can be generalized in order to fully decompose Clifford TPRs.

In Chapter 3, I present a new efficient construction of *approximate unitary t-designs*. Unitary t-designs are probability distributions on the unitary group that emulate the first t moments of the Haar distribution. The circuits arising in the construction shown here are dominated by Clifford unitaries, with a number $\tilde{O}(t^4)$ of non-Clifford gates that does not depend on the system size (e.g. the number of qubits). The proofs in this chapter build heavily on recent results which characterize the commutant of Clifford TPRs.

Part II focuses on algorithms for the numerical decomposition of representations. My motivation here comes from semi-definite programming, where such algorithms can be used to significantly reduce the dimension of the optimization problem. The working principle of the package RepLAB, which tackles this problem, is presented in Chapter 5. In that chapter, I moreover show that under certain simplifying assumptions, this algorithm is stable against numerical perturbations.

While the results in Chapter 5 strongly suggest that the output of RepLAB is correct, they do not rigorously prove this statement. In Chapter 6, I address this issue. I

provide an algorithm that certifies whether a numerical decomposition of a representation is close to exact. This certifying algorithm has rigorous performance guarantees. I have coded it into the Python package RepCert, which I present and benchmark in Chapter 7.

Taken together, RepLAB and RepCert may be used to decompose representations of compact groups in such a way that the output is guaranteed to be correct. The runtime of this combination depends on several factors: the complexity of taking products in the group, the largest dimension d of an irreducible block in the decomposition, and, most importantly, the dimension n of the representation. In the realistic scenario where the group product is cheap to compute and $d \ll n$, the runtime is dominated by its dependence on n, scaling as $\tilde{O}(n^3)$.

Acknowledgements

The only possible beginning to these acknowledgements is by thanking my thesis advisor, David Gross. Through long discussions on a board, over lunch, over coffee or with a beer, I have learned from you more than I could have imagined in my most optimistic dreams. In your research group I have learned mathematics, yes, but, more broadly, your insights have taught me how analysis and empathy play a part in science, in politics, and in personal life. I cherish these lessons as a scientist and as a person. Finally, your energy has permeated through the dynamics of the whole research group. I could not have chosen a better place to do my PhD.

This naturally leads me to the next acknowledgement: thanking the quantum research group in Cologne. I will not list all of your names out of fear that I would forgetfully leave anyone out, but you know who you are. I feel both lucky and humbled to have been part of such a group of skilled, funny, and intensely passionate people. I look up to you and am thankful for the stimulating environment you have all contributed to over these years. I hope to continue seeing all of you in the future, as scientists and as friends. Moreover, I want to specifically thank Mariela and Markus for saving the day every time that the world seemed to be burning.

Over these years, I have met and collaborated with some amazing scientists who have played a big role in the results presented here. I would particularly like to acknowledge three groups of scientists: Denis and Jean-Daniel; Jonas, Markus, Jens and Ingo; and Arne and Frank. From you I have learned many things in mathematics, in programming, and outside those fields as well. Thank you for including me in these exciting projects and for your hospitality while hosting me. There were also some people that, while we did not collaborate on any project, never failed to inspire me with their insights and put a smile on my face whenever we met. Askery, Kerstin (back at you!), Richard, this one goes out especially for you.

Ari, Arne, Chae-Yeun, David, Johan, Lorenz, Mari, Markus, Oscar and Vahide: thank you so much for your comments and feedback on this thesis. The care you took in reading through this document has made it much better in many regards. If this document is well-written, then it is very much thanks to you all.

To the close friends who saw me through thick and thin during my grad school years I send a humongous hug and an acknowledgement of the part they played in this thesis. Oscar, Lorenz, Janin, Mary, Shreyasi: without you people, the highs would have been much less beautiful and the lows would have been dreadful. Please never stop being the weird, crazy, loving people that you are. From the deepest part of my heart I thank you for always being there.

Finalmente, mi familia: Jime, Ma (Grace), y Pa (Alejandro). En estas páginas, aunque estén llenas de ecuaciones y símbolos raros, hay un poquito de cada uno de

ustedes. Gracias a ustedes soy la persona que soy. De ustedes aprendí a ser tenaz y paciente para llevar a cabo proyectos como el que está en estas páginas. Ustedes me apoyaron hasta en mis momentos más vulnerables y me celebraron en mis éxitos. No puedo resumir en un sólo párrafo todo lo que les quiero agradecer. Por eso, con todo el amor que tengo en el corazón – y hasta un poquito más – quiero dedicarles esta tésis a ustedes tres. Los amo.

Contents

Notation and conventions

- If V is a vector space, the space of linear functions $V \to V$ (*endomorphisms* of V) is denoted $\mathrm{End}(V)$.

- Schatten p-norms of operators on a finite dimensional Hilbert space are denoted $\| \cdot \|_p$, so that

$$\|X\|_p^p = \mathrm{tr}\Big(\Big(\sqrt{X^\dagger X}\Big)^p\Big) = \sum_i |x_i|^p,$$

  where $\sqrt{X^\dagger X}$ is the unique square root of the positive matrix $X^\dagger X$, and where x_i are the singular values of X. In particular, the *operator norm* is given by

$$\|X\|_\infty = \max_{|\psi\rangle} \frac{\langle\psi| X |\psi\rangle}{\langle\psi|\psi\rangle}.$$

  The *Frobenius norm* is the Schatten 2-norm and will be denoted by $\| \cdot \|_2$ and $\| \cdot \|_F$ interchangeably.

- Consider two vector spaces A and B, with corresponding norms $\| \cdot \|_A$ and $\| \cdot \|_B$. The space of linear operators $M : A \to B$ hosts an *induced norm* $\| \cdot \|_{A\to B}$ given by

$$\|M\|_{A\to B} = \max_{a\in A} \|Ma\|_B \quad \text{s.t.} \quad \|a\|_A = 1.$$

  Most relevant to this thesis will be the case where A and B are spaces of operators on a Hilbert space. These operator spaces are endowed with the Schatten 1-norm (also known as the *trace norm*). The *diamond norm* of $M \in \mathrm{End}(\mathrm{End}(\mathbb{C}^n))$ is given by

$$\|M\|_\diamond = \|M \otimes \mathbb{1}_{n^2}\|_{1\to 1}.$$

- The *max norm*, known also as the vector ℓ_∞ norm, on a vector space $\mathbb{C}^n$ or $\mathbb{R}^n$ is given by

$$\|v\|_{\max} = \max_i |v_i|.$$

- The *commutant* of a subalgebra $A \subseteq \mathrm{End}(\mathbb{C}^n)$ is

$$A' = \{X \in \mathrm{End}(\mathbb{C}^n) \mid Xa - aX = 0 \quad \forall\, a \in A\}.$$

  In the case where A is spanned by a subgroup $G \subseteq \mathrm{Gl}(n)$, we denote its com-

mutant by G'.

- The ring of integers with addition and multiplication computed modulo d is denoted $\mathbb{Z}_d := \mathbb{Z}/d\mathbb{Z}$. When d is a prime, this ring is actually a field, which will sometimes referred to alternatively as $\mathbb{F}_d$.

- If a matrix M acts on a vector space V, then for any subset $S \subseteq V$ I denote $MS = \{Ms \mid s \in S\}$.

- If G is a compact group, a subset $\{g_i\} \subseteq G \subseteq \mathbb{C}^{n\times n}$ is said to *generate* G if the group $\langle\{g_i\}\rangle$ formed by generator words of arbitrary length is dense in G.

- If ρ is a representation of some group G, I will sometimes use ρ to denote the representation *space*. Moreover, I denote by $\mathrm{Irr}\ G$ the set of equivalence classes irreducible complex representations of G. Two representations $\rho,\ \rho'$ with representation spaces $\mathcal{H}_\rho$, $\mathcal{H}_{\rho'}$ are equivalent, denoted $\rho \simeq \rho'$, if there is a invertible linear map $M : \mathcal{H}_\rho \to \mathcal{H}_{\rho'}$ for which $M\rho(g)M^{-1} = \rho'(g)$ for all $g \in G$. Any complex representation is assumed to be unitary unless otherwise stated.

Introduction to this thesis

This thesis is the result of a series of projects with a common thread: the development of representation theoretical tools for applications in quantum information science. These projects can be roughly divided into two thematic clusters.

Stabilizer representation theory. The first cluster revolves around the *stabilizer formalism*. The stabilizer formalism, arguably the most well understood approach to quantum error correction, has become a staple mathematical tool in quantum information theory. The two central objects in this formalism are the set of *stabilizer states* in $(\mathbb{C}^d)^{\otimes n}$ and their symmetry, the *Clifford group* $\mathrm{Cl} \subset \mathrm{U}(d^n)$. The stabilizer formalism lies at a useful middle ground in the trade-off between complexity and expressibility. On the one hand, Clifford unitaries and stabilizer states give rise to a wealth of quantum phenomena, while on the other hand, they have very convenient algebraic and geometric properties that allow them to be efficiently described and manipulated. In short, the stabilizer formalism is simple yet expressive. This intersection allows its use in a variety of situations—as a toy model for condensed matter [BDCP12] and high energy systems [DS21], as a foundation for quantum error correction [NC10, Chap. 10] and as a tool in classical [NRS06] error correction, as a primitive for the simulation of quantum computing [AG04, BBC$^+$19] and quantum device characterization [KR21], and as a tool in quantum entanglement theory [NW16]. Moreover, the stabilizer formalism is connected to other mathematical topics like automorphic forms [Gel06], lattice theory [NRS01] and Howe duality [GH20, How10].

Given this context, understanding the structures arising from the stabilizer formalism is, in and of itself, an interesting scientific problem. Part I of this thesis presents a contribution to this program. My emphasis here is mainly on representation theoretical aspects associated to the stabilizer formalism.

In Chap. 2, I study tensor power representations of the Clifford group and of the closely related oscillator representation μ of the finite symplectic group $\mathrm{Sp}(\mathbb{F}_d^{2n})$ (also known as Weil or Schroedinger representation). The aforementioned chapter contains two sections, the first is published as [MMG21a], while the second is based on the draft [MMG21b]. These contributions extend the recently developed theory of the η correspondence [GH17, GH20].

The formalism of the η correspondence builds on a construction that assigns a certain "rank" to each irreducible representation (in the following *irrep* for short) of $\mathrm{Sp}(\mathbb{F}_d^{2n})$ [How10]. The tensor power representation $\mu^{\otimes t}$ contains irreps with rank $\leq t$ and gives rise to a correspondence, η, between irreps of a finite orthogonal group $O(\mathbb{F}_d^t)$ to irreps of $\mathrm{Sp}(\mathbb{F}_d^{2n})$ with rank t. This allows one to fully decompose the "maximal rank" subspace, i.e., the subspace spanned by the rank t irreducible subrepresentations of $\mu^{\otimes t}$.

In the first section, based on Ref. [MMG21a], I provide a characterization of the rank deficient sectors in $\mu^{\otimes t}$, i.e. the orthocomplement to the maximal rank subspace. Specifically, it is shown that the rank-deficient $\mathrm{Sp}(\mathbb{F}_d^{2n})$-isotypes are labeled by irreps of $O(\mathbb{F}_d^k)$ for certain values of $k < t$. Furthermore, I show that these rank-deficient representations are spanned by certain types of *Calderbank-Shor-Steane* (CSS) codes introduced in [GNW21]. A central insight which enables the proof of its main theorem is that the code spaces, which are themselves $\mathrm{Sp}(\mathbb{F}_d^{2n})$-subrepresentations, are equivalent to k-th tensor powers of the oscillator representation (again, for certain values of $k < t$). This allows us to provide a full decomposition of the t-th tensor power of the oscillator representation.

The second section, based on Ref. [MMG21b], extends the "η formalism" to set up a correspondence between the irreps of a certain subgroup of $O(\mathbb{F}_d^t)$ – the *stochastic orthogonal group* studied in [GNW21] – and Cl. Here too, this correspondence is used to describe a full decomposition of t-th tensor power representations. This work answers several questions that had been left open hitherto:

1. Can the proof techniques of [GH17, MMG21a] be modified to decompose Clifford—rather than oscillator—tensor powers? In particular, for the important case of *qubits* ($d = 2$), the oscillator representation does not exist. Can we use a similar approach to understand tensor powers of the qubit Clifford group?

2. Can the results in [GNW21] be used to simplify the proofs in [MMG21a]?

3. The results in [MMG21a] provide a way to indirectly decompose Clifford t-th tensor powers as long as t is not a multiple of d. We know, moreover, that some peculiarities arise when t is a multiple of d: For instance, in this case the restriction to the subgroup of Pauli matrices is Abelian. Can we provide a uniform description of the structure of t-th tensor power Clifford representations, i.e. with a formalism that is independent of the value t? What precisely is the difference between these two types of representations?

Chap. 3, based on Ref. [HMMH$^+$20], proposes a protocol to efficiently generate *approximate unitary t-designs*. Unitary designs are certain probability distributions on $\mathrm{U}(d^n)$ that mimick the first t moments of the Haar distribution. Unitary designs have come into the spotlight in recent years due to their applications in, e.g., a variety of quantum characterization techniques [Sco08, KdSR$^+$14, KZG16a, MGE12, RKK$^+$18], in the study of quantum chaos [RY17] and black holes [HP07], and as primitives in quantum cryptography [KMK21, ABW09, BGGS21].

Finding exact unitary designs for $t > 3$ is infamously hard [ZKGG16, BNRT20]. On the other hand, it is known that random local circuits converge to approximate unitary t-designs in depth $O(n^2 t^{10})$ [BHH16]. This result was an important breakthrough

at the time—it says that while a typical Haar random unitary requires $\sim \exp(n)$ local gates to be approximated, some properties of Haar randomness may be approximated with polynomial quantum circuits. While this construction is efficient, it is out of reach for near term quantum devices, needing precise implementation of general polynomial depth circuits.

Chap. 3 leverages recent results on Clifford tensor power representations [GNW21] to improve on this situation:

It is known that the qubit Clifford group is a 3-design and not a higher t-design [Zhu17, Web16, ZKGG16]. This said, the Clifford group may be used to efficiently generate higher-order *complex projective t-designs*—sets of pure states which mimick the moments of the flat distribution on the unit sphere in $\mathbb{C}^{2^n}$. Ref. [ZKGG16] asks whether it is possible to construct, similarly, higher-order unitary t-designs based on the Clifford group. Moreover, it is natural to ask whether one can improve on the construction of [BHH16] by using knowledge about the tensor power representations of the Clifford group.

Chap. 3 answers these questions positively. The construction presented there considers *k-interleaved quantum circuits*: circuits of the form $U_1 K \cdots U_k K$, where $U_i \in \mathrm{Cl}$ are randomly sampled and K is a fixed single-qubit gate. It is shown that if $k = O(t^4 \log t)$, the circuits are sampled from an approximate t-design. In this way, the circuits require an n-independent number of non-Clifford gates. This makes this constuction easier to implement in a fault-tolerant way and efficiently simulatable in a classical computer (using the methods of Ref. [BBC$^+$19]). The proofs in [HMMH$^+$20] make extensive use of the structure of the commutant of tensor power Clifford representations studied in [GNW21]. I believe that more detailed information about these representations, and in particular the results in Chap. 2, could be used to improve on this construction.

Beyond these projects, I have also participated in the work leading up to [HMMVG21]. Here, a distance measure to the set of stabilizer states – the *stabilizer extent* [BBC$^+$19] – is considered. The extent is a convex relaxation of the *stabilizer rank*, the parameter that governs the complexity of simulating a quantum circuit via the methods in [BBC$^+$19]. The extent is known to be submultiplicative, $\mathrm{extent}(\psi \otimes \phi) \leq \mathrm{extent}(\psi)\mathrm{extent}(\phi)$. Furthermore, if ψ and ϕ are states on $n \leq 3$ qubits it is known that the extent is actually multiplicative—that is, the inequality presented above is saturated. Since the introduction of the stabilizer extent in 2018, it has remained an open question whether it is multiplicative for higher n. In this work, we answer this question negatively. My role in this last project was secondary, making my contribution hard to single out. Because of this, I have decided not to include it here.

Numerical representation theory. The second cluster of results in this thesis formulates an algorithm to decompose numerically defined representations. This algo-

rithm may decompose finite-dimensional representations of compact groups, and has a rigorous performance guarantee. Moreover its runtime beats state-of-the-art algorithms in certain regimes.

My interest in this problem comes from the solution of semi-definite programs (SDPs). Semi-definite programs are a class of optimization problems defined over the cone of positive semidefinite matrices in $\mathbb{C}^{n\times n}$. Their solution is known to be efficient. However, some applications of SDPs in quantum information require the solution of high-dimensional SDPs—that is, SDPs whose dimension n is prohibitively large. The goal is to produce a practical tool for dimensional reduction in SDPs. In this way, this research program aimed to both produce an efficient algorithm for the claimed dimensional reduction, and a performant implementation of it.

It is known that one can exploit symmetries of the SDPs to achieve dimensional reductions [Val09]. Moreover, in recent years several algorithms have been proposed for this purpose [PP20, MM11, CL20, MKKK10]. These algorithms consider the more general problem of decomposing matrix $*$-algebras. However, they have large runtimes as a function of n. Arguably, one could hope that if the symmetry being exploited arises from a *group action*, then simpler and faster algorithms could be obtained. In this case, rather than solving the more general problem of decomposing matrix $*$-algebras, one is interested in decomposing unitary group representations.

A foundational result in this regard is the variant of Dixon's algorithm [Dix70] analyzed in Ref. [BF91]. The decompositions obtained using this algorithm are guaranteed to be accurate, but its runtime also increases rather steeply, scaling as $O(n^5)$.

In Chap. 5, I present a method for decomposing representations of compact groups. This algorithm, RepLAB, is presented in [RMMB19] and was coded by my colleagues in [RB18]. RepLAB shares similarities with Dixon's algorithm, which are discussed in that same chapter. These similarities notwithstanding, RepLAB has a considerably shorter runtime of $O(n^3)$. This comes at a price: as opposed the variant of Dixon's algorithm studied in [BF91], is not guaranteed to give correct solutions. (However, RepLAB has been observed to work well in practice.) The proof of such a guarantee would be hindered by the facts that: *1.* It would require knowledge of the eigenvalue statistics of *finite dimensional* random matrices (these results are typically only known asymptotically). *2.* Although its working principle is simple, the actual workflow of RepLAB is rather involved due to runtime optimization.

To provide a fuller picture on the correctness of RepLAB outputs, I have used two approaches. One, used in Sec. 5.4.2, is to give evidence that RepLAB's working principle provides accurate decompositions. Here, this evidence requires as an assumption that finite-dimensional random matrices have sufficiently well separated eigenvalues. This is subsequently checked empirically in Sec. 5.4.4. In this way, Chap. 5 gives reason to believe that RepLAB's output are close-to-exact representation decompositions.

However, this chapter falls short of giving a rigorous proof of this.

The second approach, used in Chap. 6, is a response to this lack of a proof. In that chapter, an efficient *certifying* algorithm is provided: It takes a decomposition of RepLAB as an input and certifies whether each block is "approximately invariant" and "close to irreducible." Here, a projector is approximately invariant if there exists a second projector in its neighborhood which is exactly invariant. Moreover, it is close to irreducible if that invariant projector in the neighborhood projects onto an irreducible representation. I coded this certifying algorithm in [MM21] and present the most important features of this code in Chap. 7. Moreover, I benchmark the runtime of this algorithm on some representative examples of group representations in Sec. 7.2. These tests simultaneously benchmark the accuracy of RepLAB decompositions, which, as expected, are found to have a high quality.

Part I

Stabilizer representation theory

1 Intoduction to Part I and summary of its results

The stabilizer formalism was the first framework proposed to implement quantum error correction—to this day, stabilizer codes remain by far the best understood methods for this. From these origins, the stabilizer formalism has grown to encompass a myriad of other applications: classical simulation of quantum computing [BBC$^+$19, RLCK19, HL19, BK19, HMW20, HL21, GGKS20, Qas21, Gro06], certification of quantum systems [KLR$^+$08, RKK$^+$18, HWFW19, MGE11, HFGW18, MGJ$^+$12, KdSR$^+$14, DHW19, HF17, KL17], or matrix reconstruction [KZG16b] are some examples.

In this chapter I will introduce the results derived in Part I of this thesis.

Chapter 2 presents results on tensor power representations of the Clifford group and the closely-related oscillator representation. The results presented here extend the Theta correspondence to the Clifford group. In particular, the irreps appearing in Clifford t-th tensor powers are labeled by the characters of a sequence of orthogonal groups $O(\mathbb{Z}_d^{t_i})$ of dimensions $t_i \leq t$. In essence, these results say that the class of Calderbank-Shor-Steane (CSS) codes introduced in Ref. [GNW21], together with the η duality formalism of [GH17], may be used to understand Clifford tensor power representations.

Chapter 3 uses the structure of the commutant of Clifford tensor power representations, derived recently in Ref. [GNW21], to efficiently construct unitary t-designs. The circuits realizing these designs are random Clifford circuits with certain interleaved single-qubit non-Clifford gates. The main result is that only $\sim t^4$ non-Clifford gates are sufficient for this. Strikingly, this number is independent of the system size n. Because the number of Clifford gates scales as $\sim n^2$, this result states that a vanishingly small density of non-Clifford gates is sufficient to upgrade random Clifford circuits from 3-designs to approximate t-designs.

Throughout this introductory chapter I will focus on the simplest scenarios, leaving the discussion of results in their full generality to later chapters.

1.1 The stabilizer formalism

Here I introduce some basic concepts needed to understand the results of this thesis. A much more in-depth discussion of the stabilizer formalism is found in [Hei21].

The foundational object in the formalism is the *Pauli group* (also known as the *finite Heisenberg* or *Heisenberg-Weyl* group, or as the *extra special p-group* when $d = p$ is an odd prime). Consider the *displacement operators* acting on a Hilbert space $\mathcal{H} = (\mathbb{C}^d)^{\otimes n}$, for d prime, as

$$X(a)\left|x\right\rangle = \left|x+a\right\rangle, \quad Z(b)\left|x\right\rangle = \omega^{b\cdot x}\left|x\right\rangle,$$

where $a,\ b,\ x \in \mathbb{Z}_d^n$ (thus, $x + a$ is modulo d), and $\omega = \exp(2i\pi/d)$. Then, the Pauli group is

$$\mathcal{P} := \left\{\tau^a Z(b)X(c) \mid a \in \mathbb{Z}_D,\ b,\ c \in \mathbb{Z}_d\right\},$$

where $\tau = (-1)^d \exp(i\pi/d)$ and $D = 4$ if $d = 2$ and $D = d$ otherwise.[1] We denote $W(b, c) := Z(b)X(c)$. Pauli matrices obey the commutation relation $W(u)W(v) = \omega^{[u,v]}W(v)W(u)$, where

$$[u, v] = u_X \cdot v_Z - u_Z \cdot v_X, \qquad u, v \in \mathbb{Z}_d^{2n}.$$

This implies that $\mathcal{P}/\langle\tau\mathbb{1}\rangle \simeq \mathbb{Z}_d^{2n}$. The *Clifford group* $\mathrm{Cl} \subset \mathrm{U}(d^n)$ is a finite group which, up to phases, is the normalizer of the Pauli group in $\mathrm{U}(d^n)$. That is,

$$\{e^{i\varphi}U \mid U \in \mathrm{Cl},\ \varphi \in [0, 2\pi)\} = \{U \in \mathrm{U}(d^n) \mid U\mathcal{P}U^\dagger = \mathcal{P}\}.$$

More precisely, the Clifford group is generated by $\mathcal{P}$, together with the Hadamard gate

$$\mathrm{H} = \frac{1}{\sqrt{d^n}} \sum_{a,b \in \mathbb{Z}_d^n} \omega^{a \cdot b} \left|a\right\rangle\!\left\langle b\right|,$$

acting on any qudit, the phase gate

$$\mathrm{P}_\tau = \sum_{a \in \mathbb{Z}_d^n} \tau^{a \cdot a} \left|a\right\rangle\!\left\langle a\right|,$$

acting on any qudit as well, and the the control-add gate

$$\mathrm{CADD} = \sum_{a,b \in \mathbb{Z}_d^n} \left|a, a + b\right\rangle\!\left\langle a, b\right|$$

acting on any pair of qudits.

When d is odd, the Clifford group contains a subgroup isomorphic to the finite symplectic group $\mathrm{Sp}(\mathbb{Z}_d^{2n}) := \{g \in \mathrm{Gl}(\mathbb{Z}_d^{2n}) \mid [g \cdot\ , g \cdot\ ] = [\ \cdot\ ,\ \cdot\ ]\}$ [Gro06]. The unitaries of this subgroup, denoted $\mu(S)$ for $S \in \mathrm{Sp}(\mathbb{Z}_d^{2n})$, define a representation of the symplectic group known as the oscillator representation or the Weil representation. The oscillator representation acts on Pauli operators as

$$\mu(S)W(v)\mu(S)^\dagger = W(Sv).$$

When $d = 2$, in contrast, this construction fails: there is no subgroup G of Cl isomo-

[1] This notational convention has been chosen in order to treat the slightly different qubit and odd qudit cases uniformly.

prhic to the symplectic group [BRW61, Zhu16] which complements $\mathcal{P}$ (ie. which has a trivial intersection with $\mathcal{P}$ and for which $\mathrm{Cl} = \mathcal{P} \cdot \mathrm{G}$).

A *stabilizer group* is an Abelian subgroup of $\mathcal{P}$ with no pure phases $e^{i\varphi}\mathbb{1}$ other than the identity ($\varphi = 0$). The invariant space of a stabilizer group is a *stabilizer code*. That is, if $\mathcal{S} \subset \mathcal{P}$ is a stabilizer group, then the projector onto the corresponding code is given by

$$P_{\mathcal{S}} = \frac{1}{|\mathcal{S}|} \sum_{W \in \mathcal{S}} W.$$

If $\mathcal{S}$ has k generators, $\mathcal{S} \simeq \mathbb{Z}_d^k$ and the code has dimension $\mathrm{tr}\, P_{\mathcal{S}} = d^{n-k}$.

The oscillator representation can be defined analogously as a representation $\mu_{\mathbb{R}}$ of $\mathrm{Sp}(\mathbb{R}^{2n})$. While this thesis does not deal directly with this representation, some of the results presented have been inspired by this continuous oscillator representation. In particular, the results of Chap. 2 can be seen as generalizations of the Theta correspondence between $\mathrm{Sp}(\mathbb{R}^{2n})$ and $\mathrm{O}(\mathbb{R}^t)$. I would like to highlight this connection, so let us shortly introduce $\mu_{\mathbb{R}}$. For this, Folland's classic textbook [Fol89] will be followed.

Consider a quantum system with Hilbert space $\mathcal{H}_{\mathbb{R}} = L^2(\mathbb{R}^n)$, and define the displacement operators by

$$W_{\mathbb{R}}(v) = \exp\Big(i(v_X \cdot \hat{P} - v_P \cdot \hat{X})\Big), \quad v = (v_X, v_P) \in \mathbb{R}^{2n},$$

where $\hat{X}$, $\hat{P}$ are the canonical position and momentum operators,

$$(\hat{X}\psi)(x) = (x_1\psi(x), \ldots, x_n\psi(x)), \quad 2\pi i(\hat{P}\psi)(x) = \nabla\psi(x).$$

The displacement operators generate a representation of the (continuous) Heisenberg group. Similar to the discrete case presented above, they obey the commutation relation

$$W_{\mathbb{R}}(v)W_{\mathbb{R}}(u) = \exp(2\pi i[v,u])W_{\mathbb{R}}(u)W_{\mathbb{R}}(v),$$

where $[v,u] = v_X \cdot v_P - v_P \cdot v_X$. There exists a map $\mu_{\mathbb{R}} : \mathrm{Sp}(\mathbb{R}^{2n}) \to \mathrm{U}(\mathcal{H}_{\mathbb{R}})$, called the *metaplectic representation*, satisfying

$$\mu_{\mathbb{R}}(S)\mu_{\mathbb{R}}(S') = \pm\mu_{\mathbb{R}}(SS'), \tag{1}$$

and,

$$\mu_{\mathbb{R}}(S)W_{\mathbb{R}}(v)\mu_{\mathbb{R}}(S)^\dagger = W_{\mathbb{R}}(Sv).$$

The $\pm$ factor in eq. (1) implies that $\mu_{\mathbb{R}}$ is not a representation of $\mathrm{Sp}(\mathbb{R}^{2n})$ itself, but rather of a double cover called the *metaplectic group* (hence the name). When dealing over a finite field, in contrast, μ is a representation of $\mathrm{Sp}(\mathbb{Z}_d^{2n})$ itself. Because of this I will omit this technical detail as it is not important for the present discussion. Refs. [Fol89, KV78] give a rigorous treatment of this construction.

1.2 Tensor power representations

A considerable amount of the work in my thesis was devoted to studying two interrelated objects:

1. tensor power representations of the oscillator representation μ,

$$\mu^{\otimes(r,s)} : S \mapsto \mu(S)^{\otimes r} \otimes \bar{\mu}^{\otimes s}(S), \quad S \in \mathrm{Sp}(\mathbb{Z}_d^{2n}),$$

   where $\bar{\mu}(S)$ is the complex conjugate of $\mu(S)$,

2. tensor power representations of the defining representation of the Clifford group,

$$\Delta_{r,s} : U \mapsto U^{\otimes(r,s)}, \quad U \in \mathrm{Cl}.$$

This line of work is presented in Chap. 2. Throughout this introduction, I will focus on the case $s = 0$ for ease of exposition. Furthermore, let $\Delta_{r,0} =: \Delta_r$.

Low tensor powers, with $r + s =: t \leq 4$ have been fully understood in the past few years [ZKGG16, Zhu17, Web16]. These results have lead to several applications, for example in the simulation of quantum computing [BBC$^+$19], in quantum state distinction [KZG16a], quantum device characterization [RKK$^+$18], and low-rank matrix recovery [KZG16b]. The variety of these applications have motivated me to elucidate the structure of these representations for larger values of t.

1.2.1 The oscillator representation and Theta correspondence

The oscillator representation and its tensor powers have interested representation theorists for a long time now. These representations give rise to the *Theta correspondence* [How89a, How89b, KV78, GH17], also known as *Howe duality*, whose study has shed light on e.g. invariant theory [How89a], the structural properties of classical groups [How10] and on automorphic forms [Gel06].

Infamously, the Theta correspondence fails to hold between the finite symplectic and orthogonal groups. This has not deterred work on finding *some* correspondence that works here [AKP16, AM93, AMR96, How73, Sri79]. Given the rich theory that arises from the Theta correspondence, one would hope to extend these results to the

finite cases as well. Recent work on the finite cases has found that a similar duality – called the η *correspondence* – holds on certain "maximal rank" representations within the "stable range" $t \leq n$ [GH17, GH21]. Further work has considered generalizations of the η correspondence beyond the stable range [Pan20]. My work [MMG21a] clarified the structure of "rank-deficient" representations, which was left open in the works cited above. Furthermore, it provides a rigorous connection to the representation theory of Clifford tensor power representations. This work laid the conceptual foundation on which my second contribution, Ref. [MMG21b], was built.

Tensor powers of $\mu_{\mathbb{R}}$ and the Θ correspondence. To mathematically motivate the study of $\mu^{\otimes t}$, let us briefly recall the duality between $\mathrm{Sp}(\mathbb{R}^{2n})$ and $\mathrm{O}(\mathbb{R}^t)$. Here, the tensor power representation $\mu_{\mathbb{R}}^{\otimes t}$, acting on the Hilbert space $\mathcal{H}_{\mathbb{R}}^{\otimes t}$, is considered. The classical result from Refs. [KV78, How89a] implies that if $t \leq n$ the commutant of the representation $\mu_{\mathbb{R}}^{\otimes t}$ is spanned by the representation $R_{\mathbb{R}}$ of $\mathrm{O}(\mathbb{R}^t)$ defined on $L^2(\mathbb{R}^t \otimes \mathbb{R}^n) \simeq \mathcal{H}_{\mathbb{R}}^{\otimes t}$ by

$$(R_{\mathbb{R}}(O)\psi)(F) = \psi((O^{-1} \otimes \mathbb{1}_{\mathbb{R}^n})F), \quad F \in \mathbb{R}^t \otimes \mathbb{R}^n.$$

As a consequence, there exists an injective function $\Theta : \mathrm{Irr}\ \mathrm{O}(\mathbb{R}^t) \to \mathrm{Irr}\ \mathrm{Sp}(\mathbb{R}^{2n})$ for which

$$\mathcal{H}_{\mathbb{R}}^{\otimes t} \simeq \bigoplus_{\tau \in \mathrm{Irr}\ \mathrm{O}(\mathbb{R}^t)} \tau \otimes \Theta(\tau),$$

where the decomposition is as a $\mathrm{O}(\mathbb{R}^t) \times \mathrm{Sp}(\mathbb{R}^{2n})$ representation. In this equation τ is a representation *space* of $\mathrm{O}(\mathbb{R}^t)$, and similarly $\Theta(\tau)$ is a representation space of $\mathrm{Sp}(\mathbb{R}^{2n})$. This way, studying tensor powers of $\mu_{\mathbb{R}}$ provides a connection between the representation theory of the orthogonal and symplectic groups.

Tensor powers of μ. As mentioned before, the Theta correspondence does not hold between symplectic and orthogonal groups over a finite field. To understand this statement, it is instructive to look at the contribution to this field from the physics point of view. Ref. [GNW21] characterizes the commutant of Clifford tensor power representations. A very mild tweak of their proofs (see [MMG21b, Prop. III.1]) gives an analogous characterization of the commutant of $\mu^{\otimes t}$.

The representation $\mu^{\otimes t}$ acts on the following Hilbert space:

$$\mathcal{H}_{n,t} := ((\mathbb{C}^d)^{\otimes n})^{\otimes t} = \begin{array}{ccccccc} & \mathbb{C}^d & \otimes & \dots & \otimes & \mathbb{C}^d \\ \otimes & \mathbb{C}^d & \otimes & \dots & \otimes & \mathbb{C}^d \\ & \vdots & & \ddots & & \vdots \\ \otimes & \mathbb{C}^d & \otimes & \dots & \otimes & \mathbb{C}^d, \end{array} \tag{2}$$

where the grid has t rows and n columns. Each column corresponds to a Hilbert space $(\mathbb{C}^d)^{\otimes t}$. Two types of operators on this column space play a prominent role in [GNW21]. The first is a representation r of the orthogonal group $\mathrm{O}(\mathbb{Z}_d^t)$ given by

$$r(O)\,|x\rangle = |Ox\rangle\,, \quad x \in \mathbb{Z}_d^t.$$

The second set consists of certain projectors onto CSS codes. A subspace $N \subset \mathbb{Z}_d^t$ is called *isotropic* if $N \subset N^\perp := \{v \in \mathbb{Z}_d^t \mid v \cdot u = 0 \;\forall\; u \in N\}$. The code projectors then correspond to

$$\pi_N = \frac{1}{|N|^2} \sum_{u,v \in N} Z(u)X(v).$$

The relevant restatement of the main result in [GNW21] is the following.

Theorem 1.1. *Let $t \leq n$. The commutant of the representation $\mu^{\otimes t}$ is generated as an algebra by the set*

$$\{P_N,\; R(O) \mid N \textit{ isotropic},\; O \in \mathrm{O}(\mathbb{Z}_d^t)\}$$

where $P_N := \pi_N^{\otimes n}$, $R := r^{\otimes n}$. Furthermore, the following set

$$\{P_N R(O) \mid N \textit{ isotropic},\; O \in \mathrm{O}(\mathbb{Z}_d^t)\}$$

is a basis for this commutant.

This result provides an intuition as to how the Theta correspondence fails in this finite case. Namely, it fails precisely because of the presence of the projectors P_N. Conceptually, it is pleasing that this duality fails precisely because of an object which has no clear analog in the real case—$\mathbb{R}^t$ has no isotropic subspaces. Throughout the research program developed in Chap. 2, I have built on this intuition, using the codes P_N to understand the detailed structure of tensor power representations.

It is clear that, on the orthocomplement of all code spaces,

$$\mathcal{H}_0 := \mathrm{span}\{C_N := \mathrm{range}\; P_N \mid N \text{ isotr.}\}^\perp,$$

the commutant of $\mu^{\otimes t}$ is spanned by the matrices

$$\Pi_{\mathcal{H}_0} R(O) \Pi_{\mathcal{H}_0},$$

where $\Pi_{\mathcal{H}_0}$ is the orthogonal projector onto $\mathcal{H}_0$. Furthermore, because of the relation $R(O)P_N R(O)^\dagger = P_{ON}$, $\mathcal{H}_0$ is an invariant subspace of R. On this invariant subspace

it acts as some representation R_0,

$$R_0(O) := \Pi_{\mathcal{H}_0} R(O) \Pi_{\mathcal{H}_0}.$$

This implies that the commutant of $\mu^{\otimes t}|_{\mathcal{H}_0}$ is spanned by images of the representation R_0. Thus, a similar equation appears here: there exists an injective function $\eta : \mathrm{Irr}\ \mathrm{O}(\mathbb{Z}_d^t) \to \mathrm{Irr}\ \mathrm{Sp}(\mathbb{Z}_d^{2n})$ such that

$$\mathcal{H}_0 \simeq \bigoplus_{\tau \in \mathrm{Irr}\ \mathrm{O}(\mathbb{Z}_d^t)} \tau \otimes \eta(\tau), \tag{3}$$

as a $\mathrm{O}(\mathbb{Z}_d^t) \times \mathrm{Sp}(\mathbb{Z}_d^{2n})$ representation.

This is, in essence, equivalent to the main result of Ref. [GH17]. This reference argues differently, without mentioning the codes P_N explicitly. Instead, the "organizing principle" it uses is a notion of *rank* for representations of the symplectic group. Namely, given a representation ρ of $\mathrm{Sp}(\mathbb{Z}_d^{2n})$, it considers the restriction of ρ to the subgroup

$$\mathcal{N} := \left\{ \begin{pmatrix} \mathbb{1}_n & A \\ 0 & \mathbb{1}_n \end{pmatrix} \;\middle|\; A \in \mathrm{Sym}_n \right\} \subset \mathrm{Sp}(\mathbb{Z}_d^{2n}),$$

where Sym_n is the space of $n \times n$ symmetric matrices over $\mathbb{Z}_d$. That is, it looks at the representation $\{\rho(S) \mid S \in \mathcal{N}\}$.

Because $\mathcal{N}$ is Abelian and isomorphic to Sym_n (seen as an additive group), this restricted representation decomposes into one dimensional blocks,

$$\rho \begin{pmatrix} \mathbb{1} & A \\ 0 & \mathbb{1} \end{pmatrix} \simeq \bigoplus_{B \in \mathrm{Sym}_n} \chi_B^{\oplus m_B}(A),$$

where $\chi_B(A) = \omega^{\mathrm{tr}\, AB}$ is an additive character. The rank of ρ is then defined as

$$\mathrm{rk}\ \rho = \max_B \mathrm{rank}\, B \quad \text{s.t.} \quad m_B \neq 0.$$

Ref. [GH17] proves that $\mathrm{rk}\ \mu^{\otimes t} = t$ and its main result is the following.

Theorem 1.2. *Let $t \leq n$. Consider the subspace $\mathcal{H}^0$ spanned by all irreducible components in $\mathcal{H}_{n,t}$ which have rank t. Then, there exists an injective function η : $\mathrm{Irr}\ \mathrm{O}(\mathbb{Z}_d^t) \to \mathrm{Irr}\ \mathrm{Sp}(\mathbb{Z}_d^{2n})$ for which*

$$\mathcal{H}^0 \simeq \bigoplus_{\tau \in \mathrm{Irr}\ \mathrm{O}(\mathbb{Z}_d^t)} \tau \otimes \eta(\tau), \tag{4}$$

as a $\mathrm{O}(\mathbb{Z}_d^t) \times \mathrm{Sp}(\mathbb{Z}_d^{2n})$ representation.

Ref. [GH17] left open the question of what the structure of lower rank representations is. A strong indication that this structure can be understood through the codes P_N is the glaring similarity between eqs. (3) and (4). My first contribution [MMG21a] formalizes this idea. To understand its main result, two technical remarks are necessary.

First, given an isotropic subspace $N \subset \mathbb{Z}_d^t$ consider the subgroup

$$\mathrm{O}_N := \{O \in \mathrm{O}(\mathbb{Z}_d^t) \mid ON = N\}.$$

This subgroup acts naturally on the space $N^\perp/N$. Furthermore, the space $N^\perp/N$ inherits a quadratic form by restricting the dot product (since, for any $v \in N^\perp$ and $u \in N$, it holds that $(v+u)\cdot(v+u) = v\cdot v$). This inherited quadratic form is preserved by O_N, so the action of this group on $N^\perp/N$ gives a canonical homomorphism $\mathrm{O}_N \to \mathrm{O}(N^\perp/N)$. Because of this, we may inject $\mathrm{Irr}\ \mathrm{O}(N^\perp/N) \hookrightarrow \mathrm{Irr}\ \mathrm{O}_N$. Second, because of a result commonly known as "Witt's lemma," for every pair N, N' of isotropic subspaces of the same dimension, the groups O_N and $\mathrm{O}_{N'}$ are conjugate in $\mathrm{O}(\mathbb{Z}_d^t)$. In particular, $\mathrm{O}(N^\perp/N) \simeq \mathrm{O}(N'^\perp/N')$.

Theorem 1.3. *Let $t \leq n$. Then, every subrepresentation of $\mu^{\otimes t}$ with rank $k < t$ is contained in the span of all codes P_N, that is, in* $\mathrm{span}\{C_N \mid N \text{ isotr.}\} = \mathcal{H}_0^\perp$. *Furthermore, consider any sequence of isotropic subspaces $N_1, \ldots, N_{m(t)}$, with* $\dim N_l = l$ *and $m(t)$ being the maximal dimension of an isotropic subspace. That is, the sequence contains one representative subspace N_i for each possible dimension i.*

Then, as a representation of $\mathrm{O}(\mathbb{Z}_d^t) \times \mathrm{Sp}(\mathbb{Z}_d^{2n})$,

$$\mathcal{H}_{n,t} = \bigoplus_{l=1}^{m(t)} \bigoplus_{\tau \in \mathrm{Irr}\ \mathrm{O}(N_l^\perp/N_l)} \mathrm{Ind}_{\mathrm{O}_{N_l}}^{\mathrm{O}(\mathbb{Z}_d^t)}(\tau) \otimes \eta(\tau), \tag{5}$$

where η is as in Thm. 1.2. Here, $\mathrm{Ind}_{\mathrm{O}_{N_l}}^{\mathrm{O}(\mathbb{Z}_d^t)}(\tau) \otimes \eta(\tau)$ *is the* induced representation *of τ from O_{N_l} to* $\mathrm{O}(\mathbb{Z}_d^t)$.

While I have presented this result here in relation to Thm. 1.1, the proof of Thm. 1.3 is formally independent the former. Its proof, rather, relies on the application of two ideas. The first, is a structural property of the code spaces C_N, namely,

$$C_N \simeq \mu^{\otimes(r',s')}, \tag{6}$$

where $s' = \dim N \mod 2 \in \{0,1\}$ and $r' = t - 2 - s'$. The second, is an adaptation of a result from Fourier analysis—if $f \in \mathbb{C}^k \to \mathbb{C}$ is supported on a subspace $\mathcal{V}$, then its Fourier transform is invariant under $\mathcal{V}^\perp$ translations. Intuitively speaking, if f is sparse, its Fourier transform has a large symmetry group. The idea of the adaptation

is the following. Consider Hilbert space as a function space, $\mathcal{H}_{n,t} \simeq \mathbb{C}[\mathbb{Z}_d^{n\times t}]$. A short calculation shows that subrepresentations ρ with rank $k < t$ are only supported on matrices $M \in \mathbb{Z}_d^{n\times t}$ of rank $\leq k$. Then, for any vector $\Psi \in \rho$, its Fourier transform,

$$\mu^{\otimes t}\begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}\Psi \in \rho,$$

has a rather restricted support. Therefore, Ψ itself can be expected to be invariant under a large class of translations. In the final step of the proof, it is shown that these invariances correspond exactly to lying on the span of the CSS code spaces $\{C_N\}_N$.

This work provides a bridge between the representation theory of oscillator tensor powers, and tensor powers of the Clifford group (see [MMG21a, Prop. 4.2]). This was important for my purposes since it connected the widely-studied Theta correspondence, to the less-studied but more physically interesting representation theory of the Clifford group. This correspondence is explicitly constructed in the proof of the following proposition.

Proposition 1.1. *Let $t \mod d \neq 0$, and define $s_t = 0$ if t is a square in $\mathbb{Z}_d$, and otherwise $s_t = 1$. Recall that the representation $\Delta_{t,0}$ of Cl has the form $\Delta_{t,0}(U) = U^{\otimes t}$. Then, there is a one-to-one correspondence between the isotypes in $\Delta_{t,0}$ and those in $\mu^{\otimes(t-2-s_t,s_t)}$.*

In my later contribution [MMG21b], the insights gained in [MMG21a] were used to directly study Clifford tensor powers—thus relieving the need for Prop. 1.1. I have decided to anyway highlight this theorem as it might be mathematically interesting, connecting, e.g., the research program on Clifford invariants [NRS06, NRS01, BOZ21] – which has met success in coding theory – to Howe duality.

1.2.2 Clifford tensor powers

The representation theory of oscillator tensor powers is mathematically interesting in its own right. For quantum information purposes, however, the most relevant object is the Clifford group and its representation theory. Indeed, the question that started the work [MMG21a] was *how do tensor power representations of the Clifford group decompose?* With regard to this question, Ref. [MMG21a] left several imporant aspects unanswered:

1. *How do tensor powers of the qubit Clifford group decompose?* The qubit case – arguably the most important in quantum information theory – can not be tackled using the methods of [MMG21a]. (Recall that the qubit Clifford group does not define an oscillator representation in this case—this seemingly places a "dead end" sign in front of any such attempt.)

2. *How do t-th tensor powers of the Clifford group decompose when t is a multiple of d?* Whenever this is the case, Prop. 1.1 can not be used to decompose such representations.

3. The proof of the main theorem in Ref. [MMG21a] does not use the structure of the commutant of tensor power representations given in [GNW21] (Thm. 1.1). *How can one use Thm 1.1 to decompose Clifford tensor power representations?* Such an approach would deal directly with Clifford representations, rather than referring them back to symplectic ones (as in Prop. 1.1). Moreover, it would have a chance of working on the qubit case, given that [GNW21] holds in that case as well.

In [MMG21b], I address these questions. In particular, a full decomposition of Clifford tensor power representations – one that works for the cases highlighted in questions 1. and 2. – is provided. The proof of the main theorem, furthermore, heavily relies on the results of [GNW21].

Thm. 1.1 is a mild alteration of the main result of [GNW21], which I have chosen in order to motivate the η correspondence of [GH17]. In this section, however, it is convenient to present that result in its original form.

For this, let me first remark three technical details. Let $\mathbf{1}_t = (1, 1, \ldots, 1) \in \mathbb{Z}_d^t$ be the all-ones vector. First detail: In the following, I will denote by $\mathrm{O}(\mathbb{Z}_d^t)$ the group of matrices O for which

$$(Ov) \cdot (Ov) = v \cdot v \mod D.$$

Notice that this group is simply the orthogonal group in the case where d is odd, and if $d = 2$ it is a strict subgroup of the orthogonal group, ie. the group of matrices $O \in \mathrm{Gl}(\mathbb{Z}_2^t)$ for which

$$(Ov) \cdot (Ov) = v \cdot v \mod 2.$$

Moreover, the *orthogonal stochastic group* is the subgroup of $\mathrm{O}(\mathbb{Z}_d^t)$ which preserves the all-ones vector,

$$\mathrm{St}(\mathbb{Z}_d^t) := \{O \in \mathrm{O}(\mathbb{Z}_d^t) \mid O\mathbf{1}_t = \mathbf{1}_t\}.$$

Second, the notion of *isotropic* subspaces is slightly stronger for the qubit case. Namely, for any d, a subspace $N \subset \mathbb{Z}_d^t$ is *isotropic* if

$$v \cdot v = 0 \mod D, \qquad \forall\, v \in N.$$

In particular, isotropic vectors in $\mathbb{Z}_2^t$ are supported on $4k$ components, for some k.

Third, an isotropic subspace $N \subset \mathbb{Z}_d^t$ is *stochastic* if

$$\mathbf{1}_t \cdot v = 0 \mod d \qquad \forall\, v \in N.$$

The main result of [GNW21] reads:

Theorem 1.4. *Let the operators $R(O)$ and P_N be as in Thm 1.1. Then, the commutant of the t-th tensor power representation of the Clifford group is generated, as an algebra, by*

$$\{R(O), P_N \mid O \in \mathrm{St}(\mathbb{Z}_d^t),\ N \textit{ stoch.}\}.$$

Moreover, the set

$$\mathcal{A}_t := \{R(O)P_N \mid O \in \mathrm{St}(\mathbb{Z}_d^t),\ N \textit{ stoch.}\}$$

forms a basis for this algebra.

To motivate the results in this section, we may argue analogously to the discussion leading up to eq. (3). Let $\mathcal{H}_0$ be given by

$$\mathcal{H}_0 := \mathrm{span}\{C_N \mid N \text{ stoch. isotr.}\}^{\perp},$$

notice that we now consider only code spaces corresponding to stochastic isotropic subspaces N. As before, $\mathcal{H}_0$ is an invariant subspace of $R(O)$, and $P_N\mathcal{H}_0 = 0$ for all isotropic stochastic N. Furthermore it is an invariant subspace of the Clifford action (because each code space C_N is itself invariant). Therefore, by Thm. 1.4 the commutant of the Clifford action on $\mathcal{H}_0$ is spanned by the subrepresentation $\{R_0(O) \mid O \in \mathrm{St}(\mathbb{Z}_d^t)\}$, where $R_0 := R|_{\mathcal{H}_0}$. It follows that there exists an injective function $\eta : \mathrm{Irr}\ \mathrm{St}(\mathbb{Z}_d^t) \to \mathrm{Irr}\ \mathrm{Cl}$ such that

$$\mathcal{H}_0 \simeq \bigoplus_{\tau \in \mathrm{Irr}\ \mathrm{St}(\mathbb{Z}_d^t)} \tau \otimes \eta(\tau), \tag{7}$$

as a $\mathrm{St}(\mathbb{Z}_d^t) \times \mathrm{Cl}$ representation.

This argument tells us that it might be possible to directly decompose Clifford tensor powers by analysing the action of Cl on the code spaces C_N. It also hints at the possibility that extending the theory of rank to the Clifford group might help in understanding these tensor powers. These hints lead me to the ideas that are behind the main results in [MMG21b].

The *pièce de la resistánce* of this work is Thm. 1.5. There I use the following

notation,

$$\mathcal{G}_m := \{N \subset \mathbb{Z}_d^t \text{ stoch. isotr.} \mid \dim N = m,\ \mathbf{1}_t \notin N\},$$
$$\mathcal{G}_m^0 := \{N \subset \mathbb{Z}_d^t \text{ stoch. isotr.} \mid \dim N = m,\ \mathbf{1}_t \in N\}.$$

Moreover, let N be an stochastic isotropic subspace and $v \in N^\perp$. Then, I denote $[v]_N := \{v + a \mid a \in N\} \in N^\perp/N$. The space $N^\perp/N$ inherits an inner product through

$$[u]_N \cdot [v]_N := u \cdot v \mod D.$$

This product is well defined: for all $a, b \in N$,

$$(u + a) \cdot (v + b) = u \cdot v \mod D.$$

The group $\mathrm{St}(N^\perp/N)$ is defined as the subgroup of $\mathrm{Gl}(N^\perp/N)$ which preserves the inherited dot product and stabilizes $[\mathbf{1}_t]_N$, i.e.

$$\mathrm{St}(N^\perp/N) := \{g \in \mathrm{Gl}(N^\perp/N) \mid g[\mathbf{1}_t]_N = [\mathbf{1}_t]_N,\ (g[u]_N) \cdot (g[v]_N) = [u]_N \cdot [v]_N\ \forall\, u, v \in N^\perp\}.$$

for any stochastic isotropic subspace $N \subset \mathbb{Z}_d^t$, the space $N^\perp/N$ inherits a quadratic form q_N from $q_{r,s}$ Notice that $\mathcal{G}^0$ is empty whenever t is not a multiple of D (in which case $\mathbf{1}_t \cdot \mathbf{1}_t \neq 0 \mod D$).

Theorem 1.5. *Let $t \leq n$. For each m such that $\mathcal{G}_m$ is non-empty, let $N_m \in \mathcal{G}_m$ be arbitrary. Similarly, let $M_m \in \mathcal{G}_m^0$ whenever $\mathcal{G}_m^0$ is non-empty. Then, there exist injective functions $\eta,\ \eta_0$,*

$$\eta : \bigcup_{N_m} \mathrm{Irr}\ \mathrm{St}(N_m^\perp/N_m) \to \mathrm{Irr}\ \mathrm{Cl}, \quad \eta_0 : \bigcup_{M_m} \mathrm{Irr}\ \mathrm{St}(M_m^\perp/M_m) \to \mathrm{Irr}\ \mathrm{Cl},$$

such that

$$\mathcal{H}_{n,t} \simeq \Big(\bigoplus_{N_m} \bigoplus_{\tau \in \mathrm{Irr}\ \mathrm{St}(N_m^\perp/N_m)} \mathrm{Ind}_{\mathrm{St}_{N_m}}^{\mathrm{St}(\mathbb{Z}_d^t)}(\tau) \otimes \eta(\tau) \Big)$$
$$\oplus \Big(\bigoplus_{M_m} \bigoplus_{\tau \in \mathrm{Irr}\ \mathrm{St}(M_m^\perp/M_m)} \mathrm{Ind}_{\mathrm{St}_{M_m}}^{\mathrm{St}(\mathbb{Z}_d^t)}(\tau) \otimes \eta_0(\tau) \Big).$$

Furthermore, $\mathrm{range}\ \eta \cap \mathrm{range}\ \eta_0 = \emptyset$.

The most important ingredient to the proof of Thm. 1.5 is a generalization of the argument above, which studies the structure of the basis $\mathcal{A}_t$. This basis forms a *semi-*

group, as proven in [GNW21], with the product laws given by

$$R(O)P_N R(O)^\dagger = P_{ON}, \quad P_{N_1}P_{N_2} = R(O_{N_1,N_2})P_{N(N_1,N_2)},$$

where $N(N_1, N_2) = \langle N_1^\perp \cap N_2,\ N_2\rangle \supseteq N_2$, and O_{N_1,N_2} is given in [GNW21, eq. (4.24)]. We use this product law to prove the following lemma.

Lemma 1.1. *Let N be a stochastic isotropic subspace with $t - 2\dim N \leq n$. Consider the subgroup $\mathrm{St}_N := \{O \in \mathrm{St}(\mathbb{Z}_d^t) \mid ON = N\}$, and the space*

$$\mathcal{H}_N := C_N \cap \mathrm{span}\{\mathrm{range}\ P_{N'} \mid N \subset N'\}^\perp.$$

In words, $\mathcal{H}_N$ is the orthocomplement of all proper subcodes $\mathrm{range}\ P_{N'} \subset C_N$ within C_N itself. Then, the action of St_N on C_N preserves $\mathcal{H}_N$ and gives rise to the homomorphism $\mathrm{St}_N \to \mathrm{St}(N^\perp/N)$. Moreover, there exists an injective function $\eta_N : \mathrm{Irr}\ \mathrm{St}(N^\perp/N) \to \mathrm{Irr}\ \mathrm{Cl}$ such that, as a $\mathrm{St}_N \times \mathrm{Cl}$ representation,

$$\mathcal{H}_N \simeq \bigoplus_{\tau \in \mathrm{Irr}\ \mathrm{St}(N^\perp/N)} \tau \otimes \eta_N(\tau).$$

This lemma tells us that each code gives rise to its own "lower rank" version of eq. (7). This lemma can be seen as generalizing the spirit of eq. (6). Namely, that equation together with the formalism of [GH17] gives a similar result: each code C_N gives rise to an eta correspondence η_N of rank $t - 2\dim N$. Lem. 1.1 translates this to the Clifford case and works uniformly for all d.

Here it is important to note two things. First, the codes C_N form tensor power representations whenever d is odd or whenever $d = 2$ and t is not a multiple of 4 (see [MMG21b, Lems. III.4, III.5]). Second, the proof of this statement fails whenever $d = 2$, $t = 4k$ and $\mathbf{1}_t \in N$ (see. [Rem. III.1][MMG21b]). Because of this, this lemma allows us to analyze code spaces associated to spaces N which contain the all-ones vector—something that Lems. III.4 and III.5 did not allow for.

To obtain the main theorem from Lem. 1.1, the theory of rank – developed in [GH17, GH20] – was generalized to the Clifford group. We may thus assign a rank $\mathrm{rk}\ \rho$ to any Cl representation ρ. The details of this generalization are covered in Sec. IV of [MMG21b]. Having defined the rank of a Clifford representation, I proceeded to generalize the proof techniques used in [GH17]. In particular, Ref. [MMG21b] shows that $\mathrm{rk}\ C_N = t - 2\dim N$ and that $\mathcal{H}_N \subseteq C_N$ is the sum of all irreducible blocks with maximal rank.

This development is important in two regards. The first aspect is practical: it allows us to leverage the proof of [MMG21a, Lem. 3.3] in order to show the structure of Cl-isotypes as $\mathrm{St}(\mathbb{Z}_d^t)$-representations. Namely, the isotype corresponding to a Cl

representation $\eta(\tau)$ with rank $t - 2m$ is

$$\mathrm{Ind}_{\mathrm{St}(N^\perp/N)}^{\mathrm{St}(\mathbb{Z}_d^t)}(\tau) \otimes \eta(\tau),$$

where N is a stochastic isotropic subspace with $\dim N = m$, and $\mathbf{1}_t \in N$ if and only if $\mathcal{P} \subseteq \ker \eta(\tau)$. The second aspect is more conceptual: it suggests that the tools used to study representations of the finite symplectic groups may be generalized to study the Clifford group as well. This could be an exciting avenue—using the wide array of tools that have been developed to study the repesentation theory of $\mathrm{Sp}(\mathbb{Z}_d^{2n})$ in order to clarify the much-less studied representation theory of the Clifford group.

1.3 Efficient approximate unitary designs

The Clifford group has received attention from the quantum information and the representation theoretical communities because its property of being a *unitary design* [Zhu17, ZKGG16, Web16, KG15, GAE07], and its relation to the classification of finite group designs [BNRT20]. A set of unitaries $\{U_i \in \mathrm{U}(d)\}_{i=1}^M$ is said to be a *unitary t-design* if it holds that

$$\frac{1}{M}\sum_i U_i^{\otimes(t,t)} = \int_{\mathrm{U}(d)} \mathrm{d}\mu_{\mathrm{Haar}}(U) U^{\otimes(t,t)}, \tag{8}$$

where $U^{\otimes(t,t)} = U^{\otimes t} \otimes \bar{U}^{\otimes t}$, as before, and where μ_{Haar} is the Haar measure on $\mathrm{U}(d)$. If the set $\{U_i\}$ forms a group then it is called a *unitary t-group*.

The defining condition, eq. (8), becomes rather restrictive in the case of unitary t-groups. Indeed, in this case, the equation translates to the representation-theoretical condition that the representation $U_i \mapsto U_i^{\otimes t}$ has the same decomposition as the same tensor power representation of the unitary group. Ref. [BNRT20] formalizes this idea, showing that there are few families of unitary 2- or 3-groups, and only a finite number of instances of higher order unitary t-groups. Prominently, the Clifford group is among the few unitary 3-groups.

The fact that the Clifford group is a unitary 3-design has been of interest in the recent years in quantum computing [Zhu17, Web16, KG15]. Recently, the fourth tensor power representation of the Clifford group was studied in detail in [ZKGG16]. There, it was shown that while the Clifford group is not a 4-design, it is sufficiently similar to one to allow many applications such as matrix reconstruction [KZG16b] and quantum state discrimination [KZG16a]. This work gives the intuition that although the Clifford group fails to be a higher-order design, it nevertheless may be used as a starting point from which to arrive at higher-order designs.

Furthermore, by describing the structure of the commutant of the fourth tensor power representation of the Clifford group (Δ_4 in the notation of Sec. 1.2.2), this work

shows how to produce *complex projective* 4*-designs*. In particular, it proposes an algorithm that generates a constant number of Clifford orbits $\{U\,|\psi_i\rangle\}_{i,U}$, where i labels the different orbits, which satisfy

$$\mathbb{E}_{i,U}[(U\,|\psi_i\rangle\langle\psi_i|\,U^\dagger)^{\otimes 4}] = \frac{1}{\operatorname{tr} \mathrm{P}_{\mathrm{Sym}}}\,\mathrm{P}_{\mathrm{Sym}}, \tag{9}$$

where $\mathrm{P}_{\mathrm{Sym}}$ is the projector onto the totally symmetric vectors (i.e. vectors in $(\mathbb{C}^{2^n})^{\otimes 4}$) which are invariant under permutations of the 4 tensor factors). A follow-up work, [GNW21], studies the commutant of higher tensor power representations of the Clifford group. There, it is proven that there exist certain small families of Clifford orbits – with a number of orbits only depending on t and not n – that form complex projective designs. The question of constructing higher order *unitary* designs based on the Clifford group has, however, been left open in the literature. Chap. 3 provides an answer to this question, providing an explicit construction of *approximate* unitary t-designs.

A set $\{U_i\}$ of unitaries is an ϵ-approximate unitary t-design, if it satisfies

$$\left\|\mathbb{E}_i[\mathrm{Ad}_{U_i}^{\otimes t}] - \mathrm{P}_{\mathrm{Haar}}\right\|_\diamond \leq \epsilon, \tag{10}$$

where $\mathrm{Ad}_U(\,\cdot\,) = U\,\cdot\,U^\dagger$ is the quantum channel corresponding to U and

$$\mathrm{P}_{\mathrm{Haar}} := \mathbb{E}_{U\sim\mu_{\mathrm{Haar}}}[\mathrm{Ad}_U^{\otimes t}]$$

is the projector onto the commutant of the t-th tensor power representation of $\mathrm{U}(2^n)$. In eq. (10), we used the *diamond norm*, defined on quantum channels on an n-qubit system $\mathbb{C}^{2^n}$ by

$$\|\mathcal{A}\|_\diamond = \max_\rho \|(\mathcal{A}\otimes\mathbb{1})(\rho)\|_1,$$

where the maximization is over $2n$-qubit states $\rho \in \mathrm{End}(\mathbb{C}^{2^n}\otimes\mathbb{C}^{2^n})$, and where $\|\,\cdot\,\|_1$ is the *Schatten 1-norm* or *trace-norm*.

Approximate t-designs are arguably easier to construct explicitly. Moreover, approximate designs are sufficient for most applications in quantum information processing. A fast way to convince oneself that this is true is to notice that any application that would require implementing an *exact* design, would be out of reach as soon as even a small amount of noise is present in the quantum computer.

The seminal paper [BHH16] shows that *local random quantum circuits* of depth $\tilde{O}(n^2t^9\log(1/\epsilon))$ are ϵ-approximate unitary t-designs. In other words one may sample from approximate unitary designs effectively. This result is in contrast to the task of sampling Haar-random unitaries—here, a crude lower bound for the circuit length k is given by requiring $\log \mathrm{vol}(\mathrm{U}(2^n)) \leq \log\epsilon + k\log(\#\text{ local gates})$, giving $k \geq O(2^n/n)$.

In a nutshell: while Haar random unitaries put the "exp" on "expensive," approximate t-desings can be cheap.

In Chap. 3 this result is improved upon. Rather than starting from local random quantum circuits, the basic building blocks here are random Clifford unitaries. The main result is an explicit construction of random quantum circuits that give ϵ-approximate unitary t-designs. These quantum circuits are dominated by Clifford gates, using a number of single-qubit non-Clifford gates that depends only on t and ϵ, but not on n. Translating this into the language of quantum computing with magic states, the amount of magic required for sampling from an approximate unitary design is independent of the size n of the quantum computer. Because magic is a valuable and costly resource in quantum computers, this construction provides an advantage over the circuits considered in [BHH16], whose $\sim n^2$ gates are non-Clifford with unit probability. Furthermore, several properties of these circuits may in principle be efficiently simulated by low stabilizer-rank decomposition methods [BBC$^+$19].

Consider K-interleaved random Clifford circuits of the form $U = K_k U_k \cdots K_1 U_1$, where U_i is a random n-qubit Clifford and K_i is sampled uniformly from $\{\mathbb{1}, K, K^\dagger\}$. Here k is called the *depth* of the random circuit. Then the main result of Chap. 3 is the following.

Theorem 1.6. *Let T be a non-Clifford single-qubit gate. Then, there exist constants $C_1(K)$ and $C_2(K)$ such that, K-interleaved random Clifford circuits of depth*

$$k \geq C_1(K) \log^2(t)(t^4 + t \log(1/\epsilon))$$

are ϵ-approximate t-designs for all $n \geq C_2(K)t^2$.

The proof of Thm 1.6 relies heavily on the characterization of the commutant of Δ_t given in [GNW21] (cf. Thm. 1.4). Consider the average quantum channel Φ_k for a K-interleaved random Clifford circuit of depth k. Denoting,

$$\mathrm{P}_{\mathrm{Cl}} = \frac{1}{|\mathrm{Cl}|} \sum_{U \in \mathrm{Cl}} \mathrm{Ad}_U^{\otimes t}, \quad \mathrm{R}(K) = \frac{1}{3}(\mathrm{Ad}_K^{\otimes t} + \mathrm{Ad}_{K^\dagger}^{\otimes t} + \mathbb{1}_1) \otimes \mathbb{1}_{n-1},$$

we can see that

$$\Phi_k = (\mathrm{P}_{\mathrm{Cl}} \mathrm{R}(K))^k.$$

Notice that P_{Cl} projects onto the commutant of the t-th tensor power representation of the Clifford group. Because the Clifford group together with K generate $\mathrm{U}(2^n)$ [NRS01], it is clear that

$$\lim_{k \to \infty} (\mathrm{P}_{\mathrm{Cl}} \mathrm{R}(K))^k = \mathrm{P}_{\mathrm{Haar}}.$$

Indeed, by the same argument, the Hermitian operator $\mathrm{P_{Cl}R}(K)\mathrm{P_{Cl}}$ has $\mathrm{P_{Haar}}$ exactly as the eigenspace corresponding to the eigenvalue of one, and all of its other eigenvalues are smaller than one. From this point of view, this proof gives a bound on the speed with which this convergence happens in the diamond norm.

Here I summarize the proof. It first proceeds by finding an alternative formula for $\mathrm{P_{Cl}}$, one that is arguably more workable. The basis $\mathcal{A}_t$ for $\mathrm{range}\ \mathrm{P_{Cl}}$ found in [GNW21] (cf. Thm. 1.4) is asymptotically orthogonal for high values of n. Write the normalized basis elements in $\mathcal{A}_t$ as A_i, where $i = 1, \ldots, |\mathcal{A}_t|$. Then,

$$\mathrm{tr}\, A_i^\dagger A_j \sim 2^{-n}, \quad i \neq j.$$

Because of this, one may use Gram-Schmidt to find an orthonormal basis E_i for $\mathrm{range}\ \mathrm{P_{Cl}}$, rather close to the basis A_i. Namely, in [HMMH$^+$20, Lem. 4] a Gram-Schmidt transformation

$$\begin{pmatrix} \mathrm{G}_{1,1} & \cdots & \mathrm{G}_{1,|\mathcal{A}_t|} \\ & \ddots & \\ \mathrm{G}_{|\mathcal{A}_t|,1} & \cdots & \mathrm{G}_{|\mathcal{A}_t|,|\mathcal{A}_t|} \end{pmatrix} \begin{pmatrix} A_1 \\ \vdots \\ A_{|\mathcal{A}_t|} \end{pmatrix} = \begin{pmatrix} E_1 \\ \vdots \\ E_{|\mathcal{A}_t|} \end{pmatrix}$$

is constructed, whose off-diagonal elements are exponentially small in n.

Using this basis, we may express

$$\Phi_k = \Big(\sum_i |E_i\rangle\langle E_i|\, \mathrm{R}(K) \Big)^k.$$

The basis $\mathcal{A}_t$ contains as a subset the restricted representation $R(S_t)$, where S_t is the subgroup of permutations in $\mathrm{St}(\mathbb{Z}_d^t)$. As shown in Chap. 3, without loss of generality we may take $E_1, \ldots,\ E_{t!}$ to span this same subspace,

$$\mathrm{span}\{E_1, \ldots,\ E_{t!}\} = \mathbb{C}[R(S_t)] = \mathrm{range}\ \mathrm{P_{Haar}},$$

where the right-hand-side equality follows by Schur-Weyl duality. This way,

$$\begin{aligned} \|\Phi_k - \mathrm{P_{Haar}}\|_\diamond &= \|((\mathrm{P_{Cl}} - \mathrm{P_{Haar}})\mathrm{R}(K))^k\|_\diamond \\ &= \left\| \left(\sum_{j>t!} |E_j\rangle\langle E_j|\, \mathrm{R}(K) \right)^k \right\|_\diamond \\ &\leq \sum_{j_1,\ldots,j_k>t!} \big\| \, |E_{j_1}\rangle\langle E_{j_k}| \, \big\|_\diamond \cdot \prod_{r=1}^{k-1} \langle E_{j_r}|\, \mathrm{R}(K)\, |E_{j_{r+1}}\rangle \end{aligned}$$

where the last inequality follows by using subadditivity of the diamond norm. Our goal is, then, to bound the right-hand side of this inequality.

To prove Thm. 1.6, the bound on $|\mathrm{G}_{ij}|$ ([HMMH$^+$20, Lem. 4], as mentioned above) is carefully combined with bounds on

$$\langle A_i| \,\mathrm{R}(K)\, |A_j\rangle\,, \quad \text{and,} \quad \big\|\, |A_i\rangle\langle A_j| \,\big\|_\diamond. \tag{11}$$

The latter bounds follow from Lems. 2 and 3 in [HMMH$^+$20]. The details of these combinations are outside of the scope of this introductory chapter.

2 Representations of the Clifford and symplecitc groups

This chapter has two sections. The first was published as [MMG21a]

> Montealegre-Mora, F., Gross, D. (2021). *Rank-deficient representations in the Theta correspondence over finite fields arise from quantum codes.* Representation Theory of the American Mathematical Society, 25(8), 193-223.

The second is a manuscript entitled *The representation theory of Clifford tensor powers*, written together with David Gross. This manuscript is currently in the final stages of preparation and will be available at `https://arXiv.org` shortly after the publication of this thesis.

These two works belong to a same research program to understand the representation theory of the Clifford group and, in particular, tensor powers of its defining representation.

The first work focuses one tensor power representations of the *oscillator representation* of the finite symplectic group over odd characteristic. In particular, we extend the η *correspondence* introduced in [GH17] to understand certain "maximal rank subrepresentations," to a full decomposition of the representation (cf. Sec. 2.3 of [MMG21a] for the relevant definition of rank). The connection between Clifford and oscillator tensor powers is made in Sec. 4 of [MMG21a].

The main theorem in this first paper was inspired by the results of [GNW21]. These results say, in a nutshell, that the commutant of an oscillator tensor-power representation is generated by two kinds of operators: *1.* a representation of a finite orthogonal group, *2.* a set of projectors onto certain CSS quantum codes. Because of the presence of these CSS projectors, the Theta correspondence between the symplectic and orthogonal groups does not hold in this scenario—that is, Howe duality breaks when considering the finite orthogonal-symplectic dual pair. Our main theorem in [MMG21a] is a formalization of the idea that these CSS codes, together with the η correspondence of [GH17], can be used to fully decompose oscillator tensor powers.

While our first paper was based on the intuitions gained in [GNW21], the proof techniques were significantly different. Because of this, [MMG21a] holds in a slightly different scenario than [GNW21]. Namely, it holds for symplectic groups over arbitrary finite fields of odd charactersitic, whereas [GNW21] holds over arbitrary prime fields.

Our second work is the manuscript [MMG21b] currently in the final stages of preparation. It was initially an effort to generalize the results of [MMG21a] to the important case of characteristic 2 (that is, the case of the qubit Clifford group). In the course of this work, however, we extended our previous results in several regards. These are covered in the introduction of the paper. Importantly, the proof strategies

used in this second work differ substantially from the proof strategies in [MMG21a]. Namely, while the proof of the main theorem in [MMG21a] is self-contained, the proof of the main theorem in the second work relies heavily on [GNW21, GH17].

These works were the outcome of a collaborative effort between myself and David Gross, my thesis advisor. I was the lead researcher in both of these projects.

REPRESENTATION THEORY
An Electronic Journal of the American Mathematical Society
Volume 25, Pages 193–223 (March 25, 2021)
https://doi.org/10.1090/ert/563

RANK-DEFICIENT REPRESENTATIONS IN THE THETA CORRESPONDENCE OVER FINITE FIELDS ARISE FROM QUANTUM CODES

FELIPE MONTEALEGRE-MORA AND DAVID GROSS

Abstract. Let V be a symplectic vector space and let μ be the *oscillator representation* of $\mathrm{Sp}(V)$. It is natural to ask how the tensor power representation $\mu^{\otimes t}$ decomposes. If V is a real vector space, then the theta correspondence asserts that there is a one-one correspondence between the irreducible subrepresentations of $\mathrm{Sp}(V)$ and the irreps of an orthogonal group $O(t)$. It is well-known that this duality fails over finite fields. Addressing this situation, Gurevich and Howe have recently assigned a notion of *rank* to each $\mathrm{Sp}(V)$ representation. They show that a variant of the Theta correspondence continues to hold over finite fields, if one restricts attention to subrepresentations of maximal rank. The nature of the rank-deficient components was left open. Here, we show that all rank-deficient $\mathrm{Sp}(V)$-subrepresentations arise from embeddings of lower-order tensor products of μ and $\bar{\mu}$ into $\mu^{\otimes t}$. The embeddings live on spaces that have been studied in quantum information theory as tensor powers of *self-orthogonal Calderbank-Shor-Steane (CSS) quantum codes*. We then find that the irreducible $\mathrm{Sp}(V)$-subrepresentations of $\mu^{\otimes t}$ are labelled by the irreps of orthogonal groups $O(r)$ acting on certain r-dimensional spaces for $r \leq t$. The results hold in odd charachteristic and the "stable range" $t \leq \frac{1}{2}\dim V$. Our work has implications for the representation theory of the *Clifford group*. It can be thought of as a generalization of the known characterization of the invariants of the Clifford group in terms of self-dual codes.

1. Introduction and summary of results

The *oscillator representation* (also: *Schrödinger*, *Weil*, or *metaplectic* representation) is a representation μ_V of the symplectic group $\mathrm{Sp}(V)$ over a symplectic vector space V. It appears in many contexts, including time-frequency analysis, coding theory, and quantum mechanics.

The starting point of this work is the natural question of how tensor powers $\mu_V^{\otimes t}$ decompose into irreducible representations.

One may reformulate this problem in a more geometric and slightly more general way [11]. If U is an orthogonal space, then $U \otimes V$ is again symplectic. The tensor power $\mu_V^{\otimes t}$ is isomorphic to $\mu_{U\otimes V}$ for a suitable t-dimensional space U (Corollary 2.3). The symmetry group $O(U) \times \mathrm{Sp}(V)$ associated with the tensor factors embeds into $\mathrm{Sp}(U \otimes V)$. Clearly, the restriction of $\mu_{U\otimes V}$ to $O(U)$ commutes with the restriction to $\mathrm{Sp}(V)$. One can thus decompose the representation into a direct

Received by the editors June 24, 2020, and, in revised form, November 23, 2020.
2020 *Mathematics Subject Classification.* Primary 20C33; Secondary 20G40.
This work was by the Excellence Initiative of the German Federal and State Governments (Grant ZUK 81), the ARO under contract W911NF-14-1-0098 (Quantum Characterization, Verification, and Validation), and the DFG (SPP1798 CoSIP, project B01 of CRC 183).

sum

$$\mu_{U\otimes V} \cong \bigoplus_{\tau\in \mathrm{Irr}(O(U))} \tau \otimes \Theta(\tau), \tag{1.1}$$

where τ ranges over irreps of $O(U)$, and $\Theta(\tau)$ is a representation of $\mathrm{Sp}(V)$. If U, V are real spaces and the form on U is definite, the theta correspondence asserts that $\Theta(\tau)$ is again irreducible, and that the correspondence Θ between representations is injective [14, 15]. Over finite fields, the correspondence fails: $\Theta(\tau)$ is in general no longer irreducible, and equivalent $\mathrm{Sp}(V)$ representations might appear in $\Theta(\tau)$ for different τ's. Our goal is to understand this situation better.

The main part of this paper is presented in the basis-free notation set out in [11]. For ease of exposition, we will use more concrete (and slightly less general) constructions in this introductory section. From now on, we assume that $V = \mathbb{F}_q^{2n}$ is $2n$-dimensional over a finite field $\mathbb{F}_q$ of odd characteristic, and endowed with a symplectic form.

Reference [11] introduces a notion of *rank* for $\mathrm{Sp}(V)$ representations. To describe it, recall that the oscillator representation of $\mathrm{Sp}(V)$ can be realized over the Hilbert space $\mathcal{H} = \mathbb{C}[\mathbb{F}_q^n]$ of complex linear combinations of basis vectors δ_x labeled by vectors $x \in \mathbb{F}_q^n$ (Sec. 2). Given $x_1, \ldots, x_t \in \mathbb{F}_q^n$, we may arrange these vectors as the rows of a $t \times n$ matrix F. This way, we obtain an isormorphism

$$\mathcal{H}^{\otimes t} = \mathbb{C}[\mathbb{F}_q^n]^{\otimes t} \simeq \mathbb{C}[\mathbb{F}_q^{t\times n}] \tag{1.2}$$

via the identification

$$\delta_{x_1} \otimes \cdots \otimes \delta_{x_t} \simeq \delta_F.$$

The *rank* of an element $\psi \in \mathcal{H}^{\otimes t}$ and of a subspace $\mathcal{K} \subset \mathcal{H}^{\otimes t}$ are defined as, respectively,

$$\operatorname{rank}\psi = \sup\{\operatorname{rank} F^T F \mid (\delta_F, \psi) \neq 0\}, \qquad \operatorname{rank}\mathcal{K} = \sup\{\operatorname{rank}\psi \mid \psi \in \mathcal{K}\}.$$

The central result of [11] is this:

Theorem 1.1 ([11])**.** *Assume $t \leq n$. Then $\Theta(\tau)$ contains a unique irreducible representation $\eta(\tau)$ of rank t. The function η defines an injective map from the irreducible representations of $O(U)$ to the irreducible rank-t subrepresentations of $\mathrm{Sp}(V)$ in $\mu_{U\otimes V}$.*

The purpose of this work is to understand the *rank-deficient* $\mathrm{Sp}(V)$-subrepresentations of $\mu_{U\otimes V}$, i.e. those that have rank $r < t$. Key to this are *self-orthogonal Calderbank-Shor-Steane (CSS) quantum codes* [3, 24, 25], which are studied in the theory of quantum error correction [22]. For now, we will take $U = \mathbb{F}_q^t$ with the standard orthogonal form $\beta(u, v) = \sum_{i=1}^t u_i v_i$. Let N be an *isotropic subspace* of U, i.e. such that $N \subset N^\perp$. To each coset $[u] = u + N \subset U$ of N, one associates the *coset state*

$$e_{[u]} = \sum_{v\in[u]} \delta_v \in \mathbb{C}[\mathbb{F}_q^t].$$

Analogous to the construction in Eq. (1.2), we identify

$$\mathbb{C}[\mathbb{F}_q^t]^{\otimes n} \simeq \mathbb{C}[\mathbb{F}_q^{t\times n}], \qquad \delta_{u_1} \otimes \cdots \otimes \delta_{u_n} \simeq \delta_F,$$

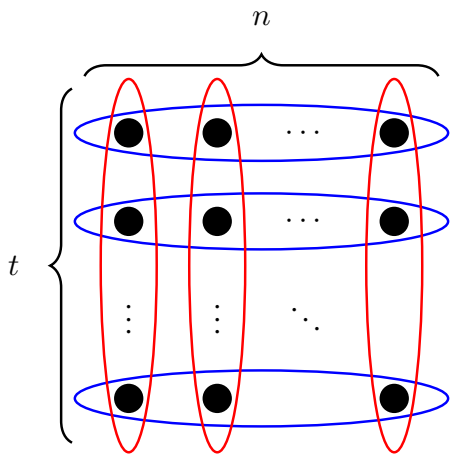

FIGURE 1. Sketch of the commuting actions of the Weil representation and tensor-power CSS codes. The each tensor factor in the representation $\mu_V^{\otimes t}(S)$ for an arbitrary S acts on a row (highlighted in blue). The code projector is an n-th tensor power of a projector supported on a column (red).

where F is now the matrix whose *columns* are given by the $u_1, \ldots, u_t \in \mathbb{F}_q^n$. The *tensor power CSS code* C_N associated with N is the space with basis

$$\left\{ e_{[u_1]} \otimes \cdots \otimes e_{[u_n]} \;\middle|\; [u_i] \in N^\perp/N \right\}, \tag{1.3}$$

the set of products of coset states corresponding to the elements of the quotient space $N^\perp/N$.

The codes C_N can be shown to be invariant subspaces of $\mu_{U\otimes V}|_{\mathrm{Sp}(V)}$. What is more, we will show:

Lemma (Lemma 2.7, simplified version)**.** *As a representation of* $\mathrm{Sp}(V)$*, the restriction of* $\mu_{U\otimes V}$ *to a tensor power CSS code* C_N *is isomorphic to* $\mu_{U'\otimes V}$*, where* $U' = N^\perp/N$*.*

Figure 1 displays graphically the commuting actions of the projector P_N onto an arbitrary CSS code C_N and $\mu_V^{\otimes t}(S)$ for an arbitrary S. There, we identify

$$\mathbb{C}[\mathbb{F}_q]^{\otimes nt} \simeq \mathbb{C}[\mathbb{F}_q^{t\times n}], \tag{1.4}$$

in an analogous way as above. Each dot in the diagram corresponds to a $\mathbb{C}[\mathbb{F}_q]$ factor in the left-hand side of (1.4). The tensor factors in the Weil representation $\mu_V^{\otimes t}$ act row-wise, highlighted in blue, whereas the projector P_N acts column-wise, highlighted in red.

(We note that in odd characteristic, there are two inequivalent orthogonal geometries in each dimension. They are distinguished by their *discriminant*, the square class of the determinant of the Gram matrix of any basis. So far, we have only considered the standard orthogonal form on $U = \mathbb{F}_q^t$. It turns out that $U' = N^\perp/N$ inherits an orthogonal form from U – however, it need not be equivalent to the standard one. We will deal with this more general situation in the main part.)

The lemma immediately implies that non-trivial CSS codes carry rank-deficient representations of the symplectic group. Our main result is that this construction is exhaustive.

Theorem 1.2 (Main theorem)**.** *Assume that $t \leq n$ and let $\mathcal{K}$ be an $\mathrm{Sp}(V)$-subrepresentation of $\mu_{U\otimes V}$ of rank r. Then $(t-r)$ is even and $\mathcal{K}$ is contained in the span of all tensor power CSS codes C_N with $\dim N = (t-r)/2$.*

The result allows us to give an explicit decomposition of $\mu_{U\otimes V}$ in terms of irreducible and inequivalent $\mathrm{Sp}(V)$ representation spaces. Indeed, we find (Sec. 3.3) that as an $O(U)\times \mathrm{Sp}(V)$ representation:

$$\mu_{U\otimes V} \simeq \bigoplus_{r\in R(U)} \bigoplus_{\tau\in \mathrm{Irr}\, O(U_r)} \mathrm{Ind}_{O_r}^{O(U)}(\tau)\otimes \eta(\tau). \tag{1.5}$$

We have used the following expressions: $R(U)$ is the set $\{t-2k\}_k$, where k ranges from 0 to the largest dimension of an isotropic subspace in U (its *isotropy index*). For each k, we choose some istropic subspace $N\subset U$ of dimension k and set $U_r = N^\perp/N$. Then U_r is an orthogonal space of dimension $r = t-2k$ and discriminant $d(U_r) = (-1)^k d(U)$. Let $O_r := O_N \subset O(U)$ be the stabilizer of N. Notice that because of a lemma proven by Witt, the group O_r is independent of the choice of N, up to isomorphism. This justifies surpressing N in our notation. The group O_r acts on U_r as $O(U_r)$.

Thus any $\tau \in \mathrm{Irr}\, O(U_r)$ can be interpreted as an O_r-representation, and the induced representation in Eq. (1.5) is hence well-defined. All $\mathrm{Sp}(V)$-irreps $\eta(\tau)$ appearing in Eq. (1.5) are indeed inequivalent: Those corresponding to different $O(U_r)$ are distinguished by their rank, whereas the inequivalence of summands of the same rank is a consequence of Theorem 1.1.

It is natural to ask whether the assumption that $t\leq n$ is necessary. We show that some constraints on t, n are indeed required, by explicitly constructing rank-0 (i.e. trivial) subrepresentations for $t = 3, n = 1$ that do not come from CSS codes (Section 3.4).

Our work was motivated by recent related observations on tensor powers of the *Clifford group* [12, 16, 19–21, 23, 26–28], the group generated by the oscillator representation of $\mathrm{Sp}(V)$ and the Weyl representation of the Heisenberg group. In [10], it has been shown that the commutant algebra of the Clifford group is generated by projections onto tensor power CSS codes whose isotropic spaces are orthogonal to the all-ones vector $\underline{1} = (1,\ldots,1)\in \mathbb{F}_q^t$; together with the elements of $O(U)$ that preserve $\underline{1}$. While it was not explicitly worked out in [10], their arguments strongly suggest that the commutant of the oscillator representation alone is generated by $O(U)$ and tensor power CSS codes, without the constraints involving the $\underline{1}$-vector. This drew our attention to the action of tensor power representations on CSS code spaces. While the present paper mostly focuses on the symplectic group alone—instead of the full Clifford group—one can in some cases relate the theory for the two groups explicitly (Sec. 4):

Proposition (Proposition 4.2, simplified version)**.** *If the characteristic of $\mathbb{F}_q$ does not divide t, there is a one-one correspondence between irreducible components of t-th tensor powers of the Clifford goup and irreducible components of $\mu_{U\otimes V}$ for a certain orthogonal space U of dimension $t-1$.*

An $\mathrm{Sp}(V)$-representation space is trivial if and only if it has rank equal to 0 [11]. The rank-0 case connects our results with prior work on the *invariants* of the Clifford group [20, 23]. Indeed, it is well-knwon that the invariants are associated with self-*dual* CSS codes, i.e. those arising from subspaces $N\subset U$ with $N^\perp = N$.

In this sense, our work can be seen as a generalization of these results to higher ranks.

Our main theorem is based on a careful analysis of the action of certain Fourier transforms in the oscillator representation. The same techniques can be used to find auxilliary results, which may be of independent interest. For example, we show that the "set of ranks" one can associate with an irreducible $\mathrm{Sp}(V)$-subrepresentation of $\mu_{U\otimes V}$ is a contiguous set of integers (Prop. 2.10).

The rest of the paper is organized as follows: We will introduce the technical background in Sec. 2. Our original contributions are in Sec. 3, where we prove the main theorem, and in Sec. 4 laying out the connections to the Clifford group.

This work is written in a basis-free language inspired by [11]. We believe that the results will be of interest to researchers in quantum information theory, who may not be familiar with this point of view. A follow-up paper [18] will address a quantum information audience, both in terms of presentation and in terms of applications. In particular, it will also treat the Clifford group in characteristic 2.

2. Technical background

In this section, we collect definitions and some technical statements. While we are not aware of references for every specific result, the material presented here seems to be known in the general literature.

2.1. **General notation.** In what follows, q is the power of an odd prime p, and $\mathbb{F}_q$ the finite field of order q. We denote the multiplicative group in $\mathbb{F}_q$ by $\mathbb{F}_q^\times$. For $\lambda \in \mathbb{F}_q^\times$, the *Legendre symbol* is $\left(\frac{\lambda}{q}\right)$, which is $+1$ if λ is a square in $\mathbb{F}_q^\times$, and -1 otherwise. If q is clear from the context, we also use the short-hand notation ℓ_λ for the Legendre symbol. We write $\mathrm{Tr} : \mathbb{F}_q \to \mathbb{F}_p$ for the the trace in $\mathbb{F}_q$ over $\mathbb{F}_p$ (also known as the *field trace*).

The *transpose* of a linear map $A : Y \to Z$ is $A^* : Z^* \to Y^*$ (not to be confused with A^T, which is defined in Eq. (2.12)). A map $A : Y \to Y^*$ is *symmetric* if $A = A^*$.

2.2. **The oscillator representation.** Let $V = X \oplus X^*$ be the direct sum of two n-dimensional dual vector spaces over $\mathbb{F}_q$. The space V carries a symplectic form

$$[x \oplus y, x' \oplus y'] = y'(x) - y(x').$$

Every symplectic vector space is (non-canonically) of this form. Indeed, the choice of a decomposition $V = X \oplus X^*$ is equivalent to fixing a *polarization* of V. From now on, we will assume that dual $X, X^* \subset V$ have been chosen.

The *oscillator representation* μ_V is a representation of $\mathrm{Sp}(V)$ on the Hilbert space $L^2(X^*)$ of complex functions on X^*. The representation depends on a parameter $m \in \mathbb{F}_q^\times$ – sometimes referred to as the *mass* of the representation in mathematical physics [7] – which defines a character

$$\omega^{(m)} : \mathbb{F}_q \to \mathbb{C}, \qquad \lambda \mapsto e^{i\frac{2\pi}{p}\mathrm{Tr}(m\lambda)}$$

of $\mathbb{F}_q$. One can show [11] that the oscillator representations $\mu_V^{(m)}$, $\mu_V^{(m')}$ are unitarily equivalent if and only if m and m' belong to the same square class. What is more, $\mu_V^{(-m)} = \bar{\mu}_V^{(m)}$, i.e. the inverting the sign of the mass corresponds to passing to the complex conjugate representation. From now on, we will write ω, μ_V for $\omega^{(1)}$ and $\mu_V^{(1)}$ respectively.

Next, we recall [8, 11] the explicit form of the oscillator representation on the following three subsets, which taken together generate $\mathrm{Sp}(V)$.

$$\mathcal{J} = \left\{ \begin{pmatrix} 0 & B \\ -B^{-1} & 0 \end{pmatrix} \,\middle|\, B : X^* \to X,\ B \text{ invertible, symmetric} \right\}, \tag{2.1}$$

$$\mathcal{N} = \left\{ \begin{pmatrix} \mathbb{1} & A \\ 0 & \mathbb{1} \end{pmatrix} \,\middle|\, A : X^* \to X,\ A \text{ symmetric} \right\}, \tag{2.2}$$

$$\mathcal{D} = \left\{ \begin{pmatrix} C & 0 \\ 0 & C^{-*} \end{pmatrix} \,\middle|\, C \in GL(X) \right\}, \tag{2.3}$$

where we used the shorthand $C^{-*} := (C^*)^{-1}$. The sets $\mathcal{N}$ and $\mathcal{D}$ are subgroups and generate the *Siegel parabolic*, with the Abelian $\mathcal{N}$ the *unipotent radical* of the parabolic group. We write, respectively, N_A, J_B, D_C for the elements of $\mathcal{N}, \mathcal{J}, \mathcal{D}$ that appear above. Let $y \in X^*$ and let $\delta_y \in L^2(X^*)$ the indicator function at y. Then the action of the oscillator representation is

$$\mu_V(J_B)\,\delta_y = \gamma(B)^{-1} \sum_{y' \in X^*} \omega\big(B(y,y')\big)\delta_{y'}, \tag{2.4}$$

$$\mu_V(N_A)\,\delta_y = \omega\big(2^{-1}A(y,y)\big)\,\delta_y, \tag{2.5}$$

$$\mu_V(D_C)\,\delta_y = \ell_{\det C}\,\delta_{C^{-*}y}, \tag{2.6}$$

where $B(y,y')$ is a less-confusing notation for $B(y)(y')$, and where

$$\gamma(B) = \sum_{y \in X^*} \omega\big(-2^{-1}B(y,y)\big)$$

is the *Gauss sum* corresponding to B.

We will frequently make use of the fact that the oscillator representation of block matrices factorizes. This factorization property is well-known – see e.g. [8, Corollary 2.5] and [13]. We give a short self-contained proof in Appendix A.1.

Lemma 2.1. *Let $X = X_1 \oplus X_2$ be a direct sum of vector spaces. Then we have an orthogonal decomposition $V = V_1 \oplus V_2$ of $V = X \oplus X^*$ into symplectic subspaces $V_i = X_i \oplus X_i^*$. As a representation of the subgroup $\mathrm{Sp}(V_1) \times \mathrm{Sp}(V_2) \subset \mathrm{Sp}(V)$, the oscillator representation factorizes*

$$\mu_V \simeq \mu_{V_1} \otimes \mu_{V_2}. \tag{2.7}$$

Let $\pi_i : X \to X_i$ be the projections onto the i-th direct summand. An isomorphism

$$L^2(X^*) \to L^2(X_1^*) \otimes L^2(X_2^*)$$

realizing Eq. (2.7) is given by

$$\delta_y \mapsto \delta_{y\pi_1} \otimes \delta_{y\pi_2}. \tag{2.8}$$

2.3. **The rank of a representation.** We consider the subgroups $\mathcal{N}, \mathcal{D}$ of $\mathrm{Sp}(V)$ given in Eqs. (2.2), (2.3).

If π is a representation of $\mathrm{Sp}(V)$ on some Hilbert space $\mathcal{H}$, then the restriction of π to the Abelian group $\mathcal{N}$ decomposes $\mathcal{H}$ into a direct sum of one-dimensional representations. Every character of $\mathcal{N}$ is of the form

$$N_A \mapsto \omega\big(\operatorname{tr} AB\big)$$

for some symmetric $B : X \to X^*$, which we will refer to as an $\mathcal{N}$-*weight*. With each irreducible subrepresentation $\mathbb{C}\Phi \subset \mathcal{H}$, we can thus associate an $\mathcal{N}$-weight B such that

$$\pi(N_A)\Phi = \omega(\operatorname{tr} AB)\Phi, \quad \forall\, N_A \in \mathcal{N}.$$

Reference [11] defines the $\mathcal{N}$-*spectrum* of π as the set of $\mathcal{N}$-weights, counted with multiplicities, that occur in the decomposition of $\mathcal{H}$.

The set of $\mathcal{N}$-weights decomposes into a union of orbits under the action $B \mapsto CBC^*$, $C \in \mathrm{GL}(X)$. This follows from the fact that $\mathcal{D}$ normalizes $\mathcal{N}$:

$$D_C N_A D_C^{-1} = N_{CAC^*},$$

so that if Φ carries the $\mathcal{N}$-weight B, then $\pi(D_C)\Phi$ is associated with the $\mathcal{N}$-weight C^*BC. From the theory of quadratic forms, it is well-known that the orbits are labelled by the rank and the discriminant of B (c.f. Section 2.5).

The *rank of* π is the maximum of the rank taken over the $\mathcal{N}$-spectrum. If all $\mathcal{N}$-weights of maximal rank have the same discriminant d, π is said to have *discriminant* or *type* d.

As an example, we compute the $\mathcal{N}$-spectrum of the oscillator representation. By Eq. (2.5), the delta functions $\{\delta_y \mid y \in X^*\}$ diagonalize the restriction of μ_V to $\mathcal{N}$. We can re-write

$$A(y, y) = A(y)(y) = \operatorname{tr} A(y \otimes y).$$

The map $B = 2^{-1}\, y \otimes y$ is the most general form of a symmetric map $X \to X^*$ of rank ≤ 1 and of discriminant ℓ_2. Since $\pm y$ lead to the same B, the $\mathcal{N}$-spectrum consists of the following $\mathrm{GL}(X)$-orbits: $\{0\}$ occurs once, and the set of non-zero rank-1 B's of discriminant ℓ_2 occurs twice.

2.4. **Orthogonal spaces and higher-rank representations.** We recall some standard facts about orthogonal spaces over finite fields (see e.g. [4, 5, 17]) and fix notation.

Let U with be a t-dimensional $\mathbb{F}_q$-vector space with non-degenerate symmetric form β. Let $\{f_i\}_{i=1}^t$ be a basis of U. The square class $d(U)$ of the determinant of the matrix with elements $\beta(f_i, f_j)$ does not depend on the basis. It is called the *discriminant* of the form β. Quadratic spaces are characterized up to isometries by their dimension and discriminant. The discrimant is multiplicative: if $U_1 \oplus U_2$ is an orthogonal sum, then

$$d(U_1 \oplus U_2) = d(U_1)d(U_2).$$

One can find an *orthogonal basis* that diagonalizes the form in that

$$\beta(f_i, f_j) = d_i\, \delta_{i,j} \tag{2.9}$$

for suitable $d_i \in \mathbb{F}_q$. From the discussion above, it follows that one can choose

$$d_i = 1 \quad (i = 1, \ldots, t-1), \qquad d_t \in d(U), \tag{2.10}$$

and we will usually do so.

An important orthogonal space is the *hyperbolic plane* $\mathbb{H}$, which has dimension $t = 2$ and discriminant $d(\mathbb{H}) = -1$.

For a subspace $N \subset U$, its orthogonal complement is $N^\perp = \{u \mid \beta(u, v) = 0\, \forall v \in N\}$. The space N is *isotropic* if $N \subset N^\perp$. From the relation $\dim N + \dim N^\perp = t$,

valid for any non-degenerate form, one finds the dimension bound for isotropic spaces:

$$N \subset N^\perp \qquad \Rightarrow \qquad \dim N \leq \frac{t}{2}. \tag{2.11}$$

We will use the symbol β both to refer to the form $U \times U \to \mathbb{F}_q$ and to the induced isomorphism

$$\beta : U \to U^*, \qquad u \mapsto \beta(u) := \beta(u, \cdot).$$

For maps $F \in \mathrm{Hom}(Y \to U)$, we will write

$$F^T := F^* \circ \beta \in \mathrm{Hom}(U \to Y^*). \tag{2.12}$$

With $U \simeq \mathrm{Hom}(\mathbb{F}_q \to U)$ and $\mathbb{F}_q^* \simeq \mathbb{F}_q$, this implies in particular

$$u^T = u^* \circ \beta = \beta(u) = \beta(u, \cdot).$$

If the form β is degenerate, then the quotient space $U/\operatorname{rad}\beta$ of U by the radical of β is non-degenerate. The *rank* and the discriminant of U are then defined to be the dimension and the discriminant of the quotient space.

A symmetric map $B : X \to X^*$ defines a quadratic form $B(x, y) = B(x)(y)$ on a linear space X. Below, we will often be concerned with forms defined as $B = F^T F$ for some $F : X \to U$. In this case, B is the pull-back of β to X via F, and so we have

$$B(x, y) = (F^* \beta F)(x)(y) = \beta(Fx, Fy), \tag{2.13}$$

so that the rank and discriminant of such B are the rank and the discriminant of $\operatorname{range} F$ as a subspace of U.

Given a space $V = X \oplus X^*$ and an orthogonal space U, the tensor product $U \otimes V$ is again a direct sum of dual spaces and thus carries a symplectic form. Indeed,

$$U \otimes V \simeq (U \otimes X) \oplus (U \otimes X^*) \tag{2.14}$$

and the pairing between (factorizing) elements of the two summands is just

$$\langle u \otimes x, v \otimes y \rangle = \beta(u, v) y(x). \tag{2.15}$$

We will usually make the identification

$$U \otimes X = \mathrm{Hom}(X^* \to U), \qquad U \otimes X^* = \mathrm{Hom}(X \to U).$$

Then the pairing (2.15) between $Z \in \mathrm{Hom}(X^* \to U)$ and $F \in \mathrm{Hom}(X \to U)$ takes the form

$$\langle Z, F \rangle = \operatorname{tr} \beta Z F^*. \tag{2.16}$$

It follows that there is an oscillator representation $\mu_{U \otimes V}$ of $\mathrm{Sp}(U \otimes V)$ on $L^2(\mathrm{Hom}(X \to U))$.

From Eq. (2.15), one sees that $O(U) \times \mathrm{Sp}(V)$ embeds into $\mathrm{Sp}(U \otimes V)$. The main goal of this work is to understand the restriction of $\mu_{U \otimes V}$ to $\mathrm{Sp}(V)$.

We compute the $\mathcal{N}$-spectrum and rank of $\mu_{U \otimes V}$ as an $\mathrm{Sp}(V)$-representation. To this end, we must find the eigenspaces of $\mu_{U \otimes V}(\mathbb{1}_U \otimes N_A)$. Under the identification (2.14),

$$N_A = \begin{pmatrix} \mathbb{1} & A \\ 0 & \mathbb{1} \end{pmatrix} \in \mathrm{Sp}(V) \quad \Rightarrow \quad \mathbb{1} \otimes N_A \simeq \begin{pmatrix} \mathbb{1} \otimes \mathbb{1} & \mathbb{1} \otimes A \\ 0 & \mathbb{1} \otimes \mathbb{1} \end{pmatrix} \in \mathrm{Sp}(U \otimes V).$$

Thus, the embedding $\mathbb{1} \otimes N_A$ of $N_A \in \mathrm{Sp}(V)$ into $\mathrm{Sp}(U \otimes V)$ is again an element of the unipotent radical. The action of $\mu_{U\otimes V}(\mathbb{1} \otimes N_A)$ is thus also given by Eq. (2.5), this time acting on $L^2(\mathrm{Hom}(X \to U))$. Let $F \in \mathrm{Hom}(X \to U))$. With Eq. (2.16), we can express the quadratic form in Eq. (2.5) as

$$(\mathbb{1} \otimes A)(F)(F) = \langle F, (\mathbb{1} \otimes A)F\rangle = \langle F, FA\rangle = \mathrm{tr}\, \beta FAF^* = \mathrm{tr}\, F^T FA. \tag{2.17}$$

The $\mathcal{N}$-weight on δ_F is thus given by

$$B = 2^{-1}F^T F.$$

Conversely, the representation space

$$\left\{\Phi \in L^2(\mathrm{Hom}(X \to U)) \;\middle|\; \mu_{U\otimes V}(N_A)\Phi = \omega(\mathrm{tr}\, AB)\Phi\right\} \tag{2.18}$$

on which $\mathcal{N} \subset \mathrm{Sp}(V)$ acts with $\mathcal{N}$-weight B is equal to the span $\langle\{\delta_F \mid F^T F = B\}\rangle$ of the δ_F's with $F^T F = B$.

Notice that $\mathrm{rank}\, F^T F \leq \min(n, t)$. From now on, we will focus on the case where $t \leq n$ (this is referred to as the *stable range* in [11]), and call a representation of rank strictly smaller than t *rank-deficient*.

2.5. **Representations associated with direct sums of orthogonal spaces.** The original motivation of this work was to understand tensor power representations $\mu_V^{\otimes t}$. The more geometric language employed e.g. in [11] relates tensor factors to direct summands of orthogonal spaces. The Corollary 2.2 of Lemma 2.1 makes the connection precise.

Corollary 2.2. *Assume $U = U_1 \oplus U_2$ is an orthogonal direct sum. Then, as a representation of* $\mathrm{Sp}(V)$*, the oscillator representation factorizes as*

$$\mu_{(U_1\oplus U_2)\otimes V} \simeq \mu_{U_1\otimes V} \otimes \mu_{U_2\otimes V}. \tag{2.19}$$

Let $\pi_i : U \to U_i$ be the projections onto the direct summands. An isomorphism

$$L^2(\mathrm{Hom}(X \to U)) \to L^2(\mathrm{Hom}(X \to U_1)) \otimes L^2(\mathrm{Hom}(X \to U_2))$$

realizing Eq. (2.19) *is defined by*

$$\delta_F \mapsto \delta_{\pi_1 F} \otimes \delta_{\pi_2 F}. \tag{2.20}$$

Proof. By assumption, both terms U_i are non-degenerate β-spaces, so we have a canonical identifications $U_i^* \cong U_i$ and $\mathrm{Hom}(X \to U_i)^* \cong \mathrm{Hom}(X^* \to U_i)$. The latter identification satisfies that for any $h \in \mathrm{Hom}(X \to U_1)^*$ and any $f \in \mathrm{Hom}(X \to U_2)$, it holds that $h(f) = 0$ (and the same statement holds if we exchange U_1 and U_2).

This way, the advertised claim is a consequence of Corollary 2.1 for the decomposition

$$\begin{aligned}\mathrm{Hom}(X \to U) &= \mathrm{Hom}(X \to U_1) \oplus \mathrm{Hom}(X \to U_2)\\ \mathrm{Hom}(X \to U)^* &= \mathrm{Hom}(X^* \to U_1) \oplus \mathrm{Hom}(X^* \to U_2)\end{aligned}$$

which give rise to the following decomposition into symplectic subspaces

$$U \otimes V = (U_1 \otimes V) \oplus (U_2 \otimes V).$$

□

Iterating this observation over an orthogonal basis gives the connection between $\mu_{U\otimes V}$ and tensor powers of μ_V.

Corollary 2.3. *As a representation of* $\mathrm{Sp}(V)$*, we have that*

$$\mu_{U\otimes V} \simeq \underbrace{\mu_V \otimes \cdots \otimes \mu_V}_{(t-1)\times} \otimes \mu_V^{(d(U))}. \tag{2.21}$$

Let $\{f_i\}_{i=1}^t$ *be an orthogonal basis of* U *as in Eq.* (2.10)*. An isomorphism*

$$L^2(\mathrm{Hom}(X \to U)) \to \left(L^2(X^*)\right)^{\otimes t}$$

realizing Eq. (2.21) *is defined by*

$$\delta_F \mapsto \delta_{f_1^T F} \otimes \cdots \otimes \delta_{f_t^T F}. \tag{2.22}$$

Proof. Set $U_i = \mathbb{F}_q f_i$, so that $d(U_i) = \beta(f_i, f_i) = d_i$. The projections $\pi_i : U \to U_i$ are given by

$$u \mapsto d_i^{-1}\, f_i\, f_i^T(u).$$

Iterating Corollary 2.2 thus gives an isomorphism

$$i_1 : L^2(\mathrm{Hom}(X \to U)) \to \bigotimes_{i=1}^t L^2(\mathrm{Hom}(X \to U_i))$$

defined by

$$\delta_F \mapsto \delta_{f_1 f_1^T F} \otimes \cdots \otimes \delta_{f_{t-1} f_{t-1}^T F} \otimes \delta_{d(U)^{-1}\, f_t f_t^T F}.$$

We may identify $\mathrm{Hom}(X \to U_i) \simeq U_i \otimes X^*$ with X^* via $f_i \otimes y \mapsto y$. This induces an isomorphism

$$i_2 : \bigotimes_{i=1}^t L^2(\mathrm{Hom}(X \to U_i)) \to \left(L^2(\mathrm{Hom}(X^*)\right)^{\otimes t}.$$

Finally, let $C = d(U)^{-1}\,\mathbb{1} \in GL(X)$ and, using Eq. (2.6), let i_3 be $\mu_V(D_C)$ acting on the t-th tensor factor. Then the advertised isomorphism is $i_3\, i_2\, i_1$. □

Note that the standard inner product $\beta(x, y) = \sum_{i=1}^t x_i y_i$ on $\mathbb{F}_q^t$ has an ortho*normal* basis, and thus discriminant $d(\mathbb{F}_q^t) = 1$. Therefore,

$$\mu_{\mathbb{F}_p^t \otimes V} \simeq \mu_V^{\otimes t}. \tag{2.23}$$

We end this section by analyzing $\mu_{\mathbb{H}\otimes V}$, where $\mathbb{H}$ is the hyperbolic plane. To this end, define the *permutation representation* π of $\mathrm{Sp}(V)$ as the map that acts on $L^2(V)$ by sending the delta function δ_v at $v \in V$ to

$$\pi(S)\delta_v = \delta_{Sv}. \tag{2.24}$$

Lemma 2.4. *Let* $\mathbb{H}$ *be the hyperbolic plane. We then have:*

1. *As a representation of* $\mathrm{Sp}(V)$*,* $\mu_{\mathbb{H}\otimes V}$ *is isomorphic to the permutation representation.*
2. *If* $I \subset \mathbb{H}$ *is a non-zero isotropic space, then* $\mathrm{Sp}(V)$ *acts trivially on*

$$\psi_I := \sum_{F \in \mathrm{Hom}(X \to I)} \delta_F \in L^2(\mathrm{Hom}(X \to \mathbb{H})).$$

The second part of the lemma makes a connection between rank-deficient subrepresentations and isotropic spaces. Generalizations of this will be the central theme in the rest of this work.

This lemma is well known, see for example [2, Thm. 3.10] for item (1). For completeness, we include a self-contained proof based on the *Weyl representation* of the Heisenberg group introduced in Sec. 2.2.

Proof. By Eq. (4.2), the adjoint representation $\mathrm{Ad}_{\mu_V} : A \mapsto \mu_V A \mu_V^\dagger$ on $\mathrm{End}(L^2(X^*))$ permutes the Weyl operators $\{W_V(v)\}_{v\in V}$ and is thus isomorphic to the permutation representation π. But by Corollary 2.3,

$$\mu_{\mathbb{H}\otimes V} \simeq \mu_V \otimes \mu_V^{(d(\mathbb{H}))} = \mu_V \otimes \bar{\mu}_V \simeq \mathrm{Ad}_{\mu_V}.$$

This proves the first claim.

Next, note that the adjoint representation acts trivially on $W_V(0) = \mathbb{1}$. Our strategy is to show that for every isotropic space $I \subset \mathbb{H}$, one can choose the isomorphisms employed in the first part, to map $W_V(0)$ to ψ_I. Indeed, the isomorphism $\mathrm{Ad}_{\mu_V} \simeq \mu_V \otimes \bar{\mu}_V$ is implemented by

$$i_1 : \mathrm{End}(L^2(X^*)) \to L^2(X^*)^{\otimes 2}, \qquad \delta_y \otimes \delta_{y'}^T \mapsto \delta_y \otimes \delta_{y'},$$

where $\delta_{y'}^T$ is the map acting on $\psi \in L^2(X^*)$ as $\psi \mapsto \psi(y')$. Choose an orthogonal basis $\{f_1, f_2\} \subset \mathbb{H}$ as in Eq. (2.10) and let i_2 be the associated isomorphism defined in Corollary 2.3. Then

$$W_V(0) = \mathbb{1}_V = \sum_{y\in X^*} \delta_y \otimes \delta_y^T \overset{i_1}{\mapsto} \sum_{y\in X^*} \delta_y \otimes \delta_y \overset{i_2^{-1}}{\mapsto} \sum_{y\in X^*} \delta_{(f_1-f_2)\otimes y} = \psi_{I_-},$$

where $I_- = \mathbb{F}_q(f_1 - f_2)$ is isotropic. Finally, any isotropic $I \subset U$ can be written this way, with a suitable choice of orthogonal basis $\{f_1, f_2\}$ and associated isomorphism i_2. □

2.6. **Quotient spaces and self-orthogonal Calderbank-Shor-Steane codes.** In this section, we will introduce the type of spaces that will turn out to contain all rank-deficient representations. In the field of quantum error correction, these spaces are called (tensor powers of) *self-orthogonal Calderbank-Shor-Steane (CSS) codes* [3, 24, 25].

Definition 2.5. Let $N \subset U$ be an isotropic space. The *self-orthogonal CSS code* associated with N is the space

$$\{\Phi \in L^2(U) \mid \operatorname{supp}\Phi \subset N^\perp, \Phi(u) = \Phi(u') \quad \forall\, u - u' \in N\}$$

of functions whose support is contained in $N^\perp$ and which are constant on cosets of N.

We will require an extension of this definition to functions on the tensor product space $U \otimes X^* \simeq \mathrm{Hom}(X \to U)$.

Definition 2.6. Let $N \subset U$ be an isotropic space. The *tensor power CSS code* associated with N is the subspace $C_N \subset L^2(\mathrm{Hom}(X \to U))$ of all functions Φ satisfying

$$\begin{cases} \Phi(F) = \Phi(F'), & \text{if } F - F' \in \mathrm{Hom}(X \to N), \\ \operatorname{supp}\Phi \subseteq \mathrm{Hom}(X \to N^\perp). \end{cases} \tag{2.25}$$

Using Lemma 2.1, one can see that the codes defined above are indeed tensor powers of the self-orthogonal CSS codes of Definition 2.5. We also note that projectors onto tensor powers of CSS codes have previously been identified in the commutant of the Clifford group [10, 20, 27].

Tensor power CSS codes carry a representation of $\mathrm{Sp}(V)$ that is associated with the orthogonal space $N^\perp/N$:

Lemma 2.7. *Let $N \subset U$ be an isotropic space.*

The quotient space $U' = N^\perp/N$ inherits an orthogonal form with dimension and discriminant given by, respectively

$$\dim U' = U - 2\dim N, \qquad d(U') = (-1)^{\dim N} d(U).$$

The stabilizer group $O_N \subset O(U)$ of N acts on U'. The maps that arise this way are exactly $O(U')$.

The restriction of $\mu_{U\otimes V}$ to $O_N \times \mathrm{Sp}(V)$ acts on C_N. As a representation of $O(U') \times \mathrm{Sp}(V)$, it is equivalent to $\mu_{U'\otimes V}$.

In view of this lemma, we will say that a tensor power CSS code C_N has *rank* r, if it carries a rank-r representation, or, equivalently, if $\dim N = (t-r)/2$.

Proof. Let $\{u_1, \ldots, u_k\}$ be a basis of N. There exist $u'_1, \ldots, u'_k$ such that

$$\beta(u_i, u'_j) = \delta_{i,j} \tag{2.26}$$

(because, for each j, Eq. (2.26) is an underdetermined system of linear equations for u'_j). Then $\mathbb{H}_i = \langle u_i, u'_i\rangle$ is a hyperbolic plane, and we arrive at an orthogonal decomposition

$$U = \mathbb{H}_1 \oplus \cdots \oplus \mathbb{H}_k \oplus U' =: H \oplus U', \tag{2.27}$$

where $U' = H^\perp$ is the orthogonal complement of the hyperbolic planes. Equation (2.27) implies: (1) The discriminant of U' is $d(U') = (-1)^k d(U)$, and (2) the orthogonal complement $N^\perp$ equals $N \oplus U'$, and we thus have $N^\perp/N \simeq U'$. Because the form on U' is inherited from the one of U, it is clear that O_N acts isometrically on U. Let $i : O_N \to O(U')$ be the homomorphism that maps elements of O_N to their action on $O(U')$. Then i is onto: If $g \in O(U')$, then, using the decomposition (2.27), we can embed g as $\mathrm{id} \oplus g$ into O_N.

By Corollary 2.2,

$$L^2(\mathrm{Hom}(X \to U)) \simeq L^2(\mathrm{Hom}(X \to H)) \otimes L^2(\mathrm{Hom}(X \to U')),$$
$$\mu_{U\otimes V} \simeq \mu_{H\otimes V} \otimes \mu_{U'\otimes V}.$$

By Lemma 2.4, $\mu_{H\otimes V}$ acts trivially on

$$\begin{aligned}\psi_{u_1} \otimes \cdots \otimes \psi_{u_k} &= \sum_{y_1,\ldots,y_k \in X^*} \delta_{u_1\otimes y_1} \otimes \cdots \otimes \delta_{u_k \otimes y_k} \\ &\simeq \sum_{F \in \mathrm{Hom}(X\to N)} \delta_F \quad \in L^2(\mathrm{Hom}(X \to H)).\end{aligned}$$

Thus

$$C_N \simeq \Big(\sum_{F\in \mathrm{Hom}(X\to N)} \delta_F\Big) \otimes L^2(\mathrm{Hom}(X \to U')),$$

on which $\mu_{U\otimes V}$ acts as $\mu_{U'\otimes V}$. □

In the remainder of this section, we introduce two concepts that will be used in Section 3 to reconstruct the codes a rank-deficient representation lives on.

A natural orthogonal basis on a tensor power CSS code is given by *coset states* (the generalization of Eq. (1.3)). Given an isotropic subspace $N \subset U$, an $F \in \mathrm{Hom}(X \to N^\perp)$, and a coset

$$[F] \in \mathrm{Hom}(X \to N^\perp)/\mathrm{Hom}(X \to N) \simeq \mathrm{Hom}(X \to N^\perp/N),$$

the associated *tensor power coset state* is

$$e_{[F]} = \sum_{G \in [F]} \delta_G \in C_N.$$

We will occassionally write $[F]_N$, if the vector space N is not unambiguously clear from context. The set

$$\{e_{[F]} \mid [F] \in \mathrm{Hom}(X \to N^\perp/N)\} \tag{2.28}$$

is an orthogonal basis for C_N. Note that if $[F] = [F']$, then $F = F' + \Delta$ for some $\Delta \in \mathrm{Hom}(X \to N)$ and thus

$$(F')^T F' = F^T F + F^T \Delta + \Delta^T F + \Delta^T \Delta = F^T F. \tag{2.29}$$

In particular, $e_{[F]}$ carries the $\mathcal{N}$-weight $B = F^T F$.

With each $F \in \mathrm{Hom}(X \to U)$, we associate the isotropic space

$$N_F = \mathrm{range}\, F \cap (\mathrm{range}\, F)^\perp, \tag{2.30}$$

which is the radical of the range of F.

Lemma 2.8. *Let $F \in \mathrm{Hom}(X \to U)$ be such that $\mathrm{rank}\, F^T F = r$. Then we have the dimension bound*

$$\dim N_F \leq \lfloor (t - r)/2 \rfloor. \tag{2.31}$$

Proof. We decompose U as $U_1 \oplus U_2 \oplus U_3$, where $U_1 = N_F$, U_2 is a complement to N_F in range F, and U_3 a complement to range F in U (c.f. Fig. 2). By construction, the space U_2 is non-degenerate and of dimenesion r, which implies that $U_2^\perp$ is $(t-r)$-dimensional and non-degenerate. Thus $N_F = U_1 \subset U_2^\perp$ is isotropic and contained in a $(t-r)$-dimensional non-degenerate space, which implies by Eq. (2.11) that $\dim N_F \leq (t-r)/2$. □

2.7. **Fourier transforms.** Central to the proof of our main result will be the fact that subrepresentations of the oscillator representation are closed under certain Fourier transforms. By a *Fourier transform*, we mean a map of the form $\mu(J_B)$ defined in Eq. (2.4), for $B : X^* \to X$ symmetric and invertible.

A standard result from harmonic analysis says that the support of a function is contained in a vector space if and only if its Fourier transform is supported on a (suitably defined) orthogonal complement. This statement can be generalized in a number of ways, the version we will require below reads:

Lemma 2.9. *Let $B : X^* \to X$ be symmetric and invertible, let $\Phi \in L^2(\mathrm{Hom}(X \to U))$, and let $U' \subset U$ be a subspace.*

Then the support of Φ is contained in the space $\mathrm{Hom}(X \to U')$ if and only if the support of the Fourier transform $\tilde{\Phi} := \mu_{U \otimes V}(J_B)\Phi$ is contained in $\mathrm{Hom}(X \to U'^\perp)$.

What is more, Φ is the indicator function on $\mathrm{Hom}(X \to U')$ if and only if $\tilde{\Phi}$ is the indicator function on $\mathrm{Hom}(X \to U'^\perp)$.

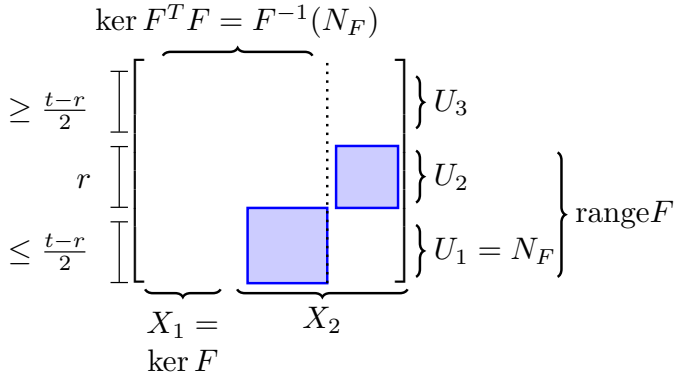

FIGURE 2. Illustration of the various subspaces we will associate with an $F \in \mathrm{Hom}(X \to U)$. In Lemma 3.1 and in the proof of the Main Theorem, the domain X will be decomposed as a direct sum of $X_1 = \ker F$ and some complement X_2. In Lemma 2.8 and in the proof of the Main Theorem, we decompose U as a direct sum of $U_1 = N_F = \mathrm{range}\, F \cap (\mathrm{range}\, F)^\perp$; U_2, some complement of U_1 within $\mathrm{range}\, F$; and U_3, some complement of $\mathrm{range}\, F$. These choices decompose $\mathrm{Hom}(X \to U)$ into six different subspaces $\mathrm{Hom}(X_i \to U_j)$, each of which can be visualized as a block in the matrix depicted. In Lemma 3.1, the map Δ lives in the lower left-hand side block, $\mathrm{Hom}(X_1 \to U_1 = N_F)$. In Lemma 3.2, we extend this to elements $\Delta = F - F'$ of the entire lower block $\mathrm{Hom}(X \to U_1)$, subject to a rank constraint. In the proof of the Main Theorem, G lives in the left block $\mathrm{Hom}(X_1 \to U)$. One could further subdivide X_2 into $X_2 \cap F^{-1}(N_F)$ (left side of the dotted line), and some complement (right side of the dotted line). We do not make use of this division in our argument. With respect to this choice, F is non-zero exactly on the two shaded blocks (where, in fact, it is invertible).

Since the proof follows the standard template for such results in harmonic analysis, we have deferred it to Appendix A.2.

Inspecting the generators in Sec. 2.2, it is clear that *only* Fourier transforms – i.e. generators from $\mathcal{J} \subset \mathrm{Sp}(V)$ – can possibly affect the rank of an element $\Phi \in L^2(\mathrm{Hom}(X \to U))$. This is the reason such maps figure prominently in our argument. By analyzing the action of Fourier transforms, one can easily derive further statements about the "rank spectrum" of representation spaces. Proposition 2.10 is one such example which to the best of our knowledge is not in the literature.

Proposition 2.10. *Let $(\mathcal{K}, \rho_{\mathcal{K}})$ be an irreducible $\mathrm{Sp}(V)$-subrepresentation of $\mu_{U \otimes V}$, where $\mathcal{K} \subset L^2(\mathrm{Hom}(X \to U))$. Let*

$$R = \{\mathrm{rank}\, B \mid B \text{ is a weight that appears in } \rho_{\mathcal{K}}|_{\mathcal{N}}\}$$

be the set of values the rank takes on the $\mathcal{N}$-spectrum of the representation. Then R is a contiguous range of integers.

As the rest of the arugment will not rely on Proposition 2.10, its proof is given in Appendix A.3.

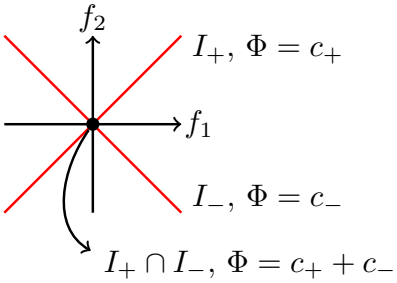

FIGURE 3. "Branch and stem" structure of rank-0 subrepresentations of $\mu_{\mathbb{H}\otimes V}$ associated with the hyerpbolic plane. The vectors f_1, f_2 denote an orthogonal basis of $\mathbb{H}$. The red lines $I_\pm$ are the two isotropic spaces. A rank-deficient subrepresentation (Eq. (3.2)) of $\mu_{\mathbb{H},V}$ takes values that are constant on the "branches" $\{F \mid N_F = I_\pm\}$, while the values add up on the "stem" $\{0\}$ where the spaces intersect.

3. The classification of rank-deficient subrepresentations

3.1. **Informal outline of the main proof.** Let $\mathcal{K} \subset L^2(\mathrm{Hom}(X \to U))$ be a representation space of rank $r < t$. We aim to show that there is some $\Phi \in \mathcal{K}$ that can be written as a linear combination

$$\Phi = \sum_{N \text{ isotropic}} \Phi_N, \tag{3.1}$$

of components Φ_N in suitable tensor power CSS code spaces C_N. This, together with Lemma 2.7, will imply the Main Theorem.

One of the defining properties of elements Φ_N of C_N is that they are constant on cosets of $\mathrm{Hom}(X \to N)$. It is not obvious how one can derive such *invariance properties* from *rank deficiency*.

To achieve this, we rely on the fact that $\mathcal{K}$ is closed under certain Fourier transforms. More precisely, if we decompose X as a direct sum $X_1 \oplus X_2$, then any $F : X \to U$ can be written as the sum of two blocks $F = F_1 + F_2$ with $F_i : X_i \to U$ (Fig. 3). Now fix some F_2 and consider the dependency $\phi : F_1 \mapsto \Phi(F_1 + F_2)$ of Φ on the first block alone. It turns out that rank deficiency imposes linear constraints on the maps F_1 that can appear in the support of ϕ. But, as we have recalled in Sec. 2.7, if the support of a function is contained in a linear subspace, then its Fourier transform is invariant under translations along the orthogonal complement. Closure of $\mathcal{K}$ under Fourier transforms then implies invariances of the type that occur in CSS codes for any $\Phi \in \mathcal{K}$. This first step of recovering a CSS code structure is made precise in Lemma 3.1.

The next challenge we are facing is that Φ is a linear combination of elements from *different* codes, so that there is no single space N under which Φ is invariant. Indeed, the symmetries found in the first step are only "local" in that they depend on the fixed block F_2. To get some feeling for what we can expect, we look at the simplest non-trivial example: $\mu_{\mathbb{H}\otimes V}$ with $\mathbb{H}$ the hyperbolic plane.

The plane has a orthogonal basis $\{f_1, f_2\}$, with

$$\beta(f_1, f_1) = 1, \qquad \beta(f_2, f_2) = d(\mathbb{H}) = -1.$$

There are two isotropic spaces, $I_\pm = \mathbb{F}_q\,(f_1 \pm f_2)$ (c.f. Figure 3). It follows that there are two tensor power CSS codes $C_{I_\pm}$ in $\mathrm{Hom}(X \to \mathbb{H})$. They are one-dimensional, proportional to the vectors $\psi_{I_\pm}$ defined in Lemma 2.4. Thus, for $c_\pm \in \mathbb{C}$, the vector

$$\Phi = c_+\psi_{I_+} + c_-\psi_{I_-} \in L^2(\mathrm{Hom}(X \to \mathbb{H})) \tag{3.2}$$

carries a rank-0 representation (and we will see that these are the only rank-deficient subrepresentation of $\mu_{\mathbb{H}\otimes V}$). Using Lemma 2.4

$$\Phi(F) = \begin{cases} c_+ & \operatorname{rank} F^T F = 0, N_F = I_+ \\ c_- & \operatorname{rank} F^T F = 0, N_F = I_- \\ c_+ + c_- & \operatorname{rank} F^T F = 0, N_F = \{0\} \\ 0 & \operatorname{rank} F^T F = 1 \end{cases},$$

a situation sketched in Fig. 3. Embracing a horticultural analogy, Φ is constant on the two "branches" $\{F \mid N_F = I_\pm\}$, while the values add up on the "stem" $\{0\}$, where the spaces intersect.

This structure generalizes to higher-dimensional orthogonal spaces U. Define the "generalized branches" to be

$$B_N := \{F \in \mathrm{Hom}(X \to U) \mid \operatorname{rank} F^T F = r, N_F = N\}.$$

Then Lemma 3.2 states that on each B_N, a vector Φ in a rank-deficient representation exhibits the invariance under $\mathrm{Hom}(X \to N)$ that is characteristic of elements of the code C_N. More precisely:

$$F, F' \in B_N, \quad (F - F') \in \mathrm{Hom}(X \to N) \qquad \Rightarrow \qquad \Phi(F) = \Phi(F').$$

Thus Φ is well-defined on sets $B_N / \mathrm{Hom}(X \to N)$.

After this, we "prune off the branches" by setting

$$\Phi' := \Phi - \sum_{\substack{N \text{ isotropic} \\ \dim N = \lfloor (t-r)/2 \rfloor}} \sum_{[F] \in B_N / \mathrm{Hom}(X \to N)} \Phi([F])\, e_{[F]}.$$

The right-hand summand involves the coset states $e_{[F]}$, which are elements of the respective code C_N. The support of the remainder Φ' is thus contained in the "stem". We conclude the argument by showing that representations with rank $< t$ do not contain non-zero vectors supported on such a stem, so in fact $\Phi' = 0$.

This final step again relies on Fourier transforms. Roughly, the "stem" is a "small" space, so that by the uncertainty principle, Fourier transforms will have "large" support – so large, in fact, that they are guaranteed to contain higher-rank elements.

3.2. **Proof of the Main theorem.**

Lemma 3.1. *Let $\mathcal{K} \subset L^2(\mathrm{Hom}(X \to U))$ be a subrepresentation of rank $r < t$. Let $\Phi \in \mathcal{K}$ and $F \in \operatorname{supp}\Phi$ such that $\operatorname{rank} F^T F = r$, and let N_F be as in Eq. (2.30). If $\Delta \in \mathrm{Hom}(X \to N_F)$ is such that*

$$\operatorname{range} F_{|\ker \Delta} = \operatorname{range} F, \tag{3.3}$$

then

$$\Phi(F) = \Phi(F + \Delta).$$

We note that in the case $r = t$ one has $N_F = \{0\}$ and so the lemma trivially holds.

Proof. Set $X_1 = \ker F$. The assumption (3.3) implies that there is a complement X_2 of X_1 contained in $\ker \Delta$. This choice induces a decomposition $\mathrm{Hom}(X \to U) = \mathrm{Hom}(X_1 \to U) \oplus \mathrm{Hom}(X_2 \to U)$ with $\Delta \in \mathrm{Hom}(X_1 \to U)$, $F \in \mathrm{Hom}(X_2 \to U)$.

Let $B : X_1^* \to X_1$ be invertible, let $\mu_{U\otimes V_1}(J_B)$ be the associated Fourier transform, and let i be the isomorphism (2.8). By Section 2.7 and Lemma 2.1, the vector

$$\tilde{\Phi} := \left(i^{-1} \left(\mu_{U\otimes V_1}(J_B) \otimes \mu_{U\otimes V_2}(\mathbb{1})\right) i \right) \Phi$$

is an element of $\mathcal{K}$. Thus, by the assumption on the rank of the representation, $\tilde{\Phi}$ has support only on maps $F' \in \mathrm{Hom}(X \to U)$ with $\mathrm{rank}(F')^T F' \leq r$.

If $F' = G + F$ for some $G \in \mathrm{Hom}(X_1 \to U)$, then

$$\mathrm{range}(G + F) = \langle \mathrm{range}\, G \cup \mathrm{range}\, F \rangle.$$

The condition $\mathrm{rank}(G+F)^T(G+F) \leq r$ is equivalent to demanding that $\mathrm{range}(G+F)$ has rank at most r as an orthogonal space. This implies

$$\mathrm{range}\, G \subset \langle \mathrm{range}\, F \cup (\mathrm{range}\, F)^\perp \rangle = N_F^\perp. \tag{3.4}$$

Set

$$\phi \in L^2(X_1 \to U), \qquad \phi(G) = \Phi(G + F), \qquad \tilde{\phi} = \mu_{U\otimes V_1}(J_B)\phi.$$

Then

$$\tilde{\Phi}(G + F) = \tilde{\phi}(G),$$

so that the preceding discussion implies that $\mathrm{supp}\, \tilde{\phi} \subset \mathrm{Hom}(X \to N_F^\perp)$. Thus Lemma 2.9 implies that ϕ is constant on cosets of $\mathrm{Hom}(X_1 \to N_F)$, a space which includes Δ. □

Lemma 3.2 extends the invariances – essentially by using the fact that there is a some freedom in choosing the complement X_2 to $\ker F$ that appears in the proof above.

Lemma 3.2. *Let $\mathcal{K}$ be a representation of rank $r < t$, and $\Phi \in \mathcal{K}$. Let $N \subset U$ be an isotropic space of dimension $\dim N \leq (t - r)/2$, and set*

$$B_N := \{F \in \mathrm{Hom}(X \to U) \mid \mathrm{rank}\, F^T F = r, N_F = N\}.$$

Then B_N is non-empty and, on B_N, Φ is invariant under $\mathrm{Hom}(X \to N)$:

$$\Phi(F) = \Phi(F') \qquad \forall\, F, F' \in B_N, (F - F') \in \mathrm{Hom}(X \to N). \tag{3.5}$$

The proof uses the *probabilistic method* [1]: The strategy is to ascertain the (deterministic) exsistence of an object by showing that a randomized construction yields one with positive probability. Presumably an explicit construction would offer us more insight into the structure of the problem. We leave such a *derandomization* for future work.

Proof. Because $t - 2\dim N \geq r$, there exists an r-dimensional non-degenerate subspace M of $N^\perp$. Then any $F \in \mathrm{Hom}(X \to U)$ satisfying $\mathrm{range}\, F = \langle M, N \rangle$ will be an element of B_N. The existence of such an F is guaranteed by $n \geq t$. Hence B_N is not empty.

Now let F, F' be as in Eq. (3.5). The aim is to show that there exists a "midpoint" G such that both F with $\Delta = (G - F)$, as well as F' with $\Delta' = (F' - G)$ fulfill the assumptions of Lemma 3.1. It then follows that $\Phi(F) = \Phi(G) = \Phi(F')$.

We claim that if Δ is chosen uniformly at random from $\mathrm{Hom}(X \to N)$, then, with probability strictly larger than $1 - \frac{1}{q-1}$, it holds that $\operatorname{range} F_{|\ker \Delta} = \operatorname{range} F$, i.e. Lemma 3.1 applies to F, Δ.

Before turning to the analysis of the randomized procedure, we state two preparatory facts. First, for each subspace $Z \subset X$, it holds that

$$\operatorname{range} F_{|Z} = \operatorname{range} F \qquad \Leftrightarrow \qquad \dim Z - \dim(Z \cap \ker F) = \operatorname{rank} F. \tag{3.6}$$

Second, for each $\Delta \in \mathrm{Hom}(X \to N)$, Lemma 2.8 gives the dimension bound

$$\dim \ker \Delta \geq n - \dim N \geq t - \dim N \geq t - (t-r)/2 = r + (t-r)/2 \geq \operatorname{rank} F. \tag{3.7}$$

Now assume Δ is distributed uniformly at random. From the previous equation, any $\operatorname{rank} F$-dimensional subspace Z will occur within $\ker \Delta$ with equal probability. By Eq. (3.6), if $\dim(Z \cap \ker F) = 0$ for some such Z, then the assumption of Lemma 3.1 is met.

There are $(q^k - 1)/(q-1)$ one-dimensional spaces in a k-dimensional vector space. Thus, the probability that any fixed one-dimesional subspace is contained in a randomly chosen z-dimensional one is $(q^z - 1)/(q^n - 1)$. By the union bound, the probability that at least one non-zero element of a fixed $(n-z)$-dimensional space is contained in a z-dimensional random one is therefore upper-bounded by

$$\frac{q^{n-z}-1}{q-1}\frac{q^z-1}{q^n-1} = \frac{1}{q-1}\frac{(q^n - q^z - q^{n-z} + 1)}{q^n - 1} < \frac{1}{q-1} \quad (\forall\, z \leq t).$$

This establishes the claim made at the beginning of the proof.

Now set $G = F + \Delta$. The distribution of $\Delta' = F' - G = (F' - F) - \Delta$ is the same as the distribution of Δ. Thus Lemma 3.1 applies to F', Δ' with the same probability.

We conlcude by the union bound that the probability of the construction working in both cases simultaneously is strictly larger than $1 - \frac{2}{q-1} \geq 0$. □

Proof of the Main Theorem. Let $\Phi \in \mathcal{K}$ carry an $\mathcal{N}$-weight B of rank r. By Lemma 3.1 and Lemma 3.2, Φ is well-defined on cosets $B_N / \mathrm{Hom}(X \to N)$ for a suitable N. Set

$$\Phi' := \Phi - \sum_{\substack{N \text{ isotropic} \\ \dim N = \lfloor (t-r)/2 \rfloor}} \sum_{[F] \in B_N / \mathrm{Hom}(X \to N)} \Phi([F])\, e_{[F]}. \tag{3.8}$$

We will prove that Φ' is actually equal to zero, using a Fourier-transform argument as in Lemma 3.1.

For the sake of reaching a contradiction, assume that $\Phi' \neq 0$ and choose an $F \in \operatorname{supp} \Phi'$ such that

$$\operatorname{rank} F = \max_{F' \in \operatorname{supp} \Phi'} \operatorname{rank} F'. \tag{3.9}$$

For such an F, we next show that $\dim N_F < \lfloor (t-r)/2 \rfloor$. For this, we use that

$$F \in \operatorname{supp} \Phi' \implies F \notin \bigcup_{\substack{N \text{ isotrop. s.t.} \\ \dim N = \lfloor (t-r)/2 \rfloor}} B_N,$$

and so $\dim N_F \neq \lfloor (t-r)/2 \rfloor$. But since $F \in \operatorname{supp} \Phi$, and $\operatorname{rank} \mathcal{K} = r$, we also know that $N_F \leq \lfloor (t-r)/2 \rfloor$ and the claim follows.

Now, as in the proof of Lemma 3.1, set $X_1 = \ker F$, choose some complement X_2 to X_1, an invertible symmetric $B : X_1^* \to X_1$, set $V_i = X_i \oplus X_i^*$, and define $\phi', \tilde{\phi}' \in L^2(X_1 \to U)$ as

$$\phi'(G) := \Phi'(F+G), \qquad \tilde{\phi}' := \mu_{U\otimes V_1}(J_B)\phi'.$$

Then, with $\tilde{\Phi}' := (\mu_{U\otimes V_1}(J_B) \otimes \mu_{U\otimes V_2}(\mathbb{1}))\Phi' \in \mathcal{K}$, it holds that

$$\tilde{\Phi}'(F+G) = \tilde{\phi}'(G) \quad \forall G \in \mathrm{Hom}(X_1 \to U). \tag{3.10}$$

We decompose U as $U_1 \oplus U_2 \oplus U_3$, where $U_1 = N_F$, U_2 is a complement to N_F in $\operatorname{range} F$, and U_3 a complement to $\operatorname{range} F$ in U (c.f. Fig. 2). Let G be an element of $\operatorname{supp}\phi'$. Write $G = G_1 \oplus G_2 \oplus G_3$ for maps $G_i \in \mathrm{Hom}(X_1 \to U_i)$. Because β restricted to U_2 is non-degenerate, it follows from $(G+F)^T(G+F) = B$ that $G_2 = 0$. By Eq. (3.9), $G_3 = 0$. From Lemma 3.1, ϕ' is invariant under $\mathrm{Hom}(X_1 \to N_F) = G_1$. Thus ϕ' is proportional to the indicator function on $\mathrm{Hom}(X_1 \to N_F)$. Hence $\tilde{\phi}'$ is proportional to the indicator function on $\mathrm{Hom}(X_1 \to N_F^\perp)$. But $N_F^\perp$ contains the r-dimensional non-degenerate space U_2 and has dimension $\dim N_F^\perp > t - \lfloor (t-r)/2 \rfloor \geq r + (t-r)/2$. Therefore, as an orthogonal space, $N_F^\perp$ has rank strictly larger than r and hence contains a non-isotropic vector $u \notin \operatorname{range} F$. If $G \in \mathrm{Hom}(X_1 \to N_F^\perp)$ has u in its range, then $\operatorname{rank}(G+F)^T(G+F) \geq r+1$. But by Eq. (3.10), $G+F$ appears in the support of $\tilde{\Phi}'$, contradicting the assumption that $\mathcal{K}$ has rank r.

It follows that Φ is in the span of the rank-r tensor power CSS codes $\mathcal{C}_r$. If $\mathcal{K}$ is irreducible, then it is spanned by the orbit $\mathrm{Sp}(V)\cdot\Phi$. But Lemma 2.7 says that $\mathcal{C}_r$ is invariant under the $\mathrm{Sp}(V)$ action, so $\mathcal{K} \subseteq \mathcal{C}_r$. By the same lemma, $t-r$ is even. The Main Theorem therefore holds for irreps, and hence for all representations. □

3.3. **The connection to the η correspondence.** Here, we will combine the respective main results of this work and of [11] to arrive at a complete decomposition of $L^2(\mathrm{Hom}(X \to U))$ in terms of irreducible $\mathrm{Sp}(V)$ subrepresentations.

One can find $\mathrm{Sp}(V)$-subrepresentations of $\mu_{U\otimes V}$ in the following way. First, choose an isotropic subspace $N \subset U$ and a $\tau \in \mathrm{Irr}(O(N^\perp/N))$. Then, use Lemma 2.7 to find that the code C_N is isomorphic to $\mu_{(N^\perp/N)\otimes V}$ as an $\mathrm{Sp}(V)$ representation. Finally, invoke the η correspondence on this code to find an irreducible representation $\eta(\tau) \subset C_N$ of rank $t - 2\dim N$.

Lemma 3.3 observes that, while in general different CSS codes may have non-trivial intersections, the representation spaces arising in the way just described are linearly independent. This allows us to identify the joint action of $U(O)$ and $\mathrm{Sp}(V)$ on their span as a certain induced representation.

Recall that by Lemma 2.7, there is a homomorphism $i : O_N \to O(N^\perp/N)$ from the stabilizer group $O_N \subset O(U)$ of an isotropic subspace onto the orthogonal group of $N^\perp/N$. Thus, if $\tau \in \operatorname{Irr} O(N^\perp/N)$, then $\tau \circ i$ represents O_N. In this section, we will implicitly make this identification and we will not distinguish notationally between τ and $\tau \circ i$.

Lemma 3.3. *Let $N \subset U$ be an isotropic space and let $\tau \in \mathrm{Irr}(N^\perp/N)$.*

Let $\mathcal{K} \subset L^2(\mathrm{Hom}(X \to U))$ be the subspace on which $\mathrm{Sp}(V)$ acts as $\eta(\tau)$. Then, as an $O(U) \times \mathrm{Sp}(V)$-representation,

$$\mathcal{K} \simeq \mathrm{Ind}_{O_N}^{O(U)}(\tau) \otimes \eta(\tau). \tag{3.11}$$

Proof. Set $U' = N^\perp/N$. By Lemma 2.7 and Theorem 1.1, there is a unique $O(U') \times \mathrm{Sp}(V)$-representation space $\mathcal{K}_0$ of type $\tau \otimes \eta(\tau)$ in C_N.

The *isotropic Grassmanian*

$$I_r = \{N \mid N \text{ isotropic}, \dim N = (t-r)/2\} \simeq O(U)/O_N$$

can be identified with the cosets $O(U)/O_N$. Let $\{g_i\}_{i=1}^{|I_r|}$ be a choice of representatives for each coset. Define

$$\mathcal{K}_{[g_i]} = \mu_{U\otimes V}(g_i)\left(\mathcal{K}_0\right).$$

As $\mathrm{Sp}(V)$-representation spaces, the $\mathcal{K}_{[g_i]}$ are all equivalent to $\eta(\tau)$. Conversely, from Theorem 1.2, every $\mathrm{Sp}(V)$-representation of type $\eta(\tau)$ is contained in their span. Therefore,

$$\mathcal{K} = \mathrm{span}\{\mathcal{K}_{[g_i]}\}_{i=1}^{|I_r|}.$$

We claim that the spaces $\mathcal{K}_{[g_i]}$ are linearly independent.

Indeed: We need to show that for each i, the space $\mathcal{K}_{[g_i]}$ intersects the span $\mathcal{K}'$ of the other spaces only at $\{0\}$. Since $O(U)$ acts transitively on the $\mathcal{K}_{[g_i]}$, it is enough to treat the case $i = 1$. As $\mathcal{K}'$ and $\mathcal{K}_{[g_1]}$ are $O_N \times \mathrm{Sp}(V)$ representation spaces, and because $\mathcal{K}_{[g_1]}$ is irreducible, we have the alternatives

$$\mathcal{K}_{[g_1]} \subset \mathcal{K}' \qquad \text{or} \qquad \mathcal{K}_{[g_1]} \cap \mathcal{K}' = \{0\}.$$

It thus suffices to show that $\mathcal{K}_{[g_1]}$ contains one vector that is not an element of $\mathcal{K}'$. Let $\Phi_1 \in \mathcal{K}_{[g_1]}$ carry an $\mathcal{N}$-weight B of rank r, let $F \in \mathrm{supp}\,\Phi_1$. There is some $F' \in [F]_N$ that is *maximal* in the sense $\mathrm{range}\,F' = N^\perp$ (rather than its range being a strict subset of $N^\perp$). Since $\mathrm{rank}\,F^T F = r$, there must be some complement W of N in $N^\perp$ for which $W \subseteq \mathrm{range}\,F$. From the decomposition $\mathrm{Hom}(X \to N^\perp) = \mathrm{Hom}(X \to N) \oplus \mathrm{Hom}(X \to W)$ it is clear that there exists a $\Delta \in \mathrm{Hom}(X \to N)$ for which $F + \Delta = F'$ is maximal. By the invariance property of CSS codes, $F' \in \mathrm{supp}\,\Phi_1$, i.e. the inner product $(\delta_{F'}, \Phi_1) \neq 0$. In contrast, let $\Phi_i \in \mathcal{K}_{[g_1]}$ for $i \neq 1$. Then $F_i \in \mathrm{supp}\,\Phi_i \Rightarrow \mathrm{range}\,F_i \subset N_i^\perp$. But $\mathrm{range}\,F' = N_1^\perp \not\subset N_i^\perp$, so that $(\delta_{F'}, \Phi_i) = 0$. It follows that $\Phi_1 \notin \mathcal{K}'$, as claimed.

The space $\mathcal{K}$ is therefore a direct sum of the $\mathcal{K}_{[g_i]}$. We will now compute the action of $O(U)$ on this direct sum. It suffices to consider vectors of the form

$$\mu_{U\otimes V}(g_i)(\phi \otimes \psi),$$

which span $\mathcal{K}$. For each $g \in O(U)$, there is a permutation $\pi \in S_{|I_r|}$ and elements $h_i \in O_N$ such that for each $gg_i = g_{\pi_i} h_i$. Thus

(3.12)

$$\mu_{U\otimes V}(g)\left(\mu_{U\otimes V}(g_i)(\phi\otimes\psi)\right) = \mu_{U\otimes V}(g_{\pi_i}h_i)(\phi\otimes\psi) = \mu_{U\otimes V}(g_{\pi_i})\left(\tau(h_i)\phi\otimes\psi\right).$$

But this is the action of the advertised induced representation. □

By Lemma 2.7, the space $N^\perp/N$ is an orthogonal space of dimension $r = t - 2k$ and discriminant $d(U_r) = (-1)^k d(U)$, with $k = \dim N$. In particular, up to orthogonal maps, $N^\perp/N$ only depends on r. With this in mind, we suppress the dependency on N in our notation, and define U_r to be $N^\perp/N$ for *some* isotropic N of dimension $(t-r)/2$. In the same vein, any two isotropic spaces of the same dimension have conjugate stabilizer groups, and thus the isomorphism class of the induced representation in Eq. (3.11) does not depend on N. Again, this justifies defining O_r to be O_N for *some* isotropic N of dimension $(t-r)/2$.

Theorem 1.1, Theorem 1.2, and Lemma 3.3 then yield the decomposition

$$\mu_{U\otimes V} \simeq \bigoplus_{r\in R(U)} \bigoplus_{\tau\in \mathrm{Irr}\, O(U_r)} \mathrm{Ind}_{O_r}^{O(U)}(\tau)\otimes\eta(\tau), \tag{3.13}$$

with

$$R(U) = \{t-2k \mid \text{ there is an isotropic } N\subset U \text{ with } \dim N = k\}.$$

All $\mathrm{Sp}(V)$-irreps $\eta(\tau)$ appearing in Eq. (3.13) are indeed inequivalent: Those corresponding to different $O(U_r)$ are distinguished by their rank, whereas the inequivalence of summands of the same rank is a consequence of Theorem 1.1.

As an $O(U_r)$-representation,

$$\mathrm{Ind}_{O_r}^{O(U)}(\tau) \simeq \tau\otimes\mathbb{C}^{|O(U)/O_N|}$$

is just τ with degeneracy equal to the number of isotropic subspaces of dimension k.

A comparison with Theorem 1.1 shows that the $\mathrm{Sp}(V)$-representations in $\Theta(\tau)$ are exactly those $\eta(\tau')$, where τ appears in $\mathrm{Ind}_{O_N}^{O(U)}(\tau')$. In terms of character inner products, and using Frobenius reciprocity:

$$\langle\Theta(\tau),\eta(\tau')\rangle_{\mathrm{Sp}(V)} = \sum_{r\in R(U)} \langle\tau,\mathrm{Ind}_{O_r}^{O(U)}(\tau')\rangle_{O(U)} = \sum_{r\in R(U)} \langle\mathrm{Res}_{O_r}^{O(U)}(\tau),\tau'\rangle_{O_r}.$$

As an example, we consider the case where $\tau = \mathrm{id}_{O(U)}$ is the trivial representation of $O(U)$. Then

$$\langle\mathrm{Res}_{O_r}^{O(U)}(\mathrm{id}_{O(U)}),\tau'\rangle_{O_r} = \langle\mathrm{id}_{O_r},\tau'\rangle_{O_r} = \delta_{\mathrm{id}_{O(U_r)},\tau'}.$$

Therefore,

$$\Theta(\mathrm{id}_U) = \bigoplus_{r\in R(U)} \eta(\mathrm{id}_{O(U_r)})$$

has a number of components equal to the isotropy index of U.

3.4. **A non-CSS type rank-deficient subrepresentation.** Our main theorem makes statements only in the regime $t \le n$. Here, we show that it indeed cannot be extended to all pairs t, n. To this end, we construct a rank-0 subrepresentation of $\mu_{\mathbb{F}_p^3\otimes\mathbb{F}_p}$, i.e. for the case of $t = 3$ and $n = 1$. Here, p is an arbitrary odd prime. This is incompatible with Theorem 1.2, which posits that $t-r$ be even. Thus, more general subrepresentations can occur for $t > n$.

Set $V = \mathbb{F}_p\oplus\mathbb{F}_p^*$ and $U = \mathbb{F}_p^3$ with the standard orthogonal form β. The oscillator representation $\mu_{U\otimes V}$ thus acts on $L^2(U\otimes\mathbb{F}_p^*) \simeq L^2(U)$.

Our construction depends[1] on the choice of an isotropic vector $x_0 \in U$. Define $\psi \in L^2(U)$ by

$$\psi(z) = \begin{cases} 0 & \beta(z,z) \neq 0 \text{ or } z = 0 \\ \ell_{\beta(x_0,z)} & \beta(z,z) = 0, z \neq \lambda x_0 \\ \ell_{2\lambda} & z = \lambda x_0. \end{cases} \tag{3.14}$$

In particular, ψ is supported on the set of isotropic vectors in U, and restricts to a Legendre symbol on every ray.

Proposition 3.4. *The representation $\mu_{U\otimes V}$ of $\mathrm{Sp}(V)$ acts trivially on ψ.*

Proof. From the explicit definitions in Section 2.2, one can easily see that ψ affords trivial actions by the subgroups $\mathcal{N}$ (using isotropy of the support) and $\mathcal{D}$ (using the multiplicativity of the Legendre symbol). All elements J_B of the subgroup $\mathcal{J}$ can be written as a product of an element from $\mathcal{D}$ with J_{id}, where $\mathrm{id} : X^* \to X$ is the canonical identification of $\mathbb{F}_p^*$ with $\mathbb{F}_p$. It therefore remains to be shown that ψ is stabilized by $\mu_{U\otimes V}(J_{\mathrm{id}})$.

We begin by deriving a more convenient expression for ψ. The standard form in $\mathbb{F}_p^3$ is isomorphic to $\mathbb{H} \oplus \langle -1 \rangle$. In other words, there exists a basis with respect to which the standard form on $\mathbb{F}_p^3$ is

$$\beta(x,y) = x_1y_2 + x_2y_1 - x_3y_3.$$

In this basis, define

$$\begin{aligned} x_a &= (1, 2^{-1}a^2, a), \qquad a \in \mathbb{F}_p \\ x_\infty &= (0,2,0). \end{aligned}$$

By enumerating all points in projective space $\mathbb{F}_p^3/\mathbb{F}_p$, one may easily convince oneself that every isotropic vector in $\mathbb{F}_p^3$ is a multiple of exactly one x_a, for $a \in \bar{\mathbb{F}}_p := (\mathbb{F}_p \cup \infty)$. We can choose the basis change such that the vector x_0 that appears in (3.14) is mapped to the vector x_0 as defined here.

For $a \neq b \in \mathbb{F}_p$,

$$\begin{aligned} \beta(x_a, x_b) &= 2^{-1}(a^2 + b^2) - ab = 2^{-1}(a-b)^2, \\ \beta(x_a, x_\infty) &= 2, \end{aligned}$$

so that the Legendre symbol of the inner products is constant:

$$\ell_{\beta(x_a,x_b)} = \ell_2 \qquad \forall a \neq b \in \bar{\mathbb{F}}_p^3.$$

With these definitions, ψ takes a simple form:

$$\begin{aligned} \psi &= \sum_{a\in\bar{\mathbb{F}}_p} \sum_{\lambda\in\mathbb{F}_p^\times} \left(\ell_{\beta(x_0,x_a)} + \ell_2\delta_{a,0}\right) \ell_\lambda e_{\lambda x_a} \\ &= \ell_2 \sum_{a\in\bar{\mathbb{F}}_p} \sum_{\lambda\in\mathbb{F}_p^\times} \ell_\lambda e_{\lambda x_a}. \end{aligned}$$

[1] Numerically, it appears that the resulting representation space is actually independent of the choice of x_0. Numerical investigations also indicate that when substituting $U = \mathbb{F}_p^3$ (which has discriminant $d(U) = 1$) by a three-dimensional U' with discriminant $d(U')$ a non-square, then $\mu_{U'\otimes V}$ will still act trivially on ψ if $p = 3$. On the other hand, for $p = 5, 7, 11, 13$, it holds that $\mu_{U'\otimes V}$ does not afford *any* trivial representation space. We will neither use, nor attempt to prove, these statements.

We evaluate the Fourier transform $\tilde{\psi} = \mu_{U\otimes V}(J_{\mathrm{id}})\psi$ on an isotropic vector, using Eq. (2.4): For $\kappa \in \mathbb{F}_p^\times, b \in \bar{\mathbb{F}}_p$, it holds that

$$\begin{aligned}\tilde{\psi}(\kappa x_b) &= \gamma^{-3}\ell_2 \sum_{a\in\bar{\mathbb{F}}_p}\sum_{\lambda\in\mathbb{F}_p^\times} \ell_\lambda \omega(\beta(\lambda x_a, \kappa x_b)) \\ &= \gamma^{-3}\ell_2 \sum_{a\in\bar{\mathbb{F}}_p, a\neq b} \ell_\kappa \ell_{\beta(x_a,x_b)} \sum_{\lambda\in\mathbb{F}_p^\times} \ell_\lambda \omega(\lambda) \\ &= \gamma^{-3} p \ell_\kappa \sum_{\lambda\in\mathbb{F}_p^\times} \ell_\lambda \omega(\lambda) \\ &= \gamma^{-2} p \ell_\kappa = \ell_2 \ell_\kappa = \psi(\kappa x_b),\end{aligned}$$

where we have used the standard properties of quadratic Gauss sums. Restricted to the support of ψ, this is the required eigenvalue equation.

In particular, we have found that $\tilde{\psi}$ coincides with ψ on the support of ψ. Because the oscillator representation acts isometrically, $\tilde{\psi}$ must thus also have the same support as ψ. □

4. The connection to the Clifford group

The motivation for this work was to understand the appearance of projections onto CSS codes in the commutant of tensor power representations of the *Clifford group* [10]. While we have opted to state our main results for representations of the symplectic group, the two cases can sometimes be precisely linked. This is the purpose of Proposition 4.2, which will be developed in this section.

We start by recalling the basic definitions. In addition to the oscillator representation, the Hilbert space $L^2(X^*)$ also carries a representation $W^{(m)}$ of the *Heisenberg group* $H(V)$ over $V = X \oplus X^*$. The Heisenberg group $H(V)$ is the set $\mathbb{F}_q \times V$ with group law

$$(\lambda, v) \circ (\lambda', v') = (\lambda + \lambda' + 2^{-1}[v, v'], v + v').$$

For $m \in \mathbb{F}_q^\times$, the *Weyl representation of mass* m on $L^2(X^*)$ is

$$W_V^{(m)}(\lambda, x \oplus y)\delta_z = \omega^{(m)}(-2^{-1}y(x) + z(x) + \lambda)\,\delta_{z+y}. \tag{4.1}$$

As is true for the oscillator representation (Sec. 2.2), we again have that $W_V^{(-m)}$ is the complex conjugate of $W_V^{(m)}$, and again we will omit the superscript for the mass-1 version. Two Weyl representations of different mass are inequivalent [8]. The Weyl and the oscillator representations are compatible in that

$$\mu_V^{(m)}(S)\, W_V^{(m)}(\lambda, v)\, \mu_V^{(m)}(S)^{-1} = W_V^{(m)}(\lambda, Sv) \tag{4.2}$$

for all $S \in \mathrm{Sp}(V), v \in V$. The semi-direct product $H(V) \rtimes \mathrm{Sp}(V)$ with automorphism

$$S\,(\lambda, v)\,S^{-1} = (\lambda, Sv)$$

is the *Jacobi group* $J(V)$. By Eq. (4.2), the map

$$\mathcal{C}l_V^{(m)} : (\lambda, v, S) \mapsto W_V^{(m)}(\lambda, v)\mu_V^{(m)}(S),$$

thus defines a representation of the Jacobi group on $L^2(X^*)$. The operators realizing this representation are known in quantum information theory as the *Clifford group* ("Clifford *representation* of the Jacobi group" would be a more consistent term,

but we have chosen here to keep with the standard use in physics). Because the maps $\{W_V(v)\}_{v\in V}$ form a basis in $L^2(X^*)$, Eq. (4.2) determines $\mu(S)$ up to a phase factor.

As $\mathrm{Sp}(V)$ embeds into $\mathrm{Sp}(U\otimes V)$ (Sec. 2.4), so too can one embed the Heisenberg group $H(V)$ into $H(U\otimes V)$. However, the embedding we will use is no longer canonical, but depends on the choice of a vector $u\in U$. Roughly, we use u to lift $y\in X^*$ to $u\otimes y\in U\otimes X^*$. Given u, define

$$\iota_u : (\lambda, v) \mapsto (\beta(u,u)\lambda, u\otimes v).$$

This is a homomorphism:

$$\begin{aligned}
&\iota_u\big((\lambda, x\oplus y)\circ(\lambda', x'\oplus y')\big)\\
=&\Big(\beta(u,u)(\lambda+\lambda')2^{-1}\beta(u,u)\big(y'(x)-y(x')\big), u\otimes\big((x+x')\oplus(y+y')\big)\Big)\\
=&\iota_u(\lambda, x\oplus y)\circ\iota_u(\lambda', x'\oplus y'),
\end{aligned}$$

from which one verifies that we have a representation

$$\mathcal{C}l^{(u)}_{U\otimes V} : (\lambda, v, S) \mapsto W_{U\otimes V}(\iota_u(\lambda, v))\,\mu_{U\otimes V}(S)$$

of the Jacobi group over V on $L^2(\mathrm{Hom}(X\to U))$. Combining this construction with [13, Prop. 2], one obtains a factorization property generalizing Corollaries 2.2 and 2.3. This is stated precisely (and with a self-contained proof) in Lemma 4.1.

Lemma 4.1. *Assume $U = U_1\oplus U_2$ is an orthogonal direct sum and let $u = u_1\oplus u_2$ with $u_i\in U_i$. Then, under the same ismorphism as introduced in Corollary 2.2,*

$$\mathcal{C}l^{(u)}_{U\otimes V} \simeq \mathcal{C}l^{(u_1)}_{U_1\otimes V}\otimes\mathcal{C}l^{(u_2)}_{U_2\otimes V}.$$

If $u=\sum_{i=1}^t f_i$ for an orthogonal basis $\{f_i\}_{i=1}^t$ as in (2.10), then

$$\mathcal{C}l^{(u)}_{U\otimes V} \simeq \underbrace{\mathcal{C}l_V\otimes\cdots\otimes\mathcal{C}l_V}_{(t-1)\times}\otimes\mathcal{C}l_V^{(d(U))}.$$

In particular, if $U=\mathbb{F}_q^t$ and f_i is the standard orthonormal basis, then

$$\mathcal{C}l^{(u)}_{\mathbb{F}_q^t\otimes V} \simeq \mathcal{C}l_V^{\otimes t}.$$

Proof of Lemma 4.1. The symplectic subgroup of the Clifford group factorizes according to Corollary 2.2. It remains to be shown that the same is true for the image of $W(\iota_u(\lambda, v))$ under the isomorphism (2.20). Using $\beta(u,u)=\beta(u_1,u_1)+\beta(u_2,u_2)$:

$$\begin{aligned}
&W(\iota_{u_1\oplus u_2}(\lambda, x\oplus y)\,\delta_F\\
\simeq\ &\big(\omega(-2^{-1}y(x)\beta(u_1,u_1)+\lambda\beta(u_1,u_1)+\beta(u_1,\pi_1 Fx))\delta_{\pi_1 F+u_1\otimes y}\big)\\
&\otimes\big(\omega(-2^{-1}y(x)\beta(u_2,u_2)+\lambda\beta(u_2,u_2)+\beta(u_2,\pi_2 Fx))\delta_{\pi_2 F+u_2\otimes y}\big)\\
=\ &W(\iota_{u_1}(\lambda, x\oplus y))\otimes W(\iota_{u_2}(\lambda, x\oplus y))\delta_{\pi_1 F}\otimes\delta_{\pi_2 F}.
\end{aligned}$$

The second part is proven analogously to Corollary 2.3. □

The lemma gives a correspondence between the tensor powers of the symplectic group and the tensor powers of the Clifford group. Assume that t is not a multiple of p. Let f_i be the standard orthonormal basis of $\mathbb{F}_q$. Then $u=\sum_{i=1}^t f_i$ is not isotropic, so we can decompose

$$\mathbb{F}_q^t = \langle u\rangle\oplus u^\perp =: U_1\oplus U_2, \qquad d(U_1)=d(U_2)=t.$$

Then Lemma 4.1 gives

$$(\mathcal{C}l_V)^{\otimes t} \simeq \mathcal{C}l^{(u)}_{\mathbb{F}_q^t \otimes V} = \mathcal{C}l^{(t)}_V \otimes \mu_{U_2 \otimes V} \simeq \mathcal{C}l^{(t)}_V \otimes \mu_V^{\otimes(t-2)} \otimes \mu^{(t)}_V, \tag{4.3}$$

where we have used that $u_2 = 0$ and that $W_{U_2\otimes V}(\iota_0(\lambda, v)) = W_{U_2\otimes V}(0) = \mathbb{1}$. In Eq. (4.3), the action of the Heisenberg group has been compressed to the first tensor factor, i.e. μ_V is seen here as a representation of $J(V)$ with kernel equal to $H(V)$.

We note the relationship between Eq. (4.3) and the discussion proceeding Prop. 2 in [13]: the latter assures us that because

$$(\mathcal{C}l_V)^{\otimes t}_{|H(V)} \simeq W_{U\otimes V} \circ \iota_u \simeq W^{(t)}_V,$$

then

$$(\mathcal{C}l_V)^{\otimes t} \simeq \mathcal{C}l^{(t)}_V \otimes \rho,$$

where ρ is some representation of $J(V)$ with $H(V) \subseteq \ker \rho$. Eq. (4.3) specifies that in fact $\rho \simeq \mu_V^{\otimes(t-2)} \otimes \mu^{(t)}_V$.

This observation is closely related to the discussion proceeding Prop. 2 from [13]. A consequence of it is:

Proposition 4.2. *Let $u \in U$ be anisotropic, let $U' = u^\perp$. There is a one-one correspondence between*

1. *representation spaces of the symplectic group acting via $\mu_{U'\otimes V}$ on $L^2(\mathrm{Hom}(X \to U'))$, and*
2. *representation spaces of the Jacobi group acting via $\mathcal{C}l^{(u)}_{U\otimes V}$ on $L^2(\mathrm{Hom}(X \to U))$.*

In particular, if the characteristic of $\mathbb{F}_q$ does not divide t, there is a one-one correspondence between irreducible $\mathrm{Sp}(V)$-subrepresentations of $(\mu_V)^{\otimes(t-2)} \otimes \mu^{(t)}_V$ and irreducible $H(V) \rtimes \mathrm{Sp}(V)$-subrepresentations of $\mathcal{C}l^{\otimes t}_V$.

Proof. As u is non-isotropic, we have the orthogonal direct sum

$$U = (\mathbb{F}_q\, u) \oplus U'.$$

As in the proof of Corollary 2.3, the ismorphism

$$i : L^2(\mathrm{Hom}(X \to U)) \to L^2(X^*) \otimes L^2(\mathrm{Hom}(X \to U')$$

defined by

$$\delta_F \mapsto \delta_{u^T F} \otimes \delta_{\pi_2 F},$$

realizes

$$\mathcal{C}l_{U\otimes V} \simeq \mathcal{C}l^{(t)}_V \otimes \mu_{U'\otimes V} \tag{4.4}$$

as representations of $H(V) \rtimes \mathrm{Sp}(V)$.

In the one direction, let $\mathcal{K} \subset L^2(\mathrm{Hom}(X \to U'))$ be invariant under $\mu_{U'\otimes V}$. Then

$$\mathcal{K}' := L^2(X^*) \otimes \mathcal{K} \tag{4.5}$$

is invariant under $\mathcal{C}l^{(t)}_V \otimes \mu_{U'\otimes V}$. In the other direction, let

$$\mathcal{K}' \subset L^2(X^*) \otimes L^2(\mathrm{Hom}(X \to U'))$$

be invariant under $\mathcal{C}l_V^{(t)} \otimes \mu_{U'\otimes V}$. Then, because the Weyl representation acting on the first tensor factor is irreducible, $\mathcal{K}'$ must factorize as

$$\mathcal{K}' = L^2(X^*) \otimes \mathcal{K}$$

with a suitable $\mathcal{K} \subset L^2(\mathrm{Hom}(X \to U'))$ invariant under $\mu_{U'\otimes V}$. Thus Eq. (4.5) defines a one-one correspondence $\mathcal{K} \to \mathcal{K}'$ as advertised.

For the second part, assume that p does not divide t. Let $U = \mathbb{F}_q^t$ with standard basis $\{f_i\}_i$, and set $u = \sum_{i=1}^t f_i$. Then $\beta(u,u) = t$, which is non-zero by assumption. The claim now follows from the first part and Lemma 4.1. □

If t is a multiple of p, the situation is more complicated. In that case, $u = \sum_i f_i$ is isotropic, which reflects the fact that in this case the representation

$$(\lambda, x \oplus y) \mapsto W_V^{\otimes t}(\lambda, x \oplus y)$$

is Abelian. The smallest non-degenerate subspace $U_1 \subset \mathbb{F}_q^t$ containing $u = \sum_i f_i$ is then a hyperbolic plane. Following the same recipe as above, we can therefore arrange for the Heisenberg group to act only on the first *two* copies of $L^2(X^*)$. However, the action of the Clifford group on these two copies is its adjoint action (as in Lemma 2.4). This action, unlike the case treated in Proposition 4.2, is reducible. We will analyze this situation elsewhere [18].

We close this section with a sample application of Proposition 4.2. Our goal is to directly see the equivalence of two well-known facts: (1) The Clifford group forms a unitary 2-design [6, 9], i.e. its second tensor power decomposes into a direct sum of two irreducible representations, supported on the $q^n(q^n+1)/2$-dimensional symmetric subspace, and the $q^n(q^n-1)/2$-dimensional anti-symmetric one. Here, the (anti-)symmetry is w.r.t. to an exchange of tensor factors. (2) The Weil representation of the symplectic group decomposes as the direct sum of two irreducible spaces, namely the $(q^n+1)/2$-dimensional subspace of $L^2(X^*)$ of functions that are symmetric under the reflection $y \mapsto -y$, and the $(q^n-1)/2$-dimensional subspace of anti-symmetric functions. The correspondence in Proposition 4.2 maps these two decompositions onto each other:

$$\begin{gathered}\{\text{(anti-)symm. tensors on } L^2(X^*) \otimes L^2(X^*)\} \\ \updownarrow \\ L^2(X^*) \otimes \{\text{(anti-)symm. functions on } L^2(X^*)\}.\end{gathered}$$

As a consistency check: the ortho-complement of $u = f_1 + f_2$ is spanned by $v = f_1 - f_2$. The interchange of tensor factors acts trivially on u, but changes the sign of v. Thus, the two notions of (anti-)symmetry are indeed mapped onto each other.

5. Summary and outlook

Reference [11] introduced a notion of rank for $\mathrm{Sp}(V)$-representations, and showed that there is a one-one correspondence between irreps of $O(U)$ and highest-rank $\mathrm{Sp}(V)$-irreps in $\mu_{U\otimes V}$. Here, we have classified the rank-deficient components and have achieved a decomposition of $\mu_{U\otimes V}$ in terms of irreducible and inequivalent $\mathrm{Sp}(V)$-representations.

A number of natural directions deserve further attention. Most importantly from the point of view of quantum information theory, one must treat the case of characteristic 2. We will pursue this in an upcoming paper, which will also be

written in a language better-suited for consumption by physicists [18]. While we have occassionally remarked on connections between this paper and previous works from quantum information (e.g. [10]) and coding theory (e.g. [20]), the relation between their respective approaches and the one taken in this paper should be made more explicit. Lastly, the joint action of $O(U) \times \mathrm{Sp}(V)$ should be worked out more explicitly.

Appendix A. Deferred proofs

A.1. **Factorization property of the oscillator representation.** It is possible to prove Lemma 2.1 by directly verifying the claim on a set of generators (as in Eqs. (2.2), (2.1), and (2.3)) for $\mathrm{Sp}(V_1) \times \mathrm{Sp}(V_2)$. Our approach is based on realizing that it suffices to check the factorization property for Weyl operators (as in Eq. (4.1)), and then use Eq. (4.2) to "lift" it to the oscillator representation.

Proof of Lemma 2.1. Assume that $X = X_1 \oplus X_2$ and that $V = V_1 \oplus V_2$ is the resulting decomposition of V. By computing the action on basis vectors, it is immediate that

$$i\, W_V(v)\, i^{-1} = W_{V_1}(v_1) \otimes W_{V_2}(v_2),$$

where

$$i : L^2(X^*) \to L^2(X_1^*) \otimes L^2(X_2^*), \qquad \delta_y \mapsto \delta_{y\pi_1} \otimes \delta_{y\pi_2}$$

is the isomoprhism introduced in Eq. (2.8).

Let $S \in \mathrm{Sp}(V_1)$, then, using Eq. (4.2),

$$\begin{aligned}
&\left(i\, \mu_V(S)\, i^{-1}\right)\left(i\, W(v)\, i^{-1}\right)\left(i\, \mu_V(S)^\dagger\, i^{-1}\right)\\
=&i\, W(Sv)\, i^{-1}\\
=&W_{V_1}(Sv_1) \otimes W_{V_2}(v_2)\\
=&\Big(\mu_{V_1}(S_1) \otimes \mathbb{1}\Big)\Big(W_{V_1}(v_1) \otimes W_{V_2}(v_2)\Big)\Big(\mu_{V_1}(S_1) \otimes \mathbb{1}\Big)^\dagger.
\end{aligned}$$

Since the Weyl operators $W_V(v)$ for a basis for $\mathrm{End}(L^2(X^*))$, this implies

$$i\, \mu_V(S_1)\, i^{-1} = \kappa(S)(\mu_{V_1}(S) \otimes \mathbb{1})$$

for some scalar function $\kappa : \mathrm{Sp}(V_1) \to \mathbb{C}^\times$. We now show $\kappa(S) = 1$ for all S.

Unitarity implies that $|\kappa(S)| = 1$. Because

$$i\, \mu_V|_{\mathrm{Sp}(V_1)}\, i^{-1}$$

is a representation, κ must also be a (one dimensional) representation of $\mathrm{Sp}(V_1)$. Let $\mathcal{N}_1 \rtimes \mathcal{D}_1$ be the Siegel parabolic of $\mathrm{Sp}(V_1)$ with $\mathcal{N}_1$ its unipotent radical.

Since κ is a one dimensional representation, it must contain only one weight associated to $\mathcal{N}_1$. This $\mathcal{N}_1$-weight must have a trivial orbit under conjugation by $\mathcal{D}_1 \cong \mathrm{GL}(X_1)$ transformations, and thus $\kappa|_{\mathcal{N}_1} = 1$. But $\mathrm{Sp}(V_1)$ is generated by $\mathcal{N}_1$-conjugates, so $\kappa = 1$. □

A.2. **Fourier transforms and invariance.** Here, we prove Lemma 2.9.

We begin by noting that B turns $\mathrm{Hom}(X \to U)$ into an orthogonal space with form β_B given by

$$\beta_B(F, G) = \operatorname{tr} F^* \beta G B.$$

Let $W = \mathrm{Hom}(X \to U')$. In the following we will show that

(1) for any $\Phi \in L^2 \mathrm{Hom}(X \to U)$, $\operatorname{supp} \Phi \subseteq W$ if and only if $\mu_{U\otimes V}(J_B)\Phi$ is invariant under $W^\perp$ translations,
(2) $W^\perp = \mathrm{Hom}(X \to U'^\perp)$.

These two claims imply the first statement of the lemma

For the first claim, start with the "only if" direction. Apply the inverse map (associated with $-B$) to a function $\tilde{\Phi}$ with the invariance stated. Then

$$\begin{aligned}
\Phi(F) &= \gamma(B, U) \sum_{F'} \omega(-\beta_B(F, F'))\tilde{\Phi}(F') \\
&= \gamma(B, U) \sum_{C \in \mathrm{Hom}(X\to U)/W^\perp} \tilde{\Phi}(C) \sum_{G \in W^\perp} \omega(-\beta_B(F, C + G)) \\
&= \gamma(B, U) \sum_{C \in \mathrm{Hom}(X\to U)/W^\perp} \tilde{\Phi}(C)\omega(-\beta_B(F, C)) \sum_{G \in W^\perp} \omega(-\beta_B(F, G)) \\
&= \gamma(B, U)|W^\perp|\delta_W(F) \sum_{C \in \mathrm{Hom}(X\to U)/W^\perp} \tilde{\Phi}(C)\omega(-\beta_B(F, C)).
\end{aligned}$$

Conversely, the set of Φ's with support in W is a vector space of dimension $|W|$. At the same time, the set of solutions we have identified in the direct direction has dimension

$$\big|\mathrm{Hom}(X \to U)/W^\perp\big| = \frac{|\mathrm{Hom}(X \to U)|}{|W^\perp|} = \frac{|\mathrm{Hom}(X \to U)|\,|W|}{|\mathrm{Hom}(X \to U)|} = |W|,$$

so we have found all solutions.

Now we prove the second claim. Assume F is such that $\beta_B(F, F') = 0$ for all $F' \in \mathrm{Hom}(X \to U')$, and choose $F' = u \otimes y$ for $y \in X^*$ and $u \in U'$. Note that

$$F^* \beta F' B = F^*\big(\beta(u)\big) \otimes (By),$$

where we used the fact that B is symmetric. Hence

$$\beta_B(F, F') = \big(F^*\beta(u)\big)(By) = \beta(u, FBy).$$

Because $y \in X^*$ and $u \in U'$ are arbitrary, and because B is surjective, it follows that $F \in \mathrm{Hom}(X \to U'^\perp)$ and with this the claim also follows.

Now on to the second statement of the lemma. Acting explicitly on the indicator function of W we get

$$\mu_{U\otimes V}(J_B) \sum_{F\in W} \delta_F = \gamma(B)^{-1} \sum_{F\in W} \sum_{F'\in W^\perp} \omega^{\beta_B(F,F')}\delta_{F'},$$

where we used the first statement to restrict the sum over F'. Notice that by the definition of $W^\perp$, every coefficient in the expression above is 1, so that

$$\mu_{U\otimes V}(J_B) \sum_{F\in W} \delta_F = \gamma(B)^{-1}|W| \sum_{F'\in W^\perp} \delta_{F'},$$

as claimed.

A.3. **Contiguity of ranks.** In this section we prove Proposition 2.10. Our strategy will be to find a set of generators for $\mathrm{Sp}(V)$ which consists of elements that either keep the rank of an $\mathcal{N}$-weight invariant or change it by at morankst one. It then follows that if the ranks of the $\mathcal{N}$-spectrum has a gap, then the representation is reducible.

Recall that $\mathcal{K}$ is an irreducible representation of rank $< t$, and

$$R := \{\, k \,|\, \text{there is an } \mathcal{N}\text{-weight with rank } k\}.$$

Let $\mathcal{K}^r$ be the subspace of $\mathcal{K} \subset L^2(\mathrm{Hom}(X \to U))$ that is spanned by $\mathcal{N}$-weights of rank r. This space is invariant under the action of the parabolic subgroup $\mathcal{N} \rtimes \mathcal{J}$. We will analyze its image under the Fourier transforms $\mathcal{J}$.

Let $\{e_i\}$ be an arbitrary basis of X and $\{\varepsilon_i\}$ be its dual basis. For any isomorphism $B : X^* \to X$, there is a $C \in \mathrm{GL}(X)$ satisfying

$$CB\varepsilon_i = e_i, \quad \forall\, i.$$

Using the isomorphism in Eq. (2.8), we find that there exist a set of $B_i : \mathrm{span}\{\varepsilon_i\} \to \mathrm{span}\{e_i\}$ for which

$$i\, \mu_{U\otimes V}(C)\mu_{U\otimes V}(J_B)\, i^{-1} = \mu_{U\otimes V_1}(J_{B_1}) \otimes \cdots \otimes \mu_{U\otimes V_n}(J_{B_n}),$$

where $V_i := \mathrm{span}\{\varepsilon_i, e_i\}$. It follows that $\mathrm{Sp}(V)$ is generated by the parabolic subgroup together with any $i^{-1}\,(\mu_{U\otimes V_1}(J_{B_1}) \otimes \mathbb{1})\, i$ (sometimes referred to as *single-system Fourier transform* in quantum information theory). Let $X_1 = \mathrm{span}\{e_1\}$, $X_2 = \mathrm{span}\{e_2, \ldots, e_n\}$, and let $V = V_1 \oplus V_1^\perp$ be the corresponding decomposition of V, where $V_1^\perp = X_2 \oplus X_2^*$ is the symplectic complement of V_1. Let π_1, π_2 be the projections associated with the decomposition $\mathrm{Hom}(X \to U) = \mathrm{Hom}(X_1 \to U) \oplus \mathrm{Hom}(X_2 \to U)$. We see that

$$\big(i^{-1}\,(\mu_{U\otimes V_1}(J_{B_1}) \otimes \mathbb{1})\, i\big)\delta_F = \gamma^{-1}(B_1) \sum_{F' \in \mathrm{Hom}(X_1 \to U)} \omega\left(\beta_B(F', \pi_1 F)\right) \delta_{\pi_2 F + F'}.$$

Throughout the rest of the argument, let $\mathrm{rank}\, F^T F = r$.

Now, range $\pi_2 F$ is either equal to range F or it is a subspace of the latter with co-dimension 1. Thus $\mathrm{rank}(\pi_2 F)^T(\pi_2 F) \in \{r, r-1\}$. Furthermore, either range $\pi_2 F + F' =$ range $\pi_2 F$ or

$$\text{range } \pi_2 F \subset \text{range } \pi_2 F + F'$$

is a subspace of co-dimension 1. Thus,

$$\mathrm{rank}(\pi_2 F + F')^T(\pi_2 F + F') \in \{r-1,\, r,\, r+1\},$$

for any $F' \in \mathrm{Hom}(X_1 \to U)$. This implies that

$$i^{-1}\,(\mu_{U\otimes V_1}(J_{B_1}) \otimes \mathbb{1})\, i : \mathcal{K}^r \to \mathcal{K}^{r-1} + \mathcal{K}^r + \mathcal{K}^{r+1}.$$

If for some $r \in R$ it held that $\mathcal{K}^l = \{0\}$ then the spaces $\sum_{r>l} \mathcal{K}^r$ and $\sum_{r<l} \mathcal{K}^r$ would be invariant under all generators (and thus subrepresentations). Since $\mathcal{K}$ is irreducible by assumptiuon, this cannot happen.

Acknowledgments

The authors thank Mateus Araujo, Victor Bankston, Markus Heinrich, Sepehr Nezami, and Michael Walter for interesting discussions.

References

[1] Noga Alon and Joel H. Spencer, *The probabilistic method*, 4th ed., Wiley Series in Discrete Mathematics and Optimization, John Wiley & Sons, Inc., Hoboken, NJ, 2016. MR3524748

[2] Anne-Marie Aubert and Tomasz Przebinda, *A reverse engineering approach to the Weil representation*, Cent. Eur. J. Math. **12** (2014), no. 10, 1500–1585, DOI 10.2478/s11533-014-0428-8. MR3224014

[3] A Robert Calderbank and Peter W Shor, *Good quantum error-correcting codes exist*, Phys. Rev. A **54** (1996), no. 2, 1098.

[4] Peter J Cameron, *Notes on classical groups*, (2000), available at `http://www.maths.qmul.ac.uk/~pjc/class_gps/cg.pdf`.

[5] Wai Kiu Chan, *Arithmetic of quadratic forms*, (2019), available at `http://wkchan.faculty.wesleyan.edu/files/2019/04/qflecturenotes.pdf`.

[6] Christoph Dankert, Richard Cleve, Joseph Emerson, and Etera Livine, *Exact and approximate unitary 2-designs and their application to fidelity estimation*, Phys. Rev. A **80** (2009), 012304.

[7] Gerald B. Folland, *Harmonic analysis in phase space*, Annals of Mathematics Studies, vol. 122, Princeton University Press, Princeton, NJ, 1989. MR983366

[8] Paul Gérardin, *Weil representations associated to finite fields*, J. Algebra **46** (1977), no. 1, 54–101, DOI 10.1016/0021-8693(77)90394-5. MR460477

[9] D. Gross, K. Audenaert, and J. Eisert, *Evenly distributed unitaries: on the structure of unitary designs*, J. Math. Phys. **48** (2007), no. 5, 052104, 22, DOI 10.1063/1.2716992. MR2326329

[10] David Gross, Sepehr Nezami, and Michael Walter, *Schur-Weyl duality for the Clifford group with applications: Property testing, a robust Hudson theorem, and de Finetti representations*, arXiv preprint `arXiv:1712.08628` (2017).

[11] Shamgar Gurevich and Roger Howe, *Small representations of finite classical groups*, Representation theory, number theory, and invariant theory, Progr. Math., vol. 323, Birkhäuser/Springer, Cham, 2017, pp. 209–234. MR3753913

[12] Jonas Helsen, Joel J. Wallman, and Stephanie Wehner, *Representations of the multi-qubit Clifford group*, J. Math. Phys. **59** (2018), no. 7, 072201, 20, DOI 10.1063/1.4997688. MR3827133

[13] Roger E. Howe, *On the character of Weil's representation*, Trans. Amer. Math. Soc. **177** (1973), 287–298, DOI 10.2307/1996597. MR316633

[14] Roger Howe, *Remarks on classical invariant theory*, Trans. Amer. Math. Soc. **313** (1989), no. 2, 539–570, DOI 10.2307/2001418. MR986027

[15] M. Kashiwara and M. Vergne, *On the Segal-Shale-Weil representations and harmonic polynomials*, Invent. Math. **44** (1978), no. 1, 1–47, DOI 10.1007/BF01389900. MR463359

[16] Richard Kueng and David Gross, *Qubit stabilizer states are complex projective 3-designs*, arXiv preprint `arXiv:1510.02767` (2015).

[17] T. Y. Lam, *Introduction to quadratic forms over fields*, Graduate Studies in Mathematics, vol. 67, American Mathematical Society, Providence, RI, 2005. MR2104929

[18] Felipe Montealegre-Mora and David Gross, *The representation theory of Clifford tensor powers* (2021), in preparation.

[19] Gabriele Nebe, E. M. Rains, and N. J. A. Sloane, *The invariants of the Clifford groups*, Des. Codes Cryptogr. **24** (2001), no. 1, 99–121, DOI 10.1023/A:1011233615437. MR1845897

[20] Gabriele Nebe, Eric M. Rains, and Neil J. A. Sloane, *Self-dual codes and invariant theory*, Algorithms and Computation in Mathematics, vol. 17, Springer-Verlag, Berlin, 2006. MR2209183

[21] Sepehr Nezami and Michael Walter, *Multipartite entanglement in stabilizer tensor networks*, arXiv preprint `arXiv:1608.02595` (2016).

[22] Michael A. Nielsen and Isaac L. Chuang, *Quantum computation and quantum information*, Cambridge University Press, Cambridge, 2000. MR1796805

[23] Bernhard Runge, *On Siegel modular forms. I*, J. Reine Angew. Math. **436** (1993), 57–85, DOI 10.1515/crll.1993.436.57. MR1207281

[24] Andrew Steane, *Multiple-particle interference and quantum error correction*, Proc. Roy. Soc. London Ser. A **452** (1996), no. 1954, 2551–2577, DOI 10.1098/rspa.1996.0136. MR1421749

[25] A. M. Steane, *Error correcting codes in quantum theory*, Phys. Rev. Lett. **77** (1996), no. 5, 793–797, DOI 10.1103/PhysRevLett.77.793. MR1398854

[26] Zak Webb, *The Clifford group forms a unitary 3-design*, Quantum Inf. Comput. **16** (2016), no. 15-16, 1379–1400. MR3616033
[27] Huangjun Zhu, *Multiqubit Clifford groups are unitary 3-designs*, Phys. Rev. A **96** (2017), no. 6, 062336, 7, DOI 10.1103/physreva.96.062336. MR3746769
[28] Huangjun Zhu, Richard Kueng, Markus Grassl, and David Gross, *The Clifford group fails gracefully to be a unitary 4-design*, arXiv preprint `arXiv:1609.08172` (2016).

Institute for Theoretical Physics, University of Cologne, 50937 Cologne, Germany
Email address: `fmonteal@thp.uni-koeln.de`

Institute for Theoretical Physics, University of Cologne, 50937 Cologne, Germany
Email address: `david.gross@thp.uni-koeln.de`

The representation theory of Clifford tensor powers

Felipe Montealegre-Mora* and David Gross
Institute for Theoretical Physics, University of Cologne, 50937 Cologne, Germany†
(Dated: November 17, 2021)

Tensor power representations of the Clifford group play an increasingly prominent role in the construction of protocols for quantum system certification, quantum simulation, quantum cryptography, among others. Beyond physics, Clifford tensor powers have made appearances in areas such as invariant theory, Theta duality and the construction of unitary designs. Here we fully decompose these tensor power representations. For this, we generalize the rank theory of symplectic representations to the case of the Clifford group. In this way, we generalize the "eta correspondence" between the symplectic and orthogonal groups to a correspondence between the Clifford and orthogonal-stochastic groups. These results go beyond our previous works in two main regards: 1. They work for arbitrary t-th tensor powers with $t \leq n$ (where n is the number of qudits). 2. They work for both the qubit and odd qudit cases. As a sample application, we provide a protocol to efficiently implement the complex conjugate of a black-box Clifford unitary evolution.

CONTENTS

* fmonteal@thp.uni-koeln.de
† We thank Oscar Garcia, Shamgar Gurevich, Markus Heinrich, Sepehr Nezami, and Michael Walter for very insightful discussions. We also thank Markus Heinrich and Pascal Bassler for bringing the problem of black box Clifford conjugation to our attention. This work has been supported by the Excellence Initiative of the German Federal and State Governments (Grant ZUK 81), the ARO under contract W911NF-14-1-0098 (Quantum Characterization, Verification, and Validation), and the DFG (SPP1798 CoSIP, project B01 of CRC 183).

I. INTRODUCTION AND SUMMARY OF RESULTS

Tensor power representations of the Clifford group Cl,

$$\mathrm{Cl} \ni U \mapsto U^{\otimes t},$$

have played an increasingly prominent role in quantum information theory in the past years. Studying these has, for example, lead to efficient protocols for quantum device characterization [1–4], quantum state distinction [5], compressed sensing [6], stabilizerness testing [7], bounding the stabilizer rank of magic states [8], as well as complex projective [7, 9] and unitary [10] t-designs.

Tensor power representations with $t \leq 4$ are well understood [9, 11–13]. An important characteristic here is that the Clifford group is a *unitary 3-design*, ie. the uniform distribution on the Clifford group has the same third order moments as the Haar distribution on the full unitary group. Moreover, the Clifford group is singled out among unitary designs as the only infinite family of finite groups forming 3-designs [14]. Generalising these results to larger t, Ref. [7] studied the commutant of general t-th tensor power representations. While a description of the commutant may in general be used to decompose the representation, this insight was left for future work in that reference. In our previous work, Ref. [15], we provided a full decomposition of these representations in the case of odd qudits and where t is not divisible by the local dimension d. The question of generalising these results to the qubit and t-multiple-of-d cases was left open.

In this paper, solve this open problem. Our contribution goes beyond previous works on various regards.

1. We generalize the main result of Ref. [15] to the case of qubits, and to the case where t is a multiple of d.

2. While the main result of Ref. [15] was inspired by Ref. [7], the proof strategies used in the former were independent of the results latter. Thus, the question of whether the proofs of Ref. [15] could be simplified by exploiting directly the main result of Ref. [7] arises. Here we answer this question in the affirmative.

3. Refs. [16, 17] develop a rank theory for representations of the symplectic group. This rank theory is crucial in the generalization of the Theta correspondence [18, 19] between the real orthogonal and symplectic groups, to the scenario over finite fields [16, 17]. Here, we extend this theory to a rank theory for Clifford representations. This way, we generalize the results of [16] to the Clifford group.

4. Refs. [20–23] study the related problem of classifying polynomial invariants of the Clifford group —equivalently, they study the trivial Cl-subrepresentation of $\mathrm{Sym}_t(\mathbb{C}^{d^n})$. There, certain self-dual codes span the space of these invariants. Here, we show that these invariants correspond to rank-0 subrepresentations of Clifford tensor powers. We furthermore show that all subrepresentations with rank $< t$ are contained in the span of certain self-orthogonal codes, of which a subset are the aforementioned self-dual codes. In this sense, our result extends Refs. [20, 21, 23] to non-zero rank representations.

5. In Ref. [9], the code space with stabilizer group $\{P^{\otimes 4} \mid P \text{ } n\text{-qubit Pauli } \}$ was studied in detail. In particular, it was found that the *real* Clifford group acts as a permutation representation on this space. Here we generalize this result to $t = 4k$, $k \geq 1$, and show that this representation gives rise to a duality between the real Clifford group and a symplectic group. We conjecture that, in analogy to [16], on a given maximal-rank subspace, this duality gives a correspondence between irreps of the real Clifford group and the symplectic group.

6. As a sample application of our results, we study the problem of using a Clifford black box U to implement $\bar{U}$ (*Clifford black box conjugation*). We provide an optimal parallel inversion protocol for Clifford black boxes and show that the number of black box uses needed for this is $\sim d$. This is in stark contrast to the problem of inverting arbitrary unitaries, which requires $d^n - 1$ black box uses.

Many of the results found in the works cited above are closely analogous to our results. Here we summarize these interrelations between the literature and this paper.

Ref. [16] develops a theory of rank for $\mathrm{Sp}(2n, q)$ representations. They show that certain "maximal rank" subrepresentations of tensor powers of the oscillator representation set up a correspondence between $\mathrm{Irr}\,\mathrm{Sp}(2n, q)$ and $\mathrm{Irr}\,\mathrm{O}^{\pm}(t, q)$. The rank theory developed in that reference is analogous to the rank theory in Sec. IV. In fact, Thm. V.1 is a generalization of the main theorem in [16] and their proofs are conceptually very similar. The analogy between both rank theories is broken in two prominent places. First, it is broken in a mild way in the qubit case, where having zero rank does not imply triviality (compare Lem. IV.3 to [16, Lem. 1.3.1]). Instead, rank zero representations of the qubit Clifford group are in general ± 1 valued. Second, while [16] requires $t \leq n$, we are able to also describe certain low-rank sectors even when $t > n$ (see the relevant condition in Thm. V.1).

Ref. [7] provides a description for the commutant of Clifford tensor powers. This amounts to computing the trivial component of "balanced" tensor powers $U \mapsto U^{\otimes t} \otimes \bar{U}^{\otimes t}$. Here we build on this result to find a description of the full representation theory of *arbitrary* tensor powers $U \mapsto U^{\otimes r} \otimes \bar{U}^{\otimes s}$. In this way, while the methods of [16] (appropriately generalized) describe the

"maximal rank" component of tensor powers, and the methods of [7] describe the "rank zero" component, this paper describes all other rank components.

Our main result, Thm. V.2 is closely analogous to [15, Thm. 1.2]. Three points are important in this regard. First, that [15] holds over finite fields of characteristic not two, while our results hold for finite fields of prime order (in particular, our results hold for the qubit case). We believe that our results can be straightforwardly generalized to arbitrary finite fields but we will not attempt to show this. Second, Ref. [15, Prop. 4.2] describes the representation theory of Clifford tensor powers indirectly, heavily relying on its symplectic counter-part [15, Thm. 1.2]. Here, our proof of Thm. V.2 directly describes the Clifford representation theory. This allows us to generalize the aforementioned result to the qubit case and the case where $t = 0 \mod d$. Third, even if the statements Thm. V.2 and [15, Thm. 1.2] are analogous, our proof strategies are considerably different. Importantly, while in [15] we resorted to probabilistic proof strategies, the proofs presented here are much more constructive.

As mentioned, Sec. VI looks at a subrepresentation which generalizes the one studied in [9, Sec. 3.4 and App. B]. Our emphasis here is different, however. That reference focuses on singling out the full Clifford action on this space, where they find that it is equivalent to an analog of the oscillator representation of $\mathrm{Sp}(2n, 2)$. Here, we focus on the action of the real Clifford group on it, which gives rise to a new correspondence between $\mathrm{O}^+(2n, 2)$ and $\mathrm{Sp}(2t', 2)$, where $t = 2t'$ is even.

Finally, [24, Thm. 3] provides a list of Clifford tensor powers which lead to an exact correspondence between the Clifford group and an orthogonal group $\mathrm{O}(\mathbb{Z}_d^t)$. In Sec. V D we show that this list is essentially complete.

II. PRELIMINARIES

Here we summarize some relevant results found in the literature.

A. The Pauli and Clifford groups

Consider n qudits, with Hilbert space $\mathcal{H}_n := \left(\mathbb{C}^d\right)^{\otimes n}$ and computational basis states $|x\rangle$ where $x \in \mathbb{Z}_d^n =: \mathrm{X}$. The main results of this paper hold only in the case where d is a prime, however the presentation within this subsection holds for arbitrary d.

Let the *shift* and *clock* operators be

$$X(q)\,|x\rangle = |x + q\rangle \qquad Z(p)\,|x\rangle = \omega^{p\cdot x}\,|x\rangle\,, \qquad p, q \in \mathrm{X},$$

where $\omega = \exp(2\pi i/d)$. Because the construction of the Pauli and Clifford groups differs slightly depending on whether d is even or odd, it is convenient to introduce $\tau = (-1)^d \exp(i\pi/d)$, and let D be the order of τ. Notice that $\tau^2 = \omega$, that $d = D$ if and only if d is odd, and otherwise $D = 2d$ if d is even. Then, the *Weyl operators* (also known as *displacement operators*) are

$$W_v := \tau^{-v_z\cdot v_x} Z(v_z)X(v_x),$$

where $v = (v_z, v_x) \in \mathbb{Z}_d^{2n} := V$. The n-qudit Pauli group $\mathcal{P}$ is generated by the displacement operators and $\tau\mathbb{1}$.[1] Its defining representation, known as the *Weyl representation*, is denoted W.

One may define several inequivalent representations of $\mathcal{P}$ acting on $\mathcal{H}_n$. For each $m \in \mathbb{Z}_d^+ := \mathbb{Z}_d\backslash\{0\}$ define the displacement operator *with mass* m to be:

$$W^{(m)}_{(v_z,v_x)} := W_{(mv_z,v_x)}.$$

Then, the Weyl representation with mass m is given by

$$W^{(m)} : \begin{cases} \tau\mathbb{1} & \mapsto \tau^m\mathbb{1}, \\ W_v & \mapsto W_v^{(m)}. \end{cases}$$

The vector space V naturally hosts the symplectic product

$$[v, u] = v_z \cdot u_x - v_x \cdot u_z,$$

[1] In the mathematical literature the Pauli group is also known as the Heisenberg group and as the extra-special p-group.

arising through the commutation relation

$$W_v^{(m)} W_u^{(m)} = \omega^{m[v,u]} W_u^{(m)} W_v^{(m)}.$$

One may further verify that

$$W_v^{(m)} W_u^{(m)} = \tau^{m[v,u]} W_{v+u}^{(m)}.$$

so that any element of $\mathcal{P}$ is equal to some W_v up to a phase.

The Clifford group [7, 25–27], denoted Cl, is the subgroup of $\mathrm{U}(\mathcal{H}_n)$ generated by the discrete Fourier transform on any given qudit (also known as *Hadamard gate* for $d = 2$)

$$\mathrm{H} = \frac{1}{\sqrt{d}} \sum_{x,y \in \mathbb{Z}_d} \omega^{xy} |x\rangle \langle y| , \tag{1}$$

the phase gate acting on any qudit,

$$\mathrm{P} = \sum_{x \in \mathbb{Z}_d} \tau^{x^2} |x\rangle \langle x| \quad (d \text{ even}), \quad \mathrm{P} = \sum_{x \in \mathbb{Z}_d} \tau^{x(x-1)} |x\rangle \langle x| \quad (d \text{ odd}), \tag{2}$$

and the controlled addition acting on any pair of qudits,

$$\mathrm{CADD} = \sum_{x,y} |x, x+y\rangle \langle x, y| . \tag{3}$$

When $d = 2$ we use, alternatively, the standard notation $\mathrm{CNOT} := \mathrm{CADD}$.

The Clifford group contains and normalizes the Pauli group, $\mathcal{P} \triangleleft \mathrm{Cl}$. Furthermore every character-preserving automorphism of the Pauli group can be realized by conjugating with some Clifford matrix, and in particular $\mathrm{Cl}/\mathcal{P} \simeq \mathrm{Sp}(V)$. Each automorphism corresponds to a coset $U\mathcal{Z}(\mathrm{Cl}) \subset \mathrm{Cl}$, where $U \in \mathrm{Cl}$. In the odd d case, $\mathcal{Z}(\mathrm{Cl}) = \langle \omega \mathbb{1} \rangle \simeq \mathbb{Z}_p$, while in the qubit case $\mathcal{Z}(\mathrm{Cl}) = \langle \omega_8 \mathbb{1} \rangle \simeq \mathbb{Z}_8$ whenever $n \geq 2$ [20], where $\omega_8 = \exp(2\pi i/8)$.[2] If d is *odd* it further holds that $\mathrm{Cl} \simeq \mathcal{P} \rtimes \mathrm{Sp}(V)$ and, up to a phase, any $U \in \mathrm{Cl}$ can be expressed [28] as

$$U = e^{i\varphi} W(v) \mu(S),$$

where μ is a unitary representation of the symplectic group $\mathrm{Sp}(V)$ known as the *oscillator representation* or the *Weil representation*. The *projective Clifford group* is $\mathrm{PCl} := \mathrm{Cl}/\mathcal{Z}(\mathrm{Cl})$. If d is odd, one can see that $\mathrm{PCl} \simeq V \rtimes \mathrm{Sp}(V)$ is the affine symplectic group. In the qubit case, $d = 2$, this separation of Cl and PCl into a semi-direct product ceases to hold [29].

The oscillator representation satisfies

$$\mu(S) W(v) \mu(S)^\dagger = W(Sv).$$

One may similarly define the oscillator representation with mass m, denoted by $\mu^{(m)}$, as the unique representation of $\mathrm{Sp}(V)$ satisfying

$$\mu^{(m)}(S) W^{(m)}(v) \mu^{(m)}(S)^\dagger = W(Sv). \tag{4}$$

It turns out that $\mu^{(m)}$ only depends on whether m is a square or not in $\mathbb{Z}_d$. In the former case, $\mu^{(m)} = \mu$, while if m is a non-square then $\mu^{(m)} = \bar{\mu}$. Finally, by (4), it follows that $W^{(m)}$ and $\mu^{(m)}$ generate a representation of Cl which we will denote by $\mathrm{Cl}_{(m)}$.

In previous studies of the oscillator representation [15–17] the following subgroup of $\mathrm{Sp}(V)$ has figured prominently,

$$\mathcal{N} := \left\{ N_B = \begin{pmatrix} \mathbb{1}_n & B \\ 0 & \mathbb{1}_n \end{pmatrix} : B = B^T \in \mathbb{Z}_d^{n \times n} \right\}.$$

One may verify [16] that

$$\mu(N_B) = \sum_{x \in \mathrm{X}} \tau^{2^{-1} x^T B x} |x\rangle\langle x| .$$

[2] For $n = 1$ the center is $\langle i\mathbb{1} \rangle \simeq \mathbb{Z}_4$. Furthermore, a careful choice of the phase in front of the H generator can make the center be $\mathbb{Z}_4$ for larger n as well [9].

B. Quadratic and bilinear forms

Here we review some results on quadratic and bilinear forms over finite fields. We mainly follow [30], taking a few results from [31] for the case $d = 2$.

We partition $\mathbb{Z}_d^+ := \mathbb{Z}_d \setminus \{0\}$ into two *square classes*: the squares are given by

$$\mathbb{Z}_d^{\times 2} := \{a^2 \,|\, a \in \mathbb{Z}_d^+\},$$

and the non-squares are their complement. The *Legendre symbol* of a number $a \in \mathbb{Z}_d$ is,

$$\ell(a) = \begin{cases} 0, & \text{if } a = 0, \\ +1, & \text{if } a \in \mathbb{Z}_d^{\times 2}, \\ -1, & \text{else.} \end{cases}$$

Let $K = \mathbb{Z}_d^k$ be a vector space. A *bilinear form* on K is a map $\beta : K \times K \to \mathbb{Z}_d$ which is linear in both arguments. It is symmetric if $\beta(v, u) = \beta(u, v)$ for all $u, v \in K$, and it is alternating if $\beta(v, v) = 0$ for all $v \in K$. The spaces of symmetric and alternating forms over K are denoted, respectively, $\mathrm{Sym}(K)$ and $\mathrm{Alt}(K)$. The *radical* of a bilinear form β is given by,

$$\mathrm{rad}\,(\beta) = \{u \in K \,|\, \beta(u, v) = 0 \,\forall\, v \in K\},$$

if $\mathrm{rad}\,(\beta) = 0$, we say β is *non-degenerate*. We denote by β_0 the non-degenerate form inherited to $K/\mathrm{rad}\,(\beta)$ and let $\mathrm{rank}\,(\beta) = \dim K/\mathrm{rad}\,(\beta)$.

A *quadratic form* is a map $q : K \to \mathbb{Z}_d$ such that for every pair $u, v \in K$, the function

$$\beta(u, v) := q(u + v) - q(u) - q(v) \tag{5}$$

is bilinear, and $q(\alpha u) = \alpha^2 q(u)$ for $\alpha \in \mathbb{Z}_d$. The form q is a *quadratic refinement* of β, and conversely, β is the *polarisation* of q. We say q is non-degenerate whenever β is. A *generalized quadratic refinement* of β is a map $q : K \to \mathbb{Z}_D$ satisfying

$$q(u + v) - q(u) - q(v) = 2\beta(u, v) \mod D, \tag{6}$$

and $q(\alpha u) = \alpha^2 q(u) \mod D$. The space of quadratic forms over K is denoted $\mathrm{Q}(K)$, while the space of generalized quadratic refinements over K is denoted $\tilde{\mathrm{Q}}(K)$. In the case where $d = D$ is odd, then these two sets coincide, but if $d = 2$ they are different.

When d is odd there's a one-to-one correspondence beween quadratic forms and symmetric bilinear forms through

$$2^{-1}\beta(u, u) = q(u). \tag{7}$$

This correspondence fails if $d = 2$. In this case, the function $u \mapsto \beta(u, u)$ is a linear *and* a quadratic form. This way, there exists a vector $v_\beta \in K/\mathrm{rad}\,(\beta)$ for which

$$\beta_0(u, u) = \beta_0(u, v_\beta), \qquad \forall\, u \in K/\mathrm{rad}\,(\beta). \tag{8}$$

The *type* of β is *even* if $v_\beta = 0$ and *odd* otherwise. That is, β is even if and only if it is alternating.

Example II.1. *Two examples of generalized quadratic refinements are the following, where $d = 2$. First, if $q \in \mathrm{Q}(K)$, then we may define the form $2q \in \tilde{\mathrm{Q}}(K)$ by*

$$(2q)(u) = 2\delta(q(u) - 1).$$

If q is a refinement of β, then $2q$ is a generalized refinement of the same form.

Second, we may define the form $q_{r,s}$ given on some basis $\{e_i\}$ of K by

$$q_{r,s}(e_i) = \begin{cases} +1 & \text{if } i \le r, \\ -1 & \text{if } i > r, \end{cases}$$

here $s := \dim K - r$. For any value of r, $q_{r,s}$ is a generalized refinement of the standard dot product, $e_i \cdot e_j = \delta_{ij}$.

A matrix representation of a symmetric form β with respect to a basis $\{e_i\}$ of K is a matrix M with components

$$M_{ij} = \beta(e_i, e_j).$$

In this way, writing $u, v \in K$ as tuples of coordinates with respect to this basis, $u^T M v = \beta(u, v)$. The representation is uniquely specified by the choice of basis.

Similarly, a matrix representation M of a quadratic form $q \in \mathrm{Q}(K)$ with respect to a basis $\{e_i\}$ is such that $u^T M u = q(u)$. The representation is no longer uniquely specified by the basis: indeed if A is an alternating matrix then $M + A$ also represents q. Next we show the converse statement.

Proposition II.1. *Let M and M' be matrix representations of $q \in \mathrm{Q}(K)$ with respect to a basis $\{e_i\}$. Then,*

$$M_{ij} + M_{ji} = M'_{ij} + M'_{ji}, \qquad M_{ii} = M'_{ii}, \qquad \forall\, i < j.$$

Proof. Let β be such that q refines it, and let M_β be its matrix representation with respect to $\{e_i\}$. Then by eq. (5), $M + M^T = M_\beta = M' + (M')^T$. If d is odd, then this finishes the proof, otherwise we still need to show that $M_{ii} = M'_{ii}$. But this follows from $e_i^T M e_i = q(e_i) = e_i^T M' e_i$. □

1. Classification of forms

Two spaces $K,\ K'$ equipped with quadratic or bilinear forms $B,\ B'$ are *equivalent as formed spaces*, denoted $K \simeq K'$, if there exists an invertible linear map $M : K \to K'$ for which $B' = B \circ M$. In the case $K = K'$, we obtain the notion of $\mathrm{Gl}(K)$-equivalence between forms. Specifically, the spaces $\mathrm{Sym}(K)$, $\mathrm{Alt}(K)$, $\mathrm{Q}(K)$ and $\tilde{\mathrm{Q}}(K)$ are naturally $\mathrm{Gl}(K)$ representations through

$$gq(\,\cdot\,) = q(g^{-1}\,\cdot\,), \tag{9}$$

$$g\beta(\,\cdot\,,\,\cdot\,) = \beta(g^{-1}\,\cdot\,, g^{-1}\,\cdot\,). \tag{10}$$

The equivalence of forms up to $\mathrm{Gl}(K)$ is obtained by the following propositions. There we denote the $\mathrm{Gl}(K)$ equivalence of forms by $\sim$, e.g. $q \sim q' \in \mathrm{Q}(K)$ if there is some $g \in \mathrm{Gl}(K)$ such that $q' = gq$.

Example II.2. *When $d = 2$ there are* two *equivalence classes of quadratic refinements to the same non-degenerate bilinear form β. This can be exemplified through the* hyperbolic plane*, defined for any d as $\mathbb{H} = \mathbb{Z}_d^2$ with basis $\{e, f\}$, equipped with a form $\beta_\mathbb{H}$ given by*

$$\beta_\mathbb{H}(e, e) = \beta_\mathbb{H}(f, f) = 0, \qquad \beta_\mathbb{H}(e, f) = 1.$$

In the case $d = 2$, the following two quadratic forms are compatible with $\beta_\mathbb{H}$,

$$q^0_\mathbb{H}(e) = q^0_\mathbb{H}(f) = 0 \qquad q^0_\mathbb{H}(e + f) = 1,$$
$$q^1_\mathbb{H}(e) = q^1_\mathbb{H}(f) = 1 \qquad q^1_\mathbb{H}(e + f) = 0.$$

The following results classify symmetric, quadratic, and generalized quadratic forms.

Proposition II.2. *Let d be odd, $\beta \in \mathrm{Sym}(K)$, and β_0 be the restriction of β to $K/\mathrm{rad}\,(\beta)$. The* discriminant *of β,*

$$\mathrm{dis}(\beta) := \ell(\det M_{\beta_0}),$$

is independent of the basis used to construct the matrix representation M_{β_0} of β_0, and so it is $\mathrm{Gl}(K)$-invariant.

Furthermore, if $\mathrm{dis}(\beta) = \mathrm{dis}(\beta')$ and $\mathrm{rank}\,(\beta) = \mathrm{rank}\,(\beta')$, then $\beta \sim \beta'$ and for any $a \in \mathbb{Z}_d$ satisfying $\ell(a) = \mathrm{dis}(\beta)$, one can choose

$$M_{\beta_0} = \mathrm{diag}(1, \ldots, 1, a).$$

By eq. (7), Prop. II.2 gives a classification of the orbits on $\mathrm{Q}(K)$ for odd characteristic as well.

Proposition II.3. *Let $d = 2$ and $\beta \in \mathrm{Sym}(K)$. Then exactly one of the following holds:*

1. *The type of β is odd and there exists an orthonormal basis $\{e_i\}$ of $K/\mathrm{rad}\,(\beta)$,*

$$\beta_0(e_i, e_j) = \delta_{ij}.$$

2. *The type of β is even,* $\dim(K/\mathrm{rad}(\beta)) = 2m$ *and* $\beta_0 \sim \beta_{\mathbb{H}}^{\oplus m}$.

This way, the $\mathrm{Gl}(K)$ *orbits on* $\mathrm{Sym}(K)$ *are labeled by rank and type.*

Proposition II.4. *Let $d = 2$ and $q \in \mathrm{Q}(K)$ be a quadratic refinement of the non-degenerate form β. Then β is even and* $\dim K = 2m$. *Furthermore, consider the* Arf invariant *of q,* $\mathrm{Arf} : \mathrm{Q}(K) \to \mathbb{Z}_2$, *defined by*

$$(-1)^{\mathrm{Arf}(q)} = \mathrm{sgn} \sum_{u \in K} (-1)^{q(u)}.$$

Then exactly one of the following holds:

1. $\mathrm{Arf}(q) = 0$ *and* $q \sim (q_{\mathbb{H}}^0)^{\oplus m}$,
2. $\mathrm{Arf}(q) = 1$ *and* $q \sim q_{\mathbb{H}}^1 \oplus (q_{\mathbb{H}}^0)^{\oplus(m-1)}$.

Proposition II.5. *Let $d = 2$ and $q \in \tilde{\mathrm{Q}}(K)$ be a generalized quadratic refinement of the non-degenerate form β. Consider the* generalized Arf invariant *of q,* $\tilde{\mathrm{Arf}} : \tilde{\mathrm{Q}}(K) \to \mathbb{Z}_8$ *defined through*

$$\exp\left(\frac{2i\pi}{8}\tilde{\mathrm{Arf}}(q)\right) = \text{phase of } \sum_{u \in K} i^{q(u)}.$$

Then, if β is of even type, $\dim K = 2m$, *and exactly one of the following holds,*

1. $\tilde{\mathrm{Arf}}(q) = 0$ *and* $q \sim (2q_{\mathbb{H}}^0)^m$,
2. $\tilde{\mathrm{Arf}}(q) = 4$ *and* $q \sim (2q_{\mathbb{H}}^1) \oplus (2q_{\mathbb{H}}^0)^{\oplus(m-1)}$.

Otherwise, β is of odd type and exactly one of the following holds,

1. $\tilde{\mathrm{Arf}}(q) = k \mod 8$ *and* $q \sim q_{k,0}$,
2. $\tilde{\mathrm{Arf}}(q) = k - 2 \mod 8$ *and* $q \sim q_{k-1,1}$,
3. $\tilde{\mathrm{Arf}}(q) = k - 4 \mod 8$ *and* $q \sim q_{k-2,2}$,
4. $\tilde{\mathrm{Arf}}(q) = k - 6 \mod 8$ *and* $q \sim q_{k-3,3}$,

where $k := \dim K$.

2. *Duality between quadratic and symmetric forms*

We now explore the relationship between quadratic forms and symmetric forms in some more detail. To the best of our knowledge the results in this section are not explicitly proven in the literature.

Evidently all quadratic forms are a refinement of some bilinear form. If $d = 2$, however, the converse fails: only symmetric forms of even type admit a quadratic refinement. Indeed, if β and q are as in (5) we see that

$$\beta(u, u) = q(u) + q(u) + q(u + u) = 0.$$

Let $\Xi : \mathrm{Q}(K) \to \mathrm{Sym}(K)$ be the linear map sending each q to β such that (5) holds. We call $\Xi(q)$ the *polarisation* of q. Eq. (7) tells us that when d is odd, Ξ is invertible. On the other hand, if $d = 2$ then $K^* \subset \mathrm{Q}(K)$ is the kernel of Ξ. The range of Ξ in this case is no longer $\mathrm{Sym}(K)$ but actually $\mathrm{range}\,\Xi = \mathrm{Alt}(K)$ as shown in App. A 1. This means that a symmetric bilinear form over $\mathbb{Z}_2$ has a quadratic refinement if and only if it is alternating.

We may similarly define a linear map $\tilde{\Xi} : \tilde{\mathrm{Q}}(K) \to \mathrm{Sym}(K)$ which polarizes generalized quadratic forms, ie. for which $\beta = \tilde{\Xi}(q)$ satisfies eq. (6). It can be seen that

$$\ker \tilde{\Xi} = 2K^* := \{2f \mid f \in K^*\}.$$

In App. A 1 we show that $\mathrm{range}\,\tilde{\Xi} = \mathrm{Sym}(K)$, that is, that *every* symmetric bilinear form has a *generalized* quadratic refinement.

In the case of odd d, the map Ξ provided a canonical identification between $\mathrm{Q}(K)$ and $\mathrm{Sym}(K)$. In the following we proposition we construct a different map, Φ, which provides a similar identification but works for the case $d = 2$ as well.

Proposition II.6. *Recall that one can canonically embed* $\mathrm{Sym}(K^*) = \langle v \otimes v \rangle \subseteq K \otimes K$. *Consider the linear function* $\Phi : \mathrm{Sym}(K^*) \to \mathrm{Q}(K)^*$ *given by*

$$\Phi(v \otimes v)(q) = q(v).$$

Then, Φ *is well defined, bijective, and* $\mathrm{Gl}(K)$*-covariant ie.,*

$$\Phi(gv \otimes gv)(q) = \Phi(v \otimes v)(g^{-1}q).$$

Proof. To prove consistency, first consider Φ evaluated on the set of tensor squares $v \otimes v$, we can see that

$$\begin{aligned}\Phi(v \otimes u + u \otimes v)(q) &= \Phi\big((v+u) \otimes (v+u) - v \otimes v - u \otimes u\big)(q) \\ &= q(u+v) - q(u) - q(v) = \Xi(q)(u,v)\end{aligned}$$

is a symmetric bilinear form. Similarly $q(av) = \Phi(av \otimes av)(q) = a^2\Phi(v \otimes v)(q) = a^2 q(v)$. On linear combinations of tensor squares, moreover, if

$$B := \sum_i a_i v_i \otimes v_i = 0,$$

then by linearity of Φ

$$\Phi(B)(q) = \sum_i a_i q(v_i) = \sum_i a_i \mathrm{tr}\left((v_i \otimes v_i^T) M_q\right) = \sum_i a_i \mathrm{tr}\left((v_i \otimes v_i) M_q^{\Gamma_2}\right) = 0,$$

where M_q is any matrix representation of q and Γ_2 is a partial transpose acting on the second tensor factor.

The function is injective: consider a $B \in \mathrm{Sym}(K)^*$ for which

$$\Phi(B)(q) = 0, \qquad \forall\, q \in \mathrm{Q}(K).$$

Then, if $B = \sum_i a_i v_i \otimes v_i$,

$$0 = \sum_i \mathrm{tr}\left(a_i (v_i \otimes v_i) M_q^{\Gamma_2}\right), \qquad \forall\, q \in \mathrm{Q}(K). \tag{11}$$

Now, any symmetric matrix M represents some quadratic form q_M through $q_M(v) = v^T M v$. This way, eq. (11) implies that B is orthogonal to all symmetric tensors in $K^* \otimes K^*$ and hence equal to zero. The function is also surjective because $\dim \mathrm{Q}(K)^* = \dim \mathrm{Sym}(K^*)$.

Finally, Φ is $\mathrm{Gl}(K)$-covariant since

$$\Phi(gv \otimes gv)(q) = q(gv) = (g^{-1}q)(v) = \Phi(v \otimes v)(g^{-1}q).$$

□

Example II.3. *Consider the symmetric form* κ_h *on* K^* *given by* $\kappa_x(u,v) = u(x)v(x)$*, where* $u, v \in K^*$ *and* $x \in K$*. Then* $\kappa_x = x \otimes x \in K \otimes K$ *and* $\Phi(\kappa_x) : q \mapsto q(x)$*. Similarly, consider a symmetric form*

$$\kappa(u,v) = \sum_i s_i u(x_i) v(x_i)$$

where $x_i \in K$ *and* $s_i \in \mathbb{Z}_d$*. Then,*

$$\Phi(\kappa) : q \mapsto \sum_i s_i q(x_i). \tag{12}$$

3. *Model quadratic space*

Throughout the remainder of the paper we take $T := \mathbb{Z}_d^{r+s}$, with $r + s = t$ and let $\beta_{r,s} : T \times T \to \mathbb{Z}_d$ be the symmetric form defined on the standard basis by

$$\beta_{r,s}(e_i, e_j) = s_i \delta_{ij},$$

where $s_i = 1$ if $i \leq r$, and $s_i = -1$ if $i > r$. In the case $d = 2$, this form is the standard dot product, however we keep the notation above for uniformity. We will denote the matrix representation of $\beta_{r,s}$ with respect to the diagonalising basis $\{e_i\}$ by $M_{r,s}$.

Let $q_{r,s} : T \to \mathbb{Z}_D$ be the (generalized) quadratic refinement of $\beta_{r,s}$ evaluating on the standard basis by

$$q_{r,s}(e_i) = \begin{cases} +1 \mod D, & i \leq r, \\ -1 \mod D, & i > r. \end{cases}$$

A subspace $N \subset T$ is *isotropic* if $q_{r,s}|_N = 0$. It is *stochastic* if $\mathbf{1}_t \in N^\perp$, where

$$N^\perp := \{u \in T \,|\, \beta(u, v) = 0, \forall\, v \in N\}$$

is the *orthocomplement* of N. Notice that for an isotropic subspace N, it holds that $N \subseteq N^\perp$. We use the following notation, $\mathbf{1}_t := (1, 1, \ldots, 1) \in T$,

$$\begin{aligned} \mathcal{G}_m &:= \{N \subset T \,|\, N \text{ stoch. isotr. }, \mathbf{1}_t \notin N, \dim N = m\} \\ \mathcal{G} &:= \bigcup_m \mathcal{G}_m, \\ \mathcal{G}^0_m &:= \{N \subset T \,|\, N \text{ stoch. isotr. }, \mathbf{1}_t \in N, \dim N = m\} \\ \mathcal{G}^0 &:= \bigcup_m \mathcal{G}^0_m, \\ T_N &:= N^\perp / N, \end{aligned}$$

The maximal m for which $\mathcal{G}_m$ is non-empty is denoted $m(T)$, the largest m for which $\mathcal{G}^0_m$ is non-empty is $m(T) - 1$.

Because

$$\operatorname{rad}\,(q_{r,s}|_{N^\perp}) = N,$$

the form $q_{r,s}$ is well defined on T_N. We write $q_N \in \tilde{\mathrm{Q}}(T_N)$ for the non-degenerate form inherited.

The orthogonal group, $\mathrm{O}(T)$, is the subgroup of $\mathrm{Gl}(T)$ leaving $q_{r,s}$ invariant,

$$q(O\,\cdot) = q(\,\cdot\,) \mod D, \qquad O \in \mathrm{O}(T).$$

The orthogonal stochastic group, $\mathrm{St}(T)$, is the subgroup of $\mathrm{O}(T)$ leaving $\mathbf{1}_t$ invariant,

$$O\mathbf{1}_t = \mathbf{1}_t \mod d, \qquad O \in \mathrm{St}(T).$$

Proposition II.7. *Let $d = 2$. Then $\mathrm{O}(T) = \mathrm{St}(T)$.*

Proof. Because $q_{r,s}$ is a generalized refinement of $\beta_{r,s}$, the following equation over $\mathbb{Z}_4$ holds for any $O \in \mathrm{O}(T)$,

$$2\beta_{r,s}(Ou, Ov) = q_{r,s}(Ou + Ov) - q_{r,s}(Ou) - q_{r,s}(Ov) = q_{r,s}(u + v) - q_{r,s}(u) - q_{r,s}(v) = 2\beta_{r,s}(u, v),$$

so that $\beta_{r,s}$ is $\mathrm{O}(T)$-invariant. But, $\beta_{r,s}(u, u) = \beta_{r,s}(\mathbf{1}_t, u)$, and so

$$\beta_{r,s}(O^{-1}\mathbf{1}_t, u) = \beta_{r,s}(Ou, Ou) = \beta_{r,s}(u, u) = \beta_{r,s}(\mathbf{1}_t, u).$$

The claim follows from non-degeneracy of $\beta_{r,s}$. □

Model spaces of the same dimension are sometimes isometric, the following proposition classifies all such isometries.

Proposition II.8. *Let $r, r', s, s' \in \mathbb{N}$ with $r + s = r' + s'$. Then $q_{r,s} \sim q_{r',s'}$ if and only if one of the following conditions holds,*

1. $d = 1 \mod 4$,
2. $d = 3 \mod 4$, $s = s' \mod 2$,
3. $d = 2$, $r - s = r' - s' \mod 8$.

Proof. Point 2. follows from $\mathrm{dis}(\beta_{r,s}) =$ sq. class of $(-1)^s$. Point 1. uses this identity together with the fact that -1 is a square over $\mathbb{Z}_d$ in this case.

For Point 3. we begin by noticing that both quadratic forms are of odd type. Therefore they are equivalent if and only if their generalized Arf invariant is equal.

$$\begin{aligned}\exp\left(\frac{2i\pi}{8}\tilde{\mathrm{Arf}}(q_{r,s})\right) &= \text{ phase of } \sum_{u_1\in\mathbb{Z}_d^r}\sum_{u_2\in\mathbb{Z}_d^s} i^{q_{r,0}(u_1)-q_{s,0}(u_2)} \\ &= \exp\left(\frac{2i\pi}{8}\left(\tilde{\mathrm{Arf}}(q_{r,0})-\tilde{\mathrm{Arf}}(q_{s,0})\right)\right) \\ &= \exp\left(\frac{2i\pi}{8}(r-s)\right), \end{aligned} \tag{13}$$

where the last line follows from Prop. II.5. □

It is convenient to define $q_{r,s}$ for possibly negative $r, s \in \mathbb{Z}$. Namely, we let $q_{r,s}$ be any generalized quadratic form equivalent to some $q_{r',s'}$ with $r', s' \geq 0$ such that one of points 1.-3. of Prop. II.8 holds.

Proposition II.9. *Let $N \in \mathcal{G}_m$, then $q_N \simeq q_{r-m,s-m}$.*

Proposition II.10. *Let $r - s = 0 \mod 4$, $d = 2$, and $N \in \mathcal{G}_m^0$. Then, $t = 0 \mod 2$, and it holds that*

$$\beta_N \simeq \beta_{\mathbb{H}}^{\oplus(t-2m)/2}.$$

Proof. The first statement follows from $r = s \mod 2$. For any $u \in T_N = N^\perp/N$, let $u_0 \in u \subset N^\perp$ be some point in the corresponding affine plane. The second claim follows from the fact that, for all $u \in T_N$,

$$\beta_N(u,u) = \beta_{r,s}(\mathbf{1}_t, u_0) = 0$$

because $\mathbf{1}_t \in N$. Then, the inherited form has even type. □

Prop. II.10 motivates us to define $\mathrm{Sp}(T_N)$, the isometry group of β_N for $N \in \mathcal{G}_m^0$. It is clear that $\mathrm{Sp}(T_N) \simeq \mathrm{Sp}(2, t-2m)$, and that $\mathrm{St}(T_N) \subseteq \mathrm{Sp}(T_N)$ where this inclusion is strict as long as $t > 2m$. Finally, notice that because $\mathbf{1}_t \in N$, $\mathrm{St}(T_N) = \mathrm{O}(T_N)$ is the isometry group of q_N.

C. Representation Theory

A representation of a group G is a matrix-valued function $\rho : G \to \mathbb{F}^{k\times k}$ for which

$$\rho(g_1)\rho(g_2) = \rho(g_1 g_2),$$

where $\mathbb{F}$ is an arbitrary number field. The space of functions $f : \mathbb{F}^k \to S$, where S is a set, is canonically a G-representation through

$$gf(\,\cdot\,) := f(\rho(g^{-1})\,\cdot\,). \tag{14}$$

Throughout we will assume that any space of functions on a representation space is itself a representation arising in this way.

The representations with which we deal with here are usually complex, $\mathbb{F} = \mathbb{C}$, and unitary, $\mathrm{range}\,\rho \subset \mathrm{U}(\mathbb{C}^k)$. The *representation space* of ρ is the space on which it acts, namely $\mathbb{C}^k$. A subspace $\mathcal{V} \subseteq \mathbb{C}^k$ is called a *G-invariant space*, or simply an invariant space, if $\rho(G)\mathcal{V} = \mathcal{V}$. A representation with no non-trivial invariant subspaces is *irreducible*, the set of all such representations is $\mathrm{Irr}\,G$. If ρ is *reducible*, there is a $U \in \mathrm{U}(\mathbb{C}^k)$ for which

$$U\rho(g)U^\dagger = \rho_1(g) \oplus \cdots \oplus \rho_m(g), \quad \forall\, g \in G,$$

where ρ_i are irreducible representations. Equivalently, we may decompose the representation space into invariant subspaces

$$\mathbb{C}^k = \bigoplus_i \mathcal{V}_i,$$

where $\rho(g)\mathcal{V}_i = \rho_i(g)\mathcal{V}_i$. We summarize this into the equation $\rho \simeq \oplus_i \rho_i$, where $\simeq$ denotes isomorphism between representations.

Let $H \trianglelefteq G$ and $G' = G/H$. Any representation ρ of G' can be extended to a representation of G which has H in its kernel. For simplicity, we will also call this representation ρ. It follows that there is a canonical embedding $\mathrm{Irr}\, G' \subseteq \mathrm{Irr}\, G$.

The *regular representation* of G is the space $\mathbb{C}[G]$ of formal linear combinations of group elements, with G acting from the left as

$$g\,|g'\rangle = |gg'\rangle\,.$$

The regular representation decomposes as

$$\mathbb{C}[G] \simeq \bigoplus_{\tau \in \mathrm{Irr}\, G} \tau \otimes \mathbb{C}^{\dim \tau},$$

where the right factors are multiplicity spaces. The commutant of the regular representation is spanned by the right-action of G,

$$g : |g'\rangle \mapsto \left|g' g^{-1}\right\rangle.$$

Proposition II.11. *Let G be a finite group acting on a finite set S with $|S| = |G|$. Furthermore, assume that for some $s \in S$, it holds that*

$$gs = s \implies g = 1.$$

Then, the space $\mathbb{C}[S]$ equipped with the G action

$$g\,|s\rangle = |gs\rangle$$

is isomorphic to the regular representation $\mathbb{C}[G]$.

Proof. The isomorphism is given explicitly by $|gs\rangle \mapsto |g\rangle$. □

III. CLIFFORD TENSOR POWERS

In this paper we study the tensor-power representations of Cl,

$$\Delta_{r,s}(U) = U^{\otimes r} \otimes \bar{U}^{s} =: U^{\otimes (r,s)}, \quad U \in \mathrm{Cl}.$$

The representation space corresponding to $\Delta_{r,s}$ is $\mathcal{H}_{n,t} := \mathcal{H}_n^{\otimes t}$ where $t := r + s$. It is useful to think of $\mathcal{H}_{n,t}$ as an n by t "grid" of qudit Hilbert spaces,

$$\mathcal{H}_{n,t} = \begin{array}{ccccc} \mathbb{C}^d & \otimes & \cdots & \otimes & \mathbb{C}^d \\ \vdots & & \ddots & & \vdots \\ \mathbb{C}^d & \otimes & \cdots & \otimes & \mathbb{C}^d. \end{array}$$

We label the computational basis of $\mathcal{H}_{n,t}$ using $t \times n$ matrices F over $\mathbb{Z}_d$, that is $|F\rangle$. In the grid picture, we use

$$|F\rangle = \begin{array}{ccccc} |F_{11}\rangle & \otimes & \cdots & \otimes & |F_{n1}\rangle \\ \vdots & & \ddots & & \vdots \\ |F_{1t}\rangle & \otimes & \cdots & \otimes & |F_{nt}\rangle \end{array}, \qquad F_{ij} \in \mathbb{Z}_d.$$

We furthermore identify columns of F with vectors in T.

Th Hilbert space $\mathcal{H}_{n,t}$ hosts its own "global" Pauli group $\mathcal{P}_{nt}$, generated by $\tau\mathbb{1}$ and tensor products of displacement operators acting on any $\mathbb{C}^d$ factor. We can see that the phase space associated to this Pauli group is

$$\mathcal{P}_{nt}/\mathcal{Z}(\mathcal{P}_{nt}) \simeq \mathbb{Z}_d^{2n\times t} \simeq T \otimes V \simeq \mathrm{Hom}(V \to T).$$

Notice that $\Delta_{1,1}$ is isomorphic to the action of Cl on $\mathrm{End}(\mathcal{H}_n)$ by conjugation. This space decomposes into two irreducible subrepresentations: the trivial component spanned by the identity and the space of traceless matrices. The latter irrep will be referred to as the *adjoint* representation of Cl, denoted $\mathrm{Ad}(\mathrm{Cl})$. It can be seen to be the restriction to Cl of the adjoint representation of $\mathrm{U}(\mathcal{H}_n)$.

The following lemma establishes equivalences between different tensor power representations, we prove it in App. A 2.[3]

[3] Compare the condition $r - s = r' - s' \mod d$ to [15, Lem. 4.2]. In the notation of that reference, consider two orthogonal bases $\{e_i\}$ and $\{f_i\}$ of the space U, satisfying $[\beta(e_i, e_j)]_{ij} = \mathbb{1}_r \oplus (-\mathbb{1}_s)$, $[\beta(f_i, f_j)]_{ij} = \mathbb{1}_{r'} \oplus (-\mathbb{1}_{s'})$. Then our $\Delta_{r,s}$ corresponds to $\mathcal{C}l^{(e)}_{U\otimes V}$ from the reference, and $\Delta_{r',s'}$ to $\mathcal{C}l^{(f)}_{U\otimes V}$, where $e = \sum_i e_i$ and $f = \sum_i f_i$. These two representations of the Clifford group are non-equivalent whenever $\beta(e, e) \neq \beta(f, f)$. To see this, one must simply compute the action of the central matrices of these representations, denoted in the proof of [15, Lem. 4.2] as $W_{U\otimes V}(\iota_e(\lambda, 0))$ and $W_{U\otimes V}(\iota_f(\lambda, 0))$.

Lemma III.1 (Equivalent tensor powers)**.** *Let $d, r, s, r', s' \in \mathbb{N}$ be such that $r + s = r' + s'$, and let $\Delta_{r,s}$ be as above. Furthermore, if d is odd, let $r - s = r' - s' \mod d$. Then for all the following cases we have that $\Delta_{r,s} \simeq \Delta_{r',s'}$:*

1. *If $d = 1 \mod 4$*
2. *If $d = 3 \mod 4$, and $s = s' \mod 2$,*
3. *If $d = 2$, $r - s = r' - s' \mod 8$.*

As we did with $q_{r,s}$, we will define $\Delta_{r,s}$ to possibly negative values of r, s. Namely, take some $r', s' \in \mathbb{N}$. Then if d is odd, for any $r, s \in \mathbb{Z}$ with

$$r + s = r' + s', \qquad r - s = r' - s' \mod d, \qquad s = s' \mod 2,$$

we let $\Delta_{r,s}$ be a representaiton equivalent to $\Delta_{r',s'}$. On the other hand, if $d = 2$, then for any $r, s \in \mathbb{Z}$ with

$$r + s = r' + s', \qquad r - s = r' - s' \mod 8,$$

we let $\Delta_{r,s}$ be equivalent to $\Delta_{r',s'}$.

A. The commutant algebra

In [7] the *commutant* $\mathrm{A}_{r,s}$ of $\Delta_{r,s}$ was studied, ie. the subalgebra of $\mathrm{End}(\mathcal{H}_{n,t})$ which commutes with all images $\Delta_{r,s}(C)$, where $C \in \mathrm{Cl}$. Important to this construction where two ingredients.

The first is a class of tensor-power CSS code projectors associated to an isotropic subspace $N \subset U$

$$P_N = d^{-2n\dim N} \sum_{M \in \mathrm{Hom}(V \to N)} W(M). \tag{15}$$

Letting $C_N := \mathrm{range}\, P_N$, these code spaces have the following *coset state basis*,

$$|[F]_N\rangle := d^{-\dim N/2} \sum_{F' \in \mathrm{Hom}(\mathrm{X} \to N)} |F + F'\rangle\,, \qquad F \in \mathrm{Hom}(\mathrm{X} \to N^\perp).$$

If the columns of F are $f_1, \ldots, f_n$, we can see that

$$|[F]_N\rangle = |[f_1]_N\rangle \otimes \cdots \otimes |[f_n]_N\rangle\,, \tag{16}$$

where,

$$|[f_i]_N\rangle = d^{-\dim N/2} \sum_{u \in N} |f_i + u.\rangle$$

Notice that $[F]_N \in \mathrm{Hom}(\mathrm{X} \to T_N)$, so that we can directly identify $C_N = \mathbb{C}[\mathrm{Hom}(\mathrm{X} \to T_N)]$.

The second is the following representation of $O_{r,s}$,

$$R(O) = \sum_{F \in \mathrm{Hom}(\mathrm{X} \to T)} |OF\rangle\langle F|\,. \tag{17}$$

We will use the following slight generalization of the main result of [7], which dealt with the representation $\Delta_{t,0}$.

Proposition III.1. *Let P_N be as in* (15)*, where N is isotropic, and R be as in* (17)*.*
Then, if $t \leq n - 1$, the commutant $\mathrm{A}_{r,s}$ of $\Delta_{r,s}$ has the following basis

$$\mathcal{A}_{r,s} = \{R(O)P_N \,|\, O \in \mathrm{St}(T),\ N \textit{ isotropic stochastic}\}.$$

Furthermore, if $d =$ odd, the commutant of $\mathrm{Res}_{Sp(V)}\, \Delta_{r,s}$ is generated as an algebra by

$$\{R(O), P_N \,|\, O \in \mathrm{O}(T), N \textit{ isotropic}\}.$$

Proof sketch. One can show (e.g. as in [7, Lem. 4.5]) that the operators $R(O)$ and P_N commute with $\Delta_{r,s}$. Following the same procedure as in [7, Lem. 4.7] one can show that $\mathcal{A}_{r,s}$ is a set of linearly independent operators (because $t-1 \leq n$). Finally, by the fact that the dimension of the commutant of $\Delta_{r,s}$ is the same as the dimension of the commutant of $\Delta_{t,0}$, the proof of [7, Thm. 4.9] shows our first claim. The second claim follows similarly. □

It can be easily checked that

$$R(O)P_N R(O)^\dagger = P_{ON}. \tag{18}$$

Conversely, if $\dim N' = \dim N$, these two subspaces are isometric and so there exists some $O \in \mathrm{O}(T)$ for which $N' = ON$. Further, if N is stochastic and $O \in \mathrm{St}(T)$, ON is also stochastic. We now show the converse.

Lemma III.2. *Let $N,\ N' \subset T$ be stochastic isotropic subspaces of the same dimension. Furthermore, let it be the case that either $\mathbf{1}_t \in N \cap N'$ or $\mathbf{1}_t \notin N \cup N'$. Then, there exists an $O \in \mathrm{St}(T)$ for which $N' = ON$.*

Proof. Consider the two isometric spaces $M = \mathrm{span}\,\{N, \mathbf{1}_t\}$, $M' = \mathrm{span}\,\{N', \mathbf{1}_t\}$. Let $\{e_i\}$ be a basis of M and $\{e'_i\}$ be a basis of M' with $e_1 = e'_1 = \mathbf{1}_t$. Then, due to a lemma by Witt (cf. for example [32, Thm. 3.3]), there is an $O \in \mathrm{O}(T)$ for which $Oe_i = e'_i$. □

Lem. III.2 motivates the definition of the following subsets of $\mathcal{A}_{r,s}$,

$$\begin{aligned}
\mathcal{A}^m_{r,s} &:= \{R(O)P_N \,|\, O \in \mathrm{St}(T),\ N \text{ stoch. isotr.}, \dim N \geq m\},\\
\mathcal{A}^{m,0}_{r,s} &:= \{R(O)P_N \,|\, O \in \mathrm{St}(T),\ N \text{ stoch. isotr.}, \dim N \geq m,\ \mathbf{1}_t \in N\},\\
\mathcal{A}^N_{r,s} &:= \{R(O)P_{N'} \,|\, O \in \mathrm{St}(T),\ N \subseteq N' \text{ stoch. isotr.},\}.
\end{aligned}$$

Each of these subsets is invariant under right and left multiplication with $\mathrm{St}(T_N)$. Furthermore, by [7, eq. (4.24)], $\mathcal{A}_{r,s}$ forms a semigroup and the subalgebras $\mathrm{A}^m_{r,s} := \mathrm{span}\,\{\mathcal{A}^m_{r,s}\}$ and $\mathrm{A}^{m,0}_{r,s} := \mathrm{span}\,\mathcal{A}^{m,0}_{r,s}$ are ideals of $\mathrm{A}_{r,s}$. Namely, for any $N_1,\ N_2$, there exists some O_{N_1,N_2} for which

$$P_{N_1}P_{N_2} = R(O_{N_1,N_2})P_N, \qquad N := \langle N_1 \cap N_2^\perp,\ N_2\rangle. \tag{19}$$

Here, by [7, eq. (4.25)], $\dim N \geq \max\{\dim N_1,\ \dim N_2\}$.

Lemma III.3. *Let N be an isotropic stochastic subspace of T. Then the code C_N is a* Cl *subrepresentation of $\Delta_{r,s}$ and a basis for the commutant of this action is contained in $\mathcal{A}^N_{r,s}$.*

Proof. The commutant of Cl in $\mathrm{End}(C_N)$ is given by $P_N \mathrm{A}_{r,s} P_N$. Using (19), for any $R(O)P_M \in \mathcal{A}_{r,s}$,

$$\begin{aligned}
P_N R^\dagger(O) P_M P_N &= R^\dagger(O) P_{ON} P_M P_N\\
&\propto R^\dagger(O) R(O_{ON,M}) P_{N'} P_N\\
&\propto R(O^{-1}O_{ON,M}O_{N',N})P_{N''},
\end{aligned}$$

where $N' := \langle ON^\perp \cap M, M\rangle$ and

$$N'' = \langle N' \cap N^\perp, N\rangle \supset N.$$

This implies that there is some $A \in \mathcal{A}^N_{r,s}$ for which $A \propto P_N R^\dagger(O) P_M P_N$. □

B. Code representation spaces

By Prop. III.1, the code space C_N corresponding to a stochastic isotropic N is a Cl representation. Furthermore, by Lem. III.2, for any pair $N,\ N' \in \mathcal{G}_m$ or $N,\ N' \in \mathcal{G}^0_m$, the corresponding code spaces C_N and $C_{N'}$ are isomorphic as Cl representations.

The following lemmas specify this representation and can be seen as a generalization of [15, Lem. 2.7]. We prove them in App. A 2.

Lemma III.4 (Code representations)**.** *Let $N \in \mathcal{G}_m$ and $C_N \subset \mathcal{H}_{n,t}$ be the associated code. Then, $\Delta_{r,s}|_{C_N} \simeq \Delta_{r-m,s-m}$.*

Lemma III.5 (Qudit $C_{\mathbf{1}_t}$ representation)**.** *Let d be odd and $r - s = 0 \mod d$. As a* Cl*-subrepresentation of $\Delta_{r,s}$, we have that* $\ker(C_{\mathbf{1}_t}) = \mathcal{P}$ *and*

$$C_{\mathbf{1}_t} \simeq \begin{cases} \mu^{\otimes(r-1,s-1)}, & s > 0, \\ \mu^{\otimes(r-3,1)}, & s = 0. \end{cases}$$

Remark III.1. *The proof of Lem. III.4 does* not *hold when $N \in \mathcal{G}_m^0$. To see this most easily, notice that in this case $\mathcal{P} \subseteq \ker(C_N)$, whereas $\mathcal{P} \not\subseteq \ker(\Delta_{r-m,s-m})$. By Lem. III.5, when d is odd this isomorphism nevertheless continues to hold if one restricts attention to the action of* $\mathrm{Sp}(V) \subset \mathrm{Cl}$*.*

While Lems. III.4, III.5 were crucial for the proof of our previous result in [15], they are not needed for the main result in this paper. Because of this, we have also left open the question of whether there exists some generalization of Lem. III.5 to the qubit case.

Now we work out the action of stochastic orthogonal matrices on the code spaces.

Lemma III.6. *Let* $\mathrm{St}(T)^N$ *be the subgroup of* $\mathrm{St}(T)$ *which preserves the stochastic isotropic subspace N. Then,* $\mathrm{St}(T)^N$ *acts on C_N, and* $\mathrm{St}(T)^N / \ker(C_N) \simeq \mathrm{St}(T_N)$.

Proof. It suffices to consider the $n = 1$ case. Because $O \in \mathrm{St}(T)^N$ acts on N-cosets, there exists some $\tilde{O} \in \mathrm{Gl}(T_N)$ for which

$$R(O)\,|[f]_N\rangle = |[Of]_N\rangle = \left|\tilde{O}[f]_N\right\rangle, \qquad f \in N^\perp.$$

Moreover, $\tilde{O}$ preserves q_N and $[\mathbf{1}_t]_N$, so $\tilde{O} \in \mathrm{St}(T_N)$. Finally, the Cahit-Arf theorem [33] implies that every $\tilde{O} \in \mathrm{St}(T_N)$ can be extended to an $O \in \mathrm{St}(T)^N$ such that $[Of]_N = \tilde{O}[f]_N$ for all $f \in N^\perp$. □

IV. RANK THEORY OF CLIFFORD REPRESENTATIONS

Recently, several notions of rank have been introduced to study the representation theory of discrete symplectic and orthogonal groups [16, 17]. Here we extend a part of this formalism to the Clifford group.

Let $\mathrm{D} \subset \mathrm{Cl}$ be the subgroup of diagonal Clifford matrices.

Lemma IV.1. *Any $U \in \mathrm{D}$ is, up to a phase, of the form*

$$U = \sum_{x \in \mathrm{X}} \tau^{q(x) + 2x' \cdot x}\, |x\rangle\langle x| \tag{20}$$

for some $q \in \tilde{\mathrm{Q}}(\mathrm{X})$ and $x' \in \mathrm{X}$. If $d = 2$ one can furthermore set $x' = 0$.

Proof. Consider the homomorphism $\varphi : \mathrm{Cl} \to \mathrm{Cl}/\mathcal{P} \simeq \mathrm{Sp}(V)$. Diagonal Cliffords preserve all Z-type Pauli matrices so that

$$\varphi(\mathrm{D}) = \left\{ \begin{pmatrix} \mathbb{1} & S \\ 0 & \mathbb{1} \end{pmatrix} \;\middle|\; S \in \mathrm{Sym}(\mathrm{X}) \right\} =: \mathcal{N}, \tag{21}$$

where we have used the basis $\{e_1, \ldots, e_n, f_1, \ldots, f_n\}$ for which $W(e_i) = Z_i$. We may further specify

$$\varphi(\mathrm{P}_i) = \mathbb{1} + e_i e_i^T$$
$$\varphi(\mathrm{CPHASE}_{ij}) = \mathbb{1} + e_i e_j^T + e_j e_i^T,$$

where $\mathrm{CPHASE}_{ij} := \mathrm{H}_i \mathrm{H}_j \mathrm{CADD}_{ij} \mathrm{H}_i^\dagger \mathrm{H}_j^\dagger$. Thus, $\langle \varphi(\mathrm{P}_i), \varphi(\mathrm{CPHASE}_{ij}) \rangle_{ij} = \mathcal{N}$, and by (21),

$$\langle \mathrm{P}_i, \mathrm{CPHASE}_{ij}, Z_i \rangle_{ij} = \mathrm{D}. \tag{22}$$

Since each of these generators is of the form (20), the first claim follows. The second claim follows from the fact that $2\mathrm{X}^* \subset \tilde{\mathrm{Q}}(\mathrm{X})$ when $d = 2$. □

Consider the subgroup $\mathrm{RD} \subset \mathrm{D}$ of elements U, as in Lem. IV.1, for which $x' = 0$ and $q \in 2\mathrm{Q}(\mathrm{X})$. By the proof of that lemma, $\mathrm{RD} \simeq \mathrm{Q}(\mathrm{X})$ and so we use quadratic forms to label group elements,

$$\mathrm{RD} = \{U_q \,|\, q \in \mathrm{Q}(\mathrm{X})\}.$$

This group will be used to define a notion of *rank* on representations of the Clifford group, in close analogy to how the subgroup $\mathcal{N} \subset \mathrm{Sp}(V)$ was used to define the rank of a symplectic representation in [16, 17].[4]

Lemma IV.2. *Let ρ be a representation of* Cl*. Then* $\mathrm{Res}_{\mathrm{RD}}\, \rho$ *decomposes into one-dimensional representations $|\psi_B\rangle$ labeled by $B \in \mathrm{Sym}(\mathrm{X}^*)$ such that*

$$\rho(U_q)\,|\psi_B\rangle = \omega^{\Phi(B)(q)}\,|\psi_B\rangle\,,$$

where $U_q \in \mathrm{RD}$ and Φ is as in Prop. II.6.

Proof. The first claim follows from the fact that D is Abelian. Thus, an irrep in this decomposition is of the form

$$U_q \mapsto \omega^{p(q)}, \qquad p \in \mathrm{Q}(X)^*.$$

Finally by Prop. II.6, we can take $p = \Phi(B)$ for some $B \in \mathrm{Sym}(\mathrm{X}^*)$. □

The eigenvectors $|\psi_B\rangle$ are called the *weight vectors* of ρ, and the symmetric matrices B their corresponding *weights.* The span of all weight vectors corresponding to the same weight B is called the *weight space* of B. The rank of a representation ρ of Cl is

$$\mathrm{rk}(\rho) = \max_{B \text{ wght. of } \rho} \mathrm{rank}\,(B). \tag{23}$$

By Prop. II.6, weight-spaces are permuted by the subgroup $\mathrm{Gl}(\mathrm{X}) \subset \mathrm{Cl}$ generated by the CADD gates. Because of this, weights in the same equivalence class – that is, with the same rank and type – appear with the same multiplicity.

The following lemma is analogous to [16, Lem. 1.3.1.] which states that the only irrep of the symplectic group with rank zero is the trivial one. We prove it in App. A 3.

Lemma IV.3 (Rank 0 irreps)**.** *If d is odd, the unique rank zero* Cl *irrep is the trivial one. If $d = 2$ and $n \geq 3$, a rank zero representation is one dimensional, ± 1 valued, and uniquely specified by its restriction to $\mathcal{Z}(\mathrm{Cl})$. Namely, if ρ, ρ' are rank zero repesentations with*

$$\rho(\omega_8 \mathbb{1}) = \rho'(\omega_8 \mathbb{1}),$$

then $\rho \simeq \rho'$.

We believe that with a closer consideration one could strengthen this result to show that the irrep is in fact trivial. In particular, we believe that choosing the phases of the generators of Cl carefully (cf. [9]) is sufficient for this. We leave this for future work.

Using (12) and Lem. IV.1, we can directly read out the weight data of the defining representation of Cl. Namely, the weight vectors are the computational basis, and the weight of $|x\rangle$ has matrix representation xx^T. In particular, $\mathrm{rk}(\text{def. rep of Cl}) = 1$.

More generally, the following statement holds.

Lemma IV.4. *The weight corresponding the the computational basis state $|F\rangle \in \mathcal{H}_{n,t}$ is*

$$w_F := \sum_{i=1}^{t} s_i f_i \otimes f_i,$$

where $s_i \in \mathbb{Z}_d$ satisfies $s_i = +1$ if $i \leq r$ and $s_i = -1$ otherwise, and where the rows of F are $f_1, \ldots, f_t \in \mathrm{X}$. Thus $\mathrm{rk}(w_F) = \mathrm{rank}\,\beta_{r,s}|_{\mathrm{range}\,(F)}$ and if $t \leq n$, then $\mathrm{rk}(\Delta_{r,s}) = t$.

Proof. We may directly compute

$$U_q^{\otimes(r,s)}\,|F\rangle = \omega^{\sum_i s_i q(f_i)}\,|F\rangle\,.$$

The first statement follows from Ex. II.3. The second statement from the fact that $w_F(\,\cdot\,,\,\cdot\,) = \beta_{r,s}(F\,\cdot\,, F\,\cdot\,)$ and so $\mathrm{rank}\, w_F = \mathrm{rank}\,\beta_{r,s}|_{\mathrm{range}\,F}$. □

[4] In [17] several notions of rank are introduced, the closest analogue to our current definition is that of "U-rank".

V. CLASSIFICATION OF SUBREPRESENTATIONS

A. Rank and duality

Here, we use the formalism developed in Sec. IV to obtain a series of dualities. Our proof techniques are an extension of those introduced in [16].

Consider some, possibly zero, stochastic isotropic subspace $N \subset T$ with $\dim N = m \geq 0$. For notational uniformity, we define $C_{N=0} = \mathcal{H}_{n,t}$. We will show that Howe duality holds on the subset of all maximal-rank irreps in C_N. This generalizes the central result of [16], which covers the case of odd d and $m = 0$.

Theorem V.1. *Let $N \subset T$ be a stochastic isotropic subspace of dimension $m \geq 0$ with $t - 2m \leq n$, and let $\Delta_{r,s}^{(k)}$ be the subrepresentation of $\Delta_{r,s}$ spanned by all* Cl *irreps with rank k. Then, there exists some injective function $\eta_N : \operatorname{Irr} \mathrm{St}(T_N) \to \operatorname{Irr} \mathrm{Cl}$ for which, as a $\mathrm{St}(T)^N \times \mathrm{Cl}$ representation,*

$$C_N \cap \Delta_{r,s}^{(t-2m)} \simeq \bigoplus_{\tau \in \operatorname{Irr} \mathrm{St}(T_N)} \tau \otimes \eta_N(\tau). \tag{24}$$

It furthermore holds that $\dim \tau \leq \dim \eta_N(\tau)$.

To show this, we require three intermediate results, whose proofs we postpone to App. A 4. Throughout this section it sometimes will be convenient to work with the identification of $C_N \simeq \mathbb{C}[\mathrm{Hom}(\mathrm{X} \to T_N)]$. This is done by identifying the coset $[F]_N \in \mathrm{Hom}(\mathrm{X} \to N^\perp)/\mathrm{Hom}(\mathrm{X} \to N)$ with the matrix $F_0 \in \mathrm{Hom}(\mathrm{X} \to T_N)$ satisfying $F_0 e_i = [f_i]_N$, where f_i are the columns of F. Furthermore, we denote by R the representation of $\mathrm{St}(T_N)$ on $\mathbb{C}[\mathrm{Hom}(\mathrm{X} \to T_N)]$ corresponding to the action of $\mathrm{St}(T)^N$ on C_N.

Lemma V.1. *Let N be a stochastic isotropic subspace of dimension m, $F \in \mathrm{Hom}(\mathrm{X} \to T_N)$ be surjective, and consider the following two subspaces of C_N,*

$$\begin{aligned}
\mathcal{H}_F &:= \operatorname{span}\{|J\rangle \mid J \in \mathrm{Hom}(\mathrm{X} \to T_N),\ q_N(Jx) = q_N(Fx)\ \forall\, x \in \mathrm{X},\ F^{-1}([\mathbf{1}_t]_N) = J^{-1}([\mathbf{1}_t]_N)\},\\
\mathcal{H}^F &:= \operatorname{span}\{|OF\rangle \mid O \in \mathrm{St}(T_N)\},
\end{aligned}$$

where $F^{-1}([\mathbf{1}_t]_N)$ is the $(n - t + 2m)$-dimensional preimage of $\mathbf{1}_t$ under F. Then $\mathcal{H}_F = \mathcal{H}^F$.

Lemma V.2. *Let C_N be as in Thm. V.1 and $\mathcal{H}^F$ as in Lem. V.1. Then $\mathcal{H}^F$ is the regular representation of $\mathrm{St}(T_N)$, that is,*

$$\mathcal{H}^F \simeq \bigoplus_{\tau \in \operatorname{Irr} \mathrm{St}(T_N)} \tau \otimes \mathbb{C}^{\dim \tau},$$

where the sum ranges over every irrep of $\mathrm{St}(T_N)$, and where right-hand side factors are multiplicity spaces.

For the last intermediate result, we define $q^F \in \tilde{\mathrm{Q}}(\mathrm{X})$ by $q^F(x) = q_N(Fx)$, where $\operatorname{range}(F) = T_N$.

Lemma V.3. *Let N and F be as in Lem. V.1. Let $G_F \subset \mathrm{Gl}(\mathrm{X}) \subset \mathrm{Cl}$ be given by*

$$G_F = \begin{cases} \{g \mid q^F(g^T\,\cdot) = q^F(\,\cdot),\ g^{-T} F^{-1}([\mathbf{1}_t]_N) = F^{-1}([\mathbf{1}_t]_N)\} & N \in \mathcal{G}_m, \\ \{g \mid q^F(g^T\,\cdot) = q^F(\,\cdot)\} & N \in \mathcal{G}_m^0. \end{cases}$$

Here, $g^{-T} F^{-1}([\mathbf{1}_t]_N) = F^{-1}([\mathbf{1}_t]_N)$ is an equality of sets. Then, the commutant of $R(\mathrm{St}(T_N))|_{\mathcal{H}^F}$ in $\mathrm{End}(\mathcal{H}^F)$ is spanned by $\Delta_{r,s}(G_F)|_{\mathcal{H}^F}$.

Proof of Thm. V.1. Because $t - 2m \leq n$, then a surjective $F \in \mathrm{Hom}(\mathrm{X} \to T_N)$ exists and so do the spaces $\mathcal{H}_F$ and $\mathcal{H}^F$ defined in Lem. V.1.

As a $\mathrm{St}(T_N) \times \mathrm{Cl}$ representation, we may decompose

$$C_N \simeq \bigoplus_{\tau \in \operatorname{Irr} \mathrm{St}(T_N)} \tau \otimes \Theta(\tau),$$

where $\Theta(\tau)$ is a (possibly reducible) representation of Cl. This is because the actions of these two groups commute. The sum in this decomposition ranges over all of $\operatorname{Irr} \mathrm{St}(T_N)$ because of Lem. V.2.

Now consider the subspace $C_N \cap \mathcal{H}_F$, where $F \in \mathrm{Hom}(\mathrm{X} \to T_N)$ is surjective. This subspace is a $\mathrm{St}(T_N) \times G_F$ representation, and so it decomposes as

$$C_N \cap \mathcal{H}_F \simeq \bigoplus_{\tau \in \mathrm{Irr}\,\mathrm{St}(T_N)} \tau \otimes \Theta^F(\tau),$$

where $\Theta^F(\tau)$ is a G_F-subrepresentation of $\Theta(\tau)$. By Lem. V.3, $\Theta^F(\tau)$ is irreducible. Acting with Cl on $\Theta^F(\tau) \subseteq \Theta(\tau)$ we obtain some irreducible representation $\eta_N(\tau)$. Because $t - 2m \leq n$, for every surjective $F' \in \mathrm{Hom}(\mathrm{X} \to T_N)$, there exists some $g \in \mathrm{Gl}(\mathrm{X}) \subset \mathrm{Cl}$ for which $Fg^T = F'$, so $\eta_N(\tau)$ is independent of the F used to construct $\Theta^F(\tau)$. Because of this, every surjective F' is such that $|F'\rangle \in \Delta_{r,s}(\mathrm{Cl})(C_N \cap \mathcal{H}^F)$ and therefore

$$\Delta_{r,s}^{(t-2m)} \cap C_N = \Delta_{r,s}(\mathrm{Cl})(C_N \cap \mathcal{H}^F) \simeq \bigoplus_{\tau \in \mathrm{Irr}\,\mathrm{St}(T_N)} \tau \otimes \eta_N(\tau).$$

By Lem. V.2, if $\tau \not\simeq \tau'$ then $\mathrm{Res}_{G_F}\, \eta_N(\tau) \not\simeq \mathrm{Res}_{G_F}\, \eta_N(\tau')$ and therefore $\eta_N(\tau) \not\simeq \eta_N(\tau')$. Finally, the same lemma implies $\dim \tau \leq \dim \Theta_F(\tau) \leq \dim \eta_N(\tau)$. □

Lemma V.4. *Let N and η_N be as in Thm. V.1. Consider some N' satisfying $\dim N' = \dim N$ and $\mathbf{1}_t \in N' \iff \mathbf{1}_t \in N$. Then, $\mathrm{St}(T)^N \simeq \mathrm{St}(T)^{N'}$ are conjugate in $\mathrm{St}(T)$ and, for any $\tau \in \mathrm{St}(T_N)$, $\eta_N(\tau) \simeq \eta_{N'}(\tau)$.*

Proof. Let O be such that $ON = N'$. Then a short calculation shows $O\mathrm{St}(T)^N O^{-1} = \mathrm{St}(T)^{N'}$. Let $P_{N,\tau}$ be the projector onto the $\tau \otimes \eta_N(\tau)$ component in C_N. Then the isomorphism is afforded by $R(O)$. This is because, by the first claim, $R(O)P_{N,\tau}R^\dagger(O) = P_{N',\tau}$ and so,

$$\begin{aligned} R(O)P_{N,\tau}\Delta_{r,s}(U)P_{N,\tau}R^\dagger(O) &= P_{N',\tau}\Delta_{r,s}(U)P_{N',\tau}, \quad \forall\, U \in \mathrm{Cl}, \\ R(O)P_{N,\tau}R(O')P_{N,\tau}R^\dagger(O) &= P_{N',\tau}R(OO'O^{-1})P_{N',\tau}, \quad \forall\, O' \in \mathrm{St}(T)^N. \end{aligned}$$

□

B. Proof of the main theorem

In Sec. V A we showed how to classify the maximal rank subrepresentations in an arbitrary CSS code C_N. Here we use this result to classify all subrepresentations of $\mathcal{H}_{n,t}$.

Theorem V.2. *Let $t \leq n$. Consider a subset $\{N_i\}_i \subset \mathcal{G} \cup \mathcal{G}^0$ such that every $\mathrm{St}(T)$ orbit on $\mathcal{G} \cup \mathcal{G}^0$ contains exactly one space N_i. In particular, $\{\mathrm{St}(T)N_i\}_i = \mathcal{G} \cup \mathcal{G}^0$. Let $T_i := N_i^\perp / N_i$. Then, there exists an injective function*

$$\eta : \bigcup_i \mathrm{Irr}\,\mathrm{St}(T_i) \to \mathrm{Irr}\,\mathrm{Cl},$$

such that

$$\mathcal{H}_{n,t} \simeq \bigoplus_i \bigoplus_{\substack{\tau \in \\ \mathrm{Irr}\,\mathrm{St}(T_i)}} \mathrm{Ind}_{\mathrm{St}(T)^{N_i}}^{\mathrm{St}(T)}(\tau) \otimes \eta(\tau).$$

If $\tau \in \mathrm{Irr}\,\mathrm{St}(T_i)$, the corresponding Clifford representation has rank $\mathrm{rk}\,\eta(\tau) = \dim T_i$.

Consider the subset of terms associated to $N_j \in \mathcal{G}^0$. This set is non-empty if and only if $r - s = 0 \mod D$, and in this case, the terms span $C_{\mathbf{1}_t}$. For any such term, $\tau \in \mathrm{Irr}\,\mathrm{St}(T_j)$, it holds that $\mathcal{P} \subset \ker \eta(\tau)$.

To prove this theorem we will build heavily on the results of the previous subsection, and derive a couple of intermediate results more. Let

$$\begin{aligned} \mathcal{C}_m &:= \mathrm{span}\,\{C_N \,|\, N \in \mathcal{G}_m\}, \\ \mathcal{D}_m &:= \mathrm{span}\,\{C_N \,|\, N \in \mathcal{G}_m^0\}, \end{aligned}$$

with $0 < m \leq m(T)$, and $\mathcal{C}_0 := \mathcal{H}_{n,t}$. The spaces $\mathcal{D}_m$ are only non-zero if $r - s = 0 \mod D$ and $m \leq m(T) - 1$. Notice that $\mathcal{D}_1 = C_{\mathbf{1}_t}$. These spaces are nested: if $N \subset N'$ then $C_{N'} \subset C_N$, and thus,

$$\mathcal{C}_{m+1} \subset \mathcal{C}_m, \quad \mathcal{D}_{m+1} \subset \mathcal{C}_m, \quad \mathcal{D}_{m+1} \subset \mathcal{D}_m. \tag{25}$$

Lemma V.5. *For arbitrary values of r, s, m, d, the algebra $\mathrm{A}_{r,s}$ acts on $\mathcal{C}_m$ and $\mathcal{D}_m$. Furthermore:*

1. *the action of $\mathrm{A}_{r,s}$ on $\langle \mathcal{D}_m, \mathcal{C}_m \rangle^\perp$ mods out the ideal $\mathrm{A}^m_{r,s}$,*
2. *if $r - s = 0 \mod d$ the action on $\mathcal{D}_m^\perp$ mods out the ideal $\mathrm{A}^{m,0}_{r,s}$.*

Proof. The first point follows from the fact that for each $R(O)P_N \in \mathcal{A}^m_{r,s}$, the code space $C_N \subseteq \langle \mathcal{D}_m, \mathcal{C}_m \rangle$. The second point follows similarly: if $R(O)P_N \in \mathcal{A}^{m,0}_{r,s}$, the code space $C_N \subseteq \mathcal{D}_m$. □

This motivates

$$\mathcal{K}_m := \mathcal{C}_m \cap \langle \mathcal{C}_{m+1}, C_{\mathbf{1}_t} \rangle^\perp$$
$$\mathcal{L}_m := \mathcal{D}_m \cap \mathcal{D}_{m+1}^\perp$$

for $m > 0$ and $\mathcal{K}_0 := \langle \mathcal{C}_1, C_{\mathbf{1}_t} \rangle^\perp$. By Lem. V.5 and $\mathcal{D}_m \subset C_{\mathbf{1}_t}$,

$$\mathrm{A}^{m+1}_{r,s} \subseteq \ker_{\mathrm{A}_{r,s}}(\mathcal{K}_m), \qquad \mathrm{A}^{m+1,0}_{r,s} \subseteq \ker_{\mathrm{A}_{r,s}}(\mathcal{L}_m). \tag{26}$$

Lemma V.6. *Let $N \in \mathcal{G}_m$ with $t - 2m \leq n$, then*

$$C_N \cap \Delta^{(t-2m)}_{r,s} = C_N \cap \mathcal{K}_m.$$

On the other hand, let $N \in \mathcal{G}^0_m$ with $t - 2m \leq n$, then

$$C_N \cap \Delta^{(t-2m)}_{r,s} = C_N \cap \mathcal{L}_m.$$

Proof. Assume that $N \in \mathcal{G}_m$, the second case follows similarly. By Lem. III.3, the commutant of Cl in $\mathrm{End}(C_N \cap \mathcal{K}_m)$ is contained in $\mathbb{C}[\mathcal{A}^N_{r,s}]$. By eq. (26), any element of $\mathrm{A}^{m+1}_{r,s}$ acts trivially on this space. This way, for any $R(O)P_{N'} \in \mathcal{A}^N_{r,s}$,

$$R(O)P_{N'}|_{C_N \cap \mathcal{K}_m} = \begin{cases} R(O)|_{C_N \cap \mathcal{K}_m}, & N' = N, \\ 0, & N' \supsetneq N. \end{cases}$$

This way, the commutant of Cl in $\mathrm{End}(C_N \cap \mathcal{K}_m)$ is generated by the subgroup of $\mathrm{St}(T)$ for which $R(O) \cdot (C_N \cap \mathcal{K}_m) = C_N \cap \mathcal{K}_m$. Because $R(O)\mathcal{K}_m = \mathcal{K}_m$ for all $O \in \mathrm{St}(T)$, it is sufficient to require $R(O)C_N = C_N$, which happens if and only if $O \in \mathrm{St}(T)^N$, where

$$\mathrm{St}(T)^N := \{O \in \mathrm{St}(T) \,|\, ON = N\}.$$

By Lem. III.6 the action of $\mathrm{St}(T)^N$ on C_N realizes the homomorphism $\mathrm{St}(T)^N \to \mathrm{St}(T_N)$.

Then, the commutant of Cl in $\mathrm{End}(C_N \cap \mathcal{K}_m)$ is generated by $R(\mathrm{St}(T)^N)|_{C_N}$, and by Schur-Weyl duality there exists some injective function $\tilde{\eta} : \mathrm{Irr}\,\mathrm{St}(T_N) \to \mathrm{Irr}\,(\mathrm{Cl})$ for which

$$C_N \cap \mathcal{K}_m \simeq \bigoplus_{\tau \in \mathrm{Irr}\,\mathrm{St}(T_N)} \tau \otimes \tilde{\eta}(\tau),$$

where, by the same argument as in the proof of Thm. V.1, the sum ranges over all $\mathrm{Irr}\,\mathrm{St}(T_N) \subseteq \mathrm{Irr}\,\mathrm{St}(T)^N$ since $t - 2m \leq n$.

Now, every $F \in \mathrm{Hom}(\mathrm{X} \to N^\perp)$ for which $\mathrm{rank}\, q_{r,s}|_{T_N} = t - 2m$ is orthogonal to all subcodes $C_{N'} \subset C_N$, and therefore

$$C_N \cap \Delta^{(t-2m)}_{r,s} \subseteq C_N \cap \mathcal{K}_m,$$

which implies that $\eta_N(\tau) \subseteq \tilde{\eta}(\tau)$ for each τ (where η_N is as in Thm. V.1). However, $\tilde{\eta}(\tau)$ is irreducible so that $\tilde{\eta}(\tau) = \eta_N(\tau)$ and

$$C_N \cap \Delta^{(t-2m)}_{r,s} \simeq C_N \cap \mathcal{K}_m,$$

and so the result follows. □

Lemma V.7. *The equation*

$$\mathcal{H}_{n,t} = \bigoplus_m \mathcal{K}_m \oplus \mathcal{L}_m, \tag{27}$$

gives an orthogonal decomposition into $\mathrm{St}(T) \times \mathrm{Cl}$-subrepresentations.

Proof. Because $\mathrm{St}(T)$ acts on $\mathcal{G}_m$ and $\mathcal{G}_m^0$, the spaces $\mathcal{K}_m$ and $\mathcal{L}_m$ are $\mathrm{St}(T) \times \mathrm{Cl}$ representations.

Because $P_{\mathbf{1}_t} : \mathcal{C}_m \to \mathcal{C}_m$,

$$\begin{aligned}\bigoplus_m \mathcal{K}_m &= \bigoplus_m \left(\mathcal{C}_m \cap \mathcal{C}_{m+1}^\perp \cap C_{\mathbf{1}_t}^\perp\right) \\ &= \left(\bigoplus_m \mathcal{C}_m \cap \mathcal{C}_{m+1}^\perp\right) \cap C_{\mathbf{1}_t}^\perp \\ &= \mathcal{C}_0 \cap C_{\mathbf{1}_t}^\perp = \mathcal{H}_{n,t} \cap C_{\mathbf{1}_t}^\perp = C_{\mathbf{1}_t}^\perp.\end{aligned}$$

Moreover, $C_{\mathbf{1}_t} = \oplus_m \mathcal{L}_m$.

By definition, $\mathcal{K}_m \perp \mathcal{K}_{m'}$ and $\mathcal{L}_m \perp \mathcal{L}_{m'}$ for each $m \neq m'$. Finally, $\mathcal{K}_m \perp C_{\mathbf{1}_t} \supseteq \mathcal{L}_{m'}$ for every m, m'. □

Lemma V.8. *Use the notation above, let* $0 < m \leq m(T)$ *and* $t - m \leq n$*. Then,*

$$\mathcal{K}_m = \bigoplus_{N \in \mathcal{G}_m} (C_N \cap \mathcal{K}_m), \qquad \mathcal{L}_m = \bigoplus_{N \in \mathcal{G}_m^0} (C_N \cap \mathcal{L}_m). \tag{28}$$

Proof. We prove the statement for $\mathcal{K}_m$. The decomposition for $\mathcal{L}_m$ can be proven analogously. Throughout the proof, we let

$$\operatorname{supp} C_N := \bigcup_{\psi \in C_N} \operatorname{supp} \psi.$$

Consider one of the terms in the right-hand side of (28). By Lem. V.6,

$$C_N \cap \mathcal{K}_m = \operatorname{span}\{\Delta_{r,s}(U)\Psi_B \,|\, U \in \mathrm{Cl},\ \Psi_B \in C_N \text{ wght. vec. with } \operatorname{rank} B = t - 2m\}.$$

We claim that all weight vectors Ψ_B with rank $t - 2m$ in C_N are linear combinations of $|[F]_N\rangle$ where $\operatorname{range} F = N^\perp$. To see this, write

$$\Psi_B = \sum_{F' \in \mathrm{Hom}(\mathrm{X} \to N^\perp)} c_{F'} \, |[F']_N\rangle \,.$$

Then, for any F' with $c_{F'} \neq 0$, according to Lem. IV.4,

$$B = (F')^T M_{r,s} F',$$

so that the rank of the right-hand side is $t - 2m$. This implies that $\langle \operatorname{range} F', N\rangle = N^\perp$. This way, because $t - m \leq n$, there exists some $F \in [F']_N$ for which $\operatorname{range} F = N^\perp$. Furthermore, for any two distinct coset vectors $|[F]_N\rangle \neq |[F']_N\rangle$ their supports are disjoint. This implies that there exists at least one $F \in \operatorname{supp} \Psi_B$ for which $\operatorname{range} F = N^\perp$.

The corresponding computational basis vector $|F\rangle$ is orthogonal to all other codes $C_{N'}$, where $N' \in \mathcal{G}_m$ and $N' \neq N$. To prove this we argue by contradiction. If $F \in \operatorname{supp}(C_N) \cap \operatorname{supp}(C_{N'})$ then $N^\perp = \operatorname{range} F \subseteq N^\perp \cap (N')^\perp$. But this can not hold because $N^\perp$ and $(N')^\perp$ are of the same dimension and distinct.

This way,

$$\{\Psi_B \in C_N \text{ wght. vec. with } \operatorname{rank} B = t - 2m\} \cap \operatorname{span}\{C_{N'} \,|\, N' \in \mathcal{G}_m,\ N' \neq N\} = \emptyset.$$

Finally, each irrep in $\mathcal{K}_m \cap C_N$ contains at least one maximal rank weight vector Ψ_B, and so is linearly independent of

$$\operatorname{span}\{C_{N'} \,|\, N' \in \mathcal{G}_m,\ N' \neq N\}.$$

Then,

$$\mathcal{K}_m = \sum_{N \in \mathcal{G}_m} (\mathcal{K}_m \cap C_N) = \bigoplus_{N \in \mathcal{G}_m} (\mathcal{K}_m \cap C_N), \tag{29}$$

where the first equality follows from each P_N commuting with the projector onto $\mathcal{K}_m$, and the second equality follows from linear independence. □

Notice that combining eq. (29) with Lem. V.6, we obtain that

$$\begin{aligned}\mathcal{K}_m \cap \Delta_{r,s}^{(t-2m')} &= \bigoplus_{N\in\mathcal{G}_m} \mathcal{K}_m \cap C_N \cap \Delta_{r,s}^{(t-2m')} \\ &= \bigoplus_{N\in\mathcal{G}_m} \mathcal{K}_m \cap C_N \cap \mathcal{K}_{m'} \\ &= \begin{cases} \mathcal{K}_m, & m = m', \\ \{0\}, & m \neq m', \end{cases}\end{aligned}$$

and similarly

$$\mathcal{L}_m \cap \Delta_{r,s}^{(t-2m')} = \begin{cases} \mathcal{L}_m, & m = m', \\ \{0\}, & m \neq m'. \end{cases}$$

This allows us to conclude that

$$\Delta_{r,s}^{(t-2m)} = \mathcal{K}_m \oplus \mathcal{L}_m. \tag{30}$$

We now look at the representation spaces on the right-hand side of eq. (30).

Theorem V.3. *Assume that m is such that $t - 2m < n$.*
Let $N \in \mathcal{G}_m$. Then, there exists an injective map $\eta : \operatorname{Irr}\operatorname{St}(T_N) \to \operatorname{Irr}\operatorname{Cl}$ such that, as a representation of $\operatorname{St}(T) \times \operatorname{Cl}$,

$$\mathcal{K}_m \simeq \bigoplus_{\tau\in\operatorname{Irr}\operatorname{St}(T_N)} \eta(\tau) \otimes \operatorname{Ind}_{\operatorname{St}_N}^{\operatorname{St}(T)}(\tau), \tag{31}$$

where $\dim\eta(\tau) \geq \dim\tau$ for all τ.
Moreover if $r - s = 0 \mod D$, let $N' \in \mathcal{G}_m^0$. Then, there exists an injective map $\eta_0 : \operatorname{Irr}\operatorname{St}(T_N) \to \operatorname{Irr}\operatorname{Sp}(V) \subset \operatorname{Irr}\operatorname{Cl}$ such that, as a representation of $\operatorname{St}(T) \times \operatorname{Cl}$,

$$\mathcal{L}_m \simeq \bigoplus_{\tau\in\operatorname{Irr}\operatorname{St}(T_N)} \eta(\tau) \otimes \operatorname{Ind}_{\operatorname{St}_N}^{\operatorname{St}(T)}(\tau), \tag{32}$$

where $\dim\eta(\tau) \geq \dim\tau$ for all τ.

Proof. We prove the first case. The second case follows similarly, with the only difference that because $\mathcal{P} \subseteq \ker(\mathcal{L}_m)$, all the irreps appearing in eq. (32) are actually $\operatorname{Sp}(V)$ irreps.

Let $\mathcal{R}_N^\tau \subset C_N \cap \mathcal{K}_m$ be the subspace isomorphic to $\tau \otimes \eta_N(\tau)$ as in Thm. V.1, where $\tau \in \operatorname{Irr}\operatorname{St}(T_N)$. Acting with $\operatorname{St}(T)$ on $\mathcal{R}_N^\tau$, we get some representation $\mathcal{R}^\tau \subseteq \mathcal{H}_{n,t}$. Then, there is some representation $I(\tau)$ of $\operatorname{St}(T)$ for which $\mathcal{R}^\tau \simeq I(\tau) \otimes \eta(\tau)$. By Lem. V.8,

$$I(\tau) \otimes \eta(\tau) \simeq \bigoplus_{O\in\operatorname{St}(T)/\operatorname{St}(T)^N} \mathcal{R}_{ON}^\tau = \left(\bigoplus_{O\in\operatorname{St}(T)/\operatorname{St}(T)^N} \tau_O\right) \otimes \eta(\tau),$$

where in the last equality we used Lem. V.4 to show that $\eta_{ON}(\tau) \simeq \eta_N(\tau) := \eta(\tau)$ for all $O \in \operatorname{St}(T)$, and where τ_O is a copy of the τ factor within $\mathcal{R}_{ON}^\tau$. Now, the space

$$\bigoplus_{O\in\operatorname{St}(T)/\operatorname{St}_N} \tau_O$$

is a $\operatorname{St}(T)$ representation and so it is isomorphic to $I(\tau)$. The action of $\operatorname{St}(T)$ on this space is $\operatorname{Ind}_{\operatorname{St}_N}^{\operatorname{St}(T)}(\tau)$, as claimed: Choose a complete set of representatives $O_i \in \operatorname{St}(T)/\operatorname{St}_N$. Then, for any $O \in \operatorname{St}(T)$ there is a permutation $\pi : i \to \pi(i)$ and an $O' \in \operatorname{St}_N$ such that $OO_i = O_{\pi(i)}$. This way, if

$$\psi = \sum_i R(O_i)\psi_i, \qquad \psi_i \in \tau_{\mathbb{1}},$$

then

$$R(O)\psi = \sum_i R(O_{\pi(i)})R(O')\psi_i.$$

Finally,

$$\mathcal{K}_m = \bigoplus_{\substack{N'\in\mathcal{G}_m,\\ \tau\in\operatorname{Irr}\operatorname{St}(T_N)}} \mathcal{R}^\tau_{N'} = \bigoplus_{\substack{O_i\in\operatorname{St}(T)/\operatorname{St}_N,\\ \tau\in\operatorname{Irr}\operatorname{St}(T_N)}} \mathcal{R}^\tau_{O_iN} \simeq \bigoplus_{\tau\in\operatorname{Irr}\operatorname{St}(T_N)} \operatorname{Ind}_{\operatorname{St}_N}^{\operatorname{St}(T)}(\tau)\otimes\eta(\tau).$$

Here the first equality follows by Lem. V.8, and the second equality follows by Lem. III.2. □

Proof of Thm. V.2. Thm. V.3 and Lem. V.7 show that the claimed equation is a valid decomposition of $\mathcal{H}_{n,t}$ as a $\operatorname{St}(T)\times\mathrm{Cl}$ representation. We now show that η is injective. Consider two spaces $T_i, T_{i'}$ with $\tau\in\operatorname{Irr}\operatorname{St}(T_i)$, $\tau'\in\operatorname{Irr}\operatorname{St}(T_{i'})$. By Thm. V.3, $\operatorname{rk}\eta(\tau)=\dim T_i$. In particular, if $\dim T_i\neq\dim T_{i'}$ then $\eta(\tau)\not\simeq\eta(\tau')$. On the other hand, $\dim T_i=\dim T_{i'}=m$, assume without loss of generality that $N_i\in\mathcal{G}_m$ and $N_{i'}\in\mathcal{G}^0_m$. In this case, by Thm. V.3, the $\eta(\tau)$-isotype is a subspace of $\mathcal{K}_m$ and the $\eta(\tau')$-isotype is a subspace of $\mathcal{L}_m$. In particular $\mathcal{P}\subseteq\ker\eta(\tau')$ and $\mathcal{P}\not\subseteq\ker\eta(\tau)$. □

As a corollary, we can rederive a classical result from the invariant theory of the Clifford group [20, 22, 23].

Corollary V.1. *If d is odd, $\Delta_{r,s}$ contains a* Cl*-trivial subrepresentation if and only if $t=2t'$ is even, $s=t' \mod 2$ and $r-s=0 \mod d$.*

If $d=2$, $\Delta_{r,s}$ contains a Cl*-trivial subrepresentation only if*

$$r-s=0 \mod 8. \tag{33}$$

Moreover if $s=0$ (so $r=t$), eq. (33) *is also sufficient for this.*

In all cases considered above, whenever the trivial component is non-zero, it is equal to $\mathcal{L}_{t'}$.

Proof. Let $\rho\subset\Delta_{r,s}$ be the trivial component. Because $\mathcal{P}$ acts trivially on ρ it is non-trivial only if $r-s=0 \mod D$, in which case $\rho\subset C_{\mathbf{1}_t}$. Furthermore $\operatorname{rk}(\rho)=0$ so that $\rho\subseteq\Delta^{(0)}_{r,s}$. A necessary condition for the latter to be non-empty is for $t=2t'$ to be even, in which case

$$\Delta^{(0)}_{r,s}=\mathcal{K}_{t'}\oplus\mathcal{L}_{t'}=\mathcal{L}_{t'}.$$

Here we used eq. (30) together with the fact that a stochastic isotropic N must contain $\mathbf{1}_t$ to be maximal.

Now, $\Delta^{(0)}_{r,s}$ is non-zero if and only if $\mathcal{G}^0_{t'}$ to be non-empty. If d is odd, this happens exactly when $T\simeq\mathbb{H}^{\oplus t'}$, or equivalently, when $\operatorname{dis}(\beta_{r,s})=(-1)^s=(-1)^{t'}$. In this case, $\Delta^{(0)}_{r,s}=\rho$ by Lem. IV.3.

If $d=2$, $\Delta^{(0)}_{r,s}$ may contain non-trivial components according to Lem. IV.3. In particular, $(\omega_8\mathbb{1})^{\otimes(r-s)}=1$ if and only if (33) holds. This way, ρ is non-trivial only if the latter equation holds, in which case $\rho=\Delta^{(0)}_{r,s}$.

Finally, if $d=2$ and $s=0$, then it can be shown [23] that stochastic isotropic subspaces of dimension t' exist if and only if $r-s=t=0 \mod 8$. □

We finally point out how the spaces corresponding to all representations with a fixed rank, $\langle\mathcal{K}_m,\mathcal{L}_m\rangle$, are "singled out" rather canonically by the action of $\mathrm{A}_{r,s}$ on them.

Lemma V.9. *The space $\langle\mathcal{K}_m,\mathcal{L}_m\rangle$ is the maximal $\mathrm{A}_{r,s}$-subrepresentation of $\mathcal{H}_{n,t}$ which has $\mathrm{A}^{m+1}_{r,s}$ in its kernel but not $\mathrm{A}^m_{r,s}$.*

Proof. Consider a distinct pair of subspaces $\mathcal{S}_i=\langle\mathcal{K}_{m_i},\mathcal{L}_{m_i}\rangle$, with $i=1,2$ and $m_1\neq m_2$. Each of these spaces is a Cl representation and, moreover, they share no irrep in common. That is, if χ_i is the character of $\mathcal{S}_i$, then the character inner product vanishes, $\langle\chi_1,\chi_2\rangle_{\mathrm{Cl}}=0$. To see this, notice that by eq. (30), every irrep in $\mathcal{S}_i$ has rank m_i.

By Schur's lemma, each space $\mathcal{K}_m$, $\mathcal{L}_m$ is preserved by $\mathrm{A}_{r,s}$. In particular one may block diagonalize any $A\in\mathrm{A}_{r,s}$ as

$$A=\bigoplus_m A_m,$$

where A_m acts on $\langle\mathcal{K}_m,\mathcal{L}_m\rangle$. By Lem. V.5, for any $A\in\mathrm{A}^m_{r,s}$ it holds that $A_{m'}=0$ for all $m'>m$, so that $\langle\mathcal{K}_m,\mathcal{L}_m\rangle$ mods out $\mathrm{A}^{m+1}_{r,s}$ but not $\mathrm{A}^m_{r,s}$. Moreover, any $\mathrm{A}_{r,s}$-irrep in some $\langle\mathcal{K}_{m'},\mathcal{L}_{m'}\rangle$ with $m'\neq m$ either mods out $\mathrm{A}^m_{r,s}$ (if $m'<m$), or does not mod out $\mathrm{A}^{m+1}_{r,s}$ (if $m<m'$). □

C. The space of stabilizer tensor powers

In [7] it was shown that the space spanned by stabilizer tensor powers is the trivial $\mathrm{St}(T)$-subrepresentation of $\mathcal{H}_{n,t}$:

$$\mathbb{C}\mathrm{STABS}_{n,t} := \mathrm{span}\left\{|s\rangle^{\otimes t} \mid |s\rangle \in \mathrm{STABS}_n\right\} = \mathcal{H}_{n,t}^{\mathrm{St}(T)}.$$

Independently, [15] decomposes the trivial $\mathrm{O}(T)$-subrepresentation of the same Hilbert space (for d odd) as

$$\mathcal{H}_{n,t}^{\mathrm{O}(T)} \simeq \bigoplus_r \eta(\mathrm{id}_{\mathrm{St}(T_r)}). \tag{34}$$

Here we show, with an analogous calculation, that the space $\mathbb{C}\mathrm{STABS}_{n,t}$ decomposes in a similar fashion to eq. (34). For simplicity, we consider only the case where $q_{r,s}(\mathbf{1}_t) \neq 0$.

Consider the decomposition of $\mathcal{H}_{n,t}$ into $\mathrm{St}(T)$-isotypes,

$$\mathcal{H}_{n,t} \simeq \bigoplus_{\tau \in \mathrm{Irr}\,\mathrm{St}(T)} \tau \otimes \Theta(\tau), \tag{35}$$

where $\Theta(\tau)$ is a (possibly reducible) Cl-representation. We are interested in decomposing $\Theta(\mathrm{id}_{\mathrm{St}(T)})$. Consider some $\tau \in \mathrm{Irr}\,\mathrm{St}(T_m)$, where T_m is as in Thm. V.2. Comparing (35) to Thm. V.2, we see that the multiplicity of $\eta(\tau)$ in $\Theta(\mathrm{id}_{\mathrm{St}(T)})$ is equal to the multiplicity of $\mathrm{id}_{\mathrm{St}(T)}$ in $\mathrm{Ind}_{\mathrm{St}(T)^{N_m}}^{\mathrm{St}(T)}(\tau)$, ie.

$$\langle \Theta(\mathrm{id}_{\mathrm{St}(T)}),\ \eta(\tau)\rangle_{\mathrm{Cl}} = \langle \mathrm{id}_{\mathrm{St}(T)},\ \mathrm{Ind}_{\mathrm{St}(T)^{N_m}}^{\mathrm{St}(T)}(\tau)\rangle_{\mathrm{St}(T)}.$$

By Frobenius reciprocity,

$$\left\langle \mathrm{id}_{\mathrm{St}(T)},\ \mathrm{Ind}_{\mathrm{St}(T)^{N_m}}^{\mathrm{St}(T)}(\tau)\right\rangle_{\mathrm{St}(T)} = \left\langle \mathrm{Res}_{\mathrm{St}(T)^{N_m}}(\mathrm{id}_{\mathrm{St}(T)}),\ \tau\right\rangle_{\mathrm{St}(T)^{N_m}} = \left\langle \mathrm{id}_{\mathrm{St}(T)^{N_m}}, \tau\right\rangle_{\mathrm{St}(T)^{N_m}} = \delta_{\tau,\mathrm{id}_{\mathrm{St}(T_m)}}.$$

This proves:

$$\mathbb{C}\mathrm{STABS}_{n,t} \simeq \bigoplus_m \eta(\mathrm{id}_{\mathrm{St}(T_m)}) \tag{36}$$

D. Exact dualities for low tensor powers

For some combinations of r, s, d, the resulting representation $\Delta_{r,s}$ gives rise to an exact correspondence between $\mathrm{St}(T)$ and Cl. This happens exactly when there are no isotropic stochastic subspaces of T. For qubits, this is exactly the case if $r = t = 3$. Here one obtains a duality between Cl and $\mathrm{St}(\mathbb{Z}_2^3) \simeq S_3$, and this proves that Cl is a unitary 3-design.

If $d > 2$, on the other hand, Ref. [24, Thm. 3] provides a list of tensor power representations which give rise to this exact duality. Here we include a short proof that this list is complete.

Lemma V.10. *Let $n \geq 2$, d be odd and $r + s \geq 2$. Then the commutant of $\Delta_{r,s}$ is spanned by $R(O)$ where $O \in \mathrm{St}(T)$ if and only if d, r, s satisfy $r + s \leq 3$, $rs = 0$, and if $r + s = 3$ then $\ell(3) = -\ell(-1)$.*

Proof. A necessary and sufficient condition for R to span the commutant of $\Delta_{r,s}$ is $|\mathcal{G}_m| = |\mathcal{G}_m^0| = 0$ for all m. This happens if and only if both $\mathbf{1}_t$ and $\mathbf{1}_t^\perp$ are anisotropic.

$$T = \langle \mathbf{1}_t \rangle \oplus \mathbf{1}_t^\perp,$$

So by the Chevalley-Warning theorem $\dim \mathbf{1}_t^\perp = t - 1 \leq 2$.

Anisotropicity of $\mathbf{1}_t$ is equivalent to $r - s \neq 0 \mod d$. If $t = 2$, this implies $s \in \{0, 2\}$ and hence $rs = 0$. Anisotropicity of $\mathbf{1}_t^\perp$ is equivalent to the following: there exist $a, b \neq 0$ for which

$$\beta_{r,s}|_{\mathbf{1}_t^\perp} \sim \begin{cases} a\beta_{1,0}, & t = 2, \\ \mathrm{diag}(1, b), & t = 3, \end{cases} \tag{37}$$

where $\ell(b) \neq \ell(-1)$ (which is equivalent to the condition that $\mathbf{1}_t^\perp$ is not a hyperbolic plane). By multiplicativity of the discriminant we have that

$$\begin{aligned}\operatorname{dis}(\beta_{r,s}) &= \ell((-1)^s) \\ &= \operatorname{dis}(\langle \mathbf{1}_t \rangle)\operatorname{dis}(\mathbf{1}_t^\perp) \\ &= \begin{cases} \ell(r-s)\ell(a), & t=2, \\ \ell(r-s)\ell(b), & t=3. \end{cases}\end{aligned}$$

If $t = 2$, there always exists an a satisfying the conditions above. We conclude that if $r + s = 2$ and $rs = 0$, then $|\mathcal{G}_m|,\ |\mathcal{G}_m^0| = 0$ for all m.

If $t = 3$, in contrast, b is subject to the following *two* conditions derived above:

$$\ell(b) = \ell(r-s)\ell((-1)^s), \qquad \ell(b) = -\ell(-1).$$

A solution to these equations exists if and only if

$$-1 = \ell\left((-1)^{s+1}(r-s)\right).$$

If $s = 1, 2$ this equation implies $1 = -1$ and does not hold. If $s = 0, 3$ it holds if and only if $\ell(3) = -\ell(-1)$. □

As a direct consequence of this, we have that, for any r, s, d as in Lem V.10,

$$\Delta_{r,s} \simeq \bigoplus_{\tau \in \operatorname{Irr} \mathrm{St}(T)} \tau \otimes \theta(\tau),$$

for some injective function $\theta : \operatorname{Irr} \mathrm{St}(T) \to \operatorname{Irr} \mathrm{Cl}$.

Remark V.1. *The statement [24, Thm. 3] deals with the case $r = t = 3$, where the condition $d = 2 \mod 3$ is obtained. This condition is equivalent to our condition $\ell(3) = -\ell(-1)$. To see this, recall that our condition is equivalent to $\mathbf{1}_3^\perp$ being anisotropic, which is equivalent to the equations*

$$1 + x_1^2 + x_2^2 = 0 = 1 + x_1 + x_2,$$

having no solution over $\mathbb{Z}_d$. A short calculation shows that these equations have a solution if and only if there exists an x for which $1 + x + x^2 = 0$ (in which case, the solution is $x_1 = x = x_2^{-1}$). Finally, as pointed out in [24], this polynomial is reducible (and hence contains a root over $\mathbb{Z}_d$) if and only if $d \neq 2 \mod 3$.

VI. REAL CLIFFORD ACTION ON $C_{\mathbf{1}_t}$

Consider the action of RCl on $C_{\mathbf{1}_t}$ when $d = 2$ and $r - s = 0 \mod 4$. Since $\mathcal{P} \in \ker(C_{\mathbf{1}_t})$, this action realizes the homomorphism $\mathrm{RCl} \to \mathrm{RCl}/\mathcal{P}$ where the orthogonal group $\mathrm{O}(V) \subset \mathrm{Sp}(V)$ preserves the form $\kappa(v) = v_z \cdot v_x$. For simplicity, we call this representation Δ.

Recall that the coset basis for $C_{\mathbf{1}_t}$ is given by

$$|[F]_{N_1}\rangle = \frac{1}{\sqrt{|N|}} \sum_{F' \in \operatorname{Hom}(\mathrm{X} \to N_1)} |F + F'\rangle$$

where $N_1 := \langle \mathbf{1}_t \rangle$ and $F \in \operatorname{Hom}(\mathrm{X} \to \mathbf{1}_t^\perp)$. This basis can be equivalently described using the columns $f_1, \ldots f_n \in T$ of F,

$$|[F]_{N_1}\rangle = \frac{1}{\sqrt{|N|}} \big(|f_1\rangle + |f_1 + \mathbf{1}_t\rangle \big) \otimes \cdots \otimes \big(|f_n\rangle + |f_n + \mathbf{1}_t\rangle \big).$$

Now, let T' be a subspace of $\mathbf{1}_t^\perp$ for which $\mathbf{1}_t^\perp = \langle \mathbf{1}_t \rangle \oplus T'$ so that

$$T' \simeq \mathbf{1}_t^\perp / \mathbf{1}_t,$$

and let $e \in (T')^\perp$ be the unique solution to $\beta_{r,s}(e, \mathbf{1}_t) = 1$. Let $\beta' = \beta|_{T'}$, and $\tilde{\mathrm{Sp}}(T')$ be the subgroup of $\mathrm{Gl}(T)$ of matrices S satisfying

$$Se = e, \quad S\mathbf{1}_t = \mathbf{1}_t, \quad ST' = T',$$

and for every $u, v \in T'$,

$$\beta'(Su, Sv) = \beta'(u, v).$$

This subgroup is an embeddig of $\mathrm{Sp}(T')$ into $\mathrm{Gl}(T)$, since, by Prop. II.10, β' is symplectic. We may define a representation $\tilde{\Delta}$ of $\tilde{\mathrm{Sp}}(T')$ on $C_{\mathbf{1}_t}$ by

$$\tilde{\Delta}(S) \left|[F]_{N_1}\right\rangle = \left|[SF]_{N_1}\right\rangle.$$

Lemma VI.1. *The representations Δ and $\tilde{\Delta}$ commute with each other.*

Proof. A straightforward but bulky calculation shows that, for any $S \in \tilde{\mathrm{Sp}}(T')$ and for any basis element $\left|[F]_{\langle \mathbf{1}_t \rangle}\right\rangle$, the following holds:

$$\begin{aligned}
\Delta(\mathrm{H}_i)\tilde{\Delta}(S) \left|[F]_{\langle \mathbf{1}_t \rangle}\right\rangle &= \tilde{\Delta}(S)\Delta(\mathrm{H}_i) \left|[F]_{\langle \mathbf{1}_t \rangle}\right\rangle \\
\Delta(X_i)\tilde{\Delta}(S) \left|[F]_{\langle \mathbf{1}_t \rangle}\right\rangle &= \tilde{\Delta}(S)\Delta(X_i) \left|[F]_{\langle \mathbf{1}_t \rangle}\right\rangle \\
\Delta(\mathrm{CNOT}_{ij})\tilde{\Delta}(S) \left|[F]_{\langle \mathbf{1}_t \rangle}\right\rangle &= \tilde{\Delta}(S)\Delta(\mathrm{CNOT}_{ij}) \left|[F]_{\langle \mathbf{1}_t \rangle}\right\rangle,
\end{aligned}$$

where i is arbitrary and $j \neq i$. This implies the claim since these group elements generate RCl. □

We now consider a second basis, dual to this first one, which we call the *Weyl basis*. The starting point is the identification

$$\mathcal{H}_{n,t} \simeq \mathcal{H}_{2n,t/2} \simeq \mathrm{End}(\mathcal{H}_n)^{\otimes t/2},$$

obtained first by grouping factors, and then using the inverse vectorization map. Then, letting $|W(a)\rangle := 2^{-n/2}\mathrm{vec}(()W(a))$ with $a \in V$, $W(a)$ a *real* Wey operator, and $t' := \frac{t}{2} - 1$, the basis is given by

$$\{|\Psi_A\rangle := |W(a_1)\rangle \otimes \cdots \otimes |W(a_{t'})\rangle \otimes |W(a_1 + \cdots + a_{t'}\rangle\}$$

where A is a $t' \times 2n$ matrix with rows a_i.

Lemma VI.2. *The set $\{\Psi_A\}$ introduced above is an orthonormal basis for $C_{\mathbf{1}_t}$.*

Proof. Orthonormality follows from $\mathrm{tr}\, W^T(a)W(b) = 2^n\delta_{a,b}$. Furthermore, this set contains one element for each matix in $\mathbb{Z}_2^{t' \times 2n}$, which gives

$$|\{\Psi_A\}| = 2^{n(t-2)} = \dim C_{\mathbf{1}_t}.$$

It is therefore sufficient to prove that

$$W^{\otimes(r,s)}(v) \left|\Psi_A\right\rangle = W^{\otimes t}(v) = \left|\Psi_A\right\rangle,$$

for all $v \in V$, where $W(v)$ is a *real* Weyl operator (this is because $\mathcal{Z}(\mathcal{P})$ is modded out because $r - s = 0 \mod 4$.

$$W^{\otimes 2}(v)\mathrm{vec}(W(a_i)) = \mathrm{vec}(W(v)W(a_i)W^T(v)) = (-1)^{[v,a_i]+\kappa(v)}\mathrm{vec}(W(a_i)).$$

This way

$$W^{\otimes t}(v) \left|\Psi_A\right\rangle = (-1)^{\sum_i [v,a_i]+[v,\sum_i a_i]} \left|\Psi_A\right\rangle = \left|\Psi_A\right\rangle$$

□

Following a similar argumentation as in [9, App. B], we can see that any $O \in O(V)$ acts by permuting the Weyl basis,

$$\Delta(O) \left|\Psi_A\right\rangle = \left|\Psi_{AO^T}\right\rangle.$$

Lemma VI.3. *Let $u', v' \in \mathbb{Z}_2^{t'}$ and*

$$u := (u', \mathbf{1}_{t'} \cdot u), \qquad v := (v', \mathbf{1}_{t'} \cdot v) \qquad \in \mathbb{Z}_2^{t/2}.$$

Then,

$$2^{-t/4} \sum_{w \in \mathbb{Z}_2^{t/2}} (-1)^{u \cdot w}\mathrm{vec}(W((w, v))) = \frac{1}{\sqrt{2}}\Big(\left|(v, v+u)^T\right\rangle + \left|(v, v+u)^T + \mathbf{1}_t\right\rangle \Big)$$

Let $u'_1, v'_2, \ldots, u'_n, v'_n$ be the columns of A,

$$u_i := (u'_i, \mathbf{1}_{t'} \cdot u_i), \qquad v_i := (v'_i, \mathbf{1}_{t'} \cdot v_i),$$

and let A_Z (resp. A_X) be the $(t/2) \times n$ matrix with columns u_i (resp. v_i). Then, as a corollary of the result above,

$$2^{-tn/4} \sum_{M \in \mathbb{Z}_2^{(t/2)\times n}} (-1)^{\mathrm{tr}\,(M^T A_Z)} \left|\Psi_{(M,A_X)}\right\rangle = \left| \left[\begin{pmatrix} A_X \\ A_X + A_Z \end{pmatrix} \right]_{N_1} \right\rangle$$

We now calculate the dimension of the commutant of Δ, which is equal to the number of orbits of $O(V)$ on $V^{\times(t')}$.

Consider some t'-tuple $\mathbf{v}$ of V vectors, let the orbit containing this point be $\mathcal{O}(\mathbf{v})$. Associated to $\mathbf{v}$ are a set of index sets $I \subset \{1, \ldots, t'\}$ for which $\{v_i\}_{i\in I}$ is l.i. Ordering these subsets lexicographically, we let $I(\mathbf{v})$ be the minimal such subset. Further, let $M(\mathbf{v})$ be the $(t' - |I(\mathbf{v})|) \times |I(\mathbf{v})|$ matrix such that,

$$v_j = \sum_{i \in I(\mathbf{v})} M(\mathbf{v})_{ji} v_i, \qquad \forall j \notin I(\mathbf{v}).$$

Lemma VI.4. *A point* $\mathbf{u}$ *is in the orbit* $\mathcal{O}(\mathbf{v})$ *if and only if the following conditions hold:*

1. $I(\mathbf{u}) = I(\mathbf{v})$,
2. $[u_i, u_j] = [v_i, v_j]$, *for all* $i < j \in I(\mathbf{v})$,
3. $\kappa(u_i) = \kappa(v_i)$, *for all* $i \in I(\mathbf{v})$,
4. $M(\mathbf{v}) = M(\mathbf{u})$.

Proof. The *only if* direction follows simply from the facts that $O \in \mathrm{O}(V)$ preserves $[\,\cdot\,,\,\cdot\,]$ and $\kappa(\,\cdot\,)$, and that any linear relation $\mathbf{v} \cdot a = 0$, where $a \in \mathbb{Z}_2^{t'}$, implies $O\mathbf{v} \cdot a = 0$.

Conversely, let O be an element of $O(V)$ for which $Ov_i = u_i$ for all $i \in I := I(\mathbf{v})$ —such an O exists by the Cahit-Arf theorem. Let $M := M(\mathbf{v})$. Then, for each $j \notin I$, it holds that

$$u_j = \sum_{i \in I} M_{ji} u_i = \sum_{i \in I} M_{ji} O v_i = O v_j,$$

and thus $\mathbf{u} = O\mathbf{v}$. □

In the regime where $t' \ll n$, the vast majority of orbits will contain t' linearly independent vectors. In each of these orbits $M(\mathbf{v}) = 0$ and $I(\mathbf{v}) = \{1, \ldots, t'\}$, so this class of orbits is labeled by the numbers $[v_i, v_j]$ and $\kappa(v_i)$. Since each of these numbers may be chosen independently, there are $2^{t'^2}$ such orbits. Up to subleading order corrections, this coincides with $|\mathrm{Sp}(T')|$. This leads to the intuition that $\mathrm{Sp}(T')$ captures most of the structure of the commutant of Δ. In analogy to Lem. V.6, one would expect that a subspace $\mathcal{L} \subseteq C_{\mathbf{1}_t}$ gives rise to an exact duality between RCl and $\mathrm{Sp}(T')$.

Conjecture VI.1. *For any fixed* t *and sufficiently large* n *the following holds. There exists a subspace* $\mathcal{L} \subset C_{\mathbf{1}_t}$ *with*

$$\frac{\dim C_{\mathbf{1}_t} - \dim \mathcal{L}}{\dim C_{\mathbf{1}_t}} = o(\exp(-n)),$$

and an injective function $\theta : \mathrm{Irr}\,\mathrm{Sp}(T') \to \mathrm{Irr}\,\mathrm{O}(V) \subset \mathrm{Irr}\,\mathrm{RCl}$ *such that, as a* $\mathrm{Sp}(T') \times \mathrm{RCl}$ *representation,*

$$\mathcal{L} \simeq \bigoplus_{\tau \in \mathrm{Irr}\,\mathrm{Sp}(T')} \tau \otimes \theta(\tau).$$

VII. BLACK BOX CONJUGATES OF CLIFFORD UNITARIES

Suppose one is given t uses of a black box Clifford unitary U. How large does t have to be in order to implement $\bar{U}$? The simplest case to analyse here is when the implementation is *parallel*, which is to say when there exist isometries V_1, V_2 for which $\bar{U} = V_2 U^{\otimes t} V_1$. This question is equivalent to asking what is the minimal t for which $\Delta_{0,1} \subset \Delta_{t,0}$, ie. that the conjugate representation of Cl is a subrepresentation of the t-th tensor power representation.

Lemma VII.1. *The minimal t for which $\Delta_{0,1}$ is a subrepresentation of $\Delta_{t,0}$ is*

1. $t = 7$ *if* $d = 2$,
2. $t = 2d - 1$ *if* $d = 1 \mod 4$,
3. $t = 4d - 1$ *if* $d = 3 \mod 4$.

Proof. It is clear that $t > 1$ is necessary, and so any subrepresentation ρ of $\Delta_{t,0}$ isomorphic to $\Delta_{0,1}$ will be rank-deficient and by Thm. V.3 is in the span of all codes C_N with $t - 2\dim N = 1$. This can only happen if t is odd. Moreover, $\mathcal{P} \not\subseteq \ker \Delta_{0,1}$ so that

$$\rho \subseteq \mathcal{K}_m, \quad m = \frac{t-1}{2}.$$

If d is odd and $t > 2$, then by the Chevalley-Warning theorem $\mathcal{G}_m$ is non-empty. If $d = 2$ and $t = 7$, then $\mathcal{G}_3$ is non-empty by the following example:

$$N = \langle (1111000),\ (0011110),\ (1010101) \rangle,$$

where vectors are written in the orthonormal basis of $q_{t,0}$.

Now, because for every $N, N' \in \mathcal{G}_m$, there is an $O \in \mathrm{St}(T)$ for which

$$R(O) P_N R^\dagger(O) = P_{N'},$$

it follows that these are isomoprhic as Clifford representations and we can assume without loss of generality that $\rho \subseteq C_N$ for some code N. Now, this N is such that $\dim T_N = 1$, so that $T_N = \langle [\mathbf{1}_t]_N \rangle$ and $\mathrm{St}(T_N) = \{\mathbb{1}\}$. By V.1, C_N is irreducible and so $C_N \simeq \Delta_{0,1}$.

We can use Lem. III.1, to re-express

$$\Delta_{t,0} \simeq \Delta_{r,s},$$

for some r and s that are subject to the conditions of that lemma. By Lem. III.4, $C_N \simeq \Delta_{r-m,s-m}$ must be isomorphic to $\Delta_{0,1}$, and thus we must be able to choose r and s such that $r - m = 0$, $s - m = 1$. Thus, t must be such that

$$\Delta_{t,0} \simeq \Delta_{m,m+1}.$$

Here we argue by cases: If $d = 1 \mod 4$ we require $t \mod d = m - (m+1) = -1$, the smallest odd t for which this equation holds is $t = 2d - 1$. If $d = 3 \mod 4$ we require furthermore that $s = \frac{t+1}{2}$ is even. The smallest t for which these two conditions hold is $t = 4d - 1$. If $d = 2$, we require instead that $t \mod 8 = -1$, in which case we can take $t = 7$. □

By encoding $\mathcal{H} \mapsto C_N$, where C_N is as in the proof of Lem. VII.1, we obtain an implementation of $\bar{U}$. Namely, if U_N is the encoding isometry, $U_N^\dagger U^{\otimes t} U_N = \bar{U}$. Using this result, one may use the teleportation trick from [34] to probabilistically implement $U^\dagger$ in a heralded fashion.

The protocol for conjugating Cliffords above is considerably simpler than the protocols studied in [34, 35] which conjugate *arbitrary unitaries*. Two properties contrast these two cases. First, the isometry U_N is a Clifford operation, while the corresponding isometry from [34] could require a high T-gate count. Second, in Ref. [34] it is shown that in order to implement $\bar{U}$, at least $t = d^n - 1$ black box uses are necessary (even if one is content with a heralded implementation with positive probability). On the other hand, our protocol requires only on the order of d black box uses to implement $\bar{U}$ deterministically.

Appendix A: Deferred proofs

1. Proofs from Sec. II

Proposition A.1. *Let $d = 2$. Then it holds that*

$$K^* \subset \mathrm{Q}(K), \qquad \ker \Xi = K^*, \qquad \mathrm{range}\, \Xi = \mathrm{Alt}(K).$$

Proof. One may immediately verify that $K^* \subset \mathrm{Q}(K)$ and that $K^* \subseteq \ker \Xi$. Furthermore, any q satisfying $q(v+u) = q(u)+q(v)$ is linear and so part of K^*. Thus $\ker \Xi = K^*$.

We now show the last statement. We claim that, if $\beta \in \mathrm{Alt}(K)$, then the following quadratic form refines β:

$$q(v) = \sum_{\substack{i<j\in \\ \mathrm{supp}\,(v)}} \beta(e_i, e_j),$$

where $\{e_i \,|\, i = 1, \dots, k\}$ is a basis of K. For this, we compute $q(u+v)$. Let $s_u := \mathrm{supp}\,(u)$,

$$I_1 := s_v \setminus (s_v \cap s_u), \quad I_2 := s_v \cap s_u, \quad I_3 := s_u \setminus (s_v \cap s_u),$$

so that $s_v = I_1 \cup I_2$, $s_u = I_2 \cup I_3$ and $s_{u+v} = I_1 \cup I_3$. Furthermore, for $a < b \in \{1, 2, 3\}$,

$$[a; b] := \sum_{\substack{i\in I_a \\ j\in I_b}} \beta(e_i, e_j)$$
$$[a; a] := \sum_{i<j\in I_a} \beta(e_i, e_j).$$

Then,

$$\begin{aligned} q(u+v) &= [1;1] + [1;3] + [3;3] \\ q(u) &= [1;1] + [1;2] + [2;2] \\ q(v) &= [2;2] + [2;3] + [3;3], \end{aligned}$$

and so $q(u+v) + q(u) + q(v) = [1;2] + [1;3] + [2;3]$. Finally,

$$\beta(u, v) = [1;2] + [1;3] + [2;3] + \sum_{i,j\in I_2} \beta(e_i, e_j),$$

and the last term vanishes because β is symmetric and alternating. □

Similarly, we define the *generalized polarisation map* $\tilde{\Xi} : \tilde{\mathrm{Q}}(K) \to \mathrm{Sym}(K)$ to be the additive map $\tilde{\Xi}(q) = \beta$ where q and β satisfy (6).

Proposition A.2. *Let $d = 2$. Then*

$$\ker \tilde{\Xi} = 2K^* := \{2f \,|\, f \in K^*\}, \qquad \mathrm{range}\, \tilde{\Xi} = \mathrm{Sym}(K).$$

Proof. If $\tilde{\Xi}(q) = 0$ for some $q \in \tilde{\mathrm{Q}}(K)$, then for all $u, v \in K$ it holds that

$$q(u+v) = q(u) + q(v). \tag{A1}$$

Using $v = u$ we see that $q(u) = 2f(u)$ for some $f : K \to \mathbb{Z}_2$. But by (A1), $f \in K^*$.

Now we constructively show that every $\beta \in \mathrm{Sym}(K)$ has a generalized quadratic refinement $q \in \tilde{\mathrm{Q}}(K)$. Throughout the rest of the proof we will use $\{\{\dots\}\}$ to denote a "$\mathbb{Z}_2$ pocket inside of a $\mathbb{Z}_4$ environment," that is if a is a $\mathbb{Z}_2$-valued expression, then $\{\{a\}\} = 1 \in \mathbb{Z}_4$ if $a = 1 \in \mathbb{Z}_2$ and $\{\{a\}\} = 0$ otherwise. We furthermore use the notation of the proof of A.1.

Pick some basis $\{e_i\}$ of K. Then, we claim that q defined by

$$q(u) = \sum_{i\in s_u} \{\{\beta(e_i, e_i)\}\} + \sum_{\substack{i<j \\ \in s_u}} 2\{\{\beta(e_i, e_j)\}\},$$

is a generalized refinement of β. Indeed,

$$\begin{aligned} q(u+v) &= \sum_{i\in I_1, I_3} \{\{\beta(e_i, e_i)\}\} + 2\Big\{\Big\{[1;1] + [1;3] + [3;3]\Big\}\Big\}, \\ q(u) &= \sum_{i\in I_1, I_2} \{\{\beta(e_i, e_i)\}\} + 2\Big\{\Big\{[1;1] + [1;2] + [2;2]\Big\}\Big\}, \\ q(v) &= \sum_{i\in I_2, I_3} \{\{\beta(e_i, e_i)\}\} + 2\Big\{\Big\{[2;2] + [2;3] + [3;3]\Big\}\Big\}, \end{aligned}$$

and so

$$\begin{aligned} q(u+v)-q(u)-q(v) &= \sum_{i\in I_2} 2\{\{\beta(e_i,e_i)\}\} + 2\Big\{\Big\{[1;2]+[2;3]+[1;3]\Big\}\Big\} \\ &= 2\Big\{\Big\{\sum_{i\in I_2}\beta(e_i,e_i)+[1;2]+[2;3]+[1;3]\Big\}\Big\} \\ &= 2\{\{\beta(u,v)\}\}, \end{aligned}$$

where the last line follows from the fact that β is symmetric. □

Proof of Prop. II.9. If d is odd,

$$\mathrm{dis}(\beta_N) = (-1)^m \mathrm{dis}(\beta_{r,s}) = (-1)^{s-m}.$$

This way $\beta_{r-m,s-m} \simeq \beta_N$ since they have the same rank and discriminant.

Now turn to the case $d = 2$. It suffices to prove the claim when $m = 1$. The form β_N is of odd type since $\mathbf{1}_t \notin N$ and so, for at least some $u \in T_N$,

$$\beta_N(u,u) = \beta_N(\mathbf{1}_t, u) \neq 0.$$

This way $q_N \simeq q_{r',s'}$ where $r' + s' = t - 2$. Now, we can write the isometry $T \simeq T_0 \oplus T_0^\perp$, where T_0 is the subspace of $N^\perp$ isometric to T_N. By (13) and the additivity of the generalized Arf invariant,

$$\tilde{\mathrm{Arf}}(q_N) + \tilde{\mathrm{Arf}}\left(q_{r,s}|_{T_0^\perp}\right) = r' - s' + \tilde{\mathrm{Arf}}\left(q_{r,s}|_{T_0^\perp}\right) = \tilde{\mathrm{Arf}}(q_{r,s}) = r - s \mod 8.$$

We end the proof by showing that $\tilde{\mathrm{Arf}}(q_{r,s}|_{T_0^\perp}) = 0$. Let $N = \langle a \rangle$, where $a \notin \langle \mathbf{1}_t \rangle$. It is sufficient to show that there exists a vector $b \in T$ such that

$$q_{r,s}(b) = 0, \qquad \beta_{r,s}(a,b) = a \cdot b = 1.$$

Indeed, these equations imply that $T_0^\perp = \langle a, b \rangle$ and so

$$q_{r,s}|_{T_0^\perp} \sim 2q^0_{\mathbb{H}}.$$

We prove this in three cases. Let α_r be $|\mathrm{supp}\,(a) \cap \{1,\dots,r\}|$ and $\alpha_s = |\mathrm{supp}\,(a)| - \alpha_r$. Then the cases are: *i)* either $s = 0$ or $r = 0$, *ii)* $r, s \geq 1$ and either $\alpha_r = 0$ or $\alpha_s = 0$, *iii)* $r, s, \alpha_r, \alpha_s \geq 1$.

Case i) Without loss of generality take $s = 0$ and write

$$a = e_1 + e_2 + \cdots + e_{\alpha_r - 1} + e_{\alpha_r}, \tag{A2}$$

where $4 \leq \alpha_r < r$, and the right-hand inequality follows from $a \notin \langle \mathbf{1}_t \rangle$. Then $b = e_1 + e_2 + e_3 + e_r$.

Case ii) Similarly, take $\alpha_s = 0$ and a as in (A2). Then $b = e_1 + e_{r+1}$.

Case iii) Finally, without loss of generality consider

$$a = e_1 + \cdots + e_{\alpha_r} + e_{r+1} + \cdots e_{r+\alpha_s}, \tag{A3}$$

where at least one of the two inequalities $\alpha_s < s$, $\alpha_r < r$ hold. Without loss of generality we can assume the first inequality holds and take $b = e_1 + e_t$.

□

2. Proofs from Sec. III

Lemma III.1 (Equivalent tensor powers)**.** *Let $d, r, s, r', s' \in \mathbb{N}$ be such that $r + s = r' + s'$, and let $\Delta_{r,s}$ be as above. Furthermore, if d is odd, let $r - s = r' - s' \mod d$. Then for all the following cases we have that $\Delta_{r,s} \simeq \Delta_{r',s'}$:*

1. *If $d = 1 \mod 4$*
2. *If $d = 3 \mod 4$, and $s = s' \mod 2$,*

3. If $d = 2$, $r - s = r' - s' \mod 8$.

Proof. In all the cases of the lemma, $q_{r,s} \sim q_{r',s'}$. Let $g \in \mathrm{Gl}(T)$ be some transformation for which $q_{r',s'}(g\,\cdot) = q_{r,s}(\cdot)$. If d is odd, the condition $r' - s' = r - s \mod d$ allows us to choose g such that $g\mathbf{1}_t = \mathbf{1}_t$. If $d = 2$, on the other hand, $\beta_{r,s} = \beta_{r',s'}$ and thus g is an isometry of $\beta_{r,s}$, that is

$$\beta_{r,s}(g\,\cdot, g\,\cdot) = \beta_{r,s}(\cdot,\cdot)$$

Because $\beta_{r,s}(u,u) = \beta_{r,s}(\mathbf{1}_t, u)$, the equation above implies that $g\mathbf{1}_t = \mathbf{1}_t$.

Then, the isomorphism U is can be expressed in the computational basis by $U : |F\rangle \mapsto |gF\rangle$, where $F \in \mathrm{Hom}(\mathrm{X} \to T)$. To verify this claim, we act on generators.

First, we can see that $\mathrm{CADD}^{\otimes(r,s)} = \mathrm{CADD}^{\otimes t} = \mathrm{CADD}^{\otimes(r',s')}$ and compute that

$$U\mathrm{CADD}^{\otimes t}U^\dagger = \mathrm{CADD}^{\otimes t}.$$

Since all the other generators are single-qudit, we set $n = 1$ for the rest of the proof. This way,

$$U\mathrm{H}^{\otimes(r,s)}U^\dagger = \sum_{u,v\in T} (-1)^{\beta_{r,s}(g^{-1}u, g^{-1}v)} |u\rangle\langle v| = \mathrm{H}^{\otimes(r',s')},$$

where $\mathbb{1}_2$ is the two dimensional identity.

If $d = 2$,

$$\mathrm{P}^{\otimes(r,s)} = \sum_{u\in T} \tau^{q_{r,s}(u)} |u\rangle\langle u|,$$

and so

$$U\mathrm{P}^{\otimes(r,s)}U^\dagger = \sum_{u\in T} \tau^{q_{r,s}(g^{-1}u)} |u\rangle\langle u| = \sum_{u\in T} \tau^{q_{r',s'}(u)} |u\rangle\langle u| = \mathrm{P}^{\otimes(r',s')}.$$

This proves point 3.

If d is odd, then using

$$\mathrm{P}^{\otimes(r,s)} = \sum_{u\in T} \omega^{2^{-1}q_{r,s}(u) + 2^{-1}\beta_{r,s}(\mathbf{1}_t, u)} |u\rangle\langle u|,$$

we conclude

$$U\mathrm{P}^{\otimes(r,s)}U^\dagger = \sum_{u\in T} \omega^{2^{-1}q_{r,s}(g^- u) + 2^{-1}\beta_{r,s}(g^{-1}\mathbf{1}_t, g^{-1}u)} |u\rangle\langle u| = \mathrm{P}^{\otimes(r',s')}.$$

Finally, in the odd d case we must check the isomorphism on the Pauli Z operator too:

$$UZ^{\otimes(r,s)}U^\dagger = \sum_{u\in T} \omega^{\beta_{r,s}(\mathbf{1}_t,u)} U\,|u\rangle\langle u|\,U^\dagger = \sum_{u\in T} \omega^{\beta_{r,s}(\mathbf{1}_t, g^{-1}u)} |u\rangle\langle u| = \sum_{u\in T} \omega^{\beta_{r',s'}(\mathbf{1}_t,u)} |u\rangle\langle u| = Z^{\otimes(r',s')},$$

where we used $g^{-1}\mathbf{1}_t = \mathbf{1}_t$. This concludes the proof of Points 1. and 2. □

Lemma III.4 (Code representations)**.** *Let* $N \in \mathcal{G}_m$ *and* $C_N \subset \mathcal{H}_{n,t}$ *be the associated code. Then,* $\Delta_{r,s}|_{C_N} \simeq \Delta_{r-m,s-m}$.

Proof. We construct an explicit isomorphism. Prop. II.9 implies that $q_N \sim q_{r-m,s-m}$ and hence $\beta_N \sim \beta_{r-m,s-m}$. Let the corresponding isometry be $\nu : T_N \to \mathbb{Z}_d^{t-2m}$. Now, for each coset $[F]_N$, where $F \in \mathrm{Hom}(\mathrm{X} \to N^\perp)$, there corresponds a $F_0 \in \mathrm{Hom}(\mathrm{X} \to \mathbb{Z}_d^{t-2m})$ such that $\nu[Fx]_N = F_0 x$ for all $x \in \mathrm{X}$. Furthermore, it is clear that

$$q_N([\mathbf{1}_t]_N) = q_{r,s}(\mathbf{1}_t) = r - s \mod D = q_{r-m,s-m}(\mathbf{1}_{t-2m}),$$

and since $\mathbf{1}_t \in N^\perp \setminus N$, we can choose $\nu[\mathbf{1}_t]_N = \mathbf{1}_{t-2m}$. Then the isomorphism is given by

$$\iota : |[F]_N\rangle \mapsto |\nu F_0\rangle \in \mathcal{H}_{n,t-2m}.$$

To show this, we evaluate the action of generators.

For the generators H, P and X we will take $n = 1$, since the code spaces C_N are n-th tensor powers. In this case,

$$\begin{aligned}
\iota \mathrm{H}^{\otimes(r,s)}|_{C_N}\iota^\dagger &= \iota\left(d^{-t/2}\sum_{u,v\in T}\omega^{\beta_{r,s}(u,v)}P_N\,|u\rangle\langle v|\,P_N\right)\iota^\dagger \\
&= d^{-m-t/2}\sum_{[u]_N,[v]_N\in T_N}\left(\sum_{u'\in[u]_N}\sum_{v'\in[v]_N}\omega^{\beta_{r,s}(u',v')}\right)\iota\,|[u]_N\rangle\langle[v]_N|\,\iota^\dagger \\
&= d^{m-t/2}\sum_{[u]_N,[v]_N\in T_N}\omega^{\beta_N([u]_N,[v]_N)}\iota\,|[u]_N\rangle\langle[v]_N|\,\iota^\dagger \\
&= d^{m-t/2}\sum_{a,b\in\mathbb{Z}_d^{t-2m}}\omega^{\beta_{r-m,s-m}(a,b)}\,|a\rangle\langle b|
\end{aligned}$$

where the third line follows from $\beta_{r,s}$ being well defined on T_N and the last follows from the identification $a := \nu[u]_N$, $b := \nu[v]_N$.

Similarly,

$$\iota \mathrm{P}^{\otimes(r,s)}|_{C_N}\iota^\dagger = \sum_{[v]_N\in T_N}\tau^{q_N([v]_N)}\iota\,|[v]_N\rangle\langle[v]_N|\,\iota^\dagger = \sum_{a\in\mathbb{Z}_d^{t-2m}}\tau^{q_{r-m,s-m}(a)}\,|a\rangle\langle a|\,.$$

For the Pauli X case, $X^{\otimes(r,s)} = X^{\otimes t}$, and thus for all $v \in N^\perp$,

$$\iota X^{\otimes t}\,|[v]_N\rangle = \iota\,|[v]_N + [\mathbf{1}_t]_N\rangle = |\nu[v]_N + \mathbf{1}_{t-2m}\rangle = X^{\otimes(t-2m)}\,|\nu[v]_N\rangle = X^{\otimes(t-2m)}\iota\,|[v]_N\rangle\,.$$

Finally, to compute the action of $\mathrm{CADD}^{\otimes(r,s)} = \mathrm{CADD}^{\otimes t}$, take $n = 2$ and act on $|[v_1]_N\rangle\,|[v_2]_N\rangle$,

$$\begin{aligned}
\mathrm{CADD}^{\otimes t}\,|[v_1]_N\rangle\,|[v_2]_N\rangle &= d^{-m}\sum_{u_1,u_2\in N}\mathrm{CADD}^{\otimes t}\,|v_1+u_1\rangle\,|v_2+u_2\rangle \\
&= d^{-m}\sum_{u_1,u_2\in N}|v_1+u_1\rangle\,|v_1+v_2+u_1+u_2\rangle \\
&= |[v_1]_N\rangle\,|[v_1+v_2]_N\rangle\,.
\end{aligned}$$

Then, using $a_i := \nu[v_i]_N$,

$$\iota\mathrm{CADD}^{\otimes t}\,|[v_1]_N\rangle\,|[v_2]_N\rangle = |a_1\rangle\,|a_1+a_1\rangle = \mathrm{CADD}^{\otimes(t-2m)}\iota\,|[v_1]_N\rangle\,|[v_2]_N\rangle\,.$$

□

Lemma III.5 (Qudit $C_{\mathbf{1}_t}$ representation)**.** *Let d be odd and $r - s = 0 \mod d$. As a* Cl*-subrepresentation of $\Delta_{r,s}$, we have that* $\ker(C_{\mathbf{1}_t}) = \mathcal{P}$ *and*

$$C_{\mathbf{1}_t} \simeq \begin{cases}\mu^{\otimes(r-1,s-1)}, & s > 0,\\ \mu^{\otimes(r-3,1)}, & s = 0.\end{cases}$$

Proof. Consider $\Delta_{r,s}|_{C_{\mathbf{1}_t}}$. Clearly $\mathcal{P}$ is contained in the kernel of this representation and so

$$\Delta_{r,s}(W\mu(S))|_{C_{\mathbf{1}_t}} = \mu^{\otimes(r,s)}(S)|_{C_{\mathbf{1}_t}}.$$

But then [15, Lem. 2.7, Cor. 2.3] imply that this representation is of the claimed form. Finally, this representation is a faithful representation of $\mathrm{Sp}(V)$ and we get $\ker(C_{\mathbf{1}_t}) = \mathcal{P}$. □

3. Proofs from Sec. IV

Lemma IV.3 (Rank 0 irreps)**.** *If d is odd, the unique rank zero* Cl *irrep is the trivial one. If $d = 2$ and $n \geq 3$, a rank zero representation is one dimensional, ± 1 valued, and uniquely specified by its restriction to $\mathcal{Z}(\mathrm{Cl})$. Namely, if ρ, ρ' are rank zero repesentations with*

$$\rho(\omega_8\mathbb{1}) = \rho'(\omega_8\mathbb{1}),$$

then $\rho \simeq \rho'$.

Proof. Let ρ have rank zero, and consider the group

$$G := \begin{cases} \ker\left(\mathrm{Res}_{\mathrm{RCl}}\,\rho\right), & d = 2, \\ \mathcal{P}\ker\left(\rho\right), & d > 2. \end{cases}$$

In the qubit case, $\mathrm{rk}(\rho) = 0$ implies $\mathrm{R}\mathcal{P} \subseteq \ker(\mathrm{Res}_{\mathrm{RCl}}\,\rho)$, where $\mathrm{R}\mathcal{P}$ is the subgroup of real multi-qubit Pauli matrices. Notice

$$G \trianglelefteq \mathrm{Cl}; \quad G \subseteq \mathrm{RCl}, \quad \text{if } d = 2.$$

Then,

$$\tilde{G} := \begin{cases} G/\mathrm{R}\mathcal{P}, & d = 2, \\ G/\mathcal{P}, & d > 2, \end{cases}$$

satisfies $\tilde{G} \trianglelefteq \mathrm{O}(V)$ in the qubit case, and $\tilde{G} \trianglelefteq \mathrm{Sp}(V)$ when $d > 2$. Furthermore, $\tilde{G}$ is non-trivial: in the qubit case it contains $\mathrm{R}\mathcal{P}\mathrm{RD}/\mathrm{R}\mathcal{P} \simeq Q(\mathrm{X})$, whereas in the $d > 2$ case it contains the subgroup

$$\mathcal{N} := \left\{ \begin{pmatrix} \mathbb{1} & A \\ 0 & \mathbb{1} \end{pmatrix} \mid A \in \mathrm{Sym}_{n\times n} \right\} \subseteq \mathrm{Sp}(V).$$

If $d = 2$, $\tilde{G} = \mathrm{O}(V)$ because $\mathrm{O}(V)$ is simple for $n \geq 3$ [36, Sec. 1.4], and thus $G = \mathrm{RCl}$. But then $\mathrm{RCl} \subseteq \ker\rho$. This implies that the subgroup $H = \ker(\rho)$ satisfies

$$\mathrm{RCl} \subseteq H \trianglelefteq \mathrm{Cl}.$$

Consider the group $\langle i\mathbb{1},\ H\rangle = \{H,\ iH\} \subseteq \mathrm{Cl}$. Then, $\mathrm{O}(V) \subseteq \tilde{H} := \langle i\mathbb{1},\ H\rangle/\mathcal{P} \trianglelefteq \mathrm{Sp}(V)$, but since $\mathrm{Sp}(V)$ is simple for $n \geq 3$ [36, Sec. 1.3] we have that $\tilde{H} = \mathrm{Sp}(V)$.

Now, we show that in fact $i\mathbb{1} \in H$. Because $\tilde{H} = \mathrm{Sp}(V)$, there is some $W \in \mathrm{R}\mathcal{P}$ and a phase α for which $\alpha W\mathrm{P}_1 \in H$. Using $\mathrm{R}\mathcal{P} \subset H$,

$$1 = \rho(X_1) = \rho(\alpha W\mathrm{P}_1 X_1 \mathrm{P}_1^\dagger W^\dagger \alpha^*) = \rho(\pm iW X_1 Z_1 W^\dagger) = \rho(i\mathbb{1})\rho(\pm X_1 Z_1) = \rho(i\mathbb{1}).$$

This way $\mathcal{P} \subset H$, and thus $\{H, \omega_8 H\} = \mathrm{Cl}$ where ω_8 is a primitive eigth root of unity. Thus, an arbitarary Clifford has the form $g = \omega_8^a h$, where $h \in H$ and $a \in \{0, 1\}$, and $\rho(g) = \rho(\omega_8^a \mathbb{1})$. This equation implies two things:

1. ρ is Abelian and thus one-dimensional,
2. if ρ' has rank zero aswell and $\rho'(\omega_8^a\mathbb{1}) = \rho(\omega_8^a\mathbb{1})$, then $\rho'(g) = \rho(g)$ for all $g \in \mathrm{Cl}$.

Moreover $\mathrm{Res}_{\mathcal{Z}(\mathrm{Cl})}\,\rho$ has as kernel equal to either $\mathcal{Z}(\mathrm{Cl})$ or $\langle i\mathbb{1}\rangle$. Because of this, $\mathrm{Res}_{\mathcal{Z}(\mathrm{Cl})}\,\rho$ is a $\mathbb{Z}_2$ representation and thus ± 1 valued.

If $d > 2$, on the other hand, $\tilde{G} = \mathrm{Sp}(V)$ because $\mathcal{N} \subseteq \ker\rho$ and $\mathrm{Sp}(V)$ is generated by $\mathcal{N}$ conjugates. This way $G = \mathrm{Cl}$ and $\mathrm{Sp}(V) \subseteq \ker(\rho)$. For any $S \in \mathrm{Sp}(V)$ let $v \in V$ be such that $Sv \neq v$. Then,

$$W^\dagger(v)\mu(S)W(v) = W^\dagger(v)W(Sv)\mu(S) = W(Sv - v)\mu(S) \in \ker\rho,$$

so that $W(Sv - v) \in \ker\rho$ and $\mathcal{P} \subset \ker(\rho)$. This implies, finally, that $\ker(\rho) = \mathrm{Cl}$. □

4. Proofs from Sec. V

Lemma V.1. *Let N be a stochastic isotropic subspace of dimension m, $F \in \mathrm{Hom}(\mathrm{X} \to T_N)$ be surjective, and consider the following two subspaces of C_N,*

$$\mathcal{H}_F := \mathrm{span}\left\{|J\rangle \mid J \in \mathrm{Hom}(\mathrm{X} \to T_N),\ q_N(Jx) = q_N(Fx)\ \forall\, x \in \mathrm{X},\ F^{-1}([\mathbf{1}_t]_N) = J^{-1}([\mathbf{1}_t]_N)\right\},$$
$$\mathcal{H}^F := \mathrm{span}\left\{|OF\rangle \mid O \in \mathrm{St}(T_N)\right\},$$

where $F^{-1}([\mathbf{1}_t]_N)$ is the $(n - t + 2m)$-dimensional preimage of $\mathbf{1}_t$ under F. Then $\mathcal{H}_F = \mathcal{H}^F$.

Proof. It is clear that $\mathcal{H}^F \subseteq \mathcal{H}_F$. To prove equality we will construct, for any $|J\rangle \in \mathcal{H}_F$, an $O \in \mathrm{St}(T_N)$ such that $J = OF$.

For this, it is useful to first establish $\ker F = \ker J$. Consider any $x \in \ker J$, then for all $y \in \mathrm{X}$,

$$q_N(Fy) = q_{r,s}(Jy) = q_N(J(x+y)) = q_N(Fx+Fy) = q_N(Fx) + q_N(Fy) + 2\beta_N(Fx, Fy).$$

Using $q_N(Fx) = q_N(Jx) = 0$, we obtain that $\beta_N(Fx, Fy) = 0$. Since F is surjective onto T_N and β_N is non-degenerate, it follows that $x \in \ker F$. This way, $\ker J \subseteq \ker F$. However, F is surjective, so $\mathrm{range}\, J \subseteq \mathrm{range}\, F$, and $\dim \ker J \geq \dim \ker F$. Equality follows.

Let $\tilde{F},\ \tilde{J} \in \mathrm{Hom}(\mathrm{X}/\ker(F) \to T_N)$ be such that

$$\tilde{F}[x]_{\ker F} = Fx, \quad \tilde{J}[x]_{\ker F} = Jx, \quad \forall\, x \in \mathrm{X}.$$

Then, $\tilde{J}\tilde{F}^{-1}F = J$. Indeed, for any $x \in \mathrm{X}$,

$$\tilde{J}\tilde{F}^{-1}Fx = \tilde{J}\tilde{F}^{-1}\tilde{F}[x]_{\ker F} = \tilde{J}[x]_{\ker F} = Jx.$$

We conclude the proof by showing that $\tilde{J}\tilde{F}^{-1} \in \mathrm{St}(T_N)$. For any $u \in T$, let $x_u \in \mathrm{X}$ be such that $Fx_u = u$. Then,

$$q_N(\tilde{J}\tilde{F}^{-1}u) = q_N\big(\tilde{J}[x_u]_{\ker F}\big) = q_N(Jx_u) = q_N(Fx_u) = q_N(u).$$

This shows that it is an element of $\mathrm{O}(T_N)$, the isometry group of q_N. Finally, the conditions on J imply that $\tilde{J}^{-1}[\mathbf{1}_t]_N = \tilde{F}^{-1}[\mathbf{1}_t]_N$ and so

$$\tilde{J}\tilde{F}^{-1}[\mathbf{1}_t]_N = [\mathbf{1}_t]_N.$$

□

Lemma V.2. *Let C_N be as in Thm. V.1 and $\mathcal{H}^F$ as in Lem. V.1. Then $\mathcal{H}^F$ is the regular representation of $\mathrm{St}(T_N)$, that is,*

$$\mathcal{H}^F \simeq \bigoplus_{\tau \in \mathrm{Irr}\,\mathrm{St}(T_N)} \tau \otimes \mathbb{C}^{\dim \tau},$$

where the sum ranges over every irrep of $\mathrm{St}(T_N)$, and where right-hand side factors are multiplicity spaces.

Proof. Because $F : \mathrm{X} \to T_N$ is surjective, it's columns contain a basis for T_N. Then, any $O \in \mathrm{St}(T_N)$ that satisfies $OF = F$ leaves these columns invariant and thus is the identity. This way, the isomorphism is afforded by the map $|OF\rangle \mapsto |O\rangle \in \mathbb{C}[\mathrm{St}(T_N)]$. □

Lemma V.3. *Let N and F be as in Lem. V.1. Let $G_F \subset \mathrm{Gl}(\mathrm{X}) \subset \mathrm{Cl}$ be given by*

$$G_F = \begin{cases} \{g \,|\, q^F(g^T\,\cdot) = q^F(\,\cdot), \ g^{-T}F^{-1}([\mathbf{1}_t]_N) = F^{-1}([\mathbf{1}_t]_N)\} & N \in \mathcal{G}_m, \\ \{g \,|\, q^F(g^T\,\cdot) = q^F(\,\cdot)\} & N \in \mathcal{G}_m^0. \end{cases}$$

Here, $g^{-T}F^{-1}([\mathbf{1}_t]_N) = F^{-1}([\mathbf{1}_t]_N)$ is an equality of sets. Then, the commutant of $R(\mathrm{St}(T_N))|_{\mathcal{H}^F}$ in $\mathrm{End}(\mathcal{H}^F)$ is spanned by $\Delta_{r,s}(G_F)|_{\mathcal{H}^F}$.

Proof. For any $g \in G_F$, we may directly read out that $\left|Fg^T\right\rangle \in \mathcal{H}_F$. Therefore, by Lem. V.1, there exists some $O_g \in \mathrm{St}(T_N)$ for which $Fg^T = O_g^{-1}F$. We may verify that the the function $g \mapsto O_g$ is a homomorphism. Because of $t - 2m$, this function is surjective: indeed, for any $O \in \mathrm{St}(T_N)$ we may find some $g \in \mathrm{Gl}(\mathrm{X})$ for which $O^{-1}F = Fg^T$, but by the latter equation it holds that in fact $g \in G_F$.

In this way, we have the following commuting actions on $\mathcal{H}^F$,

$$\Delta_{r,s}(g)R(O)\,|O'F\rangle = \left|OO'Fg^T\right\rangle = \left|OO'O_g^{-1}F\right\rangle.$$

Then the map $\iota : |O'F\rangle \mapsto |O'\rangle \in \mathbb{C}[\mathrm{St}(T_N)]$ from the proof of Lem. V.2 maps the actions of G_F and $\mathrm{St}(T_N)$. Then, by surjectivity of $g \to O_g$, ι maps the actions of G_F and on $\mathcal{H}^F$, respectively, to the right and left actions of $\mathrm{St}(T_N)$ on $\mathbb{C}[\mathrm{St}(T_N)]$. These two span each others commutant. □

Appendix B: Table of symbols used

Symbol	Meaning/Definition
K^+	K vector space, K^+ non-zero vectors
T	$\mathbb{Z}_d^t$ with quadratic form $q_{r,s}$
$\mathrm{dis}(\,\cdot\,)$	Discriminant of bilinear form
$\mathrm{Arf}(\,\cdot\,)$, $\tilde{\mathrm{Arf}}(\,\cdot\,)$	Arf invariant, generalized Arf invariant of quadratic form
$\Xi(\,\cdot\,)$, $\tilde{\Xi}(\,\cdot\,)$	Polarization and generalized polarization of quadratic form
$\mathcal{G}_m$	Set of isotropic stochastic subspaces of T not containing $\mathbf{1}_t$ with dimension m
$\mathcal{G}_m^0$	Set of isotropic stochastic subspaces of T containing $\mathbf{1}_t$ with dimension m
T_N, q_N	$N^\perp/N$ with $N \subset T$ isotr. stoch., $q_{r,s}\vert_{T_N}$
C_N, $C_{\mathbf{1}_t}$	Code spaces $\mathrm{range}\, P_N$, $\mathrm{range}\, P_{\langle \mathbf{1}_t \rangle}$
$\Delta_{r,s}^{(k)}$	Rank k component of $\Delta_{r,s}$
Δ	Action of RCl on $C_{\mathbf{1}_t}$
$\mathcal{A}_{r,s}$	$\{R(O)P_N \,\vert\, N \subset T$ stoch. isotr., $O \in \mathrm{St}(T)\}$
$\mathcal{A}_{r,s}^N$	Subset $\{R(O)P_{N'} \,\vert\, N \subseteq N' \subset T$ stoch. isotr., $O \in \mathrm{St}(T)\} \subset \mathrm{A}_{r,s}$
$\mathrm{A}_{r,s}$	span $\{\mathcal{A}_{r,s}\}$ commutant algebra of $\Delta_{r,s}$
$\mathrm{A}_{r,s}^m$	span $\{P_N R(O) \,\vert\, \dim N \geq m\}$
$\mathrm{A}_{r,s}^m$	span $\{P_N R(O) \,\vert\, \dim N \geq m,\ \mathbf{1}_t \in N\}$
$\mathcal{C}_m$	span $\{C_N \,\vert\, N \in \mathcal{G}_m\}$
$\mathcal{D}_m$	span $\{C_N \,\vert\, N \in \mathcal{G}_m^0\}$
$\mathcal{K}_m$	$\mathcal{C}_m \cap \mathcal{C}_{m+1} \cap C_{\mathbf{1}_t}$
$\mathcal{L}_m$	$\mathcal{D}_m \cap \mathcal{D}_{m+1}$

[1] Ingo Roth, Richard Kueng, Shelby Kimmel, Y-K Liu, David Gross, Jens Eisert, and Martin Kliesch. Recovering quantum gates from few average gate fidelities. *Physical review letters*, 121(17):170502, 2018.

[2] Jonas Helsen, Joel J Wallman, Steven T Flammia, and Stephanie Wehner. Multiqubit randomized benchmarking using few samples. *Physical Review A*, 100(3):032304, 2019.

[3] Martin Kliesch, Richard Kueng, Jens Eisert, and David Gross. Guaranteed recovery of quantum processes from few measurements. *Quantum*, 3:171, 2019.

[4] Shelby Kimmel and Yi-Kai Liu. Phase retrieval using unitary 2-designs. In *2017 International Conference on Sampling Theory and Applications (SampTA)*, pages 345–349. IEEE, 2017.

[5] Richard Kueng, Huangjun Zhu, and David Gross. Distinguishing quantum states using clifford orbits. *arXiv preprint arXiv:1609.08595*, 2016.

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

[7] David Gross, Sepehr Nezami, and Michael Walter. Schur-Weyl duality for the Clifford group with applications: Property testing, a robust Hudson theorem, and de Finetti representations. *arXiv preprint arXiv:1712.08628*, 2017.

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

[9] Huangjun Zhu, Richard Kueng, Markus Grassl, and David Gross. The Clifford group fails gracefully to be a unitary 4-design. 2016.

[10] Jonas Haferkamp, Felipe Montealegre-Mora, Markus Heinrich, Jens Eisert, David Gross, and Ingo Roth. Quantum homeopathy works: Efficient unitary designs with a system-size independent number of non-clifford gates. *arXiv preprint arXiv:2002.09524*, 2020.

[11] Huangjun Zhu. Multiqubit Clifford groups are unitary 3-designs. *Physical Review A*, 96(6):062336, 2017.

[12] Zak Webb. The Clifford group forms a unitary 3-design. *Quant. Inf. Comp.*, 26:1379–1400, 2016.

[13] Richard Kueng and David Gross. Qubit stabilizer states are complex projective 3-designs. 2015.

[14] Eiichi Bannai, Gabriel Navarro, Noelia Rizo, and Pham Huu Tiep. Unitary t-groups. *Journal of the Mathematical Society of Japan*, 72(3):909–921, 2020.

[15] Felipe Montealegre-Mora and David Gross. Rank-deficient representations in howe duality over finite fields arise from quantum codes. *arXiv preprint arXiv:1906.07230*, 2019.

[16] Shamgar Gurevich and Roger Howe. Small representations of finite classical groups. In *Representation Theory, Number Theory, and Invariant Theory*, pages 209–234. Springer, 2017.

[17] Shamgar Gurevich and Roger Howe. Rank and duality in representation theory. *Takagi lectures*, 19:67–100, 2017.

[18] Roger Howe. Remarks on classical invariant theory. *Transactions of the American Mathematical Society*, 313(2):539–570, 1989.

[19] Masaki Kashiwara and Michele Vergne. On the segal-shale-weil representations and harmonic polynomials. *Inventiones mathematicae*, 44(1):1–47, 1978.
[20] Gabriele Nebe, Eric M. Rains, and Neil JA Sloane. The invariants of the Clifford groups. *Designs, Codes and Cryptography*, 24(1):99–122, 2001.
[21] Eiichi Bannai, Manabu Oura, and Da Zhao. The complex conjugate invariants of clifford groups. *Designs, Codes and Cryptography*, 89(2):341–350, 2021.
[22] Gabriele Nebe, Eric M Rains, and Neil James Alexander Sloane. *Self-dual codes and invariant theory*. Springer, 2006.
[23] Bernhard Runge. Codes and siegel modular forms. *Discrete Mathematics*, 148(1-3):175–204, 1996.
[24] Sepehr Nezami and Michael Walter. Multipartite entanglement in stabilizer tensor networks. *Physical Review Letters*, 125(24):241602, 2020.
[25] JM Farinholt. An ideal characterization of the clifford operators. *Journal of Physics A: Mathematical and Theoretical*, 47(30):305303, 2014.
[26] Erik Hostens, Jeroen Dehaene, and Bart De Moor. Stabilizer states and clifford operations for systems of arbitrary dimensions and modular arithmetic. *Physical Review A*, 71(4):042315, 2005.
[27] 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.
[28] David Gross. Hudson's theorem for finite-dimensional quantum systems. *J. Math. Phys.*, 47(12):122107, 2006.
[29] Huangjun Zhu. Permutation symmetry determines the discrete wigner function. *Physical review letters*, 116(4):040501, 2016.
[30] John Willard Milnor and Dale Husemoller. *Symmetric bilinear forms*, volume 73. Springer, 1973.
[31] Stephan Klaus. *Brown-Kervaire invariants*. Shaker, 1995.
[32] Robert Wilson. *The finite simple groups*, volume 251. Springer Science & Business Media, 2009.
[33] Beverley Bolt, TG Room, and GE Wall. On the Clifford collineation, transform and similarity groups. ii. *Journal of the Australian Mathematical Society*, 2(1):80–96, 1961.
[34] Marco Túlio Quintino, Qingxiuxiong Dong, Atsushi Shimbo, Akihito Soeda, and Mio Murao. Probabilistic exact universal quantum circuits for transforming unitary operations. *Physical Review A*, 100(6):062339, 2019.
[35] Jisho Miyazaki, Akihito Soeda, and Mio Murao. Complex conjugation supermap of unitary quantum maps and its universal implementation protocol. *Physical Review Research*, 1(1):013007, 2019.
[36] Roger W Carter. *Simple groups of Lie type*, volume 22. John Wiley & Sons, 1989.

3 Approximate unitary t-designs

This chapter is the preprint [HMMH$^+$20]

> Haferkamp, J., Montealegre-Mora, F., Heinrich, M., Eisert, J., Gross, D., Roth, I. (2020). *Quantum homeopathy works: Efficient unitary designs with a system-size independent number of non-Clifford gates.* arXiv preprint arXiv:2002.09524.

It has been submitted to the Journal *Communications in Mathematical Physics* and is in the peer review process.

I would like to emphasize that the main researcher in this project was Jonas Haferkamp, the first author. The project was initially phrased as an efficient construction of unitary 4-designs. In this first phase, the main statement and its proof were worked out by Haferkamp. The current version of the paper gives an efficient construction of unitary t-designs, with $t \geq 4$ generally. The proof strategy used here is, in spirit, essentially unchanged with respect to the original document dealing with 4-designs.

That said, some technical tools needed to be developed to cover the more general t case. My role in this project has been, primarily, of generalizing several technical lemmas from the $t = 4$ case to the general $t \geq 4$ case. Specifically, Lemmas 3, 4, and 13 in [HMMH$^+$20] are the outcome of discussions in which I provided crucial technical insights. These lemmas are important stepping stones in the proof of the paper's main theorem.

I have decided to include the full article in this thesis even though I am not the lead researcher in the project. This is because I believe my contributions are best understood in relation to the main theorem of [HMMH$^+$20].

Quantum homeopathy works: Efficient unitary designs with a system-size independent number of non-Clifford gates

J. Haferkamp,[1] F. Montealegre-Mora,[2] M. Heinrich,[2] J. Eisert,[1] D. Gross,[2] and I. Roth[1]

[1]*Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, Germany*
[2]*Institute for Theoretical Physics, University of Cologne, Germany*

Many quantum information protocols require the implementation of random unitaries. Because it takes exponential resources to produce Haar-random unitaries drawn from the full n-qubit group, one often resorts to t-designs. Unitary t-designs mimic the Haar-measure up to t-th moments. It is known that Clifford operations can implement at most 3-designs. In this work, we quantify the non-Clifford resources required to break this barrier. We find that it suffices to inject $O(t^4 \log^2(t) \log(1/\varepsilon))$ many non-Clifford gates into a polynomial-depth random Clifford circuit to obtain an ε-approximate t-design. Strikingly, the number of non-Clifford gates required is independent of the system size – asymptotically, the density of non-Clifford gates is allowed to tend to zero. We also derive novel bounds on the convergence time of random Clifford circuits to the t-th moment of the uniform distribution on the Clifford group. Our proofs exploit a recently developed variant of Schur-Weyl duality for the Clifford group, as well as bounds on restricted spectral gaps of averaging operators.

Random vectors and unitaries are ubiquitous in protocols and arguments of quantum information and many-body physics. In quantum information, a paradigmatic example is the *randomized benchmarking protocol* [1–3], which aims to characterize the error rate of quantum gates. There, random unitaries are used to average potentially complex errors into a single, easy to measure error rate. In many-body physics, random unitaries are used e.g. to model the dynamics that are thought to describe the mixing process that quantum information undergoes when absorbed into, and evaporated from, a black hole [4]. In these and related cases, one is faced with the issue that unitaries drawn uniformly from the full many-body group are *unphysical* in the sense that, with overwhelming probability, they cannot be implemented efficiently. The notion of a *unitary t-design* captures an efficiently realizable version of uniform randomness [5–7]. More specifically, a probability measure on the unitary group is a t-design if it matches the uniform Haar measure up to t-th moments.

Applications abound. The randomness provided by designs is used to foil attackers in quantum cryptography protocols [8–10]. It guards against worst case behavior in various quantum [10–16] and classical [17] estimation problems. Designs allow for an efficient implementation of *decoupling* procedures, a primitive in quantum Shannon theory [18]. In quantum complexity, unitary designs are used as models for generic instances of time evolution that display a quantum computational speed-up [19]. Unitary designs are now standard tools for the quantitative study of toy models in high energy physics, quantum gravity, and quantum thermodynamics [4, 20–22].

The multitude of applications motivates the search for efficient constructions of unitary t-designs [23–27]. In particular, Brandao, Harrow and Horodecki [23] show that local random circuits on n qubits with $O(n^2 t^{10})$ many gates give rise to an approximate t-design. In practice, it is often desirable to find more structured implementations. Designs consisting of *Clifford operations* would be particular attractive from various points of view: (i) Because the Clifford unitaries form a finite group, elements can be represented exactly using a small number ($O(n^2)$) of bits. (ii) The Gottesman-Knill Theorem ensures that there are efficient classical algorithms for simulating Clifford circuits. (iii) Most importantly, in *fault-tolerant architectures* [28, 29], Clifford unitaries

tend to have comparatively simple realizations, while the robust implementation of general gates (e.g. via *magic-state distillation*) carries a significant overhead. The difference is so stark that in this context, Clifford operations are often considered to be a free resource, and the complexity of a circuit is measured solely in terms of the number of non-Clifford gates [30, 31].

The Clifford group is known to form a unitary t-design for $t = 2$ [9] and $t = 3$ [32–34], but fails to have this property for $t > 3$ [32–36]. More generally, Refs. [37, 38] together imply that *any* local gate set that realizes 4-designs must necessarily be universal (c.f. Proposition 3).

This leads us to the central question underlying this work: *How many non-Clifford gates are required to generate an approximate unitary t-design?* A direct application of the random circuit model of Ref. [23] yields an estimate of $O(n^2 t^{10})$ non-Clifford operations. In this paper we show that a polynomial-sized random Clifford circuit, together with a *system size-independent* number of $O(t^4 \log^2(t))$ non-Clifford gates – a homeopathic dose – is already sufficient.

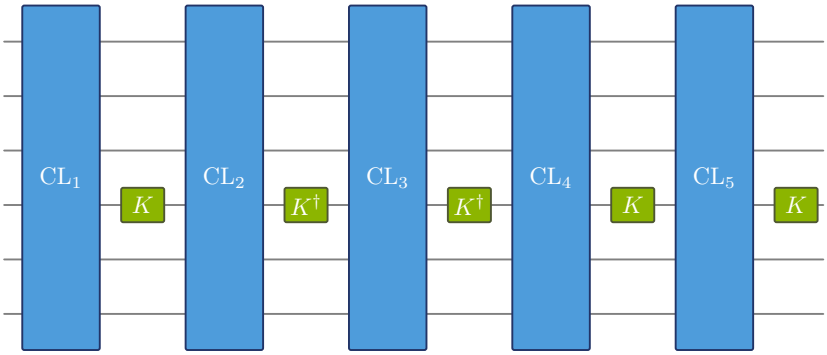

Figure 1: *K-interleaved Clifford circuits*: We consider a model where random Clifford operations are alternated with a non-Clifford gate K or its inverse $K^\dagger$.

We establish this main result for two different circuit models (Fig. 1). In Section I A, we consider alternating unitaries drawn uniformly from the Clifford group with a non-Clifford gate. This gives rise to an efficient quantum circuit, as there are classical algorithms for sampling uniformly from the Clifford group, and for producing an efficient gate decomposition of the resulting operation [39]. A somewhat simpler model is analyzed in Section I B. There, we assume that the Clifford layers are circuits consisting of gates drawn form a local Clifford gate set. These circuits will only approximate the uniform measure on the Clifford group. Theorem 3, which might be of independent interest, gives novel bounds on the convergence rate.

The key to this scaling lies in the structure of the commutant of the t-th tensor power of the Clifford group, described by a variant of Schur-Weyl duality developed in a sequence of recent works [35, 40–42]. There, it has been shown that the dimension of this commutant – which measures the failure of the Clifford group to be a t-design from a representation theoretical perspective – is independent of the system size. Refs. [35, 41] have used this insight to provide a construction for exact *spherical* t-designs that consist of a system size-independent number of Clifford orbits. It has been left as an open problem whether these ideas can be generalized from spherical designs to the more complex notion of unitary designs, and whether the construction can be made efficient [41]. The present work resolves this question in the affirmative.

Finally, we note that in Ref. [43], it has been observed numerically that adding a single T gate to a random Clifford circuit has dramatic effects on the entanglement spectrum. A relation to t-designs was suspected. Our result provides a rigorous understanding of this observation.

I. RESULTS

A. Approximate t-designs with few non-Clifford gates

To state our results precisely, we need to formalize the relevant notion of approximation, as well as the circuit model used. Let ν be a probability measure on the unitary group $U(d)$. The measure ν gives rise to a quantum channel

$$\Delta_t(\nu)(\rho) := \int_{\mathrm{U}(d)} U^{\otimes t} \rho \left(U^\dagger\right)^{\otimes t} \mathrm{d}\nu(U), \tag{1}$$

which applies $U^{\otimes t}$, with U chosen according to ν. We will refer to $\Delta_t(\nu)$ as the *t-th moment operator* associated with ν. Following Ref. [26], we quantify the degree to which a measure approximates a t-design by the diamond norm distance of its moment operator to the moment operator of the Haar measure μ_{H} on $U(d)$.

Definition 1 (Approximate unitary design)**.** *Let ν be a distribution on $\mathrm{U}(d)$. Then ν is an (additive) ε-approximate t-design if*

$$\|\Delta_t(\nu) - \Delta_t(\mu_{\mathrm{H}})\|_\diamond \leq \varepsilon. \tag{2}$$

Denote the uniform measure on the multiqubit Clifford group $\mathrm{Cl}(2^n)$ by μ_{Cl}, and let K be some fixed single-qubit non-Clifford gate. The circuit model we are considering (Figure 1) interleaves Clifford unitaries drawn from μ_{Cl}, with random gates from $\{K, K^\dagger, \mathbb{1}\}$ acting on an arbitrary qubit[1]. Note that the concatenation of two unitaries drawn from measures ν_1 and ν_2 is described by the convolution $\nu_1 * \nu_2$ of the respective measures. We thus arrive at this formal definition of the circuit model:

Definition 2 (K-interleaved Clifford circuits)**.** *Let $K \in U(2)$. Consider the probability measure ξ_K that draws uniformly from the set $\{K \otimes \mathbb{1}_{2^{n-1}}, K^\dagger \otimes \mathbb{1}_{2^{n-1}}, \mathbb{1}_{2^n}\}$. A* K-interleaved Clifford circuit of depth k *is the random circuit acting on n qubits described by the probability distribution*

$$\sigma_k := \underbrace{\mu_{\mathrm{Cl}} * \xi_K * \cdots * \mu_{\mathrm{Cl}} * \xi_K}_{k \text{ times}}. \tag{3}$$

For convenience, we work with the logarithm of base 2: $\log(x) := \log_2(x)$. We are now equipped to state the main result of this work in the form of a theorem:

Theorem 1 (Unitary designs with few non-Clifford gates)**.** *Let $K \in U(2)$ be a non-Clifford unitary. There are constants $C_1(K), C_2(K)$ such that for any $k \geq C_1(K) \log^2(t)(t^4 + t\log(1/\varepsilon))$, a K-interleaved Clifford circuit with depth k acting on n qubits is an additive ε-approximate t-design for all $n \geq C_2(K)t^2$.*

We give the proofs of this theorem in Section III. In Theorem 1, we consider uniformly drawn multiqubit Clifford unitaries. This can be achieved with $O(n^3)$ classical random bits [39] and then implemented with $O(n^2/\log(n))$ gates [44]. Combined with these results, Theorem 1 implies an overall gate count of $O(n^2/\log(n)t^4\log^2(t))$ improving the scaling compared to Ref. [23] in the

[1] We use the set $\{K, K^\dagger, \mathbb{1}\}$ instead of just $\{K\}$ for technical reasons: Making the set closed under the adjoint causes the moment operator to be Hermitian. The identity is included to ensure that the concatenation of two random elements has a non-vanishing probability of producing a non-Clifford gate—a property that will slightly simplify the proof. Of course, in a physical realization, identity gates and the following Clifford operation are redundant and need not be implemented.

dependence on both t and n. In this sense, our construction can be seen as a classical-quantum hybrid construction of unitary designs: The scaling is significantly improved by outsourcing as many tasks as possible to a classical computer. A construction in which all parts of the random unitary are local random circuits is considered in Corollary 3.

For designs generated from general random local circuits, numerical results suggest that convergence is much faster in practice than indicated by the proven bounds [45]. We expect that a similar effect occurs here, and that in fact very shallow K-interleaved Clifford circuits are sufficient to approximate t-designs. This intuition is supported by the numerical results of Ref. [43], which show that even a single T-gate has dramatic effects on the entanglement spectrum of a quantum circuit.

It is moreover noteworthy that circuits with few T-gates can be efficiently simulated [46–49]. The scaling of these algorithms is polynomial in the depth of the circuit, but exponential in the number of T-gates. Combined with our result, this implies that for fixed additive errors ε, there are families of ε-approximate unitary $O(\log(n))$-designs simulable in quasi-polynomial time. For the general random quantum circuit model, it is conjectured that a depth of order $O(nt)$ suffices to approximate t-designs [23, 50]. If such a linear scaling is sufficient in our model, the quasi-polynomial time estimate for classical simulations would improve to polynomial.

Conversely, our result has implications for the circuit complexity of quantum circuits dominated by Clifford gates. For this we combine Theorem 1 with a recent connection between unitary designs and complexity [50]. Consider the following (informal) definition of circuit complexity:

Definition 3 ([50])**.** *The complexity of a quantum circuit U is the minimal circuit size required to implement an ancilla-assisted measurement that is capable of distinguishing the map $\rho \mapsto U\rho U^\dagger$ from the completely depolarizing channel $\rho \mapsto \frac{1}{d}\mathbb{1}$.*

Then, the following result holds:

Theorem 2 ([50], informal)**.** *Consider an approximate unitary t-design. Then, a randomly selected element is very likely to have strong circuit complexity $\approx t$.*

Combined with Theorem 1, this yields an immediate corollary:

Corollary 1 (Circuit complexity scaling)**.** *A K-interleaved Clifford circuit of depth $k = O(t^4)$ contains at least $\exp(\Omega(k))$ elements with strong complexity $\Omega(k^{\frac{1}{4}})$, provided that $n \geq C_2(K)t^2$.*

In particular, the circuit complexity scales directly with the number of non-Clifford gates independent of the system size. The system size only enters in the number of Clifford gates.

For the proof of Theorem 1 we need to analyse the connection between the t-th moment operator of the Haar measure and the commutant of the diagonal action of the Clifford group. The latter was proven to be spanned by representations of so-called *stochastic Lagrangian subspaces* in Ref. [41]. In particular, we prove almost tight bounds on the overlap of the Haar operator with these basis vectors in Lemma 13 that might be of independent interest. This will allow us to invoke a powerful theorem by Varjú [51] on restricted spectral gaps of probability distributions on compact Lie groups to show that non-Clifford unitaries have a strong impact on representations of Lagrangian subspaces that are not also permutations. We combine this insight with a careful combinatorial argument about the Gram-Schmidt orthogonalization of the basis corresponding to stochastic Lagrangian subspaces to bound the difference to a unitary t-design in diamond norm.

Moreover, the bound for Theorem 1 allows us to prove a corollary about the stronger notion of *relative approximate designs*:

Definition 4 (Relative ε-approximate t-design)**.** *We call a probability ν a relative ε-approximate t-design if*

$$(1-\varepsilon)\Delta_t(\nu) \preccurlyeq \Delta_t(\mu_{\mathrm{H}}) \preccurlyeq (1+\varepsilon)\Delta_t(\nu), \tag{4}$$

where $A \preccurlyeq B$ if and only if $B - A$ is completely positive.

Corollary 2 (K-interleaved Clifford circuits as relative ε-approximate t-designs)**.** *There are constants $C_1'(K), C_2'(K)$ such that a K-interleaved Clifford circuit is a relative ε-approximate t-design in depth $k \geq C_1'(K)\log^2(t)(2nt + \log(1/\varepsilon))$ for all $n \geq C_2'(K)t^2$.*

Hence, if we drop the system-size independence, we can achieve a scaling of $O(nt)$ at least until $t \sim \sqrt{n}$.

While we believe the setting of K-interleaved Clifford circuits to be the more relevant case, the same method of proof works for *Haar*-interleaved Clifford circuits. Here, we draw not from the gate set $\{K_i, K_i^\dagger, \mathbb{1}\}$, but instead Haar-randomly from $U(2)$. The advantage is that we obtain explicit constants for the depth, while the depth in the K-interleaved setting has to depend on a constant (as K might be arbitrarily close to the identity).

Proposition 1 (Haar-interleaved Clifford circuits as relative ε-approximate t-designs)**.** *For $k \geq 36(33t^4 + 3t\log(1/\varepsilon))$, Haar-interleaved Clifford circuits with depth k form an additive ε-approximate t-design for all $n \geq 32t^2 + 7$.*

Similarly, variants of Corollary 2 for Haar-interleaved Clifford circuits can be obtained, here also without the $\log^2(t)$ dependence.

B. Local random Clifford circuits for Clifford and unitary designs

The circuits considered in the previous section require one to find the gate decomposition of a random Clifford operation. In this section, we analyze the case where the Clifford layers are circuits consisting of gates drawn from a local set of generators.

As a first step, we establish that a 2-local random Clifford circuit on n qubits of depth $O(n^2t^9\log^{-2}(t)\log(1/\varepsilon))$ constitutes a relative ε-approximate Clifford t-design, i.e., reproduces the moment operator of the Clifford group up to the t-th order with a relative error of ε. We consider local random Clifford circuits that consist of 2-local quantum gates from a finite set G with is closed under taking the inverse and generates $\mathrm{Cl}(4)$. We refer to such a set as a *closed, generating set*. A canonical example for such a closed, generating set is $\{H \otimes \mathbb{1}, S \otimes \mathbb{1}, S^3 \otimes \mathbb{1}, \mathrm{CX}\}$ where H is the Hadamard gate, S is the phase gate and CX is the cNOT-gate [52]. Such a set G induces a set of multi-qubit Clifford unitaries $\hat{G} \subset \mathrm{Cl}(n)$ by acting on any pair of adjacent qubits, where we adopt periodic boundary conditions. We then define the corresponding random Clifford circuits.

Definition 5 (Local random Clifford circuit)**.** *Let $G \subset \mathrm{Cl}(4)$ be a closed, generating set. Define the probability measure σ_G as the measure having uniform support on $\hat{G} \subset \mathrm{Cl}(n)$ acting on n qubits. A* local random Clifford circuit *of depth m is the random circuits described by the probability measure σ_G^{*m}.*

Our result on local random Clifford circuits even holds for a stronger notion for approximations of designs, namely relative approximate designs. Write $A \preccurlyeq B$ if $B - A$ is positive semi-definite.

Definition 6 (Relative approximate Clifford t-designs)**.** *Let ν be a probability measure on $\mathrm{Cl}(2^n)$. Then, ν is a* relative ε-approximate Clifford t-design *if*

$$(1-\varepsilon)\Delta_t(\mu_{\mathrm{Cl}}) \preccurlyeq \Delta_t(\nu) \preccurlyeq (1+\varepsilon)\Delta_t(\mu_{\mathrm{Cl}}). \tag{5}$$

With this definition, our result reads as follows.

Theorem 3 (Local random Clifford designs)**.** *Let $n \geq 12t$, then a local random Clifford circuit of depth $O(n \log^{-2}(t) t^8 (2nt + \log(1/\varepsilon)))$ constitutes a relative ε-approximate Clifford t-design.*

The proof of the theorem is given in Section IV. This result is a significant improvement over the scaling of $O(n^8)$, which is implicit in Ref. [9].

We can combine this result with the bounds obtained in Section III. To this end, consider a random circuit that k-times alternatingly applies a local random Clifford circuit of depth m, and a unitary drawn from the probability measure ξ_K. The corresponding probability measure is

$$\sigma_{k,m} := \underbrace{\sigma_G^{*m} * \xi_K * \cdots * \sigma_G^{*m} * \xi_K}_{k \text{ times}}. \tag{6}$$

For these local random circuits we establish the following result:

Corollary 3 (Local random unitary design)**.** *Let $K \in U(2)$ be a non-Clifford gate and let $G \subset \mathrm{Cl}(4)$ be a closed, generating set. There are constants $C_1''(K,G), C_2''(K), C_3''(K)$ such that whenever*

$$m \geq C_1''(K,G) n \log^{-2}(t) t^8 \left(2nt + \log(1/\varepsilon)\right) \quad \textit{and} \quad k \geq C_2''(K) \log^2(t)(t^4 + t\log(1/\varepsilon)),$$

the local random circuit $\sigma_{k,m}$, defined in (6)*, is an ε-approximate unitary t-design for all $n \geq C_3''(K) t^2$.*

The complete argument for the corollary is given at the end of Section IV. After introducing technical preliminaries in Section II, the remainder of the paper, Section III and Section IV, is devoted to the proofs of Theorem 1, Theorem 3 and the Corollary 3. Finally, in Section V we elaborate on and formalize as Proposition 3 the observation that there exists no non-universal family of exact 4-designs for arbitrary system size.

II. TECHNICAL PRELIMINARIES

A. Operators and superoperators

Given a (finite-dimensional) Hilbert space $\mathcal{H}$, we denote with $L(\mathcal{H})$ the space of linear operators on $\mathcal{H}$ with involution $\dagger$ mapping an operator to its adjoint with respect to the inner product on $\mathcal{H}$. $L(\mathcal{H})$ naturally inherits a Hermitian inner product, the *Hilbert-Schmidt inner product*

$$(A|B) := \mathrm{Tr}(A^\dagger B), \qquad \forall A, B \in L(\mathcal{H}). \tag{7}$$

As this definition already suggests, we will use "operator kets and bras" whenever we think it simplifies the notation. Concretely, we write $|B) = B$ and denote with $(A|$ the linear form on $L(\mathcal{H})$ given by

$$(A| : \ B \longmapsto (A|B). \tag{8}$$

Following common terminology in quantum information theory, we call linear maps $\phi : \ L(\mathcal{H}) \to L(\mathcal{H})$ on operators "superoperators". We use $\phi^\dagger$ to denote the adjoint map with respect to the Hilbert-Schmidt inner product. Note that with the above notation, $\phi = |A)(B|$ defines a rank one superoperator with $\phi^\dagger = |B)(A|$. Moreover, we will denote by the superoperator $\mathrm{Ad}_A := A \cdot A^{-1}$

the *adjoint action* of an invertible operator $A \in \mathrm{GL}(\mathcal{H})$ on $L(\mathcal{H})$. For notational reasons, we sometimes write $\mathrm{Ad}(A)$ instead of Ad_A.

We consistently reserve the notation $\|\cdot\|_p$ for the Schatten p-norms

$$\|A\|_p := \mathrm{Tr}(|A|^p)^{1/p} = \|\sigma(A)\|_{\ell_p}\,, \tag{9}$$

where $\sigma(A)$ is the vector of singular values of A. In particular, we use the *trace norm* $p = 1$, the *Frobenius* or *Hilbert-Schmidt norm* $p = 2$ and the *spectral norm* $p = \infty$. Clearly, this norms can be defined for both operators and superoperators and we will use the same symbol in both cases. For the latter, however, there is also a family of induced operator norms

$$\|\phi\|_{p\to q} := \sup_{\|X\|_p \leq 1} \|\phi(X)\|_q\,. \tag{10}$$

Note that $\|\cdot\|_{2\to 2} \equiv \|\cdot\|_\infty$. Finally, we are interested in "stabilized" versions of these induced norms, in particular the *diamond norm*

$$\|\phi\|_\diamond := \sup_{d\in\mathbb{N}} \left\|\phi \otimes \mathrm{id}_{L(\mathbb{C}^d)}\right\|_{1\to 1} = \left\|\phi \otimes \mathrm{id}_{L(\mathcal{H})}\right\|_{1\to 1}\,. \tag{11}$$

The following norm inequality will be useful [53]

$$\|\phi\|_\diamond \leq (\dim \mathcal{H})^2 \|\phi\|_\infty\,, \qquad \|\phi\|_\infty \leq \sqrt{\dim \mathcal{H}}\, \|\phi\|_\diamond\,. \tag{12}$$

B. Commutant of the diagonal representation of the Clifford group

In this section, we review some of the machinery developed in Ref. [41]. Recall that the n-qubit *Clifford group* $\mathrm{Cl}(n)$ is defined as the unitary normalizer of the Pauli group $\mathcal{P}_n$:

$$\mathrm{Cl}(n) = \left\{U \in U(2^n, \mathbb{Q}[i]) \;\middle|\; U\mathcal{P}_n U^\dagger \subset \mathcal{P}_n\right\}. \tag{13}$$

Here, we followed the convention to restrict the matrix entries to rational complex numbers. This avoids the unnecessary complications from an infinite center $U(1)$ yielding a finite group with minimal center $Z(\mathrm{Cl}(n)) = Z(\mathcal{P}_n) \simeq \mathbb{Z}_4$. The Clifford group can equivalently be defined in a less conceptual but more constructive manner: It is the subgroup of $\mathrm{U}(2^n)$ generated by CX, the controlled not gate, the Hadamard gate H and the phase gate S.

For this work, the t-th diagonal representation of the Clifford group, defined as

$$\tau^{(t)}:\ \mathrm{Cl}(n) \longrightarrow \mathrm{U}(2^{nt}), \quad U \longmapsto U^{\otimes t}, \tag{14}$$

will be of major importance. It acts naturally on the Hilbert space $((\mathbb{C}^2)^{\otimes n})^{\otimes t}$ which can be seen as t copies of a n-qubit system. However, it will turn out that the operators commuting with this representation naturally factorize with respect to a different tensor structure on this Hilbert space, namely $((\mathbb{C}^2)^{\otimes t})^{\otimes n} \simeq ((\mathbb{C}^2)^{\otimes n})^{\otimes t}$. Because of the different exponents, it should be clear from the context which tensor structure is meant. We will make ubiquitous use of the description of the commutant of the diagonal representation in terms of *stochastic Lagrangian subspaces* [41]:

Definition 7 (Stochastic Lagrangian subspaces)**.** *Consider the quadratic form* $\mathfrak{q} : \mathbb{Z}_2^{2t} \to \mathbb{Z}_4$ *defined as* $\mathfrak{q}(x, y) := x \cdot x - y \cdot y \mod 4$*. The set* $\Sigma_{t,t}$ *denotes the set of all subspaces* $T \subseteq \mathbb{Z}_2^{2t}$ *being subject to the following properties:*

1. *T is totally* $\mathfrak{q}$*-isotropic:* $x \cdot x = y \cdot y \mod 4$ *for all* $(x, y) \in T$.

2. *T has dimension* t *(the maximum dimension compatible with total isotropicity).*

3. *T is* stochastic*:* $(1, \ldots, 1) \in T$.

We call elements in $\Sigma_{t,t}$ *stochastic Lagrangian subspaces*. We have

$$|\Sigma_{t,t}| = \prod_{k=0}^{t-2}(2^k + 1) \leq 2^{\frac{1}{2}(t^2+5t)}. \tag{15}$$

With this notion, we can now state the following key theorem from Ref. [41].

Theorem 4 ([41])**.** *If* $n \geq t-1$*, then the commutant* $\tau^{(t)}(\mathrm{Cl}(n))'$ *of the* t*-th diagonal representation of the Clifford group is spanned by the linearly independent operators* $r(T)^{\otimes n}$*, where* $T \in \Sigma_{t,t}$ *and*

$$r(T) := \sum_{(x,y)\in T} |x\rangle\langle y|. \tag{16}$$

Since the representation in question is fixed throughout this paper, we will simplify the notation from now on and write $\mathrm{Cl}(n)' \equiv \tau^{(t)}(\mathrm{Cl}(n))'$. To make use of a more sophisticated characterization of the elements $r(T)$ developed in Ref. [41, Sec. 4], we need the following definitions.

Definition 8 (Stochastic orthogonal group)**.** *Consider the quadratic form* $q : \mathbb{Z}_2^t \to \mathbb{Z}_4$ *defined as* $q(x) := x \cdot x \mod 4$*. The* stochastic orthogonal group O_t *is defined as the group of* $t \times t$ *matrices* O *with entries in* $\mathbb{Z}_2$ *such that*

1. $q(Ox) = q(x)$ *for all* $x \in \mathbb{Z}_2^t$ *and*

2. $O(1, \ldots, 1)^T = (1, \ldots, 1)^T \mod d$*.*

The subspace $T_O := \{(Ox, x), x \in \mathbb{Z}_2^t\}$ is a stochastic Lagrangian subspace. Moreover, the operator $r(O) := r(T_O)$ is unitary. With this notion and the next one, all stochastic Lagrangian subspaces can be characterized.

Definition 9 (Defect subspaces)**.** *A defect subspace is a subspace* $N \subseteq \mathbb{Z}_2^t$ *subject to the following conditions:*

1. $q(x) = 0$ *for all* $x \in N$.

2. $(1, \ldots, 1)^T \in N^\perp$.

Here, $N^\perp = \{y \in \mathbb{Z}_2^t \,|\, x \cdot y = 0 \; \forall x \in N\}$.

First, note that N is totally isotropic, i.e. $N \subseteq N^\perp$. Moreover, we have that $\dim N \leq t/2$ for all defect subspaces N. Spaces $N \subseteq \mathbb{Z}_2^t$ that satisfy the first condition define *Calderbank-Shor-Sloane (CSS)* codes

$$\mathrm{CSS}(N) := \{Z(p)X(q) \mid q, p \in N\}, \tag{17}$$

where the action of the multi-qubit Pauli operators is $Z(p)\,|x\rangle := (-1)^{p\cdot x}\,|x\rangle$ and $X(q)\,|x\rangle := |x + q\rangle$ for $x \in \mathbb{Z}_2^t$. The corresponding projector is given by

$$P_N := P_{\mathrm{CSS}(N)} = \frac{1}{|N|^2}\sum_{q,p\in N} Z(p)X(q). \tag{18}$$

Since the order of the stabilizer group is $2^{2\dim N}$, P_N projects onto a $2^{t-2\dim N}$-dimensional subspace of $(\mathbb{C}^2)^{\otimes t}$. For $N = \{0\}$ we set $P_{\mathrm{CSS}(N)} := \mathbb{1}$. We summarize the findings of [41, Sec. 4] as follows:

Theorem 5 ([41]). *Consider $T \in \Sigma_{t,t}$, then*

$$r(T) = 2^{\dim N} r(O) P_{\mathrm{CSS}(N)} = 2^{\dim N'} P_{\mathrm{CSS}(N')} r(O') \tag{19}$$

for $O, O' \in O_t$ and N, N' defect subspaces with $\dim N = \dim N'$.

Lemma 1 (Norms of $r(T)$). *Suppose $r(T) = 2^{\dim N} r(O) P_N$ as in Theorem 5. Then it holds:*

$$\|r(T)\|_1 = 2^{t-\dim N}, \qquad \|r(T)\|_2 = 2^{t/2}, \qquad \|r(T)\|_\infty = 2^{\dim N}. \tag{20}$$

Proof. Since any Schatten p-norm is unitarily invariant, we have $\|r(T)\|_p = 2^{\dim N} \|P_N\|_p$. The statements follow from $\operatorname{rank} P_N = 2^{t-2\dim N}$. □

In the following, we will often work with a normalized version of the $r(T)$ operators which we define as

$$Q_T := \frac{r(T)}{\|r(T)\|_2} = 2^{-t/2} r(T). \tag{21}$$

III. APPROXIMATE UNITARY t-DESIGNS

In this section, we give a bound on the number of non-Clifford gates needed to leverage the Clifford group to an approximate unitary t-design. This is made precise by the following two theorems which rely on two distinct proof strategies and come with different trade-offs.

Theorem 1 (Unitary designs with few non-Clifford gates). *Let $K \in U(2)$ be a non-Clifford unitary. There are constants $C_1(K), C_2(K)$ such that for any $k \geq C_1(K) \log^2(t)(t^4 + t\log(1/\varepsilon))$, a K-interleaved Clifford circuit with depth k acting on n qubits is an additive ε-approximate t-design for all $n \geq C_2(K) t^2$.*

Recall from Def. 2 that a K-interleaved Clifford circuit has an associated probability measure $\sigma_K := (\mu_{\mathrm{Cl}} * \xi_K)^{*k}$ where ξ_K is the measure which draws uniformly from $\{K, K^\dagger, \mathbb{1}\}$ on the first qubit. Let us introduce the notation

$$\mathrm{R}(K) := \int_{U(2^n)} \mathrm{Ad}_U^{\otimes t} \, \mathrm{d}\xi_k(U) = \frac{1}{3}\left(\mathrm{Ad}_K^{\otimes t} + \mathrm{Ad}_{K^\dagger}^{\otimes t} + \mathrm{id}\right) \otimes \mathrm{id}_{n-1}. \tag{22}$$

Then, our goal is to bound the deviation of the moment operator

$$\Delta_t(\sigma_k) = \int_{U(2^n)} \mathrm{Ad}_U^{\otimes t} \, \mathrm{d}\sigma_k(U) = \underbrace{\Delta_t(\mu_{\mathrm{Cl}})\mathrm{R}(K) \ldots \Delta_t(\mu_{\mathrm{Cl}})\mathrm{R}(K)}_{k \text{ times}}, \tag{23}$$

from the Haar projector $P_{\mathrm{H}} \equiv \Delta_t(\mu_{\mathrm{H}})$ in diamond norm. Using that P_{H} is invariant under left and right multiplication with unitaries, we have the identity

$$A^k - P_{\mathrm{H}} = (A - P_{\mathrm{H}})^k, \tag{24}$$

for any mixed unitary channel A. Thus, we can rewrite the difference of moment operators as

$$\Delta_t(\sigma_k) - P_{\mathrm{H}} = [P_{\mathrm{Cl}}\mathrm{R}(K)]^k - P_{\mathrm{H}} = \left[(P_{\mathrm{Cl}} - P_{\mathrm{H}})\,\mathrm{R}(K)\right]^k, \tag{25}$$

where we introduced the shorthand notation $P_{\mathrm{Cl}} := \Delta_t(\mu_{\mathrm{Cl}})$.

Remark 1 (Non-vanishing probability of applying the identity)**.** *We apply K, $K^\dagger$ with equal probability in Theorem 1 such that $R(K)$ is Hermitian. The non-vanishing probability of applying $\mathbb{1}$, i.e., of doing nothing, is necessary in the proof of Lemma 2, because we require the probability distribution $\tilde{\xi}_K * \xi_K = \xi_K * \xi_K$ to have non-vanishing support on a non-Clifford gate. If ξ_K is the uniform measure on K and $K^\dagger$, then $\xi_K * \xi_K$ has support on K^2, $(K^\dagger)^2$ and $\mathbb{1}$. We can hence drop this assumption for gates that do not square to a Clifford gate. This is not the case for e.g. the T-gate.*

Our proof strategy for Theorem 1 makes use of the following two lemmas which are proven in Sec. VI A and VI B. The first lemma is key to the derivations in this section. It is based on a bound (Lemma 13) on the overlap of stochastic Lagrangian subspaces with the Haar projector and Theorem 6, a special case of a theorem about restricted spectral gaps of random walks on compact Lie groups due to Varjú [51].

Lemma 2 (Overlap bound)**.** *Let K be a single qubit gate which is not contained in the Clifford group. Then, there is a constant $c(K) > 0$ such that*

$$\eta_{K,t} := \max_{\substack{T \in \Sigma_{t,t} - S_t \\ T' \in \Sigma_{t,t}}} \frac{1}{3} \left| (Q_T| \operatorname{Ad}_K^{\otimes t} + \operatorname{Ad}_{K^\dagger}^{\otimes t} + \mathrm{id}\, |Q_{T'}) \right| \leq 1 - c(K) \log^{-2}(t). \tag{26}$$

The second lemma is of a more technical nature.

Lemma 3 (Diamond norm bound)**.** *Consider $T_1, T_2 \in \Sigma_{t,t}$ and denote with N_1, N_2 their respective defect spaces. Then, it holds that*

$$\| |Q_{T_1})(Q_{T_2}| \|_\diamond \leq 2^{\dim N_2 - \dim N_1}, \tag{27}$$

$$|\, (Q_{T_1}|Q_{T_2})\, | \leq 2^{-|\dim N_1 - \dim N_2|}. \tag{28}$$

The difficulty of using these results to bound the difference

$$\Delta_t(\sigma_k) - P_{\mathrm{H}} = \left[(P_{\mathrm{Cl}} - P_{\mathrm{H}})\, \mathrm{R}(K) \right]^k, \tag{29}$$

stems from the following reason: The range of the projector $P_{\mathrm{Cl}} - P_{\mathrm{H}}$ is the ortho-complement of the space spanned by permutations $Q_\pi^{\otimes n}$ for $\pi \in S_t$ within the commutant of the Clifford group spanned by the operators $Q_T^{\otimes n}$. Although this is a conveniently factorizing and well-studied basis, it is *non-orthogonal*. Thus, the projectors do not possess a natural expansion in this basis and we can not directly use the above bounds. However, we can write it explicitly in a suitable orthonormal basis of the commutant obtained by the Gram-Schmidt procedure from the basis $\{Q_T^{\otimes n} \,|\, T \in \Sigma_{t,t}\}$. We summarize the properties of this basis in the following lemma:

Lemma 4 (Properties of the constructed basis)**.** *Let $\{T_j\}_{j=1}^{|\Sigma_{t,t}|}$ be an enumeration of the elements of $\Sigma_{t,t}$ such that the first $t!$ spaces T_j correspond to the elements of S_t. Then, the $\{E_j\}$ constitutes an orthogonal (but not normalized) basis, where*

$$E_j := \sum_{i=1}^{j} A_{i,j}\, Q_{T_i}^{\otimes n} := \sum_{i=1}^{j} \left[\sum_{\substack{\Pi \in S_j \\ \Pi(j) = i}} \operatorname{sign}(\Pi) \prod_{l=1}^{j-1} \left(Q_{T_l} \middle| Q_{T_{\Pi(l)}} \right)^n \right] Q_{T_i}^{\otimes n}. \tag{30}$$

Denote by N_i the defect space of T_i. For $n \geq \frac{1}{2}(t^2+5t)$, we have

$$|A_{i,j}| \leq 2^{t^3+4t^2+6t-n|\dim N_i - \dim N_j|}, \qquad \forall i,j, \tag{31}$$

$$|A_{i,j}| \leq 2^{2t^2+10t-n}, \qquad \forall i \neq j. \tag{32}$$

Moreover, it holds that

$$1 - 2^{t^2+7t-n} \leq A_{j,j} \leq 1 + 2^{t^2+7t-n}. \tag{33}$$

We believe that the explicit bounds in Lemma 4 might be of independent interest in applications of the Schur-Weyl duality of the Clifford group. For the sake of readibility, and as Theorem 1 holds up to an inexplicit constant, we will bound all polynomials in t by their leading order term in the following.

Proof of Theorem 1. Notice that from (25), we have the expression

$$\|[P_{\mathrm{Cl}}\mathrm{R}(K)]^k - P_{\mathrm{H}}\|_\diamond \tag{34}$$

$$= \left\| \left[\left(\sum_{j=t!+1}^{|\Sigma_{t,t}|} \frac{1}{(E_j|E_j)} \, |E_j)(E_j| \right) \mathrm{R}(K) \right]^k \right\|_\diamond \tag{35}$$

$$= \left\| \sum_{j_1,\ldots,j_m=t!+1}^{|\Sigma_{t,t}|} \prod_{l=1}^{k} \frac{1}{(E_{j_l}|E_{j_l})} \, |E_{j_1}) \, (E_{j_1}| \, R(K) \, |E_{j_2}) \ldots (E_{j_k}| \, R(K) \right\|_\diamond \tag{36}$$

$$\leq \sum_{j_1,\ldots,j_k=t!+1}^{|\Sigma_{t,t}|} \prod_{l=1}^{k} \frac{1}{(E_{j_l}|E_{j_l})} \prod_{r=1}^{k-1} |\, (E_{j_r}| \, R(K) \, \left|E_{j_{r+1}}\right) | \cdot \left\| \, |E_{j_1})(E_{j_k}| \, \right\|_\diamond . \tag{37}$$

We now bound each of the factors in each term above.

First, we compute the squared norm of $|E_j)$,

$$(E_j|E_j) = \sum_{r,l=1}^{j} A_{rj}A_{lj} \, (Q_{T_r}|Q_{T_l})^n = A_{j,j}^2 + \sum_{k,l<j} A_{rj}A_{lj} \, (Q_{T_k}|Q_{T_l})^n \, . \tag{38}$$

Using eqs. (32) and (33), we thus bound

$$\begin{aligned} (E_j|E_j) &\leq \left(1 + 2^{t^2+7t-n}\right)^2 + (j^2-1)4^{2t^2+10t-n} \\ &\leq \left(1 + 2^{t^2+7t-n}\right)^2 + |\Sigma_{t,t}|^2 4^{2t^2+10t-n} \\ &\leq 1 + 2^{31t^2-2n}, \end{aligned} \tag{39}$$

and in the same way

$$(E_j|E_j) \geq 1 - 2^{31t^2-2n}. \tag{40}$$

Now we use that $n \geq 16t^2$. Letting $x := 2^{31t^2-2n} \in [0, \frac{1}{2}]$, the inequalities $1/(1-x) \leq 1+2x$ and $1-2x \leq 1/(1+x)$ hold. This leads to

$$\frac{1}{(E_j|E_j)} = 1 + a_j \qquad \text{with} \qquad |a_j| \leq 2^{32t^2-2n}. \tag{41}$$

We now focus on the second factor,

$$|\,(E_i|\,R(K)\,|E_j)\,| \leq \sum_{r=1}^{i}\sum_{l=1}^{j} |A_{r,i}A_{l,j}| \cdot \left|\left(Q_{T_r}^{\otimes n}\right|\,R(K)\,\left|Q_{T_l}^{\otimes n}\right)\right|. \tag{42}$$

If for $(Q_{T_r}|\,R(K)\,|Q_{T_l})$ one of the stochastic Lagrangian subspaces does not correspond to a permutation, Lemma 2 introduces a factor of $\eta_{K,t}$. If both correspond to a permutation, for the sake of beauty we redefine the factors $A_{r,i}$ and $A_{l,j}$ by multiplying it with 2, and compensate this by introducing a factor of $\frac{1}{4}$ and let

$$\bar{\eta}_{K,t} := \max\left\{\frac{1}{4}, \eta_{K,t}\right\}. \tag{43}$$

In this case $r < t! + 1 \leq i$ and $l < t! + 1 \leq j$, so the factor $|A_{r,i}A_{l,j}|$ will be exponentially suppressed according to (32) and so this redefinition will not affect the asymptotic scaling in n.

We provide two bounds for $|\,(E_i|\,R(K)\,|E_j)\,|$ that will be used later on. First, using (31), (33) and (28), we obtain

$$\begin{aligned}
|\,(E_i|\,R(K)\,|E_j)\,| &\leq \sum_{r=1}^{i}\sum_{l=1}^{j} |A_{r,i}A_{l,j}| \cdot \left|\left(Q_{T_r}^{\otimes n}\right|\,R(K)\,\left|Q_{T_l}^{\otimes n}\right)\right| && (44)\\
&\leq \bar{\eta}_{K,t}(1+2^{8t^2-n}) \sum_{r=1}^{i}\sum_{l=1}^{j} 2^{24t^3 - n|\dim N_r - \dim N_i| - n|\dim N_l - \dim N_j| - (n-1)|\dim N_l - \dim N_r|} && (45)\\
&\leq \bar{\eta}_{K,t}(1+2^{8t^2-n})|\Sigma_{t,t}|^2 2^{25t^3 - n|\dim N_j - \dim N_i|} && (46)\\
&\leq \bar{\eta}_{K,t}(1+2^{8t^2-n})2^{31t^3 - n|\dim N_j - \dim N_i|}, && (47)
\end{aligned}$$

where we have used $2^{|\dim N_l - \dim N_r|} \leq 2^t$.

The second bound follows from equations (32) and (33), and we consider two cases. If $i \neq j$, then

$$\begin{aligned}
|\,(E_i|\,R(K)\,|E_j)\,| &\leq \sum_{r=1}^{i}\sum_{l=1}^{j} |A_{r,i}A_{l,j}| \cdot |\left(Q_{T_r}^{\otimes n}\right|\,R(K)\,\left|Q_{T_l}^{\otimes n}\right)| && (48)\\
&\leq \bar{\eta}_{K,t}(1+2^{8t^2-n})|\Sigma_{t,t}|^2 2^{19t^2-n} && (49)\\
&\leq \bar{\eta}_{K,t}(1+2^{8t^2-n})2^{25t^2-n}. && (50)
\end{aligned}$$

Otherwise,

$$\begin{aligned}
|\,(E_i|\,R(K)\,|E_i)\,| \leq& \sum_{r=1}^{i}\sum_{l=1}^{i} |A_{r,i}A_{l,i}| \cdot |\left(Q_{T_r}^{\otimes n}\right|\,R(K)\,\left|Q_{T_l}^{\otimes n}\right)| && (51)\\
\leq& \bar{\eta}_{K,t}|A_{i,i}|^2 + (i^2-1)2^{12t^2-n} && (52)\\
\leq& \bar{\eta}_{K,t}(1+2^{8t^2-n})^2 + \bar{\eta}_{K,t}(1+2^{8t^2-n})2^{16t^2-n} && (53)\\
\leq& \bar{\eta}_{K,t}(1+2^{16t^2-n})^3. && (54)
\end{aligned}$$

Lastly, we obtain from (31) and (27)

$$\| \, |E_i)(E_j| \, \|_\diamond \leq \sum_{r=1}^{i} \sum_{l=1}^{j} |A_{r,i} A_{l,j}| \cdot \left\| \left| Q_{T_r}^{\otimes n} \right) \left(Q_{T_l}^{\otimes n} \right| \right\|_\diamond \tag{55}$$

$$\leq |\Sigma_{t,t}|^2 2^{24t^3 - n|\dim N_r - \dim N_i| - n|\dim N_l - \dim N_j| + n(\dim N_l - \dim N_r)} \tag{56}$$

$$\leq 2^{30t^3 + n(\dim N_j - \dim N_i)}. \tag{57}$$

We now start piecing these expressions together to bound (37). Equations (57) and (41) give

$$\|[P_{\mathrm{Cl}} \mathrm{R}(K)]^k - P_{\mathrm{H}}\|_\diamond \leq$$
$$\left(1 + 2^{32t^2 - 2n}\right)^k \sum_{j_1, \ldots, j_k = t! + 1}^{|\Sigma_{t,t}|} 2^{30t^3 + n(\dim N_{j_k} - \dim N_{j_1})} \prod_{r=1}^{k-1} |\, (E_{j_r} \,|\, R(K) \,\left| E_{j_{r+1}} \right) |. \tag{58}$$

To bound (58), we will bunch together the contribution of all terms whose sequence $\{j_1, \ldots, j_k\}$ contains l changes. Moreover, we will treat differently the cases $l \leq \lfloor t/2 \rfloor$ and $l > \lfloor t/2 \rfloor$. In the former case, we use (47) to get

$$\prod_{r=1}^{k-1} |\, (E_{j_r} \,|\, R(K) \,\left| E_{j_{r+1}} \right) | \leq \bar{\eta}_{K,t}^{k-1} (1 + 2^{16t^2 - n})^{3(k-1)} 2^{l 31 t^3 - n|\dim N_{j_k} - \dim N_{j_1}|}. \tag{59}$$

In this case, the factor of $2^{n(\dim N_{j_k} - \dim N_{j_1})}$ coming from (57) is cancelled by the last factor of $2^{-n|\dim N_{j_k} - \dim N_{j_1}|}$.

In the latter case, we turn to (50) instead to obtain

$$\prod_{r=1}^{k-1} |\, (E_{j_r} \,|\, R(K) \,\left| E_{j_{r+1}} \right) | \leq \bar{\eta}_{K,t}^{k-1} (1 + 2^{16t^2 - n})^{3(k-1)} 2^{l 25 t^2 - ln}.$$

Here, the exponential factor coming from (57) is cancelled by 2^{-ln} since $\dim N_{j_k} - \dim N_{j_1} \leq \lfloor t/2 \rfloor$. Counting the instances of sequences with l changes, we may put these considerations

together to bound

$$\begin{aligned}
\|[P_{\mathrm{Cl}}\mathrm{R}(K)]^k - P_{\mathrm{H}}\|_\diamond &\leq \left(1+2^{32t^2-2n}\right)^k \left(1+2^{16t^2-n}\right)^{3(k-1)} \bar{\eta}_{K,t}^{k-1} \Bigg[\sum_{l=0}^{\lfloor \frac{t}{2} \rfloor} \binom{k}{l} |\Sigma_{t,t}|^{l+1} 2^{l31t^3} \\
&\quad + \sum_{l=\lfloor \frac{t}{2} \rfloor+1}^{k} \binom{k}{l} |\Sigma_{t,t}|^{l+1} 2^{(l-\lfloor \frac{t}{2} \rfloor)(25t^2-n)} 2^{\lfloor \frac{t}{2} \rfloor 25t^2} \Bigg] \\
&\leq \left(1+2^{32t^2-2n}\right)^{4k} \bar{\eta}_{K,t}^{k-1} \Bigg[\frac{t}{2} \binom{k}{\lfloor \frac{t}{2} \rfloor} |\Sigma_{t,t}|^{\lfloor \frac{t}{2} \rfloor+1} 2^{\lfloor \frac{t}{2} \rfloor 31t^3} \\
&\quad + \sum_{l=1}^{k-\lfloor \frac{t}{2} \rfloor} \binom{k}{l+\lfloor \frac{t}{2} \rfloor} |\Sigma_{t,t}|^{l+1+\lfloor \frac{t}{2} \rfloor} 2^{l(25t^2-n)} 2^{13t^3} \Bigg] \\
&\overset{\ddagger}{\leq} \left(1+2^{32t^2-2n}\right)^{4k} \bar{\eta}_{K,t}^{k-1} \Bigg[2^{32t^4+t\log(k)} \\
&\quad + k^{\lfloor \frac{t}{2} \rfloor} |\Sigma_{t,t}|^{1+\lfloor \frac{t}{2} \rfloor} 2^{13t^3} \sum_{l=0}^{k} \binom{k}{l} |\Sigma_{t,t}|^{l} 2^{l(25t^2-n)} \Bigg] \\
&\leq \left(1+2^{32t^2-2n}\right)^{4k} \bar{\eta}_{K,t}^{k-1} \left[2^{32t^4+t\log(k)} + 2^{18t^3+\log(k)t} \left(1+2^{28t^2-n}\right)^k \right],
\end{aligned}$$

where we have used in ‡ that

$$\begin{aligned}
\binom{k}{l+\lfloor \frac{t}{2} \rfloor} &= \frac{(k)!}{(k-l-\lfloor \frac{t}{2} \rfloor)!(l+\lfloor \frac{t}{2} \rfloor)!} \\
&\leq (k-l-\lfloor \frac{t}{2} \rfloor+1)\dots(k-l)\frac{k!}{(k-l)!l!} \\
&\leq k^{\lfloor \frac{t}{2} \rfloor} \binom{k}{l}.
\end{aligned}$$

Combined we obtain the bound

$$\|\Delta_t(\sigma_k) - P_{\mathrm{H}}\|_\diamond \leq 2^{33t^4+t\log(k)} \left(1+2^{32t^2-n}\right)^{5k} \bar{\eta}_{K,t}^{k-1}, \tag{60}$$

where $\bar{\eta}_{K,t}$ is bounded by Lemma 2. Taking the logarithm and using the inequality $\log(1+x) \leq x$ repeatedly, this implies Theorem 1. □

With the above bound, we can also prove Corollary 2.

Proof of Corollary 2. Consider the self-adjoint superoperator $A := P_{\mathrm{Cl}} R(K) P_{\mathrm{Cl}}$. As P_{Cl} is a projector, we have with Eq. (24)

$$(A - P_{\mathrm{H}})^k = A^k - P_{\mathrm{H}} = [P_{\mathrm{Cl}} R(K)]^k - P_{\mathrm{H}} = \Delta_t(\sigma_k) - P_{\mathrm{H}}. \tag{61}$$

Using norm inequality between operator and diamond norm Eq. (12) and the previous result Eq. (60), we find

$$||A - P_H||_\infty^k = ||(A - P_H)^k||_\infty \leq 2^{nt/2} \|\Delta_t(\sigma_k) - P_{\mathrm{H}}\|_\diamond \\ \leq 2^{33t^4 + t\log(k) + nt/2} \left(1 + 2^{32t^2 - n}\right)^{5k} \bar{\eta}_{K,t}^{k-1}. \quad (62)$$

Taking the k-th square root and the limit $k \to \infty$ on both sides, this yields

$$||A - P_H||_\infty \leq \left(1 + 2^{32t^2 - n}\right)^5 \bar{\eta}_{K,t}. \quad (63)$$

Combined with Ref. [23, Lem. 4], Eq. (63) implies the result. □

The bound in Eq. (60) also suffices to prove Proposition 1:

Proof of Proposition 1. The proof follows exactly as the proof of Theorem 1, but with the factor $7/8$ instead of $\bar{\eta}_{K,t}$ (compare Lemma 13). Using $\log_2(7/8) \leq -0.19$ the result can be checked. □

IV. CONVERGENCE TO HIGHER MOMENTS OF THE CLIFFORD GROUP

In this section, we aim to prove:

Theorem 3 (Local random Clifford designs)**.** *Let $n \geq 12t$, then a local random Clifford circuit of depth $O(n \log^{-2}(t) t^8 (2nt + \log(1/\varepsilon)))$ constitutes a relative ε-approximate Clifford t-design.*

The proof of Theorem 3 follows a well-established strategy [23, 54] in a sequence of lemmas. For the sake of readibility, the proofs of these lemmas have been moved to Sec. VI D. Given a measure ν on the Clifford group $\mathrm{Cl}(n)$, recall that its t-th moment operator was defined as

$$\Delta_t(\nu) := \int_{\mathrm{Cl}(2^n)} \mathrm{Ad}_U^{\otimes t} \, \mathrm{d}\nu(U).$$

The idea of the proof is that if $\Delta_t(\nu)$ is close to the moment operator $\Delta_t(\mu_{\mathrm{Cl}}) \equiv P_{\mathrm{Cl}}$ of the uniform (Haar) measure μ_{Cl} on the Clifford group, ν is an approximate Clifford design. However, we have seen that there are different notions of closeness. We define its deviation in (superoperator) *spectral norm* as

$$g_{\mathrm{Cl}}(\nu, t) := \|\Delta_t(\nu) - \Delta_t(\mu_{\mathrm{Cl}})\|_\infty .$$

Then, we prove the following lemma in Sec. VI D:

Lemma 5 (Relative $\varepsilon 2^{2tn}$-approximate Clifford t-designs)**.** *Suppose that $0 \leq \varepsilon < 1$ is such that $g_{\mathrm{Cl}}(\nu, t) \leq \varepsilon$. Then, ν is a relative $\varepsilon 2^{2tn}$-approximate Clifford t-design.*

Recall that we have defined the measure σ_G on the Clifford group $\mathrm{Cl}(n)$ in Def. 5 by randomly drawing from a 2-local Clifford gate set G and applying it to a random qubit i, or to a pair of adjacent qubits $(i, i+1)$, respectively. For this measure, we show that it fulfills the assumptions of Lemma 5:

Proposition 2 (Clifford expander bound)**.** *Let σ_G be as in Def. 5 and $n \geq 12t$. Then, $g_{\mathrm{Cl}}(\sigma_G, t) \leq 1 - c(G) n^{-1} \log^2(t) t^{-8}$ for some constant $c(G) > 0$.*

We will prove Proposition 2 in the end of this section. From this, Theorem 3 follows as a direct consequence:

Proof of Theorem 3. First, note that $g_{\mathrm{Cl}}(\nu^{*k}, t) = g_{\mathrm{Cl}}(\nu, t)^k$ for all probability measures ν on the Clifford group. This can be easily verified using the observation

$$\Delta_t(\mu_{\mathrm{Cl}})\Delta_t(\nu) = \Delta_t(\nu)\Delta_t(\mu_{\mathrm{Cl}}) = \Delta_t(\mu_{\mathrm{Cl}}). \tag{64}$$

Hence, combining the bound given by Proposition 2 and Lemma 5, we find that the k-step random walk σ_G^{*k} is a ε-approximate Clifford t-design, if we choose $k = O\left(n \log^{-2}(t) t^8 \left(2nt + \log(1/\varepsilon)\right)\right)$. □

For the sake of readibility, let us from now on drop the dependence on G and write $\sigma \equiv \sigma_G$. In order to prove Proposition 2, we use a reformulation of $g(\sigma, t)$ based on the following observation. Since G is closed under taking inverses, the moment operator $\Delta_t(\sigma)$ is self-adjoint with respect to the Hilbert-Schmidt inner product. Due to σ being a probability measure, its largest eigenvalue is 1 with eigenspace corresponding to the operator subspace which is fixed by the adjoint action $\mathrm{Ad}(g^{\otimes t})$ of all generators. Equivalently, this is the subspace of operators which commute with any generator $g^{\otimes t}$. However, any operator commuting with all generators also commutes with every element in the Clifford group $\mathrm{Cl}(n)$ and vice versa. Hence, this subspace is nothing but the Clifford commutant $\mathrm{Cl}(n)'$ with projector $P_{\mathrm{Cl}} \equiv \Delta_t(\mu_{\mathrm{Cl}})$. Thus, the spectral decomposition is

$$\Delta_t(\sigma) = P_{\mathrm{Cl}} + \sum_{r \geq 2} \lambda_r(\Delta_t(\sigma))\Pi_r, \tag{65}$$

where $\lambda_r(X)$ denotes the r-th largest eigenvalue of an operator X. Hence, we find

$$g(\sigma, t) = \left\|\Delta_t(\sigma) - P_{\mathrm{Cl}}\right\|_\infty = \lambda_2\left(\Delta_t(\sigma)\right). \tag{66}$$

Note that since $\Delta_t(\sigma)$ is self-adjoint, we can interpret it as an Hamiltonian on the Hilbert space $L((\mathbb{C}^2)^{\otimes nt})$. In this light, it will turn out to be useful to recast Eq. (66) as the spectral gap of a suitable family of *local Hamiltonians* with vanishing ground state energy:

$$H_{n,t} := n\left(\mathrm{id} - \Delta_t(\sigma)\right) = \sum_{i=1}^{n} h_{i,i+1}, \quad \text{with} \quad h_{i,i+1} := \frac{1}{|G|}\sum_{g \in G}\left(\mathrm{id} - \mathrm{Ad}(g_{i,i+1}^{\otimes t})\right). \tag{67}$$

Let us summarize these findings in the following lemmas.

Lemma 6 (Spectral gap)**.** *Let σ be as in Def. 5 and $H_{n,t}$ the Hamiltonian from Eq.* (67)*. It holds that*

$$g(\sigma, t) = 1 - \frac{\Delta(H_{n,t})}{n}. \tag{68}$$

Lemma 7 (Ground spaces)**.** *The Hamiltonians $H_{n,t}$ are positive operators with ground state energy* 0*. The ground space is given by the Clifford commutant*

$$\mathrm{Cl}(n)' = \mathrm{span}\left\{r(T)^{\otimes n} \;\middle|\; T \in \Sigma_{t,t}\right\}, \tag{69}$$

where $\Sigma_{t,t}$ is the set of stochastic Lagrangian subspaces of $\mathbb{Z}_2^t \oplus \mathbb{Z}_2^t$.

In the remainder of this section, we will prove the existence of a uniform lower bound on the spectral gap of $H_{n,t}$. In combination with Lemma 6 and Lemma 5 this will imply Theorem 3. While it is highly non-trivial to show spectral gaps in the thermodynamic limits, we can use the fact that $H_{n,t}$ is *frustration-free* (compare Lemma 7). This allows us to apply the powerful *martingale method* pioneered by Nachtergaele [55].

Lemma 8 (Lower bound to spectral gap)**.** *Let the Hamiltonian $H_{n,t}$ be as in Eq.* (67) *and assume that $n \geq 12t$. Then, $H_{n,t}$ has a spectral gap satisfying*

$$\Delta(H_{n,t}) \geq \frac{\Delta(H_{12t,t})}{48t}. \tag{70}$$

Proof of Proposition 2. We can now combine the bound in (70) with any lower bound on the spectral gap independent of t. Let $T_\nu : L^2(\mathrm{Cl}(n)) \to L^2(\mathrm{Cl}(n))$ be given by

$$T_\nu f(g) := \int f(h^{-1}g)\mathrm{d}\nu(h). \tag{71}$$

Notice that it is the (Hermitian) averaging operator with respect to ν on the group algebra. The highest eigenvalue of T_ν is $\lambda_1(T_\nu) = 1$, its eigenspace corresponds to the trivial representation. By Ref. [56, Cor. 1] we have that

$$\lambda_2(T_\sigma) \leq 1 - \frac{\eta}{d^2}, \tag{72}$$

where η is the probability of the least probable generator (here $1/|G|n$) and d is the diameter of the associated Cayley graph (given in Ref. [57] as $d = O(n^3/\log(n))$.

According to the Peter-Weyl theorem, the spectrum of $H_{n,t}$ is contained in the spectrum of T_σ, in particular it is the same as the spectrum of the restriction of T_σ to the irreducible representations that appear in the representation $U \mapsto \mathrm{Ad}_U^{\otimes t}$. This representation contains a trivial component, so $H_{n,t}$ also has a gap of at least η/d^2. Finally, by Lemma 8 it follows that

$$\Delta(H_{n,t}) \geq \frac{\Delta(H_{12t,t})}{48t} \geq c(G)t^{-8}\log(t)^2, \tag{73}$$

for a constant $c(G)$. We note that the applicability of Ref. [56, Cor. 1] to random walks on the Clifford group has also been observed in Ref. [9].

□

We can combine Theorem 3 and Theorem 1 to obtain the following corollary:

Corollary 3 (Local random unitary design)**.** *Let $K \in U(2)$ be a non-Clifford gate and let $G \subset \mathrm{Cl}(4)$ be a closed, generating set. There are constants $C_1''(K,G), C_2''(K), C_3''(K)$ such that whenever*

$$m \geq C_1''(K,G)n\log^{-2}(t)t^8\left(2nt + \log(1/\varepsilon)\right) \quad \textit{and} \quad k \geq C_2''(K)\log^2(t)(t^4 + t\log(1/\varepsilon)),$$

the local random circuit $\sigma_{k,m}$, defined in (6)*, is an ε-approximate unitary t-design for all $n \geq C_3''(K)t^2$.*

Proof. Consider the superoperator

$$\Delta_t(\sigma_{k,m}) = \int_{U(2^n)} \mathrm{Ad}(U^{\otimes t})\,\mathrm{d}\sigma_{k,m}(U) = \underbrace{\Delta_t(\sigma^{*m})\mathrm{R}(K)\ldots\Delta_t(\sigma^{*m})\mathrm{R}(K)}_{k \text{ times}}, \tag{74}$$

where σ^{*m} denotes the probability measure of a depth m local random walk on the Clifford group (cp. Def. 5). We would like to bound the difference between the Haar random t-th moment operator $\Delta_t(\mu_{\mathrm{H}}) =: P_{\mathrm{H}}$ and $\Delta_t(\sigma_{k,m})$. Notice the following standard properties of P_{H}:

$$P_{\mathrm{H}}\Delta_t(\nu) = \Delta_t(\nu)P_{\mathrm{H}} = P_{\mathrm{H}}, \qquad \text{and} \qquad P_{\mathrm{H}}^\dagger = P_{\mathrm{H}}, \tag{75}$$

for any probability measure ν on $U(2^n)$. In particular, we have that P_{H} is an orthogonal projector. As in the last section, we make use of the spectral decomposition in Eq. (65) to decompose $\Delta_t(\sigma^{*k})$ as follows:

$$\begin{aligned}\Delta_t(\sigma_{k,m}) - P_{\mathrm{H}} &= [\Delta_t(\sigma^{*m})\mathrm{R}(K)]^k - P_{\mathrm{H}} \\ &= \left[\left(P_{\mathrm{Cl}} + \sum_{i\geq 2}\lambda_i^m \Pi_i\right)\mathrm{R}(K)\right]^k - P_{\mathrm{H}}.\end{aligned} \tag{76}$$

Recall the shorthand notation $P_{\mathrm{Cl}} := \Delta_t(\mu_{\mathrm{Cl}})$. Using the triangle inequality and the inequality (12), this implies

$$\begin{aligned}\|\Delta_t(\sigma_{k,m}) - P_{\mathrm{H}}\|_\diamond &\leq \left\|[P_{\mathrm{Cl}}\mathrm{R}(K)]^k - P_{\mathrm{H}}\right\|_\diamond + 2^{2tn}\sum_{l=1}^{k}\binom{k}{l}\lambda_2^{lm} \\ &\leq \left\|[P_{\mathrm{Cl}}\mathrm{R}(K)]^k - P_{\mathrm{H}}\right\|_\diamond + k2^{2tn+1}\lambda_2^m.\end{aligned} \tag{77}$$

Note that we bounded the second largest eigenvalue λ_2 of $\Delta_t(\sigma)$ in Proposition 2. We can now combine Proposition 2 with (60) to obtain:

$$\|\Delta_t(\sigma_{k,m}) - P_{\mathrm{H}}\|_\diamond \leq k2^{2tn+1}\lambda_2^m + 2^{33t^4+t\log(k)}\left(1+2^{32t^2-n}\right)^{5k}\bar{\eta}_{K,t}^k. \tag{78}$$

□

V. SINGLING OUT THE CLIFFORD GROUP

There are a number of ways to motivate the construction of approximate unitary t-designs from random Clifford circuits. For example, from a physical point of view, Clifford gates are often comparatively easy to implement, in particular in fault-tolerant architectures. In this section, we point out that Refs. [37, 38] together imply that the Clifford groups are also mathematically distinguished. Proposition 3 is a Corollary of the recently published classification of finite unitary subgroups which form t-designs, so-called *unitary t-groups*, by Bannai *et al.* [37] and a theorem about universality of finitely generated subgroups by Sawicki and Karnas [38].

For any subgroup $G \subseteq \mathrm{U}(d)$, we let

$$\overline{G} := \{\det(U^\dagger)U \,|\, U \in G\} \subseteq \mathrm{SU}(d).$$

Notice that $\overline{G}$ is a unitary t-design if and only if G is.

Proposition 3 refers to t-designs generated by *finite gate sets*, which we define now. The starting point is a Hilbert space $(\mathbb{C}^q)^{\otimes r}$ for some r. A finite gate set is a finite subset

$$\mathcal{G} \subset \mathrm{SU}\left((\mathbb{C}^q)^{\otimes r}\right).$$

We will denote by $\mathcal{G}_n$ the subgroup of $\mathrm{SU}\left((\mathbb{C}^q)^{\otimes n}\right)$ generated by elements of $\mathcal{G}$ acting on any r tensor factors (here $r \leq n$). The number q is called the *local dimension* of $\mathcal{G}$.

Proposition 3 (Singling out the Clifford group [37, 38])**.** *Let $t \geq 2$, and let $\mathcal{G}$ be a finite gate set with local dimension $q \geq 2$. Assume that (1) either all $\mathcal{G}_n$ are finite or they are all infinite, and (2) there is an n_0 such that for all $n \geq n_0$, $\mathcal{G}_n$ is a unitary t-design.*

Then, one of the following cases apply:

(i) If $t = 2$, we have either q prime and $\mathcal{G}_n$ is isomorphic to a subgroup of the Clifford group $\overline{\mathrm{Cl}}(q^n)$, or *$\mathcal{G}_n$ is dense in $\mathrm{SU}(q^n)$,*

(ii) If $t = 3$, we have either *$q = 2$ and $\mathcal{G}_n$ is isomorphic to the full Clifford group $\overline{\mathrm{Cl}}(2^n)$* or *$\mathcal{G}_n$ is dense in $\mathrm{SU}(q^n)$,*

(iii) If $t \geq 4$ then $\mathcal{G}_n$ is dense in $\mathrm{SU}(q^n)$.

Note that a finitely generated infinite subgroup of $\mathrm{SU}(d)$ is always dense in some compact Lie subgroup (cp. [38, Fact 2.6]). In particular, it inherits a Haar measure from this Lie subgroup which allows for a definition of unitary t-design.

a. Finite case. In the classification in Ref. [37], the non-existence of finite unitary t-groups was shown for $t \geq 4$ (and dimension $d > 2$). Already the case $t = 3$ is very restrictive, since the authors arrive at the following result:

Lemma 9 (Ref. [37, Thm. 4])**.** *Suppose $d \geq 5$ and $H < \mathrm{SU}(d)$ is a finite unitary 3-group. Then, H is either one of finitely many exceptional cases or $d = 2^n$ and H is isomorphic to the Clifford group $\overline{\mathrm{Cl}}(2^n)$.*

This establishes the finite version of (ii), the $t = 3$ case.

The classification of unitary 2-designs is however more involved, it includes certain irreducible representations of finite unitary and symplectic groups (compare [37, Thm. 3 Lie-type case]), and a finite set of exceptions. The exceptions can be ruled out in the same way as above.

The former, the Lie-type cases, happen in dimensions $(3^n \pm 1)/2$ and $(2^n + (-1)^n)/3$. There is no q for which there exists an n_0 such that for all $n \geq n_0$ there exists an $m \in \mathbb{N}$ satisfying either

$$q^n = (3^m \pm 1)/2 \qquad \text{or} \qquad q^n = (2^m + (-1)^m)/3.$$

Thus, the assumptions of Prop. 3 rule these out. This establishes the finite version of (i).

b. Infinite case. Define the commutant for a set $S \subset \mathrm{SU}(d)$ of the adjoint action as

$$\mathrm{Comm}(\mathrm{Ad}_S) := \left\{ L \in \mathrm{End}\left(\mathbb{C}^{d\times d}\right) \;\middle|\; [\mathrm{Ad}_g, L] = 0 \;\; \forall g \in S \right\}.$$

We show that the second case can be reduced to Cor. 3.5 from Ref. [38] applied to the simple Lie group $\mathrm{SU}(d)$.

Lemma 10 ([38, Cor. 3.5])**.** *Given a finite set $G \subset \mathrm{SU}(d)$ such that $\mathcal{G} = \langle G \rangle$ is infinite. Then, the group $\mathcal{G}$ is dense in $\mathrm{SU}(d)$ if and only if*

$$\mathrm{Comm}(\mathrm{Ad}_{\mathcal{G}}) \cap \mathrm{End}(\mathfrak{su}(d)) = \{\lambda \,\mathrm{id}_{\mathfrak{su}(d)} \mid \lambda \in \mathbb{R}\}. \tag{79}$$

Recall that a subgroup $\mathcal{G} \subseteq U(d)$ is a unitary 2-group if and only if $\mathrm{Comm}(U \otimes U | U \in \mathcal{G}) = \mathrm{Comm}(U \otimes U | U \in \mathrm{U}(d)) = \mathrm{span}(\mathbb{1}, \mathbb{F})$, where $\mathbb{F}$ denotes the flip of two tensor copies (see also App. A). Let us denote the partial transpose on the second system of a linear operator $A \in L(\mathbb{C}^d \otimes \mathbb{C}^d)$ by A^Γ. Then, one can easily verify that Γ induces a vector space isomorphism

between $\mathrm{Comm}(U \otimes U | U \in \mathcal{G})$ and $\mathrm{Comm}(U \otimes \overline{U} | U \in \mathcal{G})$. The image of the basis $\{\mathbb{1}, \mathbb{F}\}$ is readily computed as

$$\mathbb{1}^{\Gamma} = \mathbb{1}, \qquad \mathbb{F}^{\Gamma} = d\, |\Omega\rangle\langle\Omega|\,, \tag{80}$$

where $|\Omega\rangle = d^{-1/2} \sum_{i=1}^{d} |ii\rangle$ is the maximally entangled state vector. Next, we use that $U \otimes \overline{U} = \mathrm{mat}(\mathrm{Ad}_U)$ is the matrix representation of $\mathrm{Ad}_U = U \cdot U^{\dagger}$ with respect to the basis $E_{i,j} = |i\rangle\langle j|$ of $L(\mathbb{C}^d)$. Thus, we have $\mathrm{Comm}(\mathrm{Ad}_{\mathcal{G}}) \simeq \mathrm{Comm}(U \otimes \overline{U} | U \in \mathcal{G})$ as algebras. Pulling the above basis of $\mathrm{Comm}(U \otimes \overline{U} | U \in \mathcal{G})$ back to $\mathrm{Comm}(\mathrm{Ad}_{\mathcal{G}})$, we then find:

$$\mathrm{mat}^{-1}(\mathbb{1}) = \mathrm{id}_{L(\mathbb{C}^d)}, \quad \mathrm{mat}^{-1}(|\Omega\rangle\langle\Omega|) = \mathrm{Tr}(\bullet)\mathrm{id}_{L(\mathbb{C}^d)}. \tag{81}$$

Hence, we have shown that any element in $\mathrm{Comm}(\mathrm{Ad}_{\mathcal{G}})$ is a linear combination of these two maps. However, by restricting to $\mathfrak{su}(d)$, the second map becomes identically zero, thus we have

$$\mathrm{Comm}(\mathrm{Ad}_{\mathcal{G}}) \cap \mathrm{End}(\mathfrak{su}(d)) = \{\lambda\, \mathrm{id}_{\mathfrak{su}(d)} \,|\, \lambda \in \mathbb{R}\}. \tag{82}$$

By Lemma 10, this shows that any finitely generated infinite unitary 2-group $\mathcal{G} \leq \mathrm{SU}(d)$ is dense in $\mathrm{SU}(d)$. Since any unitary t-group is in particular a 2-group, this is also true for any $t > 2$.

VI. PROOFS

A. Proof of overlap lemmas

In this section, we prove three technical lemmas which are needed throughout this paper.

Lemma 3 (Diamond norm bound)**.** *Consider $T_1, T_2 \in \Sigma_{t,t}$ and denote with N_1, N_2 their respective defect spaces. Then, it holds that*

$$\| |Q_{T_1})(Q_{T_2}| \|_{\diamond} \leq 2^{\dim N_2 - \dim N_1}, \tag{27}$$

$$|\, (Q_{T_1} | Q_{T_2})\, | \leq 2^{-|\dim N_1 - \dim N_2|}. \tag{28}$$

Proof. First, recall that $Q_T := 2^{-t/2} r(T))$. Then, we make use of the following elementary bound on the diamond norm of rank one superoperator $|A)(B|$:

$$\begin{aligned} \| |A)(B| \|_{\diamond} &= \sup_{\|X\|_1 = 1} \|A \otimes \mathrm{Tr}_1 (B \otimes \mathbb{1} X)\|_1 \\ &\overset{\dagger}{\leq} \|A\|_1 \sup_{\|X\|_1 = 1} \|B \otimes \mathbb{1} X\|_1 \\ &\overset{\ddagger}{=} \|A\|_1 \, \|B \otimes \mathbb{1}\|_{\infty} \\ &= \|A\|_1 \, \|B\|_{\infty}\,. \end{aligned} \tag{83}$$

Here, we have used in † that the partial trace is a contraction w.r.t. $\|\cdot\|_1$ and in ‡ a version of the duality between trace and spectral norm [58]. Given stochastic Lagrangians T_1 and T_2 with defect spaces N_1 and N_2, we thus find using Lem. 1:

$$\| |Q_{T_1})(Q_{T_2}| \|_{\diamond} \leq 2^{-t} \, \|r(T_1)\|_1 \, \|r(T_2)\|_{\infty} = 2^{\dim N_2 - \dim N_1}. \tag{84}$$

To prove 2., we use Ref. [41, Eq. (4.25)] and that the transpose does not change the dimension of the corresponding defect subspace. Moreover, we assume w.l.o.g. that $\dim N_2 \geq \dim N_1$. We have

$$|\,(Q_{T_1}|Q_{T_2})\,| = 2^{-t}|\,\mathrm{Tr}[r(T_1)r(T_2)^T]| = 2^{-t+\dim(N_1\cap N_2)}|\,\mathrm{Tr}[r(T)]| \tag{85}$$

where $r(T)$ is described by a stochastic orthogonal and a defect space $N_1^\perp \cap N_2 + N_1$. Hence, we obtain (together with Hölder's inequality):

$$|\,(Q_{T_1}|Q_{T_2})\,| \leq 2^{-t+\dim(N_1\cap N_2)}2^{t-\dim(N_1^\perp\cap N_2+N_1)}. \tag{86}$$

Using $N \subseteq N^\perp$ for all defect spaces and the general identity $\dim(V+W) = \dim V + \dim W - \dim(V\cap W)$, this yields

$$|\,(Q_{T_1}|Q_{T_2})\,| \leq 2^{\dim(N_1\cap N_2)-\dim N_1} \leq 2^{\dim N_2-\dim N_1}. \tag{87}$$

□

Lemma 11 (Overlap of stochastic Lagrangian subspaces)**.** *We have* $(Q_T|Q_{T'}) \geq 0$ *for all* $T, T' \in \Sigma_{t,t}$*. Moreover, for all* $T \in \Sigma_{t,t}$ *the sum of overlaps is*

$$\sum_{T'\in\Sigma_{t,t}} (Q_T|Q_{T'})^n = (-2^{-n};2)_{t-1} \leq 1 + t2^{t-n}, \tag{88}$$

where $(-2^{-n};2)_{t-1} = \prod_{r=0}^{t-2}(1+2^{r-n})$ *and the last inequality holds for* $n+2 \geq t + \log_2(t)$.

Proof. Denote by $\mathrm{Stab}(n)$ the set of stabilizer states on n qubits. Since the operators $r(T)$ are entry-wise non-negative, we have $(Q_T|Q_{T'}) = 2^{-t}\,\mathrm{Tr}(r(T)^\dagger r(T')) \geq 0$. Note that $r(T)^\dagger = r(\tilde{T})$ for a suitable $\tilde{T} \in \Sigma_{t,t}$ (cp. Thm. 5). We obtain

$$\begin{aligned}
\sum_{T'\in\Sigma_{t,t}} (Q_T|Q_{T'})^n &= \frac{1}{2^{tn}} \sum_{T'\in\Sigma_{t,t}} \mathrm{Tr}\left[r(\tilde{T})^{\otimes n} r(T')^{\otimes n}\right] \\
&\overset{\dagger}{=} \frac{2^n \prod_{r=0}^{t-2}(2^r+2^n)}{2^{tn}}\,\mathrm{Tr}\left[r(\tilde{T})^{\otimes n}\mathbb{E}_{s\in\mathrm{Stab(n)}}(\,|s\rangle\langle s\,|^{\otimes t})\right] \\
&= \frac{2^n \prod_{r=0}^{t-2}(2^r+2^n)}{2^{tn}}\mathbb{E}_{s\in\mathrm{Stab(n)}}\left\langle s^{\otimes t}\right| r(\tilde{T})^{\otimes n}\left|s^{\otimes t}\right\rangle \\
&\overset{\ddagger}{=} \frac{2^n \prod_{r=0}^{t-2}(2^r+2^n)}{2^{tn}} \\
&= \prod_{r=0}^{t-2}(1+2^{r-n}) \\
&\leq \left(1+2^{t-2-n}\right)^{t-1} \\
&\overset{*}{\leq} \exp\left((t-1)2^{t-n-2}\right),
\end{aligned} \tag{89}$$

where we have again used [41, Thm. 5.3] in † and in ‡ that $\langle s^{\otimes t}|\, r(T)^{\otimes n}\,|s^{\otimes t}\rangle = 1$ for all $T \in \Sigma_{t,t}$ and all $s \in \mathrm{Stab}(n)$ (compare Ref. [41, Eq. (4.10)]). Finally, in $*$ we have used the "inverse Bernoulli inequality" $(1+x)^r \leq e^{rx}$ which holds for all $x \in \mathbb{R}$ and $r \geq 0$. By assumption, the following holds

$$0 \geq t + \log_2(t) - n - 2 \quad \Rightarrow \quad 1 \geq t2^{t-n-2} \geq (t-1)2^{t-n-2}. \tag{90}$$

Thus, we can use the inequality $e^x \leq 1 + 2x$ for $0 \leq x \leq 1$ to obtain

$$\begin{aligned} \sum_{T' \in \Sigma_{t,t}} (Q_T|Q_{T'})^n &\leq 1 + (t-1)2^{t-n-1} \\ &\leq 1 + t2^{t-n}. \end{aligned} \tag{91}$$

□

Definition 10 (Clifford frame operator)**.** *We define the Clifford frame operator as*

$$S_{\mathrm{Cl}} := \sum_{T \in \Sigma_{t,t}} |Q_T)(Q_T|^{\otimes n}. \tag{92}$$

Lemma 12. *Let S_{Cl} be the Clifford frame operator and Γ the corresponding Gram matrix, i. e. $\Gamma_{T,T'} = (Q_T|Q'_T)^n$. Then the following holds*

$$\|S_{\mathrm{Cl}} - P_{\mathrm{Cl}}\|_\infty = \|\Gamma - \mathbb{1}\|_\infty \leq (-2^{-n}; 2)_{t-1} - 1 \leq t2^{t-n}, \tag{93}$$

where the last inequality holds for $n + 2 \geq t + \log_2 t$.

Proof. Define the *synthesis operator* of the frame as the map

$$V : \mathbb{C}^{|\Sigma_{t,t}|} \to \mathrm{Cl}(n)', \quad V = \sum_{T \in \Sigma_{t,t}} \left|Q_T^{\otimes n}\right)\langle e_T|, \tag{94}$$

where e_T is the standard basis of the domain. Then, we have clearly $\Gamma = V^\dagger V$ and $S_{\mathrm{Cl}}|_{\mathrm{Cl}(n)'} = VV^\dagger$. Since S_{Cl} and P_{Cl} are both identically zero on $(\mathrm{Cl}(n)')^\perp$, this part does not contribute to the spectral norm. From this it is clear that

$$\|S_{\mathrm{Cl}} - P_{\mathrm{Cl}}\|_\infty = \|\Gamma - \mathbb{1}\|_\infty. \tag{95}$$

Moreover, we can compute

$$\begin{aligned} \|\Gamma - \mathbb{1}\|_\infty &= \left\| \sum_T \sum_{T,T'} (Q_T|Q_{T'})^n \, |e_T\rangle\langle e_{T'}| \right\|_\infty \\ &\leq \max_T \sum_{T' \neq T} (Q_T|Q_{T'})^n \\ &= (-2^{-n}; 2)_{t-1} - 1, \end{aligned} \tag{96}$$

where we have used that the spectral norm of Hermitian operators is bounded by the max-column norm and inserted the exact result of Lemma 11 in the last step. Finally, said lemma provides the desired bound for $n + 2 \geq t + \log_2 t$. □

B. Proof of Lemmas for Theorem 1

Lemma 2 (Overlap bound)**.** *Let K be a single qubit gate which is not contained in the Clifford group. Then, there is a constant $c(K) > 0$ such that*

$$\eta_{K,t} := \max_{\substack{T \in \Sigma_{t,t} - S_t \\ T' \in \Sigma_{t,t}}} \frac{1}{3} \left| (Q_T| \operatorname{Ad}_K^{\otimes t} + \operatorname{Ad}_{K^\dagger}^{\otimes t} + \mathrm{id} \, |Q_{T'}) \right| \leq 1 - c(K) \log^{-2}(t). \tag{26}$$

The proof of Lemma 2 is based on two results. The first states that the basis elements $r(T)$ of the commutant of tensor powers of the Clifford group either belong to the commutant of the powers of the unitary group, or else are far away from it.

Lemma 13 (Haar symmetrization)**.** *For all t and for all $T \in \Sigma_{t,t} \setminus S_t$, it holds that*

$$(Q_T|\, P_{\mathrm{H}}\, |Q_T) = 2^{-t}\, \|P_{\mathrm{H}}[r(T)]\|_2^2 \leq \frac{7}{8}, \tag{97}$$

where Q_T is as in eq. (21) and $P_{\mathrm{H}} = \Delta_t(\mu_{\mathrm{H}})$ is the t-th moment operator of the single-qubit unitary group $\mathrm{U}(2)$.

The proof is given in Section VI C. In Appendix C, we show that the constant $7/8$ cannot be improved below $7/10$, by exhibiting a T that attains this bound.

The second ingredient to Lemma 2 is a powerful theorem by Varjú [51]. Here, we specialize this theorem to the unitary group:

Theorem 6 ([51, Thm. 6])**.** *Let ν be a probability measure on* $\mathrm{U}(d)$. *Consider the averaging operator $T_v(\nu)$ on a irreducible representation $\pi_v : \mathrm{U}(d) \to \mathrm{End}(W_v)$ parameterized by highest weight $v \in \mathbb{Z}^d$:*

$$T_v(\nu) := \int_{\mathrm{U}(d)} \pi_v(U)\, \mathrm{d}\nu(U). \tag{98}$$

Then there are numbers $C(d) > 0$ and $r_0 > 0$ such that

$$\Delta_r(\nu) := 1 - \max_{0<|v|\leq r} \|T_v(\nu)\|_\infty \geq C(d)\Delta_{r_0}(\nu) \log^{-2}(r), \tag{99}$$

where $|v|^2 = \sum_i v_i^2$.

Proof of Lemma 2. Consider the probability measure ξ_K that draws uniformly from the set $\{K, K^\dagger, \mathbb{1}\}$. Moreover, define ν_K on $\mathrm{U}(2)$ as the average of the uniform measure on $\{H, S, S^3\}$ and $\xi_K * \xi_K$. Hence, the according moment operator is

$$\begin{aligned} \Delta_t(\nu_K) :=& \frac{1}{6}(\mathrm{Ad}_H^{\otimes t} + \mathrm{Ad}_S^{\otimes t} + (\mathrm{Ad}_S^3)^{\otimes t}) + \frac{1}{2} M_t(\xi_K * \xi_K) \\ =& \frac{1}{6}(\mathrm{Ad}_H^{\otimes t} + \mathrm{Ad}_S^{\otimes t} + (\mathrm{Ad}_S^3)^{\otimes t}) + \frac{1}{2} \Delta_t(\xi_K)^2. \end{aligned} \tag{100}$$

As the Clifford group augmented with any non-Clifford gate is universal [59, Thm. 6.5], so is the probability measure ν_K.

It follows from the representation theory of the unitary group (see App. B) that the representation $U \mapsto \mathrm{Ad}_U^{\otimes t}$ does not contain irreducible representations W_v with highest weight of length $|v| > \sqrt{2}t$. Thus, we can decompose into these irreducible representations as follows:

$$\begin{aligned} \|\Delta_t(\nu_K) - P_{\mathrm{H}}\|_\infty &= \left\| \bigoplus_{|v|\leq\sqrt{2}t} (T_v(\nu_K) - T_v(\mu_{\mathrm{H}})) \otimes \mathrm{id}_{m_v} \right\|_\infty \\ &\leq \left\| \bigoplus_{0<|v|\leq\sqrt{2}t} T_v(\nu_K) \right\|_\infty \\ &= \max_{0<|v|\leq\sqrt{2}t} \|T_v(\nu_K)\|_\infty \\ &= 1 - \Delta_{\sqrt{2}t}(\nu_K). \end{aligned} \tag{101}$$

Here, m_v denotes the multiplicity of the irreducible representation W_v (possibly zero). In the second step we have used that P_{H} has only support on the trivial irreducible representation $v = 0$, where both P_{H} and $\Delta_t(\nu_K)$ act as identity and thus cancel. Hence, only non-trivial irreducible representations are contributing. To bound $\Delta_{\sqrt{2}t}(\nu_K)$, we can invoke Theorem 6 combined with the fact that for any universal probability measure the restricted gap is non-zero: $\Delta_r(\nu_K) > 0$ for all $r \geq 1$ (compare e.g. Ref. [26]). Hence, we obtain

$$\Delta_{\sqrt{2}t}(\nu_K) \geq C(2)\Delta_{r_0}(\nu_K)\log^{-2}(\sqrt{2}t) \geq \frac{1}{4}C(2)\Delta_{r_0}(\nu_K)\log^{-2}(t) =: c'(K)\log^{-2}(t) > 0, \tag{102}$$

where $c(K) > 0$. Therefore, we have

$$\left\|\Delta_t(\nu_K) - P_{\mathrm{H}}\right\|_\infty \leq 1 - \Delta_{\sqrt{2}t}(\nu_K) \leq 1 - c'(K)\log^{-2}(t) =: \kappa_{t,K}, \tag{103}$$

Furthermore, consider the operator

$$X_T := \frac{(\mathrm{id} - P_{\mathrm{H}})Q_T}{\left\|(\mathrm{id} - P_{\mathrm{H}})Q_T\right\|_2}. \tag{104}$$

We obtain

$$\begin{aligned}
\left\|\Delta_t(\nu_K) - P_{\mathrm{H}}\right\|_\infty &= \max_{\|X\|_2=1} \left|(X|\,\Delta_t(\nu_K) - P_{\mathrm{H}}\,|X)\right| \\
&\geq \frac{\left|(X_T|\,\Delta_t(\nu_K) - P_{\mathrm{H}}\,|X_T)\right|}{\|X_T\|_2^2} \\
&= \frac{\left|(Q_T|\,(\mathrm{id} - P_{\mathrm{H}})\Delta_t(\nu_K)(\mathrm{id} - P_{\mathrm{H}})\,|Q_T)\right|}{(Q_T|\,(\mathrm{id} - P_{\mathrm{H}})^2\,|Q_T)} \\
&= \frac{\left|\,(Q_T|\,\Delta_t(\nu_K)\,|Q_T) - (Q_T|\,P_{\mathrm{H}}\,|Q_T)\,\right|}{1 - (Q_T|\,P_{\mathrm{H}}\,|Q_T)} \\
&\geq \frac{(Q_T|\,\Delta_t(\nu_K)\,|Q_T) - (Q_T|\,P_{\mathrm{H}}\,|Q_T)}{1 - (Q_T|\,P_{\mathrm{H}}\,|Q_T)}.
\end{aligned} \tag{105}$$

In the fourth step, we again used the properties of the Haar projector as in Eq. (75). Combining this with (103) and Lemma 13 we obtain

$$(Q_T|\,\Delta_t(\nu_K)\,|Q_T) \leq \kappa_{t,K} + (1 - \kappa_{t,K})\,(Q_T|\,P_{\mathrm{H}}\,|Q_T) \leq 1 - \frac{1}{8}c'(K)\log^{-2}(t). \tag{106}$$

We can use that $(Q_T|\,\mathrm{Ad}_S^{\otimes t}\,|Q_T) = (Q_T|\,\mathrm{Ad}_{S^3}^{\otimes t}\,|Q_T) = (Q_T|\,\mathrm{Ad}_H^{\otimes t}\,|Q_T) = 1$ for all $T \in \Sigma_{t,t}$ because $Q_T = 2^{-t/2}r(T)$ commutes with the t-th diagonal action of the single-qubit Clifford group (compare [41, Lem. 4.5]). We immediately obtain

$$(Q_T|\,\Delta_t(\xi_K)^2\,|Q_T) \leq 1 - \frac{1}{4}c'(K)\log^{-2}(t). \tag{107}$$

From the Cauchy-Schwarz inequality, we now get

$$\begin{aligned}
\left|(Q_T|\,\Delta_t(\xi_K)\,|Q_{T'})\right| &\leq \sqrt{(Q_T|\,\Delta_t(\xi_K)^2\,|Q_T)} \\
&\leq \sqrt{1 - \frac{1}{4}c'(K)\log^{-2}(t)} \\
&\leq 1 - \frac{1}{8}c'(K)\log^{-2}(t) \\
&=: 1 - c(K)\log^{-2}(t),
\end{aligned} \tag{108}$$

where we have used that $c'(K)\log^{-2}(t) \leq \Delta_{\sqrt{2}t}(\nu_K) \leq 1$ such that we can use the inequality $\sqrt{1-x} \leq 1 - x/2$ for $x \leq 1$. This shows the claimed statement. □

Remark 2 (Quantum gates with algebraic entries)**.** *If we restrict to gates K that have only algebraic entries, we can apply the result from Ref. [60] and save the additional overhead of $\log^2(t)$ in the scaling. This applies to the T-gate and for essentially all gates that might be used in practical implementations. Here, we have chosen the more general approach.*

Remark 3 (Implications for quantum information processing)**.** *Theorem 6 has miscellaneous implications for quantum information processing. E.g. we can immediately combine this bound with the local-to-global lemma in Ref. [22, Lem. 16] to extend Ref. [23, Cor. 7] to gate sets with non-algebraic entries at the cost of an additional overhead of $\log^2(t)$ in the scaling. The bottleneck to loosen the invertibility assumption as well is the local-to-global lemma which only works for Hermitian moment operators (symmetric distributions). Work to lessen the assumption of invertibility has been done in Ref. [61]. Extending this would be an interesting application which we, however, do not pursue in this work.*

Lemma 4 (Properties of the constructed basis)**.** *Let $\{T_j\}_{j=1}^{|\Sigma_{t,t}|}$ be an enumeration of the elements of $\Sigma_{t,t}$ such that the first $t!$ spaces T_j correspond to the elements of S_t. Then, the $\{E_j\}$ constitutes an orthogonal (but not normalized) basis, where*

$$E_j := \sum_{i=1}^{j} A_{i,j}\, Q_{T_i}^{\otimes n} := \sum_{i=1}^{j} \left[\sum_{\substack{\Pi \in S_j \\ \Pi(j)=i}} \operatorname{sign}(\Pi) \prod_{l=1}^{j-1} \left(Q_{T_l} \middle| Q_{T_{\Pi(l)}} \right)^n \right] Q_{T_i}^{\otimes n}. \tag{30}$$

Denote by N_i the defect space of T_i. For $n \geq \frac{1}{2}(t^2+5t)$, we have

$$|A_{i,j}| \leq 2^{t^3+4t^2+6t-n|\dim N_i - \dim N_j|}, \qquad \forall i,j, \tag{31}$$

$$|A_{i,j}| \leq 2^{2t^2+10t-n}, \qquad \forall i \neq j. \tag{32}$$

Moreover, it holds that

$$1 - 2^{t^2+7t-n} \leq A_{j,j} \leq 1 + 2^{t^2+7t-n}. \tag{33}$$

Proof. The form of (30) is up to a constant the determinant formulation of the Gram-Schmidt procedure. First, note that the number of permutations of n elements with no fixed points is known from Ref. [62] to be

$$D(n) = n! \sum_{r=0}^{n} \frac{(-1)^r}{r!} \leq 2\frac{n!}{e} \tag{109}$$

for $n \geq 1$. Here, D stands for "derangement" as permutations without fixed points are sometimes called. Then, the number of permutations having exactly k fixed points is $\binom{n}{k}$ many choices of k points times the number $D(n-k)$ of deranged permutations on the remaining $n-k$ objects:

$$p(n,k) := \binom{n}{k} D(n-k) \leq 2e^{-1}\frac{n!}{k!}. \tag{110}$$

The following estimate for certain sums involving $p(n,k)$ will shortly become useful. Note that we have for any $M, L \in \mathbb{N}$ and $m \in \mathbb{R}$ such that $2^m > M - L$ and $M \geq L \geq 1$:

$$\begin{aligned}\sum_{k=0}^{M-L} p(M,k) 2^{-m(M-k)} &\leq \frac{2}{e} \sum_{k=0}^{M-L} 2^{-mM} M! \frac{2^{mk}}{k!} \\ &\leq \frac{2}{e} 2^{-mM} (M-L+1) M! \frac{2^{m(M-L)}}{(M-L)!} \leq M^{L+1} 2^{-mL}. \end{aligned} \tag{111}$$

Here, we have used in the second inequality that $2^{mk}/k!$ is monotonically increasing for $k \leq M - L < 2^m$ and a standard bound on binomial coefficients in the last step.

We start by bounding the diagonal coefficients $A_{j,j}$. The idea is to divide the set of permutations into sets of permutations with exactly k fixed points. For any such permutation, the product of overlaps collapses to only $j - 1 - k$ non-trivial inner products. By assumption $n \geq \frac{1}{2}(t^2 + 5t) \geq t + \log_2 t$, thus we can be bound any of those using Lemma 11 as

$$\left(Q_T | Q_{T'}\right)^n \leq t 2^{t-n}, \quad \text{for all } T \neq T'. \tag{112}$$

Note that the trivial permutation (corresponding to $k = j - 1$ fixed points) contributes by exactly 1 to the sum. Thus, we find the following bound using Eq. (111) with $M = j - 1$, $L = 1$ and $m = n - t - \log_2 t$:

$$\begin{aligned} A_{j,j} = |A_{j,j}| &\leq \sum_{\pi \in S_{j-1}} \prod_{l=1}^{j-1} \left(Q_l | Q_{\pi(l)}\right)^n \\ &\leq 1 + \sum_{k=0}^{j-2} p(j-1,k) 2^{-(n-t-\log_2 t)(j-1-k)} \\ &\leq 1 + (j-1)^2 \, 2^{-n+t+\log_2 t} \\ &< 1 + 2^{t^2+7t-n}, \end{aligned} \tag{113}$$

where we have used Eq. (15) in the last step as $j - 1 < j \leq |\Sigma_{t,t}| \leq 2^{\frac{1}{2}(t^2+5t)}$. Using the reverse triangle inequality, we get a lower bound in the same way:

$$A_{j,j} = |A_{j,j}| \geq 1 - \left| \sum_{\pi \in S_{j-1} \setminus \mathrm{id}} \mathrm{sign}(\pi) \prod_{l=1}^{j-1} \left(Q_l | Q_{\pi(l)}\right)^n \right| \geq 1 - 2^{t^2+7t-n}. \tag{114}$$

Next, we will bound the off-diagonal terms $A_{i,j}$. It is well known that every permutation $\Pi \in S_j$ can be written as a product of disjoint cycles. Given a $\Pi \in S_j$ with $\Pi(j) = i$, consider the cycle $j \mapsto i \mapsto i_1 \mapsto i_2 \mapsto \ldots i_r \mapsto j$ in Π. Then, we have the bound

$$\begin{aligned} \prod_{l=1}^{j-1} \left(Q_{T_l} \middle| Q_{T_{\Pi(l)}}\right)^n &\leq \left(Q_{T_i} | Q_{T_{i_1}}\right)^n \ldots \left(Q_{T_{i_r}} | Q_{T_j}\right)^n \\ &\leq 2^{-n(|\dim N_i - \dim N_{i_1}| + \ldots + |\dim N_{i_r} - \dim N_j|)} \\ &\leq 2^{-n|\dim N_i - \dim N_j|}, \end{aligned} \tag{115}$$

where we have used Lemma 3, the triangle inequality and a telescope sum. We set $L := |\dim N_i - \dim N_j|$ and split the sum over permutations into those with more than or equal to $j - L$ many fixed points and those with less. In the first case, we use Eq. (115) to bound the overlaps, in the second case we use Eq. (111) as before. This yields the following bound

$$\begin{aligned}
|A_{i,j}| &\leq \sum_{\substack{\Pi\in S_j \\ \Pi(j)=i}} \prod_{l=1}^{j-1} \left(Q_{T_l} \middle| Q_{T_{\Pi(l)}}\right)^n \\
&\leq \sum_{k=j-L}^{j-1} p(j,k) 2^{-nL} + \sum_{k=0}^{j-L-1} p(j,k) 2^{-(n-t-\log_2 t)(j-1-k)} \\
&\leq \frac{2}{e} \sum_{k=j-L}^{j-1} \frac{j!}{k!} 2^{-nL} + 2^{n-t-\log_2 t} j^{L+2}\, 2^{-(n-t-\log_2 t)(L+1)} \\
&\leq L \frac{j!}{(j-L)!} 2^{-nL} + j^{L+2}\, 2^{-(n-t-\log_2 t)L} \\
&\leq L j^L 2^{-nL} + j^{L+2}\, 2^{-(n-t-\log_2 t)L} \\
&\leq L |\Sigma_{t,t}|^{L+2}\, 2^{-(n-t-\log_2 t)L} \\
&\leq 2^{\log_2 L} 2^{\frac{1}{2}(t^2+5t)(L+2)}\, 2^{(t+\log_2 t - n)L} \\
&= 2^{t^2+5t} 2^{(\frac{1}{2}t^2+\frac{5}{2}t+t+\log_2 t - n)L} \\
&\leq 2^{\frac{1}{4}t^3+\frac{11}{4}t^2+5t+(\frac{t}{2}+1)\log_2 t - nL} \\
&\leq 2^{t^3+4t^2+6t-n|\dim N_i - \dim N_j|},
\end{aligned} \tag{116}$$

where we have used again $j \leq |\Sigma_{t,t}|$ and $L \leq t/2$.

Note that we can alternatively bound $A_{i,j}$ for $i \neq j$ using that the identity is not an allowed permutation, i. e. only permutations with less than $j - 2$ fixed points can appear. With Eq. (111) and (112), we get the following inequality

$$\begin{aligned}
|A_{i,j}| &\leq \sum_{k=0}^{j-2} p(j,k) 2^{-(n-t-\log_2 t)(j-1-k)} \\
&\leq j^3 2^{-(n-t-\log_2 t)} \\
&\leq 2^{\frac{3}{2}t^2+\frac{15}{2}t+t+\log_2 t - n} \\
&\leq 2^{2t^2+10t-n}.
\end{aligned} \tag{117}$$

□

C. Proof of Haar symmetrization Lemma 13

Lemma 13 (Haar symmetrization)**.** *For all t and for all $T \in \Sigma_{t,t} \setminus S_t$, it holds that*

$$(Q_T|\, P_{\mathrm{H}}\, |Q_T) = 2^{-t}\, \|P_{\mathrm{H}}[r(T)]\|_2^2 \leq \frac{7}{8}, \tag{97}$$

where Q_T is as in eq. (21) and $P_{\mathrm{H}} = \Delta_t(\mu_{\mathrm{H}})$ is the t-th moment operator of the single-qubit unitary group $\mathrm{U}(2)$.

For an analysis of the tightness of the bound, see Appendix C. Recall that

$$P_{\mathrm{H}}[A] := \int_{U(2)} U^{\otimes t} A (U^\dagger)^{\otimes t} \mu_{\mathrm{H}}(U). \tag{118}$$

Let P_D be the Haar averaging operator, restricted to the diagonal unitaries. As it averages over a subgroup, P_D is a projection with range a super-set of P_{H}. By applying P_D to $r(T)$, we can turn the statement (97) from one involving *Hilbert space* geometry to one about the *discrete* geometry of stochastic Lagrangians. Indeed,

$$\begin{aligned}
2^{-t} \left\| P_{\mathrm{H}}[r(T)] \right\|_2^2 &= 2^{-t} \left\| P_{\mathrm{H}}[P_D[r(T)]] \right\|_2^2 \\
&\leq 2^{-t} \left\| P_D[r(T)] \right\|_2^2 \\
&= 2^{-t} \Big(r(T), P_D[r(T)] \Big) \\
&= 2^{-t} \sum_{(x,y)\in T} \sum_{(x',y')\in T} \big(\left|x\right\rangle\!\left\langle y\right|, P_D[\left|x'\right\rangle\!\left\langle y'\right|] \big) \\
&= 2^{-t} \sum_{(x,y)\in T} \sum_{(x',y')\in T} \big(\left|x\right\rangle\!\left\langle y\right|, \int_0^{2\pi} e^{i2\phi(h(x')-h(y'))} \left|x'\right\rangle\!\left\langle y'\right| \mathrm{d}\,\phi \big) \\
&= 2^{-t} |\{(x,y) \in T \,|\, h(x) = h(y)\}| \\
&= \Pr_{(x,y)}[h(x) = h(y)],
\end{aligned}$$

i.e. the overlap is upper-bounded by the probability that a uniformly sampled element (x, y) of T has components of equal Hamming weight.

We will bound the probability in slightly different ways for spaces T with and without defect spaces.

a. Case I: trivial defect subspaces In this case, $T = \{(Oy, y) \,|\, y \in \mathbb{F}_2^t\}$ for some orthogonal stochastic matrix O. The next proposition treats a slightly more general situation.

Proposition 4 (Hamming bound)**.** *Let $O \in \mathrm{GL}(\mathbb{F}_2^t)$. Assume O has a column of Hamming weight r. Then the probability that O preserves the Hamming weight of a vector y chosen uniformly at random from $\mathbb{F}_2^t$ satisfies the bound*

$$\Pr_y[h(Oy) = h(y)] \leq \frac{1}{2} + \begin{cases} 2^{-(r+1)} \binom{r+1}{(r+1)/2} & r \text{ odd} \\ 0 & r \text{ even}. \end{cases} \tag{119}$$

The bound in Eq. (119) decreases monotonically in r. Orthogonal stochastic matrices O satisfy $r = 1 \mod 4$, so the smallest non-trivial r that can appear is $r = 5$, for which the bound gives .65.

The proof idea is as follows: For each $y \in \mathbb{F}_2^t$, the two vectors $y, y + e_1$ differ in Hamming weight by ± 1. But, if $h(e_1) \neq 1$, then $h(Oy) - h(O(y + e_1))$ tends not to be ± 1. In such cases, O does not preserve weights for *both* y and $y + e_1$. Applying this observation to randomly chosen vectors, we can show the existence of many vectors for which O changes the Hamming weight.

Proof (of Proposition 4). Assume without loss of generality that the first r entries of Oe_1 are 1, and the remaing $t - r$ entries are 0.

Let y be a uniformly distributed random vector on $\mathbb{F}_2^t$, notice that also Oy, and $O(y+e_1)$ are uniformly distributed. Using the union bound, we find that

$$\begin{aligned}\Pr[h(Oy)=h(y)] &= 1-\Pr[h(Oy)\neq h(y)]\\ &= 1-\frac{1}{2}\big(\Pr[h(Oy)\neq h(y)]+\Pr[h(Oy+Oe_1)\neq h(y+e_1)]\big)\\ &\leq 1-\frac{1}{2}\Pr[h(Oy)\neq h(y)\ \vee\ h(Oy+Oe_1)\neq h(y+e_1)]\\ &= \frac{1}{2}+\frac{1}{2}\Pr[h(Oy)=h(y)\ \wedge\ h(Oy+Oe_1)=h(y+e_1)]\\ &\leq \frac{1}{2}+\frac{1}{2}\Pr[h(Oy)-h(Oy+Oe_1)=\pm 1].\end{aligned}$$

We would like to compute $\Pr[h(Oy)-h(O(y+e_1))=\pm 1]$. The vector $O(y+e_1)=O(y)+O(e_1)$ arises from $O(y)$ by flipping the first r components. This operation changes the Hamming weight by ± 1 if and only if the number of ones in the first r components of $O(y)$ equals $(r\pm 1)/2$. For even r, this condition cannot be met, and correspondingly $\Pr[h(Oy)-h(O(y+e_1))=\pm 1]=0$.

In case of odd r, this probability becomes

$$\begin{aligned}\Pr[h(Oy)-h(O(y+e_1))=\pm 1] &= 2^{-r}\binom{r}{(r-1)/2}+2^{-r}\binom{r}{(r+1)/2}\\ &= 2^{-r}\binom{r+1}{(r+1)/2}.\end{aligned} \tag{120}$$

□

b. Case II: non-trivial defect subspaces We now turn to Lagrangians T with a non-trivial defect subspace.

Proposition 5 (Defect Hamming bound)**.** *Let $\{0\}\neq N\subset\mathbb{F}_2^t$ be isotropic. There exists an $n\in N$ such that if x is chosen uniformly at random from $N^\perp$, then*

$$\Pr_{x\in N^\perp}[h(x)=h(x+n)]\leq\frac{3}{4}.$$

What is more, let T be a stochastic Lagrangian with non-trivial defect subspaces. Then, for an element (x,y) drawn uniformly from T, we have

$$\Pr_{(x,y)\in T}[h(x)=h(y)]\leq\frac{7}{8}.$$

Proof. Let $d=\dim N$. Consider a $t\times d$ column-generator matrix Γ for N. Permuting coordinates of $\mathbb{F}_2^t$ and adopting a suitable basis, there is no loss of generality in assuming that Γ is of the form

$$\Gamma=\begin{pmatrix}G\\ \mathbb{1}_d\end{pmatrix},\qquad G\in\mathbb{F}_2^{(t-d)\times d}.$$

Note that

$$\gamma=\big(\mathbb{1}_{t-d},\ G\big)$$

is a row-generator matrix for $N^\perp$. Indeed, the row-span has dimenion $t-d$ and the matrices fulfill

$$\gamma\Gamma = G + G = 0,$$

i.e. the inner product between any column of Γ and any row of γ vanishes. It follows that elements $n \in N$, $x \in N^\perp$ are exactly the vectors of respective form

$$n = (\underbrace{G\tilde{n}}_{t-d}, \underbrace{\tilde{n}}_{d}),\ \tilde{n} \in \mathbb{F}_2^d; \qquad x = (\underbrace{\tilde{x}}_{t-d}, \underbrace{G^T\tilde{x}}_{d}),\ \tilde{x} \in \mathbb{F}_2^{t-d}.$$

In particular, if x is drawn uniformly from $N^\perp$, then the first $t-d$ components are uniformly distributed in $\mathbb{F}_2^{t-d}$. For now, we restrict to the case where G has a column, say the first, with $r \neq 1$ non-zero entries. We then choose $n = (Ge_1, e_1)$ and argue as in Eq. (120) to obtain

$$\Pr_{x\in N^\perp}[h(x) = h(x+n)] \leq \sup_{1\neq r \text{ odd}} 2^{-r}\binom{r+1}{(r+1)/2} = \frac{3}{4} \qquad \text{(attained for } r = 3\text{)}. \tag{121}$$

We are left with the case where all columns of G have Hamming weight 1. (If N is a defect subspace, then Def. 7.1 implies that every column of Γ has Hamming weight at least 4. We treat the present case merely for completeness). As N is isotropic, the columns of Γ have mutual inner product equal to 0:

$$\Gamma^T\Gamma = 0 \qquad \Leftrightarrow \qquad G^TG = -\mathbb{1} = \mathbb{1} \mod 2.$$

It follows that all columns have to be mutually orthogonal standard basis vectors $e_i \in \mathbb{F}_2^{t-d}$. Thus, by permutating the first $t-d$ coordinates of $\mathbb{F}_2^t$, we can assume that G is of the form

$$G = \begin{pmatrix} \mathbb{1}_d \\ 0 \end{pmatrix}, \quad \Rightarrow \quad N = \{(\tilde{n} \oplus 0_{t-2d}, \tilde{n}) \,|\, \tilde{n} \in \mathbb{F}_2^d\}, \quad N^\perp = \{(\tilde{x}, \tilde{x}|_d) \,|\, \tilde{x} \in \mathbb{F}_2^{t-d}\},$$

where $\tilde{x}|_d$ denotes the restriction of $\tilde{x}$ to the first d components. Adding $n := (e_1 \oplus 0, e_1)$ to $x = (\tilde{x}, \tilde{x}|_d)$, the Hamming weight of the two parts change both by ± 1, giving $h(x+n) = h(x) \pm 2$. Thus, we have $\Pr[h(x) = h(x+n)] = 0$.

We have proven the first advertised claim. It implies the second one, as argued next. Let N be the left defect subspace of T. By Ref. [41, Prop. 4.17], we find the following.

- The restriction $\{x \,|\, (x, y) \in T \text{ for some } y\}$ equals $N^\perp$.
- The stochastic Lagrangian T contains $N \oplus 0$.

Assume that (x, y) is distributed uniformly in T. By the first cited fact, x is distributed uniformly in $N^\perp$. By the second fact, $(x+n, y)$ follows the same distribution as (x, y), for each $n \in N$. Thus, repeating the argument in the proof of Proposition 4, we find that for any fixed $n \in N$:

$$\begin{aligned}\Pr[h(x) = h(y)] &= 1 - \Pr[h(x) \neq h(y)] \\ &\leq 1 - \frac{1}{2}\Pr[h(x) \neq h(y) \ \vee \ h(x+n) \neq h(y)] \\ &\leq \frac{1}{2} + \frac{1}{2}\Pr[h(x) = h(x+n)] \leq \frac{7}{8}.\end{aligned}$$

□

D. Proof of Lemmas for Theorem 3

Lemma 5 (Relative $\varepsilon 2^{2tn}$-approximate Clifford t-designs)**.** *Suppose that $0 \leq \varepsilon < 1$ is such that $g_{\mathrm{Cl}}(\nu, t) \leq \varepsilon$. Then, ν is a relative $\varepsilon 2^{2tn}$-approximate Clifford t-design.*

Proof. This follows similar to Ref. [23, Lem. 4& Lem. 30]. Denote by $|\Omega_{2^n}\rangle$ the maximally entangled state vector on $\mathbb{C}^{2^n} \otimes \mathbb{C}^{2^n}$. The condition in (5) is equivalent to

$$(1-\varepsilon)\rho_{\mathrm{Cl}} \leq \rho_\nu \leq (1+\varepsilon)\rho_{\mathrm{Cl}}, \tag{122}$$

as an operator inequality, where

$$\rho_\nu := (\Delta_\nu \otimes \mathbb{1})(|\Omega_{2^n}\rangle\langle\Omega_{2^n}|)^{\otimes t} \qquad \text{and} \qquad \rho_{\mathrm{Cl}} := \rho_{\mu_{\mathrm{Cl}}}. \tag{123}$$

We have a decomposition of $(\mathbb{C}^{2^n})^{\otimes t}$ into irreducible representations of the Clifford group:

$$(\mathbb{C}^{2^n})^{\otimes t} \cong \bigoplus_\gamma C_\gamma \otimes L_\gamma, \tag{124}$$

where $\{C_\gamma\}$ is the set of all equivalence classes of irreducible representations of $\mathrm{Cl}(n)$ that appear in the t-th order diagonal representation, and L_γ are the corresponding multiplicity spaces (which by the double commutant theorem are irreducible representations of the commutant algebra –we have chosen L for Lagrangian). This implies that

$$|\Omega_{2^n}\rangle^{\otimes t} \cong \sum_\gamma \sqrt{\frac{\dim L_\gamma \dim C_\gamma}{2^{nt}}} |\gamma,\gamma\rangle \otimes |\Omega_{C_\gamma}\rangle \otimes |\Omega_{L_\gamma}\rangle, \tag{125}$$

where $|\Omega_{L_\gamma}\rangle$ and $|\Omega_{C_\gamma}\rangle$ denote maximally entangling state vectors on two copies of L_γ and C_γ, respectively. Indeed, observe that $|\Omega_{2^n}\rangle^{\otimes t} = 2^{-nt/2}\,\mathrm{vec}(\mathbb{1})$ and that the identity restricted to subspaces is just the identity on these subspaces. The prefactors then follow from normalizing the vectorized identity operators on the direct summands.

Since $\mathrm{Cl}(n)$ acts via multiplication on the spaces C_λ, this implies that

$$\begin{aligned} \rho_{\mathrm{Cl}} &= \int_{\mathrm{Cl}(n)} (U \otimes \mathbb{1})^{\otimes t} (|\Omega_{2^n}\rangle\langle\Omega_{2^n}|)^{\otimes t} (U^\dagger \otimes \mathbb{1})^{\otimes t} \\ &\cong \sum_\gamma \frac{\dim L_\gamma \dim C\gamma}{2^{nt}} (|\gamma\rangle\langle\gamma|)^{\otimes 2} \otimes \left(\frac{\mathbb{1}_{C\gamma}}{\dim C\gamma}\right)^{\otimes 2} \otimes |\Omega_{L_\gamma}\rangle\langle\Omega_{L_\gamma}|, \end{aligned} \tag{126}$$

where the second line follows from Schur's lemma and the fact that $\int U^{\otimes t} \bullet (U^\dagger)^{\otimes t}$ is trace preserving. The support of this operator is on the *symmetric subspace* $\vee^t(\mathbb{C}^{2^n} \otimes \mathbb{C}^{2^n})$ [23, Lem 30.1]. The minimal eigenvalue of this operator restricted to the symmetric subspace is

$$\min_\gamma \frac{\dim L_\gamma}{2^{nt} \dim C_\gamma}, \tag{127}$$

which we now lower bound. Let γ^* denote the optimizer. By Schur-Weyl duality, the diagonal action of $\mathrm{U}(2^n)$ on $(\mathbb{C}^{2^n} \otimes \mathbb{C}^{2^n})^{\otimes t}$ decomposes as $\oplus_\lambda U_\lambda \otimes S_\lambda$ where as usual U_λ are Weyl modules

and S_λ are Specht modules. Restricting this action to the Clifford group, the U_λ further decompose into irreducible representations

$$U_\lambda \simeq \bigoplus_{\gamma \in I_\lambda} C_\gamma \otimes \mathbb{C}^{d_{\lambda,\gamma}},$$

where I_λ is the spectrum of U_λ as a Clifford representation. Let Λ_0 be the set of all λ such that $\gamma^* \in I_\lambda$, then as a Clifford representation

$$(\mathbb{C}^{2^n} \otimes \mathbb{C}^{2^n})^{\otimes t} \simeq C_{\gamma^*} \otimes \Big(\bigoplus_{\lambda \in \Lambda_0} S_\lambda \otimes \mathbb{C}^{d_{\lambda,\gamma^*}} \Big) \oplus (\text{other irreducible representations }). \tag{128}$$

Thus, as a vector space, we have

$$L_{\gamma^*} = \bigoplus_{\lambda \in \Lambda_0} S_\lambda \otimes \mathbb{C}^{d_{\lambda,\gamma^*}}. \tag{129}$$

In particular, for any $\lambda \in \Lambda_0$ we have that $\dim C_{\gamma^*} \leq \dim U_\lambda$ and $\dim L_{\gamma^*} \geq \dim S_\lambda$. Thus we get the following bound for the minimal eigenvalue:

$$\frac{\dim L_{\gamma^*}}{2^{nt} \dim C_{\gamma^*}} \geq \min_{\lambda \in \mathrm{Part}(t,2^n)} \frac{\dim S_\lambda}{2^{nt} \dim U_\lambda} \geq 2^{-2nt}. \tag{130}$$

The rest of the proof follows as in Ref. [23, Lem. 4], mutatis mutandis. □

In order to prove Lemma 8 we make use of the following result by Nachtergaele [55] and Lemma 11 bounding certain sums of overlaps of the operators $r(T)$.

Lemma 14 (Nachtergaele [55, Thm. 3])**.** *Let $H_{[p,q]}$ for $[p,q] \subset [n] = \{1,\dots,n\}$ be a family of positive semi-definite Hamiltonians with support on $(\mathbb{C}^2)^{\otimes(q-p+1)} \subset (\mathbb{C}^2)^{\otimes n}$. Assume there is a constant $l \in \mathbb{N}$, such that the following conditions hold:*

1. *There is a constant $d_l > 0$ for which the Hamiltonians satisfy*

$$0 \leq \sum_{q=l}^{n} H_{[q-l+1,q]} \leq d_l H_{[1,n]}. \tag{131}$$

2. *There are $Q_l \in \mathbb{N}$ and $\gamma_l > 0$ such that there is a local spectral gap:*

$$\Delta\left(H_{[q-l+1,q]}\right) \geq \gamma_l, \quad \forall q \geq Q_l. \tag{132}$$

3. *Denote the ground state projector of $H_{[p,q]}$ by $G_{[p,q]}$. There exist $\varepsilon_l < 1/\sqrt{l}$ such that*

$$\left\| G_{[q-l+2,q+1]} \left(G_{[1,q]} - G_{[1,q+1]} \right) \right\|_\infty \leq \varepsilon_l, \quad \forall q \geq Q_l. \tag{133}$$

Then, it holds that

$$\Delta\left(H_{[1,n]}\right) \geq \frac{\gamma_l}{d_l} \left(1 - \varepsilon_l \sqrt{l}\right)^2. \tag{134}$$

While conditions 1) and 2) are merely translation-invariance with finit range of interactions and frustration-freeness in disguise, the third condition is highly non-trivial and involves knowledge of the ground-space structure. Usually, finding the ground space in a basis can be just as hard as computing the spectral gap in the first place. Fortunately, the ground space structure of the Hamiltonians $H_{n,t}$ is determined by the representation theory of the Clifford group. With little additional work, we obtain the following lemma about the ground space structure of our Hamiltonians.

Lemma 8 (Lower bound to spectral gap)**.** *Let the Hamiltonian $H_{n,t}$ be as in Eq.* (67) *and assume that $n \geq 12t$. Then, $H_{n,t}$ has a spectral gap satisfying*

$$\Delta(H_{n,t}) \geq \frac{\Delta(H_{12t,t})}{48t}. \tag{70}$$

Proof. We make use of the Nachtergaele lemma. We have to verify the three conditions of Lemma 14. As already stated in Ref. [55], the first two conditions hold directly for translation-invariant local Hamiltonians as in our case.

1. The first condition immediately follows from the fact that we consider a translation-invariant 2-local Hamiltonian. It is fulfilled for any choice of $l \geq 2$ and $d_l = l - 1$.

2. The second condition follows again for all $l \geq 2$ and the choice $Q_l = l$, since $H_{[q-l+1,q]}$ is a sum of positive semi-definite operators for all $q \geq l$ with spectrum that does not depend on q due to translation-invariance. Thus, we can set

$$\gamma_l := \Delta(H_{[q-l+1,q]}) > 0. \tag{135}$$

3. The third condition requires a calculation and a non-trivial choice of l. We have to bound the quantity

$$R_{q,l} := \left\| G_{[q-l+2,q+1]} \left(G_{[1,q]} - G_{[1,q+1]} \right) \right\|_\infty, \tag{136}$$

for all $q \geq Q_l = l$. Here, $G_{[p,q]}$ denotes the orthogonal projector onto the ground space of $H_{[p,q]}$. Note that this ground space is simply a suitable translation of the Clifford commutant $\mathrm{Cl}(k)'$ for $k = q - p + 1$ as shown in Lemma 7. Recall that it comes with a non-orthogonal basis $Q_T^{\otimes k}$, where

$$Q_T := \frac{r(T)}{\|r(T)\|_2} = 2^{-t/2} r(T), \quad T \in \Sigma_{t,t}. \tag{137}$$

Moreover, the projector $G_[p,q]$ is also simply a translation of the Clifford projector $P_{\mathrm{Cl}(k)}$ projecting onto $\mathrm{Cl}(k)'$. From the discussion in Sec. VI A, we know that the Clifford frame operator

$$S_{\mathrm{Cl}(k)} := \sum_T |Q_T)(Q_T|^{\otimes k}, \tag{138}$$

is a suitable approximation to $P_{\mathrm{Cl}(k)}$ when k is large enough. Concretely, we have by Lem. 12:

$$\left\| S_{\mathrm{Cl}(k)} - P_{\mathrm{Cl}(k)} \right\|_\infty \leq (-2^{-k}; 2)_{t-1} - 1. \tag{139}$$

Defining the shorthand notation $s_t(k) = (-2^{-k}; 2)_{t-1}$, we in particular get the bound

$$\left\| S_{\mathrm{Cl}(k)} \right\|_\infty \leq \left\| S_{\mathrm{Cl}(k)} - P_{\mathrm{Cl}(k)} \right\|_\infty + \left\| S_{\mathrm{Cl}(k)} \right\|_\infty \leq s_t(k), \tag{140}$$

Let us introduce the shorthand notation $G_q := G_{[1,q]} \equiv P_{\mathrm{Cl}(q)}$, $S_q = S_{[1,q]} \equiv S_{\mathrm{Cl}(q)}$, and $G_{q,l} := G_{[q-l+2,q+1]}$, $S_{q,l} := S_{[q-l+2,q+1]}$ for translations of the Clifford projector and frame

operator, respectively. Combining the above inequalities with the fact that $G_q - G_{q+1}$ is an orthogonal projection, we find

$$\begin{aligned} R_{q,l} &= \|G_{q,l}\,(G_q - G_{q+1})\|_\infty \\ &\leq \|(G_{q,l} - S_{q,l})(G_q - G_{q+1})\|_\infty + \|S_{q,l}(G_q - G_{q+1})\|_\infty \\ &\leq s_t(l) - 1 + \|S_{q,l}(S_q - S_{q+1})\|_\infty + \|S_{q,l}(G_q - S_q)\|_\infty + \|S_{q,l}(G_{q+1} - S_{q+1})\|_\infty \\ &\leq \|S_{q,l}(S_q - S_{q+1})\|_\infty + s_t(l) - 1 + s_t(l)\,(s_t(q) + s_t(q+1) - 2) \\ &\overset{q\geq l}{\leq} \|S_{q,l}(S_q - S_{q+1})\|_\infty + (s_t(l) - 1)\,(2s_t(l) + 1) \\ &= \left\| \sum_{T\in\Sigma_{t,t}} |Q_T)(Q_T|^{\otimes(q-l+1)} \otimes Y_T \right\|_\infty + (s_t(l) - 1)\,(2s_t(l) + 1)\,, \end{aligned} \tag{141}$$

where the operator Y_T can be straightforwardly computed as

$$Y_T := \sum_{T'\neq T} \Big((Q_{T'}|Q_T)^{l-1}\, |Q_{T'})(Q_T|^{\otimes(l-1)} \Big) \otimes \Big(|Q_{T'})(Q_{T'}| \left(\mathrm{id} - |Q_T)(Q_T|\right) \Big). \tag{142}$$

Invoking the synthesis operators

$$V_k = \sum_T \left|Q_T^{\otimes k}\right\rangle\langle e_T| : \ \mathbb{C}^{|\Sigma_{t,t}|} \longrightarrow \mathrm{Cl}(k)', \tag{143}$$

introduced in Lemma 12, one can bound the above norm as

$$\begin{aligned} \left\| \sum_T |Q_T)(Q_T|^{\otimes(q-l+1)} \otimes Y_T \right\|_\infty &= \left\| \sum_T V_{q-l+1}\, |e_T\rangle\langle e_T|\, V_{q-l+1}^\dagger \otimes Y_T \right\|_\infty \\ &\leq \left\| V_{q-l+1} V_{q-l+1}^\dagger \right\|_\infty \left\| \sum_T |e_T\rangle\langle e_T| \otimes Y_T \right\|_\infty \\ &= \|S_{q-l+1}\|_\infty \max_T \|Y_T\|_\infty \\ &\leq s_t(q-l+1)\,(s_t(l-1) - 1)\,. \end{aligned} \tag{144}$$

Thus, we arrive at

$$\begin{aligned} R_{q,l} &\leq s_t(q-l+1)\,(s_t(l-1) - 1) + (s_t(l) - 1)\,(2s_t(l) + 1) \\ &\leq s_t(1)\,(s_t(l-1) - 1) + (s_t(l) - 1)\,(2s_t(l) + 1)\,. \end{aligned} \tag{145}$$

For $l + 1 \geq t + \log_2(t)$, we can use Lemma 11 to get:

$$\begin{aligned} R_{q,l} &\leq t2^{t-l+1}\left(1 + t2^{t-1}\right) + t2^{t-l}\left(3 + t2^{t-l}\right) \\ &= t^2 2^{2t-l}\left(\frac{5}{t}2^{-t} + 2^{-l} + 1\right) \\ &\leq 4t^2 2^{2t-l}. \end{aligned} \tag{146}$$

Finally choose any $l \geq 4t + 4\log_2(t) + 6$, then we find

$$l \leq \frac{4^{l-2t}}{64t^2} \quad \Rightarrow \quad R_{q,l} \leq 4t^2 2^{2t-l} \leq \frac{1}{2\sqrt{l}} < \frac{1}{\sqrt{l}}, \quad \forall q \geq l. \tag{147}$$

In particular, we can choose $l = 12t$, $\varepsilon_l = 1/2\sqrt{l}$ to get the desired bound in Lemma 14 $\forall q \geq l$.

Hence, for the choices $l = 12t$, $d_l = l - 1$, $Q_l = l$, $\gamma_l = \Delta(H_{12t,t})$ and $\varepsilon_l = 1/2\sqrt{l}$, Lemma 14 gives the claimed bound on the spectral gap:

$$\Delta(H_{n,t}) \geq \frac{\gamma_l}{d_l}\left(1 - \varepsilon_l^2\sqrt{l}\right) \geq \frac{\Delta(H_{12t,t})}{48t}. \tag{148}$$

□

VII. SUMMARY AND OPEN QUESTIONS

We have found that a number of non-Clifford gates independent of the system size suffices to generate ε-approximate unitary t-designs.

This is surprising, conceptually interesting and practically relevant: After all, it is the main objective in quantum gate synthesis to minimize the number of non-Clifford gates in a circuit implementation of a given unitary. There are multiple open questions and ways to continue this work:

- Similar to the result in Ref. [23], the scaling in n is near to optimal, the scaling in t can probably be improved.

- Another natural open question is whether the condition $n = O(t^2)$ can be lifted. Notably, this is reminiscent to the situation discussed in Ref. [63], where the improved scaling can be proven only in the regime $t = o(n^{\frac{1}{2}})$.

- Our result holds for additive errors in the diamond norm. Our bounds can be used to obtain a quadratic advantage in the number of non-Clifford gates in Corollary 2. This still allows the density of non-Clifford gates to go to zero in the thermodynamic limit. It would be interesting to investigate whether the independence of the system size in the number of non-Clifford gates holds for relative errors.

- We strongly expect that the results can be generalized to qu$d$its for arbitrary d.

We hope the present work stimulates such endeavors.

VIII. ACKNOWLEDGEMENTS

We would like to thank Richard Kueng, Lorenz Mayer and Adam Sawicki for helpful discussions. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - Cluster of Excellence Matter and Light for Quantum Computing (ML4Q) EXC 2004/1 - 390534769, the ARO under contract W911NF-14-1-0098 (Quantum Characterization, Verification, and Validation), and the DFG (SPP1798 CoSIP, project B01 of CRC 183). The Berlin group has been supported in this work by the DFG (SPP1798 CoSIP, projects B01 and A03 of CRC 183, and EI 519/14-1) and the Templeton Foundation. This work has also

received funding from the European Union's Horizon2020 research and innovation programme under grant agreement No. 817482 (PASQuanS).

[1] J. Emerson, R. Alicki, and K. Zyczkowski, "Scalable noise estimation with random unitary operators," J. Opt. B **7**, S347–S352 (2005).

[2] E. Magesan, J. M. Gambetta, and J. Emerson, "Characterizing quantum gates via randomized benchmarking," Phys. Rev. A **85**, 042311 (2012), arXiv:1109.6887 [quant-ph].

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

[4] P. Hayden and J. Preskill, "Black holes as mirrors: quantum information in random subsystems," JHEP **0709**, 120 (2007).

[5] C. Dankert, R. Cleve, J. Emerson, and E. Livine, "Exact and approximate unitary 2-designs and their application to fidelity estimation," Phys. Rev. A **80**, 012304 (2009).

[6] C. Dankert, "MSc thesis, University of Waterloo," (2005), arXiv:quant-ph/0512217.

[7] D. Gross, K. M. R. Audenaert, and J. Eisert, "Evenly distributed unitaries: on the structure of unitary designs," J. Math. Phys. **48**, 052104 (2007).

[8] A. Ambainis, J. Bouda, and A. Winter, "Nonmalleable encryption of quantum information," J. Math. Phys. **50**, 042106 (2009).

[9] D. P DiVincenzo, D. W. Leung, and B. M. Terhal, "Quantum data hiding," IEEE, Trans. Inf Theory **48**, 3580–599 (2002).

[10] W. Matthews, S. Wehner, and A. Winter, "Distinguishability of quantum states under restricted families of measurements with an application to quantum data hiding," Commun. Math. Phys. **291**, 813–843 (2009).

[11] P. Sen, "Random measurement bases, quantum state distinction and applications to the hidden subgroup problem," IEEE Conference on Computational Complexity , 274–287 (2006).

[12] A. Hayashi, T. Hashimoto, and M. Horibe, "Reexamination of optimal quantum state estimation of pure states," Phys. Rev. A **72**, 032325 (2005).

[13] A. J. Scott, "Optimizing quantum process tomography with unitary 2-designs," J. Phys. A **41**, 055308 (2008), arXiv:0711.1017.

[14] H. Zhu and B.-G. Englert, "Quantum state tomography with fully symmetric measurements and product measurements," Phys. Rev. A **84**, 022327 (2011).

[15] I. Roth, R. Kueng, S. Kimmel, Y.-K. Liu, D. Gross, J. Eisert, and M. Kliesch, "Recovering quantum gates from few average gate fidelities," Phys. Rev. Lett. **121**, 170502 (2018).

[16] R. Kueng, H. Zhu, and D. Gross, "Distinguishing quantum states using Clifford orbits," (2016), arXiv:1609.08595.

[17] D. Gross, F. Krahmer, and R. Kueng, "A partial derandomization of PhaseLift using spherical designs," J. Fourier Anal. Appl. **21**, 229–266 (2015).

[18] O. Szehr, F. Dupuis, Marco Tomamichel, and R. Renner, "Decoupling with unitary approximate two-designs," New J. Phys. **15**, 053022 (2013).

[19] F. G. S. L. Brandao and M. Horodecki, "Exponential quantum speed-ups are generic," Quant. Inf. Comp. **13**, 0901 (2013).

[20] D. A. Roberts and B. Yoshida, "Chaos and complexity by design," JHEP **04**, 121 (2017).

[21] L. Masanes, A. J. Roncaglia, and A. Acín, "Complexity of energy eigenstates as a mechanism for equilibration," Phys. Rev. E **87**, 032137 (2013).
[22] E. Onorati, O. Buerschaper, M. Kliesch, W. Brown, A. H. Werner, and J. Eisert, "Mixing properties of stochastic quantum Hamiltonians," Commun. Math. Phys. **355**, 905–947 (2017).
[23] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, "Local random quantum circuits are approximate polynomial-designs," Commun. Math. Phys. **346**, 397–434 (2016).
[24] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, "Efficient quantum pseudorandomness," Phys. Rev. Lett. **116** (2016).
[25] R. Cleve, D. Leung, L. Liu, and C. Wang, "Near-linear constructions of exact unitary 2-designs," Quant. Inf. Comp. **16**, 0721–0756 (2015).
[26] A. W. Harrow and R. A. Low, "Random quantum circuits are approximate 2-designs," Commun. Math. Phys. **291**, 257–302 (2009), arXiv: 0802.1919.
[27] N. Hunter-Jones, "Unitary designs from statistical mechanics in random quantum circuits," (2019), arXiv:1905.12053.
[28] D. Gottesman, "An introduction to quantum error correction and fault-tolerant quantum computation," ArXiv:0904.2557.
[29] E. T. Campbell, B. M. Terhal, and C. Vuillot, "Roads towards fault-tolerant universal quantum computation," Nature **549**, 172–179 (2017).
[30] V. Veitch, A. H. Mousavian, D. Gottesman, and J. Emerson, "The resource theory of stabilizer quantum computation," New J. Phys. **16**, 013009 (2014).
[31] M. Howard and E. Campbell, "Application of a resource theory for magic states to fault-tolerant quantum computing," Phys. Rev. Lett. **118**, 090501 (2017).
[32] Z. Webb, "The clifford group forms a unitary 3-design," (2015), arXiv:1510.02769.
[33] H. Zhu, "Multiqubit clifford groups are unitary 3-designs," Phys. Rev. A **96**, 062336 (2017).
[34] R. Kueng and D. Gross, "Qubit stabilizer states are complex projective 3-designs," (2015), arXiv:1510.02767.
[35] H. Zhu, R. Kueng, M. Grassl, and D. Gross, "The Clifford group fails gracefully to be a unitary 4-design," ArXiv:1609.08172.
[36] J. Helsen, J. J. Wallman, and S. Wehner, "Representations of the multi-qubit clifford group," J. Math. Phys. **59**, 072201 (2018).
[37] Eiichi Bannai, Gabriel Navarro, Noelia Rizo, and Pham Huu Tiep, "Unitary t-groups," J. Math. Soc. Japan 10.2969/jmsj/82228222, advance publication.
[38] A. Sawicki and K. Karnas, "Universality of single qudit gates," Ann. Henri Poincaré , Volume 18, Issue 11, pp 3515–3552 (2017).
[39] R. Koenig and J. A. Smolin, "How to efficiently select an arbitrary Clifford group element," J. Math. Phys. **55**, 122202 (2014), arXiv: 1406.2170.
[40] S. Nezami and M. Walter, "Multipartite entanglement in stabilizer tensor networks," (2016), arXiv:1608.02595.
[41] D. Gross, S. Nezami, and M. Walter, "Schur-Weyl duality for the Clifford group with applications," (2017), arXiv:1712.08628.
[42] F. Montealegre-Mora and D. Gross, "Rank-deficient representations in howe duality over finite fields arise from quantum codes," (2019), arXiv:1906.07230.
[43] S. Zhou, Z.-C. Yang, A. Hamma, and C. Chamon, "Single T gate in a Clifford circuit drives transition to universal entanglement spectrum statistics," (2019), arXiv:1906.01079.
[44] S. Aaronson and D. Gottesman, "Improved simulation of stabilizer circuits," Phys. Rev. A **70**, 052328 (2004).

[45] P. Cwiklinski, M. Howodecki, M. Mozrzymas, L. Pankowski, and M. Studzinski, "Local random quantum circuits are approximate polnomial-designs - numerical results," J. Phys. A **46**, 305301 (2013).

[46] S. Bravyi, D. Browne, P. Calpin, E. Campbell, D. Gosset, and M. Howard, "Simulation of quantum circuits by low-rank stabilizer decomposition," Quantum **3**, 181 (2019).

[47] H. Pashayan, J. J. Wallman, and S. D. Bartlett, "Estimating outcome probabilities of quantum circuits using quasiprobabilities," Phys. Rev. Lett. **115**, 070501 (2015).

[48] M. Heinrich and D. Gross, "Robustness of magic and symmetries of the stabiliser polytope," Quantum **3**, 132 (2019).

[49] J. Seddon, B. Regular, H. Pashayan, Y. Ouyang, and E. Campbell, "Quantifying quantum speedups: improved classical simulation from tighter magic monotones," (2020), arXiv:2002.06181.

[50] F. G. S. L. Brandao, W. Chemissany, N. Hunter-Jones, R. Kueng, and J. Preskill, "Models of quantum complexity growth," (2019), arXiv:1912.04297.

[51] P. Varju, "Random walks in compact groups," Doc. Math. **18**, 1137–1175 (2013).

[52] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*, Cambridge Series on Information and the Natural Sciences (Cambridge University Press, 2000).

[53] R. A. Low, "Pseudo-randomness and Learning in Quantum Computation," arXiv:1006.5227 [quant-ph] (2010), arXiv: 1006.5227.

[54] W. G. Brown and L. Viola, "Convergence rates for arbitrary statistical moments of random quantum circuits," Phys. Rev. Lett. **104**, 250501.

[55] B. Nachtergaele, "The spectral gap for some spin chains with disrete symmetry breaking," Commun. Math. Phys. **175**, 565–606 (1996).

[56] P. Diaconis and L. Saloff-Coste, "Comparison techniques for random walk on finite groups," Ann. Probab. **21**, 2131–2156 (1993).

[57] S. Aaronson and D. Gottesman, "Improved simulation of stabilizer circuits," Phys. Rev. A **70**, 052328 (2004).

[58] B. Bhatia, "Matrix analysis," Springer Science & Business Media **169** (2013).

[59] G. Nebe, E. M. Rains, and N. J. A Sloane, "The invariants of the Clifford groups," (2001), arXiv:math/0001038v2.

[60] J. Bourgain and A. Gamburd, "A Spectral Gap Theorem in SU(d)," arXiv:1108.6264 [math] (2011), arXiv: 1108.6264.

[61] R. Mezher, J. Ghalbouni, J. Dgheim, and D. Markham, "Efficient approximate unitary t-designs from partially invertible universal sets and their application to quantum speedup," arXiv preprint arXiv:1905.01504 (2019).

[62] P. R. de Montmort, "Essay d'analyse sur les jex de hazard," seconde edition, Jacque Quillau, Paris (1753).

[63] Y. Nakata, C. Hirche, M. Koashi, and A. Winter, "Efficient quantum pseudorandomness with nearly time-independent Hamiltonian dynamics," Physical Review X **7** (2017).

[64] T. Bröcker and T. Dieck, *Representations of compact Lie groups*, Graduate Texts in Mathematics (Springer-Verlag).

[65] W. Fulton and J. Harris, *Representation theory*, edited by W. Fulton and J. Harris, Graduate Texts in Mathematics (Springer).

[66] R. Goodman and N. R. Wallach, *Symmetry, representations, and invariants*, edited by R. Goodman and N. R. Wallach, Graduate Texts in Mathematics (Springer).

[67] H. Zhu, R. Kueng, M. Grassl, and D. Gross, "The Clifford group fails gracefully to be a unitary 4-design," (2016), arXiv:1609.08172.

[68] G. B. Folland, "How to integrate a polynomial over a sphere," The American Mathematical Monthly **108**, 446–448 (2001).

Appendix A: Unitary t-designs

In the following, we review the concept of a *unitary t-design* [5–7], giving different but equivalent definitions which prove to be useful in different contexts. They also serve as starting point to explore connections to other mathematical fields, e. g. representation theory. To this end, let us introduce some notation. Define μ_{H} to be the (normalized) Haar measure on $\mathrm{U}(d)$ and let $\mathrm{Hom}_{(t,t)}(\mathrm{U}(d))$ be the space of homogeneous polynomials of degree t in both the entries of $U \in \mathrm{U}(d)$ as well as $\overline{U}$.

Definition 11 (Unitary t-design)**.** *A probability measure ν on $\mathrm{U}(d)$ is called a* unitary t-design *if the following holds for all $p \in \mathrm{Hom}_{(t,t)}(\mathrm{U}(d))$:*

$$\int_{\mathrm{U}(d)} p(U)\,\nu(U) = \int_{\mathrm{U}(d)} p(U)\,\mu_{\mathrm{H}}(U). \tag{A1}$$

A subset $D \subseteq \mathrm{U}(d)$ is called a unitary t-design*, if it comes with a probability measure ν_D which, continued trivially to $\mathrm{U}(d)$, is a unitary t-design. In particular, if D is finite, ν_D is usually taken to be the (normalized) counting measure.*

It might not come as a surprise that Def. 11 has not to be checked for any polynomial. Since any homogeneous polynomial $p \in \mathrm{Hom}_{(t,t)}(\mathrm{U}(d))$ can be linearized as

$$p(U) = \mathrm{Tr}\left(AU^{\otimes t,t}\right), \qquad U^{\otimes t,t} := U^{\otimes t} \otimes \overline{U}^{\otimes t}, \tag{A2}$$

the defining Eq. (A1) becomes

$$M_t(\nu) := \int_{\mathrm{U}(d)} U^{\otimes t,t}\,\nu(U) = \int_{\mathrm{U}(d)} U^{\otimes t,t}\,\mu_{\mathrm{H}}(U) =: M_t(\mu_{\mathrm{H}}). \tag{A3}$$

Thus ν is a unitary t-design if and only if its moment operator $M_t(\nu)$ agrees with the one of the Haar measure. Note that the operators $U^{\otimes t,t}$ are the matrix representation of the t-diagonal adjoint action $\mathrm{Ad}(U^{\otimes t}) = U^{\otimes t} \bullet (U^\dagger)^{\otimes t}$ with respect to the standard basis $|i\rangle\langle j|$ of $L(\mathbb{C}^d)$. Thus, this can be equivalently stated as equality of the twirls $\Delta_t(\nu) = \Delta_t(\mu_{\mathrm{H}})$ over the two measures.

A particularly fruitful theory of designs is possible in the case where the design (G, ν) itself constitutes a (locally compact) subgroup $G \subseteq \mathrm{U}(d)$ and ν is the normalized Haar measure on G. Following Ref. [37], we call these *unitary t-groups*. In this case, we see that Eq. (A3) implies that the trivial isotype of the representation $G \ni g \mapsto \mathrm{Ad}_g^{\otimes t}$ shall agree with the trivial isotype of $\mathrm{U}(d) \ni U \mapsto \mathrm{Ad}_U^{\otimes t}$. Since the trivial isotype exactly corresponds to the commutant of the respective diagonal representations $\tau_t : U \mapsto U^{\otimes t}$, this is equivalent to the statement that the commutant of the representation τ_t agrees with the commutant of the restriction $\tau_t|_G$. However, this is the case if and only if $\tau_t|_G$ decomposes into the same irreducible representations as τ_t.

Appendix B: Representations of the unitary group

The representation theory of the unitary group can be understood using the theory of highest weight for compact Lie groups, see, for example Refs. [64–66]. We present a short summary of the

part relevant to us here. Let ρ be an irreducible representation of $\mathrm{U}(d)$, and consider the restriction $\rho|_{D(d)}$ to the diagonal subgroup $D(d) \simeq (S^1)^{\times d}$ (which is a so-called *maximal torus* in $\mathrm{U}(d)$). In general, this is a reducible representation of $D(d)$. Since $D(d)$ is Abelian, $\rho|_{D(d)}$ decomposes into one-dimensional irreducible representations , i. e. characters of $D(d) \simeq (S^1)^{\times d}$. Those are of the form $\chi_u(\theta) := e^{iu^T\theta}$ for some vector $u \in \mathbb{Z}^d$, and thus we find

$$\rho|_{D(d)} \simeq \bigoplus_{u\in\mathbb{Z}^d} \chi_u \otimes \mathbb{1}_{m_u}, \tag{B1}$$

where $m_u \in \mathbb{N}$ are multiplicities. The vectors u for which $m_u \neq 0$ are called the *weights* of ρ. Introducing a lexicographical ordering of the weights, we call a weight u higher than the weight v if $u > v$. The *theorem of the highest weight* states that any irreducible representation ρ has a highest weight and that irreducible representations with the same highest weight are isomorphic. Thus, irreducible representations are unambiguously labeled by their highest weight. Next, let us consider the tensor product $\pi_u \otimes \pi_v$ of two irreducible representations labeled by their highest weights u and v. One can easily check that the weights of irreducible representations in $\pi_u \otimes \pi_v$ have to be sums of weights of π_u and π_v. In particular, the highest weight of all irreducible representations is at most $u + v$.

As a relevant example consider the (irreducible) defining representation $\rho : U \mapsto U$ of $U(2)$. Its restriction to the diagonal subgroup $S^1 \times S^1$ decomposes as

$$\rho|_{S^1\times S^1} \simeq \chi_{e_1} \oplus \chi_{e_2},$$

with highest weight $e_1 = (1, 0)$. Using $\bar{\chi}_u = \chi_{-u}$, the highest weight of the complex conjugate representation $\bar{\rho} : U \mapsto \bar{U}$ can be immediately determined as $(0, -1)$. Hence, the weights of $\rho \otimes \bar{\rho}$ are $\{(0, 0), (1, -1), (-1, 1)\}$. Here, $(0, 0)$ is the highest weight of the trivial irreducible representation and $(1, -1)$ the highest weight of the adjoint irrep. Finally, all irreducible representations appearing in $(\rho \otimes \bar{\rho})^{\otimes t}$ have weights w satisfying $(-t, t) \leq w \leq (t, -t)$ and, in particular,

$$w = \sum_{i=1}^{t} u_i$$

where $u_i \in \{(0, 0), (1, -1), (-1, 1)\}$. It follows that the Euclidean norm of these weights is at most $\sqrt{2}t$.

Appendix C: Converse bounds for estimates in Section VI C

Here, we collect various tightness results that limit the degree by which the estimates in Sec. VI C can be improved. The bound in Proposition 4 is tight in many cases. Most interestingly, the anti-identity [41]

$$\overline{\mathbb{1}} = \begin{pmatrix} 0 & 1 & \cdots & 1 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} \in O_t\,, \tag{C1}$$

meets the bound if both

$$r = t - 1 \qquad \text{and} \qquad t/2 = (r + 1)/2 \qquad \text{are odd.} \tag{C2}$$

Indeed, the anti-identity flips the components of the input if its parity is odd, and leaves the input invariant if the parity is even. The flipping step preserves the Hamming weight if and only if $h(a) = t/2$. Thus

$$\begin{aligned}\Pr[h(Oa) = h(a)] &= \Pr[h(a) \text{ even}] + \Pr[h(a) \text{ odd} \wedge h(a) = t/2] \\ &= \Pr[h(a) \text{ even}] + \Pr[h(a) = t/2] \qquad \text{(using (C2))} \\ &= \frac{1}{2} + 2^{-t}\binom{t}{t/2} \\ &= \frac{1}{2} + 2^{-(r+1)}\binom{r+1}{(r+1)/2}.\end{aligned}$$

Likewise, both estimates in Proposition 5 are tight. The first bound is saturated for $N = \{0, (1,1,1,1)\}$. Indeed, $N^\perp$ is the space of all even-weight elements of $\mathbb{F}_2^4$. The only non-trivial element of N is $(1,1,1,1)$ and adding it to an even-weight vector changes its weight if and only if the vector is in N itself. But $|N|/|N^\perp| = 1/4$. In an exactly analogous way, the second bound is tight for the stochastic Lagrangian with left and right defect spaces equal to the same N. As detailed in Example 4.27 of Ref. [41], this stochastic Lagrangian is the one identified in Ref. [67] as the sole non-trivial one in case of $t = 4$.

In contrast, we do not know (but suspect) that we pay a price by restricting from the full Haar symmetrizer to the one over diagonal matrices in Eq. (119). For the two cases that saturate the bounds in Proposition 4 and Proposition 5, we can compute the full projection explictily and show that at least there, Eq. (119) indeed fails to be tight.

One can expand the anti-id $\overline{\mathbb{1}}$ in terms of Pauli operators [41]

$$\overline{\mathbb{1}} = \frac{1}{2}\big(\mathbb{1}^{\otimes t} + X^{\otimes t} + Y^{\otimes t} + Z^{\otimes t}\big). \tag{C3}$$

Then

$$\begin{aligned}2^{-t}\big(r(\overline{\mathbb{1}}), P_H[r(\overline{\mathbb{1}})]\big) =& 2^{-t}\int \operatorname{Tr} r(\overline{\mathbb{1}})U^{\otimes t} r(\overline{\mathbb{1}})^\dagger (U^\dagger)^{\otimes t}\,\mathrm{d}U \\ =& 2^{-t-2}\sum_{i,j=0}^{3}\int \operatorname{Tr}\sigma_i^{\otimes t}U^{\otimes t}\sigma_j^{\otimes t}(U^\dagger)^{\otimes t}\,\mathrm{d}U \\ =& 2^{-t-2}\sum_{i,j}\int \Big(\operatorname{Tr}\sigma_i U\sigma_j U^\dagger\Big)^t\,\mathrm{d}U \\ =& 2^{-2} + 2^{-t-2}\sum_{i,j\neq 0}\int \Big(\operatorname{Tr}\sigma_i U\sigma_j U^\dagger\Big)^t\,\mathrm{d}U \\ =& 2^{-2} + 2^{-2}9\frac{1}{4\pi}\int_{S^2} x_1^t \mathrm{d}x \qquad \text{(C4)} \\ =& \frac{1}{4} + \frac{9}{4}\frac{1}{4\pi}\frac{4\pi}{1+t} = \frac{1}{4}\Big(1 + \frac{9}{t+1}\Big),\end{aligned}$$

where in (C4), we have interpreted the Haar integral over inner products of Paulis as an integral over the Bloch sphere and in the next line, used the formula from [68]. For $t = 2$, Eq. (C1) is just the swap operator (i.e. a permutation), and the formula gives 1, as it should. The smallest non-trivial case is $t = 6$ [41] , where we get roughly $0.571 < 0.65$.

Next, we consider the CSS code P_N for $N = (1, 1, 1, 1)$. We use the results in Section 3 of Ref. [67]. For a given partition λ, let W_λ be the associated Weyl module and S_λ the Schur module. As in Ref. [67], let $W_\lambda^+ \subset W_\lambda$ be the subspace such that

$$\big(W_\lambda \otimes S_\lambda\big) \cap \operatorname{range} P_N = W_\lambda^+ \otimes S_\lambda.$$

For the projection operators onto the various spaces, we write P_λ (Schur module), Q_λ (Weyl module), and Q_λ^+ (the subspace defined above). Then [67]

$$P_N = \sum_\lambda Q_\lambda^+ \otimes P_\lambda.$$

By Schur's Lemma,

$$P_H[P_N] = \sum_\lambda c_\lambda Q_\lambda \otimes P_\lambda,$$

for suitable coefficients c_λ, which are seen to equal $c_\lambda = D_\lambda^+ / D_\lambda$ by the fact that Haar averaging preserves the trace. Hence, using Table 1 of Ref. [67] for $d = 2$,

$$2^{-t+2\dim N}(P_N, P_{\mathrm{H}}[P_N]) = 2^{-2} \sum_\lambda \frac{d_\lambda (D_\lambda^+)^2}{D_\lambda} = \frac{7}{10} < \frac{7}{8}.$$

Part II

Numerical representation theory

4 Intoduction to Part II and summary of its results

Decomposing representations into irreducible blocks is the bread and butter of the whole field of representation theory. In many cases, as those discussed in Part I of this thesis, one may obtain full decompositions by the oldest method in the field: the pen-and-paper approach. For example this is the case when decomposing tensor power representations—be it of the unitary group, the Clifford group, or the metaplectic representation. Here, modifications of the theory of Schur-Weyl duality may be employed to describe the desired decomposition—sometimes in rather explicitly, as in the case of the duality between $\mathrm{U}(d)$ and S_t.

Roughly speaking, theoretical methods such as those presented in Part I may be used to analyse infinite sequences of representations. With regard to this, the results of Chap. 2 decompose Clifford t-th tensor powers for all numbers of qubits n no smaller than t. In many cases, however, the pen-and-paper approach might fail to give insightful results. Furthermore, sometimes the generality provided by this approach is not needed: sometimes decomposing a *single* representation is all that is required. When these two situations are encountered, *computational* methods provide a promising alternative.

The intersection of these two situations is realized in a widespread application—an application that, in particular, motivated this project: the symmetrization of semidefinite programs (SDPs). An SDP is an optimization problem of the sort

$$\max_{X\in\mathbb{C}^{n\times n}} \mathrm{tr}(XA_0) \quad \text{s.t.} \quad X \geq 0,\ \mathrm{tr}(XA_i) = a_i,\ i = 1,\ldots,k,$$

where $A_i \in \mathbb{C}^{n\times n}$ are Hermitian matrices and where $X \geq 0$ means that X is a *positive semi-definite* matrix. SDPs may be solved efficiently in principle, however the scaling of their complexity with n and k yields them impractical in many applications. This has been noted, in particular, within quantum information theory, where high-dimensional SDPs are often encountered.

For certain SDPs, the matrices A_i may be simultaneously block-diagonalized with block sizes which are much smaller than n [Val09]. In such cases, the optimization can be restricted to matrices X with the same block diagonal structure. In a nutshell, one large optimization problem is exchanged by many small ones, possibly leading to a significant improvement in runtime.

This block-diagonalization happens when the algebra generated by $\{A_i\}_i$ is *reducible*, or equivalently, when this algebra has a non-trivial commutant. Commonly in quantum information theoretical applications, such reductions are possible due to a group symmetry of the SDP problem. That is, these SDPs are such that there is a group

$G \subset \mathrm{U}(n)$ which commutes with the matrices A_i,

$$[g, A_i] = 0, \qquad \forall\, g \in G,\ i = 0, \ldots, k.$$

Assuming that this happens, let $U_G \in \mathrm{U}(n)$ be the change of basis that decomposes G, ie.

$$U_G g U_G^\dagger = \bigoplus_\rho \rho(g) \otimes \mathbb{1}_{d_\rho},$$

where ρ are irreducible representations of G, d_ρ are multiplicities, and $\otimes$ is the Kronecker product of matrices. Then, by Schur's lemma this same change of basis block diagonalizes the SDP,

$$U_G A_i U_G^\dagger = \bigoplus_\rho \mathbb{1}_\rho \otimes A_{i,\rho}.$$

Because of this, algorithms to efficiently decompose group representations offer a promising approach to solving the high-dimensional SDPs encountered in quantum problems.

In this second part of the thesis, I address this issue. I present a novel computational approach that may be used to decompose arbitrary representations of compact groups. This approach holds rigorous performance guarantees on the accuracy of the output, and its runtime asymptotically beats other state-of-the-art methods [MM11, CL20, MKKK10, MM10, dKDP11, CSX15, BFS93, BF91]. The algorithm was coded in two separate steps: first a heuristic, RepLAB [RB18], that finds a candidate decomposition, and a second *certification* step [MM21]. While the former was written by some of my colleagues, I have authored the certification step.

The results presented here come from two subsequent projects. The first concluded with the publication of RepLAB in [RMMB19, RB18]. In this project, I provided *stability calculations*, that is, calculations that suggest that RepLAB's output is stable against perturbations. These perturbations come from two different sources: First, one may only be able to specify the group elements up to floating point precision. Second, the algorithm ideally requires access to the group average operator,

$$\mathbb{C}^{n\times n} \ni X \mapsto \mathbb{E}_{g\sim G}[gXg^\dagger],$$

where the expectation value is with respect to the *Haar measure*. In practice however, there are instances where only approximations to this operator are available. These calculations, presented in Sec. 5.4, suggest that as long as these perturbations are small, RepLAB's output will be accurate.

An important note is that these results are not a rigorous proof of the claim. Indeed, the algorithm analysed in Sec. 5.4 is a simplification of the actual workflow in RepLAB. There, the algorithm is siginificantly more involved in order to optimize runtime and accuracy—this renders any analysis of this fuller version of the algorithm unwieldy. Furthermore, to prove stability of RepLAB it would be necessary to know the probability of eigenvalue near-collisions for finite-dimensional random matrices,

$$\Pr(|\lambda_1 - \lambda_2| \leq \delta \mid \lambda_i \text{ eigenval. of rand. mat. } H \in \mathbb{C}^{n\times n}). \tag{12}$$

These distributions are typically only known in the asymptotic limit, $n \to \infty$ (see e.g. [AB13a]), and obtaining such results for finite n is outside the scope of this thesis. Because of this, in Sec. 5.4 I assume for simplicity that the difference between (12) and the asymptotic distribution is negligible. I refer to this assumption as the "near asymptoticity" of H.

The goal of the second project was to obtain a rigorous proof of correctness for RepLAB's output. In this second phase, I take a dual approach to this problem. Rather than directly analysing RepLAB and proving its correctness, I propose an algorithm which, given a decomposition of G, certifies that this decomposition is close to the true decomposition. This project concluded with the draft [MMRBG21] included in Chap. 6 and the code RepCert (available at [MM21]) presented in Chap. 7.

These two projects fit together by providing an algorithm – namely the concatenation of RepLAB and RepCert – which decomposes G and has rigorously proven performance guarantees.

4.1 RepLAB's approach and its stability

The eigenspaces of a matrix in the commutant of G, say $H \in G'$, are invariant under the action of G. RepLAB's working principle is to sample the matrix H from a sufficiently well-behaved distribution. In Chap. 5 it is shown that the eigenspaces of H correspond, with probability one, to the irreducible blocks of G. In a nutshell, we could wave our hands and say that *sufficiently random samples have no accidental symmetries*.

As mentioned above, however, there are many instances where it is not feasible to exactly sample from the space G'. Rather, one typically is only able to sample matrices $\tilde{H}$ which are *close* to this space,

$$\|H - \tilde{H}\|_F \leq \alpha, \quad \text{for some } H \in G'.$$

This sample is diagonalized by RepLAB. Letting its eigenvalues be $\tilde{\lambda}_i$, these eigenval-

ues are grouped together whenever they are less than α apart, say

$$\Lambda_i = \{\tilde{\lambda}_j \mid |\tilde{\lambda}_j - \tilde{\lambda}_i| \leq \alpha\}.$$

Finally, using the definition

$$\mathcal{H}_{\Lambda_i} := \mathrm{span}\{|\psi_j\rangle \mid \tilde{H}\,|\psi_j\rangle = \tilde{\lambda}_j\,|\psi_j\rangle\,,\ \tilde{\lambda}_j \in \Lambda_i\},$$

the output decomposition is

$$\mathbb{C}^n = \bigoplus_{\Lambda_i} \mathcal{H}_{\Lambda_i}. \tag{13}$$

The stability of RepLAB's output can be roughly characterized by how slowly, as a function of α, do the eigenspaces of $\tilde{H}$ diverge from those of H. Crucial for this is that the eigenvalues of H need to be well separated—eigenvalue gaps smaller than α may lead to misidentifying the eigenspaces of H through the diagonalization of $\tilde{H}$.

Sec. 5.4 gives two bounds: *1.* a bound on α for a simple procedure that approximates the projection onto the commutant G', *2.* an approximate bound on (12) assuming that H is nearly asymptotic.

4.2 Certification of accuracy

At the beginning of the project leading up to [MMRBG21, MM21], the status was the following. We had produced a heuristic, RepLAB, which in practice could accurately decompose representations. We furthermore had strong indications (by the computations shown in Sec. 5.4) that this was to be expected: that, assuming near asymptoticity, RepLAB's working principle is stable against numerical perturbations. This notwithstanding, we did not yet have a rigorous proof of the correctness of RepLAB's output.

This is the situation which motivated the results in Chap. 6. As mentioned earlier, this chapter takes a dual formulation of the problem. Namely, given an alleged decomposition $\mathbb{C}^n = \oplus_i \mathcal{H}_i$ of the action of G, certify that it is accurate. The desired accuracy to be certified is measured by a parameter $\epsilon \in \mathbb{R}_+$, and a probability $p_{\text{thresh.}}$ of falsely certifying a representation (ie. a *false positive rate*) is allowed. Both these quantities are explicitly specified as the input to the algorithm.

Throughout Chap. 6, I only assume that one has access to imperfect projectors onto $\mathcal{H}_i$ and imperfect images $g \in G$. That is, the algorithm assumes access to these operators only up to machine precision. In this introductory section I will disregard this technical detail for the sake of clarity.

The certification algorithm involves two steps. The first step computes whether each subspace $\mathcal{H}_i \subset \mathbb{C}^n$ is *approximately invariant*, by which I mean that there exists

an invariant projector P_i such that

$$\|P_i - P_{\mathcal{H}_i}\|_F \leq \epsilon, \tag{14}$$

where $P_{\mathcal{H}_i}$ is the projector onto $\mathcal{H}_i$ and the minimization is over invariant projectors P_i. To do this, we may estimate the spectral distance between $P_{\mathcal{H}_i}$ and the G', ie. $\|P_{\mathcal{H}_i} - \mathrm{P}_{\mathrm{Haar}}(P_{\mathcal{H}_i})\|_\infty$, by computing

$$c := \left\| \frac{1}{r} \sum_{j=1}^{r} g_j P_{\mathcal{H}_i} g_j^\dagger - P_{\mathcal{H}_i} \right\|_\infty \tag{15}$$

for r Haar-random samples $g_j \in G$. Here the working principle is that if $P_{\mathcal{H}_i}$ is far from the commutant, it is exceedingly unlikely for (15) to be small. Namely, we use the Hoeffding bound to show that if

$$\|P_{\mathcal{H}_i} - \mathrm{P}_{\mathrm{Haar}}(P_{\mathcal{H}_i})\|_\infty \leq \epsilon', \tag{16}$$

fails to hold, then

$$\Pr(2c \leq \epsilon') \leq 2n \exp\left(\frac{-r}{8}\right).$$

By choosing $r \geq 8(\log(2n) + \log\big(p_{\text{thresh.}}^{-1}\big))$, we may certify – with a false-positive probability of at most $p_{\text{thresh.}}$ – that (16) holds. Using standard results from eigenvalue perturbation theory, this result yields an algorithm that certifies whether (14) holds.

The algorithm for this first step, furthermore, has a vanishing probability of *false negatives* if eq. (14) is a sufficiently loose bound. Specifically, if

$$\|P_i - P_{\mathcal{H}_i}\|_F \leq \frac{\epsilon}{2\sqrt{2 \operatorname{tr} P_i}},$$

then the probability that the algorithm *fails to certify* $P_{\mathcal{H}_i}$ up to accuracy ϵ is exactly zero.

The second step consists of certifying *irreducibility*. Namely, assuming that eq. (14) holds, we wish to certify that $\operatorname{range} P_i$ is an irreducible representation of G. Here, the dimension of the commutant of $\rho_i(g) := P_i g P_i$, given by

$$\mathbb{E}\left[|\operatorname{tr} \rho_i(g)|^2\right], \tag{17}$$

is estimated. By Schur's lemma this is a way to certify irreducibility. In Alg. 4.1, I show this working principle—a simplification of the full algorithm presented in Chap. 6. To keep matters simple in this introduction, I assume that the user has access to the exact invariant projector P_i, rather than its approximation $P_{\mathcal{H}_i}$. Furthermore, let

$n_i := \operatorname{tr} P_i$. Provided that the constants hidden by the $O(\,\cdot\,)$ notation in the algorithm are chosen correctly, I prove a performance guarantee for this algorithm. In the main text, these constants are specified explicitly (cf. Alg. IV.1 in Chap. 6).

The performance guarantee may be informally summarized as: *The probability that Alg. 4.1 falsely classifies the reducibility/irreducibility of ρ_i is at most $p_{\text{thresh.}}$.* That is, if ρ_i is reducible, the probability of the algorithm yielding *irreducible* as an output is at most $p_{\text{thresh.}}$. The same holds, vice-versa, when ρ_i is irreducible: the rate at which irreducible representations ρ_i is not certified is at most $p_{\text{thresh.}}$.

Algorithm 4.1 Irreducibility Certificate

Input: P_i, bound on the false positive rate $p_{\text{thresh.}}$.

Output: Irreducible/Reducible

1: Sample $r = O(\log n_i + \log p_{\text{thresh.}}^{-1})$ elements $g_j \in G$ and set $S = \{g_j\} \cup \{g_j^{-1}\}$.
2: Set $t = O(\log n_i)$.
3: Sample $m = O(n_i^2 \log p_{\text{thresh.}}^{-1})$ elements $\mathbf{s}_j$ from S^{2t} uniformly.
4: Set $\theta_m = n_i \sqrt{2/m} \log p_{\text{thresh.}}^{-1}$.
5: **if** $\mathbb{E}[|\operatorname{tr} \rho_i(\mathbf{s}_i)|^2] < 2(1 - \theta_m)$ **then**
6: **Return:** Irreducible
7: **end if**
8: **Return:** Reducible

As mentioned, Chap. 6 deals with a much less ideal case: one where the user only knows $P_{\mathcal{H}_i}$ (rather than P_i and thus the subrepresentation ρ_i), and can only evaluate approximations $\tilde{g} \approx g$. In this case, there is a qualitative difference in the performance guarantees obtained. Namely, while a rigorous bound on the false *positive* rate was obtained, only an approximate bound on the false *negative* rate is derived. This second approximate bound, called the *confidence parameter* $\delta_{\text{conf.}}$ in Chap. 6, depends on ϵ and $\epsilon_0 := \|\tilde{g} - g\|_{\max}$. It holds that

$$\lim_{\epsilon, \epsilon_0 \to 0} \delta_{\text{conf.}},$$

bounds the false negative rate of the algorithm in the case $\epsilon = \epsilon_0 = 0$. This way, it can be expected that for small enough values of ϵ and ϵ_0, the false negative is not much larger than $\delta_{\text{conf.}}$.

This difference is consistent with the goal of this research project. If the algorithm certifies that a decomposition is accurate, one would like this statement to be rigorously supported. On the other hand, an approximate bound on the false negative rate is used as a justification for the use of the method. Namely, it strongly suggests that not only does the method avoid falsely certifying inaccurate decompositions, but that it tends to recognize accurate ones. This way, while both bounds are useful, only the bound

on the false positive rate is relevant in terms of quality control: it is the bound on the probability of *falsely certifying* a decomposition.

5 Numerically decomposing representations

Let $G \subseteq \mathrm{U}(n)$ be a compact subgroup, seen as a representation of itself. The problem of decomposing G into irreducible blocks is classical, and several numerical methods exist for this [Dix70, BF91, MM11, CL20, MKKK10, MM10, dKDP11, AMB04, CL17, CSX15, CCS19, BFS93].

In this chapter I will explain on a high level the heuristic used by the software suite RepLAB [RMMB19, RB18] to decompose G. Additionally, I will show calculations that indicate that this heuristic is stable under perturbations coming from – for example – floating point precision.

The calculations I perform in this chapter are *not* a rigorous proof of correctness for RepLAB. Indeed this task would be unwieldy due to the fact that, while the working principle of RepLAB is simple, its actual implementation is rather complex due to runtime optimization. Furthermore, as we will see, these calculations involve rough estimates for the eigenvalue distributions of *finite-dimensional* random matrices. While these distributions are known asymptotically, there are no general guarantees on how well they approximate their finite-dimensional counterparts.

The calculations in this chapter, rather, give a glimpse of why one generally expects such a heuristic to work. This has been my main contribution to the project leading up to [RMMB19]. The problem of certifying the correctness of a RepLAB output is addressed in Chap. 6, where I provide an explicit algorithm for this.

5.1 Dixon's method

The commutant G' of G contains all the information about the decomposition of G: every projector P onto a G-invariant subspace is in this commutant. More generally, consider an arbitrary matrix $M \in G'$. Then, each eigenspace

$$\mathcal{M}_\lambda := \{\Psi \in \mathbb{C}^n \mid M\Psi = \lambda\Psi\}$$

is G-invariant, since for any $\Psi \in \mathcal{M}_\lambda$ and any $g \in G$,

$$Mg\Psi = gM\Psi = \lambda g\Psi,$$

so that $g\Psi \in \mathcal{M}_\lambda$. This gives the intuition that a full decomposition of G may be obtained by diagonalizing elements in G'.

This intuition forms the groundwork for the approach used by Dixon in his seminal paper on numerical representation decompositions [Dix70]. The basic tool here is the *Split routine*, Alg. 5.1. There, the basic assumption is that one has access to a function which projects operators onto the commutant G' of G.

Algorithm 5.1 Split routine

Input:

- $G \subseteq \mathrm{U}(n)$, matrix basis $\{E_i\}_{i=1}^{n^2} \subset \mathbb{C}^{n\times n}$.

1: For each i compute the average $\bar{E}_i = \mathbb{E}_g[gE_ig^\dagger] =: \mathrm{P}_{\mathrm{Haar}}(E_i)$ with respect to the Haar measure on G
2: **if** all $\bar{E}_i$ are proportional to the identity **then**
3: **Return:** "G Irreducible"
4: **end if**
5: Without loss of generality, let $\bar{E}_1$ be non-trivial. Let its eigenspaces be $\mathcal{H}_\lambda$ with corresponding projectors Π_λ.
6: **Return:** subrepresentations $\{G_\lambda = \Pi_\lambda G \Pi_\lambda \subseteq \mathrm{U}(\mathcal{H}_\lambda)\}_\lambda$.

Dixon's method is simply a recursion of Alg. 5.1: at any level of the recursion, the Split routine is called on each subrepresentation G_λ of the previous level's output. This is repeated until, at some recursion depth, all subrepresentations are irreducible and the algorithm terminates.

Dixon's algorithm may be generalized to the *imperfect case*, analyzed in [BF91] by Babai and Friedl. In that reference, a slight alteration of Dixon's method is shown to be *stable* under imperfections in the evaluation of group elements g and in the projection onto G' (performed in Line 1 of Alg. 5.1). Namely, it is shown that these small perturbations lead to proportionally small errors in the output decomposition. The reason why this is not obviously the case, is that at each level of Dixon's recursion, errors could accumulate. This leads to a possibly exponential increase of the error size with the recursion depth [BF91].

In this regard, [BF91] bounds the error generated in each run of Alg. 5.1 and bounds the depth needed by the recursion to recover a full decomposition. Furthermore, they provide a *self-improving* algorithm, which takes an imperfect decomposition and reduces its errors. Its main drawback, however, is runtime. In the worst case, it must (at least) diagonalize each of the n^2 projected basis elements $\bar{E}_i$, yielding a total runtime lower bounded by $O(n^5)$.

5.2 RepLAB's approach

The core of RepLAB's approach to decomposing representations is captured by Alg. 5.2. This method, presented in [RMMB19], is based on the same primitive as Alg. 5.1. Instead of recursively using the Split subroutine, however, it uses a single matrix diagonalization. The algorithm makes use of the *Gaussian unitary ensemble* (GUE), the

probability distribution on the subspace of Hermitian matrices in $\mathbb{C}^{n\times n}$ given by

$$\Pr_{\mathrm{GUE}(n)}(H) = c_n \exp\Big(-\frac{n}{2}\operatorname{tr} H^2\Big),$$

where $c_n = 2^{-n/2}\pi^{n^2/2}$ is a normalization. Equivalently: The strictly-upper triangular matrix elements of $H \sim \Pr_{\mathrm{GUE}(n)}$ are independent identically distributed (iid) complex Gaussian variables, and the diagonal elements of H are iid real Gaussian variables. These variables have a mean and variance of

$$\langle H_{ij}\rangle = 0, \quad \langle |H_{ij}|^2\rangle = \frac{1}{n}.$$

The GUE has, as the name suggests, a unitary symmetry

$$\Pr_{\mathrm{GUE}(n)}(UHU^\dagger) = \Pr_{\mathrm{GUE}(n)}(H), \qquad U \in \mathrm{U}(n).$$

Algorithm 5.2 RepLABSplit

Input: $G \subseteq \mathrm{U}(n)$.

1: Sample $H_0 \sim \Pr_{\mathrm{GUE}(n)}$
2: Project $H_0 \mapsto H \in G'$ (as in Line 1 of Alg. 5.1)
3: Diagonalize H, finding eigenspaces $\{\mathcal{H}_\lambda\}$
4: **Return:** decomposition $\mathbb{C}^n \simeq \oplus_\lambda \mathcal{H}_\lambda$.

Lemma 5.1. *With probability one, the output of Alg. 5.2 is a full decomposition of G into irreducible blocks.*

Proof. Suppose that the action of G has the following decomposition,

$$\mathbb{C}^n \simeq \bigoplus_\rho \rho \otimes \mathbb{C}^{d_\rho},$$

where ρ are irreducible representations of G and d_ρ are multiplicities. By Schur's lemma, the sample H may be unitarily block-diagonalized as

$$H \simeq \bigoplus_\rho \mathbb{1}_\rho \otimes H_\rho \simeq \bigoplus_\rho H_\rho^{\oplus \dim \rho}. \tag{18}$$

By unitary invariance, the distribution of the matrix elements $(H_0)_{ij}$ is independent of the orthonormal basis with respect to which these elements are computed. In particular, we may choose an orthonormal basis with respect to which G' block diagonalizes with irreducible blocks, as in the right-hand-side of eq. (18). In this basis, $\mathrm{P}_{\mathrm{Haar}}$ acts by setting matrix elements of H outside the block diagonal to zero, and setting the matrix elements of different blocks corresponding to the same $\rho \in \operatorname{Irr} G$ to be equal. In

particular each block H_ρ is proportional to a GUE variable, that is, there is a constant c_ρ such that $c_\rho H_\rho \sim \Pr_{\mathrm{GUE}(\dim \rho)}$.

This way, with probability one:

1. For any ρ, the matrix H_ρ has d_ρ distinct eigenvalues.

2. For any $\rho \not\simeq \rho'$, the spectra of H_ρ and $H_{\rho'}$ do not intersect.

□

Remark 5.1. *The constant c_ρ introduced above may be found for each block by requiring the variance of the matrix elements in $c_\rho H_\rho$ to be $\frac{1}{\dim \rho}$:*

$$c_\rho = \sqrt{\frac{n}{\dim \rho}}.$$

After finding a full decomposition, RepLAB uses a similar routine to group together equivalent representations: It produces a second sample $H' \in G'$, and computes

$$\Pi_\lambda H' \Pi_{\lambda'},$$

where Π_λ is the projector onto the space $\mathcal{H}_\lambda$ in the output decomposition of Alg. 5.2. By Schur's lemma, it follows that

$$\Pi_\lambda H' \Pi_{\lambda'} \neq 0 \tag{19}$$

only if these representations are equivalent. Conversely, if $\mathcal{H}_\lambda \simeq \mathcal{H}_{\lambda'}$, then with unit probability eq. (19) holds. This is shown in [RMMB19, Prop. 1].

5.3 Projecting onto the commutant

As mentioned, a primitive used in both Dixon's and RepLAB's approach is the projection onto the commutant G'. Ref. [BF91] approaches this by assuming that one has access to a sufficiently well-behaved generator set. RepLAB, on the other hand, uses a variety of heuristics for this—its choice of heuristic depends on the details of G. Describing the path to this decision taken by RepLAB is beyond the scope of this thesis. In this section, however, I will present a simple method that approximately projects onto G' assuming one can sample from the Haar measure on G.

Consider the random linear transformation given by

$$\Sigma_r : X \mapsto \frac{1}{2r} \sum_{i=1}^{r} g_i X g_i^\dagger + g_i^\dagger X g_i,$$

where g_i are independently sampled from the Haar measure. This transformation is a Markov estimate of $\mathbb{E}_g[gXg^\dagger]$. Then, a simple corollary of the matrix Bernstein inequality is the following [Gro19]:

Lemma 5.2 ([Gro19])**.** *Let $G \subset \mathrm{U}(n)$, G' and Σ_r be as above, and let $\mathrm{P}_{\mathrm{Haar}} : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ be the projector onto G', $\mathrm{P}_{\mathrm{Haar}}(X) = \mathbb{E}_g[gXg^\dagger]$. Then,*

$$\Pr\Big[\|\Sigma_r - \mathrm{P}_{\mathrm{Haar}}\|_\infty > \epsilon\Big] < 2n^2 \exp\left(\frac{-\epsilon^2 r}{3}\right).$$

Proof. Let $\Delta_{1,1}(g) = g \cdot g^\dagger$ and consider the operators

$$A_i := \frac{1}{2}(\Delta_{1,1}(g_i) + \Delta_{1,1}(g_i^\dagger)) - \mathrm{P}_{\mathrm{Haar}}.$$

Then A_i is Hermitian with zero expectation value, and by subadditivity and unitary invariance,

$$\|A_i\|_\infty \leq \frac{1}{n}, \quad \|\mathbb{E}_{g_i}[A_i^2]\|_\infty \leq \frac{1}{n^2}.$$

By [Tro12, Thm. 1.4],

$$\Pr\Big[\lambda_{\max}(\Sigma_r - \mathrm{P}_{\mathrm{Haar}}) > \epsilon\Big] \leq n^2 \exp\left(\frac{-\epsilon^2 r}{2(1+\frac{\epsilon}{3})}\right) < n^2 \exp\left(\frac{-\epsilon^2 r}{3}\right),$$

where $\lambda_{\max}$ refers to the largest eigenvalue. Then, repeating this argument for $-\Sigma_r$ and using the union bound, we obtain the claimed result. □

It should be noted that the speed of convergence of Σ_r to $\mathrm{P}_{\mathrm{Haar}}$ is rather slow, requiring $r \sim \epsilon^{-2}$ terms in order to acheive an error bound of ϵ. Because of this, several heuristic methods are used by RepLAB in order to increase its performance in practice. These alternative methods, however, do not come with guaranteed bounds on the speed of convergence and lie outside of the scope of this thesis.

5.4 Perturbations in the RepLAB approach

In any implementation of Alg. 5.2, there are two main sources of errors that could reduce the quality of the output.

First, the evaluation of group elements $g \in G$ could be limited by machine precision (denoted ϵ_0) if, e.g. the group is continuous. For finite groups, in principle representations can be computed exactly using *cyclotomic fields*, that is, fields of the form

$$\langle \mathbb{Q},\, \omega_N \rangle,$$

with $\omega_N = \exp(2i\pi/N)$. In particular, there is some basis of $\mathbb{C}^n$ with respect to which the matrix elements of any $g \in G$ are in the field $\langle \mathbb{Q}, \omega_{|G|} \rangle$ [BR90]. This fact is used by software packages such as GAP [GAP21] in order to exactly decompose the representations of finite groups. The problem with this is that computations are extremely costly for large groups, ruling out a variety of applications where reductions of representations of large finite groups are necessary. Furthermore, it is plausible that the group G would only be known in a less convenient basis, in which case one would first need to address the problem of finding the correct basis. It is fair to say that, in a wide range of applications, imperfect representations $\tilde{g}$ of group elements $g \in G$ are all but unavoidable.

Second, the step of projecting onto the commutant G' of the group will likely only be accessible in an imperfect form (e.g. using Σ_r instead of $\mathrm{P}_{\mathrm{Haar}}$ as in Sec. 5.3). While RepLAB is occasionally able to implement $\mathrm{P}_{\mathrm{Haar}}$ exactly for large finite groups (see the discussion surrounding eq. (9) in [RMMB19]), it still relies on approximations in the general case. In particular, the matrix $\tilde{H} := \Sigma_r(H_0)$ is *not* in the commutant, but rather *close to it*. This perturbs the eigenspaces and eigenvalues: the algorithm then needs to use the diagonalization of $\tilde{H}$ to infer as much information as possible about the eigenspaces of H. For example, RepLAB addresses this by grouping together eigenvalues of $\tilde{H}$ that are close enough to each other. For each one of these collections, it returns the span of the corresponding eigenvectors as the outputs. The question becomes: *how close is close enough?*

These two preceding paragraphs give the setting for this subsection. Namely, I want to address how much these two sources of errors are expected to affect the output of the algorithm. For this, I will first provide a slight rephrasing of Alg. 5.2 in order to accommodate for these sources of error. Then, I will use a simplifying assumption on the form of these errors in order to estimate the magnitude of the error in the output decomposition—i.e. how different it is from the true decomposition of G.

As mentioned in the introduction to this section, these calculations give a sense of how likely is it for RepLAB to give an accurate decomposition of G. In a second – more rigorous – step, I provide a certification algorithm which provides a guarantee that a claimed decomposition of G is close to exact. This second step is covered in Chap. 6.

5.4.1 Algorithm for approximate representations and projections

The following algorithm is a reformulation of Alg. 5.2, where the errors mentioned above are summarized by a single parameter δ. Namely, I assume that the user may

use an oracle $\Sigma \in \mathrm{End}(\mathbb{C}^{n\times n})$, which satisfies

$$\|\Sigma - \mathrm{P}_{\mathrm{Haar}}\|_\infty \leq \delta. \tag{20}$$

Algorithm 5.3 Imperfect RepLAB Split

Input: Σ, δ satisfying eq. (20).

Output: Decomposition of $\mathcal{H}$.

1: Sample a uniform random Gaussian matrix H_0
2: Compute $\tilde{H} = \Sigma(H_0)$
3: Diagonalize $\tilde{H}$, finding eigenspaces $\{\mathcal{H}_{h_i}\}_i$ corresponding to ordered eigenvalues $\{h_i\}$ (i.e. $h_i \leq h_{i+1}$ for all i).
4: **for** each $h_I = \big\{\{h_i\}_{i\in I}$ s.t. $|h_i - h_{i+1}| < \delta\|H_0\|_F,\ \forall\, i, i+1 \in I\big\}$ **do**
5: Set $\mathcal{H}_I = \mathrm{span}\{\mathcal{H}_{h_i}\}_{i\in I}$
6: **end for**
7: **Return:** decomposition $\mathcal{H} = \bigoplus_I \mathcal{H}_I$.

The map Σ can be implemented, for example, if the user can: *1.* sample uniformly from G, and, *2.* for each $g \in G$ evaluate an approximate image $\tilde{g}$ satisfying $\|g - \tilde{g}\|_F \leq \epsilon_0$. In this case, we can use $\Sigma = \tilde{\Sigma}_r$, where

$$\tilde{\Sigma}_r(X) = \frac{1}{2r}\sum_{i=1}^{r} \tilde{g}_i X \tilde{g}_i^\dagger + \tilde{g}_i^\dagger X \tilde{g}_i.$$

Vectorising $\mathbb{C}^{n\times n}$, we see that, equivalently

$$\tilde{\Sigma}_r = \frac{1}{2r}\sum_{i=1}^{r} \tilde{g}_i \otimes \tilde{g}_i^* + \tilde{g}_i^\dagger \otimes \tilde{g}_i^T.$$

Then, using subadditivity and Lem. 5.2, we see that if

$$\delta = \epsilon + 4rn\epsilon_0 + 2r(n\epsilon_0)^2,$$

then

$$\Pr\Big[\|\tilde{\Sigma}_r - \mathrm{P}_{\mathrm{Haar}}\|_\infty > \delta\Big] < 2n^2 \exp\left(\frac{-\epsilon^2 r}{3}\right).$$

5.4.2 The effect of perturbations

Up to now, I have shown an algorithm, Alg. 5.3, which uses a map Σ subject to condition (20). Furthermore, Sec. 5.3 shows one possible method for implementing, with high probability, such a Σ. We now turn to the question of quality: just how bad can

we, roughly speaking, expect the decomposition $\mathcal{H} \simeq \bigoplus_I \mathcal{H}_I$ obtained by Alg. 5.3 to be? In particular, how small must $\|\Sigma - \mathrm{P}_{\mathrm{Haar}}\|_\infty$ be for the decomposition $\mathcal{H} \simeq \bigoplus_I \mathcal{H}_I$ to be close to the true decomposition $\mathcal{H} \simeq \bigoplus_\rho \mathcal{H}_\rho^{\oplus d_\rho}$?

Consider the error matrix

$$\Delta H := \Sigma(H_0) - \mathrm{P}_{\mathrm{Haar}}(H_0),$$

which by assumption satisfies $\|\Delta H\|_F \leq \delta \|H_0\|_F$. By Weyl's perturbation theorem (see e.g. [Bha13, Chap. VI]), every eigenvalue $\tilde{\lambda}$ of $\tilde{H} := \Sigma(H_0)$ lies in a range $\lambda \pm \delta\|H_0\|_F$, where λ is an eigenvalue of $H := \mathrm{P}_{\mathrm{Haar}}(H_0)$.

For the moment let us assume that all eigenvalues of H are well separated:

$$\min_{\lambda \neq \lambda' \text{ eigs. of } H} |\lambda - \lambda'| := \mathrm{Gap}(H) \gg 2\delta\|H_0\|_F. \tag{21}$$

We will return to the validity of this assumption in the end of this subsection. Given eq. (21), each $\mathcal{H}_I$ found in Alg. 5.3 "corresponds" to an invariant space $\mathcal{H}_\lambda$ found by Alg. 5.2. Here, this correspondence is given by $I = \{\tilde{\lambda} \in \lambda \pm \delta\|H_0\|_F\}$, and the corresponding spaces satisfy

$$\dim \mathcal{H}_I = \dim \mathcal{H}_\lambda.$$

These spaces, $\mathcal{H}_I$ and $\mathcal{H}_\lambda$, are furthermore geometrically close to each other as long as $\mathrm{Gap}(H)$ is large enough. Namely, let U_I and U_λ be matrices whose columns are a basis for the subspaces $\mathcal{H}_I$ and $\mathcal{H}_\lambda$ respectively. (These matrices are in general rectangular, since their columns are vectors in $\mathbb{C}^n$.) We may quantify the distance between these two spaces by the *canonical angle matrix* [CL06],

$$\Theta(U_I, U_\lambda) = \arccos\Big(U_I^\dagger U_\lambda U_\lambda^\dagger U_I\Big)^{1/2}.$$

Then, [DK70, Sec. 2] (see also [CL06, Thm. 1.1]) implies that

$$\|\sin\Theta(U_I, U_\lambda)\|_F \leq \frac{\delta\|H_0\|_F}{\mathrm{Gap}(H) - \delta\|H_0\|_F}. \tag{22}$$

The conclusion is that, as claimed, the output of Alg. 5.3 is of high quality as long as assumption (21) is met.

5.4.3 Approximate bounds on the probability of near collisions

Let me finally turn to whether the assumption (21) is expected to be met or not. Bounds on the spectral gaps of random matrices are known, typically, only in the asymptotic regime $n \to \infty$, e.g. [AB$^+$13b, Cor. 1.6]. In finite dimensions, much less seems to be

known about these bounds.[2]

This caveat notwithstanding, I wish to give some sort of approximate upper bound on $\Pr(\mathrm{Gap}(H) \leq \alpha)$. For this, I will assume that the error incurred by assuming that the eigenvalues of H_ρ are distributed according to the limiting distribution is negligible. I refer to this assumption as *near asymptoticity*.

To complement this approach, I also provide numerical evidence that $\Pr(\mathrm{Gap}(H) \leq \alpha)$ is low in Sec. 5.4.4. This numerical evidence does not assume near asymptoticity.

Recall that if

$$\mathbb{C}^n \simeq \bigoplus_\rho \rho \otimes \mathbb{C}^{d_\rho}$$

is the decomposition as a G representation, we may block-diagonalize

$$H = \bigoplus_\rho \mathbb{1}_\rho \otimes H_\rho = \bigoplus_\rho H_\rho^{\oplus \dim \rho},$$

where H_ρ is $d_\rho \times d_\rho$. The matrix H has a small gap if either of the following situations arise: *1.* two eigenvalues λ_ρ and λ'_ρ of a single block H_ρ are close, or *2.* two eigenvalues λ_ρ and $\lambda_{\rho'}$ of distinct blocks H_ρ and $H_{\rho'}$ are close. The approximate bound on $\Pr(\mathrm{Gap}(H) \leq \alpha)$ that I derive here will have contributions from these two possibilities.

Assuming near asymptoticity, by [AB$^+$13b, Cor. 1.6] we find that for each $d_\rho \times d_\rho$ block H_ρ of H,

$$\Pr(\mathrm{Gap}(H_\rho) \leq \alpha) = \int_0^{c_\rho \alpha} \mathrm{d}x (4 - x^2)^2,$$

where c_ρ is as in Rem. 5.1. A short calculation shows that

$$\Pr(\mathrm{Gap}(H_\rho) \leq \alpha) = 16 c_\rho \alpha - 4 (c_\rho \alpha)^3 + \frac{1}{5} (c_\rho \alpha)^5 =: f_1(\alpha, \rho).$$

We now bound the probability that for two distinct blocks H_ρ, $H_{\rho'}$, with $\rho \ncong \rho'$, there exist respective eigenvalues λ_ρ, $\lambda_{\rho'}$ such that

$$|\lambda_\rho - \lambda_{\rho'}| \leq \alpha.$$

For this—relying on near asymptoticity—, we may use the *semi-circle law*, which

[2] See [Ver10, Tei20] for relatively recent accounts of the study of eigenvalue statistics of finite dimensional random matrices.

implies that any single eigenvalue of H_ρ has a marginal distribution

$$c_\rho \lambda_\rho \sim \frac{1}{2\pi}\sqrt{4-(c_\rho\lambda_\rho)^2}, \quad |c_\rho\lambda_\rho| \leq 2$$

Because λ_ρ and $\lambda_{\rho'}$ are independent, then using the variable changes $\lambda = c_\rho\lambda_\rho$ and $\lambda' = c_\rho\lambda_{\rho'}$—where the c_ρ appearing on the second definition is not a typo—,

$$\Pr(|\lambda_\rho - \lambda_{\rho'}| \leq \alpha) = \frac{1}{4\pi^2}\int_{-2}^{2} \mathrm{d}\lambda\sqrt{4-\lambda^2}\int_{I_{\lambda,\alpha}} \mathrm{d}\lambda'\sqrt{4-\frac{\dim\rho}{\dim\rho'}\lambda'^{\,2}}.$$

Here, I have used the definition

$$I_{\lambda,\alpha} := [\lambda - c_\rho\alpha, \lambda + c_\rho\alpha] \cap \left[-2\sqrt{\frac{\dim\rho'}{\dim\rho}}, 2\sqrt{\frac{\dim\rho'}{\dim\rho}}\right].$$

In the limit where $c_\rho\alpha \leq \sqrt{n}\alpha \ll 1$, the integrand of the inner integral is approximately constant as a function of λ', and evaluates to

$$\approx \sqrt{4-\frac{\dim\rho}{\dim\rho'}\lambda^2}.$$

Without loss of generality, we may take $b := \frac{\dim\rho}{\dim\rho'} \leq 1$, to obtain the bound

$$\Pr(|\lambda_\rho - \lambda_{\rho'}| \leq \alpha) \approx \frac{|I_{\alpha,\lambda}|}{4\pi^2}\int_{-2}^{2}\mathrm{d}\lambda\sqrt{4-\lambda^2}\sqrt{4-b\lambda^2} \leq \frac{c_\rho\alpha}{2\pi^2}\int_{-2}^{2}\mathrm{d}\lambda\sqrt{4-\lambda^2}\sqrt{4-b\lambda^2},$$

where I used $|I_{\alpha,\lambda}| \leq 2c_\rho\alpha$. The expression on the right-hand-side can be symbolically integrated to

$$\frac{8c_\rho\alpha}{3b\pi^2}\big((b+1)E(b) + (b-1)K(b)\big),$$

where $K(b)$ and $E(b)$ are, respectively, the *complete elliptic integrals of the first and second kind* with parameter b. This expression varies between approximately $0.64c_\rho\alpha$ for $b \approx 0$ and approximately $0.54c_\rho\alpha$ for $b = 1$. For our purpose here, it is sufficient to bound

$$\Pr(|\lambda_\rho - \lambda_{\rho'}| \leq \alpha) \leq c_\rho\alpha =: f_2(\rho, \rho', \alpha). \tag{23}$$

Note that, while this bounding strategy produces a function f_2 which does not depend on ρ', I keep the notation in eq. (23) for conceptual clarity.

Using the union bound, and taking only the linear contribution in f_1

$$\Pr\left(\min_{\lambda,\lambda'}|\lambda-\lambda'|\leq\alpha\mid\lambda,\ \lambda'\text{ eigenvals. of }H\right)\leq\sum_{\rho}f_1(\rho,\alpha)+\sum_{\rho\neq\rho'}f_2(\rho,\rho',\alpha)$$
$$\lesssim\alpha(16+N_G)\sum_{\rho}c_\rho\leq\alpha\sqrt{n}(16+N_G)N_G,$$

where N_G is the number of non-equivalent irreducible representations appearing in the decomposition of G, and where the approximate inequality "$\lesssim$" should be interpreted as *"approximately equal to a quantity which is bounded by."*

We may now assess whether it is plausible that condition (21) is met. For this, I take the *ad hoc* value of $\alpha = 20\ \delta\|H_0\|_F$, so

$$\Pr\left(\min_{\lambda,\lambda'}|\lambda-\lambda'|\leq\alpha\mid\lambda,\ \lambda'\text{ eigenvals. of }H\right)\lesssim 20\delta\sqrt{n}\|H_0\|_F(16+N_G)N_G.$$

Reorganising factors in this expression, we expect that condition (21) is satisfied with high probability whenever

$$\delta\ll\frac{1}{20\sqrt{n}\|H_0\|_F(16+N_G)N_G}.$$

For a quick rule of thumb, we may replace $\|H_0\|_F$ with its expectation

$$\langle\|H_0\|_F\rangle=\sqrt{\frac{2}{n^2}}\times\frac{\Gamma(\frac{n^2+1}{2})}{\Gamma(\frac{n^2}{2})}=:h(n),$$

where I used that $\sqrt{n}\|H_0\|_F$ is a standard chi-squared variable with parameter n^2. This replacement gives the condition:

$$\delta\ll\frac{1}{20\sqrt{n}h(n)(16+N_G)N_G}. \tag{24}$$

Such a choice of δ would guarantee that with high probability, eq. (21) holds. In the following, I provide a slightly smaller upper bound on δ, which however has a more simple expression as a function of n: A short calculation shows that $h(n)$ is strictly increasing for $n\geq 1$, and using the Stirling asymptotic formula for Γ, we find that $\lim_{n\to\infty}h(n)=1$. That is, $h(n)\leq 1$ for all n. A sufficient condition, thus, for δ to be small enough for eq. (21) to hold (assuming near asymptoticity), is

$$\delta\ll\frac{1}{20\sqrt{n}(16+N_G)N_G}. \tag{25}$$

5.4.4 Numerical benchmark for the collision probability

The argument in Sec. 5.4.3 relied on the assumption that the asymptotic eigenvalue statistics of GUE matrices approximates sufficiently well the finite-dimensional case. There, I used the asymptotic eigenvalue distributions in order to evaluate the probability that two eigenvalues of the (finite-dimensional) matrix H are too close together. Here I provide a complementary numerical approach. Specifically, I implement an algorithm that samples random block-diagonal matrices (such as H) and count how many samples contain eigenvalue gaps below a certain threshold α.

These numerical tests are summarized in Alg. 5.4. In a nutshell, this algorithm samples block-diagonal matrices (where the sizes of the blocks in each sample are themselves random), and proceeds to count how many of these sampled matrices have a pair of eigenvalues λ_1, λ_2 that satisfy $|\lambda_1 - \lambda_2| \leq \alpha$.

More specifically, Alg. 5.4 uses a function BD – the name stands for *block diagonalizer* – which takes a matrix H_0 and a partition of its dimension Λ_n, and creates a block matrix by setting certain off-block-diagonal elements to zero. This is done in the following way: for an partition $\{n_1, \ldots, n_k\} = \Lambda_n$, with $n_i \in \mathbb{Z}_+$ and $\sum_i n_i = n$, let $I_{\Lambda_n} \subset \mathbb{Z}_+^2$ be given by

$$I_{\Lambda_n} = \{(m_1, m_2) \mid m_1, m_2 \in \{n_i, \ldots,\ n_i + n_{i+1} - 1\} \text{ for some } i\ \}.$$

Then, $BD(H_0, \Lambda_n) = H$, where

$$H_{ij} = \begin{cases} (H_0)_{ij}, & \text{if } (i,j) \in I_{\Lambda_n}, \\ 0, & \text{else.} \end{cases} \tag{26}$$

Alg. 5.4 estimates the probability that $\mathrm{Gap}(H)$ is smaller than some threshold $\alpha \in \mathbb{R}_+$. Notice that if H would have been the outcome of G-averaging, i.e. $H = \mathrm{P}_{\mathrm{Haar}}(H_0)$ as above, then $|\Lambda_n| := k$ would be the number of non-equivalent irreducible blocks in G. That is, it would hold that $k = N_G$. This follows from the fact that the samples H are such that all their blocks are independent of each other.

Algorithm 5.4 Eigenvalue clash

Input:

- Dimension n,
- distribution p on set of partitions of n
- threshold $\alpha \in \mathbb{R}_+$
- number of data points N_{data}.

1: Sample N_{data} partitions $\Lambda_n \sim p$ and matrices $H_0 \sim \Pr_{\text{GUE}(n)}$
2: For each sample compute $x = \text{Gap}(BD(H_0, \Lambda_n))$
3: **Return** frequency of $x \leq \alpha$.

I have tested three values of α. Moreover, the distribution p from which I sample partitions Λ_n is defined by:

1. Sample $k_{\max} \in \{1, \ldots, n\}$ uniformly. This is the maximal number of elements in Λ_n,
2. Sample $n_1 \in \{1, \ldots, n\}$ uniformly,
3. Sample $n_i \in \{1, \ldots, n - \sum_{j<i} n_j\}$ uniformly,
4. Continue similarly until either $n - \sum_{j<i} n_j = 0$ for some i, or until $i = k_{\max}$.
5. If the latter is the case, set $n_{k_{\max}} = n - \sum_{i<k_{\max}} n_i$.

This procedure does not sample partitions of n uniformly, however I believe that the results would not significantly change if these partitions were sampled differently.

The results are shown in Table 1. There, F_α denotes the fraction of samples H_i for which $\text{Gap}(H_i) \leq \alpha$. In this table, I additionally show the sample average of the number of elements in the partition $\mathbb{E}[|\Lambda_n|]$. This gives a rough idea of how many blocks do the samples typically have.

As we can see from the table, obtaining eigenvalue gaps below 10^{-6} is rather unlikely even for high-dimensional matrices. We obtain the following rule of thumb: Choosing a Σ for which $\|\text{P}_{\text{Haar}} - \Sigma\|_\infty \ll 10^{-6}$ makes it likely that condition (21) holds. This way, one can expect that if one obtains a decomposition

$$\mathcal{H} \simeq \oplus_I \mathcal{H}_I$$

using a sample $H = \Sigma(H_0)$ as in Sec. 5.4.1, then it is likely that the subspaces $\mathcal{H}_I$ in this decomposition are close to the spaces in the true decomposition in the sense of eq. (22).

n	α	$\mathbb{E}[\lvert\Lambda_n\rvert]$	F_α
50	10^{-8}	6.175	0/2000
50	10^{-7}	6.175	0/2000
50	10^{-6}	6.175	0/2000
100	10^{-8}	6.656	0/2000
100	10^{-7}	6.656	0/2000
100	10^{-6}	6.656	0/2000
150	10^{-8}	7.432	0/2000
150	10^{-7}	7.432	0/2000
150	10^{-6}	7.432	0/2000
200	10^{-8}	7.798	0/2000
200	10^{-7}	7.798	0/2000
200	10^{-6}	7.798	1/2000
250	10^{-8}	7.962	1/2000
250	10^{-7}	7.962	1/2000
250	10^{-6}	7.962	1/2000
300	10^{-8}	8.150	0/2000
300	10^{-7}	8.150	0/2000
300	10^{-6}	8.150	1/2000
350	10^{-8}	8.387	0/2000
350	10^{-7}	8.387	1/2000
350	10^{-6}	8.387	3/2000

Table 1: Empirical estimates of the probability of a small $\mathrm{Gap}(H)$. These are the results of Alg. 5.4 for different values of n and α. For each value of n, $N_{\text{data}} = 2000$ matrices H were sampled. Here $\mathbb{E}[|\Lambda_n|]$ refers to the *sample average* of the partition length. These samples were filtered according to how many had $\mathrm{Gap}(H) \leq \alpha$, for each value of α, leading to the faction F_α.

6 Certifying numerical decompositions

This chapter is the preprint [MMRBG21]

> Montealegre-Mora, F., Rosset, D., Bancal, J. D., Gross, D. (2021). Certifying Numerical Decompositions of Compact Group Representations. arXiv preprint arXiv:2101.12244.

I have furthermore released a Python implementation of the algorithms presented here in [MM21]. This work was the result of a collaborative effort with my co-authors, I was the lead researcher and programmer in this project.

This project emerged out of the work presented in Chap. 5. Specifically, the goal was to provide a formal guarantee of correctness of the output of the RepLAB numerical representation theory software suite. As seen in Chap. 5, proving such a guarantee by directly analysing the algorithm is frustrated by two roadblocks. First, such a guarantee would require proving bounds on the eigenvalue separations for random matrices of a fixed dimension. Such results are typically only known asymptotically, in the limit of large dimensions. Second, while the working principle of RepLAB is simple, the algorithm itself is rather complex due to runtime optimization. In practice, analyzing such an algorithm is unwieldy.

These frustrations lead to the alternative method presented here. Ref. [MMRBG21] instead proposes an algorithm that – given a numerical decomposition of a representation – certifies that this decomposition is close to exact. While this algorithm was inspired by the difficulties summarized above, it is logically independent RepLAB. Moreover, it was coded independently of RepLAB and is available at [MM21]. This implementation is summarized and benchmarked in Chap. 7.

Certifying Numerical Decompositions of Compact Group Representations

Felipe Montealegre-Mora,[1] Denis Rosset,[2] Jean-Daniel Bancal,[3] and David Gross[1]

[1] *Institute of Theoretical Physics, University of Cologne, Germany*
[2] *Perimeter Institute of Theoretical Physics, Waterloo, Canada.*
[3] *Université Paris-Saclay, CEA, CNRS, Institut de Physique Théorique, 91191, Gif-sur-Yvette, France*

We present a performant and rigorous algorithm for certifying that a matrix is close to being a projection onto an irreducible subspace of a given group representation. This addresses a problem arising when one seeks solutions to semi-definite programs (SDPs) with a group symmetry. Indeed, in this context, the dimension of the SDP can be significantly reduced if the irreducible representations of the group action are explicitly known. Rigorous numerical algorithms for decomposing a given group representation into irreps are known, but fairly expensive. To avoid this performance problem, existing software packages – e.g. RepLAB, which motivated the present work – use randomized heuristics. While these seem to work well in practice, the problem of to which extent the results can be trusted arises. Here, we provide rigorous guarantees applicable to finite and compact groups, as well as a software implementation that can interface with RepLAB. Under natural assumptions, a commonly used previous method due to Babai and Friedl runs in time $O(n^5)$ for n-dimensional representations. In our approach, the complexity of running both the heuristic decomposition and the certification step is $O(\max\{n^3 \log n, D\, d^2 \log d\})$, where d is the maximum dimension of an irreducible subrepresentation, and D is the time required to multiply elements of the group. A reference implementation interfacing with RepLAB is provided.

I. INTRODUCTION

Semi-definite programming is a widely used numerical tool in science and engineering. Unfortunately, runtime and memory use of SDP solvers scale poorly with the dimension of the problem. To alleviate this issue, symmetries can often be exploited to significantly reduce the dimension [1–8] (see [9] for a review). This requires finding a common block-diagonalization of the matrices representing the symmetry group action. A large number of numerical methods for this task have been developed [10–23]. These algorithms must be compared along a number of different dimensions:

1. What is their runtime as a function of the relevant parameters? The most important parameters are the dimension n of the input matrices, the dimension of the algebra $\mathcal{A}$ they span, and the dimension d of the largest irreducible component?
2. Are they probabilistic or deterministic?
3. Do they assume a group structure, or do they work for algebras more generally?
4. Can they handle a situation where only noisy versions of the matrices representating the symmetry are available?
5. Which aspects are covered by rigorous performance guarantees?

While a detailed review of the extensive literature is beyond the scope of this paper, we summarize the performance of the approaches that come closest to the methods described here.

References [20–23] give algorithms for finding a block decomposition for general $*$-algebras and come with rigorous guarantees. Refs. [21, 22] require one to solve a polynomial optimization problem of degree 4 on $\mathbb{C}^{n\times n}$. While this might work in practice, there is no general polynomial-time algorithm for this class of problems. The procedure of [20] requires one to diagonalize "super-operators", i.e. linear maps acting on $n \times n$-matrices. This implies a runtime of $O(n^6)$.

The method of [23] exhibits a runtime of $O(\max\{n^2 \dim^2 \mathcal{A}, n^3 \dim \mathcal{A}\})$. In this scaling, the first term comes from finding an orthogonal basis for $\mathcal{A}$ and the second term arises from using this basis to project onto the commutant and to diagonalize.[1] While the method comes with a guarantee that the output decomposition is close to invariant, it does not guarantee that the components will be irreducible in the presence of noise. The runtime is particularily competitive for "small" algebras: If $\alpha \in [0, 2]$ is such that $\dim \mathcal{A} = O(n^\alpha)$, the scaling becomes $O(n^{3+\alpha})$ for the case $\alpha < 1$. On the other hand, in the regime $\alpha > 1$, the runtime $O(n^{2+2\alpha})$ is worse than other methods discussed below.

[1] This scaling refers to Alg. B from that reference. There, the scaling of the second term is presented as $O(n^4 \dim \mathcal{A})$. Upon a closer inspection of their algorithm we found that its runtime is slightly better than claimed. It seems that the origin of the difference, in their language, is that Alg. B – as opposed to Alg. A – does not require to use the subroutine *Split*. Instead, Alg. B projects a single random matrix onto the commutant of $\mathcal{A}$, using $O(n^3 \dim \mathcal{A})$ operations.

Reference [24] works on finite group representations, rather than general $*$-algebras. It generalizes Dixon's method [25] to handle noise in the group representation. This algorithm produces a full decomposition, however, for this it must project a full matrix basis onto the commutant of the representation and diagonalize each projection. This means that its runtime scales quite steeply, as $O(n^5)$.

Here, we suggest to split the problem of decomposing a unitary group representation ρ on $\mathbb{C}^n$ into three steps:

1. Use a fast heuristic to obtain a candidate decomposition $\mathbb{C}^n \simeq R_1 \oplus R_2 \oplus \ldots$. One particular randomized algorithm running in time $O(n^3)$ has been analyzed [19, 26] and implemented as part of the RepLAB [27] software package by some of the present authors. While this algorithm seems to give accurate results in practice, this is not underpinned by a formal guarantee.

2. Certify that each of the candidate spaces R_i is within a pre-determined distance ϵ of a subspace K_i that is invariant under the group.

3. Certify that the invariant spaces K_i are irreducible.

With the first step already covered in Ref. [19, 26], the present paper focuses on the two certification steps. Thus, we are faced with the situation that a heuristically obtained $n \times n$ matrix π is provided, which may or may not be close to a projection onto an invariant and irreducible space. We provide a probabilistc algorithm for this decision problem. More precisely, our main result is this:

Result 1. *Let G be a compact group. Assume that:*

1. *There exists a representation $g \mapsto \rho(g)$ in terms of unitary $n \times n$ matrices.*

2. *In time $O(n^2)$, one can draw an element $g \in G$ according to the Haar measure, and compute an approximation $\tilde{\rho}$ such that $\max_g \max_{ij} |\rho_{ij}(g) - \tilde{\rho}_{ij}(g)| = o\left(\frac{1}{n^3 \log n}\right)$.*

Then there exists an algorithm that takes as input an $n \times n$ matrix π as well as numbers $\epsilon, p_{\text{thr.}}$, and returning true *or* false *such that:*

1. *[False positive rate] The probability that the algorithm returns* true *even though π is not ϵ-close in Frobenius norm to a projection onto an invariant and irreducible ρ-space is* upper-bounded *by $p_{\text{thr.}}$.*

2. *[False negative rate] The probability that the algorithm returns* false *even though π is $(\epsilon/2)$-close in Frobenius norm to a projection onto an invariant and irreducible ρ-space is* approximately *$2p_{\text{thr.}}$.*

3. *[Runtime] As long as $\epsilon = o\left(\frac{1}{n^2 \log n}\right)$, the algorithm terminates in time*

$$O\left(\left(n^3 \log n + D \operatorname{tr}(\pi)^2 \log \operatorname{tr} \pi\right) \log \frac{1}{p_{\text{thr.}}}\right),$$

where D is time required to multiply two elements of G.

This algorithm has been implemented in Python and is available in [28].

There is an asymmetry in the way we treat false positives rates (which are bounded rigorously) and false negative rates (which are only approximated). This reflects the different roles these two parameters play in practice. Indeed, if the certification algorithm returns false, the symmetry reduction has failed, no further processing will take place, and thus no further guarantees are needed. In contrast, if the algorithm returns true, the user must be able to quantify their confidence in the result – hence the necessity to have a rigorous upper bound on the false positive rate.

In the main text, we introduce a an additional parameter δ, which can be used to tune the false negative rate independently of the false positive rate $p_{\text{thr.}}$. The interpretation is that δ is a rigorous upper bound on the false negative rate in the limiting case where $\epsilon = 0$ and the approximation $\tilde{\rho}$ is in fact exact. We have chosen $\delta = 2p_{\text{thr.}}$ in the displayed result, which turns out to simplify the formula for the runtime.

In practice, one can find appropriate values for δ numerically: In an *exploratory phase*, one can run the algorithm for increasing values of δ, until it reliably identifies valid inputs as such. One would then certify a subspace by running the procedure *once* with the δ previously obtained.

The paper is organized as follows. In Sec. II we review the mathematical setting of the paper. In Sec. III and Sec. IV we present the algorithms to certify invariance and irreducibiltity respectively. Finally, in Sec. V we discuss the runtime of the algorithms.

II. MATHEMATICAL SETTING

Let G be a compact group, and $(\mathbb{C}^n, \rho)$ be a unitary representation of G. A subset $S \subset G$ *generates* the group if $\langle S \rangle$ is dense in G, and it is *symmetric* if $S = S^{-1}$.

We assume that the user can evaluate a function $\tilde{\rho} : G \to \mathbb{C}^{n \times n}$ satifying

$$\max_{ij} |\rho(g)_{ij} - \tilde{\rho}(g)_{ij}| \leq \epsilon_0, \qquad \forall\, g \in G.$$

If $R \subset \mathbb{C}^n$ is the subspace to be certified and π_R projects onto it, we use $\tilde{\pi}_R$ to denote an approximation to π_R:

$$\max_{ij} |(\pi_R)_{ij} - (\tilde{\pi}_R)_{ij}| \leq \epsilon_0.$$

We require that $\epsilon_0 < \frac{1}{2n}$, however in practice ϵ_0 is typically of the order of machine precision.

In the context of our algorithms, the user has obtained $\tilde{\pi}_R$ as an output of their numerical procedure to decompose ρ. Using this operator as an input, the goal is to certify two statments. The first is that there exists some invariant subspace $K \subset \mathbb{C}^n$ with associated projector π_K satisfying that

$$\|\pi_R - \pi_K\|_F \leq \epsilon, \tag{1}$$

where $\|\cdot\|_F$ is the Frobenius norm and the precision parameter $\epsilon < 1/2$ is an input. We call this procedure *certifying invariance*. The second is that the subspace K is an irreducible G representation.

For this task, we assume that one *1.* knows an upper bound r_G on the number of generators of G, and *2.* can sample from the Haar measure and evaluate $\tilde{\rho}$ on the sample. In an appendix, we show how to relax the second condition and instead assume only that the user can evaluate $\tilde{\rho}$ on a well-behaved *fixed* generator set. The algorithms are probabilistic. A bound $p_{\text{thr.}}$ on the false positive rate – i.e. the probability that an input is certified even though it is not close to the projection onto an irredudcible representation – is an explicit parameter.

Bounds r_G on the number of generators of G are known for a wide variety of groups. For example it is known that $r_G \leq 2$ when G is a finite dimensional connected compact group [29]. For a wide variety of finite simple groups, furthermore, $r_G \leq 7$ (see [30] for a review).

III. THE INVARIANCE CERTIFICATE

Here we present our algorithm for the first task, that is, certifying the approximate invariance of R. Section III A treats a closely related problem: deciding whether an operator is close to the *commutant*

$$\{Y \in \mathbb{C}^{n \times n} \,|\, [\rho(g), Y] = 0 \;\forall g \;\in G\}$$

of ρ. In that section we also work in the idealized case where $\epsilon_0 = 0$. The general algorithm deciding invariance is presented in Section III B.

A. Estimating closeness to the commutant in the ideal case

As mentioned, in this section we assume $\epsilon_0 = 0$ – i.e. that the representation ρ can be evaluated *exactly* – in order to bring out the key components of the argument.

Consider an $n \times n$ matrix X (later, we will take X to be the approximate projection $\tilde{\pi}_R$ onto a candidate subspace). The randomized Algorithm III.1 tests whether

$$\|X - P_{\text{Haar}}(X)\|_\infty \leq \epsilon.$$

There, $\|\cdot\|_\infty$ is the spectral norm and P_{Haar} is the Hilbert-Schmidt projection onto the commutant

$$P_{\text{Haar}}(X) := \mathbb{E}_g[\rho(g) X \rho^\dagger(g)],$$

where the expectation value is with respect to the Haar distribution.

Algorithm III.1 Closeness to Commutant

Input:

- $X \in \mathbb{C}^{n\times n}$,
- $p_{\text{thr.}} \in (0,1)$, $\epsilon \in (0,1/2)$.

1: Set $r = 8\lceil(\log(1/p_{\text{thr.}}) + \log(2n))\rceil$
2: Sample r group elements $g_1, \ldots, g_r \in G$ Haar-randomly
3: Compute $c = \left\| \frac{1}{r}\sum_i \rho(g_i) X \rho^\dagger(g_i) - X \right\|_\infty$
4: **if** $2c \leq \epsilon$ **then**
5: **Return:** *True*
6: **end if**
7: **Return:** *False*

Proposition 1. *Let $X \in \mathbb{C}^{n\times n}$ satisfy $\|X - P_{\text{Haar}}(X)\|_\infty > \epsilon$. Then, the probability that Alg. III.1 returns* True *is at most $p_{\text{thr.}}$.*

Proof. Consider the following matrix-valued random variable with mean equal to zero,

$$Z_g := \frac{1}{r}\Big(\rho(g) X \rho^\dagger(g) - P_{\text{Haar}}(X)\Big), \qquad g \in G \text{ Haar random.}$$

Using $R := \text{Id} - P_{\text{Haar}}$ (the projector onto the orthocomplement of the commutant of ρ), we find $Z_g = \frac{1}{r}\rho(g) R(X) \rho^\dagger(g)$, and so,

$$\|Z_g Z_g^\dagger\|_\infty = \frac{1}{r^2}\|R(X)R(X)^\dagger\|_\infty = \frac{1}{r^2}\|R(X)\|_\infty^2, \qquad \forall\, g \in G.$$

This way, by the matrix Hoeffding bound [31],

$$\text{Prob}\left[\Big\|\sum_i Z_{g_i}\Big\|_\infty \geq z\|R(X)\|_\infty\right] \leq 2n \exp\left(\frac{-rz^2}{2}\right)$$

where $\{g_i\}$ are the samples in line 2 of Alg. III.1. Taking $z = 1/2$, the right-hand side above is $\leq p_{\text{thr.}}$ and so with probability at least $1 - p_{\text{thr.}}$ it holds that

$$c = \Big\|\frac{1}{r}\sum_i \rho(g_i) X \rho^\dagger(g_i) - X\Big\|_\infty = \Big\|\sum_i Z_{g_i} - R(X)\Big\|_\infty \geq \|R(X)\| - \Big\|\sum_i Z_{g_i}\Big\|_\infty \geq \frac{1}{2}\|R(X)\|_\infty > \epsilon/2.$$

□

We now show a converse result, namely, that Alg. III.1 always "detects" matrices which are close enough to the commutant.

Proposition 2. *Let X satisfy $\|X - P_{\text{Haar}}(X)\|_\infty \leq \epsilon/2$ for some $\epsilon < 1$. Then Alg. III.1 deterministically returns* True *upon the input X, ϵ.*

Proof. For any $g \in G$ it holds that

$$\|[\rho(g), X]\|_\infty = \|[\rho(g), X - P_{\text{Haar}}(X)]\|_\infty \leq 2\|X - P_{\text{Haar}}(X)\|_\infty \leq \epsilon.$$

Therefore, using standard norm relations we obtain

$$c = \Big\|\frac{1}{r}\sum_i \big(\rho(g_i) X \rho^\dagger(g_i) - X\big)\Big\|_\infty \leq \frac{1}{r}\sum_i \Big\|[\rho(g_i), X]\Big\|_\infty \leq \epsilon.$$

□

B. The full certificate

Here, we will go beyond Section III A in two ways: First, we allow for non-zero errors ϵ_0. Second, we show that a projection that is close to being invariant is close to a projection onto an invariant subspace. The goal is, given $\tilde{\pi}_R$ as an input, to certify that there is an invariant subspace K with

$$\|\pi_K - \pi_R\|_F \leq \epsilon.$$

The procedure is given in Alg. III.2.

Algorithm III.2 Invariance certificate

Input:

- $\tilde{\pi}_R \in \mathbb{C}^{n\times n}$,
- $p_{\text{thr.}} \in (0,1)$,
- $\epsilon \in (0, 1/2)$.

Output: *True/False*

1: Set $r = 8\lceil(\log(1/p_{\text{thr.}}) + \log(2n))\rceil$, $f_{\text{err}} = 8n\epsilon_0 + 6n^2\epsilon_0^2 + 2n^3\epsilon_0^3$, and $\epsilon' = \epsilon/2\sqrt{2\dim R}$
2: Sample r group elements $g_1, \ldots, g_r \in G$ Haar-randomly
3: Compute $\tilde{c} = \left\| \frac{1}{r}\sum_i \tilde{\rho}(g_i)\tilde{\pi}_R\tilde{\rho}^\dagger(g_i) - \tilde{\pi}_R \right\|_\infty$
4: **if** $2\tilde{c} + f_{\text{err}} \leq \epsilon'$ **then**
5: **Return:** *True*
6: **end if**
7: **Return:** *False*

As before, line 4 of Alg. III.2 simply takes k close to the minimum of $f_k(c)$ and does not affect the probability of falsely certifying R. Our main result in this section is the following guarantee on the invariance certificate.

Theorem 1. *Assume that for all invariant subspaces $K \subset \mathbb{C}^n$,*

$$\|\pi_K - \pi_R\|_F > \epsilon. \tag{2}$$

Then, the probability that Alg. III.2 returns True *is upper bounded by $p_{\text{thr.}}$.*

To prove Thm. 1 we will first show that if π_R is close to the commutant, then it is close to an invariant projector π_K as in eq. (1). After that, our argument will closely follow Sec. III A.

Proposition 3. *Assume that π_R satisfies $2\sqrt{2\dim R}\,\|P_{\text{Haar}}(\pi_R) - \pi_R\|_\infty \leq \epsilon$ for some $\epsilon < 1$. Then there exists an invariant subspace K with projector π_K satisfying $\|\pi_R - \pi_K\|_F \leq \epsilon$.*

Proof. Let $\lambda^\downarrow(M)$ be the vector of eigenvalues of a Hermitian matrix $M \in \mathbb{C}^{n\times n}$ in decreasing order. By Weyl's perturbation theorem (see e.g. [32, Chap. VI]),

$$\|\lambda^\downarrow(P_{\text{Haar}}(\pi_R)) - \lambda^\downarrow(\pi_R)\|_{\ell_\infty} \leq \frac{\epsilon}{2\sqrt{2\dim R}} = \epsilon'.$$

This way, the eigenvalues of $P_{\text{Haar}}(\pi_R)$ lie in $[-\epsilon', \epsilon'] \cup [1-\epsilon', 1+\epsilon']$, where $\epsilon' < 1/2$. Let π_K be the projector onto all eigenspaces corresponding to eigenvalues in $1 \pm \epsilon'$. The projector π_K is invariant and satisfies $\|\pi_K - P_{\text{Haar}}(\pi_R)\|_\infty \leq \epsilon'$. We therefore see that,

$$\begin{aligned}\|\pi_K - \pi_R\|_F &\leq \sqrt{2\dim R}\,\|\pi_K - \pi_R\|_\infty \\ &\leq \sqrt{2\dim R}\,\left(\|\pi_K - P_{\text{Haar}}(\pi_R)\|_\infty + \|P_{\text{Haar}}(\pi_R) - \pi_R\|_\infty\right) \\ &\leq 2\epsilon'\sqrt{2\dim R} = \epsilon,\end{aligned}$$

where we used that $\operatorname{rank}(\pi_K - \pi_R) \leq \dim K + \dim R = 2\dim R$ in the first step. □

From the proof above it becomes clear that certifying that R is approximately invariant is, ultimately, just certifying that π_R is close enough to the commutant.

Proof of Thm. 1. By Prop. 3 we may take

$$\frac{\epsilon}{2\sqrt{2\dim R}} < \|P_{\mathrm{Haar}}(\pi_R) - \pi_R\|_\infty.$$

Let

$$A := \frac{1}{r}\sum_i \Big(\rho(g_i)\pi_R\rho^\dagger(g_i) - \tilde{\rho}(g_i)\tilde{\pi}_R\tilde{\rho}^\dagger(g_i)\Big), \qquad \Delta_R := \pi_R - \tilde{\pi}_R,$$

then,

$$\Big\|\frac{1}{r}\sum_i \rho(g_i)\pi_R\rho^\dagger(g_i) - \pi_R\Big\|_\infty \leq \|\Delta_R\|_\infty + \|A\|_\infty + \Big\|\frac{1}{r}\sum_i \tilde{\rho}(g_i)\tilde{\pi}_R\tilde{\rho}^\dagger(g_i) - \tilde{\pi}_R\Big\|_\infty = n\epsilon_0 + \|A\|_\infty + \tilde{c}.$$

Then, by Prop. 1, with probability at least $1 - p_{\mathrm{thr.}}$ it holds that

$$\frac{\epsilon}{2\sqrt{2\dim R}} < 2(n\epsilon_0 + \|A\|_\infty + \tilde{c}).$$

We now provide an upper bound on $\|A\|_\infty$. Let $\Delta(g) := \rho(g) - \tilde{\rho}(g)$, then

$$\begin{aligned}\|A\|_\infty \leq \mathbb{E}_i\Big[&\|\Delta(g_i)\pi_R\rho^\dagger(g_i)\|_\infty + \|\rho(g_i)\Delta_R\rho^\dagger(g_i)\|_\infty + \|\rho(g_i)\pi_R\Delta^\dagger(g_i)\|_\infty \\ &+\|\Delta(g_i)\Delta_R\rho^\dagger(g_i)\|_\infty + \|\Delta(g_i)\pi_R\Delta^\dagger(g_i)\|_\infty + \|\rho(g_i)\Delta_R\Delta^\dagger(g_i)\|_\infty \\ &+\|\Delta(g_i)\Delta_R\Delta^\dagger(g_i)\|_\infty\Big].\end{aligned}$$

Submultipliciativity, together with $\max\{\|\Delta_R\|_\infty, \|\Delta(g)\|_\infty\} \leq n\epsilon_0$ for all $g \in G$, gives

$$\|A\|_\infty \leq 3(n\epsilon_0 + n^2\epsilon_0^2) + n^3\epsilon_0^3.$$

□

IV. IRREDUCIBILITY CERTIFICATE

In this section we present an algorithm that certifies irreducibility. Given $\tilde{\pi}_R$ as an input, where R holds an invariance certificate, the goal is to certify that the minimizer of

$$\min_{\substack{K\subset\mathbb{C}^n \\ K \text{ invar.}}} \|\pi_R - \pi_K\|_F \tag{3}$$

is irreducible. We first present the idea of the algorithm in an idealized setting, and then come back to the noisy scenario.

A. The ideal case

Let $(\mathbb{C}^{n_K}, \rho_K)$ be a unitary representation of G and suppose that we have access to the same primitives as in Sec III A. Namely, we can sample Haar-randomly from G and evaluate ρ_K on any sample. Our task is to certify if ρ_K is irreducible. The following algorithm uses random walks to acheive this.

Algorithm IV.1 Ideal irreducibility certificate

Input:

- $p_{\text{thr.}} \in (0,1)$, ▷ Bound on false positive rate.
- $p'_{\text{thr.}} \in (p_{\text{thr.}}, 1)$, ▷ Bound on false negative rate.

Output: *True/False.*

1: Set $r = \max\{r_G,\, 8\lceil(\log(2/p_{\text{thr.}}) + 2\log(n_K))\rceil\}$ ▷ G generated by $\leq r_G$ elements
2: Set $m = 2n_K^2 \cdot \max\{8\lceil\log\big((p'_{\text{thr.}} - p_{\text{thr.}})^{-1}\big)\rceil, \lceil\log\big(p_{\text{thr.}}^{-1}\big)\rceil\}$ ▷ m number of random walks
3: Set $t = 2 + \lceil\log_2 n_K\rceil$ ▷ $2t$ length of random walks
4: Sample r elements $g_i \in G$ Haar-randomly and set $S = \{g_i\} \cup \{g_i^{-1}\}$
5: Sample m elements $\mathbf{s}_i \in S^{2t}$ uniformly
6: Compute $E_m = \frac{1}{m}\sum_i |\operatorname{tr}\rho_K(\mathbf{s}_i)|^2$
7: Set $\theta_m = n_K\sqrt{2/m}\log(1/p_{\text{thr.}})$
8: **if** $E_m < 2(1-\theta_m)$ **then**
9: **return** *True*
10: **end if**
11: **return** *False*

Theorem 2. *Let ρ_K be* reducible*, then the probability that Alg. IV.1 returns* True *upon this input is at most $p_{\text{thr.}}$.*

Our proof of Thm. 2 will work for any value of t, i.e. it does not rely on using $t = 2 + \lceil\log_2 n_k\rceil$. However, if t is chosen too small, the algorithm could fail to recognize irreducible representations —its false negative rate would be large. We will bound this rate at the end of this subsection.

The key for the proof of Thm. 2 is Schur's lemma —if ρ_K were irreducible it would hold that $\operatorname{tr} P_{\text{Haar}} = 1$ and otherwise it holds that $\operatorname{tr} P_{\text{Haar}} \geq 2$. What the algorithm does is estimate a quantity which is larger than the dimension of the commutant of ρ_K. As we will see, if ρ_K is reducible then it is exceedingly unlikely for this estimator to fall too much below 2.

The quantity being estimated is, in fact, $\operatorname{tr} P_S^{2t}$, where P_S is the random walk operator associated to ρ_K. The connection to the dimension of the commutant is made by the following statement.

Proposition 4. *For any t it holds that* $\operatorname{tr} P_{\text{Haar}} \leq \operatorname{tr} P_S^{2t}$.

Proof. Unitarity ensures that $\|P_S\|_\infty = 1$. Because $r \geq r_G$, the probability that S generates G is one. Together with $S = S^{-1}$, this ensures that P_S is self-adjoint and that the $+1$ eigenspace corresponds exactly to the commutant of ρ_K.

Let $\{\lambda_i\}$ be all the eigenvalues of P_S different from one. The statement follows from

$$\operatorname{tr} P_S^{2t} = \operatorname{tr} P_{\text{Haar}} + \sum_i \lambda_i^{2t} \geq \operatorname{tr} P_{\text{Haar}}.$$

□

Proof of Thm. 2. It is clear that E_m is an estimator for $\operatorname{tr} P_S^{2t}$. Since ρ_K is unitary, furthermore, $|\operatorname{tr}\rho_K(g)|^2 \leq n_K^2$ for any g, and so by Chernoff's bound,

$$\Pr\left[E_m \leq (1-\theta)\operatorname{tr} P_S^{2t}\right] \leq \exp\left(\frac{-\theta^2 m \operatorname{tr} P_S^{2t}}{2n_K^2}\right),$$

for any $\theta \in (0,1)$. But by the assumption on m we may use $\theta = \theta_m$ in the equation above. Then, using Prop. 4 and $\operatorname{tr} P_{\text{Haar}} \geq 2$,

$$\Pr\left[E_m \leq 2(1-\theta_m)\right] \leq \Pr\left[E_m \leq (1-\theta_m)\operatorname{tr} P_S^{2t}\right] \leq \exp\left(\frac{-\theta_m^2 m \operatorname{tr} P_S^{2t}}{2n_K^2}\right) \leq \exp\left(\frac{-\theta_m^2 m}{n_K^2}\right) < p_{\text{thr.}}.$$

□

As mentioned, the proof above doesn't rely on the particular choice of t in line 3 of Alg. IV.1. It also only uses the bound $m > 2n_K^2\log(1/p_{\text{thr.}})$ on the number of samples (cf. line 2). In Prop. 6, we use $t > 2 + \log_2 n$ and $m > 16n_K^2\log_2(1/(p'_{\text{thr.}} - p_{\text{thr.}}))$ to bound the false negative rate of the algorithm. To prove it, it's convenient to show the following intermediate result first.

Proposition 5. *Let S be sampled as in Alg. IV.1. The probability that $\|P_{\text{Haar}} - P_S\|_\infty > 1/2$ is strictly less than*

$$2n^2 \exp\left(\frac{-r}{8}\right) \leq p_{\text{thr.}}.$$

Proof. Let σ be the representation of G acting by conjugation on $\mathbb{C}^{n\times n}$. For a group element $g \in G$ sampled Haar-randomly, the operator

$$V_g := \frac{1}{r}\Big(\frac{1}{2}\big(\sigma(g) + \sigma^\dagger(g)\big) - P_{\text{Haar}}\Big)$$

is a Hermitian random variable with zero mean. Furthermore, by unitarity of ρ and because $\sigma(g)$ and P_{Haar} are simultaneously diagonalizable, we have that

$$\|V_g\|_\infty \leq \frac{1}{r}, \qquad \|V_g^2\|_\infty \leq \frac{1}{r^2}.$$

But then, writing $S = \{g_i\}_{i=1}^r \cup \{g_i^{-1}\}_{i=1}^r$, we see that

$$P_S - P_{\text{Haar}} = \sum_{i=1}^r V_{g_i},$$

where the operators V_{g_i} are independent random variables satisfying the conditions above. Then, by the matrix Hoeffding bound [31],

$$\text{Prob}\left(\lambda_{\max}(P_S - P_{\text{Haar}}) > x\right) < n^2 e^{\frac{-rx^2}{2}},$$

where $\lambda_{\max}$ is the maximum eigenvalue. Finally, repeating the statement above for $\lambda_{\max}(P_{\text{Haar}} - P_S)$ and using the union bound, we conclude that

$$\text{Prob}\left(\|P_{\text{Haar}} - P_S\|_\infty > x\right) < 2n^2 e^{\frac{-rx^2}{2}}.$$

Using $x = 1/2$ and the fact that $r \geq 8\lceil(\log(1/p_{\text{thr.}}) + 2\log(n))\rceil$ we recover the claimed statement. □

Proposition 6. *Let ρ_K be* irreducible*, then the probability that Alg. IV.1 returns* False *upon this input is at most $p'_{\text{thr.}}$.*

Proof. By Prop. 5, with probability at least $1 - p_{\text{thr.}}$ it holds that

$$\|P_S^{2t} - P_{\text{Haar}}\|_\infty \leq 2^{-2t}, \tag{4}$$

where we used $P_S^{2t} - P_{\text{Haar}} = (P_S - P_{\text{Haar}})^{2t}$ because P_S and P_{Haar} commute. This way,

$$\text{tr}\, P_S^{2t} \leq \text{tr}\, P_{\text{Haar}} + n_K^2 2^{-2t} \leq \text{tr}\, P_{\text{Haar}} + \frac{1}{16} = \frac{17}{16}.$$

Furthermore, by our assumption in m, we have $2(1 - \theta_m) \geq 3/2$. But then, the Chernoff bound says that the probability that $E_m \geq 3/2$ is at most

$$\exp\left(\frac{-m}{n_K^2}\frac{49}{3\times 256}\right) < \exp\left(\frac{-m}{16n_K^2}\right) \leq p'_{\text{thr.}} - p_{\text{thr.}}.$$

A false positive can occur if either eq. (4) does not hold, or if conditioned on it holding, $E_m \geq 3/2$. By the union bound, this probability is at most $p_{\text{thr.}} + (1 - p_{\text{thr.}})(p'_{\text{thr.}} - p_{\text{thr.}}) < p'_{\text{thr.}}$. □

B. The noisy case

In this section we adapt the idea presented above to the noisy scenario. Suppose we have certified that a subspace $R \subset \mathbb{C}^n$ is invariant (with precision ϵ). We now wish to certify that the minimizer K of (3) is irreducible.

The algorithm for this is Alg. IV.2. As before, the algorithm has a controllable false positive rate $p_{\text{thr.}}$ as an input. This is important from the point of view of certification —if the output is *True*, then one can be rather certain that K is irreducible.

Additionally, the algorithm takes as an input a confidence parameter $p_{\text{thr.}} < \delta_{\text{conf.}} < 1$ which roughly tunes the false negative rate. In fact, this parameter is used in the same way that $p'_{\text{thr.}}$ was used in Alg. IV.1. Because Alg. IV.2 reduces to Alg. IV.1 in the limit of ϵ, $\epsilon_0 \to 0$, we expect that the false negative rate is well approximated by $\delta_{\text{conf.}}$ when ϵ and ϵ_0 are small enough. Since the runtime of the algorithm scales with $\max \log(1/p_{\text{thr.}}), \log(1/(\delta_{\text{conf.}} - p_{\text{thr.}}))$, a reasonable choice for the confidence parameter is $\delta_{\text{conf.}} = 2p_{\text{thr.}}$.

Within Alg. IV.2 and throughout this section we use the following conventions:

$$\begin{aligned} c_1 &:= 2(\epsilon + n\epsilon_0)(1 + \epsilon + n\epsilon_0) + n\epsilon_0(1 + \epsilon + n\epsilon_0)^2, \\ c_2 &:= 2c_1(1 + c_1), \\ h_t(x) &:= (1 + x)^t - 1, \\ d_t &:= h_t(c_2), \\ e_t &:= d_{2t}(\text{int}(\text{tr}\,\tilde{\pi}_R)^2 + d_{2t}). \end{aligned}$$

For the sake of clarity, we have shifted the proofs of several propositions in this subsection to App. A.

Algorithm IV.2 Irreducibility certificate

Input:

- $\tilde{\pi}_R \in \mathbb{C}^{n\times n}$, $\epsilon \in (0, 1/2)$ ▷ π_R, ϵ satisfy (1)
- $p_{\text{thr.}} \in (0, 1)$ ▷ Bound on false positive rate
- $\delta_{\text{conf.}}$ ▷ Confidence parameter

Output: *True/False.*

if $e_t \geq 2$ **then**
  return *False*
end if
Set $r = \max\{r_G,\, 12\lceil(\log(2/p_{\text{thr.}}) + 2\log(n))\rceil\}$ ▷ G generated by $\leq r_G$ elements.
Set $m = 2\left\lceil \frac{\text{int}(\text{tr}\,\tilde{\pi}_R)^2 + d_{2t}}{2 - e_t} \cdot \max\{\log\big(p_{\text{thr.}}^{-1}\big), 8\log\big((\delta_{\text{conf.}} - p_{\text{thr.}})^{-1}\big)\}\right\rceil$ ▷ m random walks
Set $t = 2 + \lceil \log_2 \text{int}(\text{tr}\,\tilde{\pi}_R)\rceil$ ▷ $2t$ random walk length
Sample r elements $g_i \in G$, set $S = \{g_i\} \cup \{g_i^{-1}\}$
Sample m words $\mathbf{s}_i \in S^{2t}$ uniformly
Compute $E = e_t + \frac{1}{m}\sum_i |\,\text{tr}\,\tilde{\rho}_R(\mathbf{s}_i)|^2$
Set $\theta_m = \sqrt{2\log(1/p_{\text{thr.}})(\text{int}(\text{tr}\,\tilde{\pi}_R)^2 + d_{2t})/m(2 - e_t)}$
if $E < 2(1 - \theta_m)$ **then**
  return *True*
end if
return *False*

Theorem 3. *Assume that the minimizer K of eq.* (3) *is* reducible*. Then the probability that Alg. IV.2 outputs* True *is at most $p_{\text{thr.}}$.*

Similar to the ideal case, the proof of this theorem relies on characterizing the *approximate* random walk operator Q_S^R given by

$$Q_S^R(\,\cdot\,) := \frac{1}{|S|}\sum_{s\in S} \tilde{\pi}_R\tilde{\rho}(s)\tilde{\pi}_R^\dagger(\,\cdot\,)\tilde{\pi}_R\tilde{\rho}^\dagger(s)\tilde{\pi}_R^\dagger.$$

Our approach uses Q_S^R to upper-bound the dimension of the commutant of ρ restricted to K, that is $\text{tr}\,P_{\text{Haar}}^K$, where

$$P_{\text{Haar}}^K(\,\cdot\,) := \int_G \text{d}\mu_{\text{Haar}}(g)\pi_K\rho(g)\pi_K(\,\cdot\,)\pi_K\rho^\dagger(g)\pi_K.$$

An important object in our proof is the *restricted random walk operator*,

$$P_S^K(\,\cdot\,) := \frac{1}{|S|}\sum_{s\in S}\pi_K\rho(s)\pi_K(\,\cdot\,)\pi_K\rho^\dagger(s)\pi_K.$$

Notice that Q_S^R is a small perturbation of P_S^K.

Proposition 7. *Use the notation above, let $Q_e := P_S^K - Q_S^R$ and γ be such that $\|Q_e\|_\infty \le \gamma$. Then, for all t it holds that*

$$\operatorname{tr} P_{\text{Haar}}^K \le \operatorname{tr}\big((Q_S^R + \gamma\mathbb{I})^{2t}\big).$$

Proof. Let $\{r_i\}$ be the eigenvalues of P_S^K. By Weyl's perturbation theorem, for each r_i, there is some eigenvalue q_i of Q_S^R satisfying $q_i \in r_i \pm \gamma$. In particular, $Q_S^R + \gamma\mathbb{I}$ has $\operatorname{tr} P_{\text{Haar}}^K$-many eigenvalues in the range $[1, 1+2\gamma]$. Then,

$$\operatorname{tr}\big((Q_S^R + \gamma\mathbb{I})^{2t}\big) \ge \operatorname{tr} P_{\text{Haar}}^K + \sum_{\substack{i \text{ s.t.}\\ r_i<1}} (q_i+\gamma)^{2t} \ge \operatorname{tr} P_{\text{Haar}}^K.$$

□

We will show that $\|Q_e\|_\infty \le c_2$ in Prop. 11 from App. A, and so we use $\gamma = c_2$ henceforth. Then, if for any t it holds that

$$\operatorname{tr}\big((Q_S^R + c_2\mathbb{I})^{2t}\big) < 2,$$

K is irreducible. We may expand

$$\operatorname{tr}\big((Q_S^R + c_2\mathbb{I})^{2t}\big) = \sum_{k=0}^{2t}\binom{2t}{k} c_2^{2t-k}\operatorname{tr}\big((Q_S^R)^k\big) \tag{5}$$

$$= \sum_{k=0}^{2t}\binom{2t}{k} c_2^{2t-k}\,\frac{1}{|S|^k}\sum_{\mathbf{s}\in S^k}|\operatorname{tr}\tilde\rho_R(\mathbf{s})|^2, \tag{6}$$

where we used,

$$\tilde\rho_R(s) := \tilde\pi_R\tilde\rho(s)\tilde\pi_R^\dagger, \qquad \tilde\rho_R(\mathbf{s}) := \tilde\rho_R(s_1)\dots\tilde\rho_R(s_k), \qquad s\in S,\ \mathbf{s}\in S^k.$$

Our approach is to bound the norm of all terms with $k < 2t$ and estimate the one with $k = 2t$. This is because in the regime of interest c_2 is small, and so terms with non-trivial powers of c_2 are of subleading order. The following proposition will be used to bound the size of subleading terms.

Proposition 8. *Let R hold an invariance certificate with precision $\epsilon < 1/2$ and let K be the minimizer in eq.* (3)*. Then, for any $\mathbf{s}\in S^k$, it holds that*

$$|\operatorname{tr}\tilde\rho_R(\mathbf{s})|^2 \le \dim^2 K + d_k.$$

The following proposition uses the previous result to bound the size of the subleading contributions to eq. (6).

Proposition 9. *Let R, K and ϵ be as in Prop. 8, and let $n\epsilon_0 < 1/2$. Then,*

$$\left|\sum_{k=0}^{2t-1}\binom{2t}{k} c_2^{2t-k}\operatorname{tr}\big((Q_S^R)^k\big)\right| \le e_t.$$

We therefore obtain

$$\operatorname{tr} P_{\text{Haar}}^K \le e_t + \operatorname{tr}\big((Q_S^R)^{2t}\big) = e_t + \frac{1}{|S|^{2t}}\sum_{\mathbf{s}\in S^{2t}}|\operatorname{tr}\tilde\rho_R(\mathbf{s})|^2.$$

All that is left to be shown is that the estimator for the second term used by Alg. IV.2 concentrates sharply around its mean. For this we will use the following proposition, a simple consequence of the Chernoff bound.

Proposition 10. *Let R, K and ϵ be as in Prop. 8, and assume that K is* reducible. *Let $\{\mathbf{s}_i\}$ be m uniformly random samples from S^{2t}. Then, for any $\theta \in (0,1)$, it holds that*

$$\Pr\left[\frac{1}{m}\sum_{i=1}^{m} |\operatorname{tr}\tilde{\rho}_R(\mathbf{s}_i)|^2 \leq (1-\theta)\operatorname{tr}\big((Q_S^R)^{2t}\big)\right] < \exp\left(\frac{-\theta^2 m(2-e_t)}{2(\dim^2 K + d_{2t})}\right).$$

We may now prove the first main result of this subsection.

Proof of Thm. 3. By our assumption on m, it holds that $\theta_m < 1$. But then using Prop. 10 with $\theta = \theta_m$,

$$\begin{aligned}\Pr\left[\frac{1}{m}\sum_i |\operatorname{tr}\tilde{\rho}_R(\mathbf{s}_i)|^2 + e_t \leq 2(1-\theta_m)\right] &\leq \Pr\left[\frac{1}{m}\sum_i |\operatorname{tr}\tilde{\rho}_R(\mathbf{s}_i)|^2 + e_t \leq (1-\theta_m)\left[\operatorname{tr}\big((Q_S^R)^{2t}\big) + e_t\right]\right] \\ &\leq \Pr\left[\frac{1}{m}\sum_i |\operatorname{tr}\tilde{\rho}_R(\mathbf{s}_i)|^2 \leq (1-\theta_m)\operatorname{tr}\big((Q_S^R)^{2t}\big)\right] \\ &< \exp\left(\frac{-\theta_m^2 m(2-e_t)}{2(\dim^2 K + d_{2t})}\right) < p_{\text{thr.}}.\end{aligned}$$

□

V. TIME COMPLEXITY

Here we analyse the runtime of the certification procedures proposed and discuss several ways to optimize it.

Alg. III.2 runs in $O(n^3 \log n)$ steps: the main sources of complexity are the $r = O(\log n)$ matrix products and the spectral norm appearing in line 3. The latter has complexity at most $O(n^3)$ through the singular value decomposition.

In practice, this last step step is significantly cheaper. Ref. [33] estimates the spectral norm in time $O(n^2 \log n)$. Note that the method of [33] is probabilistic and so it raises the false positive rate, albeit in a controllable way. Alternatively, the spectral norm can be bounded by the Frobenius norm in $O(n^2)$ operations.

To compute the runtime of Alg. IV.2 we assume that ϵ_0 and ϵ are small enough that $d_{2(2+\log_2 d)}$ and $e_{2+\log_2 d}$ are non-increasing functions of $d := \dim R$ and n. Here, d_t and e_t are defined as in the top of Sec. IV B and we use $t = 2 + \log d$. For this it is sufficient to take

$$\epsilon < \frac{1}{48(d^2+1)(2+\log_2 d)}, \qquad \epsilon_0 < \frac{1}{120n(d^2+1)(2+\log_2 d)}. \tag{7}$$

In this regime the runtime of the algorithm, as it is written in the main text, is

$$O\left(n^3 d^2 \log d \left(\log\frac{1}{p_{\text{thr.}}} + \log\frac{1}{\delta_{\text{conf.}} - p_{\text{thr.}}}\right)\right). \tag{8}$$

Because the false negative rate is of secondary importance for our certificate, a convenient choice is $\delta_{\text{conf.}} = 2p_{\text{thr.}}$ where both terms above have the same scaling.

The main bottleneck of (8) is the n^3 factor, coming from the fact that the algorithm evaluates matrix products on $\mathbb{C}^{n\times n}$. This can be significantly reduced by either: *1.* taking products in the group and then obtaining the image, or *2.* restricting matrices $\tilde{\rho}_R(s)$ to the subspace R first, and taking products in this smaller space. Letting D denote the runtime of whichever of these two is faster, the runtime becomes $O(Dd^2 \log d \log p_{\text{thr.}}^{-1})$.

ACKNOWLEDGMENTS

We thank Markus Heinrich and Frank Vallentin for insightful conversations.

This work has been supported by the DFG (SPP1798 CoSIP), Germany's Excellence Strategy – Cluster of Excellence Matter and Light for Quantum Computing (ML4Q) EXC2004/1, Cologne's Key Profile Area Quantum Matter and Materials, the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie agreement No 764759, and by the Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Economic Development, Job Creation

and Trade. This publication was made possible through the support of a grant from the John Templeton Foundation. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

[1] F. Vallentin, "Symmetry in semidefinite programs," *Linear Algebra and its Applications*, vol. 430, no. 1, pp. 360–369, 2009.
[2] F. Permenter and P. A. Parrilo, "Dimension reduction for semidefinite programs via Jordan algebras," *Mathematical Programming*, vol. 181, no. 1, pp. 51–84, 2020.
[3] M. Heinrich and D. Gross, "Robustness of magic and symmetries of the stabiliser polytope," *Quantum*, vol. 3, p. 132, 2019.
[4] A. Raymond, J. Saunderson, M. Singh, and R. R. Thomas, "Symmetric sums of squares over k-subset hypercubes," *Mathematical Programming*, vol. 167, no. 2, pp. 315–354, 2018.
[5] C. Riener, T. Theobald, L. J. Andrén, and J. B. Lasserre, "Exploiting symmetries in SDP-relaxations for polynomial optimization," *Mathematics of Operations Research*, vol. 38, no. 1, pp. 122–141, 2013.
[6] C. Śliwa, "Symmetries of the bell correlation inequalities," *Physics Letters A*, vol. 317, no. 3-4, pp. 165–168, 2003.
[7] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, "Bell inequalities for arbitrarily high-dimensional systems," *Physical review letters*, vol. 88, no. 4, p. 040404, 2002.
[8] M. O. Renou, D. Rosset, A. Martin, and N. Gisin, "On the inequivalence of the ch and chsh inequalities due to finite statistics," *Journal of Physics A: Mathematical and Theoretical*, vol. 50, no. 25, p. 255301, 2017.
[9] C. Bachoc, D. C. Gijswijt, A. Schrijver, and F. Vallentin, "Invariant semidefinite programs," in *Handbook on semidefinite, conic and polynomial optimization*, pp. 219–269, Springer, 2012.
[10] W. Eberly, *Computations for algebras and group representations*. PhD thesis, University of Toronto., 1989.
[11] W. Eberly, "Decompositions of algebras over R and C," *Computational Complexity*, vol. 1, no. 3, pp. 211–234, 1991.
[12] K. Murota, Y. Kanno, M. Kojima, and S. Kojima, "A numerical algorithm for block-diagonal decomposition of matrix $*$-algebras with application to semidefinite programming," *Japan Journal of Industrial and Applied Mathematics*, vol. 27, no. 1, pp. 125–160, 2010.
[13] T. Maehara and K. Murota, "A numerical algorithm for block-diagonal decomposition of matrix $*$-algebras with general irreducible components," *Japan journal of industrial and applied mathematics*, vol. 27, no. 2, pp. 263–293, 2010.
[14] E. de Klerk, C. Dobre, and D. V. Pasechnik, "Numerical block diagonalization of matrix $*$-algebras with application to semidefinite programming," *Mathematical programming*, vol. 129, no. 1, p. 91, 2011.
[15] K. Abed-Meraim and A. Belouchrani, "Algorithms for joint block diagonalization," in *2004 12th European Signal Processing Conference*, pp. 209–212, IEEE, 2004.
[16] Y. Cai and C. Liu, "An algebraic approach to nonorthogonal general joint block diagonalization," *SIAM Journal on Matrix Analysis and Applications*, vol. 38, no. 1, pp. 50–71, 2017.
[17] Y. Cai, D. Shi, and S. Xu, "A matrix polynomial spectral approach for general joint block diagonalization," *SIAM Journal on Matrix Analysis and Applications*, vol. 36, no. 2, pp. 839–863, 2015.
[18] Y. Cai, G. Cheng, and D. Shi, "Solving the general joint block diagonalization problem via linearly independent eigenvectors of a matrix polynomial," *Numerical Linear Algebra with Applications*, vol. 26, no. 4, p. e2238, 2019.
[19] D. Rosset, F. Montealegre-Mora, and J.-D. Bancal, "RepLAB: a computational/numerical approach to representation theory," *arXiv preprint arXiv:1911.09154*, 2019.
[20] T. Maehara and K. Murota, "Algorithm for error-controlled simultaneous block-diagonalization of matrices," *SIAM Journal on Matrix Analysis and Applications*, vol. 32, no. 2, pp. 605–620, 2011.
[21] Y. Cai and P. Li, "Identification of matrix joint block diagonalization," *arXiv preprint arXiv:2011.01111*, 2020.
[22] Y. Cai and R.-C. Li, "Perturbation analysis for matrix joint block diagonalization," *Linear Algebra and its Applications*, vol. 581, pp. 163–197, 2019.
[23] L. Babai, K. Friedl, and M. Stricker, "Decomposition of $*$-closed algebras in polynomial time," in *Proceedings of the 1993 international symposium on Symbolic and algebraic computation*, pp. 86–94, 1993.
[24] L. Babai and K. Friedl, "Approximate representation theory of finite groups," in *[1991] Proceedings 32nd Annual Symposium of Foundations of Computer Science*, pp. 733–742, IEEE, 1991.
[25] J. D. Dixon, "Computing irreducible representations of groups," *Mathematics of Computation*, vol. 24, no. 111, pp. 707–712, 1970.
[26] A. Tavakoli, E. Z. Cruzeiro, R. Uola, and A. A. Abbott, "Bounding and simulating contextual correlations in quantum theory," *arXiv preprint arXiv:2010.04751*, 2020.
[27] `https://replab.github.io/replab`.
[28] `https://github.com/felimomo/RepCert`.
[29] K. H. Hofmann and S. A. Morris, "Weight and c," *Journal of Pure and Applied Algebra*, vol. 68, no. 1-2, pp. 181–194, 1990.
[30] A. Basheer and J. Moori, "On the ranks of finite simple groups," *Khayyam Journal of Mathematics*, vol. 2, no. 1, pp. 18–24, 2016.
[31] L. Mackey, M. I. Jordan, R. Y. Chen, B. Farrell, J. A. Tropp, *et al.*, "Matrix concentration inequalities via the method of exchangeable pairs," *The Annals of Probability*, vol. 42, no. 3, pp. 906–945, 2014.
[32] R. Bhatia, *Matrix analysis*, vol. 169. Springer Science & Business Media, 2013.

[33] M. Magdon-Ismail, "A note on estimating the spectral norm of a matrix efficiently," *arXiv preprint arXiv:1104.2076*, 2011.
[34] S. Damelin and B. Mode, "A note on a quantitative form of the solovay-kitaev theorem," *arXiv preprint arXiv:1709.03007*, 2017.
[35] J. Bourgain and A. Gamburd, "A spectral gap theorem in $su(d)$," *arXiv preprint arXiv:1108.6264*, 2011.
[36] E. Breuillard and A. Lubotzky, "Expansion in simple groups," *arXiv preprint arXiv:1807.03879*, 2018.
[37] Y. Benoist and N. de Saxcé, "A spectral gap theorem in simple lie groups," *Inventiones mathematicae*, vol. 205, no. 2, pp. 337–361, 2016.
[38] F. Montealegre-Mora, "RepCert documentation," 2021. In preparation.
[39] P. P. Varjú, "Random walks in compact groups," *Documenta Mathematica*, vol. 18, pp. 1137–1175, 2013.

Appendix A: Proofs

Proposition 11. *Let Q_e be as in Prop. 7, and c_2 be as in the beginning of Sec. IV B. Then $\|Q_e\|_\infty \leq c_2$.*

Proof. Let $\rho_K(s) := \pi_K \rho(s) \pi_K$ and $D(s) := \tilde{\rho}_R(s) - \rho_K(s)$. Using subadditivity, we bound

$$\|Q_e\|_\infty \leq \max_s \|D(s) \otimes \bar{\rho}_K(s) + \rho_K(s) \otimes \bar{D}(s) + D(s) \otimes \bar{D}(s)\|_\infty \leq \max_s (2\|D(s)\|_\infty + \|D(s)\|_\infty^2).$$

Further writing $\Delta := \tilde{\pi}_R - \pi_K$ and $\Delta(s) := \tilde{\rho}(s) - \rho(s)$, we observe that

$$D(s) = \Delta\rho(s)(\pi_K + \Delta)^\dagger + (\pi_K + \Delta)\rho(s)\Delta^\dagger + (\pi_K + \Delta)\Delta(s)(\pi_K + \Delta)^\dagger,$$

and so,

$$\|D(s)\|_\infty \leq 2\|\Delta\|_\infty(1 + \|\Delta\|_\infty) + \|\Delta(s)\|_\infty(1 + \|\Delta\|_\infty)^2.$$

We can directly bound $\|\Delta(s)\|_\infty \leq n\epsilon_0$. Then, becaus R holds an invariance certificate with precision ϵ, we deduce

$$\|\Delta\|_\infty \leq n\epsilon_0 + \epsilon.$$

It follows that $\|D(s)\|_\infty \leq c_1$, where c_1 is defined as in the top of Sec. IV B, and the claim follows. □

Proof of Prop. 8. As in the proof of Prop. 11, let $D(s) := \tilde{\rho}_R(s) - \rho_K(s)$. For the sake of convenience, let us introduce the following notation: $B_1(s) = D(s)$, $B_0(s) = \rho_K(s)$, and for any bit string $v \in \mathbb{F}_2^k$ and $\mathbf{s} \in S^k$,

$$B_v(\mathbf{s}) = B_{v_1}(s_1)B_{v_2}(s_2)\cdots B_{v_k}(s_k).$$

Then, using submultiplicativity, subadditivity and unitary invariance we find that

$$\begin{aligned}
|\mathrm{tr}(\tilde{\rho}_R(\mathbf{s}))|^2 &\leq \sum_{v \in \mathbb{F}_2^k} |\mathrm{tr}\, B_v(\mathbf{s})|^2 \\
&\leq \sum_{v \in \mathbb{F}_2^k} \|B_v(\mathbf{s})\|_F^2 \\
&\leq \dim^2 K + \sum_{v \neq 0} \max_s \|D(s)\|_F^{\mathrm{wt}(v)} \\
&\leq \dim^2 K + \sum_{w=1}^{k} \binom{k}{w} \max_s \|D(s)\|_F^w \\
&\leq \dim^2 K + \Big(1 + \max_s \|D(s)\|_F\Big)^k - 1,
\end{aligned}$$

where $\mathrm{wt}(v)$ denotes the *Hamming weight* of v. Then, because R holds an invariance certificate with precision ϵ, we may use an argument analogous to the proof of Prop. 11 to bound $\max_s \|D(s)\|_F$ by c_1 (defined in the top of Sec. IV B). This finalizes the proof. □

Proof of Prop. 9. We begin by observing that $d_k \leq d_{2t}$ for all $k \leq 2t$, and so Prop. 8 implies

$$\left| \sum_{k=0}^{2t-1} \binom{2t}{k} c_2^{2t-k} \,\mathrm{tr}\big((Q_S^R)^k\big) \right| \leq [(1 + c_2)^{2t} - 1](\dim^2 K + d_{2t}).$$

Since $\epsilon < 1/2$, $\dim K = \dim R$. Finally, $n\epsilon_0 < 1/2$ implies that $\mathrm{int}(\mathrm{tr}\,\tilde{\pi}_R) = \mathrm{tr}\,\pi_R = \dim R$. □

Proof of Prop. 10. By Prop. 8, $|\operatorname{tr}\tilde{\rho}_R(\mathbf{s}_i)|^2/(\dim^2 K + d_{2t})$ is a random variable in $[0,1]$, so Chernoff's bound gives

$$\Pr\left[\frac{1}{m}\sum_i |\operatorname{tr}\tilde{\rho}_R(\mathbf{s}_i)|^2 \leq (1-\theta)\operatorname{tr}\big((Q_S^R)^{2t}\big)\right] < \exp\left(\frac{-\theta^2 m \operatorname{tr}\big((Q_S^R)^{2t}\big)}{2(\dim^2 K + d_{2t})}\right)$$

But by Prop. 7 $\operatorname{tr}\big((Q_S^R + c_2\mathbb{I})^{2t}\big) \geq \operatorname{tr} P_{\text{Haar}}^K \geq 2$, and by Prop. 9 $\operatorname{tr}\big((Q_S^R)^{2t}\big) \geq 2 - e_t$, which finishes the proof. □

Appendix B: Extension to a weaker scenario

Here we show how to modify our algorithms to a setting in which the user has considerably less control over the group than is assumed in the main text. To keep the the line of argument clean, we provide only short proof sketches for the claimed statements, and include these at the end of the appendix. In the following, the Lie algebra $\mathfrak{g}$ of G is endowed with a G-invariant inner product $\langle\cdot,\cdot\rangle_{\mathfrak{g}}$ and a corresponding 2-norm $\|\cdot\|_{\mathfrak{g}}$.

In the current setting, the user is assumed to know $\tilde{\rho}$ evaluated on a *fixed* symmetric generator set S. The set S and the representation ρ must also satisfy two requirements.

The first is that S is not too 'ill-conditioned': We say that S is (δ,k)-*dense*, if for any $g \in G$ there exists a word $s_1\cdots s_k$ of length k in S for which

$$\left\|\log g^{-1}s_1\cdots s_k\right\|_{\mathfrak{g}} \leq \delta.$$

The second requirement is that the ρ-images of close-by group elements are also close-by. That is, we say that ρ is q-bounded if it holds that

$$\|\mathrm{d}\rho(X)\|_F \leq q\,\|X\|_{\mathfrak{g}}\,, \qquad \forall X \in \mathfrak{g},$$

where $\mathrm{d}\rho$ is the representation of $\mathfrak{g}$ corresponding to ρ. In summary, we assume that the user knows some numbers (δ,k,q) such that S is (δ,k)-dense and ρ is q-bounded (we say that (G,S,ρ) is (δ,k,q)-well conditioned).

In the case G is finite, one may take k to be the Cayley diameter and $q=\delta=0$. When G is continuous, to the best of our knowledge there are no explicit generator sets S known to be (δ,k)-dense. For special unitary groups, the Solovay-Kitaev theorem provides an asymptotic result: certain generator sets are $(\delta, O(\log^4\delta^{-1}))$-dense. In the case of SU(2), some progress towards an explicit scaling for the Solovay-Kitaev theorem has been made in [34].

Remark 1. *One can modify the algorithms presented here to use a bound on the spectral gap $\|P_S - P_{\text{Haar}}\|_\infty$ as an input instead of (δ,k,q). However, such a bound is rarely known without diagonalizing P_S. While results stating the* existence *of a gap exist for a variety of compact groups, these do not quantify how large it is (e.g. [35–37]). Because of this, we do not present such a modification.*

1. Invariance certificate

The invariance certificate in this scenario is given by Alg. B.1, where we use

$$f(x) = 2\sqrt{2\dim R}\,(xk + 2kn\epsilon_0(n\epsilon_0+1) + 2q\delta\exp(q\delta) + 2n\epsilon_0)\,. \tag{B1}$$

Algorithm B.1 Modified invariance certificate

Input:

- $\{\tilde{\rho}(s) : s \in S\} \subset \mathbb{C}^{n\times n}$,
- $\delta \in (0,1)$, $k \in \mathbb{N}$, $q \in \mathbb{R}_+$, ▷ (G,S,ρ) is (δ,k,q)-well conditioned.
- $\tilde{\pi}_R \in \mathbb{C}^{n\times n}$,
- $\epsilon \in (0,1/2)$.

Output: *True/False*

1: Let f be defined as in eq. (B1)
2: **if** $f(\max_{s\in S}\|[\tilde{\rho}(s),\tilde{\pi}_R]\|_F) \leq \epsilon$ **then**
3: **Return:** *True*
4: **end if**
5: **Return:** *False*

As in the main text, the key quantity to be bounded is $\|P_{\text{Haar}}(\pi_R) - \pi_R\|_F$. This is acheived by the following two propositions.

Proposition 12. *Let (G, S, ρ) be (δ, k, q)-well conditioned and assume that*

$$\|[\tilde{\rho}(s), \tilde{\pi}_R]\|_F \leq c_3, \qquad \forall\, s \in S.$$

Then, for all $g \in G$ we have that

$$\|[P_{\text{Haar}}(\pi_R) - \pi_R\|_F \leq kc_3 + 2kn\epsilon_0(n\epsilon_0 + 1) + 2q\delta\exp(q\delta) =: c_4(c_3).$$

Putting this together with Prop. 3 shows that if Alg. B.1 returns *True*, then R is approximately invariant up to precision ϵ.

2. Irreducibility certificate

We now move on to the irreducibility certificate. For simplicity we only present the procedure in the ideal case, given by Alg. B.2. The certificate is in essence the same as Alg. IV.1, with the prominent difference that S is not sampled at the start. The proof of Thm. 2 carries over exactly to the current case showing that this algorithm's false positive rate is at most $p_{\text{thr.}}$.

Alg. B.2 furthermore includes the parameter t as an input (compare line 3 of Alg. IV.1). This choice is made for the sake of performance. Specifically, in Prop. 13 we bound the false negative rate whenever t is large enough —this is in the same spirit as Prop 6. Here, though, the bound on t is too large to be useful in many practical settings.

Rather than using Prop. 13 to choose t, we have instead tested the performance of the algorithm for different values of t (see [38]). There it is found that, for a variety of finite group representations, taking $t \gtrsim k$ is sufficient to bring the empirical false negative rate down to zero.

Algorithm B.2 Modified ideal irreducibility certificate

Input:

- $\{\rho_K(s) : s \in S\} \subset \mathbb{C}^{n_K \times n_K}$,
- $p_{\text{thr.}} \in (0, 1)$,
- $t \in \mathbb{N}$.

Output: *True/False.*

1: Set $m = 3\lceil n_K^2 \log(1/p_{\text{thr.}})\rceil + 1$
2: Set $\theta_m = n_K\sqrt{\frac{2\log(1/p_{\text{thr.}})}{m}}$
3: Compute $E_m = \frac{1}{m}\sum_{i=1}^{m} |\operatorname{tr}\rho_K(\mathbf{s}_i)|^2$, with $\mathbf{s}_i \in S^{2t}$ sampled uniformly
4: **if** $E_m < 2(1 - \theta_m)$ **then**
5: **return** *True*
6: **end if**
7: **return** *False*

We thus conclude by analysing the false negative rate of Alg. B.2. This probability is intimately related to the spectral gap of P_S^K, —the mixing time of random walks in S. Here, we show how to obtain a bound on this spectral gap from the parameters (δ, k, q). This result follows from Ref. [39, Lemma 5] up to some minor technical detail.

Proposition 13. *There exists a constant c_0 such that for any compact group G, generator set $S \subset G$ and* irreducible *representation ρ_K the following holds. If (G, S, ρ_K) is (δ, k, q)-well conditioned with $\delta \leq (c_0 q)^{-c_0}$, then for any*

$$t \geq \frac{1}{2}\frac{\log n - 1}{\log\frac{1}{1-1/|S|k^2}},$$

it holds that the probability that Alg. B.2 returns False *upon this input is at most*

$$\exp\left(\frac{-m}{3\dim^2 K}\left(\frac{2-\theta_m}{1+(n-1)(1-k^{-2}|S|^{-1})^{2t}} - 1\right)^2\right).$$

Our approach is the following. Ref. [39, Lemma 5] gives a bound on this spectral gap as a funtion of δ, k and a third parameter, the *maximal weight length* defined by

$$\max\left\{ \|\omega\|_{\mathfrak{g}^*}^2 \;\middle|\; \omega \text{ weight in } \rho_K \right\}.$$

The following proposition relates this quantity to our parameter q, which in turn allows us to obtain a bound on the mixing time in terms of (δ, k, q).

Proposition 14. *Let (K, ρ_K) be a unitary representation of G with maximal weight length equal to w. Then*

a) ρ_K is $\sqrt{w \dim K}$-bounded,

b) if ρ_K is q-bounded, then q must satisfy $q \geq w$.

3. Proofs

Proof of Prop. 12. We directly compute that for all $s \in S$

$$\|[\rho(s), \pi_R]\|_F \leq c_3 + 4n\epsilon_0 + 2n^2\epsilon_0^2 =: c_5.$$

Similarly, for any $\mathbf{s} \in S^k$,

$$\|[\rho(\mathbf{s}), \pi_R]\|_F \leq k c_5,$$

where we used the identity $[AB, C] = A[B, C] + [A, C]B$ iteratively.

Now, let $g \in G$ be arbitrary. By assumption, there exists a word $g_s := s_1 \cdots s_k$ in S, together with an element $g_X := \exp(X)$ for which

$$\begin{aligned} g &= g_s g_X, \\ \|X\|_{\mathfrak{g}} &\leq \delta. \end{aligned}$$

Subadditivity and submultiplicativity imply that

$$\begin{aligned} \|\rho(g) - \rho(g_s)\|_F &= \|\exp \mathrm{d}\rho(X) - \mathbb{I}\|_F \\ &\leq \|\mathrm{d}\rho(X)\|_F \exp(\|\mathrm{d}\rho(X)\|_F) \\ &\leq q\delta \exp(q\delta), \end{aligned}$$

and so

$$\|[\rho(g), \tilde{\pi}]\|_F \leq 2q\delta \exp(q\delta) + k c_5 = c_4, \qquad \forall\, g \in G.$$

Finally, we may use the unitarity of ρ to obtain

$$\|P_{\mathrm{Haar}}(\pi_R) - \pi_R\|_F \leq \mathbb{E}_{g\sim G}\left[\|[\rho(g), \pi_R]\|_F\right],$$

which proves the claim. □

Proof of Prop. 14. Let $\{\omega_i\}$ be the set of weights appearing in ρ_K, let ω_0 be a weight in that set with maximal length (so $\|\omega_0\|_{\mathfrak{g}^*}^2 = w$) and let $\mathfrak{t}$ be the Lie algebra of the maximal torus in G. We begin by noting that because $\|\cdot\|_{\mathfrak{g}}$ is invariant under the adjoint G-action, we know that

$$\sup_{X\in\mathfrak{g}} \frac{\|\mathrm{d}\rho_K(X)\|_F^2}{\|X\|_{\mathfrak{g}}^2} = \sup_{X\in\mathfrak{t}} \frac{\|\mathrm{d}\rho_K(X)\|_F^2}{\|X\|_{\mathfrak{g}}^2}.$$

For any $X \in \mathfrak{t}$,

$$\|\mathrm{d}\rho_K(X)\|_F^2 = \sum_i |\omega_i(X)|^2 = \sum_i |\langle \omega_i^*, X\rangle_{\mathfrak{g}}|^2, \tag{B2}$$

where ω_i^* is the dual of ω_i with respect to the invariant inner product. Using Cauchy-Schwartz on eq. (B2) we obtain

$$\|\mathrm{d}\rho_K(X)\|_F^2 \leq \|X\|_{\mathfrak{g}}^2 \sum_i \|\omega_i^*\|_{\mathfrak{g}}^2 \leq (w \dim K) \|X\|_{\mathfrak{g}}^2 ,$$

which proves the first statement.

For the second statement, let us choose $X = \omega_0^* / \|\omega_0^*\|_{\mathfrak{g}}$ in eq. (B2). We obtain

$$\|\mathrm{d}\rho_K(X)\|_F^2 = \sum_i \frac{|\langle \omega_i^*, \omega_0^* \rangle_{\mathfrak{g}}|^2}{\|\omega_0^*\|_{\mathfrak{g}}^2} \geq \|\omega_0^*\|_{\mathfrak{g}}^2 = w .$$

But $\|X\|_{\mathfrak{g}} = 1$ so any $q \leq w$ would be inconsistent with the equation above. □

Proof of Prop. 13. By Prop. 14, the maximal weight-length r of ρ_K can be at most q. Consider the random walk operator P_S associated to ρ_K and let λ be the spectral norm of the restriction of P_S to the traceless subspace, —by the assumption that ρ_K is irreducible, we know that $\lambda < 1$.

Ref. [39, Lemma 5] implies that there exists a universal constant $c_0 > 0$ such that if $\delta \leq (c_0 q)^{-c_0}$, then

$$1 - \lambda \geq \frac{1}{|S| k^2}.$$

Hence,

$$\operatorname{tr} P_S^{2t} \leq 1 + (n-1) \left(1 - \frac{1}{|S| k^2}\right)^{2t}. \tag{B3}$$

Then, for any $x \leq 1$, the right-hand side is smaller than $2 - x$ if and only if

$$t \geq \frac{1}{2} \frac{\log \frac{n-1}{1-x}}{\log \frac{1}{1-1/|S|k^2}} =: t_x.$$

Equivalently, for any t given as in the assumption of the theorem, the right-hand side of (B3) is at most $2 - x_t$, where

$$x_t := 1 - (n-1)(1 - 1/|S|k^2)^{2t}$$

The Chernoff bound implies that for any $\alpha > 0$, if $\{\mathbf{s}_i\}$ are m uniform samples from S^{2t}, then

$$\operatorname{Prob}\left[\frac{1}{m} \sum_i |\operatorname{tr} \rho_K(\mathbf{s}_i)| \geq (2 - x_t)(1 + \alpha)\right] \leq \exp\bigl(-\alpha^2 m / 3 \dim^2 K\bigr). \tag{B4}$$

Consider the choice

$$\alpha = \frac{2 - \theta_m}{2 - x_t} - 1,$$

where θ_m is as in Line 1 of Alg.IV.1. Then, eq. (B4) becomes

$$\operatorname{Prob}\left[\frac{1}{m} \sum_i |\operatorname{tr} \rho_K(\mathbf{s}_i)| \geq (2 - x_t)(1 + \alpha)\right] \leq \exp\left(\frac{-m}{3 \dim^2 K} \left(\frac{2 - \theta_m}{2 - x_t} - 1\right)^2\right). \tag{B5}$$

□

7 Implementation of certification algorithm

The algorithms discussed in Chap. 6 were coded into the Python package RepCert [MM21]. This code may interface with RepLAB: the output decomposition obtained with RepLAB may be directly fed as an input to RepCert. Here, I will showcase the main components of the latter code, together with benchmarks on the performance of the full pipeline RepLAB+RepCert. I was the main researcher in this project.

RepCert may handle the two mathematical settings discussed in [MMRBG21] (included as Chap. 6). Namely, it can handle on the one hand the *random generator case*, where one has oracle access to a Haar-random sampler from the group G. On the other hand, it can handle the harder *fixed generator case* discussed in the appendix of [MMRBG21], where one has access to a fixed set of generators and knows their mixing time. I omit the presentation of this second case for brevity.

7.1 The certification algorithm

Recall from Chap. 6 the following definition. Let π be the orthogonal projector onto a subspace $V \subset \mathbb{C}^n$. We say that V is ϵ-close to an irreducible subrepresentation $K \subseteq \mathbb{C}^n$ if $\|\pi_K - \pi\|_2 \leq \epsilon$, where π_K is the orthogonal projector onto K.

The general structure of the algorithm in the random generator case is the following. As an input it receives the following three types of parameters: 1. An $\epsilon_0 \in \mathbb{R}_+$ and a set of matrices $\mathcal{G} = \{\tilde{g}_i\} \subset \mathbb{C}^{n\times n}$ for which there exist $g_i \in G$ satisfying $\max_i \|\tilde{g}_i - g_i\|_{\max} \leq \epsilon_0$. 2. A set of subspaces $V_i \subseteq \mathbb{C}^n$, such that $\mathbb{C}^n = \oplus_i V_i$, and, for each i, V_i is claimed to be an irreducible subrepresentation of G. Let $d_i = \dim V_i$ and π_i be the orthogonal projector onto V_i. 3. A set of accuracy parameters: a bound ϵ_0 on $\max_{g\in G} \|\tilde{g} - g\|_{\max}$, the accuracy ϵ with which to certify invariance (see Alg. IV.2 in Chap. 6), the threshold *false positive rate* $p_{\text{thresh.}}$ and the approximate bound on the *false negative rate* $p'_{\text{thresh.}}$. The output is a list of booleans, say β_i, stating whether the space V_i was certified to be ϵ-close to an irreducible subrepresentation. The false *positive* rate is the probability that $\beta_i =$`True` conditioned on V_i *not being* ϵ-close to an irreducible subrepresentaiton. The false *negative* rate is the converse.

The algorithm then calls, for each V_i, a function `invariance` (presented in Sec. 7.1.2) to certify invariance. Let $\{b_i^j\}_j$ be an orthonormal basis of V_i and $B_i = (b_i^1, \ldots, b_i^{d_i}) \in \mathbb{C}^{n\times d_i}$. If this certification succeeds then, for each i the restricted generator images are constructed,

$$g_i := B_i^T g B_i \in \mathbb{C}^{d_i\times d_i}.$$

These images are subsequently used to certify irreducibility of V_i.

7.1.1 Objects

The basic data structure used by RepCert is the class `rep_by_generators` shown in a simplified form below. As suggested by the name, it is a representation defined on a set of generators of the group. Accordingly, the main properties of an object in this class are: the dimension n, a list of names for generators, and the set of generator images ($n \times n$ matrices). These generators are assumed to have been sampled from the Haar measure on G, as mentioned above. While in this simpler case the choice of an object-oriented approach migh seem overkill, this approach was taken for its flexibility and improved readability. In particular, it allows the same data structure to handle the fixed generator case.

```
class rep_by_generators():
  def __init__(self, dimension, generatorSet = [], genImages =
     [], **kwargs):
     self.dimension = dimension #dimension of representation
     self.nGens = len(generatorSet) #number of generators
     self.Images = dict()
     if len(genImages)>0:
        self.Images = {generatorSet[i] : genImages[i] for i in
           range(len(genImages))}

  def add_generator_image(self,element,repImage):
     assert isinstance(element,group_element)
     self.Images.update({element.name : repImage})
     self.generatorSet.append(element)
     self.nGens += 1

  def image_list(self):
     return [self.Images[g] for g in self.generatorList]
```

7.1.2 The certification step

To certify invariance, the following code was used. There, `repr` is the `rep_by_generators` object defined by the input, `proj` is the projector $P_i := B_i B_i^T$ for some i, and `fl` is ϵ_0 and P_i (typically, `fl` is roughly the *floating point precision*). Moreover, the code uses the subroutine `averaging`, which computes

$$\left\| \mathbb{E}_{g \in \texttt{generatorSet}} [g P_i \, g^\dagger - P_i] \right\|_F .$$

```
def invariance(repr,proj,epsilon,pthresh,fl):
  epsprime =
     epsilon/(2*math.sqrt(2.*math.ceil(lin.trace(proj).real)))
  c = averaging(repr,proj)
  n = repr.dimension
  f_err = 8*n*fl + 6*(n*fl)**2 + 2*(n*fl)**3
  if 2*c + f_err <= epsprime:
     return True
  return False
```

As mentioned above, once the blocks B_i have been certified to be approximately invariant, the representation is restricted to each one of the blocks, giving generator images $g_i = B_i^T g B_i \in \mathbb{C}^{d_i \times d_i}$. These images are used to define a new `rep_by_generators` object using the function below. Here, the function `lin.restrict` performs the restricting step $g \mapsto g_i$.

```
def restrict_to_subrep(repr,basis):
  new_ims = [lin.restrict(im,basis) for im in
     repr.image_list()] # new rep images of generators
  dim = len(basis) # new dimension
  return rep.rep_by_generators(dim, repr.generatorList,
     new_ims)
```

Finally, this restricted approximate representation is tested for irreducibility. Here, I use the following conventions: `pthresh` is the bound on the false positive rate (denoted $p_{\text{thresh.}}$ in Chap. 6), `conf` is the approximate bound on the false negative rate (dentoed δ_{conf} in Chap. 6), and `epsilon` is the accuracy of the invariance test. The functions `rwalk.set_t` and `rwalk.number_samples` set the length of the random walk and the number of random walk samples used (as in lines 5 and 6 of Alg. IV.2 in Chap. 6). Similarly, the functions `const.et` and `const.dt` compute the constants e_t and d_t defined in Sec. IV of that same chapter. Finally, let $V_i = \text{span}\mathcal{B}_i$. Then, the function `rwalk.repRandWalkEstimator` computes

$$\frac{1}{m}\sum_{i=1}^{m} |\operatorname{tr} \mathbf{s}_i|^2,$$

where $\mathbf{s}_i \in \text{End}(V_i)$ is the outcome of a random walk of length $2t$ on the *restricted generator images* $\{g_i\}_{g \in \texttt{generatorSet}}$. In other words,

$$\mathbf{s}_i = g_i^{(1)} \dots g_i^{(2t)},$$

where $g_i^{(j)}$ are sampled uniformly from $\{g_i\}_{g \in \texttt{generatorSet}}$.

```
def irr_cert(repr,epsilon,pthresh,pthresh_pr):
  # parameters and constants:
  dim = repr.dimension
  if dim==1:
     return True

  # Random walk parameters #
  t = rwalk.set_t(repr)
  m =
     rwalk.number_samples(repr,dim,epsilon,pthresh,t,pthresh_pr)

  # other constants
  et = const.et(repr,epsilon,t,dim)
  dt = const.dt(repr,epsilon,t)
  aux = dim**2+dt
  aux*= 2*math.log(pthresh**(-1))

  if et >=2 or m <= aux:
     return False

  # Character length estimation:
  theta = math.sqrt(aux * (m*(2-et))**(-1))
  E = et + rwalk.repRandWalkEstimator(repr,m,t)

  # Irreducibility condition:
  if E < 2*(1-theta):
     return True
  return False
```

7.2 Numerical benchmarks on RepLAB+RepCert

Here I present some simple benchmarks on the performance of the "full monty" algorithm: the sequential combination of RepLAB and RepCert. Namely, I consider a set of abstract groups G and representations ρ of G. In a first step, I use RepLAB to decompose ρ. Subsequently, I use RepCert to certify that the components of this decomposition are ϵ-close to irreducible subrepresentations.

Two things are tested in these benchmarks: the ability of RepLAB to find accurate decompositions of representations on the one hand, and on the other the runtime required by RepCert.

Many of the known instances of symmetries in SDPs arising from quantum problems are either permutation groups, or wreath products thereof [TCUA20, RRMG17, TFR$^+$21]. I have thus focused on benchmarking the performance of RepLAB and RepCert on these types of groups. Namely, I consider representations of permutation groups S_a and of the following wreath products:

$$G_{ab} := S_a \wr S_b,$$
$$G_{abc} := S_a \wr S_b \wr S_c,$$

where S_a is the symmetric group of degree a, and where for any group G, $G \wr S_x$ denotes the wreath product with respect to the natural representation of the symmetric group S_x. The groups G_{abc} correspond to the symmetries of the Bell inequalities for a scenario with c parties, b measurement settings, and a measurement outcomes for each setting.

Let R be an arbitrary group, then the group elements of $G' := R \wr S_d$ are of the form

$$g' = ((r_1, \ldots, r_d), s), \tag{27}$$

where $r_i \in R$ and $s \in S_d$. The group law is given by

$$((r_1, \ldots, r_d), s) \cdot ((h_1, \ldots, h_d), q) = ((r_1 h_{s^{-1}(1)}, \ldots, r_d h_{s^{-1}(d)}), sq).$$

Two types of representations of G' are relevant for our purposes: Let $(\sigma, \mathcal{H}_\sigma)$ be a representation of G. The *S_d-imprimitive representation of σ*, ξ_σ, acts on the space

$$\mathcal{H}_\sigma \otimes \mathbb{C}^d,$$

as

$$\xi_\sigma(g') \left|\psi\right\rangle \otimes \left|i\right\rangle = \left(\sigma(r_i) \left|\psi\right\rangle\right) \otimes \left|s(i)\right\rangle,$$

where $g' \in G'$ is given as in eq. (27), $\left|\psi\right\rangle \in \mathcal{H}_\sigma$ and where $\{\left|i\right\rangle \mid i = 1, \ldots, x\}$ is the computational basis of $\mathbb{C}^d$. The *S_d-primitive representation of σ*, Ξ_σ, acts on the space

$$(\mathcal{H}_\sigma)^{\otimes d},$$

where it acts as

$$\Xi_\sigma(g') = (\sigma(r_{s^{-1}(1)}) \otimes \cdots \otimes \sigma(r_{s^{-1}(x)}))\pi_s,$$

where π_s permutes the d tensor factors of $\mathcal{H}_\sigma$. An in-depth discussion of wreath products and the representations of wreath product groups may be found in Ref. [CSST14].

Let $(\pi_a, \mathbb{C}^a)$ be the natural representation of S_a. That is, its action on the computational basis is given by

$$\pi_a(s)\,|i\rangle = |s(i)\rangle\,, \quad i = 1, \ldots, a,$$

where $s \in S_a$. The following representations were decomposed using RepLAB and subsequently certified using RepCert. *1.* The third tensor power $\rho_a := \pi_a^{\otimes 3}$ of the natural representation of S_a. *2.* The S_b-primitive representation of the S_a natural representation, ρ_{ab}, of G_{ab}. *3.* The S_c-primitive representation of the S_b-imprimitive representation of the S_a natural representation, denoted ρ_{abc}, of G_{abc}. The latter is precisely the symmetry appearing in Bell inequality scenarios [RRMG17].

7.3 Benchmark results

Here I display the results of the benchmark tests introduced above. All the computations were run on the commercial desktop computers used by the network of the Institute of Theoretical Physics, at the University of Cologne.

For each benchmarked representation ρ, RepLAB was used to find an alleged decomposition. In displaying these decompositions, I use the following notation: If $\rho : G \to \mathbb{C}^{n\times n}$ is a representation, I write

$$\rho \simeq \bigoplus_i C(d_i) \otimes \mathbb{C}^{m_i}$$

to denote a claimed decomposition of ρ. Here, $C(d_i)$ is an irreducible representation of G with dimension d_i. If $d_i = d_j$ for two distinct terms $i \neq j$ in the sum above, the representations are non-equivalent. In a second step, I used RepCert to certify these blocks: Each block was tested for invariance, and all blocks of dimension < 150 were tested for irreducibility. The reason for the latter limitation was the long runtimes expected for higher dimensions.

In these benchmarks, I have the following parameter values of the certification algorithms (using the notation of Chap. 6, Algs. III.2 and IV.2):

$$\begin{aligned} p_{\text{thresh.}} &= 10^{-7} \\ \delta_{\text{conf.}} &= 2p_{\text{thresh.}} \\ \epsilon &= 10^{-8} \\ \epsilon_0 &= 2^{-52} \quad \text{(machine precision).} \end{aligned}$$

While most of the decompositions obtained were fully certified with these parameters, the invariance certification failed for some of the higher dimensional blocks. Here, I subsequently ran an invariance certification of these blocks with a lower accuracy: $\epsilon = 10^{-7}$.

Three observations can be highlighted from the following results: First, every tested block obtained from RepLAB was certified for invariance with an accuracy of at most $\epsilon = 10^{-7}$. Blocks with dimensions ≤ 108 were all certified with the higher accuracy of $\epsilon = 10^{-8}$. This gives further evidence which points at the accuracy of RepLAB decompositions. Second, the runtimes are typically dominated by the irreducibility certifation step: in this step, one first restricts the generators $g \in G$ to the component, and then performs a the random walk of Alg. IV.2 from Chap. 6. The most "expensive" cases tested were components with dimensions $d \gtrsim 100$, whose certification took on the order of magnitude of an hour. Third, the representations ρ_{abc} associated to the symmetries of Bell inequalities [RRMG17] decompose into irreps with dimension *much* smaller than $\dim \rho_{abc}$. This way, these results open the possibility of solving the previously inaccessible SDPs associated to multi-party Bell scenarios.

Permutation groups S_a. The values of $a = 6,\ 7,\ 8,\ 9$, and 10 were used as benchmarks. Replab produced the following decompositions:

$$\begin{aligned}
\rho_6 = {} & (C(16)\otimes\mathbb{C}^2)\oplus(C(10)\otimes\mathbb{C}^6)\oplus(C(9)\otimes\mathbb{C}^6)\oplus(C(5)\otimes\mathbb{C}^{10}) \\
& \oplus C(5)\oplus(C(1)\otimes\mathbb{C}^5), \\
\rho_7 = {} & (C(35)\otimes\mathbb{C}^2)\oplus C(20)\oplus(C(15)\otimes\mathbb{C}^6)\oplus(C(14)\otimes\mathbb{C}^6)\oplus(C(6)\otimes\mathbb{C}^{10}) \\
& \oplus(C(1)\otimes\mathbb{C}^5), \\
\rho_8 = {} & (C(64)\otimes\mathbb{C}^2)\oplus C(35)\oplus C(28)\oplus(C(21)\otimes\mathbb{C}^6)\oplus(C(20)\otimes\mathbb{C}^6) \\
& \oplus(C(7)\otimes\mathbb{C}^{10})\oplus(C(1)\otimes\mathbb{C}^5), \\
\rho_9 = {} & (C(105)\otimes\mathbb{C}^2)\oplus C(56)\oplus C(48)\oplus(C(28)\otimes\mathbb{C}^6)\oplus(C(27)\otimes\mathbb{C}^6) \\
& \oplus(C(8)\otimes\mathbb{C}^{10})\oplus(C(1)\otimes\mathbb{C}^5), \\
\rho_{10} = {} & (C(160)\otimes\mathbb{C}^2)\oplus C(84)\oplus C(75)\oplus(C(36)\otimes\mathbb{C}^6)\oplus(C(35)\otimes\mathbb{C}^6) \\
& \oplus(C(9)\otimes\mathbb{C}^{10})\oplus(C(1)\otimes\mathbb{C}^5).
\end{aligned}$$

These decompositions were certified with RepCert, the results are shown in Tab. 2.

(a,b,c)	$\dim \rho_{abc}$	Irr. d	Inv. Time (s)	Restr. Time (s)	Irr. Time (s)	Inv.	Irr.
6	216	16	0.375	3.522	9.817	Yes	Yes
6	216	16	0.296	3.672	9.809	Yes	Yes
6	216	10	0.458	0.964	3.864	Yes	Yes
6	216	10	0.398	1.006	3.788	Yes	Yes
6	216	10	0.365	0.975	3.716	Yes	Yes

6	216	10	0.297	1.110	3.712	Yes	Yes
6	216	10	0.297	0.984	3.650	Yes	Yes
6	216	10	0.297	0.979	3.800	Yes	Yes
6	216	10	0.296	1.182	3.684	Yes	Yes
6	216	9	0.434	0.750	3.039	Yes	Yes
6	216	9	0.300	0.751	2.998	Yes	Yes
6	216	9	0.297	0.819	3.011	Yes	Yes
6	216	9	0.296	0.923	2.966	Yes	Yes
6	216	9	0.296	0.890	3.053	Yes	Yes
6	216	9	0.296	0.751	3.044	Yes	Yes
6	216	5	0.426	0.197	0.780	Yes	Yes
6	216	5	0.376	0.198	0.758	Yes	Yes
6	216	5	0.297	0.363	0.786	Yes	Yes
6	216	5	0.297	0.198	0.790	Yes	Yes
6	216	5	0.297	0.198	0.786	Yes	Yes
6	216	5	0.297	0.197	0.778	Yes	Yes
6	216	5	0.297	0.197	0.765	Yes	Yes
6	216	5	0.297	0.196	0.777	Yes	Yes
6	216	5	0.297	0.195	0.781	Yes	Yes
6	216	5	0.297	0.195	0.775	Yes	Yes
6	216	5	0.297	0.194	0.774	Yes	Yes
6	216	1	0.421	0.018	1.907	Yes	Yes
6	216	1	0.300	0.018	2.145	Yes	Yes
6	216	1	0.300	0.018	2.145	Yes	Yes
6	216	1	0.296	0.018	2.145	Yes	Yes
6	216	1	0.097	0.007	2.145	Yes	Yes
7	343	35	1.318	40.22	71.22	Yes	Yes
7	343	35	1.158	41.88	70.04	Yes	Yes
7	343	20	1.157	8.432	18.57	Yes	Yes
7	343	15	1.373	3.841	8.839	Yes	Yes
7	343	15	1.366	3.914	8.937	Yes	Yes
7	343	15	1.336	4.037	8.888	Yes	Yes
7	343	15	1.330	3.756	8.708	Yes	Yes
7	343	15	1.298	4.072	8.675	Yes	Yes
7	343	15	1.159	3.936	8.969	Yes	Yes
7	343	14	1.271	3.233	7.570	Yes	Yes
7	343	14	1.176	3.337	7.739	Yes	Yes
7	343	14	1.159	3.474	7.735	Yes	Yes
7	343	14	1.159	3.407	7.652	Yes	Yes

7	343	14	1.157	3.377	7.666	Yes	Yes
7	343	14	1.156	3.398	7.605	Yes	Yes
7	343	14	1.156	3.383	7.621	Yes	Yes
7	343	6	1.353	0.442	1.158	Yes	Yes
7	343	6	1.333	0.437	1.140	Yes	Yes
7	343	6	1.323	0.447	1.169	Yes	Yes
7	343	6	1.313	0.445	1.149	Yes	Yes
7	343	6	1.302	0.440	1.144	Yes	Yes
7	343	6	1.282	0.446	1.134	Yes	Yes
7	343	6	1.281	0.442	1.147	Yes	Yes
7	343	6	1.259	0.440	1.154	Yes	Yes
7	343	6	1.157	0.445	1.148	Yes	Yes
7	343	6	1.154	0.447	1.151	Yes	Yes
7	343	1	1.282	0.041	2.384	Yes	Yes
7	343	1	1.197	0.041	2.145	Yes	Yes
7	343	1	1.159	0.041	2.145	Yes	Yes
7	343	1	1.158	0.041	2.384	Yes	Yes
7	343	1	0.343	0.013	2.145	Yes	Yes
8	512	64	2.841	296.4	329.8	Yes	Yes
8	512	64	2.813	293.4	327.9	Yes	Yes
8	512	35	2.752	49.83	75.75	Yes	Yes
8	512	28	2.750	26.19	39.80	Yes	Yes
8	512	21	2.828	12.09	21.29	Yes	Yes
8	512	21	2.819	12.40	21.80	Yes	Yes
8	512	21	2.768	12.12	21.34	Yes	Yes
8	512	21	2.753	12.35	22.03	Yes	Yes
8	512	21	2.740	12.33	20.96	Yes	Yes
8	512	21	2.732	12.17	21.68	Yes	Yes
8	512	20	2.865	10.91	19.40	Yes	Yes
8	512	20	2.790	10.71	19.44	Yes	Yes
8	512	20	2.766	10.99	19.34	Yes	Yes
8	512	20	2.753	11.22	19.75	Yes	Yes
8	512	20	2.747	10.97	19.39	Yes	Yes
8	512	20	2.745	11.23	19.92	Yes	Yes
8	512	7	2.889	1.001	1.713	Yes	Yes
8	512	7	2.826	0.970	1.611	Yes	Yes
8	512	7	2.803	0.997	1.658	Yes	Yes
8	512	7	2.787	0.981	1.654	Yes	Yes
8	512	7	2.774	1.087	1.672	Yes	Yes

8	512	7	2.761	1.004	1.631	Yes	Yes
8	512	7	2.756	1.059	1.730	Yes	Yes
8	512	7	2.733	1.038	1.666	Yes	Yes
8	512	7	2.721	0.995	1.671	Yes	Yes
8	512	7	2.719	1.008	1.686	Yes	Yes
8	512	1	2.850	0.089	1.907	Yes	Yes
8	512	1	2.770	0.089	1.668	Yes	Yes
8	512	1	2.747	0.088	2.384	Yes	Yes
8	512	1	2.707	0.087	2.145	Yes	Yes
8	512	1	0.926	0.022	1.668	Yes	Yes
9	729	105	12.01	1459.	2039.	Yes	Yes
9	729	105	11.92	1457.	2005.	Yes	Yes
9	729	56	11.88	242.0	237.9	Yes	Yes
9	729	48	11.89	163.0	163.4	Yes	Yes
9	729	28	12.12	36.48	40.96	Yes	Yes
9	729	28	12.09	36.71	41.09	Yes	Yes
9	729	28	12.05	37.30	41.13	Yes	Yes
9	729	28	12.02	36.07	41.21	Yes	Yes
9	729	28	11.93	36.76	41.27	Yes	Yes
9	729	28	11.90	36.83	41.26	Yes	Yes
9	729	27	11.95	33.37	38.17	Yes	Yes
9	729	27	11.93	33.41	37.95	Yes	Yes
9	729	27	11.92	33.70	37.90	Yes	Yes
9	729	27	11.91	33.82	37.91	Yes	Yes
9	729	27	11.90	33.61	38.17	Yes	Yes
9	729	27	11.88	33.74	37.67	Yes	Yes
9	729	8	12.15	2.914	2.259	Yes	Yes
9	729	8	12.05	2.924	2.236	Yes	Yes
9	729	8	12.02	2.801	2.251	Yes	Yes
9	729	8	12.01	3.147	2.229	Yes	Yes
9	729	8	11.96	2.987	2.239	Yes	Yes
9	729	8	11.95	2.964	2.275	Yes	Yes
9	729	8	11.92	3.070	2.232	Yes	Yes
9	729	8	11.91	2.983	2.255	Yes	Yes
9	729	8	11.89	3.055	2.261	Yes	Yes
9	729	8	11.88	2.920	2.250	Yes	Yes
9	729	1	3.277	0.054	3.337	Yes	Yes
9	729	1	12.16	0.256	2.622	Yes	Yes
9	729	1	12.10	0.254	2.145	Yes	Yes

9	729	1	11.91	0.251	1.907	Yes	Yes
9	729	1	11.90	0.338	2.384	Yes	Yes

Table 2: Benchmark for the representations ρ_a of S_a for several such groups. The parameter d is the dimension of the space being certified for invariance and irreducibility, *Restr. Time* is the time necessary to restrict the sampled group elements to the corresponding block (i.e. using the notation above, the map $g \mapsto g_i$ for the i-th block). The last two columns specify whether invariance and irreducibility were certified for the representation. The irreducibility certification algorithm was only run for blocks with $d < 150$ due to the long expected runtime for larger blocks.

Wreath product groups G_{ab}. The values of $(a, b) = (2, 7)$, $(2, 8)$, $(3, 5)$, $(3, 6)$, and $(4, 4)$ were used as benchmarks. Replab produces the following decompositions:

$$\begin{aligned}
\rho_{27} &\simeq C(35) \oplus C(35) \oplus C(21) \oplus C(21) \oplus C(7) \oplus C(7) \oplus C(1) \oplus C(1),\\
\rho_{28} &\simeq C(70) \oplus C(56) \oplus C(56) \oplus C(28) \oplus C(28) \oplus C(8) \oplus C(8) \oplus C(1)\\
&\quad \oplus C(1),\\
\rho_{35} &\simeq C(80) \oplus C(80) \oplus C(40) \oplus C(32) \oplus C(10) \oplus C(1),\\
\rho_{36} &\simeq C(240) \oplus C(192) \oplus C(160) \oplus C(64) \oplus C(60) \oplus C(12) \oplus C(1),\\
\rho_{44} &\simeq C(108) \oplus C(81) \oplus C(54) \oplus C(12) \oplus C(1).
\end{aligned}$$

These decomposition were fed as input into RepCert. The results of the certification step are shown in Tabs. 3 (with $\epsilon = 10^{-8}$). As can be seen in the aforementioned table, for certain high-dimensional blocks the algorithm failed to certify invariance at this precision. Because of this, the invariance certification algorithm was run again on these blocks, this time with the lower precision of $\epsilon = 10^{-7}$. The results are displayed in Tab. 4.

The dimensions of the blocks in Tab. 4 are ≥ 160, which leads us to expect that a long runtime of the irreducibility certification algorithm on these blocks. In particular, Tab. 6 shows that, the runtime required to restrict to a 120-dimensional subrepresentation and to certify that it is irreducible, can be well above $1h$. Because of this, I have decided to not certify irreducibility of the blocks in Tab. 4, but rather only to certify invariance. I leave a more comprehensive testing of RepCert – including the certification of irreducibility of the blocks in Tab. 4 – for future work.

(a, b, c)	$\dim \rho_{abc}$	Irr. d	Inv. Time (s)	Restr. Time (s)	Irr. Time (s)	Inv.	Irr.
(2,7)	128	35	0.074	31.31	70.09	Yes	Yes
(2,7)	128	35	0.074	31.42	70.78	Yes	Yes
(2,7)	128	21	0.074	7.286	20.56	Yes	Yes
(2,7)	128	21	0.074	7.373	20.73	Yes	Yes
(2,7)	128	7	0.075	0.372	1.592	Yes	Yes

(2,7)	128	7	0.074	0.370	1.562	Yes	Yes
(2,7)	128	1	0.075	0.011	1.430	Yes	Yes
(2,7)	128	1	0.028	0.005	1.668	Yes	Yes
(2,8)	256	70	0.512	270.4	463.5	Yes	Yes
(2,8)	256	56	0.408	142.5	224.5	Yes	Yes
(2,8)	256	56	0.406	142.7	226.9	Yes	Yes
(2,8)	256	56	0.408	142.5	224.5	Yes	Yes
(2,8)	256	28	0.406	18.69	40.07	Yes	Yes
(2,8)	256	28	0.408	18.62	39.32	Yes	Yes
(2,8)	256	8	0.509	0.674	2.123	Yes	Yes
(2,8)	256	8	0.405	0.768	2.149	Yes	Yes
(2,8)	256	1	0.455	0.026	1.668	Yes	Yes
(2,8)	256	1	0.136	0.009	2.384	Yes	Yes
(3,5)	243	80	0.426	433.6	672.7	Yes	Yes
(3,5)	243	80	0.425	434.1	673.8	Yes	Yes
(3,5)	243	40	0.427	57.81	99.53	Yes	Yes
(3,5)	243	32	0.426	30.34	52.41	Yes	Yes
(3,5)	243	10	0.562	1.167	3.971	Yes	Yes
(3,5)	243	1	0.145	0.009	2.384	Yes	Yes
(3,6)	729	240	7.642	-	-	No	-
(3,6)	729	192	7.790	-	-	No	-
(3,6)	729	160	7.721	-	-	No	-
(3,6)	729	64	7.816	361.2	327.0	Yes	Yes
(3,6)	729	60	7.737	285.7	266.2	Yes	Yes
(3,6)	729	12	7.747	4.867	5.896	Yes	Yes
(3,6)	729	1	2.110	0.040	2.145	Yes	Yes
(4,4)	256	108	0.521	1000.	1689.	Yes	Yes
(4,4)	256	81	0.416	423.1	733.8	Yes	Yes
(4,4)	256	54	0.411	130.4	210.8	Yes	Yes
(4,4)	256	12	0.416	1.953	5.862	Yes	Yes
(4,4)	256	1	0.138	0.009	2.145	Yes	Yes

Table 3: Benchmark for the representations ρ_{ab} of G_{ab} for several such groups. The parameter d is the dimension of the space being certified for invariance and irreducibility, *Restr. Time* is the time necessary to restrict the sampled group elements to the corresponding block (i.e. using the notation above, the map $g \mapsto g_i$ for the i-th block). The last two columns specify whether invariance and irreducibility were certified for the representation. The irreducibility certification algorithm was only run for blocks with $d < 150$ due to the long expected runtime for larger blocks.

(a,b,c)	$\dim \rho_{abc}$	Irr. d	Inv. Time (s)	Inv.
(3,6)	729	240	7.657	Yes
(3,6)	729	192	7.762	Yes
(3,6)	729	160	7.653	Yes

Table 4: Certification of invariance of the two high dimensional blocks which failed the invariance test in Tab. 3. These were certified using $\epsilon = 10^{-7}$, as opposed to the lower value of $\epsilon = 10^{-8}$ used in Tab. 3. The other parameters are the same as those used in that table.

Double wreath product groups G_{abc}. The values of $(a,b,c) = (2,2,3)$, $(2,2,4)$, $(2,3,3)$, and $(3,3,3)$ were used as benchmarks. RepLAB produces the following decompositions:

$$
\begin{aligned}
\rho_{223} \simeq\ & C(12) \oplus C(12) \oplus C(8) \oplus C(6) \oplus C(6) \oplus C(3) \oplus C(3) \oplus C(1) \oplus C(1), \\
\rho_{224} \simeq\ & C(48) \oplus C(32) \oplus C(32) \oplus C(24) \oplus C(24) \oplus C(24) \oplus C(24) \oplus C(16) \\
& \oplus C(8) \oplus C(8) \oplus C(6) \oplus C(4) \oplus C(4) \oplus C(1) \oplus C(1), \\
\rho_{225} \simeq\ & C(160) \oplus C(120) \oplus C(120) \oplus C(80) \oplus C(80) \oplus C(80) \oplus C(80) \oplus C(60) \\
& \oplus C(40) \oplus C(40) \oplus C(40) \oplus C(40) \oplus C(32) \oplus C(10) \oplus C(10) \oplus C(10) \\
& \oplus C(10) \oplus C(5) \oplus C(5) \oplus C(1) \oplus C(1) \\
\rho_{233} \simeq\ & C(54) \oplus C(36) \oplus C(36) \oplus C(27) \oplus C(27) \oplus C(27) \oplus C(12) \oplus C(9) \\
& \oplus C(8) \oplus C(6) \oplus C(1), \\
\rho_{233} \simeq\ & C(216) \oplus C(216) \oplus C(108) \oplus C(72) \oplus C(18) \oplus C(12) \oplus C(6) \oplus C(8) \\
& \oplus C(1),
\end{aligned}
$$

These decompositions were fed as an input to RepCert. The results and runtimes for these are shown in Tabs. 5. As in the case of Tab. 3, the precision for the invariance certificate is set to $\epsilon = 10^{-8}$ in these tests. For certain high-dimensional blocks, the algorithm fails to certify invariance at such a high precision. Analogously to the previous tests on the groups G_{ab}, here we run the certification algorithm with precision $\epsilon = 10^{-7}$ on these high-dimensional blocks. The results obtained are shown in Tab. 5.

(a,b,c)	$\dim \rho_{abc}$	Irr. d	Inv. Time (s)	Restr. Time (s)	Irr. Time (s)	Inv.	Irr.
(2,2,3)	64	1	0.025	0.007	1.668	Yes	Yes
(2,2,3)	64	1	0.010	0.003	1.430	Yes	Yes
(2,2,3)	64	3	0.017	0.041	0.227	Yes	Yes
(2,2,3)	64	3	0.026	0.042	0.225	Yes	Yes
(2,2,3)	64	6	0.018	0.282	1.087	Yes	Yes
(2,2,3)	64	6	0.017	0.198	1.084	Yes	Yes
(2,2,3)	64	8	0.017	0.491	1.988	Yes	Yes

(2,2,3)	64	12	0.017	1.296	5.220	Yes	Yes
(2,2,3)	64	12	0.017	1.209	5.231	Yes	Yes
(2,2,3)	64	12	0.018	1.312	5.256	Yes	Yes
(2,2,4)	256	1	0.384	0.024	1.668	Yes	Yes
(2,2,4)	256	1	0.128	0.008	2.384	Yes	Yes
(2,2,4)	256	4	0.383	0.152	0.407	Yes	Yes
(2,2,4)	256	4	0.382	0.234	0.416	Yes	Yes
(2,2,4)	256	6	0.381	0.333	1.142	Yes	Yes
(2,2,4)	256	8	0.382	0.638	1.964	Yes	Yes
(2,2,4)	256	8	0.380	0.708	1.997	Yes	Yes
(2,2,4)	256	16	0.381	3.704	9.830	Yes	Yes
(2,2,4)	256	24	0.380	11.61	26.30	Yes	Yes
(2,2,4)	256	24	0.379	11.45	26.13	Yes	Yes
(2,2,4)	256	24	0.379	11.08	26.54	Yes	Yes
(2,2,4)	256	24	0.385	11.49	26.36	Yes	Yes
(2,2,4)	256	32	0.385	25.59	49.49	Yes	Yes
(2,2,4)	256	32	0.380	25.97	49.17	Yes	Yes
(2,2,4)	256	48	0.381	85.49	142.1	Yes	Yes
(2,2,5)	1024	1	20.60	0.459	2.861	Yes	Yes
(2,2,5)	1024	1	5.866	0.074	2.145	Yes	Yes
(2,2,5)	1024	5	20.26	2.565	0.825	Yes	Yes
(2,2,5)	1024	5	20.58	2.635	0.836	Yes	Yes
(2,2,5)	1024	10	20.50	6.227	3.931	Yes	Yes
(2,2,5)	1024	10	20.70	6.317	3.889	Yes	Yes
(2,2,5)	1024	10	20.51	6.302	3.905	Yes	Yes
(2,2,5)	1024	10	20.71	6.408	3.907	Yes	Yes
(2,2,5)	1024	32	20.46	68.52	51.56	Yes	Yes
(2,2,5)	1024	40	20.47	119.7	97.26	Yes	Yes
(2,2,5)	1024	40	20.59	116.5	97.77	Yes	Yes
(2,2,5)	1024	40	20.70	117.5	97.68	Yes	Yes
(2,2,5)	1024	40	20.41	118.2	97.35	Yes	Yes
(2,2,5)	1024	60	20.62	337.1	263.5	Yes	Yes
(2,2,5)	1024	80	20.66	782.5	665.1	Yes	Yes
(2,2,5)	1024	80	20.57	779.2	654.8	Yes	Yes
(2,2,5)	1024	80	20.62	784.1	660.0	Yes	Yes
(2,2,5)	1024	80	20.65	785.3	656.7	Yes	Yes
(2,2,5)	1024	120	20.58	-	-	No	-
(2,2,5)	1024	120	20.48	-	-	No	-
(2,3,3)	216	1	0.078	0.007	1.907	Yes	Yes

(2,3,3)	216	6	0.229	0.315	1.132	Yes	Yes
(2,3,3)	216	8	0.231	0.613	2.076	Yes	Yes
(2,3,3)	216	9	0.231	0.809	3.095	Yes	Yes
(2,3,3)	216	12	0.230	1.761	5.542	Yes	Yes
(2,3,3)	216	27	0.228	16.42	34.92	Yes	Yes
(2,3,3)	216	27	0.230	16.54	34.75	Yes	Yes
(2,3,3)	216	36	0.230	37.08	73.78	Yes	Yes
(2,3,3)	216	36	0.231	36.58	73.12	Yes	Yes
(2,3,3)	216	54	0.313	120.8	200.2	Yes	Yes
(3,3,3)	729	1	2.144	0.039	2.384	Yes	Yes
(3,3,3)	729	6	7.533	1.536	1.173	Yes	Yes
(3,3,3)	729	8	7.554	2.278	2.073	Yes	Yes
(3,3,3)	729	12	7.533	4.662	5.596	Yes	Yes
(3,3,3)	729	18	7.527	10.77	14.91	Yes	Yes
(3,3,3)	729	72	7.611	456.5	487.6	Yes	Yes
(3,3,3)	729	72	7.593	462.3	484.0	Yes	Yes
(3,3,3)	729	108	7.505	1401.	1664.	Yes	Yes

Table 5: Benchmark for the representations ρ_{abc} of G_{abc} for several such groups. The parameter d is the dimension of the space being certified for invariance and irreducibility, *Restr. Time* is the time necessary to restrict the sampled group elements to the corresponding block (i.e. using the notation above, the map $g \mapsto g_i$ for the i-th block). The last two columns specify whether invariance and irreducibility were certified for the representation. Only blocks with $d < 150$ were attempted to be certified due to the long runtime expected for higher dimensions.

(a, b, c)	$\dim \rho_{abc}$	Irr. d	Inv. Time (s)	Restr. Time (s)	Irr. Time (s)	Inv.	Irr.
(2,2,5)	1024	120	20.76	2348.	2387.	Yes	Yes
(2,2,5)	1024	120	20.57	2351.	2379.	Yes	Yes

Table 6: Certification of the two blocks which failed the invariance test in Tab. 5. These were certified using $\epsilon = 10^{-7}$, as opposed to the lower value of $\epsilon = 10^{-8}$ used in Tab. 5. The other parameters are the same as those used in that table.

Conclusions, outlook and open questions

Summary

Representation theory has a long and rich history of finding applications within the physical sciences. Throughout my thesis, I have worked on expanding this palette by developing representation theoretical tools for quantum information theory. My focus has been on two research directions: *1.* exploring the representation theory arising from the stabilizer formalism, together with its relation to the Theta duality and to t-designs, and *2.* proposing efficient algorithms for the numerical decomposition of representations.

Clifford and oscillator tensor powers. The Clifford group, and the closely-related oscillator representation of the symplectic group, play a prominent role both in quantum information science as well as outside it, in fields such as convex reconstruction, automorphic forms and classical coding theory. In particular, tensor powers representations of the form,

$$\mathrm{Cl}^{\otimes t} : U \mapsto U^{\otimes t}, \quad U \in \mathrm{Cl} \tag{28}$$

$$\mu^{\otimes t} : S \mapsto \mu^{\otimes t}(S), \quad S \in \mathrm{Sp}(\mathbb{Z}_d^{2n}), \tag{29}$$

have attracted the attention of these communities. For instance, in [GH17], oscillator tensor powers are studied in order to estimate the *character ratios* of the symplectic group, that is, expressions of the form

$$\sum_{\rho \in \mathrm{Irr}\ \mathrm{Sp}(\mathbb{Z}_d^{2n})} \frac{\chi_\rho(S)}{\dim \rho}, \quad S \in \mathrm{Sp}(\mathbb{Z}_d^{2n}),$$

where χ_ρ is the character of the irrep ρ. These ratios are important in a variety of applications of harmonic analysis [GH20], not least of which being their role in the characterization random walks on finite groups [DS81]. From a physics perspective, detailed understanding of tensor powers with low order [ZKGG16, Zhu17, Web16] has found a wealth of applications already. A prominent example, Ref. [BBC$^+$19], provides an algorithm for classically simulating quantum computing. It uses information about $t = 5$ tensor powers to bound the *stabilizer rank* of magic states—the quantity on which their algorithm's runtime depends. It can be expected, furthermore, that a detailed understanding of higher-order tensor power representations can lead to further improvements in this regard.

The list of applications in this regard goes on, from quantum device characterization [RKK$^+$18, KR21], to coding theory [NRS06] and matrix recovery [KZG16b]. The rich research that has emanated from studying tensor power representations with

$t \leq 5$ is a strong motivation for studying higher-order tensor power representations. In Chap. 2 I have presented my contribution to this subject.

Higher order tensor powers have been studied from three perspectives in the literature. In [GNW21], the *commutant* of Clifford tensor powers is studied. There, a basis for this commutant is specified, with basis elements $R(T)$ being labeled by certain subspaces $T \subset \mathbb{Z}_d^t \times \mathbb{Z}_d^t$,

$$R(T) = \left(\sum_{(x,y)\in T} |x\rangle\langle y| \right)^{\otimes n}.$$

These operators furthermore form a semi-group. In [NRS06, NRS01] the *invariant polynomials* of the Clifford group are studied. In this regard, they find that the space of homogenous invariants of degree (t, t) is spanned by polynomials p_T defined as

$$p_T(\mathbf{x}) = (\mathbf{x}^{\otimes t})^\dagger R(T) \mathbf{x}^{\otimes t},$$

where $\mathbf{x} = (x_1, \ldots, x_{d^n}) \in \mathbb{C}^{d^n}$. Finally Refs. [GH17, GH20] use the notion of *rank* introduced in [How10] to extend the well-known theory of the Theta correspondence [How89a, KV78] to the case of finite dual pairs of groups. Specifically, the authors find a subspace of $(\mathbb{C}^{d^n})^{\otimes t}$ – the subspace of *maximal rank* representations – in which there is a pairing between the irreps of symplectic-orthogonal dual pair. More specifically, the action of $\mu^{\otimes t}$ on this subspace decomposes as

$$\mu^{\otimes t}|_{\text{max. rank subspace}} \simeq \bigoplus_{\tau \in \text{Irr } \mathrm{O}(\mathbb{Z}_d^t)} \tau \otimes \eta(\tau),$$

where $\eta : \text{Irr } \mathrm{O}(\mathbb{Z}_d^t) \to \text{Irr } \mathrm{Sp}(\mathbb{Z}_d^{2n})$ is an injective function. This η correspondence can be seen as a generalization of the Θ duality between $\mathrm{O}(\mathbb{R}^t)$ and $\mathrm{Sp}(\mathbb{R}^{2n})$.

Chap. 2 generalizes the formalism of the η correspondence in two regards. The first section, published as [MMG21a], shows that this formalism can be used to decompose the *full* representation $\mu^{\otimes t}$ rather than just the maximal rank component. A key feature here is that lower-rank sectors correspond to the CSS code spaces introduced in [GNW21]. Moreover, there is a certain "self-similarity" between different rank layers: the aforementioned CSS codespaces are themselves isomorphic to lower tensor power representations, $\mu^{\otimes k}$ with $k < t$. The second section, the manuscript [MMG21b], extends this formalism in order to decompose *Clifford* tensor power representations. Importantly, this generalization covers the qubit case ($d = 2$) which was not addressed by the original formulation of the η duality.

As a sample application of these results, Chap. 2 shows that the problem of complex-conjugating a black box Clifford evolutions is vastly simpler than the more general case

involving arbitrary black box unitary evolutions. If one is given a black box under the promise that it implements some unitary evolution $U \in \mathrm{U}(d^n)$, then one requires at least $d^n - 1$ queries of the black box in order to implement $\bar{U}$ [QDS$^+$19, MSM19]. This means that performing $\bar{U}$ is not much simpler than performing full tomography on U (which would require $\gtrsim d^{2n}$ samples). On the other hand, we show that if the promise is strengthened by guaranteeing that $U \in \mathrm{Cl}$, then $\sim d$ queries are sufficent.

Approximate unitary t-designs. Chap. 3 gives an exciting application of the representation theory of Clifford tensor powers. There, an efficient construction of *approximate unitary t-designs* is proposed. Recall that an approximate nitary t-design is a probability distribution p on $\mathrm{U}(2^n)$ for which

$$\left\| \mathbb{E}_{U\sim p}[U^{\otimes(t,t)}] - \mathbb{E}_{U\sim \mathrm{Haar}}[U^{\otimes(t,t)}] \right\|_\diamond \leq \epsilon,$$

where $U^{\otimes(t,t)} = U^{\otimes t} \otimes \bar{U}^{\otimes t}$ and $\| \cdot \|_\diamond$ is the *diamond norm*,

$$\|A\|_\diamond = \sup_{\rho \in \mathrm{End}\ \mathbb{C}^{2^n} \otimes \mathrm{End}\ \mathbb{C}^{2^n}} \frac{\|(A \otimes \mathbb{1}_{2^{2n}})(\rho)\|_1}{\|\rho\|_1}, \quad A \in \mathrm{End}\ \mathrm{End}\ \mathbb{C}^{2^n}.$$

Exact unitary t-designs have $\epsilon = 0$.

Designs and approximate designs appear in a variety of fields: they are primitives and quantum cryptography and quantum Shannon theory, they have found a variety of uses in quantum and classical estimation problems, and they even serve as models for quantum chaos and certain high energy physics phenomena. For $t \leq 3$, the Clifford group is the unitary design "par excellence" in quantum information theory [Zhu17, Web16]. Additionally, while the Clifford group fails to be a 4-design, it is sufficiently similar to a 4-design for many applications [ZKGG16, KZG16b, KZG16a].

Generally speaking, explicitly constructing exact higher order unitary designs is not a simple task. One rather richly structured class of exact designs are those for which p is the flat distribution on a finite group. These *unitary t-groups* are rather uncommon though: there are only finitely many instances of unitary 4-groups, for example [BNRT20]. Even more, Ref. [BNRT20] singles out the qubit Clifford group as the *only* infinite family of finite 3-groups.

In contrast, *approximate* unitary t-designs can be constructed in a straightforward way for any t [BHH16]. In that seminal paper, it is proven that local random quantum circuits converge to approximate t-designs in depth $k = O(t^{10} n^2 \log(1/\epsilon))$. Conceptually, this result is beautiful: Most Haar random unitaries require exponentially deep local quantum circuits to be approximated. Thus, while Haar randomness is often used as a model for quantum chaos or as a tool for several quantum information processing tasks, it is unphysical. On the other hand, the randomness provided by approximate designs is the outcome of polynomially deep quantum circuits and, in this sense, is

physical.

The construction in [BHH16] is, however, not too practical if one wishes to sample from a unitary t-design using a quantum computer. Specifically, this construction would require the implementation of $O(n^2)$ arbitrary 2-qubit gates—this is beyond the capabilities of most, if not all, current quantum devices.

Chap. 3 presents a construction of unitary t-designs out of quantum circuits whose overwhelming majority of local gates are Clifford gates. Local Clifford gates are typically the first gates that can be performed to high fidelity on quantum devices. Furthermore Clifford gates have an easy fault-tolerant implementation and Clifford dominated circuits may be efficiently simulated on a classical computer [BBC$^+$19]. This makes our construction highly relevant for near-term quantum computers. This result, moreover, answers the question raised in [ZKGG16] of whether the Clifford group may be used to generate high-order unitary designs.

Consider *k-interleaved random Clifford circuits*, namely, circuits of the form

$$U_1 K U_2 K \dots U_k K,$$

where $U_i \in \mathrm{Cl}$ are uniformly random and K is a fixed single qubit unitary (acting, say, on the first qubit). The main theorem in Chap. 3 states that if $k = \Omega(t^4 \log^2 t \log(1/\epsilon))$ and $n = \Omega(t^2)$, then the distribution of k-interleaved random Clifford circuits are an ϵ-approximate t-design. In the case that K is sampled uniformly from $\mathrm{U}(2)$, explicit constants in the scalings for k and n are found.

Numerical decomposition of representations. There are instances in life where one must find an explicit decomposition of a given representation—that is, one must numerically provide a set of projectors onto the irreducible components of a numerically defined representaiton. This happens, for example, in the context of symmetrizing semi-definite programs (SDPs). Recall that an SDP is an optimization problem of the form

$$\max_{X \in \mathbb{C}n \times n} \mathrm{tr}(X A_0) \quad \text{s.t.} \quad X \geq 0,\ \mathrm{tr}(X A_i) = a_i,\ i = 1, \dots, k, \tag{30}$$

where $A_i \in \mathbb{C}^{n \times n}$ are Hermitian matrices and $a_i \in \mathbb{R}$.

SDPs are a widespread tool in science and engineering. Quantum information theory is not an exception here, with SDPs appearing on a range of applications from the simulation of quantum computing, to the study of quantum correlations. As is usual for quantum systems, SDPs arising here typically suffer from the curse of dimensionality, leading many interesting problem instances to be hard to directly solve. All hope is not lost, however: many of these SDPs are highly symmetric problems whose dimension can thus be considerably reduced. An extreme example is the linear program (LP) used

to compute the *robustness of magic* (a measure of distance to the polytope of stabilizer states). Ref. [HG19] uses symmetry to exponentially reduce the dimension of the LP, from $4^N - 1$ to N (where N is the number of qubits).

Reducing the dimension of symmetric SDPs is nothing new [Val09]. In the most general case, one may aim to directly block-diagonalize the algebra $\mathcal{A} = \langle \{A_i\}_i \rangle$— because this algebra is semi-simple, the Artin-Wedderburn theorem may be used for this.. This allows one to restrict the SDP to positive semidefinite matrices *with the same block structure* as $\mathcal{A}$. Algorithms for this exist [MM11, CL20, MKKK10, MM10, dKDP11, AMB04, CSX15, CCS19, BFS93], however their runtime scales rather steeply with n which limits their applicability.

One would hope that considering more structured symmetries might increase the efficiency of the block diagonalizing algorithms. In the context of quantum information, many interesting SDPs have a symmetry described by a group representation. In order to exploit such symmetries, it is necessary to explicitly decompose the corresponding representaiton.

Consider an SDP, as in eq. (30), which has the following symmetry under a subgroup $G \subset \mathrm{U}(n)$,

$$gA_ig^\dagger = A_i, \quad i = 0, 1, \ldots, k,\ g \in G.$$

Then, if $U \in \mathrm{U}(n)$ block diagonalizes G as $UgU^\dagger = \oplus_{j=1}^k \rho_j(g) \otimes \mathbb{C}^{m_j}$, it is such that

$$UA_iU^\dagger = \bigoplus_j \mathbb{1}_{\rho_j} \otimes A_{ij}.$$

This way, one may replace the optimization (30) with the following series of optimizations for each j

$$\max_{X_j \in \mathbb{C}^{m_j \times m_j}} \operatorname{tr}(X_j A_{0j}) \quad \text{s.t.} \quad \operatorname{tr}(X_j A_{ij}) = a_i,\ X_j \succeq 0, \tag{31}$$

$$A_{ij} = A_{ij}^\dagger,\ \forall\, i, \tag{32}$$

so that if X^* optimizes (30) and X_j^* optimizes (31), it holds that

$$\operatorname{tr}(X^* A_0) = \sum_{j=1}^k (\dim \rho_j) \operatorname{tr}\big(X_j^* A_{0j}\big).$$

Algorithms for this problem – that of obtaining the block-diagonalizing transformation U out of G – is, of course, itself nothing new either. When G is finite, one may in principle exactly decompose the representation over the *cyclotomic field* $\mathbb{Q}[\omega_{|G|}]$,

where

$$\omega_{|G|} = \exp\left(\frac{2i\pi}{|G|}\right).$$

This is the standard approach used by, e.g., the software suite GAP [GAP21]. While this approach produces high-quality results, it is not suitable for large finite groups nor continuous groups.

A second approach is the Dixon algorithm [Dix70]. This method is based on the idea that generic elements of the commutant of a reducible group G are not multiples of the identity (and thus contain some eigenvalue gap). This method may be adapted to accomodate for numerical errors in the specification of group elements $g \in G$ [BF91]. Its runtime, $O(\sum_i m_i n^5)$, is still too high for many interesting applications.

In Part II of this thesis, a novel algorithm for this task is discussed.

The method is divided into two tasks: It first uses a fast heuristic to propose a decomposition of G [RMMB19], and after this it runs a certification algorithm to probe the accuracy of the claimed decomposition [MMRBG21].

The first step, which I summarized in Chap. 5, runs in time $O(n^3)$. Denis Rosset and Jean-Daniel Bancal, with whom I collaborated on this research line, have coded the software suite RepLAB [RB18] based on the aforementioned method. This method has already been used on the SDPs arising from Bell experiments [TFR$^+$21], quantum contextuality experiments [TCUA20], and quantum communication scenarios [FST21].

Anecdotally, RepLAB has been observed to work well. More systematically, in Chap. 5 I have provided evidence that RepLAB's results are expected to be accurate. The evidence is not a rigorous proof of correctness, but rather relies on several simplifying assumptions. These calculations served as a "sanity check," i.e. they indicated that the working principle of RepLAB is likely robust enough to be useful in practice.

The algorithm for the certification step was proposed in [MMRBG21], which is Chap. 6 in this thesis. If d is the maximal dimension of a block in the decomposition obtained with RepLAB and D is the complexity of multiplying group elements together, the runtime of this certifying algorithm is $O(n^3 \log n + Dd^2 \log d)$. The two terms in this runtime come from the two subroutines in the algorithm, *invariance certification* and *irreducibility certification*, whose functions are suggested by their names. If a projector P is certified by this algorithm, then – barring the unlikely event of a false positive – we may conclude that there exists a projector P_0 onto an irreducible G-subrepresentation such that $\|P - P_0\|_F \leq \epsilon$.

I coded the certification algorithm and it is available at [MM21]. I discuss several key features in the code and benchmarks on its runtime in Chap. 7.

Outlook and open questions

Through the projects I have participated in, I have been able to answer several interesting questions. Equally exciting, however, is the doors they open for future work. Here I comment on some of these new possibilities.

Explicit descriptions of special blocks in the η correspondence. A very useful identity in the study of Clifford tensor powers is that the subspace of $\mathcal{H}_{n,t} := ((\mathbb{C}^d)^{\otimes n})^{\otimes t}$ invariant under the orthogonal stochastic group is spanned by stabilizer tensor powers:

$$\mathcal{H}_{n,t}^{\mathrm{St}} := \mathrm{range}\left(\sum_{O\in \mathrm{St}(\mathbb{Z}_d^t)} R(O)\right) = \mathrm{span}\{|\psi\rangle^{\otimes t} \mid |\psi\rangle \in \mathrm{STAB}_n\}.$$

Given the use that this identity has found in e.g. [GNW21, HMMH$^+$20], it is natural to ask whether other representation spaces have similarly explicit characterizations.

One possibility is to further characterize the irreducible blocks in $\mathcal{H}_{n,t}^{\mathrm{St}}$. By eq. (36) in [MMG21b], whenever t is not a multiple of d this subspace decomposes as

$$\mathcal{H}_{n,t}^{\mathrm{St}} \simeq \bigoplus_r \eta(1_{\mathrm{St}^r}), \tag{33}$$

where $\mathrm{St}^r \subseteq \mathrm{St}$ is the stabilizer of an r-dimensional isotropic stochastic subspace in $\mathbb{Z}_d^t$. This decomposition must be "skewed" with respect to stabilizer tensor powers: indeed stabilizer tensor powers form a single Cl-orbit, and so no such state is contained in any invariant subspace $\eta(1_{\mathrm{St}^r})$. With decomposition (33) in mind, it is natural to ask for an explicit characterization of each irreducible subrepresentation of $\mathcal{H}_{n,t}^{\mathrm{St}}$.

Consider the highest rank subspace, $\eta(1_{\mathrm{St}})$. This block is equivalently characterized as

$$\mathbb{C}\{U\,|\text{max. rank}\rangle \mid U \in \mathrm{Cl}\},$$

where

$$|\text{max. rank}\rangle = \sum_{F\in\mathbb{Z}_d^{t\times n},\ \mathrm{rank}\, F=t} |F\rangle\,.$$

On this state, the subgroup $\mathrm{Gl}(\mathbb{Z}_d^n) \subset \mathrm{Cl}$ generated by CADD gates acts trivially. Can this description of the block be useful? Additionally, is $\mathrm{Gl}(\mathbb{Z}_d^n) \subset \mathrm{Cl}$ the maximal subgroup of Cl acting trivially on $|\text{max. rank}\rangle$? If the latter question were answered

positively, then the following identity would hold,

$$\eta(1_{\mathrm{St}}) \simeq \mathrm{Ind}_{\mathrm{Gl}_n}^{\mathrm{Cl}}(1_{\mathrm{Gl}_n}).$$

Explicit full decomposition of $\mathrm{Cl}^{\otimes t}$ for fixed t. As has been discussed, Clifford tensor powers of degree up to $t = 4$ are rather well understood and this understanding has found several applications. A natural next step would be to find explicit decompositions of other fixed values of t, for example $t = 5$.

In this case, $\mathbb{Z}_2^5$ contains only five non-trivial isotropic stochastic subspaces, namely

$$\begin{aligned} N_1 :=& \langle (01111) \rangle, \quad N_2 := \langle (10111) \rangle, \\ N_3 :=& \langle (11011) \rangle, \quad N_4 := \langle (11101) \rangle, \\ N_5 :=& \langle (11110) \rangle. \end{aligned}$$

These are all one dimensional and have a trivial intersection. By the proof of Lem. V.6 in [MMG21b], every subrepresentation of $\Delta_{5,0}$ with rank < 5 is contained in $\mathrm{span}\{C_{N_i}\}_i$. Then,

$$\Delta_{5,0} \simeq \left(\bigoplus_{\tau \in \mathrm{Irr\ St}(\mathbb{Z}_2^5)} \tau \otimes \eta(\tau) \right) \oplus C_{N_1} \oplus \cdots \oplus C_{N_5}.$$

By Lem. III.4 in [MMG21b], $C_{N_i} \simeq \Delta_{0,3}$ for every i. Given that Cl is a unitary 3-design, $\Delta_{0,3}$ cab be explicitly decomposed using Schur-Weyl duality. The outstanding question is to find explicit expressions for the subrepresentations of rank 5.

Estimating non-stabilizerness. A prominent application of Schur-Weyl duality in quantum information is for the problem of *spectrum estimation*, i.e. given t copies of a state ρ, estimate its spectrum. For t large enough, one can of course simply perform individual measurements on each copy of the state, reconstruct it, and obtain the spectrum. A considerable advantage, however, can be obtained by allowing measurements on $\rho^{\otimes t}$ which do not factorize across the copies. Specifically, Keyl and Werner [KW05] use a projective measurement arising from Schur-Weyl duality for this. Schur-Weyl duality looks at two commuting actions on $\mathcal{H}_{n,t}$: the t-th tensor power representation of $\mathrm{U}(\mathcal{H}_n)$ and the representation of S_t which permutes tensor factors. As a $\mathrm{U}(\mathcal{H}_n) \times S_t$ representation,

$$\mathcal{H}_{n,t} \simeq \bigoplus_{\lambda} U_\lambda \otimes S_\lambda, \tag{34}$$

where λ is a partition of t into at most d^n parts, $U_\lambda \in \mathrm{Irr\ U}(\mathcal{H}_n)$ and $S_\lambda \in \mathrm{Irr}\ S_t$. The Keyl-Werner approach proposes the measurement $\{P_\lambda\}_\lambda$ on $\rho^{\otimes t}$, where P_λ is the

orthogonal projector onto the λ-th term in (34). The spectrum is estimated by λ/d^n (possibly padded with zeros).

This is exciting: suddenly the *Young diagram* λ – the abstract label classifying the isotypes in (34) – gains a *physical* meaning. Moreover, building on this result one may obtain a sample-optimal algorithm for quantum tomography [HHJ$^+$17].

Now, coming back to the topics covered in this thesis, I have provided a decomposition of a different pair of dual actions on $\mathcal{H}_{n,t}$—the actions of Cl and $\mathrm{St}(T)$. It would be interesting to see which quantum estimation problems could be approached using this decomposition. Particular suspects for such a problem are estimating the several measures of non-stabilizerness (e.g. stabilizer rank, extent or fidelity [BBC$^+$19]).

Consider for example the problem of estimating the stabilizer fidelity of a state,

$$F(|\psi\rangle) := \max_{|s\rangle\in\mathrm{Stabs}_n} |\langle s|\psi\rangle|^2,$$

given t copies of the state. A bound may at least be obtained by estimating the expected value

$$\left\langle\psi^{\otimes t}\right| P_{\mathrm{Stabs}} \left|\psi^{\otimes t}\right\rangle,$$

where P_{Stabs} projects onto the space spanned by stabilizer tensor powers, or, equivalently, the trivial $\mathrm{St}(T)$ isotype in $\mathcal{H}_{n,t}$. This expected value upper bounds the quantity of interest as

$$\begin{aligned} F_t\left(\left|\psi^{\otimes t}\right\rangle\right) :&= \max_{|s\rangle\in\mathrm{Stabs}_n} |\left\langle s^{\otimes t}\middle|\psi^{\otimes t}\right\rangle|^2 \\ &= F(|\psi\rangle)^t \leq \left\langle\psi^{\otimes t}\right| P_{\mathrm{Stabs}} \left|\psi^{\otimes t}\right\rangle. \end{aligned}$$

The problem with this approach is that it is currently not known how tight this bound is.

Now suppose that one implements some projective measurement $\{P_k\}$ on $|\psi^{\otimes t}\rangle$ in order to estimate $F(|\psi\rangle)$. Then,

$$F_t\left(\left|\psi^{\otimes t}\right\rangle\right) = F\left(R(O)U^{\otimes t}\left|\psi^{\otimes t}\right\rangle\right), \quad \forall\, U\in\mathrm{Cl},\ O\in\mathrm{St}(T),$$

so that the measurement operators are subject to

$$\left\langle\psi^{\otimes t}\right| U^{\otimes t,\,\dagger} R^\dagger(O) P_k R(O) U^{\otimes t} \left|\psi^{\otimes t}\right\rangle = \left\langle\psi^{\otimes t}\right| P_i \left|\psi^{\otimes t}\right\rangle, \quad \forall\ |\psi\rangle\in\mathcal{H}_n.$$

It is natural to simply require the operators P_k to commute with $U^{\otimes t}$ and $R(O)$.

By [MMG21b, Thm. V.2],

$$P_k = \bigoplus_i \bigoplus_{\tau \in \mathrm{Irr}\ \mathrm{St}(T_i)} P_k^{(T_i,\tau)} \otimes \mathbb{1}_{\eta(\tau)},$$

where $P_k^{(T_i,\tau)}$ commutes with $\mathrm{Ind}_{\mathrm{St}(T)^{N_i}}^{\mathrm{St}(T)}(\tau)$.

At this point one can ask two questions. First, could the simple choice of $P_k^{(T_i,\tau)} = \delta_{k,i}\mathbb{1}_{\mathrm{Ind}(\tau)}$ be useful for the estimation of the stabilizer fidelity? Second, can one provide a physical meaning to the symbols τ classifying $\mathrm{St}(T)$ irreps, in the same way that Keyl and Werner provide a physical meaning to Young diagrams?

Duality arising from real Clifford group. An important take-home message from the results in [GH17, GH20] is that, while the Theta correspondence fails to hold exactly over finite fields, it does hold on a fairly large portion of Hilbert space. Specifically it holds over the span $\mu_{\text{max. rk}}^{\otimes t}$ of all maximal rank subrepresentations in $\mu^{\otimes t}$. A short calculation using [MMG21a, Thm. 1.2] shows that, for any given t, this subspace accounts for most of Hilbert space as n grows,

$$\frac{\dim \mu_{\text{max. rk}}^{\otimes t}}{\dim \mu^{\otimes t} - \dim \mu_{\text{max. rk}}^{\otimes t}} = O(\exp(-n)).$$

This property, of the η correspondence "covering" most of Hilbert space, was already noted in [GH17] using different techniques.

Now, consider a qubit Clifford tensor power representation $\Delta_{t,0}$ with $t = 0 \mod 4$. An object of interest in this case is the Pauli-trivial subrepresentation,

$$\mathrm{C}_{\mathbf{1}_t} = \mathrm{range} \sum_{p \in \mathcal{P}} P^{\otimes t}.$$

This representation, for example, plays a prominent role in Ref. [ZKGG16] which looks at the case $d = 2,\ t = 4$.

In [MMG21b, Sec. VI] it is shown that the *real* Clifford group RCl acts on $\mathrm{C}_{\mathbf{1}_t}$ as a permutation representation of

$$\mathrm{RCl}/\mathrm{R}\mathcal{P} \simeq \mathrm{O}(\mathbb{H}^n),$$

where $\mathrm{R}\mathcal{P}$ is the real Pauli group and $\mathbb{H}$ is a hyperbolic plane. Notice the following coincidence: when studying the commutant of $\mu^{\otimes t}$ for odd d, a permutation representation R of the orthogonal group $\mathrm{O}(\mathbb{Z}_d^t)$ arises rather naturally. I show that this coincidence is not vain: in [MMG21b, Lem. VI.1], I obtain an action $\tilde{\Delta}$ of a certain symplectic group $\mathrm{Sp}(\mathbb{Z}_2^{t-2})$ which commutes with the real Clifford action $\mathrm{Ind}_{\mathrm{RCl}}(\Delta_{t,0})|_{\mathrm{C}_{\mathbf{1}_t}}$.

Thus,

$$\mathrm{C}_{\mathbf{1}_t} \simeq \bigoplus_{\tau \in \mathrm{Irr\ O}(\mathbb{H}^n)} \tau \otimes \Theta(\tau),$$

where $\Theta(\tau)$ is a possibly reducible $\mathrm{Sp}(\mathbb{Z}_2^{t-2})$ representation. The question is then: can one find an exact duality on a large subspace of $\mathrm{C}_{\mathbf{1}_t}$? This is the content of Conjecture VI.1 in [MMG21b]: that there exists a subspace $\mathcal{L} \subset \mathrm{C}_{\mathbf{1}_t}$ such that

$$\frac{\dim \mathcal{L}}{\dim \mathrm{C}_{\mathbf{1}_t} - \dim \mathcal{L}} = O(\exp(-n)),$$

and such that there exists an injective function $\eta : \mathrm{Irr\ O}(\mathbb{H}^n) \to \mathrm{Irr\ Sp}(\mathbb{Z}_2^{t-2})$ for which

$$\mathcal{L} \simeq \bigoplus_{\tau \in \mathrm{Irr\ O}(\mathbb{H}^n)} \tau \otimes \eta(\tau).$$

A positive answer to this question would likely involve a further generalization of the η correspondence formalism, which might then be used to better understand the representation theory of RCl.

Clifford representation theory: whereto next? I am very fond of the results found in Chap. 2. These extend the η correspondence formalism to the Clifford group, which could *in principle* be used to provide a better understanding of the representation theory of Cl. The catch is the words "*in principle*:" at the moment we do not have an explicit-enough grasp of the $\mathrm{St}(T)$-Cl duality in order to realize this possibility.

This situation can be contrasted to Schur-Weyl duality. There, the understanding of the representaiton theory of the symmetric group can be leveraged to describe in great detail the tensor power representations of the unitary group. These details can be, for example, useful in quantum tomography [HHJ$^+$17] through the use of the Keyl-Werner spectrum estiamtion procedure [KW05].

Analogously, understanding $\mathrm{Irr\ St}(T)$ could lead to handier results on the decomposition of $\Delta_{t,0}$.

A simple case in which to start working d =odd and $t \neq 0 \mod d$, in which case $\mathrm{St}(T) \simeq \mathrm{O}(\mathbb{Z}_d^{t-1})$ is an orthogonal group. The representation theory of finite orthogonal groups has been widely studied in mathematics through Deligne-Lusztig theory (see e.g. [Gec17, DM20] or [Car85, Chap. 7]). An important tool to study these is the "philosophy of cusp forms" [Bum04, Chap. 47].

The η correspondence has been introduced as an approach to the representation theory of linear algebraic groups over finite fields which is complementary to the philosophy of cusp forms and, more generally, Deligne-Lusztig theory [GH20]. However,

recent work combining the insights of both approaches has lead to some progress, for example in extending the η correspondence beyond the "stable regime" $t \leq n$ [Pan20]. With this motivation, it seems plausible that combining these two formalisms can continue to shed light on the Clifford representation theory. This paints a rough picture for a future research direction: *leverage the results from Deligne-Lusztig theory to obtain more explicit information about* $\Delta_{t,0}$.

A first place in which this research direction can be made concrete is the following. The representations of the form

$$\mathrm{Ind}_{\mathrm{St}(T)^{N_i}}^{\mathrm{St}(T)}(\tau), \quad \tau \in \mathrm{Irr}\ \mathrm{St}(T_i),$$

which appear quite naturally in the context of Clifford tensor powers, are a central object in the philosophy of cusp forms. There, they are known as *parabolic inductions* and are used as the "building blocks" that generate $\mathrm{Irr}\ \mathrm{St}(T)$. A better understanding of these representations would, for example, to a better understanding of the $\mathrm{St}(T) \times \mathrm{Cl}$ decomposition of $\mathcal{H}_{n,t}$

Namely, consider the decomposition

$$\mathrm{Ind}_{\mathrm{St}(T)^{N_i}}^{\mathrm{St}(T)}(\tau) \simeq \bigoplus_{\tau' \in \mathrm{Irr}\ \mathrm{St}(T)} \mathbb{C}^{m(\tau,\tau')} \otimes \tau',$$

which entails

$$\mathcal{H}_{n,t} \simeq \bigoplus_{i} \bigoplus_{\substack{\tau \in \mathrm{Irr}\ \mathrm{St}(T_i) \\ \tau' \in \mathrm{Irr}\ \mathrm{St}(T)}} \mathbb{C}^{m(\tau,\tau')} \otimes \tau' \otimes \eta(\tau).$$

One possibly interesting question is *when is* $m(\tau, \tau') = 1$*?* Or, in a similar note, *when is* $(\mathrm{St}(T)^{N_i}, \mathrm{St}(T))$ *a Gelfand pair?*[3] If it were the case that $m(\tau, \tau') \in \{0, 1\}$ for all τ and τ', then any $\mathrm{Cl} \times \mathrm{St}(T)$-symmetric measurement would essentially be projective. Indeed, any such measurement would be a coarse-graining of the measurement $\{P_{\tau,\tau'}\}$, where the operator $P_{\tau,\tau'}$ projects onto the $\tau' \otimes \eta(\tau)$ component. This would, rather elegantly, provide a physical meaning to the pairs of symbols (τ, τ').

Unifying the codes RepLAB and RepCert. The projects comprising Part II of this thesis have been mostly self-contained solutions to practical issues which were encountered while developing RepLAB [RB18]. As opposed to Part I, they aim for concreteness rather than generality. In this way, these projects leave little room for open questions to in which to direct further research. That said, there remains one practical aspect that should be addressed.

The code of RepCert was developed independently of RepLAB—they are two in-

[3]For a definition and discussion on Gelfand pairs, see [Chap. 45][Bum04].

dependent packages which are even written in different languages, RepCert being in Python and RepLAB being in MatLAB. As a first stage of algorithm development this makes sense: It eases the use of data structures specifically designed for RepCert's needs, without having any dependencies on RepLAB's larger repertoire of data structures. This was especially true since, throughout the period in which RepCert was coded, RepLAB was during a phase of very active developing. Moreover, this separation allows one to easily benchmark RepCert's performance independently of RepLAB.

From the point of view of user experience, however, it would make sense to unify both packages. That is, to code RepCert as a particular functionality of RepLAB, so that the user can both decompose a representation and certify the decomposition within the same piece of software. This will be the subject of future work.

Erklärung zur Dissertation

Hiermit versichere ich an Eides Statt, dass ich die vorliegende Dissertation selbständig und ohne die Benuzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Alle Stellen, die Wörtlich oder sinngemäss aus veröffentlich und nicht veröffentlicher Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation, noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten - noch nicht veröffentlich worden ist sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung des Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt. Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlicher Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der Dissertation zugrundeliegenden Arbeiten und der schriftlich verfassten Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht.

Teilpublicationen

1. Montealegre-Mora, F., Gross, D. (2021) Rank-deficient representations in the Theta correspondence over finite fields arise from quantum codes. *Representation Theory of the American Mathematical Society, 25(8)*

2. Montealegre-Mora, F., Gross, D. (2021) The representation theory of Clifford tensor powers. *(in preparation).*

3. Haferkamp, J., Montealegre-Mora, F., Heinrich, M., Eisert, J., Gross, D., Roth, I. (2020) Quantum homeopathy works: Efficient unitary designs with a system-size independent number of non-Clifford gates. *ArXiv preprint, arXiv:2002.09524* (submitted to *Communications in Mathematical Physics*).

4. Montealegre-Mora, F., Rosset, D., Bancal, J.-D., Gross, D. (2021) Certifying numerical decompositions of compact gorup representations. *ArXiv preprint, arXiv:2101.12244.*

Felipe Montealegre Mora, 09.12.2021 Köln

References

[AB13a] Gérard Ben Arous and Paul Bourgade, *Extreme gaps between eigenvalues of random matrices*, The Annals of Probability **41** (2013), no. 4, 2648–2681.

[AB+13b] Gérard Ben Arous, Paul Bourgade, et al., *Extreme gaps between eigenvalues of random matrices*, The Annals of Probability **41** (2013), no. 4, 2648–2681.

[ABW09] Andris Ambainis, Jan Bouda, and Andreas Winter, *Nonmalleable encryption of quantum information*, Journal of Mathematical Physics **50** (2009), no. 4, 042106.

[AG04] Scott Aaronson and Daniel Gottesman, *Improved simulation of stabilizer circuits*, Physical Review A **70** (2004), no. 5, 052328.

[AKP16] AM Aubert, W Kraskiewicz, and T Przebinda, *Howe correspondence and springer correspondence for dual pairs over a finite field*, Proceedings of Symposia in Pure Mathematics, vol. 92, 2016, pp. 17–44.

[AM93] Jeffrey Adams and Allen Moy, *Unipotent representations and reductive dual pairs over finite fields*, Transactions of the american mathematical society **340** (1993), no. 1, 309–321.

[AMB04] Karim Abed-Meraim and Adel Belouchrani, *Algorithms for joint block diagonalization*, 2004 12th European Signal Processing Conference, IEEE, 2004, pp. 209–212.

[AMR96] Anne-Marie Aubert, Jean Michel, and Raphaël Rouquier, *Correspondance de howe pour les groupes réductifs sur les corps finis*, Duke Mathematical Journal **83** (1996), no. 2, 353–397.

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

[BDCP12] Hector Bombin, Guillaume Duclos-Cianci, and David Poulin, *Universal topological phase of two-dimensional stabilizer codes*, New Journal of Physics **14** (2012), no. 7, 073048.

[BF91] László Babai and Katalin Friedl, *Approximate representation theory of finite groups*, [1991] Proceedings 32nd Annual Symposium of Foundations of Computer Science, IEEE, 1991, pp. 733–742.

[BFS93] László Babai, Katalin Friedl, and Markus Stricker, *Decomposition of $*$-closed algebras in polynomial time*, Proceedings of the 1993 international symposium on Symbolic and algebraic computation, 1993, pp. 86–94.

[BGGS21] Anne Broadbent, Carlos E González-Guillén, and Christine Schuknecht, *Quantum private broadcasting*, arXiv preprint arXiv:2107.11474 (2021).

[Bha13] Rajendra Bhatia, *Matrix analysis*, vol. 169, Springer Science & Business Media, 2013.

[BHH16] Fernando GSL Brandao, Aram W Harrow, and Michał Horodecki, *Local random quantum circuits are approximate polynomial-designs*, Communications in Mathematical Physics **346** (2016), no. 2, 397–434.

[BK19] Kaifeng Bu and Dax Enshan Koh, *Efficient classical simulation of clifford circuits with nonstabilizer input states*, Physical review letters **123** (2019), no. 17, 170502.

[BNRT20] Eiichi Bannai, Gabriel Navarro, Noelia Rizo, and Pham Huu Tiep, *Unitary t-groups*, Journal of the Mathematical Society of Japan **72** (2020), no. 3, 909–921.

[BOZ21] Eiichi Bannai, Manabu Oura, and Da Zhao, *The complex conjugate invariants of clifford groups*, Designs, Codes and Cryptography **89** (2021), no. 2, 341–350.

[BR90] László Babai and Lajos Rónyai, *Computing irreducible representations of finite groups*, Mathematics of computation **55** (1990), no. 192, 705–722.

[BRW61] Beverley Bolt, TG Room, and GE Wall, *On the Clifford collineation, transform and similarity groups. ii.*, Journal of the Australian Mathematical Society **2** (1961), no. 1, 80–96.

[Bum04] Daniel Bump, *Lie groups*, vol. 8, Springer, 2004.

[Car85] Roger William Carter, *Finite groups of lie type: Conjugacy classes and complex characters*, Pure Appl. Math. **44** (1985).

[CCS19] Yunfeng Cai, Guanghui Cheng, and Decai Shi, *Solving the general joint block diagonalization problem via linearly independent eigenvectors of a matrix polynomial*, Numerical Linear Algebra with Applications **26** (2019), no. 4, e2238.

[CL06] Xiaoshan Chen and Wen Li, *A note on the perturbation bounds of eigenspaces for hermitian matrices*, Journal of computational and applied mathematics **196** (2006), no. 1, 338–346.

[CL17] Yunfeng Cai and Chengyu Liu, *An algebraic approach to nonorthogonal general joint block diagonalization*, SIAM Journal on Matrix Analysis and Applications **38** (2017), no. 1, 50–71.

[CL20] Yunfeng Cai and Ping Li, *Identification of matrix joint block diagonalization*, arXiv preprint arXiv:2011.01111 (2020).

[CSST14] Tullio Ceccherini-Silberstein, Fabio Scarabotti, and Filippo Tolli, *Representation theory and harmonic analysis of wreath products of finite groups*, vol. 410, Cambridge University Press, 2014.

[CSX15] Yunfeng Cai, Decai Shi, and Shufang Xu, *A matrix polynomial spectral approach for general joint block diagonalization*, SIAM Journal on Matrix Analysis and Applications **36** (2015), no. 2, 839–863.

[DHW19] Bas Dirkse, Jonas Helsen, and Stephanie Wehner, *Efficient unitarity randomized benchmarking of few-qubit clifford gates*, Physical Review A **99** (2019), no. 1, 012315.

[Dix70] John D Dixon, *Computing irreducible representations of groups*, Mathematics of Computation **24** (1970), no. 111, 707–712.

[DK70] Chandler Davis and William Morton Kahan, *The rotation of eigenvectors by a perturbation. iii*, SIAM Journal on Numerical Analysis **7** (1970), no. 1, 1–46.

[dKDP11] Etienne de Klerk, Cristian Dobre, and Dmitrii V Pasechnik, *Numerical block diagonalization of matrix $*$-algebras with application to semidefinite programming*, Mathematical programming **129** (2011), no. 1, 91.

[DM20] François Digne and Jean Michel, *Representations of finite groups of lie type*, vol. 95, Cambridge University Press, 2020.

[DS81] Persi Diaconis and Mehrdad Shahshahani, *Generating a random permutation with random transpositions*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete **57** (1981), no. 2, 159–179.

[DS21] Anatoly Dymarsky and Alfred Shapere, *Quantum stabilizer codes, lattices, and cfts*, Journal of High Energy Physics **2021** (2021), no. 3, 1–84.

[Fol89] Gerald B Folland, *Harmonic analysis in phase space*, Princeton University Press, 1989.

[FST21] Omar Fawzi, Ala Shayeghi, and Hoang Ta, *A hierarchy of efficient bounds on quantum capacities exploiting symmetry*, 2021 IEEE International Symposium on Information Theory (ISIT), IEEE, 2021, pp. 272–277.

[GAE07] D. Gross, K. Audenaert, and J. Eisert, *Evenly distributed unitaries: On the structure of unitary designs*, J. Math. Phys. **48** (2007), no. 5.

[GAP21] The GAP Group, *GAP – Groups, Algorithms, and Programming, Version 4.11.1*, 2021.

[Gec17] Meinolf Geck, *A first guide to the character theory of finite groups of lie type*, arXiv preprint arXiv:1705.05083 (2017).

[Gel06] Stephen S Gelbart, *Weil's representation and the spectrum of the metaplectic group*, vol. 530, Springer, 2006.

[GGKS20] David Gosset, Daniel Grier, Alex Kerzner, and Luke Schaeffer, *Fast simulation of planar clifford circuits*, arXiv preprint arXiv:2009.03218 (2020).

[GH17] Shamgar Gurevich and Roger Howe, *Small representations of finite classical groups*, Representation Theory, Number Theory, and Invariant Theory, Springer, 2017, pp. 209–234.

[GH20] _____, *Rank and duality in representation theory*, Japanese Journal of Mathematics **15** (2020), 223–309.

[GH21] _____, *Ranks for representations of gln over finite fields, their agreement, and positivity of fourier transform*, Indagationes Mathematicae (2021).

[GNW21] David Gross, Sepehr Nezami, and Michael Walter, *Schur–weyl duality for the clifford group with applications: Property testing, a robust hudson theorem, and de finetti representations*, Communications in Mathematical Physics (2021), 1–69.

[Gro06] David Gross, *Hudson's theorem for finite-dimensional quantum systems*, J. Math. Phys. **47** (2006), no. 12, 122107.

[Gro19] ———, *Private communication*, 2019.

[Hei21] Markus Heinrich, *On stabiliser techniques and their application to simulation and certification of quantum devices*, Ph.D. thesis, Universität zu Köln, 2021.

[HF17] Robin Harper and Steven T Flammia, *Estimating the fidelity of t gates using standard interleaved randomized benchmarking*, Quantum Science and Technology **2** (2017), no. 1, 015008.

[HFGW18] AK Hashagen, ST Flammia, D Gross, and JJ Wallman, *Real randomized benchmarking*, Quantum **2** (2018), 85.

[HG19] Markus Heinrich and David Gross, *Robustness of magic and symmetries of the stabiliser polytope*, Quantum **3** (2019), 132.

[HHJ+17] Jeongwan Haah, Aram W Harrow, Zhengfeng Ji, Xiaodi Wu, and Nengkun Yu, *Sample-optimal tomography of quantum states*, IEEE Transactions on Information Theory **63** (2017), no. 9, 5628–5641.

[HL19] Yifei Huang and Peter Love, *Approximate stabilizer rank and improved weak simulation of clifford-dominated circuits for qudits*, Physical Review A **99** (2019), no. 5, 052307.

[HL21] ———, *Feynman-path-type simulation using stabilizer projector decomposition of unitaries*, Physical Review A **103** (2021), no. 2, 022428.

[HMMH+20] Jonas Haferkamp, Felipe Montealegre-Mora, Markus Heinrich, Jens Eisert, David Gross, and Ingo Roth, *Quantum homeopathy works: Efficient unitary designs with a system-size independent number of non-clifford gates*, arXiv preprint arXiv:2002.09524 (2020).

[HMMVG21] Arne Heimendahl, Felipe Montealegre-Mora, Frank Vallentin, and David Gross, *Stabilizer extent is not multiplicative*, Quantum **5** (2021), 400.

[HMW20] Stefan Hillmich, Igor L Markov, and Robert Wille, *Just like the real thing: Fast weak simulation of quantum computation*, 2020 57th ACM/IEEE Design Automation Conference (DAC), IEEE, 2020, pp. 1–6.

[How73] Roger Howe, *Invariant theory and duality for classical groups over finite fields with applications to their singular representation theory*, preprint (1973).

[How89a] ———, *Remarks on classical invariant theory*, Transactions of the American Mathematical Society **313** (1989), no. 2, 539–570.

[How89b] ———, *Transcending classical invariant theory*, Journal of the American Mathematical Society **2** (1989), no. 3, 535–552.

[How10] ———, *On a notion of rank for unitary representations of the classical groups*, Talamanca A.F. (eds) Harmonic Analysis and Group Representation. C.I.M.E. Summer Schools, vol. 82, Springer, 2010, pp. 224–331.

[HP07] Patrick Hayden and John Preskill, *Black holes as mirrors: quantum information in random subsystems*, Journal of high energy physics **2007** (2007), no. 09, 120.

[HWFW19] Jonas Helsen, Joel J Wallman, Steven T Flammia, and Stephanie Wehner, *Multiqubit randomized benchmarking using few samples*, Physical Review A **100** (2019), no. 3, 032304.

[KdSR+14] Shelby Kimmel, Marcus P da Silva, Colm A Ryan, Blake R Johnson, and Thomas Ohki, *Robust extraction of tomographic information via randomized benchmarking*, Physical Review X **4** (2014), no. 1, 011050.

[KG15] Richard Kueng and David Gross, *Qubit stabilizer states are complex projective 3-designs*.

[KL17] Shelby Kimmel and Yi-Kai Liu, *Phase retrieval using unitary 2-designs*, 2017 International Conference on Sampling Theory and Applications (SampTA), IEEE, 2017, pp. 345–349.

[KLR+08] Emanuel Knill, Dietrich Leibfried, Rolf Reichle, Joe Britton, R Brad Blakestad, John D Jost, Chris Langer, Roee Ozeri, Signe Seidelin, and David J Wineland, *Randomized benchmarking of quantum gates*, Physical Review A **77** (2008), no. 1, 012307.

[KMK21] Niraj Kumar, Rawad Mezher, and Elham Kashefi, *Efficient construction of quantum physical unclonable functions with unitary t-designs*, arXiv preprint arXiv:2101.05692 (2021).

[KR21] Martin Kliesch and Ingo Roth, *Theory of quantum system certification*, PRX Quantum **2** (2021), no. 1, 010201.

[KV78] Masaki Kashiwara and Michele Vergne, *On the segal-shale-weil representations and harmonic polynomials*, Inventiones mathematicae **44** (1978), no. 1, 1–47.

[KW05] Michael Keyl and Reinhard F Werner, *Estimating the spectrum of a density operator*, Asymptotic Theory Of Quantum Statistical Inference: Selected Papers, World Scientific, 2005, pp. 458–467.

[KZG16a] Richard Kueng, Huangjun Zhu, and David Gross, *Distinguishing quantum states using clifford orbits*, arXiv preprint arXiv:1609.08595 (2016).

[KZG16b] _____, *Low rank matrix recovery from clifford orbits*, arXiv preprint arXiv:1610.08070 (2016).

[MGE11] Easwar Magesan, Jay M Gambetta, and Joseph Emerson, *Scalable and robust randomized benchmarking of quantum processes*, Physical review letters **106** (2011), no. 18, 180504.

[MGE12] _____, *Characterizing quantum gates via randomized benchmarking*, Physical Review A **85** (2012), no. 4, 042311.

[MGJ+12] Easwar Magesan, Jay M Gambetta, Blake R Johnson, Colm A Ryan, Jerry M Chow, Seth T Merkel, Marcus P Da Silva, George A Keefe, Mary B Rothwell, Thomas A Ohki, et al., *Efficient measurement of quantum gate error by interleaved randomized benchmarking*, Physical review letters **109** (2012), no. 8, 080505.

[MKKK10] Kazuo Murota, Yoshihiro Kanno, Masakazu Kojima, and Sadayoshi Kojima, *A numerical algorithm for block-diagonal decomposition of matrix $*$-algebras with application to semidefinite programming*, Japan Journal of Industrial and Applied Mathematics **27** (2010), no. 1, 125–160.

[MM10] Takanori Maehara and Kazuo Murota, *A numerical algorithm for block-diagonal decomposition of matrix $*$-algebras with general irreducible components*, Japan journal of industrial and applied mathematics **27** (2010), no. 2, 263–293.

[MM11] _____, *Algorithm for error-controlled simultaneous block-diagonalization of matrices*, SIAM Journal on Matrix Analysis and Applications **32** (2011), no. 2, 605–620.

[MM21] Felipe Montealegre-Mora, *RepCert*, 2021, `https://github.com/felimomo/RepCert`.

[MMG21a] Felipe Montealegre-Mora and David Gross, *Rank-deficient representations in the theta correspondence over finite fields arise from quantum codes*, Representation Theory of the American Mathematical Society **25** (2021), no. 8, 193–223.

[MMG21b] ———, *The representation theory of Clifford tensor powers*, In preparation.

[MMRBG21] Felipe Montealegre-Mora, Denis Rosset, Jean-Daniel Bancal, and David Gross, *Certifying numerical decompositions of compact group representations*, arXiv preprint arXiv:2101.12244 (2021).

[MSM19] Jisho Miyazaki, Akihito Soeda, and Mio Murao, *Complex conjugation supermap of unitary quantum maps and its universal implementation protocol*, Physical Review Research **1** (2019), no. 1, 013007.

[NC10] Michael A Nielsen and Isaac L Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2010.

[NRS01] Gabriele Nebe, Eric M. Rains, and Neil JA Sloane, *The invariants of the Clifford groups*, Designs, Codes and Cryptography **24** (2001), no. 1, 99–122.

[NRS06] Gabriele Nebe, Eric M Rains, and Neil James Alexander Sloane, *Self-dual codes and invariant theory*, Springer, 2006.

[NW16] Sepehr Nezami and Michael Walter, *Multipartite entanglement in stabilizer tensor networks*.

[Pan20] Shu-Yen Pan, *On theta and eta correspondences for finite symplectic/orthogonal dual pairs*, arXiv preprint arXiv:2006.06241 (2020).

[PP20] Frank Permenter and Pablo A Parrilo, *Dimension reduction for semidefinite programs via Jordan algebras*, Mathematical Programming **181** (2020), no. 1, 51–84.

[Qas21] Hammam Qassim, *Classical simulations of quantum systems using stabilizer decompositions*.

[QDS+19] Marco Túlio Quintino, Qingxiuxiong Dong, Atsushi Shimbo, Akihito Soeda, and Mio Murao, *Probabilistic exact universal quantum circuits*

for transforming unitary operations, Physical Review A **100** (2019), no. 6, 062339.

[RB18] Denis Rosset and Jean-Daniel Bancal, *RepLAB*, 2018, `https://replab.github.io/replab`.

[RKK+18] Ingo Roth, Richard Kueng, Shelby Kimmel, Y-K Liu, David Gross, Jens Eisert, and Martin Kliesch, *Recovering quantum gates from few average gate fidelities*, Physical review letters **121** (2018), no. 17, 170502.

[RLCK19] Patrick Rall, Daniel Liang, Jeremy Cook, and William Kretschmer, *Simulation of qubit quantum circuits via pauli propagation*, Physical Review A **99** (2019), no. 6, 062337.

[RMMB19] Denis Rosset, Felipe Montealegre-Mora, and Jean-Daniel Bancal, *RepLAB: a computational/numerical approach to representation theory*, arXiv preprint arXiv:1911.09154 (2019).

[RRMG17] Marc Olivier Renou, Denis Rosset, Anthony Martin, and Nicolas Gisin, *On the inequivalence of the ch and chsh inequalities due to finite statistics*, Journal of Physics A: Mathematical and Theoretical **50** (2017), no. 25, 255301.

[RY17] Daniel A Roberts and Beni Yoshida, *Chaos and complexity by design*, Journal of High Energy Physics **2017** (2017), no. 4, 1–64.

[Sco08] Andrew James Scott, *Optimizing quantum process tomography with unitary 2-designs*, Journal of Physics A: Mathematical and Theoretical **41** (2008), no. 5, 055308.

[Sri79] Bhama Srinivasan, *Weil representations of finite classical groups*, Inventiones mathematicae **51** (1979), no. 2, 143–153.

[TCUA20] Armin Tavakoli, Emmanuel Zambrini Cruzeiro, Roope Uola, and Alastair A Abbott, *Bounding and simulating contextual correlations in quantum theory*, arXiv preprint arXiv:2010.04751 (2020).

[Tei20] Pedro Abdalla Teixeira, *Non asymptotic random matrix theory and the small ball method*, Ph.D. thesis, PUC-Rio, 2020.

[TFR+21] Armin Tavakoli, Máté Farkas, Denis Rosset, Jean-Daniel Bancal, and Jedrzej Kaniewski, *Mutually unbiased bases and symmetric informationally complete measurements in bell experiments*, Science Advances **7** (2021), no. 7, eabc3847.

[Tro12] Joel A Tropp, *User-friendly tail bounds for sums of random matrices*, Foundations of computational mathematics **12** (2012), no. 4, 389–434.

[Val09] Frank Vallentin, *Symmetry in semidefinite programs*, Linear Algebra and its Applications **430** (2009), no. 1, 360–369.

[Ver10] Roman Vershynin, *Introduction to the non-asymptotic analysis of random matrices*, arXiv preprint arXiv:1011.3027 (2010).

[Web16] Zak Webb, *The Clifford group forms a unitary 3-design*, Quant. Inf. Comp. **26** (2016), 1379–1400.

[Zhu16] Huangjun Zhu, *Permutation symmetry determines the discrete wigner function*, Physical review letters **116** (2016), no. 4, 040501.

[Zhu17] ———, *Multiqubit Clifford groups are unitary 3-designs*, Physical Review A **96** (2017), no. 6, 062336.

[ZKGG16] Huangjun Zhu, Richard Kueng, Markus Grassl, and David Gross, *The Clifford group fails gracefully to be a unitary 4-design*.