nature portfolio

https://doi.org/10.1038/s41567-022-01539-6

## Supplementary information

# Linear growth of quantum circuit complexity

In the format provided by the authors and unedited

# Supplementary Material for "Linear growth of quantum circuit complexity"

J. Haferkamp, P. Faist, N. B. T. Kothakonda, J. Eisert and N. Yunger Halpern

### Appendix A: Algebraic and semialgebraic sets

For convenience, we review elementary aspects of algebraic geometry over the real numbers. We apply these properties in the proof of Theorem 1. Ref. [45] contains a more comprehensive treatment.

**Definition 4** (Algebraic set)**.** *A subset $V \subseteq \mathbb{R}^m$ is called an* algebraic set*, or an* algebraic variety*, if, for a set of polynomials $\{f_j\}_j$,*

$$V = \{x \in \mathbb{R}^m | f_j(x) = 0\}. \tag{A1}$$

*A subset $V' \subseteq V$ is called an* algebraic subset *if $V'$ is an algebraic set. We call a subset $W \subseteq \mathbb{R}^m$ a* semialgebraic set *if, for sets $\{f_j\}_j$ and $\{g_k\}_k$ of polynomials,*

$$W = \{x \in \mathbb{R}^m | f_j(x) = 0, g_k(x) \leq 0\}. \tag{A2}$$

A natural topology on algebraic sets is the Zariski topology.

**Definition 5** (Zariski topology)**.** *Let $V$ denote an algebraic set. The Zariski topology is the unique topology whose closed sets are the algebraic subsets of $V$.*

A traditional definition of "dimension" for algebraic sets involves irreducible sets.

**Definition 6** (Irreducible set)**.** *Let $X$ denote a topological space. $X$ is called* irreducible *if it is not the union of two proper closed subsets.*

**Definition 7** (Dimension of algebraic sets)**.** *Let $V$ be an algebraic set that is irreducible with respect to the Zariski topology. The dimension of $V$ is the maximal length $d$ of any chain $V_0 \subset V_1 \subset \cdots \subset V_d$ of distinct nonempty irreducible algebraic subsets of $V$.*

The relevant algebraic sets in the proof of Theorem 1 are $\mathrm{SU}(4)^{\times R}$ and $\mathrm{SU}(2^n)$. Our interest in semialgebraic sets stems from the following principle. In the following, we refer to a function $F : \mathbb{R}^n \to \mathbb{R}^m$ as a *polynomial map* if its entries are polynomials in the entries of its input.

**Theorem 2** (Tarski-Seidenberg principle)**.** *Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be a polynomial map. If $W$ is a semialgebraic set, so is $F(W)$.*

The Tarski-Seidenberg principle applies to the map that contracts sets of quantum gates. This application is important for us, because it provides a natural notion of dimension for the contraction map's image.

All semialgebraic sets (and hence all algebraic sets) decompose into smooth manifolds.

**Theorem 3** (Stratification of semialgebraic sets)**.** *If $W$ is a semialgebraic set, then $W = \bigcup_{j=1}^N M_j$, wherein each $M_j$ denotes a smooth manifold. If $W$ is an algebraic set of dimension $d$ in the sense of Definition 7, then $\max_j\{\dim(M_j)\} = d$.*

This $\max_j\{\dim(M_j)\}$ does not depend on the decomposition chosen. This independence motivates the following definition:

**Definition 8** (Dimension of semialgebraic sets)**.** *Let $W$ denote a semialgebraic set, such that $W = \bigcup_{j=1}^N M_j$, wherein each $M_j$ denotes a manifold. The greatest dimension of any manifold, $\max_j\{\dim(M_j)\}$, is the semialgebraic set's dimension.*

This definition generalizes Definition 7, due to Theorem 3. One more fact about semialgebraic sets' dimensions will prove useful:

**Lemma 2** (Dimension of an image)**.** *Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be a polynomial map. If $W$ is a dimension-$d$ semialgebraic set, $F(W)$ is of dimension $\leq d$.*

The bound follows from combining the results of Ref. [45, Prop. 2.8.7] with the results of Ref. [45, Prop. 2.8.6]. (Ref. [45] invokes a semialgebraic mapping, which encompasses polynomial maps.)

### Appendix B: Proof of the main theorem and lemmata

In this appendix, we prove Lemma 3, Lemma 4, and the main theorem. The proofs rely on the topics reviewed in Appendix A, as well as the following notation and concepts. In differential geometry, the rank of $F^A$ at the point $x = (U_1, U_2, \ldots, U_R)$ is defined as the rank of the derivative $D_x F^A$. Mapping lists of gates to unitaries, $F$ is a complicated object. We can more easily characterize a map from real numbers to real numbers. Related is a map from Hermitian operators to Hermitian operators: An $n$-qubit state evolves under a Hamiltonian represented by a $2^n \times 2^n$ Hermitian operator, which has $(2^n)^2 = 4^n$ real parameters. Therefore, for convenience, we shift focus from unitaries to their Hermitian generators. We construct a map whose domain is the algebra $\mathfrak{su}(4)^{\times R} \simeq \mathbb{R}^{15R}$ that generates $\mathrm{SU}(4)^{\times R}$. The range is the set of $n$-qubit Hermitian operators, $\mathfrak{su}(2^n) \simeq \mathbb{R}^{4^n}$. We construct such a map from three steps, depicted by the dashed lines in Fig. 1.

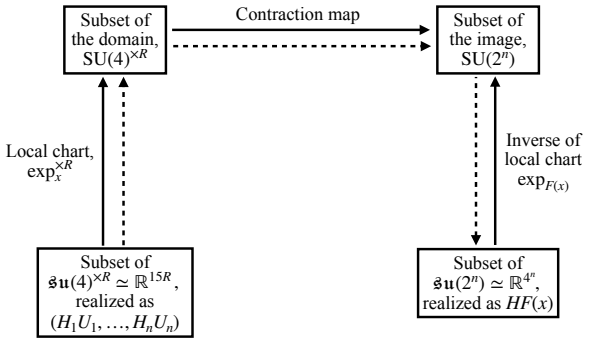


Figure 1. Three-part map used in the proof of Lemma 3. $H_j$ denotes the $j^{\text{th}}$ two-qubit Hermitian operator, $U_j$ denotes the $j^{\text{th}}$ two-qubit unitary, and $H$ denotes an $n$-qubit Hermitian operator.

The first step is a chart, a diffeomorphism that maps one manifold to another invertibly. Our chart acts on the algebra $\mathfrak{su}(4)^{\times R}$ that generates $\mathrm{SU}(4)^{\times R}$. To define the chart, we parameterize an element $H$ of the $j^{\text{th}}$ copy of $\mathfrak{su}(4)$:

$$H = \sum_{\substack{\alpha,\beta \in \{\mathbb{1},X,Y,Z\} \\ (\alpha,\beta) \neq (\mathbb{1},\mathbb{1})}} \lambda_{j,\alpha,\beta}\ \alpha \otimes \beta, \tag{B1}$$

wherein $\lambda_{j,\alpha,\beta} \in \mathbb{R}$. For each point $x = (U_1, U_2, \ldots, U_R) \in \mathrm{SU}(4)^{\times R}$, we define the local exponential chart $\exp_x^{\times R} : \mathfrak{su}(4)^{\times R} \to \mathrm{SU}(4)^{\times R}$ as $\exp_x^{\times R}(H_1, \ldots, H_R) := (e^{\mathrm{i}H_1}U_1, \ldots, e^{\mathrm{i}H_R}U_R)$, and we define the analogous $\exp_U : \mathfrak{su}(2^n) \to \mathrm{SU}(2^n)$ as $\exp_U(H) := e^{\mathrm{i}H}U$. These charts are standard for matrix Lie groups. Both are locally invertible in small neighbourhoods around $x$ and $U$, by a standard result in Lie-group theory [46]. The three-part map, represented by the dashed lines in Fig. 1, has the form $\exp^{-1}_{F^A(x)} \circ F^A \circ \exp_x^{\times R}$.

We now characterize the map's derivative, to characterize the derivative of $F^A$, to characterize the rank of $F^A$. Denote by $D_0$ the derivative evaluated where the Hermitian operators are set to zero, such that each chart reduces to the identity operation. The image of $D_0\big(\exp^{-1}_{F^A(x)} \circ F^A \circ \exp_x^{\times R}\big)$ is spanned by the operators

$$\partial_{\lambda_{j,A,B}} \left(\exp^{-1}_{F^A(x)} \circ F^A \circ \exp_x^{\times R}\right)\Big|_0. \tag{B2}$$

These operators have the form

$$U_R \ldots U_{j+1} P U_j \ldots U_1, \tag{B3}$$

wherein $P$ denotes a two-qubit Pauli operator. We apply the setting above to prove the following restatement of Lemma 1 in the methods section.

**Lemma 3** (Low-rank locus)**.** *The low-rank locus $E_{<r_{\max}}$ is an algebraic set of measure 0 and so is closed (in the Lie-group topology). Equivalently, $E_{r_{\max}}$ is an open set of measure 1. Consequently, $d_A = r_{\max}$.*

*Proof.* Consider representing an operator (B3) as a matrix relative to an arbitrary tensor-product basis. To identify the matrix's form, we imagine representing the unitaries in $\mathrm{SU}(4)^{\times R}$ as matrices relative to the corresponding tensor-product basis for $\mathbb{C}^2 \otimes \mathbb{C}^2$. Combining the unitary matrices' elements polynomially yields the matrix elements of (B3).

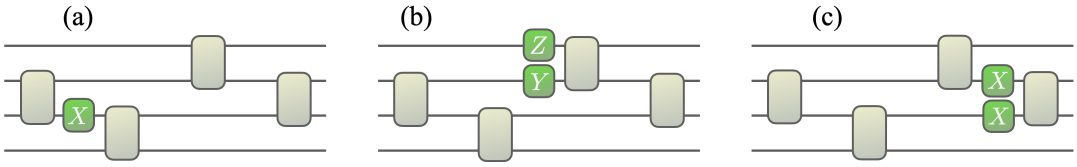


Figure 2. Examples of partial derivatives $\partial_{\lambda_{j,\alpha,\beta}}(\exp^{-1}_{F^A(x)} \circ F^A \circ \exp_x^{\times 4})\big|_{\lambda_{j,\alpha,\beta}=0}$ that span the image of $D_0(\exp_{F^A(x)} \circ F^A \circ \exp_x^{\times 4})$.

$D_x F^A$ has the same rank as $D_0(\exp^{-1}_{F^A(x)} \circ F^A \circ \exp_x^{\times R})$, because $\exp_x^{\times R}$ and $\exp_{F^A(x)}$ are local charts [47]. Recall that $E_{<r_{\max}}$ denotes the locus of points, in $\mathrm{SU}(4)^{\times R}$, where $F^A$ has a rank $< r_{\max}$. Equivalently, by the invertible-matrix theorem, $E_{<r_{\max}}$ consists of the points where certain minors of $D_0(\exp^{-1} \circ F^A \circ \exp_x^{\times R})$—the determinants of certain collections of $r_{\max} \times r_{\max}$ matrix elements—vanish. The determinants' vanishing implies a set of equations polynomial in the matrix elements of $D_0(\exp^{-1} \circ F^A \circ \exp_x^{\times R})$—and so, by the last paragraph, polynomial in the entries of matrices in $\mathrm{SU}(4)^{\times R}$. $\mathrm{SU}(4)^{\times R}$ is a real algebraic set, being the set of operators that satisfy the polynomial equations equivalent to $UU^\dagger = \mathbb{1}$ and $\det U = 1$. Thus, by Definition 4, the points of rank $< r$ form an algebraic subset of $\mathrm{SU}(4)^{\times R}$.

We can now invoke properties of algebraic subsets, reviewed in Appendix A. First, we prove that $\mathrm{SU}(4)^{\times R}$ is irreducible in the Zariski topology. The Zariski topology of $\mathrm{SU}(4)^{\times R}$ is coarser than the topology inherited from $(\mathbb{C}^{4\times 4})^{\times R}$, identified with $\mathbb{R}^{32R}$. As $\mathrm{SU}(4)^{\times R}$ is connected in the finer topology, so is $\mathrm{SU}(4)^{\times R}$ connected in the Zariski topology. This connectedness implies that $\mathrm{SU}(4)^{\times R}$ is irreducible, as $\mathrm{SU}(4)^{\times R}$ is an algebraic group [48, Summary 1.36]. Being irreducible, $\mathrm{SU}(4)^{\times R}$ has a dimension à la Definition 7. If the low-rank locus $E_{<r_{\max}}$ is not all of $\mathrm{SU}(4)^{\times R}$, then it is, by Definition 7, a lower-dimensional algebraic subset. Every dimension-$N$ algebraic subset decomposes into a collection of submanifolds, each of which has dimension $\leq N$ [45, Prop. 9.1.8]. As a proper submanifold has measure 0, $E_{<r_{\max}}$ has measure 0. As an algebraic subset, $E_{<r_{\max}}$ is closed in the Lie-group topology.

Finally, we prove that $d_A = r_{\max}$. In a small open neighborhood $V$ of a point $x \in E_{r_{\max}}$, the contraction map's rank is constant, by Lemma 3. By the constant-rank theorem [47, Thm 5.13], therefore, $F^{A_T}$ acts locally as a projector throughout $E_{r_{\max}}$—and so throughout $\mathrm{SU}(4)^{\times R}$ (except on a region of measure 0, by Lemma 3). The projector has a rank, like $F^{A_T}$, of $r_{\max}$. A rank-$r_{\max}$ projector has an image that is a dimension-$r_{\max}$ manifold. Hence $r_{\max} \leq d_A$. The other direction, $d_A \leq r_{\max}$, follows directly from Sard's theorem [49]. Let $X_r$ denote the set of points where $F^A$ is rank-$r$. As $F^A$ is a smooth map, Sard's theorem ensures that $r$ upper-bounds the Hausdorff dimension of the image $F^A(X_r)$. As $F^A\big(\mathrm{SU}(4)\big)$ is a semialgebraic set, it stratifies into manifolds, by Theorem 3. Therefore, the Hausdorff dimension coincides with the semialgebraic set's dimension. □

Lemma 3, combined with the following lemma, implies Proposition 1.

**Lemma 4** (Existence of a high-rank point)**.** *Let $T \in \mathbb{Z}_{>0}$ denote any nonnegative integer. Consider any architecture $A_T$ formed from $T$ $L$-gate, backwards-light-cone–containing blocks. The map $F^{A_T}$ has the greatest rank possible, $r_{\max} \geq T$.*

*Proof.* Without loss of generality, we assume that all $T$ blocks have identical architectures. This assumption will simplify the notation below. We can lift the assumption by complicating the notation.

Consider an arbitrary point $x = (U_1, U_2, \ldots, U_R) \in \mathrm{SU}(4)^{\times R}$. For all $x$, the contraction map $F^{A_T}$ has a derivative characterized, in the proof of Lemma 3, with local charts $\exp_{F^{A_T}(x)}$ and $\exp_x^{\times R}$. The number of gates in $A_T$ is $R \leq TL$. The map $F^{A_T}$ has an image spanned by the partial derivatives $\partial_{\lambda_{j,\alpha,\beta}}(\exp^{-1}_{F^{A_T}(x)} \circ F^{A_T} \circ \exp_x^{\times R})\big|_{\lambda_{j,\alpha,\beta}=0}$. Each partial derivative has the form

$$U_R U_{R-1} \ldots U_{j+1} (\alpha \otimes \beta) U_j U_{j-2} \ldots U_1 \tag{B4}$$

(Fig. 2). $\alpha$ and $\beta$ denote Pauli operators; each acts nontrivially on just one of the two qubits on which $U_j$ acts nontrivially. We implicitly pad operators with identities wherever necessary, such that the operators act on the appropriate Hilbert space.

We aim to lower-bound the greatest possible rank, $r_{\max}$, of the map $F^{A_T}$. To do so, we construct a point

$$x_T = \Big( \underbrace{C_1^{(1)}, \ldots, C_1^{(L)}}_{L \text{ gates}}, \ldots, \underbrace{C_T^{(1)}, \ldots, C_T^{(L)}}_{L \text{ gates}} \Big) \in \mathrm{SU}(4)^{\times R} . \tag{B5}$$

We will choose for the $C_j^{(i)}$'s to be Clifford gates. A gate's subscript, $j$, labels the blocks to which the gate belongs. The superscript, $i$, labels the gate's position within the block. The gates constitute a block as $C_j^{(L)} C_j^{(L-1)} \ldots C_j^{(1)} =: C_j$. Our

construction of $C_j$ relies on a property of an arbitrary Pauli operator $Q_j$: We can choose the Clifford gates $C_j^{(i)}$ such that block $C_j$ maps $Q_j$ to a $Z$ on qubit $t$: $C_j Q_j C_j^\dagger = Z_t \equiv \mathbb{1}^{\otimes(t-1)} \otimes Z \otimes \mathbb{1}^{\otimes(n-t)}$. We now show how the existence of such a Clifford unitary $C_j$ implies Lemma 4. Afterward, we show to construct $C_j$.

Let us choose the Pauli strings $Q_j$ that guide our construction of the Clifford block $C_j$. We choose the $Q_j$'s inductively over $T$ such that $\{(C_T C_{T-1} \dots C_j) Q_j (C_{j-1} C_{j-1} \dots C_1)\}_{1 \leq j \leq T}$ is linearly independent. We start with an arbitrary Pauli string $Q_1$. The form of $Q_1$ guides our construction of $C_1$. Second, we choose for $Q_2$ to be an arbitrary Pauli string $\neq C_1 Q_1 C_1^\dagger$. $Q_2$ guides our construction of $C_2$. Third, we choose for $Q_3$ to be an arbitrary Pauli string outside $\mathrm{span}\{C_1 C_2 Q_1 C_2^\dagger C_1^\dagger, C_2 Q_2 C_2^\dagger\}$. This $Q_3$ guides our construction of $C_3$. After $T$ steps, we have constructed all the $Q_j$'s and $C_j$'s. If $T < 4^n - 1$, enough Pauli strings exist that, at each step, a Pauli string lies outside the relevant span.

The operators $(C_T C_{T-1} \dots C_j) Q_j (C_{j-1} C_{j-2} \dots C_1)$, for $j \in [1, T]$, are in the image of $D_0(\exp^{-1}_{F^{A_T}(x_T)} \circ F^{A_T} \circ \exp^{\times(R)}_{x_T})$:

$$\begin{aligned} \partial_{\lambda_{jL,\mathbb{1}_1,\mathbb{Z}}} \left( \exp^{-1}_{F^{A_T}(x_T)} \circ F^{A_T} \circ \exp^{\times R}_{x_T} \right) \Big|_0 &= (C_T C_{T-1} \dots C_{j+1})(\mathbb{1}_{t-1} \otimes Z_t \otimes \mathbb{1}_{n-t})(C_j C_{j-1} \dots C_1) \\ &= (C_T C_{T-1} \dots C_j) Q_j (C_{j-1} C_{j-2} \dots C_1). \end{aligned} \tag{B6}$$

We have assumed, without loss of generality, that each block's final gate acts on qubit $t$. For all $j \in [1, T]$, the operators $(C_T C_{T-1} \dots C_j) Q_j (C_{j-1} C_{j-2} \dots C_1)$ are in the image of $D_0(\exp^{-1}_{F^{A_T}(x_T)} \circ F^{A_T} \circ \exp^{\times R}_{x_T})$ and are linearly independent. Therefore, the rank of $F^{A_T}$ at the point $x_T$ is $\geq T$.

In the remainder of this proof, we provide the missing link: We show that, for every Pauli string $P$, we can construct a backwards-light-cone–containing block that implements a Clifford unitary $C = C^{(L)} C^{(L-1)} \dots C^{(1)}$ such that $CPC^\dagger = Z_t$. We drop subscripts because subscripts index blocks and this prescription underlies all blocks. By definition, each block contains a qubit $t$ to which each other qubit $t'$ connects via gates in the block. The path from a given qubit $t'$ depends on $t'$, and multiple paths may connect a $t'$ to $t$. Also, one path may connect $t$ to multiple qubits. We choose an arbitrary complete set of paths (which connect all the other qubits to $t$) that satisfies the merging property described below. To introduce the merging property, we denote by $m$ the number of paths in the set. Let $p \in [1, m]$ index the paths. Path $p$ contacts the qubits in the order $i_{p,1} \mapsto i_{p,2} \mapsto \dots \mapsto i_{p,l_p} = t$, reaching $l_p \in [1,\ L+1]$ qubits. We choose the paths such that they merge whenever they cross: If $i_{p,j} = i_{p',j'}$, then $i_{p,j+k} = i_{p',j'+k}$ for all $k \in \{1, 2, \dots, l_p - j = l_{p'} - j'\}$. We choose for all the gates outside these paths to be identities. Next, we choose the nontrivial gates in terms of an arbitrary Pauli string.

Let $P = \bigotimes_{j=1}^n P_j$ denote an arbitrary nontrivial $n$-qubit Pauli string. Some Clifford unitary $C$ maps $P$ to a Pauli string that acts nontrivially on just one qubit, which we choose to be $t$. Indeed, consider conjugating an arbitrary $n$-qubit Pauli operator $P$ with a uniformly random Clifford operator $C$. The result, $C^\dagger P C$, is a uniformly random $n$-qubit Pauli operator [50]. Therefore, for every initial Pauli operator $P$ and every final Pauli operator, some Clifford operator $C$ maps one to the other.

We arbitrarily choose for the string's nontrivial single-qubit Pauli operator to be $Z$. Let $i_{p,k_p}$ denote the first index $j$ in $i_{p,1} \mapsto i_{p,2} \mapsto \dots \mapsto i_{p,l_p}$ for which $P_j \neq \mathbb{1}$. By definition, $P_{i_{p,k}} \otimes P_{i_{p,k+1}}$ is a nontrivial Pauli string. There exists a two-local Clifford gate $C^{p,(0)}$ that transforms $P_{i_{p,k}} \otimes P_{i_{p,k+1}}$ into a $Z$ acting on qubit $i_{p,k+1}$:

$$C^{p,(0)} (P_{p,i_k} \otimes P_{p,i_{k+1}}) (C^{p,(0)})^\dagger = \mathbb{1}_{i_k} \otimes Z_{i_{k+1}}. \tag{B7}$$

Operating with $C^{p,(0)}$ (padded with $\mathbb{1}$'s) on the whole string $P$ yields another Pauli string:

$$C^{p,(0)} P (C^{p,(0)})^\dagger = \bigotimes_{j=1}^n P_j^{p,(1)}. \tag{B8}$$

Let $i_{p,\ell}$ denote the first index $j$ for which $P_j^{p,(1)}$ is a nontrivial Pauli operator. Since

$$P^{(1)}_{i_{p,k+1}} \otimes P^{(1)}_{i_{p,k+2}} = Z_{i_{p,k+1}} \otimes P_{i_{p,k+2}}, \tag{B9}$$

$i_{p,\ell} = i_{p,k+1}$. There exists a two-local Clifford gate $C^{p,(1)}$ that shifts the $Z$ down the path:

$$C^{p,(1)} \left( P^{p,(1)}_{i_{p,k+1}} \otimes P^{p,(1)}_{i_{p,k+2}} \right) (C^{p,(1)})^\dagger = \mathbb{1}_1 \otimes Z_{i_{p,k+2}}. \tag{B10}$$

We perform this process—of shifting the $Z$ down the path and leaving an $\mathbb{1}_1$ behind—for every path simultaneously. For example, if we begin with two equal-length paths, $C^{(2)} = C^{p,(2)} C^{p',(2)}$. This simultaneity is achievable until two paths merge. Whenever paths merge, we choose the next Clifford gate such that we proceed along the merged path. Every qubit is visited, and every path ends at qubit $t$. Therefore, we have constructed a circuit that implements a Clifford operation $C$ such that $CPC^\dagger = Z_t$. Figure 4(d) depicts an example of this construction. □

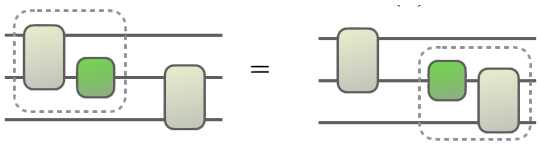

Figure 3. Along every contraction is a redundant copy of the gauge group $\mathrm{SU}(2)$.

The foregoing proof has a surprising implication: A map's rank is somewhat divorced from a circuit's complexity. The rank of $F^{A_T}$ at $x_T$ is at least $T$, which could be a large number. Yet, the contracted unitary corresponding to this circuit is Clifford. Hence the extended circuit's complexity surpassed the original circuit's complexity only a little—by, at most, $O\big(n^2/\log(n)\big)$ [51].

Finally, we combine Lemmata 3 and 4 to prove Theorem 1:

**Theorem 1** (Linear growth of complexity)**.** *Let $U$ denote a unitary implemented by a random quantum circuit in an architecture formed by concatenating $T$ blocks of $\leq L$ gates each, each block containing a backwards light cone. The unitary's circuit complexity is lower-bounded as*

$$\mathcal{C}_{\mathrm{u}}(U) \geq \frac{R}{9L} - \frac{n}{3} \, , \tag{1}$$

*with unit probability, until the number of gates grows to $T \geq 4^n - 1$. The same bound holds for $\mathcal{C}_{\mathrm{state}}(U|0^n\rangle)$, until $T \geq 2^{n+1} - 1$.*

*Proof of Theorem 1.* We reuse the notation introduced in Lemmata 3 and 4. Examples include $A_T$, an arbitrary architecture that satisfies the assumptions in Lemma 4 and that consists of $R \leq TL$ gates. $F^{A_T}$ denotes the corresponding contraction map. $E_{r_{\max}}$ denotes the locus of points at which $F^{A_T}$ achieves its greatest rank, $r_{\max}$. In a small open neighborhood $V$ of a point $x \in E_{r_{\max}}$, the contraction map's rank is constant, by Lemma 3. By the constant-rank theorem [47, Thm 5.13], therefore, $F^{A_T}$ acts locally as a projector throughout $E_{r_{\max}}$—and so throughout $\mathrm{SU}(4)^{\times R}$ (except on a region of measure 0, by Lemma 3). The projector has a rank, like $F^{A_T}$, of $r_{\max}$. Therefore, in the open set $V \subseteq \mathrm{SU}(4)^{\times R}$, $F^{A_T}$ is equivalent, up to a diffeomorphism, to the projection

$$\big(x_1, \ldots, x_{\dim(\mathrm{SU}(4)^{\times R})}\big) \mapsto (x_1, \ldots, x_{r_{\max}}, \underbrace{0, \ldots, 0}_{\dim(\mathrm{SU}(2^n)) - r_{\max}}). \tag{B11}$$

For simplicity of notation, we identify $V$ with its image under the local diffeomorphism (we do not distinguish $V$ from its image notationally).

The open subset $V$ contains, itself, an open subset that decomposes as a product: $V_1 \times V_2 \subseteq V$, such that $x \in V_1 \times V_2$ and, as suggested by Eq. (B11),

$$V_1 \subseteq \mathbb{R}^{r_{\max}}, \quad \text{and} \quad V_2 \subseteq \mathbb{R}^{\dim(\mathrm{SU}(4)^{\times R}) - r_{\max}}. \tag{B12}$$

(Again to simplify notation, we are equating the local sets $V_{j=1,2}$ with their images, under local charts, in $\mathbb{R}^m$, for $m \in \mathbb{Z}_{>0}$.) From now on, $V_1 \times V_2$ is the open subset of interest. The contraction map's equivalence to a projector, in $V_1 \times V_2$, will help us compare high-depth circuits with low-depth circuits: Consider a circuit whose contraction map takes some local neighborhood to an image of some dimension. How does the dimension differ between high-depth circuits and low-depth circuits? We start by upper-bounding the dimension for low-depth circuits.

We have been discussing an $R$-gate architecture $A_T$. Consider any smaller architecture $A'$ of $R' < R$ gates. $A'$ is encoded in a contraction map $F^{A'}$ whose domain is $\mathrm{SU}(4)^{\times R'}$. As explained in the proof of Lemma 3, $F^{A'}$ is a polynomial map. Therefore, $F^{A'}$ has a property prescribed by the Tarski-Seidenberg principle [45] (Theorem 2): The image $F^{A'}\big(\mathrm{SU}(4)^{\times R'}\big)$ is a semialgebraic set of dimension $\leq \dim\big(\mathrm{SU}(4)^{\times R'}\big) = R' \dim\big(\mathrm{SU}(4)\big) = 15R'$.

We can strengthen this bound: Consider contracting two gates that share a qubit. The shared qubit may undergo a one-qubit gate specified by three parameters (one parameter per one-qubit Pauli). The one-qubit gate can serve as part of the first two-qubit gate or as part of the second two-qubit gate; which does not affect the contraction. Hence the contraction contains 3 fewer parameters than expected. In other words, the contraction has a redundant copy of the gauge group $\mathrm{SU}(2)$: Every unitary $U \in \mathrm{SU}(4)$ decomposes as $(U_1 \otimes U_2) K (V_1 \otimes V_2)$, wherein $K = e^{\mathrm{i}(aZ \otimes Z + bY \otimes Y + cX \otimes X)}$ and the $U_j$ and the $V_j$ denote single-qubit unitaries [52]. Let us classify the shared qubit as an input of the second two-qubit gate. A two-qubit gate in a circuit's bulk accepts two input qubits outputted by earlier gates. So we might expect an $R'$-gate circuit to have $\dim\big(F^{A'}(\mathrm{SU}(4)^{\times R'})\big) \leq 15R' - 2 \times 3R' = 9R'$. But the first $n/2$ gates [the leftmost vertical line of gates in Fig. 2(a)] receive

their input qubits from no earlier gates. So we must restore $3 \times 2$ parameters for each of the $n/2$ initial gates, or restore $3n$ parameters total:

$$\dim\left(F^{A'}(\mathrm{SU}(4)^{\times R'})\right) \leq 9R' + 3n. \tag{B13}$$

We have upper-bounded the dimension for low-depth circuits.

Technically, this bound on the dimension does not follow from Lemma 2 as it does for the bound $\dim F^{A'}(\mathrm{SU}(4)^{\times R'}) \leq 15R'$. The reason is, the quotient space $\mathrm{SU}(4)^{\times R'}/SU(2)^{\times(2R'-n)}$ is not necessarily semialgebraic. This difficulty can be resolved via Sard's theorem [49], which asserts, as a special case, that the Hausdorff dimension of a smooth map's image is bounded by its domain's dimension. A semialgebraic set's dimension is the greatest dimension in its stratification and so agrees with the Hausdorff dimension.

We now lower-bound the corresponding dimension for high-depth circuits. We can do so by lower-bounding the greatest possible rank, $r_{\max}$, of a high-depth architecture's contraction map, $F^{A_T}$: In an open neighborhood of $x \in \mathrm{SU}(4)^{\times R}$, $F^{A_T}$ is equivalent to a projector, which has some rank. The neighborhood's image, under the projector, is a manifold. The manifold's dimension equals the projector's rank. Therefore, we bound the rank to bound the dimension.

Augmenting an architecture with $T$ ($\leq L$)-gate blocks increases the contraction map's greatest possible rank, $r_{\max}$, by $\geq T - 1$. Therefore, for an architecture-$A_T$ circuit of $R \leq TL$ gates, we have constructed a point of rank $T \geq R/L$. Therefore,

$$r_{\max} \geq R/L. \tag{B14}$$

We have lower-bounded the dimension of the image of a high-depth architecture's contraction map [the rank in Ineq. (B14)] and have upper-bounded the analogous dimension for a low-depth architecture [Ineq. (B13)]. The high-depth-architecture dimension upper-bounds the low-depth-architecture dimension,

$$\dim\left(F^{A'}(\mathrm{SU}(4)^{\times R'})\right) < r_{\max}, \tag{B15}$$

if

$$9R' + 3n < r_{\max}, \tag{B16}$$

by Ineq. (B13). Furthermore, by Ineqs. (B13) and (B14), Ineq. (B15) holds if $9R' + 3n < R/L$, or $R' < \frac{R}{9L} - \frac{n}{3}$.

$$R' < \frac{R}{9L} - \frac{n}{3}. \tag{B17}$$

holds. We have upper-bounded the short circuit's gate count in terms of the deep circuit's gate count.

Let us show that, if Ineq. (B17) holds, the points in $\mathrm{SU}(4)^{\times R}$ resulting in unitaries that can be prepared with short circuits form a set of measure 0. We will begin with a point $x \in E_{r_{\max}}$; apply the short-architecture contraction map $F^{A'}$; and follow with the deep-architecture contraction map's inverse, $(F^{A_T})^{-1}$. The result takes up little space in $\mathrm{SU}(4)^{\times R}$, we will see.

To make this argument rigorous, we recall the small open neighborhood $V_1 \times V_2$ of $x \in E_{r_{\max}}$. In $V_1 \times V_2$, $F^{A'}(\mathrm{SU}(4)^{\times R'})$ has the preimage, under $F^{A_T}$, of

$$\left(F^{A_T}|_{V_1 \times V_2}\right)^{-1}\left(F^{A'}(\mathrm{SU}(4)^{\times R'})\right) \simeq \left[F^{A'}(\mathrm{SU}(4)^{\times R'}) \cap V_1\right] \times V_2. \tag{B18}$$

The $\simeq$ represents our identification of the map $F^A$ with its representation in local charts. By the proof of Lemma 3, $F^{A'}(\mathrm{SU}(4)^{\times R'})$ is a semialgebraic set. Therefore, by Theorem 3, $F^{A'}(\mathrm{SU}(4)^{\times R'})$ is a union of smooth manifolds. Each manifold is of dimension $\leq 9R' + 3n$, by Theorem 3 and Ineq. (B13). By Eq. (B12), $V_2$ is of dimension $\dim(\mathrm{SU}(4)^{\times R}) - r_{\max}$. Therefore, $[F^{A'}(\mathrm{SU}(4)^{\times R'}) \cap V_1] \times V_2$ consists of manifolds of dimension $\leq 9R' + 3n + \dim(\mathrm{SU}(4)^{\times R}) - r_{\max}$. Using Ineq. (B16), we can cancel the $9R' + 3n$ with the $-r_{\max}$, at the cost of loosening the bound: $[F^{A'}(\mathrm{SU}(4)^{\times R'}) \cap V_1] \times V_2$ consists of manifolds of dimension $< \dim(\mathrm{SU}(4)^{\times R})$. As a collection of manifolds of submaximal dimension, the points in $SU(4)^{\times R}$ that contract to unitaries implementable by short circuits (satisfying (B17)), restricted to a small open neighborhood $V_1$, form a set of measure 0.

Let us extend this conclusion about $n$-qubit unitaries—about images of maps $F^{A'}$—to a conclusion about preimages—about lists of gates. By Lemma 3, $E_{r_{\max}}$ is of measure 1. Therefore, for every $\varepsilon > 0$, there exists a compact subset $K \subseteq E_{r_{\max}}$ of measure $1 - \varepsilon$. Since $K$ is compact, for any cover of $K$ by open subsets, a finite subcover exists. The foregoing paragraph shows that, restricted to each open set in this finite subcover, the preimage of the unitaries reached by lower-depth circuits is of measure 0. Therefore, the preimage of the $R'$-gate, architecture-$A'$ circuits is of measure $\leq \varepsilon$. Since $\varepsilon > 0$ is arbitrary, the preimage is of measure 0. The foregoing argument holds for each architecture $A'$ of $R'$ gates. Hence each preimage forms a set of measure 0. The total measure is subadditive. So the union of the preimages, over all architectures with $\leq R'$ gates, is of measure 0. We have proven the circuit-complexity claim posited in Theorem 1. The state-complexity claim follows from tweaks to the proof (Appendix C). □

### Appendix C: Proof of the linear growth of state complexity

At the end of Appendix B, we proven part of Theorem 1—that circuit complexity grows linearly with the number of gates. Here, we prove rest of the theorem—that state complexity grows linearly. We need only tweak the proof presented in Appendix B.

Consider instead of the contraction map $F^{A_T}$, the map that contracts a list of gates, forming an architecture-$A_T$ circuit, and applies the circuit to $|0^n\rangle$, to get

$$G^{A_T} : \mathrm{SU}(4)^{\times R} \to S^{2\times 2^n - 1} \subseteq \mathbb{C}^{2^n} . \tag{C1}$$

The argument works the same as in Appendix B, with one exception: The derivative $D_x G^{A_T}$ has an image that does not contain $4^n - 1$ nontrivial linearly independent Pauli operators. Rather, the image contains the computational basis $\{\mathrm{i}^\kappa |x\rangle\}_{x\in\{0,1\}^n, \kappa\in\{0,1\}}$ formed by applying tensor products of $Z, X$ and $Y$ to $|0^n\rangle$. (We denote the imaginary number $\sqrt{-1}$ by i.) The proof of Lemma 3 ports over without modification, as $G^{A_T}$ is a polynomial map between algebraic sets.

The proof of Lemma 4 changes slightly. We must prove the existence of a point $x \in \mathrm{SU}(4)^{\times R}$ at which $G^{A_T}$ has a rank at least linear in the circuit depth. The only difference in the proof is, we must choose the operators $Q_j$ inductively such that the states $(C_T C_{T-1} \ldots C_j) Q_j (C_{j-1} C_{j-2} \ldots C_1)|0^n\rangle$ are linearly independent. Such a choice is possible if $T < 2 \times 2^n - 1$, the number of real parameters in a pure $n$-qubit state vector.

### Appendix D: Randomized architectures

From Theorem 1 follows a bound on the complexity of a doubly random circuit: Not only the gates, but also the gates' positions, are drawn randomly. This model features in Ref. [22]. Our proof focuses on nearest-neighbor gates, but other models (such as all-to-all interactions) yield similar results.

**Corollary 1** (Randomized architectures)**.** *Consider drawing an $n$-qubit unitary $U$ according to the following probability distribution: Choose a qubit $j$ uniformly randomly. Apply a Haar-random two-qubit gate to qubits $j$ and $j+1$. Perform this process $R$ times. With high probability, the unitary implemented has a high complexity: For all $\alpha \in [0, 1)$,*

$$\Pr\left( \mathcal{C}_{\mathrm{u}}(U) \geq \alpha \frac{R}{9n(n-1)^2} - \frac{n}{3} \right) \geq 1 - \frac{1}{1-\alpha}(n-1)e^{-n} . \tag{D1}$$

*Proof.* The proof relies on the following strategy: We consider constructing blocks randomly to form a circuit. If the blocks contain enough gates, we show, many of the blocks contain backwards light cones. This result enables us to apply Theorem 1 to bound the circuit's complexity.

Consider drawing $L$ gates' positions uniformly randomly. For each gate, the probability of drawing position $(j, j+1)$ is $1/(n-1)$. The probability that no gates act at position $(j, j+1)$ is $(1 - 1/(n-1))^L$. Let us choose for each block to contain $L = n(n-1)^2$ gates. Define a binary random variable $I_j$ as follows: If one of the gates drawn during steps $(j-1)n(n-1), (j-1)n(n-1)+1, \ldots, jn(n-1)$ acts at $(j, j+1)$, then $I_j = 1$. Otherwise, $I_j = 0$. With high probability, gates act at all positions:

$$p := \Pr\left( \bigwedge_{j=1}^{n-1} (I_j = 1) \right) = \left( 1 - \left( 1 - \frac{1}{n-1} \right)^{n(n-1)} \right)^{n-1} \geq \left( 1 - e^{-n} \right)^{n-1} \geq 1 - (n-1)e^{-n}. \tag{D2}$$

We have invoked the inverse Bernoulli inequality and the Bernoulli inequality. We will use this inequality to characterize blocks that contain backwards light cones.

Consider drawing $T$ $L$-gate blocks randomly, as described in the corollary. Denote by $X$ the number of blocks in which at least one position is bereft of gates: For some $j$, $I_j = 0$. With high probability, $X$ is small: For all $a \in (0, T]$,

$$\Pr(X \geq a) \leq \frac{T(1-p)}{a} \leq T(n-1)\, e^{-L/n}/a \,, \tag{D3}$$

by Markov's inequality. Let us choose for the threshold to be $a = (1-\alpha)T$. With overwhelming probability, $\alpha T$ blocks satisfy $\bigwedge_j (I_j = 1)$ and so contain gates that act at all positions $(j, j+1)$ in increasing order. Therefore, these blocks contain a staircase architecture and so contain backwards light cones. Therefore, a slight variation on Theorem 1 governs the $\alpha T \times L = \alpha R$ gates that form the blocks. Strictly speaking, Theorem 1 governs only consecutive backwards-light-cone–containing blocks. In contrast, extra gates may separate the blocks here. However, the extra gates can only increase the contraction map's image.

Therefore, the additional $(1-\alpha)TL$ gates cannot decrease the accessible dimension $d_{A_T}$. Therefore, the bound from Theorem 1 holds. With probability $\geq 1 - \frac{1}{1-\alpha}(n-1)\, e^{-n}$ over the choice of architecture,

$$\mathcal{C}_{\mathrm{u}}(U) \geq \frac{R - (1-\alpha)R}{9n^2(n-1)} - \frac{n}{3}, \tag{D4}$$

with probability one over the choice of gates. This bound is equivalent to Ineq. (D1). □

### Appendix E: Linear growth of slightly robust circuit complexity

Corollary 2 extends Theorem 1 to accommodate errors in the target unitary's implementation. We prove Corollary 2 by drawing on the proof of Theorem 1 and reusing notation therein.

**Corollary 2** (Slightly robust circuit complexity)**.** *Let $U$ denote the $n$-qubit unitary implemented by any random quantum circuit in any architecture $A_T$ that satisfies the assumptions in Theorem 1. Let $U'$ denote the $n$-qubit unitary implemented by any circuit of $R' \leq R/(9L) - n/3$ gates. For every $\delta \in (0,1]$, there exists an $\varepsilon := \varepsilon(A_T, \delta) > 0$ such that the Frobenius distance $d_{\mathrm{F}}(U, U') \geq \varepsilon$, with probability $1-\delta$, unless $R/L > 4^n - 1$.*

*Proof of Corollary 2.* The proof of Theorem 1 can be modified to show that, for every $\delta > 0$, there exists an open set $B \subseteq \mathrm{SU}(2^n)$ that contains $F^{A'}(\mathrm{SU}(4)^{\times R'})$, such that the preimage $(F^{A_T})^{-1}(B)$ is small—of measure $\leq \delta$. The modification is as follows. For every $\delta' > 0$, there exists a measure-$(1-\delta')$ compact subset $K$ of $E_{r_{\max}}$. As $K$ is compact, there exists a finite cover of $K$ that has the following properties: $K$ is in the union $\cup_j V^j$ of subsets $V^j$. On the $V^j$, the contraction map $F^{A_T}$ is equivalent to a projector, up to a local diffeomorphism. As in the proof of Theorem 1, we can assume, without loss of generality, that $V^j = V_1^j \times V_2^j$. The $V_1^j$ and $V_2^j$ are defined analogously to the $V_1$ and $V_2$ in the proof of Theorem 1. For each $V^j$, there exists an open neighborhood $W^j$ of $F^{A'}(\mathrm{SU}(4)^{\times R'}) \cap V_1^j$ such that $W^j$ has an arbitrarily small measure $\delta_j'' > 0$. Therefore, $B := \cup_j W^j$ has a preimage of measure $\leq \delta' + \sum_j \delta_j'' = \delta$. Each of the summands, though positive, can be arbitrarily small.

The Frobenius norm induces a metric $d_{\mathrm{F}}$ on $\mathrm{SU}(4)^{\times R}$. In terms of $d_{\mathrm{F}}$, we define the function

$$d_{\mathrm{F}}\left(\,.\,,\, F^{A_T}(\mathrm{SU}(4)^{\times R}) \setminus B\right) : F^{A'}(\mathrm{SU}(4)^{\times R'}) \to \mathbb{R}_{\geq 0}. \tag{E1}$$

This function is continuous, and $F^{A'}(\mathrm{SU}(4)^{\times R'})$ is compact. Therefore, the function achieves its infimum at a point $x_{\min} \in F^{A'}(\mathrm{SU}(4)^{\times R'})$. Therefore, the minimal distance to $F^{A_T}(\mathrm{SU}(4)^{\times R}) \setminus B$ is $d_{\mathrm{F}}(x_{\min}, F(\mathrm{SU}(4)^{\times R}) \setminus B)$. Since $B$ is open, $F^{A}(\mathrm{SU}(4)^{\times R}) \setminus B$ is closed and so compact. By the same argument,

$$\varepsilon(A_T, \delta) \coloneqq d_{\mathrm{F}}\left(x_{\min}, F(\mathrm{SU}(4)^{\times R}) \setminus B\right) = \inf_{y \in F(\mathrm{SU}(4)^{\times R}) \setminus B} \{d_{\mathrm{F}}(x_{\min}, y)\} = d_{\mathrm{F}}(x_{\min}, y_{\min}) > 0. \tag{E2}$$

We have identified an $\varepsilon > 0$ that satisfies Corollary 2. □

### Appendix F: Notions of circuit complexity

As circuit complexity is a widely popular concept, there is a zoo of quantities that measure it. We prove our main theorem for the straightforward definition of exact circuit implementation—the most straightforward notion of a circuit complexity—and for a version of approximate circuit complexity (Corollary 2) with an uncontrollably small error. In this appendix, we briefly mention other notions of complexity, partially to review other notions and partially to place the main text's findings in a wider context. Let $U \in \mathrm{SU}(2^n)$ denote a unitary. Ref. [10] discusses notions of approximate circuit complexity.

**Definition 9** (Approximate circuit complexity)**.** *The approximate circuit complexity $\mathcal{C}_{\mathrm{u}}(U, \eta)$ is the least number of 2-local gates, arranged in any architecture, that implements $U$ up to an error $\eta > 0$ in operator norm $||.||$.*

This definition is similar in mindset to the above (slightly) robust definition of a circuit complexity. For every pair $U, U' \in \mathrm{SU}(2^n)$ of circuits, the Frobenius distance between them satisfies

$$\frac{1}{2^n} d_{\mathrm{F}}(U, U') \leq ||U - U'|| \leq d_{\mathrm{F}}(U, U'). \tag{F1}$$

A widely used proxy for quantum circuit complexity—one that is increasingly seen as a complexity measure in its own right—is Nielsen's geometric approach to circuit and state complexity [9, 10, 12]. This approach applies geometric reasoning to

circuit complexity and led to many intuitive insights, including Brown and Susskind's conjectures about the circuit complexity's behavior under random evolution. To connect to cost functions as considered in Nielsen's framework, consider 1-local and 2-local Hamiltonian terms $H_1, H_2, \ldots, H_m$ in the Lie algebra $\mathrm{su}(2^n)$ of traceless Hermitian matrices, normalized as $\|H_j\| = 1$ for $j = 1, 2, \ldots, m$. Consider generating a given unitary, by means of a control system, following Schrödinger's equation:

$$\frac{d}{dt}U(t) = -iH(t)U(t), \text{ wherein } H(t) = \sum_{j=1}^{m} h_j(t)H_j. \tag{F2}$$

The control function $[0, \tau] \to \mathbb{R}^m$ is defined as $t \mapsto \big(h_1(t), \ldots, h_m(t)\big)$ and satisfies $U(0) = \mathbb{1}$. That is, a quantum circuit results from time-dependent control. In practice, not all of $\mathbb{R}^m$ reflects meaningful control parameters; merely a control region $\mathcal{R} \subset \mathbb{R}^m$ does. With each parameterized curve is associated a cost function $c : \mathcal{R} \to \mathbb{R}$, so that the entire cost of a unitary $U \in \mathrm{SU}(2^n)$ becomes

$$C(U) := \inf_{T,\, t \mapsto H(t)} \int_0^\tau dt\; c\big(H(t)\big). \tag{F3}$$

We take the infimum over all time intervals $[0, \tau]$ and over all control functions $t \mapsto H(t)$ such that the control parameters are in $\mathcal{R}$ for all $t \in [0, \tau]$ and scuh that $U(\tau) = U$. Several cost functions are meaningful and have been discussed in the literature. A common choice is

$$c_p\big(H(t)\big) = \left(\sum_{j=1}^{m} h_j(t)^p\right)^{1/p}. \tag{F4}$$

In particular, $c_2$ gives rise to a sub-Riemannian metric. For the resulting cost $C_2(U)$, Ref. [11] establishes a connection between the approximate circuit complexity and the cost: Any bound on the approximate circuit complexity, with an approximation error bounded from below independently of the system size, immediately implies a lower bound on the cost.

**Theorem 4** (Approximate circuit complexity and cost [11])**.** *For every integer $n$, every $U \in \mathrm{SU}(2^n)$ and every $\eta > 0$,*

$$\mathcal{C}_{\mathrm{u}}(U, \eta) \leq c\, \frac{C_2(U)^3 n^6}{\eta^2}\,. \tag{F5}$$

The quantity on the right-hand side can, in turn, be upper-bounded: $C_2(U) \leq C_1(U)$. This $C_1$ has a simple interpretation in terms of a weighted gate complexity [13].

**Definition 10** (Weighted circuit complexities)**.** *Let $U \in \mathrm{SU}(2^n)$ denote a unitary. The weighted circuit complexity $\mathcal{C}_{\mathrm{w}}(U)$ equals the sum of the weights of 2-local gates, arranged in any architecture, that implement $U$, wherein each gate $U_j$ is weighted by its strength $W(U_j)$, defined through*

$$W(U) := \inf\left\{\|h\| : U = e^{ih}\right\}. \tag{F6}$$

The weighted circuit complexity $\mathcal{C}_{\mathrm{w}}(U)$ turns out to equal the cost $C_1(U)$ for any given unitary. We can grasp this result by Trotter-approximating the time-dependent parameterized curve in the definition of $C_1(U)$.

**Lemma 5** (Weighted circuit complexity and cost)**.** *If $n$ denotes an integer and $U \in \mathrm{SU}(2^n)$, then*

$$\mathcal{C}_{\mathrm{w}}(U) = C_1(U). \tag{F7}$$

Therefore, the weighted circuit complexity grows like the cost $C_1$. By implication, the circuit complexity's growth will be reflected by a notion of circuit complexity that weighs the quantum gates according to their strengths. Again, once the main text's approximate circuit complexity is established with an $n$-independent approximation error, one finds bounds on the weighted circuit complexity, as well.

The last important notion of circuit complexity that has arisen in the recent literature is that of Ref. [19]. Denote by $\mathcal{G}_a \subset \mathrm{SU}(2^{2n})$ the set of $2n$-qubit unitary circuits comprised of $\leq a$ elementary quantum gates, wherein the first $n$ qubits form the actual system and the next $n$ qubits form a memory. Let $\mathcal{M}_b$ denote the class of all two-outcome measurements, defined on $2n$ qubits, that require quantum circuits whose implementation requires $\leq b$ elementary quantum gates. Define

$$\beta(r, U) := \text{maximize } \left|\mathrm{tr}\left(M\left\{[U \otimes \mathbb{1}]|\phi\rangle\langle\phi|[U \otimes \mathbb{1}]^\dagger - [\mathbb{1}/2^n \otimes \mathrm{tr}_1(|\phi\rangle\langle\phi|)]\right\}\right)\right|, \tag{F8}$$

$$\text{subject to } M \in \mathcal{M}_b,\ |\phi\rangle = V|0^{2n}\rangle,\ V \in \mathcal{G}_a,\ r = a + b. \tag{F9}$$

In terms of this quantity, Ref. [19] defined strong unitary complexity.

**Definition 11** (Strong unitary complexity [19]). *Let $r \in \mathbb{R}$ and $\delta \in (0,1)$. A unitary $U \in SU(2^n)$ has strong unitary complexity $\leq r$ if*

$$\beta(r,U) \geq 1 - \frac{1}{2^{2n}} - \delta, \tag{F10}$$

*denoted by $\tilde{\mathcal{C}}(U,\delta) \geq r$.*

While seemingly technically involved, the definition is operational. The definition is also more stringent and demanding than more-traditional definitions of approximate circuit complexity. To concretize this statement, we denote the diamond norm by $\|.\|_\diamond$ [53].

**Lemma 6** (Implications of strong unitary complexity [19]). *Suppose that $U \in U(2^n)$ obeys $\tilde{\mathcal{C}}(U,\delta) \geq r+1$ for some $\delta \in (0,1)$, $r \in \mathbb{R}$, arbitrary measurement procedures that include the Bell measurement. Then*

$$\min_{\mathcal{C}_{\mathrm{u}}(V) \leq r} \frac{1}{2}\|\mathcal{U} - \mathcal{V}\|_\diamond > \sqrt{\delta}\,. \tag{F11}$$

*That is, it is impossible to accurately approximate $U$ with circuits $V$ of $< r$ elementary quantum gates.*

$\mathcal{U}$ and $\mathcal{V}$ denote the unitary quantum channels defined by $\mathcal{U}(\rho) = U\rho U^\dagger$ and $\mathcal{V}(\rho) = V\rho V^\dagger$. The diamond norm between them is

$$\begin{aligned}\frac{1}{2}\|\mathcal{U} - \mathcal{V}\|_\diamond &= \frac{1}{2}\sup_\rho \|(U\otimes\mathbb{1})\rho(U\otimes\mathbb{1})^\dagger - (V\otimes\mathbb{1})\rho(V\otimes\mathbb{1})^\dagger\|_1 \\ &\leq \frac{1}{2}\sup_\rho \|[(U-V)\otimes\mathbb{1}]\rho(U\otimes\mathbb{1})^\dagger\|_1 + \frac{1}{2}\sup_\rho \|(U\otimes\mathbb{1})\rho[(U-V)\otimes\mathbb{1}]^\dagger\|_1.\end{aligned} \tag{F12}$$

We have added and subtracted a term and have used the triangle inequality. Therefore,

$$\frac{1}{2}\|\mathcal{U} - \mathcal{V}\|_\diamond \leq \frac{1}{2}\|U - V\|_\infty \left(\sup_\rho \|\rho(U\otimes\mathbb{1})^\dagger\|_1 + \sup_\rho \|(V\otimes\mathbb{1})\rho\|_1\right) \leq \|U - V\|_\infty, \tag{F13}$$

as the operator norm is a weakly unitarily invariant norm. Therefore, $\tilde{\mathcal{C}}(U,\delta) \geq r+1$ implies that $\mathcal{C}(U,\delta) \geq r$. That is, the strong unitary complexity of Ref. [19] is tighter than approximate circuit complexity. A topic of future work will be the exploration of the growth of approximate notions of complexity with an approximation error independent of the system size.

## Appendix G: Figures for the methods section

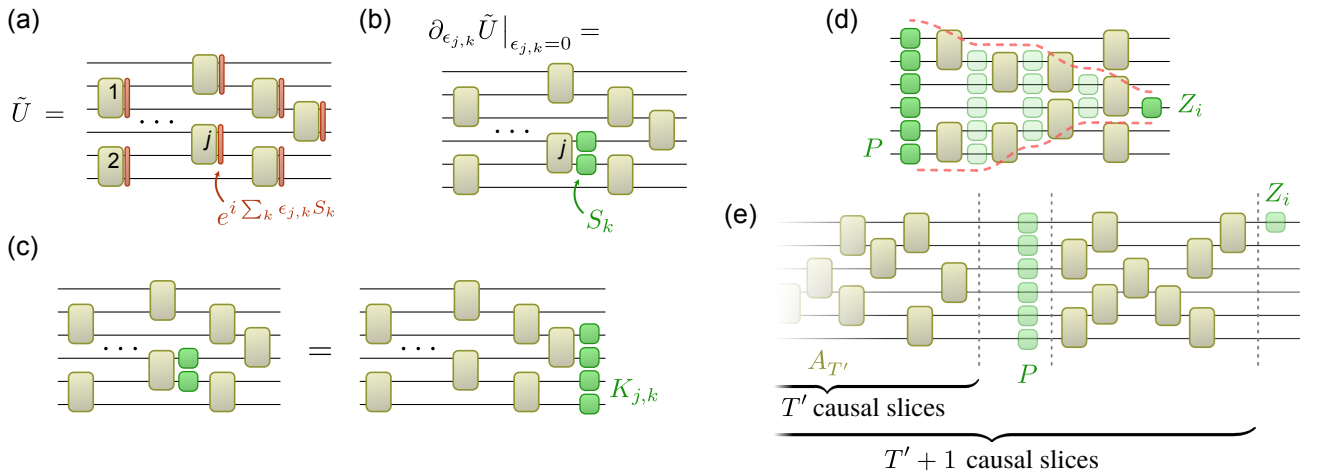


Figure 4. Our core technical result is a lower bound on the accessible dimension. **(a)** Each gate $U_j$ is perturbed with a unitary $e^{i\epsilon_{j,k}S_k}$, generated by a 2-qubit Pauli operator $S_k$ and parameterized with an infinitesimal $\epsilon_{j,k} \in \mathbb{R}$. Perturbing the gate perturbs the $n$-qubit unitary, turning $U$ into $\tilde{U} \approx U$. **(b)** A key quantity is the derivative of $\tilde{U}$ with respect to a parameter $\epsilon_{j,k}$, evaluated at $U$. Taking this derivative is equivalent to inserting the Pauli string $S_k$ immediately after the gate $U_j$. **(c)** The derivative depicted in panel (b) is equivalent to following the circuit with a Hermitian operator $K_{j,k}$ [Eq. (1)]. The operator $K_{j,k}$ results from conjugating $S_k$ with the gates after $U_j$. If the circuit consists of Clifford gates, then $K_{j,k}$ is a Pauli string, since Clifford gates map the Pauli strings to Pauli strings. Therefore, a perturbation of $U_j$ in the direction of $S_k$ results in a perturbation of the resulting unitary $U$ in the direction of $K_{j,k}$ in $\mathrm{SU}(2^n)$. **(d)** The following is true of every backwards-light-cone–containing block and every Pauli string $P$ (leftmost green squares): The block's gates can be chosen to be Cliffords that map $P$ to a single-site $Z$. The Clifford gates first map $P$ to a Pauli string that acts nontrivially on fewer qubits (pale green squares), then to a Pauli string on fewer qubits, and so on until the Pauli string dwindles to one qubit (rightmost green square). **(e)** Our lower bound is proven by recursion as explained in the main text.