Efficient approximate unitary designs from random Pauli rotations

Jeongwan Haah
Microsoft Quantum
Redmond, WA, USA
jwhaah@microsoft.com

Yunchao Liu
Department of EECS
University of California, Berkeley
Berkeley, CA, USA
yunchaoliu@berkeley.edu

Xinyu Tan
Department of Mathematics
Massachusetts Institute of Technology
Cambridge, MA, USA
norahtan@mit.edu

2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS) | 979-8-3315-1674-1/24/$31.00 ©2024 IEEE | DOI: 10.1109/FOCS61266.2024.00036

***Abstract*—We construct random walks on simple Lie groups that quickly converge to the Haar measure for all moments up to order t. Specifically, a step of the walk on the unitary or orthogonal group of dimension 2^{n} is a random Pauli rotation $e^{\mathrm{i}\theta P/2}$. The spectral gap of this random walk is shown to be $\Omega(1/t)$, which coincides with the best previously known bound for a random walk on the permutation group on $\{0,1\}^{\mathsf{n}}$. This implies that the walk gives an ε-approximate unitary t-design in depth $\mathcal{O}(\mathsf{n}t^2 + t\log\frac{1}{\varepsilon})d$ where $d = \mathcal{O}(\log \mathsf{n})$ is the circuit depth to implement $e^{\mathrm{i}\theta P/2}$. Our simple proof uses quadratic Casimir operators of Lie algebras.**

***Index Terms*—pseudorandomness, quantum computing, unitary designs, spectral methods, derandomization**

I. Introduction

An approximate unitary t-design [1], [2] is an ensemble of unitaries that behaves similarly to the Haar random ensemble up to t-th moments. For n-qubit ($\mathbb{C}^2$) systems, there have been constructions of approximate unitary t-designs with circuit size $\mathrm{poly}(\mathsf{n}, t)$ [3], [4], which have found wide applications in quantum information theory. However, existing constructions using local random quantum circuits had rather steep dependence on t. In this paper, we consider random Pauli rotations, which are $\exp(\mathrm{i}\theta P/2)$ where θ is a random angle and P is a random n-qubit Pauli operator. We show that the product of k independent random Pauli rotations $e^{\mathrm{i}\theta_k P_k/2}\cdots e^{\mathrm{i}\theta_2 P_2/2}e^{\mathrm{i}\theta_1 P_1/2}$ converges to a unitary t-design as k increases.

Theorem I.1. *For any integers $\mathsf{n}, t \geq 1$, it holds that*

$$\left\| \mathop{\mathbb{E}}_{\theta\sim(-\pi,\pi)} \mathop{\mathbb{E}}_{P\sim\mathbf{P}_{\mathsf{n}}} \left(e^{\mathrm{i}\frac{\theta}{2}P}\otimes e^{-\mathrm{i}\frac{\theta}{2}\bar{P}}\right)^{\otimes t} - \mathop{\mathbb{E}}_{U\sim\mathsf{SU}(2^{\mathsf{n}})} (U\otimes\bar{U})^{\otimes t} \right\| \leq 1 - \frac{1}{4t} - \frac{1}{4^{\mathsf{n}}-1}. \quad (1)$$

Here, $\mathbf{P}_{\mathsf{n}} = \{\mathbf{1}_2, \sigma^x, \sigma^y, \sigma^z\}^{\otimes \mathsf{n}} \setminus \{\mathbf{1}_{2^{\mathsf{n}}}\}$ is the set of all nonidentity n-qubit Pauli operators, the norm denotes the greatest singular value, $\bar{U}$ denotes the complex conjugate of U, and the distributions for P, θ, and U are uniform in the designated domains.

In addition, for any finite dimensional unitary representation ρ of $\mathsf{SU}(2^{\mathsf{n}})$, we have

$$\left\| \mathop{\mathbb{E}}_{\theta\sim(-2\pi,2\pi)} \mathop{\mathbb{E}}_{P\sim\mathbf{P}_{\mathsf{n}}} \rho(e^{\mathrm{i}\frac{\theta}{2}P}) - \mathop{\mathbb{E}}_{U\sim\mathsf{SU}(2^{\mathsf{n}})} \rho(U) \right\| < 1 - \frac{1}{4^{\mathsf{n}}-1}. \quad (2)$$

Corollary I.2. *Consider two mixed unitary channels*

$$\mathcal{C}_t : \eta \mapsto \mathop{\mathbb{E}}_{P\sim\mathbf{P}_{\mathsf{n}},\, \theta\sim(-\pi,\pi)} \left(e^{\mathrm{i}\frac{\theta}{2}P}\right)^{\otimes t} \eta \left(e^{-\mathrm{i}\frac{\theta}{2}P}\right)^{\otimes t}$$

and

$$\mathcal{H}_t : \eta \mapsto \mathop{\mathbb{E}}_{U\sim\mathsf{SU}(2^{\mathsf{n}})} U^{\otimes t}\eta U^{\dagger\otimes t}$$

using the same distributions of P, θ and U as in Eq. (1).

1) If $k \geq (4\log 2)\mathsf{n}t^2 + 4t\log\frac{1}{\varepsilon}$, then

$$\left\|\mathcal{C}_t^k - \mathcal{H}_t\right\|_\diamond \leq \varepsilon.$$

2) If $k \geq (4\log 8)\mathsf{n}t^2 + 4t\log\frac{1}{\varepsilon}$, then

$$(1-\varepsilon)\mathcal{H}_t \preceq \mathcal{C}_t^k \preceq (1+\varepsilon)\mathcal{H}_t.$$

Here, $\|\cdot\|_\diamond$ denotes the diamond norm (completely bounded trace norm). Every instance $e^{\mathrm{i}\theta_k P_k/2}\cdots e^{\mathrm{i}\theta_1 P_1/2}$ can be implemented using $\mathcal{O}(k\mathsf{n})$ 1-qubit and any-to-any CNOT *gates in depth $\mathcal{O}(k\log\mathsf{n})$.*

We also give similar results for the special orthogonal groups in Section VI.

A. Previous spectral gap bounds

Unless otherwise noted, N stands for 2^{n}.

For a distribution ν over $\mathsf{SU}(\mathsf{N})$, the spectral gap $\Delta(\nu, t)$ at t-th order[1] is given by

$$1 - \Delta(\nu,t) = \left\| \mathop{\mathbb{E}}_{U\sim\nu} (U\otimes\bar{U})^{\otimes t} - \mathop{\mathbb{E}}_{U\sim\mathsf{SU}(\mathsf{N})} (U\otimes\bar{U})^{\otimes t} \right\|.$$

[1] In [5], the spectral gap means

$$1 - \sup_\rho \left\| \mathop{\mathbb{E}}_{U\sim\nu} \rho(U) - \mathop{\mathbb{E}}_{U\sim\mathsf{SU}(\mathsf{N})} \rho(U) \right\| = 1 - \sup_\rho \left\| \mathop{\mathbb{E}}_{U\sim\nu} \rho(U) \right\|$$

where the supremum is over all nontrivial finite dimensional unitary irreps of $\mathsf{SU}(\mathsf{N})$. This is an immediate consequence of Proposition II.4 and the Peter–Weyl theorem. See also [6, Thm 3.9]. Not all irreps of $\mathsf{SU}(\mathsf{N})$ appear in $U \mapsto \bigoplus_{t\geq 1}(U\otimes\bar{U})^{\otimes t}$.

Y.L. is supported by DOE Grant No. DE-SC0024124, NSF Grant No. 2311733, and DOE Quantum Systems Accelerator. X.T. is supported by NSF Grant No. CCF-1729369.

2575-8454/24/$31.00 ©2024 IEEE
DOI 10.1109/FOCS61266.2024.00036

Consider the distribution of the product of k independent draws from ν, which corresponds to the k-fold convolution ν^{*k}. Then, since $\mathcal{H}_t = \mathbb{E}_{U\sim \mathsf{SU}(\mathsf{N})}(U\otimes \bar{U})^{\otimes t}$ is a projector and $\mathbb{E}_{U\sim\nu}(U\otimes\bar{U})^{\otimes t}$ contains the image of $\mathcal{H}_t$ in the eigenspace of eigenvalue $+1$ (see Proposition II.4), the spectral gap amplifies as $1-\Delta(\nu^{*k},t) = (1-\Delta(\nu,t))^k$. This allows us to exponentially improve the accuracy at the cost of linear blow-up in circuit size. More generally, proving lower bounds on the spectral gap of the t-th moment operator is a standard approach to show that a random walk on a group quickly converges to a t-wise independent distribution (often referred to as "designs").

Hence, a primary goal in efficient approximate unitary designs is to find ν with $\mathrm{poly}(\mathsf{n})$ circuit size with a large spectral gap, for example, $1/\mathrm{poly}(\mathsf{n},t)$. A simple brickwall "spacetime" geometry of random unitary circuit has been shown to achieve this goal [3], whose analysis was recently improved [4]. Once the operator norm distance is bounded, one can convert it to additive or relative diamond distance.

As far as we know, the best previous spectral gap for any efficient approximate unitary design on an n-qubit system was $\Omega(t^{-4-o(1)})$ [4]. This work takes the ensemble of [3], where the circuit geometry is brickwall that uses local gates in a one-dimensional lattice. Our ensemble does not have any geometric locality. Note that Theorem I.1 gives a lower bound $4^{-\mathsf{n}}$ on the spectral gap independent of t. Such a t-independent bound was also given in [4, Theorem 1], which reads $\Omega(\mathsf{n}^{-5}4^{-\mathsf{n}})$.

Similarly to unitary designs, the best previous spectral gap lower bound for the special orthogonal group $\mathsf{SO}(\mathsf{N})$ had a large inverse-polynomial dependence on t [6], while the best previous spectral gap for the symmetric group S_{N} was $\Omega(t^{-1})$ [7].[2] Our spectral gap bounds for the special unitary and orthogonal groups are thus the strongest in terms of t dependence, and they coincides with the best known spectral gap for the symmetric group.

Some results on unitary designs bypass spectral gap analysis. Aiming to minimize non-Clifford resources, [8] analyzed alternating "Clifford+K" circuits and the diamond distance of the associated mixed unitary channel to the Haar random channel directly. Compared with the brickwall circuits, our construction is conceptually closer to [8]. However, their result is only applicable in the regime when $t = \mathcal{O}(\sqrt{\mathsf{n}})$.

While the previous best approximate unitary designs on n qubits have circuit depth $\mathcal{O}(\mathsf{n}t^{5+o(1)})$ for exponentially-large moments $t = \mathcal{O}(2^{0.4\mathsf{n}})$ [4], there exist constructions with linear dependence in t but with major restrictions. A family of stochastic Hamiltonians constructed in [9] were proved to be linear designs in the regime when $t = \mathcal{O}(\sqrt{\mathsf{n}})$. It was showed in [10] that the spectral gap of local random circuits becomes t-independent $\Omega(\mathsf{n}^{-1})$ when the local dimension of a qudit is at least $6t^2$. [11] argued that certain random time-dependent Hamiltonian evolution converges to t-designs at a linear rate.

[2]Here we ignore polynomial dependence in n as it can be eliminated by taking powers.

B. Implications

1) Circuit complexity: By known reductions [12], our result directly implies a lower bound for robust quantum circuit complexity of a product of k random Pauli rotations. Specifically, let U be a product of $k \ll 2^{\mathsf{n}}$ random Pauli rotations, which can be implemented by $\mathcal{O}(k\mathsf{n})$ gates. Then with high probability over the choice of U, any unitary V satisfying $\|U-V\| \le 0.01$ must have quantum circuit complexity (the minimum number of gates to implement V) $\tilde{\Omega}(\sqrt{k\mathsf{n}})$. Note that a robust square root circuit complexity lower bound was also established in [13]; however, the family of quantum circuits considered there used a non-universal gate set, and therefore does not form an approximate unitary design. A major open question is whether it is possible to construct distributions on $\mathsf{SU}(2^{\mathsf{n}})$ using $\mathrm{poly}(\mathsf{n})$ size quantum circuits, such that the spectral gap is at least $1/\mathrm{poly}(\mathsf{n})$ and independent of t. Such a result would imply a robust linear growth of quantum circuit complexity.

2) Seed length: Our unitary design requires sampling from a continuous interval $(-\pi,\pi)$; however, for given t, we can instead sample uniformly from a discrete set $\{m\pi/t : m\in\mathbb{Z}\cap[-t,t-1]\}$ (see Appendix A). Therefore, our distribution for ε-approximate unitary t-design is samplable using only $\mathcal{O}(t(\mathsf{n}t+\log 1/\varepsilon)(\mathsf{n}+\log t))$ random bits. Furthermore, instead of sampling each random Pauli rotation independently and uniformly at random, we can sample them in a pseudorandom way using a technique of [6] which is a generalization of the derandomized graph squaring [14]. We can thus reduce the seed length to only $\mathcal{O}(\mathsf{n}t+\log 1/\varepsilon)$ by applying [6, Theorem 6.21]. While this has the same scaling as the main result of [6], our construction has the advantage of having explicit constants, as we do not rely on the implicit spectral gap of [5].

3) Orthogonal designs and more: Our approach to unitary designs can be adapted to the special orthogonal groups $\mathsf{SO}(\mathsf{N})$ with parallel arguments. The results are found in Section VI. The seed length can be similarly reduced to $\mathcal{O}(\mathsf{n}t+\log 1/\varepsilon)$ with explicit constants. This has been used to construct pseudorandom generators for halfspaces [6]. Finally, we discuss quantum state designs in Appendix B, where we obtain better bounds than what would be obtained by directly applying our unitary design.

The analysis of the orthogonal groups is so similar to that of the unitary groups that one might desire to have unified statements for all simple finite dimensional Lie groups. However, as the representation theory of Lie groups is tackled in a case-by-case fashion in detail, we find it best to analyze them separately. Beyond the unitary and orthogonal groups, there is a family of symplectic groups, which might have applications in classical Hamiltonian dynamics and quantum optics as one often encounters symplectic spaces in these subjects.

C. Overview of the argument

We start by rewriting the tensor product in a different form, $\left(e^{\mathrm{i}\frac{\theta}{2}P}\otimes e^{-\mathrm{i}\frac{\theta}{2}\bar{P}}\right)^{\otimes t} = e^{\mathrm{i}\theta\tau_*(P/2)}$, where

$$\tau_*(P/2) = \frac{1}{2}\sum_{j=1}^{t}(\mathbf{1}_{\mathsf{N}}\otimes\mathbf{1}_{\mathsf{N}})^{\otimes(j-1)}\otimes(P\otimes\mathbf{1}_{\mathsf{N}}-\mathbf{1}_{\mathsf{N}}\otimes\bar{P})\otimes(\mathbf{1}_{\mathsf{N}}\otimes\mathbf{1}_{\mathsf{N}})^{\otimes(t-j)}. \tag{3}$$

Note that for every $P\in\mathbf{P}_{\mathsf{n}}$, the eigenvalues of $\tau_*(P/2)$ are exactly the integers in $[-t,t]$. Thus, the averaging over θ gives $\mathbb{E}_{\theta\sim(-\pi,\pi)}\, e^{\mathrm{i}\theta\tau_*(P/2)} = K_P$, where K_P denotes the orthogonal projector onto the kernel of $\tau_*(P/2)$. Our goal is now reduced to analyzing the spectrum of $\mathbb{E}_{P\sim\mathbf{P}_{\mathsf{n}}} K_P$. We calculate the norm of this exactly for the special case of $\mathsf{n}=1$ in Section IV. In general cases, we first block-diagonalize K_P using the observation that $P\mapsto\tau_*(P)$ is a Lie algebra representation. In each irrep, we upper bound the kernel projector by a quadratic approximation:

$$K(H)\preceq\mathbf{1}-H^2/\|H\|^2 \tag{4}$$

where $K(H)$ is the kernel projector for a Hermitian operator H, which holds for any nonzero H. This inequality is useful because the kernel projector sum is then bounded by a sum of squares of represented operators. A nice property of Pauli operators is that this sum of squares of represented operators is a scalar multiple of the identity for any irrep. We then bound the scalar by t,N.

We use the representation theory of Lie algebras, but our exposition is elementary for the core bound in Theorem I.1; we assume no prior knowledge beyond the representation theory of $\mathfrak{su}(2)$ for the main bound.

Note added. We recently became aware of independent related work of C. Chen, J. Docter, M. Xu, A. Bouland, and P. Hayden achieving similar results via a different construction [15].

II. Lie algebras and probability distributions

We begin with an observation that any unitary design can be regarded as a distribution on the linear space of a Lie algebra. This will allow us to analyze spectral properties of a unitary design by looking at certain hermitian operators in irreducible representations of $\mathfrak{su}(\mathsf{N}=2^{\mathsf{n}})$. We will find the latter more convenient since our unitary design will have the most succinct description as a distribution on the Lie algebra, rather than on the Lie group.

Often a Lie algebra is described by very concrete data, called structure constants, f^a_{bc}, that enter in the commutation relations as $[J_b,J_c]=\mathrm{i}\sum_a f^a_{bc}J_a$ where J_a are said to span the Lie algebra. While this is mostly correct and causes no trouble in practice, the appearance of the imaginary unit i might bring some confusion. So, we would like to clarify the complex and real coefficients. The tangent space at the origin of a Lie group, taken as a real manifold, is a real Lie algebra $\mathfrak{su}(\mathsf{N};\mathbb{R})$. This is a linear space over real numbers of traceless *anti*hermitian matrices where a Lie bracket is defined by matrix commutator; after all, the commutator of two hermitian operators is antihermitian, which is *not* in the $\mathbb{R}$-linear space of hermitian operators. However, since a representation space is taken to be a complex vector space, there is no reason not to allow complex coefficients in the span of antihermitian operators. This extension of the coefficient field is formally called the complexification of the Lie algebra: $\mathbb{C}\otimes_{\mathbb{R}}\mathfrak{su}(\mathsf{N};\mathbb{R})$. This complexified space consists of all $\mathbb{C}$-linear combinations of traceless antihermitian operators, which is the $\mathbb{C}$-linear space of *all* traceless matrices. Hence, the complexification is perhaps better denoted as $\mathfrak{sl}(\mathsf{N};\mathbb{C})=\mathbb{C}\otimes_{\mathbb{R}}\mathfrak{su}(\mathsf{N};\mathbb{R})$, the Lie algebra of special linear group. In this paper, we take a liberal convention that

- when $\mathfrak{su}(\mathsf{N})$ appears in the context of representation, we mean its complexificiation $\mathfrak{sl}(\mathsf{N};\mathbb{C})$, and
- when we discuss a probability distribution on $\mathfrak{su}(\mathsf{N})$, we mean a distribution on the real vector space of hermitian, rather than antihermitian, operators, with insertion of the imaginary unit i, whenever needed, understood.

Suppose we have an M-dimensional representation $\rho:\mathsf{SU}(\mathsf{N})\to\mathsf{U}(M)$ of $\mathsf{SU}(\mathsf{N})$ for some $M\geq 1$, which may be reducible. The representation map ρ is a Lie group homomorphism, and we have a corresponding commutative diagram [16, §8.3] by the exponential map:

$$\begin{array}{ccc}\mathfrak{su}(\mathsf{N}) & \xrightarrow{\rho_*} & \mathfrak{u}(M)\\ \downarrow{\scriptstyle\exp} & & \downarrow{\scriptstyle\exp}\\ \mathsf{SU}(\mathsf{N}) & \xrightarrow{\rho} & \mathsf{U}(M)\end{array} \tag{5}$$

where ρ_* is the induced, natural, Lie algebra homomorphism (a representation). In the context of unitary designs, we are interested in the tensor representation $\tau:\mathsf{SU}(\mathsf{N})\ni U\mapsto(U\otimes\bar{U})^{\otimes t}$ so $M=\mathsf{N}^{2t}$, whose induced Lie algebra representation τ_* is given by Eq. (3) for all traceless $\mathsf{N}\times\mathsf{N}$ matrix P.

We define

$$\mathbf{P}_{\mathsf{n}}=\{\mathbf{1}_2,\sigma^x,\sigma^y,\sigma^z\}^{\otimes\mathsf{n}}\setminus\{\mathbf{1}_{2^{\mathsf{n}}}\},$$

the set of all nonidentity tensor products of Pauli matrices

$$\sigma^x=\begin{pmatrix}0&1\\1&0\end{pmatrix},\quad \sigma^y=\begin{pmatrix}0&-\mathrm{i}\\\mathrm{i}&0\end{pmatrix},\quad \sigma^z=\begin{pmatrix}1&0\\0&-1\end{pmatrix}.$$

There are $\mathsf{N}^2-1=4^{\mathsf{n}}-1$ elements, each of which is called a Pauli operator. The factor of half in Eq. (3) is immaterial in that equation since τ_* is $\mathbb{C}$-linear, but we will keep it because

Lemma II.1. *For any Pauli operator $P\in\mathbf{P}_{\mathsf{n}}$, all the eigenvalues of $\tau_*(P/2)$ are integers in $[-t,t]$, and any integer in that range appears as an eigenvalue of $\tau_*(P/2)$.*

Proof. All the summands of Eq. (3) are commuting with each other, so they are simultaneously diagonalizable, which

amounts to setting $P = \sigma^z \otimes \mathbf{1}_{2^{n-1}}$, that is a diagonal matrix with ± 1 on the diagonal. The lemma follows. Alternatively, we can think of τ_* as $\tau_* = (\mathbf{1} \oplus \mathsf{ad})^{\otimes t} = \bigoplus_{k=0}^{t} \binom{t}{k} \mathsf{ad}^{\otimes k}$ where $\mathbf{1}$ denotes the trivial representation and ad is the adjoint representation. This is because the tensor product of the defining representation of $\mathfrak{su}(\mathsf{N})$ and its dual is a direct sum of the trivial and the adjoint, each of which is irreducible.[3] The represented operator $\mathsf{ad}(P/2)$ has eigenvalues ± 1, and therefore $\mathsf{ad}^{\otimes t}$ has integer eigenvalues in $[-t, t]$. □

We note that all Pauli operators P are equivalent to one another in any representation:

Lemma II.2. *For any representation ρ_* of $\mathfrak{su}(\mathsf{N})$ that is possibly reducible, and any nonidentity hermitian Pauli operator P, the eigenvalue spectrum of the represented operator $\rho_*(P)$ is independent of P. In particular, the eigenvalue spectrum of $\rho_*(P)$ is inversion symmetric about the origin; that is, $\rho_*(P)$ and $-\rho_*(P)$ have the same spectrum.*

Proof. Any two nonidentity Pauli operator P, Q on n qubits are congruent: $P = UQU^\dagger$ for some $U \in \mathsf{SU}(\mathsf{N})$. Exponentiating with $\theta \in \mathbb{R}$ we have $e^{\mathrm{i}\theta P} = Ue^{\mathrm{i}\theta Q}U^\dagger$, and thus $\rho(e^{\mathrm{i}\theta P}) = \rho(U)\rho(e^{\mathrm{i}\theta Q})\rho(U)^\dagger$. By Eq. (5) this translates to $e^{\mathrm{i}\theta\rho_*(P)} = \rho(U)e^{\mathrm{i}\theta\rho_*(Q)}\rho(U)^\dagger$. Differentiating with respect to θ, we finally have $\rho_*(P) = \rho(U)\rho_*(Q)\rho(U)^\dagger$. The last claim is because P and $-P$ are congruent by some anticommuting Pauli operator. □

Now, we can consider probability distributions on $\mathfrak{su}(\mathsf{N})$ and their induced distributions on $\mathsf{SU}(\mathsf{N})$. For example, to assess a unitary design we have to analyze the distribution on $\mathsf{U}(\mathsf{N}^{2t})$ for various values of t induced by the tensor representation τ. For a probability distribution μ on the top left of the diagram Eq. (5), we have corresponding distributions on all three other entries. For any distribution μ on $\mathfrak{su}(\mathsf{N})$, we denote an average with respect to μ by $\int_{\mathfrak{su}(\mathsf{N})} \cdots \mu(X)\mathrm{d}X$ where X denotes any hermitian operator.[4] In other words, $X \mapsto \mu(X)$ is the probability density "function." For any distribution μ on $\mathfrak{su}(\mathsf{N})$ and any integer $t \geq 1$ we consider a linear operator on $(\mathbb{C}^{\mathsf{N}})^{\otimes 2t}$

$$\mathcal{C}_{\mu,t} = \int_{\mathfrak{su}(\mathsf{N})} \exp(\mathrm{i}X)^{\otimes t} \otimes \exp(-\mathrm{i}\bar{X})^{\otimes t} \mu(X)\mathrm{d}X. \quad (6)$$

An obvious lemma will be useful:

Lemma II.3. *If $\phi : \mathsf{SU}(\mathsf{N}) \to \mathrm{Aut}(V)$ for $V \subseteq (\mathbb{C}^{\mathsf{N}})^{\otimes 2t}$ is a subrepresentation of $\tau : U \mapsto (U \otimes \bar{U})^{\otimes t}$, then*

$$\mathcal{C}_{\mu,t}|_V = \int_{\mathfrak{su}(\mathsf{N})} \exp(\mathrm{i}\phi_*(X))\mu(X)\mathrm{d}X.$$

[3] On the trivial representation the representation map is zero, and on the adjoint we have $\mathsf{ad}(P/2) = \frac{1}{2}(P \otimes \mathbf{1} - \mathbf{1} \otimes \bar{P})|_{\mathfrak{su}(\mathsf{N})}$ where the restriction is on the linear space of all (vectorized) traceless N-by-N matrices. Almost always, the adjoint representation map is explained by $\mathsf{ad}(P/2)(X) = \frac{1}{2}[P, X]$ for all $X \in \mathfrak{su}(\mathsf{N})$.

[4] Another prevalent notation is $\int \cdots \mathrm{d}\mu(X)$.

Proof. ϕ is a Lie group representation, so the claim follows from the commutative diagram Eq. (5). The assumption that ϕ is a subrepresentation of ρ is irrelevant. □

The Haar probability distribution on $\mathsf{SU}(\mathsf{N})$, which we denote as $\mathrm{d}U$, does give a distribution on $\mathfrak{su}(\mathsf{N})$ using the fact that the exponential map is one-to-one on the open ball of radius π at the origin in the Schatten ∞-norm and is almost onto from that restricted domain, but this is not very enlightening. However, relevant averages can be succinctly described in terms of subrepresentations.

Proposition II.4. *For any finite dimensional unitary representation ρ of a compact Lie group G, the integral $\int_G \rho(U)\mathrm{d}U$ with respect to the Haar measure is the orthogonal projector onto the trivial subrepresentation subspace of ρ. In particular, for any integer $t \geq 1$ the Haar average*

$$\mathcal{H}_t = \int_{\mathsf{SU}(\mathsf{N})} (U \otimes \bar{U})^{\otimes t} \mathrm{d}U$$

is the orthogonal projector onto the trivial subrepresentation subspace of $\tau : U \mapsto (U \otimes \bar{U})^{\otimes t}$ within $(\mathbb{C}^{\mathsf{N}})^{\otimes 2t}$.

Note that in Theorem I.1 we denoted the Haar random mixed unitary channel by $\mathcal{H}_t$, but here we overload the notation to mean its vectorized map. The representation ρ does not have to be finite dimensional, but we do not discuss any infinite dimensional representations in this paper.

Proof. Let $\mathcal{H} = \int_G \rho(U)\mathrm{d}U$. Since the Haar measure is left invariant, the represented unitary $\rho(V)$ for any $V \in G$ acts by the identity on the image of $\mathcal{H}$. It follows that $\mathcal{H}^2 = \mathcal{H}$. Since ρ is a unitary representation, $\mathcal{H}^\dagger = \int_G \rho(U^{-1})\mathrm{d}U$. Since $U \mapsto U^{-1}$ is a measure-preserving homeomorphism of G onto itself, $\mathcal{H}^\dagger = \mathcal{H}$. Let $\mathcal{V}$ be the representation space. The trivial representation subspace is $\mathcal{T} = \{v \in \mathcal{V} : \rho(U)v = v,\ \forall U \in G\}$. If $v = \mathcal{H}w$ for some $w \in \mathcal{V}$, then $\rho(U)v = \rho(U)\mathcal{H}w = \mathcal{H}w = v$, and hence $v \in \mathcal{T}$. So, the image of $\mathcal{H}$ is contained in the trivial representation. If $v \in \mathcal{T}$, then $\mathcal{H}v = v$, showing that v is in the image of $\mathcal{H}$. □

III. Random Pauli rotations

Now we consider more concrete distributions on $\mathfrak{su}(\mathsf{N})$ where $\mathsf{N} = 2^{\mathsf{n}}$ for an integer $\mathsf{n} \geq 1$.

Definition III.1. For any $P \in \mathbf{P}_{\mathsf{n}}$ we define a distribution, called a *random Pauli rotation by P*, as the uniform probability measure on $\{\mathrm{i}\theta P/2 \in \mathfrak{su}(\mathsf{N}; \mathbb{R}) \,|\, \theta \in (-\pi, \pi) \subset \mathbb{R}\}$. A *random Pauli rotation* with respect to a discrete probability distribution $\{(P, \Pr[P]) | P \in \mathbf{P}_{\mathsf{n}}\}$ on $\mathbf{P}_{\mathsf{n}}$ is the probabilistic mixture $\sum_P \Pr[P]\mu_P$ of random Pauli rotations μ_P by P.

We will only use the uniform distribution over $\mathbf{P}_{\mathsf{n}}$, but it may be helpful to proceed with a general distribution on $\mathbf{P}_{\mathsf{n}}$.

Lemma III.2. *For a random Pauli rotation $\mu = \sum_P \Pr[P]\mu_P$, the average operator $\mathcal{C}_{\mu,t}$ in Eq. (6) restricted*

to a subrepresentation $\phi_* : \mathfrak{su}(\mathsf{N}) \to \mathrm{Aut}(V)$ *of* $\mathfrak{su}(\mathsf{N})$ *within the tensor representation* $\tau : U \mapsto (U \otimes \bar{U})^{\otimes t}$*, simplifies as*

$$\begin{aligned}\mathcal{C}_{\mu,t}|_V &= \int_{\mathfrak{su}(\mathsf{N})} \exp(\mathrm{i}\phi_*(X))\mu(X)\mathrm{d}X \\ &= \sum_P \Pr[P]K(\phi_*(P/2))\end{aligned}$$

where $K(H)$ *for any hermitian operator* H *is the orthogonal projector onto* $\ker H$*, the eigenspace of eigenvalue zero.*

Proof. The first equality is noted in Lemma II.3. For the second equality, it suffices to evaluate $\mathcal{C}_{\mu_P,t}|_V$ for a random Pauli rotation μ_P by P. We have observed in Lemma II.1 that all the eigenvalues of $\tau_*(P/2)$ are integers. A subrepresentation of τ_* is nothing but a block-diagonal piece of τ_* after a unitary basis change on $(\mathbb{C}^{\mathsf{N}})^{\otimes 2t}$, so the eigenvalues of $\phi_*(P/2)$ can only be a subset of those of $\tau_*(P/2)$. Hence, the average of $e^{\mathrm{i}\theta\phi_*(P/2)}$ over θ eliminates all eigenspaces of nonzero eigenvalues. □

Remark III.3. A Pauli operator P, a tensor product of hermitian Pauli matrices, is a traceless unitary of eigenvalues ± 1. Observe that $\mathrm{i}P$ is a member of $\mathsf{SU}(\mathsf{N})$ and also of $\mathfrak{su}(\mathsf{N};\mathbb{R})$ where N is a power of 2. For example, $\mathrm{i}\sigma^z \in \mathsf{SU}(2) \cap \mathfrak{su}(2;\mathbb{R})$ and $\mathrm{i}\sigma^z \otimes \sigma^z \in \mathsf{SU}(4) \cap \mathfrak{su}(4;\mathbb{R})$. If ρ is a Lie group representation map, and ρ_* is the derived Lie algebra representation map, then we may consider $\rho(\mathrm{i}P)$ and $\rho_*(\mathrm{i}P)$ both of which are some matrices of the same dimension. Generally, $\rho(\mathrm{i}P) \neq \rho_*(\mathrm{i}P)$. However, since $P^2 = \mathbf{1}$, we instead have $\exp(\mathrm{i}\pi P/2) = \cos(\pi/2)\mathbf{1} + \mathrm{i}\sin(\pi/2)P = \mathrm{i}P$ and therefore by Eq. (5) we have

$$\begin{aligned}\exp(\mathrm{i}\pi\rho_*(P)/2) &= \exp(\rho_*(\mathrm{i}\pi P/2)) \\ &= \rho(\exp(\mathrm{i}\pi P/2)) = \rho(\mathrm{i}P).\end{aligned}$$

Proposition III.4. *For any random Pauli rotation with respect to* $\{(P, \Pr[P])\}$ *at order* t*, we have*

$$\|\mathcal{C}_{\mu,t} - \mathcal{H}_t\| = \max_\phi \left\| \sum_P \Pr[P]K(\phi_*(P/2)) \right\|$$

where ϕ *ranges over all irreducible nontrivial subrepresentations of the tensor representation* $\tau : U \mapsto (U \otimes \bar{U})^{\otimes t}$*.*

Proof. Immediate from Lemma III.2 and Proposition II.4. □

It is known [17, Lemma 3.7] that the spectral gap, $1 - \|\mathcal{C}_{\mu,t} - \mathcal{H}_t\|$, is positive if $\{(P, \Pr[P])\}$ induces a "densely generating" distribution on $\mathsf{SU}(\mathsf{N})$.

The motivation for us to consider the quantum circuit of random Pauli rotations is its simple implementation:

Proposition III.5. *Suppose that for an* n*-qubit system,* CNOT *can be applied only across a set of unordered pairs of qubits. This defines an undirected simple graph ("connectivity graph") over qubits, which we assume is connected. For any* $P \in \mathbf{P}_\mathsf{n}$ *and* $\theta \in \mathbb{R}$*, a unitary* $e^{\mathrm{i}\frac{\theta}{2}P}$ *can be implemented using (1) one* 1*-qubit Pauli* X *rotation* $e^{\mathrm{i}\frac{\theta}{2}\sigma^x}$*, (2) at most* $2\mathsf{n}$ 1*-qubit Hadamard and Phase gates, and (3) at most* $2\mathsf{n} - 2$ CNOT *and* SWAP *gates.*

Proof. It suffices to find a sequence of gates that maps $e^{\mathrm{i}\frac{\theta}{2}P}$ to $e^{\mathrm{i}\frac{\theta}{2}\sigma^x} \otimes \mathbf{1}_2^{\otimes(\mathsf{n}-1)}$ by conjugation. We first apply Hadamard and Phase gates by conjugation to obtain $e^{\mathrm{i}\frac{\theta}{2}Q}$ where $Q = Q_1 \otimes \cdots \otimes Q_\mathsf{n}$ is a tensor product of σ^x's and $\mathbf{1}_2$'s. In the connectivity graph, we assign each node a binary value corresponding to the support of Q, *i.e.*, $v_i = 1$ if and only if $Q_i = \sigma^x$. Every connected graph has a spanning tree. For each edge $(v_{\text{parent}}, v_{\text{child}})$ in the spanning tree such that all the children of v_{child} are zeros, if $v_{\text{parent}} = v_{\text{child}} = 1$, apply a CNOT gate by conjugation on the corresponding two qubits. If $v_{\text{parent}} = 0$ and $v_{\text{child}} = 1$, apply a SWAP gate by conjugation on the corresponding two qubits. Both operations will result in $v_{\text{parent}} = 1$, $v_{\text{child}} = 0$, *i.e.*, all the children of v_{parent} are zeros. This procedure terminates when the only nonzero node is the root, which corresponds to $e^{\mathrm{i}\frac{\theta}{2}\sigma^x} \otimes \mathbf{1}_2^{\otimes(n-1)}$. The total number of CNOT and SWAP gates applied is at most twice the number of edges in the spanning tree, which is $2(\mathsf{n} - 1)$. □

Corollary III.6. *With all-to-all connectivity, for any* $P \in \mathbf{P}_\mathsf{n}$ *and* $\theta \in \mathbb{R}$*, the unitary* $e^{\mathrm{i}\frac{\theta}{2}P}$ *can be implemented using* $\mathcal{O}(\mathsf{n})$ 1*-qubit and* CNOT *gates in circuit depth* $\mathcal{O}(\log \mathsf{n})$*. See Fig. 1 for an example.*

IV. The special case of $\mathfrak{su}(2)$

There are only three Pauli operators $\sigma^x, \sigma^y, \sigma^z$ (up to real scalars) in $\mathfrak{su}(2)$ so a random Pauli rotation is specified by $\Pr[\sigma^x], \Pr[\sigma^y], \Pr[\sigma^z]$. The goal is clear in Proposition III.4. With an $\mathfrak{su}(2)$-irrep ϕ_* in mind, we just write $J_{x,y,z}$ to mean $\phi_*(\sigma^{x,y,z}/2)$. Note the factor of 2 in the denominator, which gives $[J_a, J_b] = \mathrm{i}J_c$ where (a, b, c) is a cyclic permutation of (x, y, z). We have to calculate the spectral norm of

$$\Pr[\sigma^x]K(J_x) + \Pr[\sigma^y]K(J_y) + \Pr[\sigma^z]K(J_z)$$

for all irreps that appear in the tensor representation $\tau : \mathsf{SU}(2) \ni U \mapsto (U \otimes \bar{U})^{\otimes t}$. Since the dual of the defining irrep of $\mathfrak{su}(2)$ is equivalent to itself, the representation τ is simply the $2t$-fold tensor product of the defining irrep of $\mathfrak{su}(2)$. It is a standard fact that every irrep that appears in τ is odd dimensional because $2t$ is an even number, and every irrep of dimension $2\ell + 1$ appears in τ where $\ell \leq t$.

Lemma IV.1. *Each of* $K(J_x), K(J_y), K(J_z)$ *has rank* 1 *on any nontrivial* $\mathfrak{su}(2)$*-irrep of odd dimension* $2\ell + 1$ *for an integer* $\ell \geq 1$*. There exist normalized vectors* $|x\rangle \in \ker J_x, |y\rangle \in \ker J_y, |z\rangle \in \ker J_z$ *such that*

$$\langle x|y\rangle = \langle y|z\rangle = \langle z|x\rangle = \begin{cases} \frac{(-1)^{\ell/2}}{2^\ell}\binom{\ell}{\ell/2} & \text{if } \ell \text{ is even,} \\ 0 & \text{otherwise.} \end{cases}$$

In view of $\mathfrak{so}(3) \cong \mathfrak{su}(2)$, an odd dimensional irrep is often called an "integer spin" representation, and an even dimensional irrep a "half-integer spin" representation. For

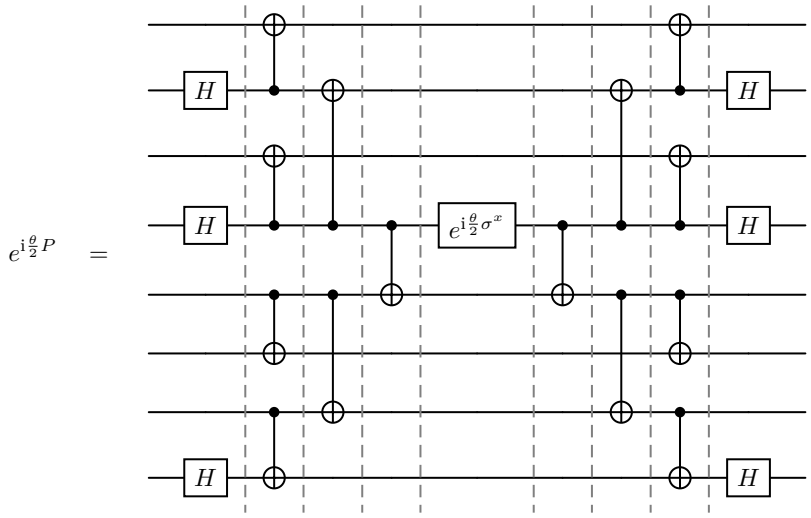

Fig. 1. Implementation of $e^{\mathrm{i}\frac{\theta}{2}P}$ ($P \in \mathbf{P}_\mathsf{n}$) by an $\mathcal{O}(\log \mathsf{n})$ depth circuit. The example corresponds to the Pauli string XZXZXXXZ. Gates between two dashed lines are implemented in parallel.

even dimensional irreps, it is well known that $K(J_x) = K(J_y) = K(J_z) = 0$.

Proof. Let ϕ_* be an $\mathfrak{su}(2)$-irreducible representation map acting on $V \cong \mathbb{C}^{2\ell+1}$. It is a standard fact that J_z has eigenvalues $\ell, \ell-1, \ldots, -\ell+1, -\ell$, each with multiplicity 1. Hence, the kernel is one-dimensional.

(*First method by symmetric powers*) Note that $U = \exp(-\mathrm{i}\frac{\pi}{3\sqrt{3}}(\sigma^x + \sigma^y + \sigma^z)) \in \mathsf{SU}(2)$ acts by conjugation as $\sigma^x \mapsto \sigma^y \mapsto \sigma^z \mapsto \sigma^x$. For a normalized vector $|z\rangle \in \ker J_z \subset V$, the vector $|x\rangle = \phi(U)\,|z\rangle$ spans $\ker J_x$ and $|y\rangle = \phi(U^2)\,|z\rangle$ spans $\ker J_y$. So, the three inner products in the claim are the same. It remains to calculate $\langle z|x\rangle = \langle z|\,\phi(U)\,|z\rangle$.

A concrete expression for ϕ is obtained by considering the 2ℓ-fold symmetric power of the defining representation of $\mathfrak{su}(2)$ [16, (11.8)]. Let $|0\rangle$ and $|1\rangle$ be a basis of $\mathbb{C}^2$ such that $\sigma^z\,|0\rangle = +\,|0\rangle$ and $\sigma^z\,|1\rangle = -\,|1\rangle$. Then,

$$U = \frac{e^{-\mathrm{i}\pi/4}}{\sqrt{2}}\begin{pmatrix}1 & -\mathrm{i}\\ 1 & \mathrm{i}\end{pmatrix}.$$

Written in terms of vectors in $(\mathbb{C}^2)^{\otimes 2\ell}$, a set of basis vectors of V can be chosen to be

$$\binom{2\ell}{k}^{-1/2} \sum_{w\in\{0,1\}^{2\ell}:|w|=k} |w\rangle$$

for $k = 0, 1, 2, \ldots, 2\ell$, where $|w|$ is the number of 1's in the bitstring w of length 2ℓ. These are eigenvectors of J_z with eigenvalues $\ell - k$. So,

$$|z\rangle = \binom{2\ell}{\ell}^{-1/2} \sum_{w:|w|=\ell} |w\rangle\,.$$

Applying $\phi(U) = U^{\otimes 2\ell}\big|_V$, we will obtain $|x\rangle$. With $|+\rangle = 2^{-1/2}(|0\rangle + |1\rangle)$ and $|-\rangle = 2^{-1/2}(|0\rangle - |1\rangle)$, we have $U\,|0\rangle = e^{-\mathrm{i}\pi/4}\,|+\rangle$ and $U\,|1\rangle = -e^{\mathrm{i}\pi/4}\,|-\rangle$. Hence,

$$\begin{aligned}\langle z|x\rangle &= \langle z|\,U^{\otimes 2\ell}\,|z\rangle\\ &= \sum_{w\in\{0,1\}^{2\ell}:|w|=\ell} \langle w|\,(-1)^\ell\,|+\rangle^{\otimes\ell}|-\rangle^{\otimes\ell}\end{aligned}$$

where the second equality is because both $|z\rangle$ and $\langle z|$ are invariant under permutations of tensor factors. For a bitstring w of length 2ℓ with $|w| = \ell$, let m be the number of 1's in the last ℓ bits. Then, $\langle w|\,+^\ell\,-^\ell\rangle = 2^{-\ell}(-1)^m$. There are $\binom{\ell}{\ell-m}\binom{\ell}{m}$ such bitstrings, and m ranges from 0 to ℓ. So,

$$\langle z|x\rangle = \frac{(-1)^\ell}{2^\ell}\sum_{m=0}^{\ell}(-1)^m\binom{\ell}{\ell-m}\binom{\ell}{m}.$$

The sum is the coefficient of h^ℓ in a polynomial $(1+h)^\ell(1-h)^\ell = (1-h^2)^\ell$ in a variable h. There is no h^ℓ term if ℓ is odd, implying that the sum is zero. If ℓ is even, then the coefficient is $(-1)^{\ell/2}\binom{\ell}{\ell/2}$. This completes the proof.

(*Second method using raising and lowering operators*) Let $|\ell\rangle$ be an eigenstate of J_z with eigenvalue ℓ: $J_z\,|\ell\rangle = \ell\,|\ell\rangle$. Define $J^+ = J_x + \mathrm{i}J_y$ and $J^- = J_x - \mathrm{i}J_y$, and inductively $J^-\,|k\rangle = a_{-k}\,|k-1\rangle$ for $k = \ell, \ell-1, \ldots, -\ell+1$ where

$$a_k = \sqrt{\ell(\ell+1) - k(k+1)} = a_{-k-1}.$$

Here, $|0\rangle = |z\rangle$. It is straightforward to check that $J^+\,|k\rangle = a_k\,|k+1\rangle$; the vector $|\ell+1\rangle$ is never defined, but $a_\ell = 0$.

Since $J_y\,|y\rangle = 0$, we have $J^+\,|y\rangle = J^-\,|y\rangle$. This implies that $|\ell-1\rangle$ cannot be a nonzero component of $|y\rangle$, which implies, in turn, that $|\ell-2j-1\rangle$ for any integer j cannot be a nonzero component of $|y\rangle$. Hence, $|y\rangle$ is in the span

of $|\ell\rangle, |\ell-2\rangle, \ldots, |-\ell+2\rangle, |-\ell\rangle$. In particular, if ℓ is odd, $\langle z|y\rangle = \langle 0|y\rangle = 0$.

Suppose $\ell = 2p$, an even integer, and put $|y\rangle = \sum_{k=-p}^{p} c_k |2k\rangle$. Then, the equation $J^+|y\rangle = J^-|y\rangle$ implies that

$$c_k a_{2k} = c_{k+1} a_{-2k-2} = c_{k+1} a_{2k+1}. \tag{7}$$

One can verify that $c_k = c_{-k}$ and

$$\begin{aligned}\frac{a_{2k}^2}{a_{2k+1}^2} &= \frac{2p(2p+1)-2k(2k+1)}{2p(2p+1)-(2k+1)(2k+2)} \\ &= \frac{(p-k)(2p+2k+1)}{(p+k+1)(2p-2k-1)} \\ &= \frac{(p-k)^2(2p+2k+2)(2p+2k+1)}{(2p-2k)(2p-2k-1)(p+k+1)^2} \\ &= \frac{\binom{2p-2k-2}{p-k-1}\binom{2p+2k+2}{p+k+1}}{\binom{2p-2k}{p-k}\binom{2p+2k}{p+k}}.\end{aligned}$$

Therefore,

$$\frac{|c_k|^2}{|c_0|^2} = \frac{\binom{2p-2k}{p-k}\binom{2p+2k}{p+k}}{\binom{2p}{p}^2}.$$

Since $\langle y|y\rangle = 1$, we must have

$$\begin{aligned}1 &= \sum_{k=-p}^{p} |c_k|^2 \\ &= \frac{|c_0|^2}{\binom{2p}{p}^2} \sum_{k=-p}^{p} \binom{2p-2k}{p-k}\binom{2p+2k}{p+k} \\ &= \frac{|c_0|^2}{\binom{2p}{p}^2} 4^{2p},\end{aligned}$$

where the last equality follows from a combinatorial identity[5] [18]

$$\sum_{i=0}^{n} \binom{2i}{i}\binom{2n-2i}{n-i} = 4^n.$$

This shows that $|\langle y|z\rangle| = |c_0| = \frac{1}{4^p}\binom{2p}{p}$ if $\ell = 2p$.

The complex phase α of $\langle y|z\rangle = \alpha \frac{1}{4^p}\binom{2p}{p}$ is not fixed by the normalization, but $\langle x|y\rangle\langle y|z\rangle\langle z|x\rangle$ is well defined regardless of α. To evaluate this product of the three inner products, we may use any normalized vectors in the kernels. A vector $|x\rangle \in \ker J_x$ can be computed by solving $J^+|x\rangle = -J^-|x\rangle$. By completely parallel calculation, we find a solution $|x\rangle = \sum_{k=-p}^{p} (-1)^k c_k |2k\rangle$. Then,

$$\begin{aligned}\beta = \langle x|y\rangle\langle y|z\rangle\langle z|x\rangle &= \langle x|y\rangle\langle y|0\rangle\langle 0|x\rangle \\ &= |c_0|^2 \sum_{k=-p}^{p} (-1)^k |c_k|^2.\end{aligned}$$

[5]This can be proved by, for example, a formula $(1-4h)^{-1/2} = \sum_{n=0}^{\infty}\binom{2n}{n}h^n$ and its square, where h is a variable.

This is a real number, which means that we may take $\alpha = \pm 1 = \beta/|\beta|$. We know that $|\beta| = |c_0|^3 < 1$, so $\sum_{k=-p}^{p}(-1)^k|c_k|^2 = \pm|c_0|$.

From Eq. (7) we know that $|c_0| < |c_1| < \cdots < |c_p|$. Suppose p is odd and $\beta > 0$. Then, we must have $\sum_{k=-p}^{p}(-1)^k|c_k|^2 = |c_0|$, which gives a contradiction: $0 = (-|c_0| - |c_1|^2) + (|c_0|^2 - |c_1|^2) + 2(|c_2|^2 - |c_3|^2) + \cdots + 2(|c_{p-1}|^2 - |c_p|^2) < 0$. Therefore, $\beta < 0$ if p is odd. Similarly, suppose p is even and $\beta < 0$. Then, we must have $\sum_{k=-p}^{p}(-1)^k|c_k|^2 = -|c_0|$, which gives a contradiction: $0 = (|c_0|+|c_0|^2)+2(-|c_1|^2+|c_2|^2)+\cdots+2(-|c_{p-1}|^2+|c_p|^2) > 0$. Therefore, $\beta > 0$ if p is even. This completes the proof. □

Corollary IV.2. *For a random Pauli rotation μ with respect to $\{(\sigma^x, \frac{1}{3}), (\sigma^y, \frac{1}{3}), (\sigma^z, \frac{1}{3})\}$ we have*

$$\|\mathcal{C}_{\mu,t} - \mathcal{H}_t\| = \frac{1}{12} \cdot \begin{cases} 4 & (t=1) \\ 6 & (t=2,3) \\ 7 & (t \geq 4) \end{cases}.$$

Proof. We calculated $f^2 = \mathrm{Tr}(K(J_x)K(J_y))$, etc, in Lemma IV.1, where $(-1)^{\ell/2} f = \binom{\ell}{\ell/2}/2^\ell$ is a decreasing function in even integer ℓ. Hence by Proposition III.4, there are only three cases to check: $f = 0$ if ℓ is odd, $f = -1/2$ if $\ell = 2$, and $f = 3/8$ if $\ell = 4$.

Let $M = K(J_x) + K(J_y) + K(J_z)$ and $a, b, c \in [0,3] \subset \mathbb{R}$ be the eigenvalues of M. Then $\mathrm{Tr}(M) = a+b+c = 3$, $\mathrm{Tr}(M^2) = a^2+b^2+c^2 = 3+6f^2$, and $\mathrm{Tr}(M^3) = a^3+b^3+c^3 = 3+18f^2+6f^3$. Calculation gives $a = b = 1-f$ and $c = 1+2f$ up to permutations, so $\|M\| = \max(1-f, 1+2f)$. This norm reaches the maximum $7/4$ when $\ell = 4$ and $f = 3/8$. □

V. A spectral gap bound by quadratic Casimir invariants

Theorem V.1. *Let μ be the random Pauli rotation with respect to the uniform distribution on $\mathbf{P}_\mathsf{n}$. Then, with $\mathsf{N} = 2^\mathsf{n}$ and for any integer $t \geq 1$,*

$$\|\mathcal{C}_{\mu,t} - \mathcal{H}_t\| \leq 1 - \frac{1}{4t}\frac{\mathsf{N}^2}{\mathsf{N}^2-1} - \frac{1}{\mathsf{N}^2-1}. \tag{8}$$

If $t \leq \mathsf{N}/2$, then

$$\|\mathcal{C}_{\mu,t} - \mathcal{H}_t\| \leq 1 - \frac{1}{2t}\frac{\mathsf{N}(\mathsf{N}-t+1)}{\mathsf{N}^2-1}. \tag{9}$$

Comparing with Corollary IV.2, we see that with $\mathsf{N} = 2$ the inequality in Eq. (8) is saturated if and only if $t \in \{1,2,4\}$.

Proof. It follows from Proposition III.4 that

$$\|\mathcal{C}_{\mu,t} - \mathcal{H}_t\| = \max_J \left\| \mathop{\mathbb{E}}_{P \in \mathbf{P}_\mathsf{n}} K(J_P) \right\|$$

where the maximum is over all nontrivial $\mathfrak{su}(\mathsf{N})$-irreps that appear in the tensor representation $\tau : U \mapsto (U \otimes \bar{U})^{\otimes t}$. By Eq. (3), the norm ℓ of a represented operator J_P in a

nontrivial $\mathfrak{su}(\mathsf{N})$-irrep is a nonzero integer and is at most t. We use an operator inequality

$$K(H) \preceq \mathbf{1} - H^2/\|H\|^2 \tag{10}$$

which holds for any nonzero hermitian operator H, where $K(H)$ is the orthogonal projector onto $\ker H$. Averaging over $\mathbf{P}_\mathsf{n}$, we have that

$$\mathop{\mathbb{E}}_{P\in\mathbf{P}_\mathsf{n}} K(J_P) \preceq \mathbf{1} - \mathop{\mathbb{E}}_{P\in\mathbf{P}_\mathsf{n}} \frac{J_P^2}{\|J_P\|^2}. \tag{11}$$

It is well known (by a more general argument) that the last term is a scalar multiple of the identity, called a Casimir operator or invariant, where the scalar depends only on the irrep. We give an elementary calculation to this end in Lemma V.2 below. By the lower bound in Lemma V.3 below,

$$\mathbf{1} - \mathop{\mathbb{E}}_{P\in\mathbf{P}_\mathsf{n}} \frac{J_P^2}{\|J_P\|^2} \preceq \mathbf{1}\left(1 - \frac{1}{4\ell}\frac{\mathsf{N}^2}{\mathsf{N}^2-1} - \frac{1}{\mathsf{N}^2-1}\right)$$

where $\ell = \|J_P\|$ is independent of P. Since $\ell \le t$, we complete the proof of Eq. (8). For Eq. (9), we use Lemma V.4 further below. □

Lemma V.2. *For any $\mathfrak{su}(\mathsf{N})$-irrep ϕ_*, the quadratic sum $\sum_{P\in\mathbf{P}_\mathsf{n}} \phi_*(P/2)^2$ is a scalar multiple of the identity.*

Proof. Abbreviate $\phi_*(P/2)$ by J_P. We show by direct calculation that $\sum_P J_P^2$ commutes with J_Q for all $Q \in \mathbf{P}_\mathsf{n}$. Since the commutator obeys the Leibniz rule, we have $[J_Q, J_P^2] = J_P[J_Q, J_P] + [J_Q, J_P]J_P$. This may be nonzero only if $PQ = -QP = \pm \mathrm{i} R$ for some $R \in \mathbf{P}_\mathsf{n}$. For an anticommuting pair P, Q, the Pauli operator R also anticommutes with each of P, Q. So, the subset of all Pauli operators that anticommute with Q is partitioned into unordered pairs $\{P, R\}$ where $[J_P, J_Q] = \mathrm{i} J_R$. That is, for each pair $\{P, R\}$, the three elements J_P, J_Q, J_R span $\mathfrak{su}(2)$. Then,

$$\begin{aligned}
&[J_Q, J_P^2 + J_R^2] \\
&= J_P[J_Q, J_P] + [J_Q, J_P]J_P + J_R[J_Q, J_R] + [J_Q, J_R]J_R \\
&= J_P(-\mathrm{i}J_R) + (-\mathrm{i}J_R)J_P + J_R(\mathrm{i}J_P) + (\mathrm{i}J_P)J_R \\
&= 0.
\end{aligned} \tag{12}$$

Since we are working with an irrep, Schur's lemma implies that $A = \sum_P J_P^2$ is proportional to the identity. □

Lemma V.3. *Let $J_P = \phi_*(P/2)$ be the represented operator in an $\mathfrak{su}(\mathsf{N})$-irrep ϕ_* for any Pauli operator $P \in \mathbf{P}_\mathsf{n}$ where $\mathsf{N} = 2^\mathsf{n} \ge 2$, and let $\ell = \|J_P\|$ be the Schatten ∞-norm, which is independent of P. Then,*

$$\left(\frac{\mathsf{N}^2\ell}{4} + \ell^2\right)\mathbf{1} \preceq \sum_{P\in\mathbf{P}_\mathsf{n}} J_P^2 \preceq \left(\frac{\mathsf{N}(\mathsf{N}-1)\ell}{2} + (\mathsf{N}-1)\ell^2\right)\mathbf{1}.$$

For any $\mathsf{n} \in \mathbb{Z}_{>0}$ and $\ell \in \frac{1}{2}\mathbb{Z}_{>0}$, there is an $\mathfrak{su}(2^\mathsf{n})$-irrep that saturates the upper bound. For any $\mathsf{n}, k \in \mathbb{Z}_{>0}$, there is an $\mathfrak{su}(2^\mathsf{n})$-irrep with $\ell = 2^{\mathsf{n}-2}k$ that saturates the lower bound. The saturating irreps are unique up to isomorphisms.

In this lemma it is not required that ϕ_* is a subrepresentation of a tensor representation $\tau : U \mapsto (U \otimes \bar{U})^{\otimes t}$.

The lower bound proof can be understood using just the representation theory of $\mathfrak{su}(2)$.

Proof of the lower bound. The norm $\ell = \|J_P\|$ is independent of P by Lemma II.2. Lemma V.2 says that $A = \sum_P J_P^2$ is a scalar multiple of the identity. We have to estimate *the* eigenvalue of A. It suffices to examine the action of A on any vector.

Let $Z_1 = \sigma^z \otimes \mathbf{1}_2^{\otimes(\mathsf{n}-1)} \in \mathbf{P}_\mathsf{n}$. Let $|\psi\rangle$ be any vector such that $J_{Z_1}|\psi\rangle = \ell|\psi\rangle$, where ℓ is the greatest eigenvalue. There are $4^{\mathsf{n}-1}$ unordered pairs $\{\sigma^x \otimes W, \sigma^y \otimes W\}$ where $W \in \{\mathbf{1}_2, \sigma^x, \sigma^y, \sigma^z\}^{\otimes(\mathsf{n}-1)}$ such that $\mathbb{C}\{Z_1, \sigma^x \otimes W, \sigma^y \otimes W\} \cong \mathfrak{su}(2)$ as Lie algebras. We know if X, Y, Z are a triple generating $\mathfrak{su}(2)$ such that $[X, Y] = \mathrm{i}Z$ (and cyclic permutations thereof), then

$$\rho_*(X)^2 + \rho_*(Y)^2 + \rho_*(Z)^2 = \ell_\rho(\ell_\rho + 1)\mathbf{1}$$

for any irrep ρ_* where ℓ_ρ is the greatest eigenvalue of $\rho_*(Z)$. The linear span of all vectors obtained by acting with $J_{Z_1}, J_{\sigma^x\otimes W}, J_{\sigma^y\otimes W}$ on $|\psi\rangle$ is an $\mathfrak{su}(2)$-irrep because J_{Z_1} assumes the greatest eigenvalue ℓ on $|\psi\rangle$, and hence

$$\begin{aligned}
(J_{Z_1}^2 + J_{\sigma^x\otimes W}^2 + J_{\sigma^y\otimes W}^2)|\psi\rangle &= (\ell^2 + \ell)|\psi\rangle, \\
(J_{\sigma^x\otimes W}^2 + J_{\sigma^y\otimes W}^2)|\psi\rangle &= \ell|\psi\rangle.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\langle\psi| A |\psi\rangle &\ge \langle\psi| J_{Z_1}^2 |\psi\rangle + \sum_W \langle\psi| (J_{\sigma^x\otimes W}^2 + J_{\sigma^y\otimes W}^2)|\psi\rangle \\
&= \ell^2 + 4^{\mathsf{n}-1}\ell.
\end{aligned}$$

This proves the lower bound. □

The remainder of the proof uses highest weights.

Proof of the rest of the claims in Lemma V.3. To prove the upper bound we take $|\psi\rangle$ to be a highest weight vector. By definition, this means that $|\psi\rangle$ is annihilated by all positive roots of $\mathfrak{su}(\mathsf{N})$, which span the $\mathbb{C}$-linear space of all strictly upper triangular N-by-N matrices. This means that

$$(J_X + \mathrm{i}J_Y)|\psi\rangle = 0$$

for any $X + \mathrm{i}Y \in \mathfrak{su}(\mathsf{N})$ that is upper triangular in a standard basis for Pauli operators. Then,

$$\begin{aligned}
0 &= \langle\psi| (J_X - \mathrm{i}J_Y)(J_X + \mathrm{i}J_Y)|\psi\rangle \\
&= \langle\psi| (J_X^2 + J_Y^2 - J_Z)|\psi\rangle
\end{aligned} \tag{13}$$

where $\mathrm{i}Z = [X, Y]$. It follows that

$$\langle\psi| J_X^2 + J_Y^2 |\psi\rangle = \langle\psi| J_Z |\psi\rangle \le \ell. \tag{14}$$

To use this we partition $\mathbf{P}_n$ as follows. For a bitstring $z \in \{0,1\}^{k-1}$ of length $k-1$, define $Z(z) = \bigotimes_{j=1}^{k-1}(\sigma^z)^{z_j} \in \mathbf{P}_{k-1}$. For a $\mathbb{Z}_4$-string $w \in \{0,1,2,3\}^{\mathsf{n}-k}$, define $W(w) =$

$\bigotimes_{j=1}^{n-k} \sigma^{w_j}$ where $\sigma^0 = \mathbf{1}_2$, $\sigma^1 = \sigma^x$, $\sigma^2 = \sigma^y$, and $\sigma^3 = \sigma^z$. Then, a triple of

$$Z(z) \otimes \sigma^x \otimes W(w),$$
$$Z(z) \otimes \sigma^y \otimes W(w),$$
$$\mathbf{1}_2^{\otimes(k-1)} \otimes \sigma^z \otimes \mathbf{1}_2^{\otimes(n-k)}$$

forms an $\mathfrak{su}(2)$ subalgebra. So we have identified $\sum_{k=1}^{n} 2^{k-1}4^{n-k} = 4^n(2^{-1} - 2^{-n-1})$ subalgebras. The pairs $\{Z(z)\otimes\sigma^x\otimes W(w), Z(z)\otimes\sigma^y\otimes W(w)\}$ account for $4^n - 2^n$ elements of $\mathbf{P}_n$, and the remaining $2^n - 1$ operators of $\mathbf{P}_n$ are tensor products of $\mathbf{1}_2$ and σ^z. Therefore,

$$\sum_{P\in\mathbf{P}_n} \langle\psi| J_P^2 |\psi\rangle \le 4^n(2^{-1} - 2^{-n-1})\ell + (2^n - 1)\ell^2.$$

The tightness of the bounds can be shown by considering specific highest weights and using the partition above. The upper bound is saturated if and only if Eq. (14) is saturated for all nonidentity tensor product Z of $\mathbf{1}_2$ and σ^z. So, we need to show that such a linear functional on the Cartan subalgebra is a valid point on the weight lattice. This is easy: if L_j denotes the dual vector of the diagonal matrix with a sole 1 at the j-th diagonal, $2\ell L_1$ is the desired weight.

To prove that the lower bound can be saturated, we take a highest weight in which only J_{Z_1}, where $Z_1 = \sigma^z \otimes \mathbf{1}_2^{\otimes(n-1)}$, takes the greatest eigenvalue ℓ, but J_Z for any other diagonal Z assumes zero. This amounts to the weight $k\sum_{j=1}^{N/2} L_j$, giving $\ell = kN/4$ for some positive integer k. Then, on the highest weight vector $|\psi\rangle$, the generator Z_1 gives ℓ^2, and the 4^{n-1} pairs $\{\sigma^x\otimes W, \sigma^y\otimes W\}$ gives $4^{n-1}\ell$, but all other $2^n - 2$ diagonal generators give zero by the choice of the highest weight. Finally, Eq. (13) implies that the $4^n - 2^n - 2\cdot 4^{n-1}$ generators $Z(z)\otimes\sigma^x\otimes W(w), Z(z)\otimes\sigma^y\otimes W(w)$ give zero.

In those saturating conditions, we are forced to choose a unique highest weight given ℓ, which in turn determines the irrep (up to equivalence). □

Next, we present an alternative and slightly tighter bound, which assumes familiarity with highest weights for Lie algebra representations (e.g. [16, §14-15]). Note, again, that our proof of the lower bound in Lemma V.3 only uses the representation theory of $\mathfrak{su}(2)$.

A finite-dimensional irreducible representation of $\mathfrak{su}(N)$ is labeled by its highest weight $\sum_j \mu_j L_j$, which is labeled by a sequence of N integers $\mu = (\mu_1, \mu_2, \ldots, \mu_N)$ where $\mu_i \ge \mu_{i+1}$ modulo integer multiples of $(1, 1, \ldots, 1)$. We choose a representative μ such that $\sum_{i=1}^{N} \mu_i = 0$, which is possible for subrepresentations of $\tau : U \mapsto (U\otimes\bar{U})^{\otimes t}$, and henceforth the Killing form $\langle\cdot,\cdot\rangle$ on the dual of Cartan subalgebra is given by $\langle\mu,\mu'\rangle = \sum_i \mu_i\mu_i'$. (The normalization here is different from that in [16, §15].)

Let

$$H = \tfrac{1}{2}\sigma^z \otimes \mathbf{1}_2^{\otimes n-1} = \frac{1}{2}\operatorname{diag}(1, 1, \ldots, 1, -1, -1, \ldots, -1)$$

be an element of the Cartan subalgebra. We would like to determine $\ell = \|\phi_*(H)\|$. Since the set of weights are in the convex hull of the Weyl group orbit of μ, the maximum eigenvalue of $\phi_*(H)$ is given by $\ell = \max_{w\in\mathrm{Weyl}}(w\cdot\mu)(H) = \mu(H) = \frac{1}{2}\big(\mu_1 + \cdots + \mu_{N/2} - \mu_{N/2+1} - \cdots - \mu_N\big)$.

Next, we invoke a formula for the quadratic Casimir operator (e.g. [16, (25.14)]): For any finite-dimensional irrep ϕ_* with highest weight μ, and for any basis $\{X_j\}$ of $\mathfrak{su}(N)$ that is orthonormal with respect to the Killing form,

$$\sum_j \phi_*(X_j)^2 = (\langle\mu,\mu\rangle + \langle\mu,\delta\rangle)\mathbf{1},$$

where δ is the sum of all positive roots, which can be written in a vector form as $\delta = (N-1, N-3, \ldots, -(N-1))$. Therefore, $\frac{4}{N}\sum_{P\in\mathbf{P}_n} J_P^2 = (\langle\mu,\mu\rangle + \langle\mu,\delta\rangle)\mathbf{1}$, where the factor $4/N$ comes from renormalizing J_P to an orthonormal basis ($\mathrm{Tr}((P/2)^2) = N/4$). Now, our quantity of interest is exactly determined by the highest weight of a given irrep:

$$\frac{1}{\ell^2}\sum_{P\in\mathbf{P}_n} J_P^2 = \frac{N(\langle\mu,\mu\rangle + \langle\mu,\delta\rangle)}{\big(\mu_1 + \cdots + \mu_{N/2} - \mu_{N/2+1} - \cdots - \mu_N\big)^2}\mathbf{1}. \tag{15}$$

Lemma V.4. *Assume $t \le N/2$. Let ϕ_* be an irreducible $\mathfrak{su}(N)$-subrepresentation of $\tau : U \mapsto (U\otimes\bar{U})^{\otimes t}$. Let $\ell = \|\phi_*(P/2)\|$ be the Schatten ∞-norm, which is independent of $P \in \mathbf{P}_n$.*

$$\frac{1}{\ell^2}\sum_{P\in\mathbf{P}_n} J_P^2 \succeq \frac{N(N-t+1)}{2t}\mathbf{1}.$$

There exists an irreducible subrepresentation ϕ_* of τ achieving the equality.

Roughly speaking, in the small-t regime where $t \ll N$, this gives a factor of 2 improvement relative to Lemma V.3. Note that this is consistent with the tightness stated in Lemma V.3, because the irreps that saturate the lower bound of Lemma V.3 are not subrepresentations of τ when $t \ll N$. If $t = N/2$, then this lemma gives the same bound as Lemma V.3.

The strategy in the proof below applies for $t > N/2$ and gives an alternative proof of the lower bound of Lemma V.3, but we will omit such calculation.

Proof. It only remains to perform an elementary calculation: minimize the right-hand side of Eq. (15) over the highest weights that correspond to irreducible subrepresentations of τ.

The decomposition of τ into irreps is well understood (e.g. [19, Theorem 4]). An irrep with highest weight μ is a subrepresentation of τ if and only if $\sum_i \mu_i = 0$ and $\sum_i |\mu_i| \le 2t$. When $t \le N/2$, the minimizer is given by

$$\mu_* = (1, 1, \ldots, 1, 0, 0, \ldots, 0, -1, -1, \ldots, -1)$$

with the first t entries being 1 and last t entries being -1. Calculating the right-hand side of Eq. (15) with respect to μ_* gives the stated bound. □

Remark V.5. If Eq. (11) is saturated, then $t \in \{1, 2, 4\}$. If $\ell \notin \{1, 2, 4\}$, then we know by comparing Corollary IV.2 with Eq. (8) that $K(J_x)+K(J_y)+K(J_z)+(J_x^2+J_y^2+J_z^2)/\ell^2$ has norm strictly smaller than 3 in any nontrivial irrep of $\mathfrak{su}(\mathsf{N})$ with $\|J_x\| = \ell$ for any triple J_x, J_y, J_z that form an $\mathfrak{su}(2)$ subalgebra.

Remark V.6. It may be significant underestimation of the spectral gap using $K(H) \preceq \mathbf{1} - H^2/\|H\|^2$. Take the case of $\mathfrak{su}(2)$ and consider a random Pauli rotation by $\{(\sigma^x, \frac{1}{2}), (\sigma^y, \frac{1}{2})\}$. In the $(2\ell+1)$-dimensional $\mathfrak{su}(2)$-irrep, since $J_x^2 + J_y^2 + J_z^2 = (\ell^2+\ell)\mathbf{1}$ and $J_z^2 \preceq \ell^2\mathbf{1}$, we see $(J_x^2 + J_y^2)/\ell^2 \succeq \mathbf{1}/\ell$, which is best possible since $\langle\psi| (J_x^2 + J_y^2) |\psi\rangle / \ell^2 = 1/\ell$ if $J_z |\psi\rangle = \ell |\psi\rangle$. This gives a lower bound on the spectral gap $\Omega(1/t)$. However, by the exact calculation in Lemma IV.1 above, we know that the spectral gap of this design is independent of t.

Proof of Theorem I.1. The spectral gap bound is proved in Theorem V.1. Note that a random Pauli rotation is defined by a uniform probability distribution on $\mathbf{P}_\mathsf{n} \times (-\pi, \pi)$ while the second statement of Theorem I.1 takes a uniform distribution on $\mathbf{P}_\mathsf{n} \times (-2\pi, 2\pi)$.

The discrepancy of a factor of 2 here is completely dismissable for the first statement of Theorem I.1 since we only need a range of θ in $\exp(\mathrm{i}\theta\phi_*(P/2))$ such that the average over θ of a represented operator $\phi_*(P/2)$ for any Pauli operator $P \in \mathbf{P}_\mathsf{n}$ is the projection onto the kernel of $\phi_*(P/2)$; for irreps ϕ that appear in $\tau : U \mapsto (U \otimes \bar{U})^{\otimes t}$, the eigenvalues of $\phi_*(P/2)$ are integers.

However, if we consider an arbitrary finite dimensional unitary representation ρ of $\mathsf{SU}(\mathsf{N})$, we are no more guaranteed that $\rho_*(P/2)$ has integer eigenvalues. Fortunately, every eigenvalue of $\rho_*(P/2)$ is half an integer for any unitary representation ρ because every finite dimensional unitary irrep of $\mathsf{SU}(\mathsf{N})$ is a subrepresentation of $U \mapsto U^{\otimes m}$ for some integer $m \geq 0$ [16, §15.3].

Note that for an irrep ρ_* where $2\|\rho_*(P/2)\|$ is an odd integer, the kernel of $\rho_*(P/2)$ is zero, implying that averaging over θ eliminates this irrep. □

Proof of Corollary I.2. Let $\mathcal{D} = \mathcal{C}_t^k - \mathcal{H}_t$ be the difference of the channels. With $k \geq (4\log 2)\mathsf{n}t^2 + 4t\log\frac{1}{\varepsilon}$, we have $\|\mathcal{D}\otimes\mathcal{I}\|_{2\to2} = \|\mathcal{D}\|_{2\to2} \leq (1-\frac{1}{4t})^k \leq \varepsilon 2^{-\mathsf{n}t}$ by Theorem I.1, where $\mathcal{I}$ means the identity channel on any auxiliary system. Since the diamond norm is obtained by taking an equal dimensional auxiliary system, $\|\mathcal{D}\|_\diamond = \|\mathcal{D} \otimes \mathcal{I}_{2^{\mathsf{n}t}}\|_{1\to1} \leq 2^{\mathsf{n}t}\|\mathcal{D} \otimes \mathcal{I}_{2^{\mathsf{n}t}}\|_{2\to2} \leq \varepsilon$, where the first equality is by [20, Theorem 11.1]. (The use of $2 \to 2$ norm for this purpose has appeared in [3], [21].)

Similarly, if $k \geq (4\log 8)\mathsf{n}t^2 + 4t\log\frac{1}{\varepsilon}$, we have $\left\|\mathcal{C}_t^k - \mathcal{H}_t\right\|_\diamond \leq \varepsilon 4^{-\mathsf{n}t}$ which implies $(1-\varepsilon)\mathcal{H}_t \preceq \mathcal{C}_t^k \preceq (1+\varepsilon)\mathcal{H}_t$ by [3, Lemma 3].

The gate complexity follows from Corollary III.6. □

VI. Orthogonal designs

Our approach can be adapted for special orthogonal groups. This section uses arguments parallel to those in the analysis for $\mathsf{SU}(\mathsf{N})$, so we will be rather brief.

A. Skew-symmetric Pauli operators

We consider the special orthogonal group $\mathsf{SO}(\mathsf{N})$ in a fashion similar to our random Pauli rotations. We directly use the inclusion $\mathsf{SO}(\mathsf{N}) = \mathsf{SU}(\mathsf{N}) \cap \mathbb{R}^{\mathsf{N}\times\mathsf{N}} \subset \mathsf{SU}(\mathsf{N})$ for $\mathsf{N} = 2^\mathsf{n}$.

Define a set $\mathbf{Y}_\mathsf{n}$ of Pauli operators with entries in $\mathrm{i}\mathbb{R}$:

$$\mathbf{Y}_\mathsf{n} = \{P \in \mathbf{P}_\mathsf{n} \mid \text{An odd number of } \sigma^y \text{ tensor factors appear in } P.\}$$

We first verify that the $\mathbb{R}$-linear span of $\mathrm{i}\mathbf{Y}_\mathsf{n}$ is precisely the real Lie algebra $\mathfrak{so}(\mathsf{N} = 2^\mathsf{n})$ consisting of all antisymmetric real matrices. It is clear that $\mathrm{i}\mathbf{Y}_\mathsf{n}$ is $\mathbb{R}$-linearly independent and consists of skew-symmetric real matrices. We can check that $|\mathbf{Y}_\mathsf{n}| = \mathsf{N}(\mathsf{N}-1)/2$ by solving a recursion equation as follows. Let $e(\mathsf{n})$ be the number of all Pauli operators $\{\mathbf{1}, \sigma^x, \sigma^y, \sigma^z\}^{\otimes\mathsf{n}}$ that contain an even number of tensor factors σ^y. The identity operator $\mathbf{1}_2^{\otimes\mathsf{n}}$ contributes 1 to $e(\mathsf{n})$. Consider a subset of Pauli "strings" whose first "letter" is σ^y, and another set of Pauli strings whose first letter is one of $\mathbf{1}, \sigma^x, \sigma^z$. It is then clear that $|\mathbf{Y}_{\mathsf{n}+1}| = 3|\mathbf{Y}_\mathsf{n}| + e(\mathsf{n})$ and $e(\mathsf{n}+1) = 3e(\mathsf{n}) + |\mathbf{Y}_\mathsf{n}|$ with initial conditions $e(1) = 3$ and $|\mathbf{Y}_1| = 1$. The claim $|\mathbf{Y}_\mathsf{n}| = 2^{\mathsf{n}-1}(2^\mathsf{n}-1)$ follows by induction in n.

Theorem VI.1. *Suppose* $\mathsf{N} = 2^\mathsf{n} > 4$. *For any integer* $t \geq 1$, *we have*

$$\left\| \mathop{\mathbb{E}}_{\theta\sim(-2\pi,2\pi),\, P\in\mathbf{Y}_\mathsf{n}} (e^{\mathrm{i}\theta P/2})^{\otimes t} - \mathop{\mathbb{E}}_{O\sim\mathsf{SO}(\mathsf{N})} O^{\otimes t} \right\| \leq 1 - \frac{1}{2t}\frac{\mathsf{N}-2}{\mathsf{N}-1} - \frac{2}{\mathsf{N}(\mathsf{N}-1)}.$$

The small orthogonal groups, $\mathsf{SO}(2)$ and $\mathsf{SO}(4)$, are excluded for simplicity of the proof as they are not simple Lie groups. Note that every finite dimensional unitary irrep[6] of $\mathsf{SO}(\mathsf{N})$ for $\mathsf{N} > 4$ appears as a subrepresentation of the tensor representation $O \mapsto O^{\otimes t}$ for some integer $t \geq 1$. This is explained in [16, §19]. Therefore, Theorem VI.1 implies that for any irrep of $\mathsf{SO}(\mathsf{N} = 2^\mathsf{n} > 4)$ the spectral gap is at least $1/\dim\mathsf{SO}(\mathsf{N})$.

A representation of $\mathsf{SU}(\mathsf{N})$ gives a representation of $\mathsf{SO}(\mathsf{N})$, but a representation of $\mathsf{SO}(\mathsf{N})$ does not in general give a representation of $\mathsf{SU}(\mathsf{N})$. So, this theorem cannot be thought of as a corollary of Theorem I.1. The proof below is however almost identical to that of Theorem I.1, mainly because we use only the common aspects of the representation theories of $\mathsf{SO}(\mathsf{N})$ and $\mathsf{SU}(\mathsf{N})$. Thus, we will assume a reader's familiarity with the proof of Theorem I.1,

[6] A representation of the Lie group $\mathsf{SO}(\mathsf{N})$ induces a representation of its Lie algebra $\mathfrak{so}(\mathsf{N})$, but not every representation of $\mathfrak{so}(\mathsf{N})$ arises in this way, and the seed for the remaining $\mathfrak{so}(\mathsf{N})$-irreps is spin representations. This phenomenon does not happen for $\mathsf{SU}(\mathsf{N})$. It does not seem to make sense to consider spin representations in orthogonal designs since the Haar average over $\mathsf{SO}(\mathsf{N})$ of a spin representation is ill-defined.

or rather the proofs of Theorem V.1 and Lemma V.3, and omit some detail.

Proof. (Step 0: to irreps) The Haar average (the second term in the norm) is the projector onto the trivial subrepresentation (Proposition II.4). Hence, we consider an irreducible nontrivial $\mathfrak{so}(\mathsf{N})$-subrepresentation ρ_* of $O \mapsto O^{\otimes t}$.

(Step 1: random angles give kernel projectors.) It is clear that $\rho_*(P/2)$ has half-integer eigenspectrum for any P, and therefore the average over $\theta \in (-2\pi, 2\pi)$ eliminates all nonzero eigenvalues: $\mathbb{E}_{\theta\sim(-2\pi,2\pi)} e^{\mathrm{i}\theta\rho_*(P/2)} = K(\rho_*(P/2))$ is the projector onto the kernel of $\rho_*(P/2)$.[7]

(Step 2: identical eigenspectra for all represented operators) The norm of $\rho_*(P/2)$ is independent of $P \in \mathbf{Y}_{\mathsf{n}}$: this is an analog of Lemma II.2 for $\mathfrak{so}(\mathsf{N})$, and the proof is similar. Note that for any matrix M, we have $\det(M \otimes \mathbf{1}_2) = \det(M \oplus M) = (\det M)^2$. So, if $O \in \mathsf{O}(\mathsf{N}/2)$, then $O \otimes \mathbf{1}_2 \in \mathsf{SO}(\mathsf{N})$. The Clifford unitaries, CNOT and Hadamard, are in $O(4)$. If there is a tensor factor $\sigma^y \otimes \sigma^y$ in some $P \in \mathbf{Y}_{\mathsf{n}}$, then we must have $\mathsf{n} \geq 3$ since P must contain an odd number of σ^y's. Hence, using CNOT and Hadamard Clifford unitary acting on those two 2×2 tensor factors, we can turn P by $\mathsf{SO}(\mathsf{N})$ conjugation into a Pauli that where $\sigma^y \otimes \sigma^y$ is replaced by $\sigma^x \otimes \sigma^x$ while not changing any other tensor factor of P. Inductively, we turn all pairs of σ^y tensor factors into pairs of σ^x. By the same argument, we can turn any σ^z tensor factor into σ^x. Now, the conjugation of $\sigma^y \otimes \mathbf{1}_2$ by CNOT is $\sigma^y \otimes \sigma^x$. Therefore, any $P \in \mathbf{Y}_{\mathsf{n}}$ with $\mathsf{n} \geq 3$ is congruent to $\sigma^y \otimes \mathbf{1}_2^{\otimes(\mathsf{n}-1)}$ by some element of $\mathsf{SO}(\mathsf{N})$. Hence for any $P \in \mathbf{Y}_{\mathsf{n}}$ where $\mathsf{n} \geq 3$, there exists $O \in \mathsf{SO}(\mathsf{N})$ such that $\rho(O)\rho_*(P)\rho(O)^{-1} = \rho_*(\sigma^y \otimes \mathbf{1}_2^{\otimes(\mathsf{n}-1)})$. Put $\ell = \|\rho_*(P/2)\| \leq t/2$ for any $P \in \mathbf{Y}_{\mathsf{n}}$.

(Step 3: to quadratic Casimir) Bounding the kernel projector by the quadratic operator (Eq. (4)), we are left with the problem of lower bounding

$$\mathop{\mathbb{E}}_{P\in\mathbf{Y}_{\mathsf{n}}} \rho_*(P/2)^2/\ell^2.$$

By a completely analogous calculation as in Eq. (12), this average is a scalar multiple of the identity.

(Step 4: find a large number of $\mathfrak{su}(2)$'s) To estimate the unique eigenvalue of this average, we look at a vector $|\psi\rangle$ such that $\rho_*(\frac{1}{2}\sigma^y \otimes \mathbf{1}_2^{\otimes(\mathsf{n}-1)})\,|\psi\rangle = \ell\,|\psi\rangle$. We can find $|\mathbf{Y}_{\mathsf{n}-1}|$ $\mathfrak{su}(2)$-subalgebras:

$$\sigma^y \otimes \mathbf{1}_2^{\otimes(\mathsf{n}-1)}, \quad \sigma^x \otimes W, \quad \sigma^z \otimes W,$$

where each triple of $\mathfrak{su}(2)$ generators is uniquely labeled by $W \in \mathbf{Y}_{\mathsf{n}-1}$. Hence,

$$\sum_{P\in\mathbf{Y}_{\mathsf{n}}} \rho_*(P/2)^2 \succeq \mathbf{1}\big(\ell^2 + |\mathbf{Y}_{\mathsf{n}-1}|\ell\big).$$

The theorem is proved since $\ell \leq t/2$. □

[7]We implicitly allowed complex coefficients for $\mathfrak{so}(\mathsf{N})$, i.e., the complexification. Exponentiated matrices are all real.

B. Skew-symmetric elementary matrix basis

We give another orthogonal design. In this subsection, we will not require N to be a power of 2.

Let $\mathsf{N} \geq 3$ be any integer. For any integers a, b $(1 \leq a, b \leq \mathsf{N})$, let $E_{a,b} = |a\rangle\langle b| - |b\rangle\langle a|$ denote the skew-symmetric $\mathsf{N} \times \mathsf{N}$ matrix in which there are only two nonzero matrix entries ± 1. Define

$$\mathbf{E}_{\mathsf{N}} = \big\{E_{a,b} \in \mathbb{R}^{\mathsf{N}\times\mathsf{N}} \;\big|\; 1 \leq a < b \leq \mathsf{N}\big\}$$

Clearly, $\mathbf{E}_{\mathsf{N}}$ is a linear basis for $\mathfrak{so}(\mathsf{N})$. We see that $[E_{a,b}, E_{b,c}] = E_{a,c}$ for any a, b, c. This basis is convenient because different elements are orthogonal with respect to the Killing form.

Theorem VI.2. *Let $\mathsf{N} \geq 3$ and $t \geq 1$ be any integers. Then,*

$$\Big\| \mathop{\mathbb{E}}_{\theta\sim(-\pi,\pi),\, E\in\mathbf{E}_{\mathsf{N}}} (e^{\theta E})^{\otimes t} - \mathop{\mathbb{E}}_{O\sim\mathsf{SO}(\mathsf{N})} O^{\otimes t} \Big\| \leq 1 - \frac{1}{t}\frac{2(\mathsf{N}-2)}{\mathsf{N}(\mathsf{N}-1)} - \frac{2}{\mathsf{N}(\mathsf{N}-1)}.$$

This can be used to generate an approximately Haar random $\mathsf{N} \times \mathsf{N}$ orthogonal matrix fast. The exponential of an element $E \in \mathbf{E}_{\mathsf{N}}$ is a 2×2 matrix, direct summed with an $\mathsf{N} - 2$ dimensional identity matrix. Hence, multiplying a dense $\mathsf{N} \times \mathsf{N}$ matrix by $e^{\theta E}$ takes $\mathcal{O}(\mathsf{N})$ arithmetic operations. For some applications, this method can be better than generating $\mathsf{N} \times \mathsf{N}$ Gaussian random entries and running the Gram–Schmidt process.

Proof. As before, we consider an $\mathsf{SO}(\mathsf{N})$-irrep ρ, and estimate the norm of $\mathbb{E}_{\theta,E}\,\rho(e^{\theta E})$.

(Step 1: random angles give kernel projectors.) The eigenvalues of any $E \in \mathbf{E}_{\mathsf{N}}$ are $0, \pm\mathrm{i}$. So, the average over θ gives $\mathbb{E}_{\theta,E}\,\rho(e^{\theta E}) = \mathbb{E}_E\,K(\rho_*(E))$.

(Step 2: identical eigenspectra for all represented operators) It is obvious that two different elements of $\mathbf{E}_{\mathsf{N}}$ are related by some row and column permutations. A transposition is not in $\mathsf{SO}(\mathsf{N})$, but the product of a transposition and a diagonal matrix with $\mathsf{N} - 1$ entries being 1 and the remaining being -1 is. Since $\mathsf{N} \geq 3$, we can always find such diagonal matrix that leaves a given $E_{a,b}$ fixed — just look at a zero column or a row. Hence, any two elements of $\mathbf{E}_{\mathsf{N}}$ are congruent by $\mathsf{SO}(\mathsf{N})$, and the eigenspectrum of represented operators $\rho_*(E)$ is independent of $E \in \mathbf{E}_{\mathsf{N}}$. Let $\ell = \|\rho_*(E)\| \leq t$ for any $E \in \mathbf{E}_{\mathsf{N}}$.

(Step 3: to quadratic Casimir) We need to check that $\sum_{E\in\mathbf{E}_{\mathsf{N}}} \rho_*(E)^2$ commutes with every $E \in \mathbf{E}_{\mathsf{N}}$. A moment's thought shows that it suffices to check the commutation of $\rho_*(E_{a,b})$ with $\rho_*(E_{a,c})^2 + \rho_*(E_{b,c})^2$ for any c, but this is exactly the same calculation as for the $\mathfrak{su}(2)$ case. Hence, an upper bound $\mathbf{1} - \mathbb{E}_E\,\rho_*(E)^2/\ell^2$ on $\mathbb{E}_E\,K(\rho_*(E))$ is a scalar multiple of the identity.

(Step 4: find a large number of $\mathfrak{su}(2)$'s) We focus on a vector with an eigenvalue ℓ for $\rho_*(E_{1,2})$. For any $c \geq 3$, we

have an $\mathfrak{su}(2)$-subalgebra generated by $E_{1,2}, E_{1,c}, E_{2,c}$. So, $\sum_{E\in\mathbf{E}_{\mathsf{N}}} \rho_*(E)^2 \succeq \mathbf{1}(\ell^2 + (\mathsf{N}-2)\ell)$. □

Acknowledgment

We thank Thiago Bergamaschi, Jonas Haferkamp, Aram Harrow, Zeph Landau, Ryan O'Donnell, and Peter Shor for helpful discussions. This work was done in part while X.T. was visiting the Simons Institute for the Theory of Computing, and while J.H., Y.L., and X.T. were at the Park City Mathematics Institute 2023 Graduate Summer School.

Appendix A
Discrete angles

In the proof of Theorem I.1, the only place where we use averaging over $\theta \sim (-\pi, \pi)$ is in the following context: for a nonzero hermitian operator H with integer eigenvalues in $[-t, t]$, we have

$$\mathop{\mathbb{E}}_{\theta\sim(-\pi,\pi)} e^{\mathrm{i}\theta H} = K(H) \preceq \mathbf{1} - \frac{H^2}{\|H\|^2},$$

where $K(H)$ is the orthogonal projector onto the kernel of H (see Eq. (10)). Since H has bounded norm, we can instead consider averaging over the angles in the discrete set $\Theta_t = \{m\pi/t : m \in \mathbb{Z} \cap [-t, t-1]\}$. Then, for any hermitian operator H with integer eigenvalues in $[-t, t]$, we have

$$\mathop{\mathbb{E}}_{\theta\sim\Theta_t} e^{\mathrm{i}\theta H} = \mathop{\mathbb{E}}_{\theta\sim(-\pi,\pi)} e^{\mathrm{i}\theta H} = K(H) \tag{16}$$

because for any integer k,

$$\sum_{m=-t}^{t-1} \exp(\mathrm{i}mk\pi/t) = \begin{cases} 2t & \text{if } k = 0, \\ 0 & \text{if } 0 < |k| \le t. \end{cases}$$

The rest of the proof is exactly the same as the proof of Theorem I.1.

Appendix B
State designs

We consider distributions ν on a complex projective space $\mathbb{CP}^{\mathsf{N}-1}$ where $\mathsf{N} = 2^{\mathsf{n}}$ is a power of 2. This is often called a state design because $\mathbb{CP}^{\mathsf{N}-1}$ is the set of all normalized state vectors modulo global phase factors in an n-qubit system, or equivalently the set of all rank-1 projectors $|\psi\rangle\langle\psi|$ on $(\mathbb{C}^2)^{\otimes \mathsf{n}}$. There is a natural (left) action of a unitary group given by $|\psi\rangle\langle\psi| \mapsto U|\psi\rangle\langle\psi|U^\dagger$ for $U \in \mathsf{SU}(\mathsf{N})$. The Haar measure of $\mathsf{SU}(\mathsf{N})$ induces an $\mathsf{SU}(\mathsf{N})$-invariant measure on $\mathbb{CP}^{\mathsf{N}-1}$. This is the target distribution we wish to approximate. A natural metric to measure the quality of approximation is closeness in t-th moments, maximized over all possible measurements. This is succinctly described by the trace distance:

$$\frac{1}{2}\Big\| \underbrace{\mathop{\mathbb{E}}_{\psi\sim\nu} (|\psi\rangle\langle\psi|)^{\otimes t}}_{\mathcal{S}_{\nu,t}} - \underbrace{\mathop{\mathbb{E}}_{U\sim\mathsf{SU}(\mathsf{N})} (U|\alpha\rangle\langle\alpha|U^\dagger)^{\otimes t}}_{\mathcal{S}_{\mathrm{Haar},t}} \Big\|_1 \tag{17}$$

where $|\alpha\rangle$ can be any normalized vector in $(\mathbb{C}^2)^{\otimes \mathsf{n}}$ due to the right invariance of the Haar measure. Any approximate unitary design can be used for state designs, and a bound on the t-th moment trace distance directly comes from the analysis of the approximate unitary design. For example, the result of Corollary I.2 serves the purpose. However, this is not necessarily the best one can show.

Theorem B.1. *Let $\|\cdot\|_1$ denote the Schatten 1-norm of a matrix, the sum of all singular values. For any integers $t, k, \mathsf{n} \ge 1$ and a normalized vector $|\alpha\rangle \in \mathbb{C}^{\mathsf{N}}$, we have*

$$\begin{aligned} &\Big\| \mathcal{C}_t^k(|\alpha\rangle^{\otimes t}\langle\alpha|^{\otimes t}) - \mathcal{S}_{\mathrm{Haar},t} \Big\|_1 \\ &\le \binom{\mathsf{N}+t-1}{t}^{1/2} \left(1 - \frac{1}{2t}\frac{\mathsf{N}}{\mathsf{N}+1} - \frac{\mathsf{N}}{2(\mathsf{N}^2-1)}\right)^k. \end{aligned} \tag{18}$$

The last term in the parenthesis is $\approx (2\mathsf{N})^{-1}$ for large N, which contrasts to the last term $\approx \mathsf{N}^{-2}$ in Theorem I.1.

Proof. The input $(|\alpha\rangle\langle\alpha|)^{\otimes t}$ is invariant under tensor factor permutations either on the ket or bra factors. The action by $\mathsf{SU}(\mathsf{N})$ commutes with this permutation symmetry, and hence the input vector is in an $\mathsf{SU}(\mathsf{N})$-representation $\Sigma = \mathrm{Sym}^t(\mathbb{C}^{\mathsf{N}}) \otimes \mathrm{Sym}^t((\mathbb{C}^{\mathsf{N}})^*)$. By the Littlewood–Richardson rule (actually its special case [16, 15.25(i)]), we have a decomposition of Σ into irreps:

$$\begin{aligned} \Sigma &= \mathrm{Sym}^t(\mathbb{C}^{\mathsf{N}}) \otimes \mathrm{Sym}^t((\mathbb{C}^{\mathsf{N}})^*) \\ &= \bigoplus_{s=0}^{t} \underbrace{\text{highest weight } s(L_1 - L_{\mathsf{N}})}_{\Gamma_s} \end{aligned}$$

Here, L_i is the dual of the diagonal matrix (an element of the Cartan subalgebra) where there is a sole nonzero entry that is 1 at the i-th position. Note that all the multiplicities of the irreps are 1, and Γ_0 is a one-dimensional trivial representation. Decompose $(|\alpha\rangle\langle\alpha|)^{\otimes t}$ into $\bigoplus_{s=0}^t \gamma_s(\alpha)$ according to the irrep decomposition Γ_s.[8] Proposition II.4 applied to Σ says that the Haar average of $\mathcal{S}_{\mathrm{Haar},t}$ projects $(|\alpha\rangle\langle\alpha|)^{\otimes t}$ onto $\gamma_0(\alpha)$. This projection is independent of α because $\gamma_0 = \gamma_0(\alpha)$ is uniquely determined by the trace-preserving property.[9]

It is now clear that

$$\mathcal{C}_t^k((|\alpha\rangle\langle\alpha|)^{\otimes t}) - \mathcal{S}_{\mathrm{Haar},t} = \mathcal{C}_t^k\left(\bigoplus_{s=1}^{t} \gamma_s(\alpha)\right). \tag{19}$$

We are going to bound the Schatten 2-norm of Eq. (19) by the factor in the parenthesis of Eq. (18). Since it is after

[8] Here, $|\alpha\rangle^{\otimes t}\langle\alpha|^{\otimes t}$ is a vector of the representation space Σ. The inner product is inherited from $(\mathbb{C}^{\mathsf{N}})^{\otimes t} \otimes (\mathbb{C}^{\mathsf{N}})^{*\otimes t}$ and is thus the Hilbert–Schmidt inner product. Since Σ is a unitary representation, the components $\gamma_s(\alpha) \in \Sigma$ for $s = 0, 1, \ldots, t$ are orthogonal to each other, and have length (defined by the inner product) equal to the Schatten 2-norm of the corresponding matrix in $(\mathbb{C}^{\mathsf{N}})^{\otimes t} \otimes (\mathbb{C}^{\mathsf{N}})^{*\otimes t}$.

[9] The other components $\gamma_s(\alpha)$ with $s > 0$ depend on α, but $\|\gamma_s(\alpha)\|_2$ are independent of α, since they are invariant under $\mathsf{SU}(\mathsf{N})$ whose action is transitive on $\mathbb{CP}^{\mathsf{N}-1}$.

all a matrix acting on $\mathrm{Sym}^t(\mathbb{C}^{\mathsf{N}})$ of dimension $\binom{\mathsf{N}+t-1}{t}$, conversion to the Schatten 1-norm gives the theorem.

Lemma III.2 says that

$$\left\|\mathcal{C}_t^k|_{\Gamma_s}\right\|_{2\to 2} = \left\|\mathop{\mathbb{E}}_{P\in\mathbf{P}_{\mathsf{n}}} K(\Gamma_{s*}(P/2))\right\|^k$$

where Γ_{s*} is the induced Lie algebra representation. Invoking Eq. (10), Lemma V.2, and most importantly Eq. (15) with highest weights $s(L_1 - L_{\mathsf{N}})$ where $s = 1, 2, \ldots, t$, we find that

$$\left\|\bigoplus_{s=1}^{t} \mathop{\mathbb{E}}_{P} K(\Gamma_{s*}(P/2))\right\| \leq 1 - \frac{1}{2t}\frac{\mathsf{N}}{\mathsf{N}+1} - \frac{\mathsf{N}}{2(\mathsf{N}^2-1)}. \quad \square$$

One can perform similar calculation for a real projective space $\mathbb{RP}^{\mathsf{N}-1}$, which is the same as the hemisphere (spherical cap) of dimension $\mathsf{N}-1$, excluding the equator of measure zero.

References

[1] C. Dankert, R. Cleve, J. Emerson, and E. Livine, "Exact and approximate unitary 2-designs and their application to fidelity estimation," *Phys. Rev. A*, vol. 80, p. 012304, 7 2009. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.80.012304

[2] D. Gross, K. Audenaert, and J. Eisert, "Evenly distributed unitaries: On the structure of unitary designs," *J. Math. Phys.*, vol. 48, no. 5, p. 052104, 05 2007.

[3] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, "Local random quantum circuits are approximate polynomial-designs," *Commun. Math. Phys.*, vol. 346, no. 2, pp. 397–434, 8 2016. [Online]. Available: https://doi.org/10.1007%2Fs00220-016-2706-8

[4] J. Haferkamp, "Random quantum circuits are approximate unitary t-designs in depth $O\left(nt^{5+o(1)}\right)$," *Quantum*, vol. 6, p. 795, Sep. 2022. [Online]. Available: https://doi.org/10.22331/q-2022-09-08-795

[5] J. Bourgain and A. Gamburd, "A spectral gap theorem in su(d)," *Journal of the European Mathematical Society (EMS Publishing)*, vol. 14, no. 5, 2012.

[6] R. O'Donnell, R. A. Servedio, and P. Paredes, "Explicit orthogonal and unitary designs," in *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*. Los Alamitos, CA, USA: IEEE Computer Society, 11 2023, pp. 1240–1260. [Online]. Available: https://doi.ieeecomputersociety.org/10.1109/FOCS57990.2023.00073

[7] A. Brodsky and S. Hoory, "Simple permutations mix even better," *Random Structures & Algorithms*, vol. 32, no. 3, pp. 274–289, 2008. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/rsa.20194

[8] J. Haferkamp, F. Montealegre-Mora, M. Heinrich, J. Eisert, D. Gross, and I. Roth, "Efficient unitary designs with a system-size independent number of non-clifford gates," *Commun. Math. Phys.*, vol. 397, no. 3, pp. 995–1041, 11 2022. [Online]. Available: https://doi.org/10.1007%2Fs00220-022-04507-6

[9] Y. Nakata, C. Hirche, M. Koashi, and A. Winter, "Efficient quantum pseudorandomness with nearly time-independent hamiltonian dynamics," *Phys. Rev. X*, vol. 7, p. 021006, 4 2017. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevX.7.021006

[10] J. Haferkamp and N. Hunter-Jones, "Improved spectral gaps for random quantum circuits: Large local dimensions and all-to-all interactions," *Phys. Rev. A*, vol. 104, p. 022417, 8 2021. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.104.022417

[11] S.-K. Jian, G. Bentsen, and B. Swingle, "Linear growth of circuit complexity from brownian dynamics," *Journal of High Energy Physics*, vol. 2023, no. 8, p. 190, 8 2023.

[12] F. G. Brandão, W. Chemissany, N. Hunter-Jones, R. Kueng, and J. Preskill, "Models of quantum complexity growth," *PRX Quantum*, vol. 2, p. 030316, 7 2021. [Online]. Available: https://link.aps.org/doi/10.1103/PRXQuantum.2.030316

[13] J. Haferkamp, "On the moments of random quantum circuits and robust quantum complexity," 2023.

[14] E. Rozenman and S. Vadhan, "Derandomized squaring of graphs," in *Approximation, Randomization and Combinatorial Optimization. Algorithms and Techniques*, C. Chekuri, K. Jansen, J. D. P. Rolim, and L. Trevisan, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 436–447.

[15] C.-F. Chen, J. Docter, M. Xu, A. Bouland, and P. Hayden, "Efficient unitary t-designs from random sums," 2024.

[16] W. Fulton and J. Harris, *Representation theory: a first course*. Springer Science & Business Media, 2013, vol. 129.

[17] A. W. Harrow and R. A. Low, "Random quantum circuits are approximate 2-designs," *Commun. Math. Phys.*, vol. 291, no. 1, pp. 257–302, 10 2009. [Online]. Available: https://doi.org/10.1007/s00220-009-0873-6

[18] M. Sved, "Counting and recounting: The aftermath," *The Mathematical Intelligencer*, vol. 6, no. 4, p. 44–46, 1984.

[19] A. Roy and A. J. Scott, "Unitary designs and codes," *Designs, Codes and Cryptography*, vol. 53, no. 1, pp. 13–31, 10 2009.

[20] A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and Quantum Computation*. American Mathematical Society, 2002, vol. 47.

[21] R. A. Low, "Pseudo-randomness and learning in quantum computation," 2010.