Short-depth designs beyond the unitary group

– Supplementary Information –

Supplementary Note 1 — Notation and definitions

Let $\mathcal{H} = \mathbb{C}^D$ be a D-dimensional Hilbert space. In this paper we focus on the qubit case, in which $D = 2^n$, where n denotes the number of qubits. Let X, Y, Z denote the Pauli matrices and let Z_i denote the operator $\mathbb{1}^{\otimes(i-1)} \otimes Z \otimes \mathbb{1}^{\otimes(n-i)}$, i.e., the Pauli-$Z$ acting on the i-th qubit and identity elsewhere on n qubits and analogously for X_i, Y_i. We let $[n]$ denote the $n+1$ integers between 0 and n, $[a:b]$ the $b-a+1$ integers between a and b, and $[a:t:b]$ the $(b-a)/t+1$ integers between a and b with steps of t, so $[0:2:n]$ denotes all even numbers between 0 and n.

Subgroups of the unitary group and Haar measure Let $G(D)$ be a subgroup of the unitary group $\mathrm{U}(D)$ on $\mathcal{H}$. The Haar-measure on $G(D)$ is defined as the unique probability measure that is both left and right invariant under multiplication with elements of $G(D)$, i.e., that for all $V \in G(D)$ and for every integrable function f we have that

$$\int_{G(D)} f(U) d\mu_H(D) = \int_{G(D)} f(VU) d\mu_H(D) = \int_{G(D)} f(UV) d\mu_H(D). \tag{1}$$

We will often write $\mathbb{E}_{U\sim G}(f(U)) \coloneqq \int_{G(D)} f(U) d\mu_H(D)$ for integration over the Haar measure. Likewise, we can define the Haar measure on the state space $W_G \coloneqq \{|\psi\rangle \mid \exists g \in G; |\psi\rangle = g\,|0_n\rangle\}$, as the unique probability measure that is invariant under multiplication with elements of G. By a slight abuse of notation we will write $|\psi\rangle \sim G$ for sampling from the Haar measure on W_G.

Ensembles and circuit depth We call distributions of elements of a group *ensembles* (typically denoted by $\mathcal{E}$). These group elements will always be implemented by quantum circuits, so we will also sometimes speak of an ensemble of circuits. When we talk about the *depth of an ensemble* (often just "the depth") we mean the maximum over the circuit depths (conventionally defined) of all the circuits in an ensemble.

Moment operators and group designs The *k-wise twirl* (also called the k-th moment operator) $\Phi_{\mathcal{E}}^{(k)}$ of a ensemble of operators $\mathcal{E}$ is defined as

$$\Phi_{\mathcal{E}}^{(k)}(A) \coloneqq \mathbb{E}_{U\sim\mathcal{E}} U^{\otimes k} A U^{\dagger,\otimes k}, \tag{2}$$

for a matrix $A \in \mathcal{L}(D^k)$. When it is clear from context what k is, we will just write the twirl as $\Phi_{\mathcal{E}}$, omitting the k from the notation. We write Φ_G for the twirl over the Haar measure of group G.

Definition 1. *A unitary k-design over the group G is an ensemble $\mathcal{E}$ of operators in G such that*

$$\Phi_G^{(k)}(A) = \Phi_{\mathcal{E}}^{(k)}(A), \tag{3}$$

for every $A \in \mathcal{L}(\mathcal{H}^{\otimes k})$.

The commutant of the *k-th diagonal representation* $U \to U^{\otimes k}$ of G (from now on we will just call this the commutant) is defined as $\mathrm{Comm}(G,k) \coloneqq \{A \in \mathcal{L}(\mathcal{H}^{\otimes k}) \mid AU^{\otimes k} = U^{\otimes k}A\}$.

Theorem 2. *The Haar twirl is the projector on the commutant. More precisely, if b_k is a basis of the commutant, then we can compute the Haar twirl as*

$$\Phi_G^{(k)}(A) = \sum_{\sigma,\tau\in b_k} \mathrm{Wg}_G(\sigma,\tau)\,\mathrm{tr}\Big(\sigma^\dagger A\Big)\tau. \tag{4}$$

Here the Weingarten function $\mathrm{Wg}_G(\sigma.\tau)$ *is defined as the pseudo-inverse of the Gram matrix* $\mathrm{G}_{\sigma,\tau} := \mathrm{tr}\left(\sigma^\dagger \tau\right)$.

If we know the commutant of a group, we thus know how to compute the twirl. See Ref.[1] for more details on the Haar measure.

Approximate group designs As it is difficult to find unitary k-designs for higher values of k, we introduce different notions of approximate group designs. All the following definitions can be defined for any subgroup of the unitary group.

Definition 3 (Relative-error approximate unitary k-designs)**.** *An ensemble* $\mathcal{E}$ *is an* ϵ*-approximate relative-error unitary* k*-design over* G *for* $\epsilon > 0$ *if*

$$(1-\epsilon)\Phi_G^{(k)} \preceq \Phi_{\mathcal{E}}^{(k)} \preceq (1+\epsilon)\Phi_G^{(k)}. \tag{5}$$

The notation $\Phi \preceq \Phi'$ *means that* $\Phi - \Phi'$ *is a completely positive map.*

Definition 4 (Additive-error approximate unitary k-designs)**.** *An ensemble* $\mathcal{E}$ *is an* ϵ*-approximate additive-error unitary* k*-design over* G *for* $\epsilon > 0$ *if*

$$||\Phi_G^{(k)} - \Phi_{\mathcal{E}}^{(k)}||_\diamond \leq \epsilon. \tag{6}$$

Here, the diamond norm is defined as $||\Phi||_\diamond = \max_{\{X;||X||_1\leq 1\}} ||(\Phi \otimes \mathbb{1}_n)(X)||_1$. A relative error ϵ always implies an additive error 2ϵ[2]. Even more, given a relative-error ϵ-approximate unitary k-design $\mathcal{E}$, any quantum experiment that queries a unitary U from $\mathcal{E}$ at most k times outputs a state that is, on average, 2ϵ-close in trace distance to the output state obtained if U were sampled from the Haar measure instead. For an additive-error ϵ-approximate unitary k-design, the same statement holds but then for non-adaptive quantum experiments[3]. An additive error is thus a weaker requirement than a relative error.

Approximate state designs We will now introduce some notions of state designs.

Definition 5 (Relative-error approximate state design)**.** *An ensemble* $\mathcal{E}$ *of pure states is an* ϵ*-approximate relative-error state* k*-design over* G *for* $\epsilon > 0$ *if*

$$(1-\epsilon)\mathbb{E}_{|\psi\rangle\sim G} |\psi\rangle\langle\psi|^{\otimes k} \preceq \mathbb{E}_{|\psi\rangle\sim\mathcal{E}} |\psi\rangle\langle\psi|^{\otimes k} \preceq (1+\epsilon)\mathbb{E}_{|\psi\rangle\sim G} |\psi\rangle\langle\psi|^{\otimes k}. \tag{7}$$

Definition 6 (Additive-error approximate state design)**.** *An ensemble* $\mathcal{E}$ *is an* ϵ*-approximate additive-error state* k*-design for* $\epsilon > 0$ *if*

$$\left|\left|\mathbb{E}_{|\psi\rangle\sim G} |\psi\rangle\langle\psi|^{\otimes k} - \mathbb{E}_{|\psi\rangle\sim\mathcal{E}} |\psi\rangle\langle\psi|^{\otimes k}\right|\right|_1 \leq \epsilon. \tag{8}$$

Like with group designs, every quantum experiment that queries a state $|\psi\rangle$ at most k times from a relative-error state k-design $\mathcal{E}$, outputs a state that is, on average, 2ϵ-close to the state we would have gotten if we would have sampled from the Haar random distribution. For additive-error a similar statement holds, but then for non-adaptive algorithms. If $\mathcal{E}$ is a relative-error unitary k-design, then $\{U|0_n\rangle \mid U \sim \mathcal{E}\}$ is a relative-error state k-design. A similar statement holds for additive error. State designs are thus a weaker notion than group designs.

Anti-concentration An even weaker notion of randomness is anti-concentration. As mentioned above, we here adopt a common misusage of this word in the random circuit community and mean that certain moments of the ensemble $\mathcal{E}$ approximate the Haar value over G. For some groups G, not even the outcome distribution induced by Haar-random unitaries anti-concentrates in the conventional sense , and thus also not ensembles $\mathcal{E}$ obeying our definition. For the here studied groups, this is the case for the matchgate group.

Concretely, the definition we will use throughout this paper is the following:

Definition 7. *An ensemble $\mathcal{E}$ anti-concentrates up to order k if*

$$\mathbb{E}_{U\sim\mathcal{E}} \sum_{x\in\{0,1\}^n} |\langle x|U|0_n\rangle|^{2k} \leq \alpha \cdot \mathbb{E}_{U\sim G} \sum_{x\in\{0,1\}^n} |\langle x|U|0_n\rangle|^{2k} \tag{9}$$

for some constant $\alpha \geq 1$.

We note that if $\mathcal{E}$ forms a relative-error state k-design, then it anti-concentrates.

Architectures Qubits are ordered in an *architecture*. An architecture is described by a graph, where the vertices of the graph are the qubits, and an edge between two vertices indicates that a 2-qubit gate can be applied on those 2 qubits. Commonly studied architectures are 1D and 2D lattices (with open or closed boundary conditions), or all-to-all connectivity (the complete graph). We define the *diameter* (often denoted as dia) of an architecture as the largest possible distance between two vertices, e.g. the diameter of 1D is $\text{dia} = \Theta(n)$, for 2D $\text{dia} = \Theta(\sqrt{n})$, and for the all-to-all connectivity $\text{dia} = \Theta(1)$. Note, however, that many known constructions of designs in the all-to-all connectivity actually use architectures of higher diameter. For example, the construction of Clifford 3-designs in[3] in $\mathcal{O}(\log\log(n))$ depth implicitly uses an architecture of diameter $\mathcal{O}(n/\log(n))$.

Supplementary Note 2 — Clifford group

1 Preliminaries

The n-qudit Clifford group Cl_n is defined as the group of unitaries that normalize the Pauli group $\mathcal{P}_n$, i.e., the group of all unitaries U such that $U\mathcal{P}_nU^\dagger \subset \mathcal{P}_n$. Here, we will only consider qubits, in which case the Clifford group can equivalently be defined as the group generated by the CNOT-gate, the Hadamard gate H and the phase gate S. The n-qubit Clifford group forms an exact unitary 3-design, but not a 4-design[4]. We can thus directly extend the gluing lemma from[3] to $k = 3$.

We note that the results from this section can be readily generalized to prime-dimensional qudits, in which case the Clifford group only forms a unitary 2-design, but not a 3-design[5]. The analogous statements of our theorems would then concern bounds on the circuit depth of Clifford 3-designs.

We are interested in the diagonal representation of the Clifford group, $U \to U^{\otimes k}$. The commutant of this representation can be described using *Stochastic Lagrangian subspaces*[6].

Definition 8. *The subspace $T \subset \mathbb{Z}_2^{2k}$ is a Stochastic Lagrangian subspace if the following three properties hold.*

1. *For every $(x, y) \in T$, we have that $x \cdot x = y \cdot y \mod 4$*
2. *T has dimension k*
3. *T is stochastic, which means that $(1, \ldots, 1) \in T$.*

We define $\Sigma_{k,k}$ as the set of all Stochastic Lagrangian subspaces.

We note that condition 1 implies that T has dimension at most k. Condition 2 could thus equivalently be stated as T having the largest possible dimension. The commutant of the diagonal representation can now be described as follows.

Theorem 9 (Commutant of the Clifford group[6])**.** *If $n \geq k-1$, then the commutant of the diagonal representation is spanned by the operators $R(T) = r(T)^{\otimes n}$ for $T \in \Sigma_{k,k}$, where*

$$r(T) \coloneqq \sum_{(x,y)\in T} |x\rangle\langle y| . \tag{10}$$

The Clifford group is a subgroup of the unitary group. This means that the commutant of the Clifford group should at least include the commutant of the unitary group, which we know to be the permutations. Indeed, for $\sigma \in S_k$ a permutation, we see that $x \cdot x = (\sigma x) \cdot (\sigma x) \mod 4$. Hence, the space $T_\sigma = \{(\sigma x, x) \mid x \in \mathbb{Z}_2^k\} \in \Sigma_{k,k}$. In the following, we will hence view S_k as a subset of $\Sigma_{k,k}$ via this identification, i.e., $S_k \subseteq \Sigma_{k,k}$. Indeed, for $k \leq 3$, $S_k = \Sigma_{k,k}$. For $k = 4$ however, there exist other subspaces and so $S_k \subset \Sigma_{k,k}$. An example of such a space is T_4 with representation

$$R(T_4) = 2^{-n} \sum_{P \in \mathrm{P}_n} P^{\otimes 4}, \tag{11}$$

where the sum runs over all n-qubit Pauli matrices.

2 Clifford group designs

In this section we prove a lower bound for the circuit depth of additive-error Clifford (unitary) designs based on lightcone arguments. This bound is remarkably general, but depends on the underlying architecture.

For generalized random brickwork circuits (see Def. **??**), we obtain a much stronger bound, namely $\Omega(n)$, independent of the architecture. This intriguing result follows from a different argument exploiting the structure of such circuits and the discreteness of the Clifford group, and is derived for relative-error state designs in the next section.

Theorem 10. *Let $\mathcal{E}$ be an approximate Clifford unitary k-design (with $k \geq 4$) with additive error $\varepsilon < 1/4$. Then the lightcone of at least one Clifford in the ensemble must cover a constant fraction of the qubits. This implies $\mathcal{E}$ must have circuit depth $d = \Omega(n)$ in* 1D *circuits and $d = \Omega(\log n)$ in all-to-all connected circuits.*

Proof. To prove the theorem, we will show that any Clifford unitary with lightcone smaller than the system size can be distinguished from a Haar-random Clifford unitary using four parallel queries and constant-depth quantum and classical operations.

Let C be a short-depth Clifford unitary and $\chi_n = \frac{1}{8^n} R(T_4) = \frac{1}{16^n} \sum_{P \in P_n} P^{\otimes 4}$ be the mixed state with stabilizers $P^{\otimes 4}$ for all Pauli operators P. The state χ_n can be prepared in constant-depth since it is a tensor product of n 4-qubit stabilizer states, $\chi_n = \chi_1^{\otimes n}$. Our distinguishing protocol proceeds as follows. We first perturb χ_n by a single-qubit Pauli operator Z_1 on the first qubit (taking some arbitrary but fixed order of the qubits) of the first copy. We then apply C to all four copies in parallel. Finally we choose $i \in [n]$ uniformly at random and we measure the expectation value of the local tensor power Pauli operator $Z_i^{\otimes 4}$.

This procedure yields an expectation value,

$$\mathbb{E}_{i\sim[n]} \operatorname{tr}\left[Z_i^{\otimes 4} C^{\otimes 4} (Z_1 \otimes \mathbb{1}^{\otimes 3}) \chi_n (Z_1 \otimes \mathbb{1}^{\otimes 3}) C^{\dagger,\otimes 4}\right] = \mathbb{E}_{i\sim[n]} \operatorname{tr}\left[Z_i^{\otimes 4} (C Z_1 C^\dagger \otimes \mathbb{1}^{\otimes 3}) \chi_n (C Z_1 C^\dagger \otimes \mathbb{1}^{\otimes 3})\right].$$

In the second expression, we use that $\chi_n = C^{\dagger,\otimes 4} \chi_n C^{\otimes 4}$ since χ_n commutes with every Clifford unitary C. We then cancel the Clifford unitaries on the final three copies. Now imagine all Cliffords C in $\mathcal{E}$ have a lightcone of size less than $n/4$. Then the resulting expectation value is larger than $1/2$ because $Z_i^{\otimes 4}$ and $CZ_1C^\dagger$ commute whenever i is outside of the lightcone and $Z_i^{\otimes 4}$ is a stabilizer of χ_n (which is true for a 3/4'th fraction of the qubits). For the qubits inside the lightcone (a less than 1/4'th fraction of the qubits) the expectation value cannot be more negative than -1, making the overall expectation value larger than $1/2$. Meanwhile, in the Haar-random Clifford unitary ensemble, the expectation value of the expectation value is exponentially small in n. This implies that $\|\Phi_{\mathcal{E}}^{(4)} - \Phi_{\mathrm{Cl}_H}^{(4)}\|_\diamond \geq 1/4$ for any Clifford ensemble $\mathcal{E}$ of Cliffords whose lightcone is of size less than $n/4$. This completes the proof.

3 Clifford state designs: relative error

Theorem 11. *Let $\mathcal{E}$ be an ensemble such that $\Phi_{\mathcal{E}}\left(|0^n\rangle\langle 0^n|^{\otimes 4}\right)$ is a relative-error ϵ-approximate Clifford state 4-design of depth d for $\epsilon < 1$. The following 2 statements hold:*

1. *If the architecture has diameter* dia *and* $\mathcal{E}$ *is locally independent (Def.* **??***), then* $d = \Omega(\text{dia})$.

2. *If* $\mathcal{E}$ *is a generalized random brickwork (Def.* **??***), then* $d = \Omega(n)$.

Proof. Let U be a Clifford unitary and consider $R(T_4)$ as in Eq. (11). We find:

$$\begin{aligned}
&\operatorname{tr}\left[(Z_i \otimes \mathbb{1}^{\otimes 3}) R(T_4) (Z_i \otimes \mathbb{1}^{\otimes 3}) U^{\otimes 4} (|0000\rangle\langle 0000|) U^{\dagger,\otimes 4}\right] && (12)\\
&= \langle 0000 | (U Z_i U^\dagger \otimes \mathbb{1}^{\otimes 3}) R(T_4) (U Z_i U^\dagger \otimes \mathbb{1}^{\otimes 3}) |0000\rangle && (13)\\
&= \frac{1}{2^n} \langle 0 | (U Z_i U^\dagger)(\mathbb{1} + Z)^{\otimes n} (U Z_i U^\dagger) |0\rangle && (14)\\
&= \langle 00 | U^{\otimes 2} Z_i^{\otimes 2} U^{\dagger,\otimes 2} |00\rangle . && (15)
\end{aligned}$$

In step two, we performed the partial contraction on the last three systems and used that $\langle 0 | P | 0 \rangle = 1$ if P is a diagonal Pauli and zero else. In the last step, we resumed the stabilizers of the all-zero state. If U is sampled from the Haar measure on the Clifford group, we get

$$\mathbb{E}_{U \sim \mathrm{Cl}_H} \langle 00 | U^{\otimes 2} Z_i^{\otimes 2} U^{\dagger,\otimes 2} |00\rangle = \frac{1}{2^n(2^n+1)} \operatorname{tr}\left[Z_i^{\otimes 2}(\mathbb{1} + \mathbb{F})\right] = \frac{1}{2^n+1}. \tag{16}$$

For sampling from an ensemble $\mathcal{E}$ of Clifford circuits of depth d, we find:

$$\mathbb{E}_{U \sim \mathcal{E}} \langle 00 | U^{\otimes 2} Z_i^{\otimes 2} U^{\dagger,\otimes 2} |00\rangle = \mathbb{E}_{U \sim \mathcal{E}} \langle 00 |_{l_i} U_{l_i}^{\otimes 2} Z_i^{\otimes 2} U_{l_i}^{\dagger,\otimes 2} |00\rangle_{l_i} , \tag{17}$$

where the subscript l_i indicates that the unitary only acts on the qubits inside the lightcone starting from qubit i. As $\mathcal{E}$ is a 4-design, Eq. (17) is within relative error ϵ to the Haar random value $\frac{1}{2^n+1}$.

1. Choose i and j as two qubits that have distance dia from each other. If $\mathcal{E}$ has depth $d < \text{dia}/4$, the following equations need to hold:

$$\begin{aligned}
\mathbb{E}_{U \sim \mathcal{E}} \langle 00 |_{l_i} U_{l_i}^{\otimes 2} Z_i^{\otimes 2} U_{l_i}^{\dagger,\otimes 2} |00\rangle_{l_i} &\leq (1+\epsilon)\frac{1}{2^n+1}\\
\mathbb{E}_{U \sim \mathcal{E}} \langle 00 |_{l_j} U_{l_j}^{\otimes 2} Z_j^{\otimes 2} U_{l_j}^{\dagger,\otimes 2} |00\rangle_{l_j} &\leq (1+\epsilon)\frac{1}{2^n+1}\\
\mathbb{E}_{U \sim \mathcal{E}} \langle 00 | U^{\otimes 2} (Z_i Z_j)^{\otimes 2} U^{\dagger,\otimes 2} |00\rangle &= \mathbb{E}_{U \sim \mathcal{E}} \langle 00 |_{l_i} U_{l_i}^{\otimes 2} Z_i^{\otimes 2} U_{l_i}^{\dagger,\otimes 2} |00\rangle_{l_i}\\
&\quad \times \langle 00 |_{l_j} U_{l_j}^{\otimes 2} Z_j^{\otimes 2} U_{l_j}^{\dagger,\otimes 2} |00\rangle_{l_j} \geq (1-\epsilon)\frac{1}{2^n+1},
\end{aligned} \tag{18}$$

where we used that $d < \text{dia}/4$, so that the lightcones starting from i and j do not intersect. We directly see that these equations contradict the locally independent assumption. This proves point 1.

2. Let $\mathcal{E}$ be a generalized random brickwork (see Def. **??**) of depth d. By construction Eq. (15) is non-negative, hence we can lower bound its average by the probability that $\langle 00 | U^{\otimes 2} Z_1^{\otimes 2} U^{\dagger,\otimes 2} |00\rangle = 1$ over all circuit instances U. This probability can be equivalently phrased as the probability that $U Z_1 U^\dagger$ is a diagonal Pauli operator. Here, we lower bound this probability exploiting that the local Clifford gates in every layer are uniformly distributed – we however note that such an argument can likely be made for much more general distributions (cf. the comment below Def. **??**). Because each brick has constant-size support, there is a constant probability p that the first brick maps Z_1 to a weight-one Pauli operator (not necessarily supported on the first qubit anymore). More precisely, if the support is of size r, then $p = 3r/(4^r - 1)$. The probability that the bricks in the remaining layers preserve the weight is also precisely p. Since the ensemble is invariant under local Cliffords, the probability that the final Pauli operator after d layers is diagonal is $p^d/3$. Hence, we have $\langle 00 | U^{\otimes 2} Z_1^{\otimes 2} U^{\dagger,\otimes 2} |00\rangle = 1$ with probability is at least $p^d/3$, resulting in the lower bound

$$\mathbb{E}_{U \sim \mathcal{E}} \langle 00 | U^{\otimes 2} Z_i^{\otimes 2} U^{\dagger,\otimes 2} |00\rangle \geq \frac{1}{2^{\mathcal{O}(d)}}. \tag{19}$$

This, together with Eq. (16), implies a lower bound of $d = \Omega(n)$ for $\mathcal{E}$ to be a relative-error Clifford state 4-design.

We believe the assumption of local independence in the first point of the theorem can be dropped, but we do not know how to prove this. In particular, our proof cannot easily be generalized. Define for instance $\mathcal{E}$ as the ensemble that draws $\mathbb{1}$ with probability $\frac{1}{2^n+1}$, and draws $H^{\otimes n}$ with probability $1-\frac{1}{2^n+1}$. Then all three values in Eq. (18) are exactly $\frac{1}{2^n(2^n+1)}$, the Haar random value. The ensemble $\mathcal{E}$ is however not locally independent, and it does not define a Clifford state design.

4 Clifford state designs: additive error

Above we have shown that the Clifford group does not allow for short-depth state k-designs (for $k > 3$) in the relative-error model. However, somewhat surprisingly, it turns out to be possible to construct short-depth state k-designs for $k = 4$ and $k = 5$ with small (i.e. $\mathcal{O}(2^{-n})$) additive error. To prove this we need the following lemma that shows that for $k \leq 5$ Clifford unitaries applied to a PPT input state (for all possible partial transposes of n qubits) generate $U(2^n)$ state k-designs with small additive error:

Lemma 12. *Consider a state ρ on $\mathcal{H}^{\otimes k}$. Suppose that ρ obeys the PPT condition on each copy of $\mathcal{H}$, i.e., $\rho^{\Gamma_i} \geq 0$ for all $i = 1, \ldots, k$, where Γ_i denotes the partial transpose on the i-th copy. Then, if $k \leq 5$, the Clifford twirl of ρ is equal to the Haar unitary twirl of ρ up to additive error:*

$$\left\| \Phi^{(k)}_{\mathrm{Cl}_H}(\rho) - \Phi^{(k)}_{U_H}(\rho) \right\|_1 \leq \mathcal{O}(2^{-n}). \tag{20}$$

Proof. The lemma is trivial for $k \leq 3$. For $k = 4, 5$, the set of stochastic Lagrangian subspaces is given by $\Sigma_{k,k} = S_k \cup \overline{S_k}$ where $\overline{S_k} = \Sigma_{k,k} \setminus S_k$ denotes the complement of S_k. Hence, the commutant of the Clifford group is spanned by the operators

$$\{R(T)\}_{T \in \Sigma_{k,k}} = \{R(\pi) \mid \pi \in S_k\} \cup \{R(T) \mid T \in \overline{S_k}\}.$$

Specifically for $k = 4, 5$, all $T \in \overline{S_k}$, $R(T)$ can be written as $R(T) = R(T_k)R(\pi)$ for a suitable permutation $\pi \in S_k$ with T_4 defined in Eq. 11 and $T_5 = T_4 \otimes I$. The $R(T)$ for $T \in \overline{S_k}$ are not unitary, and we have $\|R(T)\|_1 = 2^{n(k-1)}$. We will now show that for each of the operators $R(T_k)R(\pi)$ there is a partial transpose such that $\left\|(R(T_k)R(\pi))^{\Gamma_i}\right\| = 1$. Note that this is certainly true for $R(T_k)$. It can be checked by direct computation that $R(T_k)^{\Gamma_i}$ for $i \in \{1, 2, 3, 4\}$ is unitary. Now choose $i \in \{1, 2, 3, 4\}$ such that $\pi(i) \neq 5$ (which can always be done) and apply Γ_i. Up to a permutation σ (which does not increase the operator norm of $R(T_k)R(\pi))^{\Gamma_i}$ we have

$$(R(T_k)R(\pi))^{T_i} R(\sigma) = 2^n \sum_{P \in \mathrm{P}_n} (I \otimes)^{\delta_{k=5}} P^{\otimes 2} \otimes \left(I \otimes P \left|\Phi\right\rangle\!\left\langle\Phi\right| I \otimes P^\dagger\right) \tag{21}$$

where Φ is the maximally entangled state. This operator factorizes over n and can be checked numerically to have operator norm 1. As a consequence, we can show that for $k \leq 5$,

$$\left|\mathrm{tr}\left(R(T)\rho\right)\right| \leq 1, \quad \forall\, T \in \Sigma_{k,k}. \tag{22}$$

To see this, note that using Hölder's inequality, we have

$$\left|\mathrm{tr}\left(R(T)\rho\right)\right| = \left|\mathrm{tr}\left(R(T)^{\Gamma_i}\rho^{\Gamma_i}\right)\right| \leq \left\|R(T)^{\Gamma_i}\right\|_\infty \left\|\rho^{\Gamma_i}\right\|_1 = 1. \tag{23}$$

Now, we can bound the distance of the twirls directly, as follows

$$\left\| \Phi^{(k)}_{\mathrm{Cl}_H}(\rho) - \Phi^{(k)}_{U_H}(\rho) \right\|_1 \tag{24}$$

$$= \left\| \sum_{T,T' \in \Sigma_{k,k}} \mathrm{Wg}^{(k)}_{\mathrm{Cl}}(T,T')\, \mathrm{tr}\left(R(T)^\dagger \rho\right) R(T') - \sum_{T,T' \in S_k} \mathrm{Wg}^{(k)}_{\mathrm{U}}(T,T')\, \mathrm{tr}\left(R(T)^\dagger \rho\right) R(T') \right\|_1 \tag{25}$$

$$\leq \sum_{T,T' \in \Sigma_{k,k}} \left|\mathrm{Wg}^{(k)}_{\mathrm{Cl}}(T,T') - \mathrm{Wg}^{(k)}_{\mathrm{U}}(T,T')\right| \left|\mathrm{tr}\left(R(T)^\dagger \rho\right)\right| \left\|R(T')\right\|_1 \tag{26}$$

$$\leq \sum_{T,T' \in \Sigma_{k,k}} \left|\mathrm{Wg}^{(k)}_{\mathrm{Cl}}(T,T') - \mathrm{Wg}^{(k)}_{\mathrm{U}}(T,T')\right| \left\|R(T')\right\|_1 \tag{27}$$

where we slightly abuse notation by setting $\mathrm{Wg}_{\mathrm{U}}^{(k)}(T,T')=0$ if $T\notin S_k$ or $T'\notin S_k$. In the last step, we used Eq. 22. Next, we split the sum over $\Sigma_{k,k}$.

In particular, we can upper bound this sum by four terms by choosing T,T' in S_k or $\overline{S_k}$. For visual help, we will write $\pi\in S_k$ and $T\in\overline{S_k}$. In order we have

$$(1)\leq\sum_{\pi,\pi'\in S_k}2^{nk}|\,\mathrm{Wg}_{\mathrm{Cl}}^{(k)}(\pi,\pi')-\mathrm{Wg}_{\mathrm{U}}^{(k)}(\pi,\pi')|\leq\mathcal{O}(2^{-n}) \tag{28}$$

where we used that $\|R(\pi)\|_1=2^{nk}$ for all $\pi\in S_k$ and $|\,\mathrm{Wg}_{\mathrm{Cl}}^{(k)}(\pi,\pi')-\mathrm{Wg}_{\mathrm{U}}^{(k)}(\pi,\pi')|\leq\mathcal{O}(2^{-n(k+1)})$ (via a straightforward Weingarten calculation, see also Ref.[7]). Similarly we can see that

$$(2)\leq\sum_{\pi\in S_k,T'\in\overline{S_k}}|\,\mathrm{Wg}_{\mathrm{Cl}}^{(k)}(\pi,T')|\|R(T')\|_1=\mathcal{O}(2^{-n}), \tag{29}$$

since $|\,\mathrm{Wg}_{\mathrm{Cl}}^{(k)}(\pi,T')|\leq 2^{-(k+1)n}$ and $\|R(T')\|_1\leq 2^{n(k-1)}$. Moving on we have

$$(3)\leq\sum_{T\in\overline{S_k},\,\pi'\in S_k}|\,\mathrm{Wg}_{\mathrm{Cl}}^{(k)}(T,\pi')|\|R(\pi')\|_1=\mathcal{O}(2^{-n}), \tag{30}$$

since again $|\,\mathrm{Wg}_{\mathrm{Cl}}^{(k)}(T,\pi')|\leq 2^{-(k+1)n}$. Finally we have

$$(4)\leq\sum_{T\in\overline{S_k},\,T'\in\overline{S_k}}|\,\mathrm{Wg}_{\mathrm{Cl}}^{(k)}(T,T')|\|R(T')\|_1\leq\mathcal{O}(2^{-n}) \tag{31}$$

which follows from $\|R(T')\|_1=2^{n(k-1)}$ and $|\,\mathrm{Wg}_{\mathrm{Cl}}^{(k)}(T,T')|\leq O(2^{-nk})$. This completes the proof.

We note that Eq. (27) can be computed exactly using Weingarten calculus. This allows to bound the constant in Lemma 12 by $1/4$ (for $n>2$)[8].

From this construction and the gluing lemma due to[3] we get in a straightforward way that the superblock construction applied to the all-zero initial state is an additive error state design.

Corollary 13. *Consider the two-layer superblock brickwork circuit composed of uniformly random Clifford circuits acting on 2ξ qubits, applied to the all-zero initial state. There is a constant $C\geq 1$ such that for any $\epsilon=\Omega(2^{-n})$, $k=4,5$, and $\xi\geq\log_2(Cn/\epsilon)$, this ensemble forms an ϵ-approximate additive-error Clifford state k-design with depth $\mathcal{O}(\log(n/\epsilon))$.*

Proof. Let C' be the constant in Lemma 12. As the input state $|0\rangle^{\otimes k}$ is separable, we can replace the uniformly random Cliffords in the first layer with Haar random unitaries (using Lemma 12), incurring a total additive error $\leq C'n2^{-2\xi}\leq C'C^{-2}\epsilon^2k^{-2}n^{-1}\leq C'C^{-1}\epsilon$ (as there are $n/2\xi\leq n$ many gates). Similarly, since the state after this first layer of unitaries is still separable, we can replace the uniformly random Cliffords in the second layer by Haar-random unitaries, incurring the same amount of error again. We now have a superblock circuit with Haar-random unitaries for which we can invoke the gluing lemma of Ref.[3] to bound the relative error of the whole circuit (to a global Haar-random unitary). More precisely, this error is computed in Ref.[3]. Inserting $k\geq 4$ in the derivation of[3], we find that $\xi\geq\log_2(Cnk^2/\epsilon)$ implies a relative error $\leq\epsilon/(2C)$. This converts into an additive error ϵ/C^2. Finally, exchanging the global Haar unitary back to a random Clifford yields another additive error $\leq C'2^{-n}$. The total error is then $\frac{C'+1}{C}\epsilon+C'2^{-n}$. If the target error is larger than $C'2^{-n}$, then choosing $C=C'+1$ and redefining ϵ to be the difference between the target error and $C'2^{-n}$ shows that the ensemble is the required ϵ-design. Since one can implement arbitrary Clifford circuits in 1D in depth $O(\xi)$ the overall depth of the circuit is $O(\log(nk/\epsilon))$, which is what we set out to prove.

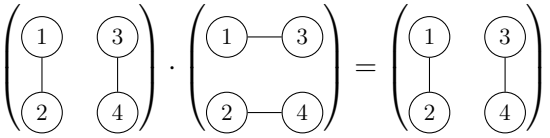
Supplementary Note 3 — Orthogonal group

1 Preliminaries

The orthogonal group $\mathrm{O}(D)$ is defined as the set of all unitary matrices that have real entries, or, equivalently, the set of real matrices O such that $OO^T = O^T O = \mathbb{1}$. It is known that orthogonal unitary k-designs can be constructed in depth linear in n and polynomial in k[9].

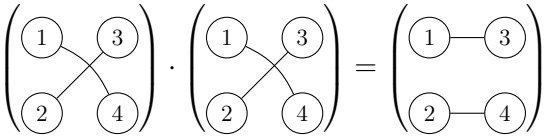
We quickly review orthogonal Weingarten calculus here. The commutant of the orthogonal group is a representation of the *Brauer algebra*.

Let M_k denote the set of all possible pairings of $2k$ vertices. If we number the vertices $1, 2, 3, 4$, the elements of M_2 are thus $\{\{1,2\},\{3,4\}\}, \{\{1,3\},\{2,4\}\}$ and $\{\{1,4\},\{2,3\}\}$. We define a multiplication $\sigma \cdot \tau$ for $\sigma, \tau \in M_k$ as concatenating the last k vertices of σ with the first k vertices of τ. For $k = 2$ this means for example that $\sigma \cdot \tau \coloneqq \{\{1,2\},\{3,4\}\} \cdot \{\{1,3\},\{2,4\}\} = \{\{1,2\},\{3,4\}\}$, or, in graphical notation



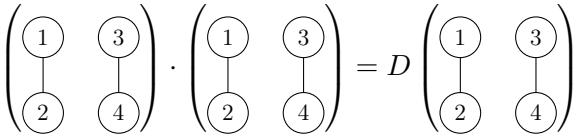
(32)

We can consider the permutations as being a subset of the Brauer algebra. If, for example, we multiply the swap operator with itself, then we get



(33)

However, sometimes it happens that by concatenating elements from M_k, we get one or more closed loop. For every closed loop, we multiply by some scalar D. An example of this for $k = 2$ is



(34)

The Brauer algebra is now defined as the set of formal linear combinations of pairings:

$$\mathcal{B}_k(D) = \left\{ \sum_{\sigma \in M_k} c_\sigma \sigma \mid c_\sigma \in \mathbb{Z}(D) \right\}. \tag{35}$$

For an element $\sigma \in M_k$, let us define a map $f_\sigma : [2k] \to [2k]$ which sends a vertex x to the vertex y that it is paired with under σ. We can now define a representation $R_\mathrm{O} : \mathcal{B}_k(D) \to \mathcal{U}(D^k)$ of the Brauer algebra. For $\sigma \in M_k$ we define

$$R_\mathrm{O}(\sigma) \coloneqq \sum_{i_1, \ldots i_{2k}} \left(\prod_{j=1}^{2k} \delta_{i_j, i_{f_\sigma(j)}} \right) |i_1 \ldots i_k\rangle\langle i_{k+1} \ldots i_{2k}| \,. \tag{36}$$

We chose D to be the dimension of one copy of our Hilbert space. If we define $\sigma = \{\{1,2\},\{3,4\}\}$, we indeed get that

$$R_\mathrm{O}(\sigma) \cdot R_\mathrm{O}(\sigma) = \left(\sum_{i_1, i_2} |i_1 i_1\rangle\langle i_2 i_2| \right) \left(\sum_{i_1, i_2} |i_1 i_1\rangle\langle i_2 i_2| \right) = D \sum_{i_1, i_2} |i_1 i_1\rangle\langle i_2 i_2| = D \cdot R_\mathrm{O}(\sigma). \tag{37}$$

We note that all permutations are in the representation of the Brauer algebra.

The elements of M_k under representation R_O form a basis of the commutant of the orthogonal group[10]. We can thus use orthogonal Weingarten calculus to compute the moments of the orthogonal group as

$$\Phi_{\mathrm{O}}^{(k)}(A) = \sum_{\sigma,\tau\in M_k} \mathrm{Wg}_{\mathrm{O}}(\sigma,\tau)\,\mathrm{tr}\Big(A\sigma^T\Big)\tau, \tag{38}$$

where Wg_O is the inverse of the orthogonal Gram matrix $(G_O)_{\sigma,\tau} = \mathrm{tr}\Big(\sigma\tau^T\Big)$. We note that, as M_k does not form a group and in particular does not have inverses, the orthogonal Gram and Weingarten matrix do not have a constant row sum.

It will be useful to have a closer look at the case of $k = 2$. We define $|\Omega_n\rangle \coloneqq \frac{1}{2^{n/2}}\sum_{x\in\{0,1\}^n}|xx\rangle$ as the maximally entangled state on two copies of $(\mathbb{C}^2)^{\otimes n}$. The basis for the commutant will now consist of the identity operator $\mathbb{1}$, the Flip operator $\mathbb{F}$ and $\Pi_o \coloneqq 2^n\,|\Omega_n\rangle\langle\Omega_n|$. Computing the Gram matrix and its inverse we get

$$G_{\mathrm{O}} = D\begin{pmatrix} D & 1 & 1\\ 1 & D & 1\\ 1 & 1 & D\end{pmatrix},\ \mathrm{Wg}_{\mathrm{O}} = \frac{1}{D(D-1)(D+2)}\begin{pmatrix} D+1 & -1 & -1\\ -1 & D+1 & -1\\ -1 & -1 & D+1\end{pmatrix}. \tag{39}$$

Hence,

$$\begin{aligned}\mathbb{E}_{U\sim\mathrm{O}}U^{\otimes 2}AU^{\dagger,\otimes 2} = \frac{1}{D(D-1)(D+2)}\Big(&[(D+1)\,\mathrm{tr}(A) - \mathrm{tr}(A\mathbb{F}) - \mathrm{tr}(A\Pi_o)]\,\mathbb{1}\\ &+[-\mathrm{tr}(A) + (D+1)\,\mathrm{tr}(A\mathbb{F}) - \mathrm{tr}(A\Pi_o)]\,\mathbb{F}\\ &+[-\mathrm{tr}(A) - \mathrm{tr}(A\mathbb{F}) + (D+1)\,\mathrm{tr}(A\Pi_o)]\,\Pi_o\Big).\end{aligned} \tag{40}$$

2 Orthogonal group designs

The following argument was already given by[3], we repeat the main idea of the proof here for completeness. For every orthogonal matrix $O \in \mathrm{O}(2^n)$, we have $O\otimes O\,|\Omega_n\rangle = OO^T\otimes\mathbb{1}\,|\Omega_n\rangle = |\Omega_n\rangle$, i.e., $|\Omega_n\rangle$ is an invariant state under the tensor square representation. We then study the action on the initial state $Z_1\otimes\mathbb{1}\,|\Omega_n\rangle$ where Z_1 denotes the Pauli Z operator on the first qudit. If we act with a Haar-random $O \in \mathrm{O}(2^n)$, we obtain the state $(OZ_1O^T\otimes\mathbb{1})\,|\Omega_n\rangle$ which will be highly entangled between the first n qudits with high probability. In contrast, if we act with a local random O-circuit, then all gates outside the lightcone of the first qudit will cancel. In particular, if the lightcone does not encompass all qudits, we are left with local maximally entangled states $|\Omega_{\bar{l}}\rangle$ in the complement of the lightcone l. These are unentangled with the remaining qubits and can be detected, for instance by measuring their fidelity. Hence, short local random O-circuits cannot reproduce this second (psd) moment correctly. In a 1D nearest-neighbor architecture, they need at least $\Omega(n)$ depth, and in an all-to-all architecture $\Omega(\log(n))$ depth is required. We provide a formal proof here:

Theorem 14. *Let $\mathcal{E}$ be an approximate orthogonal unitary k-design with additive error $\varepsilon < 1/4$. Then at least one orthogonal unitary must have a lightcone covering a $1/4$'th fraction of the qubits. This implies $\mathcal{E}$ must have circuit depth $d = \Omega(n)$ in 1D circuits and $d = \Omega(\log n)$ in all-to-all connected circuits for any $k \geq 2$.*

Proof. To prove the theorem, we will show that any orthogonal unitary with lightcone size smaller than $n/4$ can be distinguished from a Haar-random orthogonal unitary using two parallel queries and constant-depth quantum and classical operations.

Let O be a short-depth orthogonal unitary and $|\Omega_n\rangle$ the EPR state on two copies. Our distinguishing protocol proceeds as follows. We first perturb the EPR state by a single-qubit Pauli operator Z_1 on the first qubit (defined in a fixed but arbitrary order) of the first copy. We then apply O to both copies in parallel. Finally, we choose $i \in [n]$ uniformly at random and take the trace inner product with a Pauli-Z operator on both copies of qubit i, $P^{(i)} = Z_i^{\otimes 2}$.

This procedure yields an expectation value,

$$\mathbb{E}_{i\sim[n]} \langle\Omega_n| (Z_1\otimes\mathbb{1})(O^T\otimes O^T)P^{(i)}(O\otimes O)(Z_1\otimes\mathbb{1}) |\Omega_n\rangle = \mathbb{E}_{i\sim[n]} \langle\Omega_n| (OZ_1O^T\otimes\mathbb{1})P^{(i)}(OZ_1O^T\otimes\mathbb{1}) |\Omega_n\rangle .$$

Here, we have used that $O^\dagger = O^T$ and $(\mathbb{1}\otimes O)|\Omega_i\rangle = (O^T\otimes\mathbb{1})|\Omega_i\rangle$. Now consider an orthogonal unitary O with a lightcone of size less than $n/4$. Note that the internal expectation value is equal to one whenever i lies outside of the lightcone of OZ_1O^T, since then $P^{(i)}$ and OZ_1O^T commute. Since the expectation value is otherwise no less than -1 it follows that the overall expectation value is larger than $1/2$. Meanwhile, in a Haar-random orthogonal unitary, the expectation value of the expectation value is exponentially small in n. This implies that $\|\Phi^{(2)}_{\mathcal{E}} - \Phi^{(2)}_{O_H}\|_\diamond \geq 1/4$, which contradicts $\mathcal{E}$ being an additive error design. The rest of the statement follows.

3 Orthogonal state designs: relative error

Theorem 15. *Let $\mathcal{E}$ be an ensemble such that $\Phi_{\mathcal{E}}(|00\rangle\langle 00|)$ is a relative-error ϵ-approximate orthogonal state 2-design of depth d for $\epsilon < 1$. The following 2 statements hold:*

1. *If the architecture has diameter* dia *and $\mathcal{E}$ is locally independent (Def. **??**), then $d = \Omega(\text{dia})$.*
2. *If $\mathcal{E}$ is a generalized random brickwork (Def. **??**), then $d = \Omega(n)$.*

Proof. We follow a similar strategy as in the proof of Theorem 11. We know that the state $|\Omega\rangle\langle\Omega|$ lies in the commutant. To prove the theorem we take the trace inner product of the relative-error state design with the psd $(Z_i\otimes\mathbb{1})|\Omega\rangle\langle\Omega|(Z_i\otimes\mathbb{1})$. For an orthogonal O we get

$$\operatorname{tr}\left[(Z_i\otimes\mathbb{1})|\Omega_n\rangle\langle\Omega_n|(Z_i\otimes\mathbb{1})O^{\otimes 2}|00\rangle\langle 00|O^{T,\otimes 2}\right] = \frac{1}{2^n}\langle 00|O^{T,\otimes 2}Z_i^{\otimes 2}O^{\otimes 2}|00\rangle . \tag{41}$$

For part 1 we compute that

$$\begin{aligned}
\frac{1}{2^n}\mathbb{E}_{O\sim\mathrm{O}}\langle 00|O^{T,\otimes 2}Z_i^{\otimes 2}O^{\otimes 2}|00\rangle &= \frac{1}{2^n}\operatorname{tr}\left[Z_i^{\otimes 2}\Phi_{\mathrm{O}}(|00\rangle\langle 00|)\right] \\
&= \frac{1}{2^{2n}(2^n+1)}\operatorname{tr}\left[Z_i^{\otimes 2}(\mathbb{1}+\mathbb{F}+\Pi_o)\right] \\
&= \frac{1}{2^{2n}(2^n+1)}\operatorname{tr}\left[\mathbb{F}+\Pi_o\right] \\
&= \frac{2}{2^n(2^n+2)},
\end{aligned} \tag{42}$$

and similar for the perturbations Z_j and Z_iZ_j. We get a contradiction similar as in Eq. (18), provided that $\mathcal{E}$ is locally independent.

For part 2 we note that the real Clifford group is an orthogonal unitary 2-design[11] (and is moreover finite), and that in every layer we thus have a constant probability that the real Clifford factorizes in the first qubit. This implies the lower bound $d = \Omega(n)$ if $\mathcal{E}$ is a generalized random brickwork.

4 Orthogonal state designs: additive error

We prove the following theorem, which shows that the action of the Haar orthogonal twirl and the Haar unitary twirl are equal (up to small additive error) when applied to states that are unentangled between copies.

Theorem 16 (Equivalence of orthogonal and unitary designs on unentangled states)**.** *Consider a state ρ on $\mathcal{H}^{\otimes k}$. Suppose that ρ obeys the PPT condition on each copy of $\mathcal{H}$, i.e., $\rho^{\Gamma_i} \geq 0$ for all $i = 1,\ldots,k$, where Γ_i denotes the partial transpose on the i-th copy. Then the Haar orthogonal twirl of ρ is equal to the Haar unitary twirl of ρ up to additive error:*

$$\|\Phi_{\mathrm{O}}(\rho) - \Phi_{\mathrm{U}}(\rho)\|_1 \leq \mathcal{O}(k^{7/2}/2^n) + \mathcal{O}(k/2^{n/2}). \tag{43}$$

Before presenting the proof of the theorem, we briefly discuss its consequences for the construction of low-depth orthogonal state designs. Concretely we have the following corollary:

Corollary 17. *Consider the two-layer superblock brickwork circuit composed of orthogonal $O(\epsilon)$ additive error unitary k-designs acting on $\xi = \Omega(\log(nk/\epsilon))$ qubits. This ensemble, with depth $\mathcal{O}(\log(n)\operatorname{poly}(k)\log(1/\epsilon))$ is an ϵ-approximate additive-error orthogonal state k-design.*

Proof. Theorem 16 immediately implies that the two-layer brickwork circuit composed of small orthogonal unitaries acting on $\xi = \Omega(\log(nk/\varepsilon))$ qubits, is an orthogonal state design generator.

In order for a unitary ensemble to be an orthogonal state design generator, we require that it replicate the action of a Haar orthogonal unitary on states of the form $\rho = |\psi\rangle\langle\psi|^{\otimes k}$. Such states clearly obey the PPT condition in Theorem 16. In the context of the two-layer brickwork ensemble, this allows us to replace each small Haar orthogonal unitary with a small Haar unitary up to additive error $\mathcal{O}(nk^{7/2}/2^{\xi}) + \mathcal{O}(nk/2^{\xi/2})$. From Theorem 1 of Ref.[3], the two-layer brickwork circuit with small Haar unitaries forms a unitary design with relative error $\mathcal{O}(nk^2/2^{\xi})$. This translates to an additive error of the same magnitude. Finally, we can apply Theorem 16 again, to show that the Haar unitary twirl applied to ρ is equal to the Haar orthogonal twirl up to additive error $\mathcal{O}(k^{7/2}/2^n) + \mathcal{O}(k/2^{n/2})$. The total error from this sequence of steps is less than ε provided that $\xi = \Omega(\log(nk/\varepsilon))$, as claimed. Note that we can construct orthogonal designs in depth $\mathcal{O}(n\operatorname{poly}(k)\log(1/\epsilon))$ by Ref.[9], so we can replace the superblocks on ξ qubits by linear-depth orthogonal group designs. This all adds up to additive-error orthogonal state designs in depth $\mathcal{O}(\log(n/\epsilon)\operatorname{poly}(k))$.

We can now prove Theorem 16:

Proof. We prove that $\|\Phi_{\mathrm{O}}(\rho) - \Phi_{\mathrm{U}}(\rho)\|_1 \leq \mathcal{O}(k^{7/2}/2^n) + \mathcal{O}(k/2^{n/2})$. Our proof proceeds in five steps.

1. The Haar orthogonal twirl squares to itself, $\Phi_{\mathrm{O}} = \Phi_{\mathrm{O}} \circ \Phi_{\mathrm{O}}$.

2. One can insert projectors onto the distinct subspace after the first and second Haar orthogonal twirls, up to small additive error. Specifically, we will show that

$$\left\| [\mathbf{\Pi}_{\mathrm{dist}} \circ \Phi_{\mathrm{O}} \circ \mathbf{\Pi}_{\mathrm{dist}} \circ \Phi_{\mathrm{O}}](\rho) - [\Phi_{\mathrm{O}} \circ \Phi_{\mathrm{O}}](\rho) \right\|_1 \leq 5k/2^{n/2}, \tag{44}$$

   where $\mathbf{\Pi}_{\mathrm{dist}}(\cdot) \equiv \Pi_{\mathrm{dist}}(\cdot)\Pi_{\mathrm{dist}}$ denotes conjugation by the distinct subspace projector.

3. Within the distinct subspace, the Haar orthogonal twirl and Haar unitary twirl are equal up to small relative error. Specifically, we will show that

$$(1-\varepsilon')\left[\mathbf{\Pi}_{\mathrm{dist}} \circ \Phi_{\mathrm{U}} \circ \mathbf{\Pi}_{\mathrm{dist}}\right] \preceq \left[\mathbf{\Pi}_{\mathrm{dist}} \circ \Phi_{\mathrm{O}} \circ \mathbf{\Pi}_{\mathrm{dist}}\right] \preceq (1+\varepsilon')\left[\mathbf{\Pi}_{\mathrm{dist}} \circ \Phi_{\mathrm{U}} \circ \mathbf{\Pi}_{\mathrm{dist}}\right], \tag{45}$$

   with $\varepsilon' = \mathcal{O}(k^{7/2}/2^n)$. This allows us to replace the second Haar orthogonal twirl above with a Haar unitary twirl.

4. One can remove the projectors onto the distinct subspace after the Haar orthogonal and Haar unitary twirls, up to small additive error,

$$\left\| [\mathbf{\Pi}_{\mathrm{dist}} \circ \Phi_{\mathrm{U}} \circ \mathbf{\Pi}_{\mathrm{dist}} \circ \Phi_{\mathrm{O}}](\rho) - [\Phi_{\mathrm{U}} \circ \Phi_{\mathrm{O}}](\rho) \right\|_1 \leq 5k/2^{n/2}, \tag{46}$$

   similar to step 2.

5. The composition of the Haar unitary and Haar orthogonal twirl is equal to the Haar unitary twirl, $\Phi_{\mathrm{U}} \circ \Phi_{\mathrm{O}} = \Phi_{\mathrm{U}}$.

In total, these steps show that $\|\Phi_{\mathrm{O}}(\rho) - \Phi_{\mathrm{U}}(\rho)\|_1 \leq \mathcal{O}(k^{7/2}/2^n) + 10k/2^{n/2}$, as claimed.

Let us begin with step 2. From the triangle inequality, we have

$$\begin{aligned} &\left\| [\mathbf{\Pi}_{\mathrm{dist}} \circ \Phi_{\mathrm{O}} \circ \mathbf{\Pi}_{\mathrm{dist}} \circ \Phi_{\mathrm{O}}](\rho) - [\Phi_{\mathrm{O}} \circ \Phi_{\mathrm{O}}](\rho) \right\|_1 \\ &\qquad \leq \left\| [\mathbf{\Pi}_{\mathrm{dist}} \circ \Phi_{\mathrm{O}} \circ (\mathbf{\Pi}_{\mathrm{dist}} - \mathbb{1}) \circ \Phi_{\mathrm{O}}](\rho) \right\|_1 \\ &\qquad + \left\| [(\mathbf{\Pi}_{\mathrm{dist}} - \mathbb{1}) \circ \Phi_{\mathrm{O}} \circ \mathbb{1} \circ \Phi_{\mathrm{O}}](\rho) \right\|_1 . \end{aligned} \tag{47}$$

To proceed, note that the operation $\mathbf{\Pi}_{\text{dist}} \circ \Phi_O$ cannot increase the 1-norm of an operator. Thus, the first term above is upper bounded by $\left\| [(\mathbf{\Pi}_{\text{dist}} - \mathbb{1}) \circ \Phi_{\text{O}}](\rho) \right\|_1$. This is in fact equal to the second term, since $\Phi_{\text{O}} \circ \mathbb{1} \circ \Phi_{\text{O}} = \Phi_{\text{O}}$. These give,

$$\left\| [\mathbf{\Pi}_{\text{dist}} \circ \Phi_{\text{O}} \circ \mathbf{\Pi}_{\text{dist}} \circ \Phi_{\text{O}}](\rho) - [\Phi_{\text{O}} \circ \Phi_{\text{O}}](\rho) \right\|_1 \leq 2 \left\| [(\mathbf{\Pi}_{\text{dist}} - \mathbb{1}) \circ \Phi_{\text{O}}](\rho) \right\|_1 . \tag{48}$$

We will now bound the right side.

From the gentle measurement lemma, the right side is bounded as,

$$2 \left\| [(\mathbf{\Pi}_{\text{dist}} - \mathbb{1}) \circ \Phi_{\text{O}}](\rho) \right\|_1 \leq 4\sqrt{1 - \text{tr}(\Pi_{\text{dist}} \Phi_{\text{O}}(\rho))}, \tag{49}$$

where $\text{tr}(\Pi_{\text{dist}} \Phi_{\text{O}}(\rho))$ is the probability of measuring $\Phi_{\text{O}}(\rho)$ in the distinct subspace. We can now apply the operator inequality, $\mathbb{1} - \Pi_{\text{dist}} \leq \sum_{i<j} \Pi_{ij}$, where $\Pi_{ij} = \sum_{x \in \{0,1\}^n} |x\rangle\langle x|_i \otimes |x\rangle\langle x|_j$ projects onto states with the same bitstring in copy i as in copy j. This yields,

$$2 \left\| [(\mathbf{\Pi}_{\text{dist}} - \mathbb{1}) \circ \Phi_{\text{O}}](\rho) \right\|_1 \leq 4\sqrt{\sum_{i<j} \text{tr}(\Pi_{ij} \Phi_{\text{O}}(\rho))}. \tag{50}$$

We can bound each probability as follows. Note that $\text{tr}(\Pi_{ij} \Phi_{\text{O}}(\rho)) = \text{tr}(\Phi_{\text{O}}(\Pi_{ij})\rho)$, since $\Phi_{\text{O}}^T = \Phi_{\text{O}}$. The twirl of Π_{ij} over an orthogonal 2-design yields, by Eq. (40),

$$\Phi_{\text{O}}(\Pi_{ij}) = \frac{1}{2^n + 2} \left(\mathbb{1}_{ij} + \mathbb{F}_{ij} + \Pi_{o,ij} \right). \tag{51}$$

For the first two terms, we have $\text{tr}(\mathbb{1}_{ij}\rho) = 1$, and $|\text{tr}(\mathbb{F}_{ij}\rho)| \leq 1$ since $\|\mathbb{F}_{ij}\|_\infty \leq 1$. To bound the third term, we apply the PPT condition on subsystem i. We have,

$$\text{tr}(\Pi_{o,ij}\rho) = \text{tr}\left((\Pi_{o,ij})^{\Gamma_i} \rho^{\Gamma_i} \right) \leq \|(\Pi_{o,ij})^{\Gamma_i}\|_\infty \cdot \|\rho^{\Gamma_i}\|_1 = \|\mathbb{F}_{ij}\|_\infty \cdot 1 = 1. \tag{52}$$

In the second step, we apply Holder's inequality. In the third step, we use that $(\Pi_{o,ij})^{\Gamma_i} = \mathbb{F}_{ij}$, and that $\|\rho^{\Gamma_i}\|_1 = \text{tr}\left(\rho^{\Gamma_i}\right) = \text{tr}(\rho) = 1$, where the first equality follows from the PPT condition, $\rho^{\Gamma_i} \geq 0$. In total, these bounds yield

$$\text{tr}(\Pi_{ij} \Phi_{\text{O}}(\rho)) \leq \frac{3}{2^n + 2}. \tag{53}$$

Inserting into Eq. (50) and performing the sum over $i < j$, we find

$$2 \left\| [(\mathbf{\Pi}_{\text{dist}} - \mathbb{1}) \circ \Phi_{\text{O}}](\rho) \right\|_1 \leq 4\sqrt{\frac{k(k-1)}{2} \frac{3}{2^n + 2}} \leq \frac{5k}{2^{n/2}}, \tag{54}$$

which completes our proof of step 2.

We can prove step 4 in an extremely similar manner. Following the same logic as for step 2, we have

$$\begin{aligned} &\left\| [\mathbf{\Pi}_{\text{dist}} \circ \Phi_{\text{U}} \circ \mathbf{\Pi}_{\text{dist}} \circ \Phi_{\text{O}}](\rho) - [\Phi_{\text{U}} \circ \Phi_{\text{O}}](\rho) \right\|_1 \\ &\qquad \leq \left\| [(\mathbf{\Pi}_{\text{dist}} - \mathbb{1}) \circ \Phi_{\text{O}}](\rho) \right\|_1 + \left\| [(\mathbf{\Pi}_{\text{dist}} - \mathbb{1}) \circ \Phi_{\text{U}}](\rho) \right\|_1 . \end{aligned} \tag{55}$$

From the previous paragraph, the first term is upper bounded by $(5k/2^{n/2})/2$. Moreover, the second term is upper bounded by the same value. This follows since $\Phi_U = \Phi_{\text{O}} \circ \Phi_{\text{U}}$, and $\Phi_{\text{U}}(\rho)$ obeys the PPT condition because ρ obeys it. Thus, Eq. (55) is bounded by $5k/2^{n/2}$, as claimed.

It remains only to prove step 3. To show this, we first establish a technical lemma, which allows one to easily bound the relative error *within the distinct subspace* between the twirl over any unitary ensemble $\mathcal{E}$ and the Haar unitary twirl. The lemma is closely adapted from Lemma 2 of Ref.[3]. To more easily bound the error, it leverages the "approximate Haar unitary twirl"[3],

$$\Phi_a(\cdot) = \frac{1}{2^{nk}} \sum_{\pi \in S_k} \text{tr}\left((\cdot)\pi^{-1} \right) \cdot \pi, \tag{56}$$

which is close to the exact Haar unitary twirl up to relative error $k^2/2^n$ (Lemma 1 of Ref.[3]).

Lemma 18 (Relative error on the distinct subspace from EPR states)**.** *Consider a unitary ensemble $\mathcal{E}$ and its twirl $\Phi_{\mathcal{E}}$. Within the distinct subspace, the twirl is approximated by Φ_U up to relative error,*

$$\varepsilon' = \frac{4^{nk}}{k!}\left(1+\frac{k^2}{2^n}\right)\left\|\left[\Pi_{\text{dist}}\otimes\Pi_{\text{dist}}\right]\left[(\Phi_{\mathcal{E}}-\Phi_a)\otimes\mathbb{1}\right](|\Omega_n\rangle\langle\Omega_n|)\left[\Pi_{\text{dist}}\otimes\Pi_{\text{dist}}\right]\right\|_\infty + \frac{k^2}{2^n}. \tag{57}$$

That is, we have $(1-\varepsilon')\,\mathbf{\Pi}_{\text{dist}}\circ\Phi_{\text{U}}\circ\mathbf{\Pi}_{\text{dist}} \preceq \mathbf{\Pi}_{\text{dist}}\circ\Phi_{\mathcal{E}}\circ\mathbf{\Pi}_{\text{dist}} \preceq (1+\varepsilon')\,\mathbf{\Pi}_{\text{dist}}\circ\Phi_{\text{U}}\circ\mathbf{\Pi}_{\text{dist}}$, with ε' as above.

Proof. The proof follows identically to the proof of Lemma 2 and Lemma 7 in Ref.[3]. The only difference is the presence of the distinct subspace projector. Crucially, in the proof of Lemma 2, this modification does not change the eigenvalue of $[\Phi_a\otimes\mathbb{1}](|\Omega_n\rangle\langle\Omega_n|) = \frac{1}{4^{nk}}\sum_\pi \pi\otimes\pi$ (which has a flat spectrum), owing to the fact that $\Pi_{\text{dist}}\otimes\Pi_{\text{dist}}$ and $[\Phi_a\otimes\mathbb{1}](|\Omega_n\rangle\langle\Omega_n|)$ commute, since Π_{dist} and π commute for all π.

We can now apply the lemma to bound the relative error between $\mathbf{\Pi}_{\text{dist}}\circ\Phi_{\text{O}}\circ\mathbf{\Pi}_{\text{dist}}$ and $\mathbf{\Pi}_{\text{dist}}\circ\Phi_U\circ\mathbf{\Pi}_{\text{dist}}$. To do so, we first compute $[\Phi_O\otimes\mathbb{1}](|\Omega_n\rangle\langle\Omega_n|)$,

$$[\Phi_{\text{O}}\otimes\mathbb{1}](|\Omega_n\rangle\langle\Omega_n|) = \frac{1}{2^{nk}}\sum_{\sigma,\tau\in M_k}\text{Wg}_{\text{O}}(\sigma,\tau)[\tau\otimes\sigma], \tag{58}$$

where σ,τ are summed over the generators of the commutant of the orthogonal group. Multiplying by the distinct subspace projectors yields,

$$\left[\Pi_{\text{dist}}\otimes\Pi_{\text{dist}}\right]\left[\Phi_{\text{O}}\otimes\mathbb{1}\right](|\Omega_n\rangle\langle\Omega_n|)\left[\Pi_{\text{dist}}\otimes\Pi_{\text{dist}}\right] = \frac{1}{2^{nk}}\sum_{\pi,\tilde\pi\in S_k}\text{Wg}_{\text{O}}(\pi,\tilde\pi)\left[\Pi_{\text{dist}}\pi\,\Pi_{\text{dist}}\otimes\Pi_{\text{dist}}\tilde\pi\,\Pi_{\text{dist}}\right], \tag{59}$$

where $\pi,\tilde\pi$ are summed over the generators of the commutant of the unitary group. This follows since all other members of the commutant vanish upon multiplication with the projector, i.e., $\sigma\,\Pi_{\text{dist}} = 0$ if $\sigma\notin S_k$. Taking the difference with $\left[\Pi_{\text{dist}}\otimes\Pi_{\text{dist}}\right]\left[\Phi_a\otimes\mathbb{1}\right](|\Omega_n\rangle\langle\Omega_n|)\left[\Pi_{\text{dist}}\otimes\Pi_{\text{dist}}\right]$ gives,

$$\begin{aligned}&\left[\Pi_{\text{dist}}\otimes\Pi_{\text{dist}}\right]\left[(\Phi_{\text{O}}-\Phi_a)\otimes\mathbb{1}\right](|\Omega_n\rangle\langle\Omega_n|)\left[\Pi_{\text{dist}}\otimes\Pi_{\text{dist}}\right]\\ &\qquad= \frac{1}{2^{nk}}\sum_{\pi,\tilde\pi\in S_k}\left(\text{Wg}_{\text{O}}(\pi,\tilde\pi)-\frac{1}{2^{nk}}\delta_{\pi,\tilde\pi}\right)\left[\Pi_{\text{dist}}\pi\,\Pi_{\text{dist}}\otimes\Pi_{\text{dist}}\tilde\pi\,\Pi_{\text{dist}}\right].\end{aligned} \tag{60}$$

Applying the triangle inequality, the spectral norm of the above is bounded by

$$\|\ldots\|_\infty \le \frac{1}{2^{nk}}\sum_{\pi,\tilde\pi\in S_k}\left|\text{Wg}_{\text{O}}(\pi,\tilde\pi)-\frac{1}{2^{nk}}\delta_{\pi,\tilde\pi}\right| = \frac{k!}{2^{nk}}\sum_{\pi\in S_k}\left|\text{Wg}_{\text{O}}(\pi)-\frac{1}{2^{nk}}\delta_{\pi,\mathbb{1}}\right|, \tag{61}$$

where in the third expression we use that the Weingarten function depends only on the difference, $\pi\tilde\pi^{-1}\to\pi$, if $\pi,\tilde\pi$ are permutations and thus invertible. Separating out the identity term, we have

$$\sum_{\pi\in S_k}\left|\text{Wg}_{\text{O}}(\pi)-\frac{1}{2^{nk}}\delta_{\pi,\mathbb{1}}\right| \le \sum_{\pi\in S_k\setminus\{\mathbb{1}\}}|\text{Wg}_{\text{O}}(\pi)| + \left|\text{Wg}_{\text{O}}(\mathbb{1})-\frac{1}{2^{nk}}\right|. \tag{62}$$

By Theorem 4.11 of[12], we deduce that $\left|\text{Wg}_{\text{O}}(\mathbb{1})-\frac{1}{2^{nk}}\right| = \frac{1}{2^{nk}}\mathcal{O}\left(\frac{k^7}{2^{2n}}\right)$. We prove a lemma for the other term.

Lemma 19. *If $2^n > 13k^{7/2}$, then*

$$\sum_{\pi\in S_k\setminus\{\mathbb{1}\}}|\text{Wg}_{\text{O}}(\pi)| = \frac{1}{2^{nk}}\,\mathcal{O}\left(\frac{k^2}{2^n}\right). \tag{63}$$

Proof of Lemma 19. We know by Theorem 4.11 of[12] that for a permutation σ we have

$$|\operatorname{Wg}_{\mathrm{O}}(\sigma)| \leq \frac{1}{2^{n(k+|\sigma|)}} \frac{1}{1-\frac{144k^7}{2^{2n}}} \prod_{i=1}^{\#\sigma} \frac{(2\mu_i-2)!}{(\mu_i-1)!\mu_i!}, \tag{64}$$

where $\#\sigma$ is the number of cycles of σ. We defined μ_i as the cycle type of σ (the sizes of the cycles). We defined $|\sigma| \coloneqq k - \#\sigma$. We note that $\sum_i \mu_i = k$, while $\sum_i 1 = \#\sigma$, which implies that $\sum_i(\mu_i - 1) = k - \#\sigma = |\sigma|$. We can now compute

$$\begin{aligned}
\prod_{i=1}^{\#\sigma} \frac{(2\mu_i-2)!}{(\mu_i-1)!\mu_i!} &= \prod_{i=1}^{\#\sigma} \frac{1}{\mu_i} \binom{2(\mu_i-1)}{\mu_i-1} \\
&\leq \prod_{i=1}^{\#\sigma} 2^{2(\mu_i-1)} \\
&= 2^{2\sum_i(\mu_i-1)} \\
&= 2^{2|\sigma|}.
\end{aligned} \tag{65}$$

The second step follows from standard bounds on the central binomial coefficient: $\binom{2m}{m} \leq 2^{2m}$. By Eq. (64) we then have that

$$\sum_{\sigma \in S_k \setminus \{\mathbb{1}\}} |\operatorname{Wg}_{\mathrm{O}}(\sigma)| \leq \frac{1}{2^{nk}(1-\frac{144k^7}{2^{2n}})} \sum_{\sigma \in S_k \setminus \{\mathbb{1}\}} 2^{-(n-2)|\sigma|}. \tag{66}$$

Note that the number of permutations with exactly j cycles is given by the unsigned Stirling numbers $c(k,j)$. Moreover, it is well-known that their generating function over j is given by the rising factorial such that we find:

$$\begin{aligned}
\sum_{\sigma \in S_k \setminus \{\mathbb{1}\}} |\operatorname{Wg}_{\mathrm{O}}(\sigma)| \leq& \frac{1}{2^{nk}(1-\frac{144k^7}{2^{2n}})} \left(2^{-(n-2)k} \sum_{j=0}^{k} c(k,j) 2^{(n-2)j} - 1 \right) \\
\leq& \frac{1}{2^{nk}(1-\frac{144k^7}{2^{2n}})} \left(\prod_{j=0}^{k-1} \frac{2^{n-2}+j}{2^{n-2}} - 1 \right) \\
=& \frac{1}{2^{nk}(1-\frac{144k^7}{2^{2n}})} \left((1+k2^{-n+2})^k - 1 \right) \\
\leq& \frac{1}{2^{nk}(1-\frac{144k^7}{2^{2n}})} \left(e^{k^2 2^{-n+2}} - 1 \right) \\
\leq& \frac{1}{2^{nk}(1-\frac{144k^7}{2^{2n}})} \frac{1}{\log(2)} k^2 2^{-n+2},
\end{aligned} \tag{67}$$

where in the final step we used that $k^2 2^{-n+2} \leq \log(2)$ such that we can use the inequality $e^x \leq 1 + x/\log(2)$. As we assumed that $2^n > 13k^{7/2}$, we see that $\frac{1}{1-\frac{144k^7}{2^{2n}}} = \mathcal{O}(1)$. We get $\sum_{\sigma \in S_k \setminus \{\mathbb{1}\}} |\operatorname{Wg}_{\mathrm{O}}(\sigma)| = \frac{1}{2^{nk}} \mathcal{O}\left(\frac{k^2}{2^n}\right)$, which is what we wanted to prove.

Using Lemma 19 we get

$$\sum_{\pi \in S_k} \left| \operatorname{Wg}_{\mathrm{O}}(\pi) - \frac{1}{2^{nk}} \delta_{\pi,\mathbb{1}} \right| \leq \frac{1}{2^{nk}} \cdot \left(\mathcal{O}\left(\frac{k^7}{2^{2n}}\right) + \mathcal{O}\left(\frac{k^2}{2^n}\right) \right) = \frac{1}{2^{nk}} \mathcal{O}\left(\frac{k^{7/2}}{2^n}\right), \tag{68}$$

using that $k^{7/2} < 2^n$. From Eq. (61) and Lemma 18, this leads to a relative error $\mathcal{O}\left(\frac{k^{7/2}}{2^n}\right)$, as claimed. This completes our proof of step 3, and hence of Theorem 16.

5 Orthogonal anti-concentration

Theorem 20. *Let $\mathcal{E}$ be the ensemble of orthogonal matrices drawn from the two-layer superblock architecture. This gives us two layers of $m/2 = n/(2\xi)$ superblocks each, where $\xi = \mathcal{O}(\log(n/\epsilon))$. This ensemble anti-concentrates for $k = 2$, i.e.,*

$$\mathbb{E}_{O\sim\mathcal{E}} \left|\langle 0| O |0_n\rangle\right|^4 \leq (1+\epsilon)\ \mathbb{E}_{O\sim \mathrm{O}} \left|\langle 0| O |0_n\rangle\right|^4. \tag{69}$$

Proof. Let $\mathcal{E}_1, \mathcal{E}_2$ be the ensembles consisting of orthogonal matrices sampled from the first and second layer of superblocks respectively. We thus have $\Phi_{\mathcal{E}} = \Phi_{\mathcal{E}_2} \circ \Phi_{\mathcal{E}_1}$. We compute that

$$\begin{aligned}
\mathbb{E}_{O\sim\mathcal{E}} \left|\langle 0^n| O |0^n\rangle\right|^4 &= \operatorname{tr}\left[\Phi_{\mathcal{E}_1}(|0^n\rangle\langle 0^n|^{\otimes 2})\Phi_{\mathcal{E}_2}(|0^n\rangle\langle 0^n|^{\otimes 2})\right] \\
&= \frac{1}{2^{2n}(2^{2\xi}+2)^m} \operatorname{tr}\left[(\mathbb{1}+\mathbb{F}+\Pi_o)^{\otimes m/2} \overrightarrow{(\mathbb{1}+\mathbb{F}+\Pi_o)^{\otimes m/2}}\right] \\
&\leq \frac{1}{2^{2n}(2^{2n}+2)} \operatorname{tr}\left[(\mathbb{1}+\mathbb{F}+\Pi_o)^{\otimes m/2} \overrightarrow{(\mathbb{1}+\mathbb{F}+\Pi_o)^{\otimes m/2}}\right],
\end{aligned} \tag{70}$$

where the arrow indicates that the operators $\mathbb{1}+\mathbb{F}+\Pi_o$ act to the patches $(2,3),\ldots,(m,1)$. Let us now expand the product as a sum

$$\operatorname{tr}\left[(\mathbb{1}+\mathbb{F}+\Pi_o)^{\otimes m/2} \overrightarrow{(\mathbb{1}+\mathbb{F}+\Pi_o)^{\otimes m/2}}\right] = \sum_{a,b\in\{\mathbb{1},\mathbb{F},\Pi_o\}^{m/2}} \operatorname{tr}\Big[a_1 b_{m/2} \otimes a_1 b_1 \otimes \cdots \otimes a_{m/2} b_{m/2}\Big], \tag{71}$$

where $a_i b_j$ acts on two copies of ξ qubits. We can easily compute that

$$\operatorname{tr}[a_i b_j] = \begin{cases} 2^{2\xi} & \text{if } a_i = b_j \\ 2^{\xi} & \text{if } a_i \neq b_j. \end{cases} \tag{72}$$

This means that a term from the sum only depends on the number of *walls*, i.e. the number of i such that $a_i \neq b_i$ plus the number of i such that $a_i \neq b_{i+1}$. Let us define $w_{a,b}$ as the number of walls for given $a, b \in \{\mathbb{1},\mathbb{F},\Pi_o\}^{m/2}$. We define a *room* as the set of patches in between two walls. We thus get that

$$\frac{1}{2^{2n}} \sum_{a,b\in\{\mathbb{1},\mathbb{F},\Pi_o\}^{m/2}} \operatorname{tr}\Big[a_1 b_{m/2} \otimes a_1 b_1 \otimes \cdots \otimes a_{m/2} b_{m/2}\Big] = \sum_{a,b\in\{\mathbb{1},\mathbb{F},\Pi_o\}^{m/2}} \frac{1}{2^{\xi\cdot w_{a,b}}}. \tag{73}$$

To upper bound the expression, we will count the number of configurations a, b, that give a number of walls w. There are $\binom{2m}{w}$ possible positions where the walls could be placed. Furthermore, inside a room (an interval bounded by two walls), we need to have $a_i = b_j$. This can be achieved in three different ways (i.e. a_i equals $\mathbb{1}, \mathbb{F}$ or Π_o). We thus get a factor of $3^{\min\{1,w\}}$, as $\min\{1,w\}$ is the total number of rooms. As we are only interested in upper bounding the expression, we can ignore the fact that two adjacent rooms need to have a different sign ($\mathbb{1}$, $\mathbb{F}$ or Π_o). We can also ignore the fact that there exists no configuration with exactly 1 wall. Let us upper bound our equation as

$$\begin{aligned}
\sum_{a,b\in\{\mathbb{1},\mathbb{F},\Pi_o\}^{m/2}} \frac{1}{2^{\xi\cdot w_{a,b}}} &\leq \sum_{w=0}^{2m} 3^{\min\{1,w\}} \binom{2m}{w} \frac{1}{2^{\xi\cdot w}} \\
&\leq 3 \sum_{w=0}^{2m} \binom{2m}{w} \frac{1}{2^{\xi\cdot w - \log_2(3) w}} \\
&= 3\left(1 + \frac{1}{2^{\xi-\log_2(3)}}\right)^{2m} \\
&\leq 3 \exp\left(\frac{2m}{2^{\xi-\log_2(3)}}\right) \\
&\leq 3\left(1 + \frac{2m}{\log_e(2)\cdot 2^{\xi-\log_2(3)}}\right).
\end{aligned} \tag{74}$$

In the final step we used that $\exp(x) \leq 1 + x/\log_e(2)$ for $0 \leq x \leq \log_e(2)$. We use $\xi = \mathcal{O}(\log(n/\epsilon))$ to get

$$\frac{3 \cdot 2m}{\log_e(2) \cdot 2^{\xi - \log_2(3)}} \leq \epsilon. \tag{75}$$

Combining all the things we learned, we find that

$$\mathbb{E}_{O\sim\mathcal{E}} \left|\langle 0_n| O |0_n\rangle\right|^4 \leq (1+\epsilon)\frac{3}{2^n(2^n+2)} = (1+\epsilon)\ \mathbb{E}_{O\sim \mathrm{O}} \left|\langle 0_n| O |0_n\rangle\right|^4, \tag{76}$$

which is what we set out to prove.

We note that Theorem 20 also follows from the computation for random MPS in Ref.[13] along the lines of Ref.[14]. In this way, higher-order anticoncentration can be shown as well.

Supplementary Note 4 — Unitary symplectic group

1 Preliminaries

We define a *symplectic form* as a $D \times D$ matrix J such that $J^2 = -\mathbb{1}$ and $J^T = J^\dagger = -J$. Note that for such an operator to exist, the dimension of the underlying space has to be even. We say that a unitary $U \in \mathrm{U}(D)$ is symplectic if $UJU^T = J$ and denote the group of unitary symplectic matrices as $\mathrm{USp}(D, J) \equiv \mathrm{USp}(D)$. If the context is clear, we will simply call $\mathrm{USp}(D)$ the symplectic group.

We note that we could make different choices of J, but that they are all equal up to a basis change. The unitary symplectic group acts transitively on the set of pure quantum states, i.e., for every pair of states (ψ, φ) there is a $U \in \mathrm{USp}(D)$ such that $U|\psi\rangle = |\varphi\rangle$[15]. Clearly, this does not depend on the choice of J. A consequence of transitivity is that $\mathrm{USp}(D)$ induces the same measure as $\mathrm{U}(D)$ on pure quantum states and thus the ensemble $\{U|\psi\rangle\}_{U\sim\mathrm{USp}(D)}$ is an exact state design of any order, see also Ref.[16].

Another important note is that the mentioned basis change does not need to be local, so when studying the depth of circuits consisting of local gates, it is important to take into account what J is. Here, we consider symplectic forms over n qubits, $D = 2^n$, of the form $J = J_x := i^{|x|}\bigotimes_{j=1}^n Y^{x_j}$, where $x \in \{0,1\}^n$ is such that the Hamming weight $|x|$ of x is odd. For instance, if we choose $x = (100\ldots0)$, then $J_1 = iY_1$ is what is usually considered the standard choice of symplectic form and what has been studied in the context of local circuits before[16,17]. In this paper, we will however choose $x = (11\ldots11)$, and assume that n is odd. We thus define $J = i^n\bigotimes_{j=1}^n Y_j$.

Let $U = U_r \otimes \mathbb{1}_{n-r}$ be a local gate acting on r qubits. For U to be symplectic for our choice of J, we need that $U_r Y^{\otimes r} U_r^T = Y^{\otimes r}$. This means that U_r can only be symplectic as a unitary on r qubits, if r is odd. It is therefore not natural to look at symplectic circuits with local gates that act on 2 qubits. A more natural approach is to use local gates on 3 qubits. For example, it is known that the 3-qubit gate $U_3 := \exp(i(Z_1 \otimes Z_2 \otimes Z_3))$, together with all 1-qubit gates, span the symplectic group[18]. The number of qubits one wants to apply local gates on, is however really dependent on the choice of J.

We quickly review some symplectic Weingarten calculus. The commutant of the symplectic representation $U \to U^{\otimes k}$ is a representation of the Brauer algebra, but by a different representation than in the orthogonal case. We can now define the representation R_{USp} as

$$R_{\mathrm{USp}}(\sigma) := \left(\prod_{j=1}^{2k} J_j^{a_\sigma(j)}\right) \sum_{i_1,\ldots i_{2k}} \left(\prod_{j=1}^{2k} \delta_{i_j, i_{f_\sigma(j)}}\right) |i_1 \ldots i_k\rangle\langle i_{k+1}\ldots i_{2k}| \left(\prod_{j=1}^{2k} J_j^{b_\sigma(j)}\right), \tag{77}$$

where we defined $a_\sigma(j) = 1$ if $j < f_\sigma(j) \leq k$ and 0 otherwise. Likewise, $b_\sigma(j) = 1$ if $k < j < f_\sigma(j)$ and 0 otherwise. J_j is the operator J acting on subsystem j. R_{USp} is a representation of $\mathcal{B}_k(-D)$, with D the dimension.

We note that for $k = 2$ the commutant of the symplectic group is given by $\mathbb{1}, \mathbb{F}$ and $\Pi_s := (J \otimes \mathbb{1}) |\Omega_n\rangle\langle\Omega_n| (J \otimes \mathbb{1})$, where we recall that $|\Omega_n\rangle = \frac{1}{2^{-n/2}} \sum_{i \in \{0,1\}^n} |ii\rangle$. We compute the Gram and Weingarten matrices

$$G_{\mathrm{USp}} = D \begin{pmatrix} D & 1 & -1 \\ 1 & D & 1 \\ -1 & 1 & D \end{pmatrix}, \ \mathrm{Wg}_{\mathrm{USp}} = \frac{1}{D(D+1)(D-2)} \begin{pmatrix} D-1 & -1 & 1 \\ -1 & D-1 & -1 \\ 1 & -1 & D-1 \end{pmatrix}. \tag{78}$$

Hence,

$$\begin{aligned} \mathbb{E}_{U \sim \mathrm{USp}} U^{\otimes 2} X U^{\dagger,\otimes 2} = \frac{1}{D(D+1)(D-2)} \Big(& [(D-1)\operatorname{tr}(X) - \operatorname{tr}(X\mathbb{F}) + \operatorname{tr}(X\Pi_s)]\, \mathbb{1} \\ & + [-\operatorname{tr}(X) + (D-1)\operatorname{tr}(X\mathbb{F}) - \operatorname{tr}(X\Pi_s)]\, \mathbb{F} \\ & + [\operatorname{tr}(X) - \operatorname{tr}(X\mathbb{F}) + (D-1)\operatorname{tr}(X\Pi_s)]\, \Pi_s \Big). \end{aligned} \tag{79}$$

2 Symplectic group designs

Our lower bound on the depth of additive-error symplectic group designs proceeds in an exactly parallel manner to our lower bound on additive-error orthogonal group designs. We consider the canonical choice $J = i^n \bigotimes_{j=1}^n Y_j$, for which we prove the following theorem.

Theorem 21. *Let $\mathcal{E}$ be an approximate symplectic unitary k-design with additive error $\varepsilon < 1/4$. Then at least one symplectic unitary must have a lightcone covering a $1/4$'th fraction of the qubits. This implies $\mathcal{E}$ must have circuit depth $d = \Omega(n)$ in 1D circuits and $d = \Omega(\log n)$ in all-to-all connected circuits for any $k \geq 2$.*

For completeness, we can also consider the alternative choice $J = iY_1$ introduced in[17]. For this choice, all the local gates needed to construct global symplectic gates are all orthogonal, except the gates acting on the first qubit[17]. This implies that the depth lower bounds for orthogonal group designs (in Section 2, based on[3]) immediately hold for this choice of symplectic group designs as well.

Proof. To prove the theorem, we will show that any symplectic unitary with lightcone smaller than the system size can be distinguished from a Haar-random symplectic unitary using two parallel queries and constant-depth quantum and classical operations.

Let U be a short-depth symplectic unitary and $|J_n\rangle \equiv (J \otimes \mathbb{1}) |\Omega_n\rangle$ the EPR state acted on by J on the first copy. The state $|J_n\rangle$ can be prepared in constant depth since it is a tensor product of n two-qubit stabilizer states, $|J_n\rangle = |J_1\rangle^{\otimes n}$. Our distinguishing protocol proceeds as follows. We first perturb the J-state by a single-qubit Pauli operator Z_1 on the first qubit (taking an arbitrary but fixed ordering of the qubits) of the first copy. We then apply U to both copies in parallel. Finally, we choose $i \in [n]$ uniformly at random and take the trace inner product with a Pauli-Z operator on both copies of qubit i, $P^{(i)} = Z_i^{\otimes 2}$.

This procedure yields an expectation value,

$$\mathbb{E}_{i \sim [n]} \langle\Omega_n| (JZ_1 \otimes \mathbb{1})(U^\dagger \otimes U^\dagger) P^{(i)} (U \otimes U)(Z_1 J \otimes \mathbb{1}) |\Omega_n\rangle = \mathbb{E}_{i \sim [n]} \langle\Omega_n| (JUZ_1U^\dagger \otimes \mathbb{1}) P^{(i)} (UZ_1U^\dagger J \otimes \mathbb{1}) |\Omega_n\rangle .$$

In the second expression, we used $(\mathbb{1} \otimes U) |\Omega_n\rangle = (U^T \otimes \mathbb{1}) |\Omega_n\rangle$ and $JU^T = U^\dagger J$ and $U^* J = JU$. Note that $P^{(i)}$ and UZ_1U^T commute whenever i is not in the lightcone of UZ_1U^T. The proof now finished in the same manner as Theorem 14.

Moreover, we find a linear lower bound for generalized random brickworks.

Theorem 22. *Let $\mathcal{E}$ be an approximate symplectic unitary k-design with relative error $\varepsilon < 1$. If $\mathcal{E}$ is a generalized random brickwork, then $d = \Omega(n)$.*

Proof. We prepare our system in the state $(Z_1 \otimes \mathbb{1})\,|J_n\rangle$ and then apply a symplectic unitary on both systems. The probability that we measure $(\mathbb{1} \otimes J)\,|0^n\rangle^{\otimes 2}$ is

$$|\,\langle J_n|\,(Z_1 \otimes \mathbb{1})U^{\otimes 2}(\mathbb{1} \otimes J)\,|0^n\rangle^{\otimes 2}\,|^2 = |\,\langle 0^n|\,U^\dagger Z_1 U\,|0^n\rangle\,|^2 \tag{80}$$

For U drawn from the symplectic Haar measure, we get

$$\begin{aligned}\mathbb{E}_{U\sim \mathrm{USp}}|\,\langle J_n|\,(Z_1 \otimes \mathbb{1})U^{\otimes 2}(\mathbb{1} \otimes J)\,|0^n\rangle^{\otimes 2}\,|^2 &= \frac{\mathrm{tr}[(Z_1 \otimes \mathbb{1})\,|J_n\rangle\langle J_n|\,(Z_1 \otimes \mathbb{1})(2^n\mathbb{1} - 2\mathbb{F} + 2^n\Pi_s)]}{2^n(2^n+1)(2^n-2)} \\ &= \frac{1}{2^n(2^n+1)}.\end{aligned} \tag{81}$$

We note that the symplectic Clifford group (i.e. all Clifford gates that are symplectic) form a symplectic 2-design for any number of qubits (this group is also called the semi-real Clifford group[19], and corresponds to case B in Theorem 1.6 in[20] which proves this group is a 3-design for the unitary symplectic group). In the case that $\mathcal{E}$ is a generalized random brickwork, we can follow the same argument as for the Clifford and orthogonal case. We thus get that

$$\mathbb{E}_{U\sim\mathcal{E}}|\,\langle 0^n|\,U^\dagger Z_1 U\,|0^n\rangle\,|^2 \geq \Omega(p^d), \tag{82}$$

where p is the probability that a local gate (in the symplectic Clifford group) is not entangling. We get an architecture-independent linear lower bound on the depth of symplectic relative-error designs.

3 Symplectic state designs: relative error

In this section we prove a lower bound on the depth of relative-error state designs with local symplectic gates.

Theorem 23. *Let $\mathcal{E}$ be an ensemble of local symplectic circuits with respect to $J = i^n \bigotimes_{j=1}^n Y_j$ of depth d, such that $\Phi_{\mathcal{E}}(|0^n\rangle\langle 0^n|^{\otimes 2})$ is a relative-error ϵ-approximate state 2-design for $\epsilon < 1$. The lightcone starting from any qubit should contain all other qubits. In particular this implies a depth lower bound of $\Omega(n)$ in* 1D *architectures and $\Omega(\log n)$ in the all-to-all architecture.*

Note that the theorem makes a stronger statement than Theorems 11 and 15 in the all-to-all connectivity. Also, we do not require the circuit to be locally independent.

Proof. Let $\mathcal{E}$ be an ensemble of locally symplectic circuits of depth $d < n$. We define $|J\rangle := (J \otimes \mathbb{1}_{n\times n})\,|\Omega_n\rangle$, where $|\Omega_n\rangle$ is the n qubit Bell state.

The idea of the proof is that we evaluate the symplectic twirl Φ_{USp} and the $\Phi_{\mathcal{E}}$ on initial state $|0^n\rangle\langle 0^n|^{\otimes 2}$, and we show that they differ when we take the trace inner-product with the psd operator $(Y_1 \otimes \mathbb{1})\,|J\rangle\langle J|\,(Y_1 \otimes \mathbb{1})$. More precisely, we show the following two claims.

Claim 1: When evaluating the twirl over $\mathcal{E}$ we have that

$$\mathrm{tr}\Big[(Y_1 \otimes \mathbb{1})\,|J\rangle\langle J|\,(Y_1 \otimes \mathbb{1})\Phi_{\mathcal{E}}\left(|0^n\rangle\langle 0^n|^{\otimes 2}\right)\Big] = 0. \tag{83}$$

Claim 2: When evaluating the symplectic Haar twirl we get

$$\mathrm{tr}\Big[(Y_1 \otimes \mathbb{1})\,|J\rangle\langle J|\,(Y_1 \otimes \mathbb{1})\Phi_{\mathrm{USp}}\left(|0^n\rangle\langle 0^n|^{\otimes 2}\right)\Big] > 0. \tag{84}$$

These two claims together directly imply that $(1-\epsilon)\Phi_{\mathrm{USp}}\left(|0^n\rangle\langle 0^n|^{\otimes 2}\right) \preceq \Phi_{\mathcal{E}}\left(|0^n\rangle\langle 0^n|^{\otimes 2}\right)$ cannot be true for $\epsilon < 1$, thus proving the theorem.

Proof of Claim 1: We note that $U^{\otimes 2}|J\rangle = |J\rangle$ for every symplectic U. For $U \in \mathcal{E}$, we then have $U^{\otimes 2}(Y_j \otimes \mathbb{1})|J\rangle = \left(U_l^{\otimes 2}(Y_j \otimes \mathbb{1})|J_{x_l}\rangle\right) \otimes |J_{\overline{x_l}}\rangle$, where U_l is the part of the circuit U that lies in the lightcone of qubit j. All local gates outside of this lightcone vanish, which leaves us with $|J_{\overline{x_l}}\rangle$ on the qubits outside of the lightcone. We can thus compute

$$\begin{aligned}
&\mathrm{tr}\Big[(Y_1 \otimes \mathbb{1})|J\rangle\langle J|(Y_1 \otimes \mathbb{1})\Phi_{\mathcal{E}}\left(|0^n\rangle\langle 0^n|^{\otimes 2}\right)\Big] \\
&= \langle J|(Y_1 \otimes \mathbb{1})\Phi_{\mathcal{E}}\left(|0^n\rangle\langle 0^n|^{\otimes 2}\right)(Y_1 \otimes \mathbb{1})|J\rangle \\
&= \mathbb{E}_{U\sim\mathcal{E}}\left[\langle 0^n|^{\otimes 2}\left(\left(U_l^{\otimes 2}(Y_1 \otimes \mathbb{1})|J_{x_l}\rangle\langle J_{x_l}|(Y_1 \otimes \mathbb{1})U_l^{\dagger,\otimes 2}\right) \otimes |J_{\overline{x_l}}\rangle\langle J_{\overline{x_l}}|\right)|0^n\rangle^{\otimes 2}\right] \\
&= 0.
\end{aligned} \tag{85}$$

In the final step, we used that $|\langle 00|J_1\rangle| = \frac{1}{\sqrt{2}}|\langle 00|(|10\rangle - |01\rangle)| = 0$.

Proof of Claim 2: Uniformly random ensembles of symplectic and unitary states are indistinguishable[16]. We can thus replace the symplectic twirl with the unitary twirl:

$$\Phi_{\mathrm{USp}}\left(|0^n\rangle\langle 0^n|^{\otimes 2}\right) = \Phi_{\mathrm{U}}\left(|0^n\rangle\langle 0^n|^{\otimes 2}\right) = \frac{2}{2^n(2^n+1)}P_{\mathrm{sym}}, \tag{86}$$

where $P_{\mathrm{sym}} = \frac{1}{2}(\mathbb{1}+\mathbb{F})$ is the projector onto the symmetric subspace. Note that $Y_1 J$ is a symmetric operator, i.e., $Y_1 J = (Y_1 J)^T$. This means that

$$P_{\mathrm{sym}}(Y_1 \otimes \mathbb{1})|J\rangle = P_{\mathrm{sym}}(Y_1 J \otimes \mathbb{1})|\Omega_n\rangle = (Y_1 \otimes \mathbb{1})|J\rangle\,. \tag{87}$$

Hence

$$\begin{aligned}
\mathrm{tr}\Big[(Y_1 \otimes \mathbb{1})|J\rangle\langle J|(Y_1 \otimes \mathbb{1})\Phi_{\mathrm{USp}}\left(|0^n\rangle\langle 0^n|^{\otimes 2}\right)\Big] &= \langle J|(Y_1 \otimes \mathbb{1})\Phi_{\mathrm{USp}}\left(|0^n\rangle\langle 0^n|^{\otimes 2}\right)(Y_1 \otimes \mathbb{1})|J\rangle \\
&= \langle J|(Y_1 \otimes \mathbb{1})\left(\frac{2}{2^n(2^n+1)}P_{\mathrm{sym}}\right)(Y_1 \otimes \mathbb{1})|J\rangle \\
&= \frac{2}{2^n(2^n+1)} > 0,
\end{aligned} \tag{88}$$

which is what we set out to prove.

4 Symplectic state designs: additive error

In this section we prove that we can construct additive-error state designs in $\mathcal{O}(\log(n))$ depth, using local symplectic gates. We use a similar approach as in the orthogonal case in Section 4.

Theorem 24 (Equivalence of symplectic and unitary designs on unentangled states)**.** *Consider a state ρ on $\mathcal{H}^{\otimes k} \otimes \mathcal{A}$ that obeys the PPT condition on each copy of $\mathcal{H}$. Then the Haar symplectic twirl of ρ is equal to the Haar unitary twirl of ρ up to additive error*

$$\|\Phi_{\mathrm{USp}}(\rho) - \Phi_{\mathrm{U}}(\rho)\|_1 \leq \mathcal{O}(k^{7/2}/2^n) + \mathcal{O}(k/2^{n/2}). \tag{89}$$

Just as in the orthogonal case, the theorem directly implies that we can construct additive-error symplectic state designs in depth $\mathcal{O}(\log(kn/\epsilon))$.

Proof. We define the *symplectic distinct subspace* as the space spanned by strings $(x_1, \ldots, x_k)$ such that $|x_i\rangle\langle x_i| \neq J|x_j\rangle\langle x_j|J$ for all $i \neq j$. Let us define $\widehat{\Pi}_{\mathrm{dist}}$ as the projector onto this space. The proof of the theorem follows similar steps as the proof of Theorem 16, except that we project onto the symplectic distinct subspace.

1. The Haar symplectic twirl squares to itself, $\Phi_{\mathrm{USp}} = \Phi_{\mathrm{USp}} \circ \Phi_{\mathrm{USp}}$.

2. We prove that

$$\left\| [\widehat{\Pi}_{\text{dist}} \circ \Phi_{\text{USp}} \circ \widehat{\Pi}_{\text{dist}} \circ \Phi_{\text{USp}}](\rho) - \Phi_{\text{USp}} \circ \Phi_{\text{USp}}(\rho) \right\|_1 \leq 6k/2^{n/2}, \tag{90}$$

where $\widehat{\Pi}_{\text{dist}}(\cdot) \equiv \widehat{\Pi}_{\text{dist}}(\cdot)\,\widehat{\Pi}_{\text{dist}}$ denotes conjugation by the symplectic distinct subspace projector.

3. Within the symplectic distinct subspace, the Haar symplectic twirl and Haar unitary twirl are equal up to small relative error,

$$(1-\varepsilon')\,[\,\widehat{\Pi}_{\text{dist}} \circ \Phi_{\text{U}} \circ \widehat{\Pi}_{\text{dist}}\,] \preceq [\,\widehat{\Pi}_{\text{dist}} \circ \Phi_{\text{USp}} \circ \widehat{\Pi}_{\text{dist}}\,] \preceq (1+\varepsilon')\,[\,\widehat{\Pi}_{\text{dist}} \circ \Phi_{\text{U}} \circ \widehat{\Pi}_{\text{dist}}\,], \tag{91}$$

with $\varepsilon' = \mathcal{O}(k^{7/2}/2^n)$.

4. One can remove the projectors onto the symplectic distinct subspace after the Haar symplectic and Haar unitary twirls, up to small additive error,

$$\left\| [\widehat{\Pi}_{\text{dist}} \circ \Phi_{\text{U}} \circ \widehat{\Pi}_{\text{dist}} \circ \Phi_{\text{USp}}](\rho) - \Phi_{\text{U}} \circ \Phi_{\text{USp}}(\rho) \right\|_1 \leq 6k/2^{n/2}. \tag{92}$$

5. The composition of the Haar unitary and Haar symplectic twirl is equal to the Haar unitary twirl, $\Phi_{\text{U}} \circ \Phi_{\text{USp}} = \Phi_{\text{U}}$.

We start by proving the symplectic version of step 2. Similarly as in the orthogonal case, we can compute that

$$\left\| [\widehat{\Pi}_{\text{dist}} \circ \Phi_{\text{USp}} \circ \widehat{\Pi}_{\text{dist}} \circ \Phi_{\text{USp}}](\rho) - [\Phi_{\text{USp}} \circ \Phi_{\text{USp}}](\rho) \right\|_1 \leq 4\sqrt{1 - \operatorname{tr}\Big(\Phi_{\text{USp}}(\rho)\,\widehat{\Pi}_{\text{dist}}\Big)}. \tag{93}$$

We note that $\mathbb{1} - \widehat{\Pi}_{\text{dist}} \leq \sum_{i<j} \widehat{\Pi}_{ij}$, where $\widehat{\Pi}_{ij} = \sum_{x\in\{0,1\}^n} |x\rangle\langle x|_i \otimes (J\,|x\rangle\langle x|\,J)_j$. It thus follows that

$$4\sqrt{1 - \operatorname{tr}\Big(\Phi_{\text{USp}}(\rho)\,\widehat{\Pi}_{\text{dist}}\Big)} \leq 4\sqrt{\sum_{i<j} \operatorname{tr}\Big(\widehat{\Pi}_{ij}\Phi_{\text{USp}}(\rho)\Big)} = 4\sqrt{\sum_{i<j} \operatorname{tr}\Big(\Phi_{\text{USp}}(\widehat{\Pi}_{ij})\rho\Big)}. \tag{94}$$

We know that $\{\mathbb{1}_{ij}, S_{ij}, \Pi_{s,ij}\}$ is a basis of the commutant of the symplectic group over the subsystem i,j. We note that $\operatorname{tr}\Big(\Pi_{s,ij}\widehat{\Pi}_{ij}\Big) = 2^n$. Using Eq. (78) we compute that

$$\Phi_{\text{USp}}(\widehat{\Pi}_{ij}) = \frac{2^n - 1}{(2^n+1)(2^n-2)}(\mathbb{1}_{ij} + \mathbb{F}_{ij} + \Pi_{s,ij}). \tag{95}$$

We can use the PPT condition to bound the last term,

$$\begin{aligned}
\operatorname{tr}(\Pi_{s,ij}\rho) &= \operatorname{tr}((\mathbb{1}\otimes J)_{ij}\Pi_{o,ij}(\mathbb{1}\otimes J)_{ij}\rho) \\
&= \operatorname{tr}\Big((\Pi_{o,ij})^{\Gamma_i}(\mathbb{1}\otimes J)_{ij}\rho^{\Gamma_i}(\mathbb{1}\otimes J)_{ij}\Big) \\
&\leq \|(\Pi_{o,ij})^{\Gamma_i}\|_\infty \cdot \|(\mathbb{1}\otimes J)_{ij}\rho^{\Gamma_i}(\mathbb{1}\otimes J)_{ij}\|_1 \\
&= \|\mathbb{F}_{ij}\|_\infty \cdot 1 \\
&= 1.
\end{aligned} \tag{96}$$

This all adds up to

$$\operatorname{tr}\Big(\Phi_{\text{USp}}(\widehat{\Pi}_{ij})(\rho)\Big) \leq \frac{3(2^n-1)}{(2^n+1)(2^n-2)} \leq \frac{4}{2^n}. \tag{97}$$

Combining the things we learned gives

$$\left\| [\widehat{\Pi}_{\text{dist}} \circ \Phi_{\text{USp}} \circ \widehat{\Pi}_{\text{dist}} \circ \Phi_{\text{USp}}](\rho) - [\Phi_{\text{USp}} \circ \Phi_{\text{USp}}](\rho) \right\|_1 \leq 4\sqrt{\frac{k(k-1)}{2}\frac{4}{2^n}} \leq 6\frac{k}{2^{n/2}}. \tag{98}$$

Step 4 is done similarly.

To prove step 3 we see that

$$[\Phi_{\mathrm{USp}} \otimes \mathbb{1}](|\Omega_n\rangle\langle\Omega_n|) = \frac{1}{2^{nk}} \sum_{\sigma,\tau \in M_k} \mathrm{Wg}_{\mathrm{USp}}(\sigma,\tau)[\tau \otimes \sigma], \tag{99}$$

where σ, τ are summed over the generators of the commutant of the symplectic group.

Multiplying by the symplectic distinct subspace projectors yields

$$[\widehat{\Pi}_{\mathrm{dist}} \otimes \widehat{\Pi}_{\mathrm{dist}}]\,[\Phi_{\mathrm{USp}} \otimes \mathbb{1}](|\Omega_n\rangle\langle\Omega_n|)\,[\widehat{\Pi}_{\mathrm{dist}} \otimes \widehat{\Pi}_{\mathrm{dist}}] = \frac{1}{2^{nk}} \sum_{\pi,\tilde{\pi} \in S_k} \mathrm{Wg}_{\mathrm{USp}}(\pi,\tilde{\pi})\,[\widehat{\Pi}_{\mathrm{dist}}\,\pi\,\widehat{\Pi}_{\mathrm{dist}} \otimes \widehat{\Pi}_{\mathrm{dist}}\,\tilde{\pi}\,\widehat{\Pi}_{\mathrm{dist}}]. \tag{100}$$

Note that $\pi, \tilde{\pi}$ sum over S_k. This follows since the elements of the commutant that are not a permutation vanish when multiplied with the projector. Similarly as in the orthogonal case we now get that

$$\|[\widehat{\Pi}_{\mathrm{dist}} \otimes \widehat{\Pi}_{\mathrm{dist}}]\,[(\Phi_{\mathrm{USp}} - \Phi_a) \otimes \mathbb{1}](|\Omega_n\rangle\langle\Omega_n|)\,[\widehat{\Pi}_{\mathrm{dist}} \otimes \widehat{\Pi}_{\mathrm{dist}}]\|_\infty \leq \sum_{\pi \in S_k \setminus \mathbb{1}} \left|\mathrm{Wg}_{\mathrm{USp}}(\pi)\right| + \left|\mathrm{Wg}_{\mathrm{USp}}(\mathbb{1}) - \frac{1}{2^{nk}}\right|. \tag{101}$$

We use that $|\,\mathrm{Wg}_{\mathrm{USp}}(\mathbb{1}) - \frac{1}{2^{nk}}| \leq \frac{1}{2^{nk}} \mathcal{O}\left(\frac{k^{7/2}}{2^{2n}}\right)$, which follows from Theorem 4.10 in[12]. An argument similar as the proof of Lemma 19 tells us that $\sum_{\pi \in S_k \setminus \mathbb{1}} \left|\mathrm{Wg}_{\mathrm{USp}}(\pi)\right| \leq \frac{1}{2^{nk}} \mathcal{O}\left(\frac{k}{2^n}\right)$. We get that

$$\|[\widehat{\Pi}_{\mathrm{dist}} \otimes \widehat{\Pi}_{\mathrm{dist}}]\,[(\Phi_{\mathrm{USp}} - \Phi_a) \otimes \mathbb{1}](|\Omega_n\rangle\langle\Omega_n|)\,[\widehat{\Pi}_{\mathrm{dist}} \otimes \widehat{\Pi}_{\mathrm{dist}}]\|_\infty \leq \mathcal{O}(k^{7/2}/2^n). \tag{102}$$

An argument similar to the proof of Lemma 18 concludes the proof of step 3.

5 Symplectic anti-concentration

We know that the symplectic twirl equals the unitary twirl on pure states. This directly implies that the gluing lemma[3] also works for the symplectic twirl, provided that the input and the output are both pure states, that factorize in the qubit-direction. We thus have anti-concentration in $\mathcal{O}(\log(n))$ depth for locally symplectic circuits.

Theorem 25. *Let $\mathcal{E}$ be the ensemble of symplectic matrices drawn from the two-layer superblock architecture. This gives us two layers of $m/2 = n/(2\xi)$ superblocks each, where $\xi = \mathcal{O}(\log(k^2 n/\epsilon))$. This ensemble anti-concentrates, i.e.,*

$$\mathbb{E}_{U \sim \mathcal{E}}\,|\langle 0|\,U\,|0_n\rangle|^{2k} \leq (1+\epsilon)\;\mathbb{E}_{U \sim \mathrm{USp}}\,|\langle 0|\,U\,|0_n\rangle|^{2k}\,. \tag{103}$$

Proof. Let $\mathcal{E}$ be the ensemble of superblocks with the superblocks drawn from the unitary symplectic group. We define $\mathcal{E}_1, \mathcal{E}_2$ as the ensemble of superblocks in the first and second layer respectively. We have that

$$\mathbb{E}_{U \sim \mathcal{E}}|\,\langle 0^n|\,U\,|0^n\rangle\,|^{2k} = \mathrm{tr}\left[\Phi_{\mathcal{E}_1}(|0^n\rangle\langle 0^n|^{\otimes k})\Phi_{\mathcal{E}_2}(|0^n\rangle\langle 0^n|^{\otimes k})\right]. \tag{104}$$

As the superblocks are locally drawn independently from Haar random symplectic unitaries, and symplectic unitaries are indistinguishable from unitaries on pure states[16], we can replace the Haar random symplectic unitaries in Eq. (104) with Haar random unitaries. By the gluing lemma, we now have that

$$\begin{aligned}\mathbb{E}_{U \sim \mathcal{E}}|\,\langle 0^n|\,U\,|0^n\rangle\,|^{2k} &\leq (1+\epsilon)\mathbb{E}_{U \sim \mathrm{U}}|\,\langle 0^n|\,U\,|0^n\rangle\,|^{2k} \\ &= (1+\epsilon)\mathbb{E}_{U \sim \mathrm{USp}}|\,\langle 0^n|\,U\,|0^n\rangle\,|^{2k},\end{aligned} \tag{105}$$

if the superblocks act on $\mathcal{O}\left(\log(nk^2/\epsilon)\right)$ qubits.

Supplementary Note 5 — Matchgate group

In this section, we prove our results on the n-qubit matchgate group $\mathcal{M}_n$, which models the evolution of free fermions (without particle number preservation). Matchgate designs have been applied to shadow estimation of fermionic observables[21] , scalable benchmarking of quantum computers[22] , and the development of quantum advantage schemes[23]. Here we will prove that the superblock construction for matchgate designs requires depth $\Omega(n^{1/3})$ to anti-concentrate (though this is probably not sufficient). We will furthermore provide stronger lightcone-based lower bounds for relative and additive-error matchgate state designs.

1 Preliminaries

The matchgate group $\mathrm{M}(n)$ is generated by 2-qubit unitaries of the form $e^{\theta P_i Q_{i+1}}$ where $P, Q \in \{X, Y\}$, supplemented with single qubit X gates, where the qubits are laid out on a line. In this specific geometry these generators conspire to generate a small subgroup of the unitary group with a rich structure. To discuss this structure further we need to introduce the so called Majorana operators.

The Majorana operators (on n fermionic modes) are a set of operators γ_i for $i \in [2n]$ with the properties that $\gamma_i^2 = I$ and $\gamma_i\gamma_j = -\gamma_j\gamma_i$ for $i \neq j$. They form an algebra that can be represented faithfully on the space $(\mathbb{C}^2)^{\otimes n}$ of qubit states using Hermitian Pauli operators. There are several ways of doing this, but a convenient way is the so called Jordan-Wigner (JW) isomorphism, which maps

$$\begin{aligned} \gamma_{2i-1} &\to Z_1 \dots Z_{i-1} X_i, \\ \gamma_{2i} &\to Z_1 \dots Z_{i-1} Y_i. \end{aligned} \tag{106}$$

It turns out that any matchgate circuit effects an orthogonal transformation at the level of Majoranas; we can write a matchgate unitary as $U(O)$ with $O \in \mathrm{O}(2n)$ an orthogonal operator, which acts on Majorana operators as

$$U(O)\gamma_i U(O)^\dagger = \sum_{j\in[n]} O_{ij}\gamma_j. \tag{107}$$

Through the JW isomorphism, we can thus think of the matchgate group as irreducible projective representation of the Lie group $\mathrm{O}(2n)$ on $(\mathbb{C}^2)^{\otimes n}$ (sometimes called the Pin representation).[1] The $k = 2$ diagonal representation $(U(O) \cdot U(O)^\dagger)^{\otimes 2}$ was evaluated in[22], where it was shown that the average over the twofold tensor product projects onto a space spanned by the operators

$$V_k = \binom{2n}{k}^{-1/2} \sum_{S\subset[2n],\ |S|=k} 2^{-n} {\gamma_S}^{\otimes 2}, \tag{108}$$

with $k \in [0 : 2n]$ and where the Majorana operator products

$$\gamma_S = \prod_{i\in S} \gamma_i, \tag{109}$$

are normally ordered (that is, in increasing order in the set $S \subset [2n]$). Note that the operators V_k are orthonormal under the trace inner product.

For the purposes of this paper it is also important to evaluate averages over random matchgates that only act on a subset of the qubits. It is important to note that such embedded matchgates only exist if the subset is a sub-interval $[a : b]$ of the interval $[n]$ (with $a < b \in [n]$). These matchgates form a Pin representation of the orthogonal group $O(2(a-b))$ acting on the a'th to b'th qubit. The

[1] Often the name matchgate group is reserved for the subgroup $\mathrm{SO}(2n)$. The results we discuss here hold for this group as well with minor alterations to the proofs.

second moment average can again be characterized; it has support on the operators $A \otimes V_k^{(a,b)} \otimes B$ where A (resp. B) is an arbitrary operator on $(\mathbb{C}^2)^{\otimes a-1}$ (resp. $(\mathbb{C}^2)^{\otimes n-b}$), and $V_k^{(a,b)}$ is of the form

$$V_k^{(a,b)} = I^{\otimes a-1} \otimes \binom{2(a-b)}{k}^{-1/2} \sum_{S \subset [2(a-b)],\ |S|=k} 2^{-(a-b)} \gamma_S^{(a,b)\otimes 2} \otimes I^{\otimes n-b}, \tag{110}$$

where $\gamma_i^{(a,b)}$ is the operator obtained by taking the JW representation of γ_i and replacing the first $a-1$ Z operators with identity I operators. Note that if k is even this replacement does not change the Majorana product operators, i.e., if $S \subset [2(a-b)]$ and $|S|$ even, then

$$\prod_{i \in S} \gamma_i^{(a,b)} = \prod_{i \in S} \gamma_i. \tag{111}$$

2 Matchgate anti-concentration

We prove the following theorem, which constrains how much superblock matchgate circuits anti-concentrate in the generalized sense (i.e. they approximate the value for uniformly random matchgates up to a constant factor).

Theorem 26. *Consider the ensemble of matchgate random circuits in a 1D superblock construction with block-size 2ξ, with uniformly random matchgates inside the blocks (and thus total depth 4ξ). If this ensemble anti-concentrates, that is*

$$2^{2n} \mathbb{E}_{U(O)\sim\mathcal{E}} |\langle 0| U(O) |0\rangle|^4 \leq (1+\epsilon) 2^{2n} \int dO |\langle 0| U(O) |0\rangle|^4, \tag{112}$$

for some constant ϵ, then $\xi = \Omega(n/\log(n))$.

Proof. We begin by computing the anti-concentration value of random n-qubit matchgates, i.e.,

$$\chi = 2^{2n} \int_{SO(2n)} dO |\langle 0| U(O) |0\rangle|^4. \tag{113}$$

This is a second moment quantity, easily computed using the orthonormal basis $\{V_k\}$ described above. Concretely, from the JW isomorphism it is easy to see that $|0\rangle\langle 0|^{\otimes 2}$ is trace orthogonal to V_k with k odd and that for even k we have

$$\mathrm{tr}\left(|0\rangle\langle 0|^{\otimes 2} V_k\right)^2 = 2^{-2n} \binom{n}{k/2}^2 \binom{2n}{k}^{-1}. \tag{114}$$

This follows by noting that through JW, only Majorana strings composed of $\gamma_{2i-1}\gamma_{2i}$ products (i.e. Z_i operators) have non-zero expectation value on $|0\rangle\langle 0|$. It's not hard to see (using standard bounds for the binomial) that

$$C_1 \sqrt{n/((2n-k)k)} \leq \binom{n}{k/2}^2 \binom{2n}{k}^{-1} \leq C_2 \sqrt{n/((2n-k)k)}, \tag{115}$$

for $k \in [2 : 2n-2]$, for some constants C_1, C_2. From this we immediately obtain that (for some constant C):

$$\chi = 2^{2n} \sum_{k \in [0:2:2n]} \mathrm{tr}\left(|0\rangle\langle 0|^{\otimes 2} V_k\right)^2 \leq 2 + \sum_{k \in [2:2:2n-2]} C_2 \sqrt{n/((2n-k)k)} \leq C\sqrt{n}. \tag{116}$$

We will now compute the same quantity for a 1D superblock circuit of random matchgates. We divide n qubits into m equal patches of size $n/m = \xi$, and have random matchgates $U(O)_{(i)}$

act on the patches $(i, i+1)$ (assuming periodic boundary conditions for convenience). Defining $W_k^{(a,b)} := V_k^{(a,b)} \binom{2(a-b)}{k}^{1/2}$ and using the embedded Majorana notation defined above we can write

$$\int_{O_i} \mathbb{1} \otimes U(O_i)^{\otimes 2} |0\rangle\langle 0|^{\otimes 2} U(O_i)^{\otimes 2^\dagger} \otimes \mathbb{1} = \sum_{k=0,\, k\text{even}}^{4\xi} 2^{-2\xi} \binom{2\xi}{k/2} \binom{4\xi}{k}^{-1} (\mathbb{1} \otimes W_k^{(\xi i+1, \xi(i+2))} \otimes \mathbb{1}). \quad (117)$$

With this we can partially evaluate the anti-concentration value as

$$\begin{aligned} \chi &= 2^{2n} \int_{O_1,\ldots O_n \in O(4\xi)} |\langle 0| (\bigotimes_{i=1}^{m/2} U(O_{2i-1}) \bigotimes_{i=1}^{m/2} U(O_{2i}) |0\rangle|^4 \\ &= \sum_{\substack{k_1,\ldots k_{m/2} \in [0:2:4\xi] \\ l_1,\ldots l_{m/2} \in [0:2:4\xi]}} \prod_{i=1}^{m/2} \binom{2\xi}{k_i/2} \binom{2\xi}{l_i/2} \binom{4\xi}{k_i}^{-1} \binom{4\xi}{l_i}^{-1} \\ &\times \mathrm{Tr}\left[(W_{k_1}^{(1,2\xi)} \otimes \ldots \otimes W_{k_{m/2}}^{(n-2\xi+1,n)}) (W_{l_1}^{(\xi+1,3\xi)} \otimes \ldots \otimes W_{l_{m/2}}^{(n-\xi+1,\xi)}) \right] \end{aligned} \quad (118)$$

We note that W factorizes in the qubit-direction when taken inside the trace, i.e.,

$$\begin{aligned} &\mathrm{Tr}_{(\xi+1,3\xi)} \left[(W_{k_1}^{(1,2\xi)} \otimes W_{k_2}^{(2\xi+1,4\xi)}) W_{l_1}^{(\xi+1,3\xi)} \right] \\ &= \mathrm{Tr}_{(\xi+1,3\xi)} \left[(W_{k_1}^{(1,2\xi)} \otimes W_{k_2}^{(2\xi+1,4\xi)}) \left(\sum_{v_1=0}^{l_1} (W_{v_1}^{(\xi+1,2\xi)} \otimes W_{l_1-v_1}^{(2\xi+1,3\xi)}) \right) \right]. \end{aligned} \quad (119)$$

Note that the equation does not hold without the trace when v_1 is odd.

Equation 118 is a rather complicated equation, but because of its 1D nature, its periodic boundary conditions and Equation 119 we can write it as the trace of an m-fold power of a $2\xi \times 2\xi$ matrix:

$$\chi_\xi = \mathrm{Tr}((AB)^m), \quad (120)$$

where

$$A_{vw} = \binom{2\xi}{(v+w)/2} \binom{4\xi}{v+w}^{-1} \delta_{v+w}^{\text{even}} \quad (121)$$

$$B_{vw} = 2^{-2\xi} \mathrm{Tr}\left[W_v^{(1,\xi)} W_w^{(1,\xi)} \right] = \delta_{v,w} \binom{2\xi}{v}, \quad (122)$$

where $\delta_{v+w}^{\text{even}}$ is 1 for $v+w$ even, and 0 otherwise. We define $M = (AB)^2$ to get

$$M_{vw} = \sum_k \binom{2\xi}{(v+k)/2} \binom{2\xi}{(w+k)/2} \binom{4\xi}{v+k}^{-1} \binom{4\xi}{w+k}^{-1} \binom{2\xi}{v} \binom{2\xi}{k} \delta_{v+k}^{\text{even}} \delta_{w+k}^{\text{even}}. \quad (123)$$

It remains to provide a lower bound for the largest eigenvalue of this matrix. As it is a point-wise positive matrix, any element on the diagonal provides a lower bound for the largest eigenvalue. In particular we can calculate the entry M_{00}

$$\begin{aligned} M_{00} &= \sum_{k \text{ even}} \binom{2\xi}{k/2}^2 \binom{4\xi}{k}^{-2} \binom{2\xi}{k} \\ &\geq 1 + (2\xi)^2 \binom{4\xi}{2}^{-2} \binom{2\xi}{2} = 1 + \Omega(1), \end{aligned} \quad (124)$$

where for the inequality we evaluated the sum for $k = 0, 2$. It follows that for some constant $c > 0$

$$\chi_\xi \geq (1+c)^{m/2}. \quad (125)$$

In order to obtain anti-concentration, we require

$$(1+c)^{n/2\xi} \leq (1+\epsilon)C\sqrt{n}, \tag{126}$$

which directly leads to $\xi = \Omega(n/\log(n))$, which in turn implies the same scaling for the circuit depth.

3 Matchgate state designs

We now prove lower bounds for matchgate state designs. We note that due to the explicit one dimensional nature of the matchgate group these results are simpler to state and prove than the analogous statements for the other groups we have discussed. This is so because limiting the depth of a 1D circuit implies the existence of a lightcone that does not depend on the internal structure of the circuit. Hence we have no need for extra assumptions to make our proofs work.

Theorem 27. *Let $\mathcal{E}$ be an ensemble of matchgate circuits of depth d. The following now holds:*

1. *If $\Phi_{\mathcal{E}}(|0\rangle\langle 0|^{\otimes 2})$ is an additive-error (with $\epsilon < 1/(2n-1)$) state design, then $d = \Omega(n)$*
2. *If $\Phi_{\mathcal{E}}(|0\rangle\langle 0|^{\otimes 2})$ is a relative-error (with $\epsilon < 1/(4n-2)$) state design, then $d = \Omega(n)$.*

Proof. Consider the Majorana operators γ_1, γ_{2n}. For uniformly random matchgates we have (using the definition of V_k):

$$\int_{O(2n)} dO |\langle 0| U(O)\gamma_1\gamma_{2n}U(O)^\dagger |0\rangle|^2 = \frac{1}{2n-1}. \tag{127}$$

On the other hand, if $U(O)$ is a matchgate circuit of depth $d \leq n/4$, then clearly

$$|\langle 0| U(O)\gamma_1\gamma_{2n}U(O)^\dagger |0\rangle|^2 = |\langle 0| \sum_{i=1}^{n/2} \alpha_i\gamma_i \sum_{j=3n/2}^{2n} \beta_j\gamma_j |0\rangle|^2 = 0, \tag{128}$$

since $\langle 0|\gamma_i\gamma_j|0\rangle = 0$ unless $i = j$ or i is odd and $j = i+1$. Now since the operator $(\gamma_1\gamma_{2n})^{\otimes 2}$ is hermitian and of operator norm 1, one clearly requires $d \geq n/4$ to obtain any additive error state design of error smaller than $1/(2n-1)$. Since the existence of relative-error state designs with error ϵ imply the existence of additive-error state designs with error 2ϵ the same conclusion holds for relative-error designs.

References

1. Mele, A. A. Introduction to Haar measure tools in quantum information: A beginner's tutorial. *Quantum* **8.** Publisher: Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, 1340 (2024).
2. Brandao, F. G., Harrow, A. W. & Horodecki, M. Local random quantum circuits are approximate polynomial-designs. *Communications in Mathematical Physics* **346.** Publisher: Springer, 397–434 (2016).
3. Schuster, T., Haferkamp, J. & Huang, H.-Y. Random unitaries in extremely low depth. *Science* **389,** 92–96 (2025).
4. Zhu, H., Kueng, R., Grassl, M. & Gross, D. The Clifford group fails gracefully to be a unitary 4-design. *arXiv:1609.08172 [quant-ph].* arXiv: 1609.08172 (2016).
5. Zhu, H. Multiqubit Clifford groups are unitary 3-designs. *Physical Review A* **96,** 062336 (2017).
6. Gross, D., Nezami, S. & Walter, M. Schur–Weyl Duality for the Clifford Group with Applications: Property Testing, a Robust Hudson Theorem, and de Finetti Representations. *Communications in Mathematical Physics* **385,** 1325–1393. ISSN: 1432-0916 (2021).

7. Helsen, J. & Walter, M. Thrifty shadow estimation: reusing quantum circuits and bounding tails. *Physical Review Letters* **131,** 240602 (2023).
8. Turkeshi, X. Private Communication. 2025.
9. O'Donnell, R., Servedio, R. A. & Paredes, P. *Explicit orthogonal and unitary designs* in *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, 2023), 1240–1260.
10. Aharonov, D., Cotler, J. & Qi, X.-L. Quantum algorithmic measurement. *Nature communications* **13,** 1–9 (2022).
11. Hashagen, A., Flammia, S., Gross, D. & Wallman, J. Real randomized benchmarking. *Quantum* **2.** Publisher: Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, 85 (2018).
12. Collins, B. & Matsumoto, S. Weingarten calculus via orthogonality relations: new applications. eprint: 1701.04493 (2017).
13. Sauliere, A., Magni, B., Lami, G., Turkeshi, X. & Nardis, J. D. *Universality in the Anticoncentration of Chaotic Quantum Circuits* 2025. arXiv: 2503.00119.
14. Magni, B., Christopoulos, A., De Luca, A. & Turkeshi, X. Anticoncentration in clifford circuits and beyond: From random tensor networks to pseudomagic states. *Physical Review X* **15,** 031071 (2025).
15. Zimborás, Z., Zeier, R., Schulte-Herbrüggen, T. & Burgarth, D. Symmetry criteria for quantum simulability of effective interactions. *Phys. Rev. A* **92,** 042309 (4 2015).
16. West, M., Mele, A. A., Larocca, M. & Cerezo, M. Random ensembles of symplectic and unitary states are indistinguishable. *arXiv preprint arXiv:2409.16500* (2024).
17. García-Martín, D., Braccia, P. & Cerezo, M. *Architectures and random properties of symplectic quantum circuits* 2024.
18. Bremner, M. J., Dodd, J. L., Nielsen, M. A. & Bacon, D. Fungible dynamics: There are only two types of entangling multiple-qubit interactions. *Physical Review A* **69,** 012313 (2004).
19. Bolt, B., Room, T. G. & Wall, G. E. On the Clifford collineation, transform and similarity groups. II. *Journal of the Australian Mathematical Society* **2.** Publisher: Cambridge University Press, 80–96 (1961).
20. Guralnick, R. & Tiep, P. Decompositions of small tensor powers and Larsen's conjecture. *Representation Theory of the American Mathematical Society* **9,** 138–208 (2005).
21. Wan, K., Huggins, W. J., Lee, J. & Babbush, R. Matchgate shadows for fermionic quantum simulation. *Communications in Mathematical Physics* **404,** 629–700 (2023).
22. Helsen, J., Nezami, S., Reagor, M. & Walter, M. Matchgate benchmarking: Scalable benchmarking of a continuous family of many-qubit gates. *arXiv preprint ArXiv:2011.13048* **98** (2023).
23. Oszmaniec, M., Dangniam, N., Morales, M. E. & Zimborás, Z. Fermion sampling: a robust quantum computational advantage scheme using fermionic linear optics and magic input states. *PRX Quantum* **3,** 020328 (2022).