Short remarks on shallow unitary circuits

Jeongwan Haah

Leinweber Institute for Theoretical Physics, Stanford University,
Google Quantum AI

(i) We point out that every local unitary circuit of depth smaller than the linear system size is easily distinguished from a global Haar random unitary if there is a conserved quantity that is a sum of local operators. This is always the case with a continuous onsite symmetry or with a local energy conservation law.

(ii) We explain a simple algorithm for a formulation of the shallow unitary circuit learning problem and relate it to an open question on strictly locality-preserving unitaries (quantum cellular automata).

(iii) We show that any translation-invariant quantum cellular automaton in D-dimensional lattice of volume V can be implemented using only $O(V)$ local gates in a staircase fashion using invertible subalgebra pumping.

1 Symmetric shallow circuit $\neq$ symmetric global random unitary

Recently it has been shown [LL24b, SHH24] that an ensemble of random local quantum circuits U on a one-dimensional chain of n qubits of depth $k\,\mathrm{polylog}(nk)$ is indistinguishable from the Haar random global unitary on n qubits if we query U no more than k times.[1] Such an ensemble of unitaries is called a *unitary k-design*. The existence of a shallow unitary design is rather surprising since the depth is much smaller than the linear system size n, so the absolute lightcone of an observable at one end of the chain does not even reach the other end. Nonetheless, the ensemble contains sufficient randomness which prohibits any predetermined correlation function from probing the difference between Haar random global unitaries and shallow local circuits.

Here we point out that a certain symmetry condition rules out all shallow unitary k-designs for any $k \geq 1$ because there is a useful predetermined autocorrelation function. This extends the list of conditions under which shallow random unitary designs are impossible. Previously it was known [HFCL25] that a brickwork random circuit of depth $o(L^2)$ in Euclidean lattice of linear size L that conserves $U(1)$ charge remains distinguishable from the global $U(1)$-symmetric Haar random unitaries; it was also known [SHH24] that any ensemble of shallow circuits is distinguishable from Haar random unitaries if there is time-reversal symmetry or if we can access the inverse of the unitary. In essence, our argument is that if we have a local charge density whose integral (sum) is conserved, then a shallow circuit and a symmetric global Haar random unitary are very different since local unitaries may spread charges only to a small distance while a symmetric Haar random

[1] The indistinguishability can also be established under a cryptographic assumption with a milder restriction on the number of queries to the unitary [SHH24].

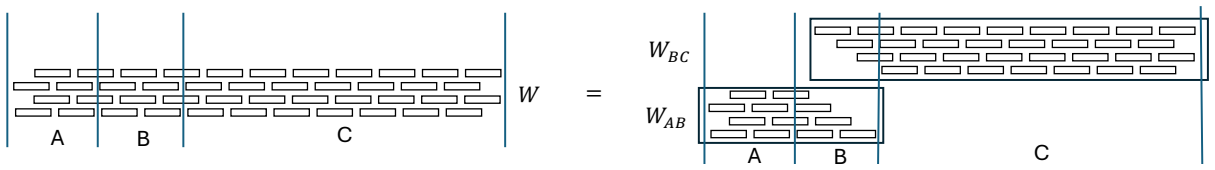

Figure 1: Any circuit W of depth smaller than the linear system size is a product of W_{AB} and W_{BC}.

unitary spreads charges across the whole system.[2] This summary is good enough as a take-away, but we need an extra ingredient for a sound protocol.

We work with a one-dimensional array of n qudits $\mathbb{C}^p$. So, the total Hilbert space is $(\mathbb{C}^p)^{\otimes n}$. Consider a symmetry (generator)

$$Q = \sum_{i=1}^{n} Q_i \tag{1}$$

where Q_i is a nonzero hermitian operator supported on a qudit i. We assume that every Q_i has an eigenvalue zero.[3] This condition is trivially fulfilled by adding a real multiple of identity to Q_i. Let $\mathsf{S}_{i,j} = \mathsf{S}_{j,i}$ be the swap operator that interchanges qudit i and j. For simplicity we assume that $Q_i = \mathsf{S}_{i,j} Q_j \mathsf{S}_{i,j}^\dagger$. The symmetric unitary group of interest consists of all $U \in \mathsf{U}(p^n)$ that commutes with Q. A symmetric local unitary circuit is a composition of layers of nonoverlapping nearest-neighbor 2-qudit unitaries, *each of which commutes with* Q. Of course, any symmetric local unitary circuit belongs to the symmetric unitary group.[4] The depth is the number of layers. This is a standard setting in the context of symmetry-protected topological phases [CGW10].

The distinguishing protocol is the following. We partition the system into three disjoint intervals A, B, C such that $\frac{1}{5}n \le |A| = |B| \le \frac{1}{4}n \le |C|$. Define $Q_S = \sum_{i \in S} Q_i$ for any subset S of qudits. We initialize a state ρ_0 so that $\mathrm{Tr}(Q_A \rho_0) = q > 0$ but $\mathrm{Tr}(Q_B \rho_0) = 0 = \mathrm{Tr}(Q_C \rho_0)$. Then, we apply a symmetric Haar random unitary U_{AB} supported on AB. We next apply a given instance W from an unknown ensemble of symmetric unitaries. We finally measure Q_A.

Suppose that W is drawn from an ensemble of symmetric local circuits of depth $\le \frac{1}{100}n$. Since A and B are sufficiently big ($\ge \frac{1}{5}n$), the instance W can be written as a product of two supergates: $W = W_{BC} W_{AB}$. See Figure 1. Note that W_{AB} commutes with Q_{AB} and W_{BC} commutes with Q_{BC}. Due to the randomization U_{AB} just before W, the component W_{AB} does nothing statistically; the probability distributions of U_{AB} and of $W_{AB} U_{AB}$ are identical. Therefore, immediately after U_{AB}, the partial sum Q_A has an expectation value $\overline{Q_A} = \frac{1}{2}q$ (over the random choices of U_{AB}) because the distribution of U_{AB} is invariant under the swap of A and B. The final measurement of Q_A is unaffected by W_{BC}.

On the other hand, if $W = U_{ABC}$ is drawn from the symmetric Haar distribution over the entire system, then the randomization U_{AB} of the protocol does nothing statistically; the probability distribution of U_{ABC} and of $U_{ABC} U_{AB}$ are identical. The final measurement gives an expectation

[2] See [KVH18, RPvK18] for related results.

[3] One may further require that each Q_i have an integer eigenvalue spectrum so that $e^{\mathrm{i}\theta} \mapsto e^{\mathrm{i}\theta Q}$ is a representation of $\mathsf{U}(1)$, but this condition is not used in our argument.

[4] The converse is false; there exists a symmetric unitary that is not expressible by any symmetric local unitary circuit [Mar22]. This is however unimportant since we focus on a lower bound on the depth, which is still meaningful since symmetric local unitary circuits with sufficient depths can form a symmetric unitary k-design for k that grows with the system size [HFCL25].

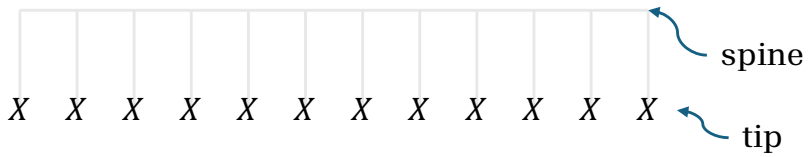

Figure 2: Comb connectivity. A qudit is placed on every tip and spine vertex. A two-qudit gate may only act on an edge. A discrete symmetry acts on the tip qudits.

value $\overline{Q_A} = \frac{|A|}{n}q \leq \frac{1}{4}q$. The measured Q_A is a random variable valued in $[-q_0, q_0]$ where $q_0 = |A|\|Q_i\|$. If we set q to be of order $|A|\|Q_i\|$, the expectation values of Q_A in the two scenarios are comparable to or larger than the statistical fluctuations, where the latter will become small, after a few repetitions, compared to the difference $\frac{1}{2}q - \frac{1}{4}q = \frac{1}{4}q$. We can therefore distinguish whether W is from the symmetric Haar random ensemble or from a shallow symmetric local circuit with confidence $1-\delta$ after $O(\log\frac{1}{\delta})$ repetitions of the protocol. Each run queries the given W once.

Since our protocol is explicit, it is at least as strong a notion of distinguishability as the additive error for unitary designs, which is more of an information-theoretic measure. Under the latter notion where two mixed unitary channels are sufficiently far apart in the diamond distance, input states and output measurements for the task of distinguishing channels may be implicit and difficult to find.

This completes the analysis of the protocol. In conclusion, for any ensemble of symmetric local circuits to form an approximate symmetric unitary k-design with $k \geq 1$ and a small constant additive error, the depth of the circuit must be at least linear in the linear system size. Some comments are below.

The randomization by U_{AB} is important to make the protocol sound. If we omitted it, the measurement of Q_A may give a misleading answer for the distinguishing task because symmetric circuits can conspire to transport an abelian charge from anywhere to anywhere instantaneously by creating and annihilating charge-anticharge pairs.[5]

The existence of a conserved total charge Q is implied by a continuous symmetry, but that is not the only source. The conserved quantity Q can be a local Hamiltonian itself, in which case our protocol is simply to measure the energy density.[6] Even with a discrete symmetry, a conserved quantity that is a sum of local operators may appear in the following "dilute symmetry" scenario. Suppose that we have a *comb*[7] connectivity graph as shown in Figure 2. A qudit is placed on every vertex, and 2-qudit local unitaries are allowed only over the edges. Label each qudit by (i, tip) or (i, spine). We impose a symmetry by a tensor product operator $Q' = \prod_i X_{(i,\text{tip})}$. Here the symbol X is reminiscent of Pauli X as in the onsite $\mathbb{Z}_2$ symmetry on qubits. Then, every local unitary gate sees at most one tensor factor of Q' and hence $Q = \sum_i X_{(i,\text{tip})} + X^\dagger_{(i,\text{tip})}$, in addition

[5]Note that U_{AB} is the only component for which the implementation gate complexity as a function of n might be high; other components beyond querying a blackbox W are preparation of eigenstates of Q_i and measurement of Q_i, both of which are local operations. Since the protocol is to measure $\overline{Q_A} = \mathbb{E}_{U_{AB},W} \operatorname{Tr}(Q_A W U_{AB} \rho_0 U_{AB}^\dagger W^\dagger)$ with a small additive error, we only need U_{AB} to be drawn from an approximate symmetric unitary 1-design. It appears that the lower bound on the convergence rate of $U(1)$-symmetric circuits to $U(1)$-symmetric approximate designs is largely an open problem.

[6]Some other anticoncentration measures [TTS24] were numerically observed to grow qualitatively differently depending on whether the dynamics conserves energy.

[7]I thank Ehud Altman for suggesting a comb to explain this dilute symmetry.

to Q', is preserved by any symmetric local circuit dynamics. A symmetric global unitary is assumed to preserve Q' but not necessarily Q. This relaxed symmetry condition only widens the gap in the expectation values of Q_A and the protocol above still disambiguates shallow symmetric local circuits from global symmetric Haar random unitaries.

We have assumed that the sizes of subsystems $|A|$ and $|B|$ are equal, that the total system is a one-dimensional array of qudits, and that all the local operators Q_i are the same up to relabeling of the qudits. These assumptions are for simplicity in presentation and are unimportant. The subsystem A is where we put some known value of charge. If A and B are comparable in size, then a symmetric Haar random unitary U_{AB} on AB will evenly distribute the charge over AB and A will retain a significant fraction of the initial charge. In the symmetric shallow circuit case, the measurement of Q_A will reveal this large fraction of the initial charge, but in the global symmetric Haar random case, the measured value of Q_A will be suppressed by a factor of the overall system size. It remains to investigate the case of discrete symmetries represented by product operators that are not necessarily dilute.

We would like to conclude this section with a rather philosophical remark. Physics — an endeavor to describe complex phenomena by simple principles — is possible because the physical principles are testable despite of all the inaccuracies in our measurements and descriptions. If those inaccuracies resemble, hypothetically, the shallow random circuit that appears in [LL24b, SHH24], then everything would look completely dull that we would not be able to speak of any phase of matter. But, fundamentally, energies and symmetry charges do exist and are conserved, based on which we have separated any shallow circuits from a global featureless unitary.

2 Shallow circuit learning

It was asked in [HLB+24] how to output a description of a shallow unitary circuit U in a D-dimensional lattice Λ of qudits given oracle access to U, while minimizing the number of queries to U as a function of the system size and other parameters. The answer in [HLB+24] does not give a description of U but instead a circuit on a doubled system that implements $U \otimes U^\dagger$. A well-known identity [ANW11] was used in [HLB+24]:

$$U \otimes U^\dagger = (U \otimes \mathbf{1})\mathsf{S}(U^\dagger \otimes \mathbf{1})\mathsf{S} \tag{2}$$

where S is an operator that swaps the entire system Λ and its copy Λ'. Since the underlying operator algebra is a tensor product of those of individual qudits, we have

$$\mathsf{S} = \prod_i \mathsf{S}_{i,i'} \tag{3}$$

where $\mathsf{S}_{i,i'}$ is the swap operator between $i \in \Lambda$ and its copy $i' \in \Lambda'$. It follows [ANW11] that,

$$U \otimes U^\dagger = \left[\prod_i (U \otimes \mathbf{1})\mathsf{S}_{i,i'}(U^\dagger \otimes \mathbf{1})\right] \prod_i \mathsf{S}_{i,i'} \,. \tag{4}$$

Since U is shallow, the conjugation by U maps a local operator to a potentially slightly bigger operator nearby. We will refer to this blow up in support as the **spread** of U. Therefore, $U \otimes U^\dagger$ is a shallow depth quantum circuit consisting of $\mathsf{S}_{i,i'}$ and their conjugates. It is now clear that the circuit learning problem can be solved, if we are fine with using an auxiliary system, by figuring

out the conjugates of $\mathsf{S}_{i,i'}$. Note that since $\mathsf{S}_{i,i'}$ for different i's are commuting, the ordering of the factors in the product of (4) is immaterial.

For the last step, we recall a simple strategy that appeared in [HKT24] where it was necessary to estimate $\mathrm{Tr}(UXU^\dagger Y)$ for various local (Pauli) operators X and Y given query access to U. Take any density matrix $\rho \succeq 0$ that, as an operator, is supported on some small set of qudits. This mixed state ρ can be prepared by some local quantum channel on its support and some other channel that prepares all the qudits elsewhere in a random product state. The conjugate $U\rho U^\dagger$ is a density operator supported on a slightly bigger set of qudits, and usual state tomography can determine $U\rho U^\dagger$ whose complexity does not depend on the system size but only on the depth of U.[8] We can use this to learn $(U \otimes \mathbf{1})\mathsf{S}_{i,i'}(U^\dagger \otimes \mathbf{1})$ even though the unitary $\mathsf{S}_{i,i'}$ is *not* a density operator. Write $\mathsf{S}_{i,i'} = \frac{1}{p}\sum_P P_i \otimes P_{i'}^\dagger$ as a linear combination of (generalized) Pauli operators, where p is the local qudit dimension. Clearly, it suffices to learn $UPU^\dagger$ for every (generalized) Pauli operator P, but since $P \mapsto UPU^\dagger$ is an algebra homomorphism, it suffices to learn just two operators $UXU^\dagger$ and $UZU^\dagger$ where $X = \sum_{j\in\mathbb{Z}/p} |j+1\rangle\langle j|$ and $Z = \sum_{j\in\mathbb{Z}/p} e^{2\pi \mathbf{i} j/p} |j\rangle\langle j|$. Any operator $O = \frac{O+O^\dagger}{2} + \mathbf{i}\frac{O-O^\dagger}{2\mathbf{i}}$ is a linear combination of two hermitian operators, and any hermitian operator H is a linear combination of two normalized density operators:

$$H = \sum_j \lambda_j |j\rangle\langle j| = \alpha \left[\frac{1}{\alpha}\sum_{j:\lambda_j>0} \lambda_j |j\rangle\langle j|\right] - \beta \left[\frac{1}{\beta}\sum_{j:\lambda_j<0} |\lambda_j| |j\rangle\langle j|\right] \tag{5}$$

where $\alpha = \sum_{j:\lambda_j>0}\lambda_j$ and $\beta = \sum_{j:\lambda_j<0}|\lambda_j|$. Hence, O is a linear combination of four density operators. It follows that, by multiplying and summing the results of tomography of eight local states, we infer $(U \otimes \mathbf{1})\mathsf{S}_{i,i'}(U^\dagger \otimes \mathbf{1})$. The needed accuracy of the tomography is $O(\epsilon/n\,\mathrm{poly}(p))$ in operator norm to ensure that the error in the reconstructed $U \otimes U^\dagger$ is at most ϵ. If each site is a qubit, it suffices to determine just two density operators $U\frac{\mathbf{1}+\sigma_i^x}{2}U^\dagger = \frac{1}{2}U\sigma_i^x U^\dagger + \frac{1}{2}\mathbf{1}$ and $U\frac{\mathbf{1}+\sigma_i^z}{2}U^\dagger$ for each i, which are U-evolved Pauli stabilizer states. The qubit case is also found in [HLB+24, WZWY25].

As used in [HKT24], the learning of $UX_iU^\dagger$ for a set of i's that are sufficiently separated can be performed simultaneously. This is because the tomography of the state $U(\prod_i \rho_i)U^\dagger$, where ρ_i are density matrices whose supports are separated by distance twice the spread of U, can be performed independently around each i. This reduces the number of queries to U by a factor of roughly n/v, where n is the total number of qudits and v is the number of qudits in a ball of radius equal to the spread of U, compared to the naive strategy where we determine each $UX_iU^\dagger$ at a time.

QCA. We used only the condition that U maps a local operator to a local operator in the above algorithm, and hence U could have been any quantum cellular automaton (QCA) — by definition, a QCA α is an automorphism[9] of an operator algebra such that there exists a constant $r > 0$ such that for any single-site operator X, the image $\alpha(X)$ is supported on the r-neighborhood of the support of X. The minimal constant r here is the spread of a QCA. If a circuit has depth d,

[8] If U is a shallow circuit consisting of gates from a discrete set (*e.g.*, Clifford + T), then the complexity depends only on the depth. If the gates comprising U are from, say, $\mathsf{SU}(4)$, then $U\rho U^\dagger$ can only be determined up to some error with some high confidence, where the accuracy and confidence must be boosted for various global requirements, and system size dependence may become necessary. If the tomography has $\sim \epsilon^{-2}\log(1/\delta)$ sample complexity for accuracy ϵ in some state distance and confidence $1-\delta$, then one would want to consume $\sim n^2\epsilon^{-2}\log(n/\delta)$ samples of $U\rho U^\dagger$ to ensure that the global accuracy of $U \otimes U^\dagger$ is at most ϵ with confidence $1-\delta$.

[9] If the operator algebra is finite dimensional, then the Skolem–Noether theorem says that there exists a unitary U such that $\alpha(X) = UXU^\dagger$ for all operators X. Hence, it is safe to regard α as some global unitary.

then we know that its spread (of the QCA defined by the conjugation by the circuit) is $r = O(d)$. A primary question in the context of QCA is the converse: given an automorphism α of a local operator algebra of spread r, does there exist a quantum circuit of depth d that implements α? If so, how does d depend on the system size and the spread? If d is a constant independent of the system size, then the QCA is said to be trivial. It is quite a widely open problem how to determine whether a given QCA, or an infinite sequence of QCA of a fixed spread on finite systems of diverging sizes, is trivial.[10] The above algorithm outputs a shallow circuit that implements $\alpha \otimes \alpha^{-1}$, but does not address the QCA triviality problem. A related question for the circuit learning problem is if it is possible to output a circuit (of a similar depth to the input circuit's) that does not use any ancillas. Interestingly, a related state learning problem uses an arbitrarily small density of ancillas to output a circuit that creates a given shallow-circuit-evolved product state [LL24a, KKR24].

3 Gate complexity of QCA

As discussed above, a nontrivial QCA is a quantum circuit whose spread is small but depth is large. For example, a shift QCA in one dimension can be implemented by a staircase circuit of swap gates on any circle, where the spread is 1, but the depth is the circumference of the circle. Though it is a deep circuit, the gate count is only the number of sites. Such a circuit has been considered under the name of sequential quantum circuits [CDH+24]. It is shown in [CDH+24] using a characterization of QCA [SW04] by generalized Margolus partitioning that *for any translation invariant*[11] *QCA α of spread r on a D-dimensional system of n qudits $\mathbb{C}^p$ with periodic boundary conditions, there exists a unitary circuit U consisting of $O(n)$ gates such that $\alpha(x) = UxU^\dagger$ for all operator x.* Hence, in terms of gate count, every QCA is similar to a constant depth quantum circuit.

Here, we present another proof of the theorem of [CDH+24] by invertible subalgebra pumping. Our proof is inductive in the spatial dimension D where the base case is $D = 1$.[12]

With $D = 1$, a QCA is completely classified [GNVW12] by a rational index that captures the "flow." If the qudit dimension p is a prime, the rational index is any integer (zero, positive, or negative) power of p. An index p^e is achieved by a shift QCA that sends every qudit along the positive axis direction by e sites; if $e < 0$, the shift is to the left. The spread is $|e|$. A theorem of [GNVW12] is that if the index is $1 = p^0$, then the QCA is a constant depth quantum circuit. The shift QCA of index p^1 is implemented by a staircase of swap gates as one can check easily. Hence, the shift QCA of index p^e has a staircase circuit that is the $|e|$-fold composition of the swap-staircase. So, any QCA α is a composition of a shift and a quantum circuit of depth depending only on the spread of α, and the total gate count is $O(n)$ where the big-O notation hides a spread-dependent constant. If p is a composite number, then we can implement a shift QCA for each prime factor of p. The gate count is still $O(n)$.

With $D = 2$, one can use a classification result [FH20] (see also [Haa21]), but we give a more

[10] This is well understood in one dimension with qudits without tails [GNVW12], with tails [RWW22], with fermions [FPPV19], or with symmetry [MLC24]; and in two dimensions with qudits without tails [FH20] (see also [Haa21]); and in any dimensions in the translation invariant prime Clifford category [Haa25], but other situations remain open. See the introduction of [Haa25] and references therein.

[11] We believe that the assumption of translation invariance is merely a technical convenience, but we do not know of a simple enough argument that would cover cases without translation invariance.

[12] We could take $D = 0$ as the base case, but it would be more instructive to start with $D = 1$.

general argument, which will complete the inductive step of the proof. We will use a result of [Haa23] that says every QCA in any dimension is basically a shift in disguise. To explain this, we recall the following definition. Let Λ_{D-1} be a $(D-1)$-dimensional lattice of qudits $\mathbb{C}^p$. An **invertible subalgebra** $\mathcal{A}$ is a subalgebra of the tensor product algebra $\mathsf{Mat}(\Lambda_{D-1}) = \bigotimes_{i\in\Lambda_{D-1}} \mathsf{Mat}(p;\mathbb{C})_i$ such that the natural multiplication map

$$\mathcal{A}\otimes\mathcal{B} \ni a\otimes b \mapsto ab \in \mathsf{Mat}(\Lambda_{D-1}), \tag{6}$$

where $\mathcal{B}$ is the commutant of $\mathcal{A}$ within $\mathsf{Mat}(\Lambda_{D-1})$, *has a locality-preserving inverse*. By locality preserving we mean that there exists $r > 0$ such that every operator $x \in \mathsf{Mat}(\Lambda_{D-1})$ can be written as $x = \sum_j a_j b_j$ where $a_j \in \mathcal{A}$, $b_j \in \mathcal{B}$, and both $\mathrm{Supp}(a_j)$ and $\mathrm{Supp}(b_j)$ are contained in the r-neighborhood of $\mathrm{Supp}(x)$. The parameter r here is independent of the system size, and is analogous to the spread of a QCA. In zero dimensions, the local operator algebra $\mathsf{Mat}(\Lambda_0)$ is a finite dimensional matrix algebra, and every invertible subalgebra is a tensor factor. Because of the inverse, the multiplication map is an locality-preserving isomorphism of algebras and we may write $\mathcal{A}\otimes\mathcal{B} \cong \mathsf{Mat}(\Lambda_{D-1})$.

Given an invertible subalgebra $\mathcal{A}$ of $\mathsf{Mat}(\Lambda_{D-1})$, we construct an associated QCA in D dimensions as follows. Consider an (infinite) stack of $(D-1)$-dimensional sheets. Each sheet supports a copy of $\mathsf{Mat}(\Lambda_{D-1})$. The full algebra of the stack, $\mathsf{Mat}(\Lambda_D)$, is the (infinite) tensor product of those of the sheets. To define an algebra endomorphism on $\mathsf{Mat}(\Lambda_D)$, it suffices to specify the images of generators, which we take to be single-qudit operators. A single-qudit operator x is supported on, say, the k-th sheet, and we have a decomposition $x^{(k)} = \sum_j a_j^{(k)} b_j^{(k)}$ according to the definition of $\mathcal{A}$. Declare that $x^{(k)} \mapsto \sum_j a_j^{(k)} b_j^{(k+1)}$, where b_j are sent to the next sheet. It can be verified [Haa23, 3.9] that this is indeed an algebra automorphism that preserves the locality with a spread bounded by a function of r that appears in the definition of $\mathcal{A}$. We denote by $\beta_{\mathcal{A}}$ the QCA defined from $\mathcal{A}$.

A theorem [Haa23, 3.14] is that for every QCA α of a constant spread r on $\mathsf{Mat}(\Lambda_D)$ there exists a shift QCA η on $\mathsf{Mat}(\Lambda_D)$ and an invertible subalgebra $\mathcal{A}$ of $\mathsf{Mat}(\Lambda_{D-1})$ such that

$$\zeta = \alpha\circ\beta_{\mathcal{A}}\circ\eta \tag{7}$$

can be defined on a half space $\mathsf{Mat}(\Lambda_D^{z<0})$ by modifying its action only within distance $O(r)$ from the boundary, while keeping the spread bounded. From now on we use z as the coordinate along a fixed direction of the lattice. That is, there exists a QCA ζ' on $\mathsf{Mat}(\Lambda_D^{z<0})$ of spread $r' = O(r)$ such that $\zeta(x) = \zeta'(x)$ for all $x \in \mathsf{Mat}(\Lambda_D^{z<-r'}) \subset \mathsf{Mat}(\Lambda_D^{z<0})$. This property is called "blending" [FHH19] or "crossover" [GNVW12] to the vacuum.

Using the assumption of translation invariance, we now show that this blending into the vacuum implies that the QCA ζ is a composition of two layers of nonoverlapping $(D-1)$-dimensional QCA. The translation invariance implies that ζ can be defined on a half space with a boundary positioned at $z = c$ for *any* c. Let ζ_c be such a QCA on $\mathsf{Mat}(\Lambda_D^{z<c})$. By definition, ζ_c agrees with ζ on the region where $z < c - r'$. We trivially extend the domain of definition of ζ_c by identity so that the extension is defined on the whole $\mathsf{Mat}(\Lambda_D)$; the extended ζ_c acts by the identity for all operators on the region where $z \geq c$. For $c + 5r' < c'$, we let $\tau_{c,c'} = \zeta_{c'} \circ \zeta_c^{-1}$. By construction, $\tau_{c,c'}$ acts by identity on the region where $z < c - r'$ or $z \geq c'$, and acts by ζ on the region where $c \leq z < c' - r'$.

See Figure 3. Now, consider a trivial equation:

$$\zeta = \zeta \circ \underbrace{\left[\prod_{j\in\mathbb{Z}} \tau^{-1}_{20r'j,20r'j+10r'}\right]}_{\mu_1} \circ \underbrace{\left[\prod_{j\in\mathbb{Z}} \tau_{20r'j,20r'j+10r'}\right]}_{\mu_2} \tag{8}$$

where the constants 10 and 20 are some convenient choices. Because $\tau^{-1}_{20r'j,20r'j+10r'}$ cancels the action of ζ on the interior of its support, the first part μ_1 is a product of QCA supported on disjoint regions. Hence, both μ_1 and μ_2 are each a product of QCA over a collection of separated regions, each of which can be regarded as a $(D-1)$-dimensional QCA. By induction hypothesis, ζ has gate complexity $O(n)$. Since we know a circuit for a shift QCA η with gate count $O(n)$, the problem of implementing α using a number of gates linear in the system volume is now reduced to that of $\beta_{\mathcal{A}}$.

We show that $\beta_{\mathcal{A}}$ can be implemented in a staircase fashion where each step of the staircase is a $(D-1)$-dimensional QCA γ of a constant spread, analogous to the swap gate in the $D=1$ case.

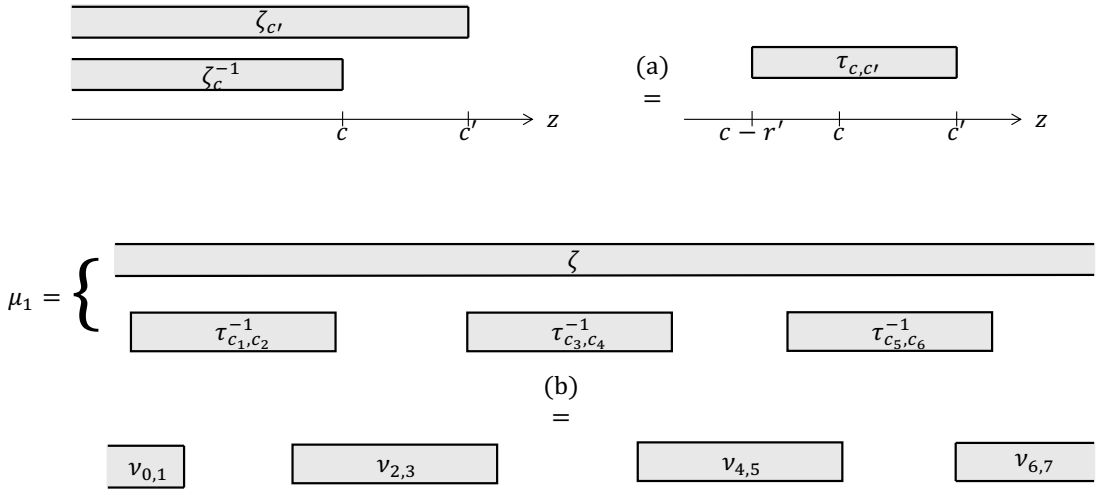

Figure 3: The QCA ζ_c acts as ζ on $z < c - r'$ but as identity on $z \geq c$. The composition $\tau_{c,c'} = \zeta_{c'} \circ \zeta_c^{-1}$ is then supported on the region $c - r' \leq z < c'$ which is compact along z-direction but otherwise not. This is the equality (a). Since $\tau_{c,c'}$ acts by ζ on the interior of the region $c - r' \leq z < c'$, a composition μ_1 of ζ and a product of separated τ's gives a product of separated QCA, each of which can be regarded as a QCA in one dimension lower. This is the equality (b).

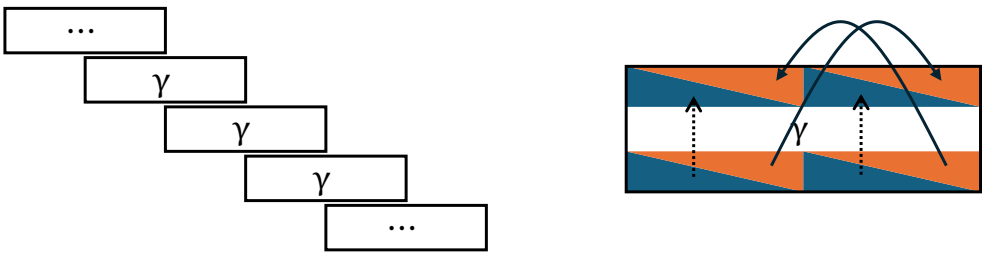

Figure 4: Implementation of $\beta_{\mathcal{A}}$ that pumps an invertible subalgebra $\mathcal{B}$ using a swap-like QCA γ. The blue triangle on the right diagram is $\mathcal{A}$ and the orange is $\mathcal{B}$. By the staircase, only $\mathcal{B}$ is pumped to the right, which is exactly the action of $\beta_{\mathcal{A}}$.

Take k-th and $(k+1)$-st sheets supporting $\mathsf{Mat}(\Lambda_{D-1})^{(k)} \otimes \mathsf{Mat}(\Lambda_{D-1})^{(k+1)}$. By definition of the invertible subalgebra $\mathcal{A}$, we have

$$\mathsf{Mat}(\Lambda_{D-1})^{\otimes 2} \cong \mathcal{A}^{(k)} \otimes \mathcal{B}^{(k)} \otimes \mathcal{A}^{(k+1)} \otimes \mathcal{B}^{(k+1)}. \tag{9}$$

The swap-like QCA γ is defined as the following. See Figure 4.

$$\begin{array}{rccc} \gamma: & \mathcal{A}^{(k)} \otimes \mathcal{B}^{(k)} \otimes \mathcal{A}^{(k+1)} \otimes \mathcal{B}^{(k+1)} & \ni & a \otimes b \otimes c \otimes d \\ & \downarrow & & \downarrow \\ & \mathcal{A}^{(k)} \otimes \mathcal{B}^{(k)} \otimes \mathcal{A}^{(k+1)} \otimes \mathcal{B}^{(k+1)} & \ni & a \otimes d \otimes c \otimes b \end{array} \tag{10}$$

Obviously, γ is an automorphism of $\mathsf{Mat}(\Lambda_{D-1})^{\otimes 2}$ with spread given by a function of r where r appears in the definition of the invertible subalgebra $\mathcal{A}$. The promised implementation of $\beta_{\mathcal{A}}$ using the QCA γ is shown in Figure 4. Therefore, the QCA $\beta_{\mathcal{A}}$ is a staircase circuit by the induction hypothesis using $O(n)$ gates. This completes the proof.

We have converted a QCA α into a constant number (independent of n) of staircase unitary circuits, so that every qubit goes through a constant number of local unitary gates. We did not answer whether the staircase circuit could be arranged so that all the gates were contained in a single hyperplane in spacetime of small thickness. The final circuit from our inductive proof consists of a constant number of hyperplanes in spacetime of constant thickness that contain all the local gates.

In connection to the previous section on the shallow circuit learning problem, an open problem is whether there is an efficient query algorithm for the conversion from a QCA to a staircase circuit. The two types of data that define a QCA α — one a table of images of single-qudit operators under a QCA and the other a local unitary circuit in a staircase fashion — contain the same number $O(n)$ of parameters up to a fixed multiplicative constant. The former is perhaps the most obvious type of data for α, but the latter is more appropriate if we are to apply α to some arbitrary operator whose decomposition in terms of single-qudit operators is complicated.

Acknowledgments

I thank Hsin-Yuan (Robert) Huang for discussions and Sumner Hearth, Anushya Chandran, and Chris Laumann for explaining [HFCL25].

References

[ANW11] Pablo Arrighi, Vincent Nesme, and Reinhard Werner. Unitarity plus causality implies localizability. *Journal of Computer and System Sciences*, 77:372–378, 2011. arXiv:0711.3975, doi:10.1016/j.jcss.2010.05.004.

[CDH+24] Xie Chen, Arpit Dua, Michael Hermele, David T. Stephen, Nathanan Tantivasadakarn, Robijn Vanhove, and Jing-Yu Zhao. Sequential quantum circuits as maps between gapped phases. *Phys. Rev. B*, 109(7):075116, 2024. arXiv:2307.01267, doi:10.1103/physrevb.109.075116.

[CGW10] Xie Chen, Zheng-Cheng Gu, and Xiao-Gang Wen. Local unitary transformation, long-range quantum entanglement, wave function renormalization, and topological order. *Phys. Rev. B*, 82(15):155138, 2010. arXiv:1004.3835, doi:10.1103/physrevb.82.155138.

[FH20] M. Freedman and M. B. Hastings. Classification of quantum cellular automata. *Commun. Math. Phys.*, 376(2):1171–1222, 2020. arXiv:1902.10285, doi:10.1007/s00220-020-03735-y.

[FHH19] Michael Freedman, Jeongwan Haah, and Matthew B. Hastings. The group structure of quantum cellular automata. *Commun. Math. Phys. 389, 1277-1302 (2022)*, 389(3):1277–1302, 2019. arXiv:1910.07998, doi:10.1007/s00220-022-04316-x.

[FPPV19] Lukasz Fidkowski, Hoi Chun Po, Andrew C. Potter, and Ashvin Vishwanath. Interacting invariants for floquet phases of fermions in two dimensions. *Phys. Rev. B*, 99(8):085115, February 2019. arXiv:1703.07360, doi:10.1103/physrevb.99.085115.

[GNVW12] D. Gross, V. Nesme, H. Vogts, and R. F. Werner. Index theory of one dimensional quantum walks and cellular automata. *Commun. Math. Phys.*, 310(2):419–454, 2012. arXiv:0910.3675, doi:10.1007/s00220-012-1423-1.

[Haa21] Jeongwan Haah. Clifford quantum cellular automata: Trivial group in 2d and Witt group in 3d. *J. Math. Phys.*, 62(9):092202, 2021. arXiv:1907.02075, doi:10.1063/5.0022185.

[Haa23] Jeongwan Haah. Invertible subalgebras. *Commun. Math. Phys.*, 403(2):661–698, 2023. arXiv:2211.02086, doi:10.1007/s00220-023-04806-6.

[Haa25] Jeongwan Haah. Topological phases of unitary dynamics: Classification in Clifford category. *Commun. Math. Phys.*, 406(76), 2025. arXiv:2205.09141, doi:10.1007/s00220-025-05239-z.

[HFCL25] Sumner N. Hearth, Michael O. Flynn, Anushya Chandran, and Chris R. Laumann. Unitary k-designs from random number-conserving quantum circuits. *Phys. Rev. X*, 15(021022), 2025. arXiv:2306.01035, doi:10.1103/PhysRevX.15.021022.

[HKT24] Jeongwan Haah, Robin Kothari, and Ewin Tang. Learning quantum Hamiltonians from high-temperature Gibbs states and real-time evolutions. *Nature Physics*, 20(6):1027–1031, 2024. arXiv:2108.04842, doi:10.1038/s41567-023-02376-x.

[HLB+24] Hsin-Yuan Huang, Yunchao Liu, Michael Broughton, Isaac Kim, Anurag Anshu, Zeph Landau, and Jarrod R. McClean. Learning shallow quantum circuits. *In Proceedings of the 56th Annual ACM Symposium on Theory of Computing (STOC 2024)*, pages 1343–1351, 2024. arXiv:2401.10095, doi:10.1145/3618260.3649722.

[KKR24] Hyun-Soo Kim, Isaac H. Kim, and Daniel Ranard. Learning state preparation circuits for quantum phases of matter. October 2024. arXiv:2410.23544, doi:10.48550/ARXIV.2410.23544.

[KVH18] Vedika Khemani, Ashvin Vishwanath, and D. A. Huse. Operator spreading and the emergence of dissipative hydrodynamics under unitary evolution with conservation laws. *Phys. Rev. X*, 8(3):031057, September 2018. arXiv:1710.09835, doi:10.1103/physrevx.8.031057.

[LL24a] Zeph Landau and Yunchao Liu. Learning quantum states prepared by shallow circuits in polynomial time. 2024. arXiv:2410.23618, doi:10.48550/ARXIV.2410.23618.

[LL24b] Nicholas LaRacuente and Felix Leditzky. Approximate unitary k-designs from shallow,

low-communication circuits. July 2024. arXiv:2407.07876, doi:10.48550/ARXIV.2407.07876.

[Mar22] Iman Marvian. Restrictions on realizable unitary operations imposed by symmetry and locality. *Nat. Phys.*, 18(3):283–289, 2022. arXiv:2003.05524, doi:10.1038/s41567-021-01464-0.

[MLC24] Ruochen Ma, Yabo Li, and Meng Cheng. Quantum cellular automata on symmetric subalgebras. November 2024. arXiv:2411.19280, doi:10.48550/ARXIV.2411.19280.

[RPvK18] Tibor Rakovszky, Frank Pollmann, and C. W. von Keyserlingk. Diffusive hydrodynamics of out-of-time-ordered correlators with charge conservation. *Phys. Rev. X*, 8(3):031058, September 2018. arXiv:1710.09827, doi:10.1103/physrevx.8.031058.

[RWW22] Daniel Ranard, Michael Walter, and Freek Witteveen. A converse to lieb-robinson bounds in one dimension using index theory. *Annales Henri Poincaré*, 23(11):3905–3979, July 2022. arXiv:2012.00741, doi:10.1007/s00023-022-01193-x.

[SHH24] Thomas Schuster, Jonas Haferkamp, and Hsin-Yuan Huang. Random unitaries in extremely low depth. 2024. arXiv:2407.07754, doi:10.48550/ARXIV.2407.07754.

[SW04] B. Schumacher and R. F. Werner. Reversible quantum cellular automata. May 2004. arXiv:quant-ph/0405174, doi:10.48550/ARXIV.QUANT-PH/0405174.

[TTS24] Emanuele Tirrito, Xhek Turkeshi, and Piotr Sierant. Anticoncentration and magic spreading under ergodic quantum dynamics. December 2024. arXiv:2412.10229, doi:10.48550/ARXIV.2412.10229.

[WZWY25] Yusen Wu, Yukun Zhang, Chuan Wang, and Xiao Yuan. Hamiltonian dynamics learning: A scalable approach to quantum process characterization. 2025. arXiv:2503.24171, doi:10.48550/ARXIV.2503.24171.