

Probing Local Many-Body Dynamics with Random Quantum Circuits

by

Laura L. Cui

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Bachelor of Science in Physics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2023

© 2023 Laura L. Cui. All rights reserved.

The author hereby grants to MIT a nonexclusive, worldwide, irrevocable, royalty-free license to exercise any and all rights under copyright, including to reproduce, preserve, distribute and publicly display copies of the thesis, or release the thesis under an open-access license.

Author

Department of Physics

May 12, 2023

Certified by

Aram W. Harrow

Professor of Physics

Thesis Supervisor

Accepted by

Lindley Winslow

Associate Professor of Physics

Associate Head, Department of Physics

Probing Local Many-Body Dynamics with Random Quantum Circuits

by

Laura L. Cui

Submitted to the Department of Physics
on May 12, 2023, in partial fulfillment of the
requirements for the degree of
Bachelor of Science in Physics

Abstract

Random quantum circuits are an attractive model for the behavior of complex many-body physics, due to their analytic tractability as well as ability to reproduce the behavior of chaotic quantum systems. Recent progress in elucidating their structure has led to an improved understanding of quantum complexity, both in the context of quantum circuit complexity and state preparation, as well as for the task of measuring and distinguishing quantum states. However, previous results in the literature rely on the properties of specific theoretical models, or require unrealistic assumptions about the dynamics and experimental realization of large systems.

In this thesis we discuss the application of random circuits to probe local dynamics in a broad class of many-body systems, focusing on the regime in which the depth of the circuit is small compared to the size of the system. Motivated by recent results, we provide a definition for local scrambling based on the difficulty of distinguishing the resulting distribution from a Haar-random transformation in the case where only a region of fixed size may be accessed. We prove that up to the second moment, local scrambling of a product state input by a 1D random circuit occurs in log depth, i.e. requiring circuit depth asymptotically proportional to the logarithm of the size of the region, and is independent of the total system size.

In addition, we consider models for classifying topological phases and characterizing the entanglement structure of quantum matter. In particular, we describe the immediate application of our above result to bounds on the detection of these phases. We then discuss the topological entanglement entropy (TEE), a quantity related to the quantum conditional mutual information. Under standard assumptions, we prove that in the trivial phase, spurious contributions to the TEE decay in the limit as the size of the system goes to infinity, suggesting that the TEE is a robust indicator of topological order.

Thesis Supervisor: Aram W. Harrow
Title: Professor of Physics

Acknowledgments

This thesis is the product of various class and research experiences spanning my undergraduate studies, and would not have been possible without the support of many people. First, I would like to thank my thesis supervisor, Aram Harrow, for his guidance in this process. I took a quantum mechanics class with him in my very first semester at MIT, and I continue to appreciate his enlightening explanations, enthusiasm, and patience in teaching quantum physics and information theory.

I would also like to thank Daniel Ranard, Alex Dalzell, and Robert Huang, who contributed significantly to the results presented. Their insights and advice have been instrumental throughout developing this work, from framing the problem, to connecting it to existing literature, to brainstorming new ideas. Chapter 3 contains previously unpublished joint work on local scrambling [1] with Alex and Robert, along with new results including an analytic lower bound [2] on scrambling depth. Chapter 4 is joint work with Aram and Daniel. This work was funded by the NSF Quantum Leap Challenge Institutes program, including grant OMA-2016245.

I am grateful to have had the opportunity to learn from so many people, including John Preskill, whom I thank for enabling me to visit the IQIM at Caltech in 2021 and helping me to navigate both virtual and in-person research. In addition, I would like to thank Jake Taylor and Daniel Carney for encouraging me to pursue physics, and for introducing me to the field of quantum information all those years ago.

Finally, thank you to my parents, classmates, and many friends and mentors including Allison Borton, Grace Cai, Zawad Chowdhury, Matthew Cox, Jiahai Feng, Alex Gu, Matt Hodel, Serina Hu, Elliot Kienzle, Andrew Lin, Jacob Stavrianos, Ryan Tse, and others for serving as a endless source of motivation and inspiration throughout my time in college.

Contents

1	Introduction	13
1.1	Outline of thesis	15
2	Random quantum circuits	17
2.1	Motivation	17
2.2	Circuit architectures	18
2.3	Properties of the unitary group	19
2.3.1	Moments of the unitary group	19
2.3.2	Haar-random states	20
2.4	Design property	20
2.5	Unitary designs from statistical mechanics	22
2.6	Symmetric circuit models	24
2.7	Entanglement dynamics	25
2.7.1	Tensor network methods	26
3	Local information dynamics	29
3.1	Models of quantum complexity	30
3.1.1	Anti-concentration in random circuits	30
3.2	Local scrambling	31
3.2.1	Proof outline	33
3.3	Proof of upper bound for product state input	33
3.3.1	Spin lattice weights	33
3.3.2	Reduced trapezoidal circuit	35

3.3.3	Combinatorial argument	37
3.4	Proof of lower bound for product state input	42
3.4.1	Quantum Wasserstein distance of order 1	42
3.4.2	Circuit reduction for lower bound	43
3.5	Numerical results	45
3.6	Discussion	46
4	Classification of topological phases	49
4.1	Topological phases in random circuit models	50
4.1.1	Bounds from quantum complexity	51
4.2	Topological entanglement entropy	52
4.2.1	Stabilizer state case	54
4.2.2	Outline of trivial phase case	55
4.3	Proof of decay in trivial phase	56
4.3.1	Tensor network representation	56
4.3.2	Conditions for decay of quantum CMI	59
4.3.3	Bounds on forgetful component	60
4.3.4	Upper bound on decay in trivial phase	62
4.4	Numerical model	63
4.5	Discussion	66
A	Representations of the unitary group	69
A.1	Overview of representation theory	69
A.2	Schur-Weyl duality	71
A.3	Weingarten calculus	72
A.3.1	$U(1)$ gauge symmetry	74
B	Properties of the quantum W_1 norm	75
B.1	Classical Wasserstein distance	75
B.1.1	Foundations of optimal transport	75
B.1.2	Continuity bound for Shannon entropy	76

B.2	Quantum Wasserstein distance of order 1	78
B.2.1	Properties	80

List of Figures

2-1	Local one-dimensional random circuit architecture with $n = 10$, $L = 7$	18
2-2	Matrix product state representation of state on n qudits	27
3-1	Example of a domain wall configuration in the spin lattice model (right) which corresponds to one of the terms in the output (left), where each gate is replaced by a vertex and domain walls are drawn in between regions of identity and swap.	34
3-2	Trapezoidal circuit consisting of gates in the “lightcone” of the region A	35
3-3	Comparison of the natural logarithm of numerically computed values for the upper bounds given in Eq. 3.25, with two different values of q	46
4-1	Illustration of annulus construction for TEE	54
4-2	Illustration of decomposition of random circuit on coarse-grained sites of boundary region, where V_i is a polyhedral region of gates applied after the square prism U_i , with their projection onto the 2D surface on the left. The remaining polyhedral regions act only on the interior or exterior of R , and do not contribute to the entanglement entropy. On the right is the tensor network representation of the circuit, where \mathcal{U}_i and \mathcal{V}_i are the rotated tensors corresponding to $U_i(0\rangle^{\otimes 16d^2})$ and V_i , respectively.	56
4-3	Illustration of the Y-shaped transfer channel \mathcal{T}_i , in which we have traced out the E_i output legs to obtain a channel with one input and two outputs.	57
4-4	Illustration of annulus construction with four boundaries of B labeled	63

4-5	Density matrices of boundary state in original and noisy model	64
4-6	Distribution of γ computed exactly with von Neumann entropy for noisy chains with $l = 2, 3, 4, 5$ and $\varepsilon = 0.5$	65
4-7	Comparison of Rényi and von Neumann entropies for chains with $\epsilon = 0.5$	65

Chapter 1

Introduction

The relationship between physics, computation, and probability theory extends as far back as the Industrial Revolution, during which statistical theories of physics were first developed to explain the thermodynamics of steam engines. The central ideas of statistical mechanics were in turn cemented into modern computer science by the seminal works of Nyquist, Harley, and Shannon in the 1920s and 1940s which pioneered the study of classical information theory [3].

Quantum information theory offers new tools and perspectives for understanding both computational theory and physical phenomena in the regime where quantum effects become important. On one hand, quantum computation provides an alternative model of the Church-Turing thesis, and takes advantage of the mathematical properties of quantum mechanical hardware to deliver theoretical speedups over classical computers for certain tasks, most notably the unstructured query and integer factorization problems. On the other, the application of quantum information-theoretic approaches has prove fruitful for a number of applications in fundamental physics. These include characterizing out-of-equilibrium physics and the observation of uniquely quantum effects such as many-body localization and exotic phases of matter [4–7], as well as in resolving the apparent conflict between unitarity and classical predictions of general relativity in models of black hole physics [8–11].

One area of both theoretical and practical interest is in the efficient and accurate description of quantum mechanical systems, and understanding the conditions for

which an advantage is obtained over classical computers. Feynman first proposed quantum computation for this application [12], citing the fact that in general, q^n parameters are required to describe the state of a quantum system composed of n q -level particles, thereby utilizing computational resources which scale exponentially in the size of the system. The task of quantum simulation has wide-ranging applications, such as in the design of novel chemical structures and materials, random number generation, and potentially more general numerical optimization methods [13].

To this end, several recent experimental and theoretical advances have contributed to an improved understanding of quantum complexity, particularly in the case of local quantum dynamics, in which the interaction terms of the Hamiltonian are restricted to nearest-neighbor pairs. A key step in these achievements was the introduction of random quantum circuit models, which give ensembles of unitary gates sampled classically from some underlying distribution, typically uniformly random over all operations [14]. Random circuits are attractive as models for many-body physics due to their ability to reproduce universal properties of a broad class of quantum phenomena [15–18], and offer a powerful yet tractable framework for quantifying information complexity in quantum systems.

In particular, one direction which has been fruitful is relating quantum circuit complexity to the complexity of algorithms both for distinguishing quantum states or phases [19–21], as well as classically simulating quantum processes [22–24], for which it gives strong bounds on quantum advantages. Notably, it has been shown that for a local one-dimensional model, both exact circuit complexity and the difficulty of distinguishing the resulting distribution of states grow linearly in time [25–27].

However, several of these results implicitly require strong assumptions regarding access to measurements the entire system. On the other hand, prior to the spread of information throughout the entire system, chaotic quantum processes are expected to exhibit *local scrambling*, so that it is difficult to distinguish the resulting distribution of states when only a local subsystem is accessed. As far as we know, local scrambling has not been studied analytically in the existing literature, but is important for applications in condensed matter and black hole physics in which both access to the entire

system is impractical, and the relevant time scales are typically small in comparison to the total system size [9, 16, 17]. This phenomenon is also related to the measurement complexity and performance of classical shadow methods for learning operators on quantum states, in the general case where the operators are of intermediate rank and have support on some subregion of the system [28].

In addition, while classical shadow tomography provides a manner of efficiently reconstructing arbitrary properties of a many-body quantum state from the measurement of certain operators, it does not always provide information about which operators to choose or what types of properties may be measured in a generic system. For instance, while it is known that no linear operator can serve as a robust indicator for topological phases, which arise from patterns of long-range entanglement [29], in general the practicality of classifying topological order in physical systems is not well-understood.

The goal of our work is twofold: first, we address the fundamental issue of local scrambling by providing analytic bounds on the difficulty of distinguishing the output of a random circuit given a fixed subset of the system. Second, we build upon existing results which apply random circuit models to characterize topological phases of matter, in an attempt to provide guarantees on robust methods of distinguishing these phases.

1.1 Outline of thesis

In Chapter 2, we review the background on random quantum circuits, with a focus on the mathematical structure of Haar-random unitary circuits, and describe techniques such as the statistical mechanics mapping for computing the moments of random circuits, which we relate to measures of information scrambling and quantum complexity. In addition, we discuss tensor network methods for analyzing entanglement and information complexity.

In Chapter 3, we provide motivation for a definition of local scrambling in random circuits, based on the difficulty of distinguishing the distribution of the circuit

output when only a local subsystem of fixed size may be accessed. We relate this definition to a previous result on anti-concentration in random circuits [22] to provide analytic bounds for its convergence in linear 1D circuits. We show that up to the second moment, local scrambling of a product state input occurs in depth scaling logarithmically in the size of region accessed.

In Chapter 4, we provide an overview of the application of higher dimensional quantum circuits as a model for topological order, and previous proposals for the detection of topological phases. We then focus on the topological entanglement entropy (TEE), a quantity related to the quantum conditional mutual information, in two dimensions. Building upon recent progress in understanding this quantity, we prove that spurious contributions in the trivial phase, i.e. with product state input, decay roughly exponentially in the system size, so that the TEE is generically a robust indicator of topological order.

Chapter 2

Random quantum circuits

2.1 Motivation

We begin by discussing basic definitions and properties as well recent theoretical progress for random circuits, and provide intuition for these structures. Throughout this work, we consider circuits on n qudits of local dimension q , for which the system is represented by a Hilbert space \mathcal{H} of dimension q^n . Notably, many platforms for quantum computation as well as physical systems of interest are restricted to local interactions. We are interested in the simplest models which incorporate both locality as well as information about generic quantum dynamics, and therefore focus on ensembles of unitary circuits without additional structure. Two-qudit gates in this model are chosen at random from the unitary group $U(q^2)$.

Physically, these circuits can be interpreted as describing the possible trajectories of a system which experiences time-dependent interactions in the high-temperature limit, where they reproduce universal features such as the spectral statistics of a chaotic quantum system [18]. Computationally, these circuits represent the set of possible quantum protocols on a given platform, providing a method to bound the typical performance of these protocols. For instance, we note that quantum measurements are completely determined by the ensemble statistics. For a given circuit ensemble, the difficulty of distinguishing the distribution of output states from a uniformly random state is therefore bounded by the difference of the moments.

We first provide in Section 2.2 a description of the brickwork circuit in one and two dimensions, which is our main circuit architecture of interest. We then review well-known properties of the unitary group in Section 2.3, including its Weingarten calculus which is used to compute the moments of the Haar measure. In Sections 2.4 and 2.5, we apply these properties to describe other useful ensembles, and summarize a statistical mechanics mapping method for computing integrals over local random circuits. We also include a note on models with conserved charges in Section 2.6. Finally, we offer an overview of tensor network methods for analyzing entanglement dynamics in quantum circuits in Section 2.7.

2.2 Circuit architectures

In our analysis we typically restrict to the one- and two- dimensional brickwork circuit architectures consisting of two-site gates on neighboring qudits, each of which are sampled uniformly over the Haar distribution. In the 1D case, we have a *linear* geometry in which each qudit interacts with two nearest neighbors, up to unspecified boundary conditions. On the other hand, in the 2D case we adopt a square lattice geometry, so that each qudit interacts with four nearest neighbors, two in each direction.

We are interested in studying the scaling relationship between the behavior of random circuits and their depth, which is given by the minimum number of non-overlapping layers in the circuit. We can therefore consider circuits with alternating layers of gates up to some depth L : Throughout most of this work, we assume the

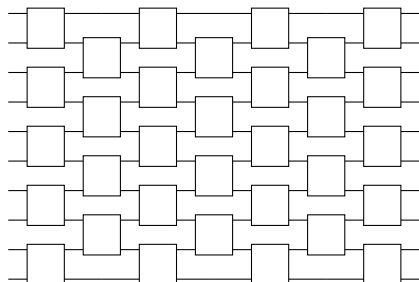


Figure 2-1: Local one-dimensional random circuit architecture with $n = 10$, $L = 7$

regime in which $L \ll n$. Thus whenever the accessible region of the system A is local, i.e. connected, and satisfies $|A| \ll n$, we are guaranteed that the “lightcone” of A is also much smaller than the rest of the system, and takes on a trapezoidal shape which does not overlap with itself.

2.3 Properties of the unitary group

2.3.1 Moments of the unitary group

In order to compute statistics over the unitary group, we must first describe the equal-volume integration measure. Formally, the Haar measure is the unique left- and right-invariant measure over $U(N)$, and therefore represents the uniform distribution over the group. The resulting ensemble is also known as the circular unitary ensemble (CUE). Using the symmetries of the unitary group, it can be shown that

$$\int_{\text{Haar}} dU U_{i_1 j_1} \dots U_{i_k j_k} U_{l_1 m_1}^\dagger \dots U_{l_k m_k}^\dagger = \sum_{\sigma, \tau \in S_k} \delta(\vec{i}, \sigma(\vec{m})) \delta(\vec{j}, \tau(\vec{l})) Wg(\sigma^{-1} \tau, d), \quad (2.1)$$

where we sum over permutation operators σ, τ . The Weingarten coefficients Wg are rational functions of q with degree at most $-k$, and can be computed exactly by applying the Schur-Weyl duality and combinatorially constructing the irreps of the symmetric group, then solving the resulting system of equations. For more detail and comments on generalizations to other groups, see Appendix A. We then define the moments of the unitary group as follows:

Definition 1 (Moments of the Haar distribution) *For any linear operator X acting on $\mathcal{H}^{\otimes k}$, the k -th moment of X over the Haar distribution can be defined via the twirling operators*

$$T_{\text{Haar}}^{(k)}(X) = \int_{\text{Haar}} dU U^{\otimes k} X (U^\dagger)^{\otimes k}. \quad (2.2)$$

As $T_{\text{Haar}}^{(k)}(X)$ commutes with any transformation on a single qudit, it must consist of a linear combination of permutation operators, which we interpret as acting on k copies

of the system. For instance, the second moment is a linear combination of identity ($\mathbb{1}$) and swap (S) operations on the two copies of the Hilbert space and is given by

$$\begin{aligned} T_{\text{Haar}}^{(2)}(X) &= \int_{\text{Haar}} dU U^{\otimes 2} X (U^\dagger)^{\otimes 2} \\ &= \frac{1}{d^2 - 1} \left(\mathbb{1} \text{Tr } X - \frac{1}{d} \mathbb{1} \text{Tr } SX + S \text{Tr } SX - \frac{1}{d} S \text{Tr } X \right). \end{aligned} \quad (2.3)$$

2.3.2 Haar-random states

It is possible to extend the above definitions to discuss distributions over states. In particular, the uniform distribution over all state vectors is equal to that obtained by applying a uniformly random unitary to a computational basis state. We can therefore define the moments of this distribution via the action of the twirling operator on the product state $T_{\text{Haar}}^{(k)}(|0\rangle\langle 0|^{\otimes k})$, which can be interpreted as an ensemble average over k copies of a given state, or of the output of k copies of a random circuit [30, 31]. We emphasize that while these two definitions are very similar, they are in general not equivalent, as it is simpler to specify the domain of independent transformations on k copies of a state than to consider arbitrary entangling transformations on $\mathcal{H}^{\otimes k}$. We provide further intuition for this fact and some of its consequences in Section 3.6.

2.4 Design property

It is often useful to consider ensembles of unitaries which capture some of the statistical properties of the Haar distribution:

Definition 2 (Unitary designs) *An ensemble $\varepsilon = \{p_i, U_i\}$ is called a unitary k -design if it reproduces the first k moments of the Haar distribution, so that*

$$T_\varepsilon^{(k')}(X) = \int_\varepsilon dU U^{\otimes k'} X (U^\dagger)^{\otimes k'} = T_{\text{Haar}}^{(k')}(X) \quad (2.4)$$

for all X acting on $\mathcal{H}^{\otimes k'}$ and $1 \leq k' \leq k$.

For instance, it is known that the Clifford group, which are the set of unitaries which

map Pauli operators to other Pauli operators, form a 2-design. In addition, it is shown in [32] that they form a 3-design in the case $q = 2^m$ for integer m , though they fail to form a 4-design for any $q \geq 2$.

In general it is difficult to find explicit constructions of k -designs for arbitrary k ; however, constructions have been given for approximate designs in connection to stochastic Hamiltonians and random circuits [21]. More precisely, we consider the following condition on ensemble averages:

Definition 3 (Approximate design property) *An ensemble $\varepsilon = \{p_i, U_i\}$ is called an ϵ -approximate k -design if the diamond norm of the difference of the k -th moment is bounded by*

$$\left\| T_\varepsilon^{(k)} - T_{\text{Haar}}^{(k)} \right\|_\diamond \leq \epsilon. \quad (2.5)$$

Recall the diamond norm of a channel is defined as

$$\|T\|_\diamond \equiv \sup_d \|T \otimes \mathbb{1}_d\|_{1 \rightarrow 1}. \quad (2.6)$$

where the superoperator norm $\|\cdot\|_{p \rightarrow p}$ for $p \geq 1$ is given in terms of a supremum over non-trivial operators

$$\|T\|_{p \rightarrow p} \equiv \sup_{O \neq 0} \frac{\|T(O)\|_p}{\|O\|_p} = \sup_{O \neq 0} \frac{(\text{Tr } |T(O)|^p)^{1/p}}{(\text{Tr } |O|^p)^{1/p}}. \quad (2.7)$$

Since integrating over the Haar measure erases the effect of any unitary applied on the system, this property can be interpreted as a condition for information scrambling in random circuits. From Eq. A.5, we observe that information contained in V is completely scrambled by a circuit which perfectly samples from the Haar distribution. Otherwise, information accessible via POVM measurements is bounded via the trace distance between the ensemble and the Haar moments. It can be shown via a statistical mechanics mapping that local random circuits acting on n qudits form approximate k -designs in $O(n \cdot \text{poly}(k))$ [26]. We reproduce the key ideas of this argument in the next section.

2.5 Unitary designs from statistical mechanics

Hunter-Jones [26] showed the unitary design condition for scrambling is achieved in linear time by applying a mapping from integrals over local random circuits to an Ising-like model. While the diamond norm is difficult to work with directly, it can be related to another quantity known as the frame potential. For any ensemble of unitaries ε , the k -th frame potential is given by

$$\mathcal{F}_\varepsilon^{(k)} = \int_\varepsilon dU dV |\text{Tr } U^\dagger V|^{2k}. \quad (2.8)$$

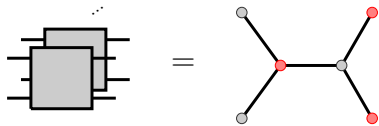
The frame potential for any ensemble is lower bounded by the Haar value, and furthermore the difference between the values bounds the diamond norm distance via

$$\left\| T_\varepsilon^{(k)} - T_{\text{Haar}}^{(k)} \right\|_\diamond \leq d^{2k} (\mathcal{F}_\varepsilon^{(k)} - \mathcal{F}_{\text{Haar}}^{(k)}). \quad (2.9)$$

When ε is a random quantum circuit of depth L , the frame potential can be written in terms of an integral over a circuit of depth $(2L-1)$ with periodic boundary conditions:

$$\mathcal{F}_{\mu_L}^{(k)} = \int d\mu_{2L-1} |\text{Tr } U|^{2k}, \quad (2.10)$$

where we denote by μ_L the ensemble generated by a local random circuit of depth L . This is a k -th moment quantity, and can be computed by expanding a product of Weingarten functions. In order to compute this expansion, we consider a mapping which sends the $2k$ copies of each gate to a pair of effective vertices with two incoming and outgoing legs, respectively.



The diagram shows a mapping between two representations. On the left, there is a stack of three gray rectangular boxes, each with four horizontal lines extending from its sides. An ellipsis (...) is placed above the stack, indicating it can be extended. On the right, there is a graph structure consisting of two vertices connected by a horizontal line. Each vertex has two incoming lines from the left and two outgoing lines to the right. The leftmost and rightmost vertices are gray, while the two inner vertices are red. The entire diagram is followed by an equals sign and the label (2.11).

We identify each vertex with a “spin” which is assigned one of the possible permutation operator terms $\sigma \in S_k$. Note that these configurations do not correspond to physical

instances of the circuit, and in general are not valid quantum states. We can visualize the result of this process as an Ising-like model on a triangular lattice where sites are given by the red vertices in Eq. 2.11, with domain walls separating regions of different spins. The frame potential can then be identified with the partition function of this model via

$$\mathcal{F}_{\mu L}^{(k)} = \sum_{\{\sigma\}} \prod_{\triangleleft \in \gamma} J_{\sigma_2 \sigma_3}^{\sigma_1}, \quad (2.12)$$

where for a plaquette with spins $\sigma_1, \sigma_2, \sigma_3$ on its three vertices, the $J_{\sigma_2 \sigma_3}^{\sigma_1}$ are computed by applying Eq. 2.1 and summing over the $\text{Tr } \tau_{j,j+1} \rho_{j,j+1}^{\otimes 2}$ coefficients, which yields

$$J_{\sigma_2 \sigma_3}^{\sigma_1} = \sum_{\tau \in S_k} \begin{array}{c} \sigma_2 \\ \nearrow \\ \sigma_1 \text{---} \tau \\ \searrow \\ \sigma_3 \end{array} = \sum_{\tau \in S_k} W g(\sigma_1^{-1} \tau, q^2) q^{\ell(\tau^{-1} \sigma_2)} q^{\ell(\tau^{-1} \sigma_3)}. \quad (2.13)$$

For $k = 2$ we have $\sigma_i, \tau \in \{\mathbb{1}, S\}$, so the coefficients for all non-zero plaquette configurations are given by

$$\begin{array}{ccc} \begin{array}{c} \mathbb{1} \\ \nearrow \\ \mathbb{1} \text{---} \mathbb{1} \\ \searrow \\ \mathbb{1} \end{array} = 1, & \begin{array}{c} \mathbb{1} \\ \nearrow \\ \mathbb{1} \text{---} S \\ \searrow \\ S \end{array} = \begin{array}{c} S \\ \nearrow \\ \mathbb{1} \text{---} \mathbb{1} \\ \searrow \\ \mathbb{1} \end{array} = \frac{q}{q^2 + 1} \\ \begin{array}{c} S \\ \nearrow \\ S \text{---} S \\ \searrow \\ S \end{array} = 1, & \begin{array}{c} S \\ \nearrow \\ S \text{---} S \\ \searrow \\ \mathbb{1} \end{array} = \begin{array}{c} \mathbb{1} \\ \nearrow \\ S \text{---} S \\ \searrow \\ S \end{array} = \frac{q}{q^2 + 1} \end{array} \quad (2.14)$$

Since plaquettes with $\sigma_2 = \sigma_3$ and $\sigma_1 \neq \sigma_2$ result in a coefficient of 0, only lattice configurations which do not have pairs of annihilating domain walls contribute to the second moment. Furthermore, a domain wall contributes a factor of $\frac{q}{q^2+1}$ for each layer it propagates, so that configurations with fewer domain walls have larger weights. Due to the periodic boundary condition, only configurations with the same domain wall locations at the leftmost and rightmost layers contribute, so the frame potential is upper bounded by a binomial expansion in the width of the circuit and thus decays asymptotically as

$$\mathcal{F}_{\mu_L}^{(2)} \approx 2 \left(1 + \frac{1}{q^{2L}} \right)^n. \quad (2.15)$$

Applying Eq. 2.9, it can then be shown that random circuits form an ϵ -approximate 2-design for depth $O(n + \log 1/\epsilon)$. For higher k , the analysis is more complicated, as some plaquettes can contribute a negative weight. However, in the limit of large local dimension, it can be shown that the depth required to form a k -design scales as $O(nk + \log 1/\epsilon)$.

The statistical mechanics mapping is a surprisingly powerful technique for analyzing random circuit dynamics, and can also be applied to study the behavior of out-of-time-order correlators [21], anti-concentration metrics [22], and even phase transitions in monitored quantum circuits [33].

2.6 Symmetric circuit models

While random circuit ensembles provide strong bounds on the complexity generic quantum dynamics and protocols, they are unphysical in the sense that they do not obey basic requirements such as energy conservation. Here we comment briefly on recent progress in understanding circuits with conserved quantities.

The presence of a conserved charge has profound effects on the structure of the circuit, including additional implications for the accessible state space. Requiring the entire circuit to commute with an Abelian or non-Abelian quantity restricts the ensemble of unitary gates to ones which are block diagonal in the different charge sectors. It is possible to apply hydrodynamics approaches to study the diffusion in the charge sectors and the resulting operator spreading, as in [34,35]. However, while it is known that in general a local random circuit of arbitrary depth can approximate any unitary transformation on the entire system [21,27], it is not the case that a local circuit which obeys a conservation law can approximate any unitary transformation on the entire system which does the same [36]. In particular, it was recently shown that non-Abelian symmetries restrict the circuit to the subgroup of symplectic or orthogonal transformations on each charge sector [37].

2.7 Entanglement dynamics

One of the hallmarks of quantum mechanical systems is entanglement, which arises from the Hilbert space structure and can be understood as a source of non-classical correlations. It is possible to quantify the amount of entanglement across any bipartition of a quantum system using the entanglement entropy, which we define in terms of the von Neumann entropy as follows: given a bipartite quantum state $|\psi\rangle_{A\bar{A}}$, the entanglement entropy of the state with respect to the partition A, \bar{A} is given by

$$S_A = -\text{Tr}(\rho_A \ln \rho_A), \quad (2.16)$$

where $\rho_A = \text{tr}_{\bar{A}} |\psi\rangle\langle\psi|$ is the reduced density matrix on A . In general, the von Neumann entropy can be interpreted as the smallest possible Shannon entropy which can be obtained for a projective measurement on a given state [38]. Note that when the state of the entire system is not pure, this quantity also includes classical correlations which do not correspond to entangled degrees of freedom.

Unfortunately, the von Neumann entropy is difficult to compute for an arbitrary state, as it requires us to diagonalize a density matrix with dimension exponential in $|A|$. It is possible to approximate this quantity with the Rényi entropies, which are more analytically and computationally tractable. Recall that for any state on system A , the α -th order Rényi entropy of ρ_A is given by

$$S_A^{(\alpha)} = \frac{1}{1-\alpha} \ln \text{Tr} \rho_A^\alpha. \quad (2.17)$$

The von Neumann entropy can be recovered in the limit as $\alpha \rightarrow 1$. These entropic quantities can be used to define a measure of both quantum and classic correlations between arbitrary subsystems whose union is not necessarily the entire system:

Definition 4 (Quantum mutual information) *For a general bipartite state ρ_{AB} , the mutual information $I(A : B)$ is given by*

$$I(A : B) = S_A + S_B - S_{AB}. \quad (2.18)$$

In addition, for a tripartite state ρ_{ABC} , we can define the conditional mutual information of the subsystems A and C , conditioned on B :

$$I(A : C|B) = S_{AB} + S_{BC} - S_B - S_{ABC}. \quad (2.19)$$

As in the classical case, the quantum mutual information and CMI are non-negative. Note that in the presence of unitary time evolution, all contributions to the von Neumann entropy are due to quantum entanglement. These quantities thus provide a measure of entanglement dynamics, describing the degree of non-classical behavior in various systems in terms of their statistical correlations.

2.7.1 Tensor network methods

Tensor networks provide a framework for the visualization and analysis of entanglement dynamics in systems with local interactions. These methods can be understood as an extension of classical techniques for solving isotropic spin lattice models under the relaxation of translation invariance [39], and enable efficient classical representations of many-body states which are generated by local dynamics.

To see this, recall that for any state $|\psi\rangle$ on a multipartite system with subsystem A , the Schmidt decomposition guarantees a representation of $|\psi\rangle$ of the form

$$|\psi\rangle = \sum_i a_i |i\rangle_A \otimes |i\rangle_{\bar{A}}, \quad (2.20)$$

where the sets $\{|i\rangle_A\}$ and $\{|i\rangle_{\bar{A}}\}$ form orthonormal bases on A and \bar{A} respectively, and $\sum |a_i|^2 = 1$. This is equivalent to the statement that the reduced states on A and \bar{A} have the same spectrum and entropy. For any 1D system with linear geometry and bounded bipartite Schmidt rank, it is possible to decompose the state via a matrix product state construction:

Definition 5 (Matrix product state) *Suppose there exists b such that for any*

$m \in [n]$, $|\psi\rangle$ can be written in the form

$$|\psi\rangle = \sum_{i=1}^b a_i |i\rangle_{[m]} |i\rangle_{[m+1,n]}. \quad (2.21)$$

Then we can write $|\psi\rangle$ in a matrix product state form as a product of tensors

$$|\psi\rangle = A_{i_1, q'_1}^{(1)} A_{i_1, i_2, q'_2}^{(2)} \dots A_{i_{n-1}, q'_n}^{(n)}, \quad (2.22)$$

where except for the first and last site each tensor is of dimension qb^2 . This corresponds to a chain of n tensors connected by bonds representing the contracted virtual indices, and dangling legs representing the physical indices as in Figure 2-2.

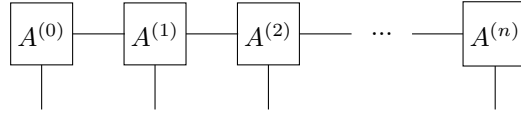


Figure 2-2: Matrix product state representation of state on n qudits

For example, it is known that all local 1D Hamiltonians which are gapped, i.e. admit a finite energy gap between the ground state and excited spectrum, have ground states with this property and can therefore be written as a matrix product state [40]. As this representation requires only $O(nq)$ parameters, these systems can be efficiently simulated on a classical computer.

While the higher-dimensional picture is more complicated, it is still possible to use a tensor network decomposition to analyze entanglement dynamics in any dimension. In particular, note that given a circuit on any input state which admits a matrix product state decomposition with tensors $A^{(m)}$, the output can also be expressed as a matrix product state with tensors $A'^{(m)}$ which are computed by factorizing the gates via the singular value decomposition and contracting with the qudit legs of the original state. A similar statement holds for density operators, which can be visualized as two copies of a matrix product state with a partial trace denoted by contraction of pairs of dangling legs. Notably, since the total bond dimension is upper bounded by the product of the bond dimensions of the corresponding contracted tensors, applying a

fixed-depth local circuit can increase the bond dimension by at most a constant factor. In addition, we remark that for a random circuit, the typical bond dimension is equal to the maximum, as there is zero probability of applying multiple transformations with the same eigenbasis.

Chapter 3

Local information dynamics

We now consider a restricted model of measurement, in which only a geometrically local subregion of fixed size may be accessed. Inspired by recent results relating complexity to the approximate design condition in random circuit models [21, 22], we define a similar condition for the achievement of local scrambling, focusing on the one-dimensional case. This corresponds to the generic information-theoretical task of distinguishing quantum states given access to some subsystem. While these conditions are not well-understood, they are directly relevant to a wide range of applications, including many systems of interest in condensed matter physics [41]. Our main results are Theorems 3 and 4, showing that in the regime where the sizes of the accessible region and circuit are both small compared to the size of the entire system, the distinguishability given a product state input decays exponentially with the depth of the circuit. This implies that, up to the second moment, the depth required to achieve local scrambling in a 1D system scales logarithmically in the size of the region accessed, and crucially is independent of the total system size.

We begin in Section 3.1 by motivating such a definition with a review of frameworks for analyzing quantum complexity, focusing specifically on state complexity. We then provide a formal definition of local scrambling in Section 3.2, and summarize our proof strategy. We provide full proofs of our upper and lower bounds in Sections 3.3 and 3.4, respectively.

3.1 Models of quantum complexity

In the absence of any global symmetries, 2-qudit gates chosen from any universal gate set are sufficient to reproduce any unitary transformation on the system [21, 36]. The minimum circuit size required to perform a given quantum operation, also known as quantum circuit complexity, is therefore a robust and natural measure of its implementation cost. The geometric approach for analyzing complexity is well-known [42–44], and provides a bound for the circuit cost of a given unitary transformation U by considering geodesics between U and the identity.

An alternative definition of quantum complexity based on distinguishable measurements is given in [21]. Informally, the measurement complexity of a quantum circuit U is defined as the minimum circuit size required to implement an ancilla-assisted measurement which can distinguish $\rho \mapsto U\rho U^\dagger$ from the completely depolarizing channel $\rho \mapsto \frac{1}{d}\mathbb{1}$. This description also invokes a definition of scrambling as the inability to distinguish information carried by a quantum state. It can be shown that this condition is in general an upper bound for the naive notion of complexity based on the circuit size required to approximate a quantum operation.

Random circuits are a natural model for the complexity of generic quantum systems and operations. More, precisely, recall that a random ensemble $\mathcal{E} = \{p_i, U_i\}$ is said to form an unitary k -design if it reproduces the first k moments of the Haar measure. It is shown in [21] that any ensemble which forms an approximate unitary k -design also has measurement complexity which scales as $O(k)$, and that random quantum circuits satisfy this property in depth scaling with $O(n \cdot \text{poly}(k))$. Furthermore, it is known that the exact circuit complexity of a sequence of Haar-random gates grows linearly with probability 1 before saturating after exponential depth [27], suggesting that random circuits are efficient scramblers of information.

3.1.1 Anti-concentration in random circuits

Dalzell, Hunter-Jones, and Brandão [22] studied the related property of anti-concentration in various local random circuit architectures, where an ensemble of states is defined

to be anti-concentrated if the expected collision probability is at most a constant factor times that for a uniform distribution. Here collision probability is defined as the probability that two identical copies of the circuit produce the same outcome when measured. For measurements with respect to the computational basis, it is given by

$$Z = q^n \mathbb{E}[p(|0\rangle)^2], \quad (3.1)$$

where $|0\rangle$ denotes the computational basis state in which all qudits are initialized to 0. The collision probability for the Haar-random distribution, which samples uniformly over all unitary transformations on the system, is given by

$$Z_H = \frac{2}{q^n + 1}. \quad (3.2)$$

Since the expected collision probability is a second moment quantity, satisfying the approximate 2-design property guarantees that the output of the circuit also satisfies the anti-concentration property; however, in general anti-concentration is weaker than the approximate 2-design condition. It is shown that for the one-dimensional ring as well as complete graph architectures, the output of a local random circuit reaches anti-concentration in log depth, faster than the circuit forms an approximate 2-design. In particular, for a one-dimensional random circuit with depth L , the collision probability is bounded by

$$Z_L = Z_H \left(1 + (e - 1)n \left(\frac{2q}{q^2 + 1} \right)^{L-1} \right). \quad (3.3)$$

We attempt to provide intuition for the difference in these two measures of scrambling and discuss further connections as well as potential for future work in Section 3.6.

3.2 Local scrambling

The observation that the k -th moment of an operator with respect to the Haar distribution must commute with and thereby absorb the effect of any unitary motivates

the definition of local scrambling via absorption of a unitary acting on a local region. More concretely, for a local random quantum circuit consisting of L layers, and A a region which is geometrically local. For a constant U_A , we are interested in the condition

$$\int d\mu_L U^{\otimes k} X(U^\dagger)^{\otimes k} \approx \int d\mu_L (U_A U)^{\otimes k} X(U^\dagger U_A^\dagger)^{\otimes k}. \quad (3.4)$$

Note that for the case $k = 1$, the twirling channel produces the completely mixed state, and thus information in the system is scrambled at constant depth $d = 1$.

We restrict to the simplest nontrivial case $k = 2$ on the 1D brickwork circuit, for which $T_{\mu_L}^{(2)}$ consists of terms with either the identity or swap operation on each qudit, and consider the action of the random circuit on an input state ρ . We therefore consider the following definition for local scrambling, up to the second moment:

Definition 6 (Local scrambling) *Consider a random circuit of depth L on n qudits with input ρ and a geometrically connected subsystem A . Let μ_L denote the ensemble consisting of instances of the random circuit. Then, up to some $\epsilon > 0$, the region A of ρ is scrambled by the circuit if for any unitary U_A which acts trivially on the subsystem \bar{A} ,*

$$\left\| \int d\mu_L U^{\otimes 2} \rho^{\otimes 2} (U^\dagger)^{\otimes 2} - \int d\mu_L (U_A U)^{\otimes 2} \rho^{\otimes 2} (U^\dagger U_A^\dagger)^{\otimes 2} \right\|_1 \leq \epsilon.$$

Recall that the trace norm or Schatten 1-norm of an operator A is defined via

$$\|A\|_1 := \text{Tr} |A| = \text{Tr} \sqrt{AA^\dagger}. \quad (3.5)$$

Intuitively, we expect that this conditions holds for $|A| \ll L$, and in particular that the maximum size of the region absorbed scales asymptotically with L , independent of the size of the entire system. We show in the following sections that the depth required is logarithmic in the size of $|A|$.

3.2.1 Proof outline

The key step in proving both the upper and lower bounds is reducing the problem to one which is asymptotically independent of n . This is relatively straightforward for the upper bound, where we use the fact that the trace norm contracts under the application of completely positive trace-preserving (CPTP) maps to reduce the random circuit into a trapezoidal region of size $|A| + 2L$. We then apply the triangle inequality to the quantity of interest in Definition 6 and bound the coefficients in terms of L , using a combinatorial argument reduce the number of terms to a sum which is asymptotically independent of n . From there, we adapt the lattice mapping technique in order to upper bound the coefficients of contributing terms.

The case of the lower bound is trickier, as we cannot directly eliminate regions outside of the lightcone of A since the appearance of additional domain walls in the circuit decreases the overall weight. We avoid this issue via an intermediate application of the quantum Wasserstein distance of order 1, trading a prefactor linear in $|A|$ in order to give a bound which is explicitly independent of the rest of the circuit, for a given U_A . We then use a similar combinatorial argument to give a lower bound in terms of the spin lattice weights for generic U_A .

3.3 Proof of upper bound for product state input

3.3.1 Spin lattice weights

We are interested in the action of a random circuit on an input state and in conditions for local scrambling which are asymptotically independent of system size. While the frame potential analysis is not directly useful for describing local scrambling due to the factor of $d^{2k} = q^{2kn}$, we can apply a similar lattice mapping technique as that described in Section 2.5 in order to bound the second moment, corresponding to accessing two copies of the input state.

From Eq. 2.3, integrating over a two-site unitary on qudits $j, j + 1$ which are not

entangled with the rest of the system yields

$$\frac{1}{q^4 - 1} \left(\mathbb{1} \text{Tr} \rho_{j,j+1}^{\otimes 2} - \frac{1}{q^2} \mathbb{1} \text{Tr} S \rho_{j,j+1}^{\otimes 2} + S \text{Tr} S \rho_{j,j+1}^{\otimes 2} - \frac{1}{q^2} S \text{Tr} \rho_{j,j+1}^{\otimes 2} \right) \otimes \rho_{\{i\}_1^n \setminus \{j,j+1\}}^{\otimes 2},$$

where $\rho_{\{i\}_1^n \setminus \{j,j+1\}}$ is the reduced density matrix on the rest of the system. Applying the statistical mechanics mapping technique in [26], we can represent these factors by replacing each of the gates with a “spin” which is assigned one of the possible permutation operators, in this case the $\mathbb{1}$ or S operators on the two copies of the system. Integrating over the entire circuit, each term in the full expansion is then represented by a configuration of domain walls separating regions with the same spin. See Figure 3-1 for an example of one of these configurations.

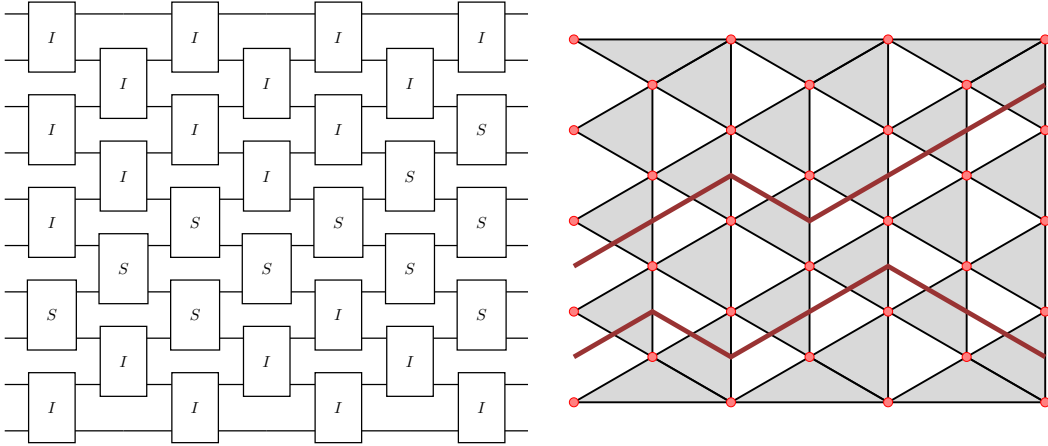


Figure 3-1: Example of a domain wall configuration in the spin lattice model (right) which corresponds to one of the terms in the output (left), where each gate is replaced by a vertex and domain walls are drawn in between regions of identity and swap.

The weighting for each term can therefore be computed via a partition function on the lattice. Let Γ denote all possible configurations γ . Then the second moment is given by

$$T_{\mu_L}^{(2)}(\rho^{\otimes 2}) = \int d\mu_L U^{\otimes 2} \rho^{\otimes 2} (U^\dagger)^{\otimes 2} = \sum_{\gamma \in \Gamma} W(\gamma) \gamma_L, \quad (3.6)$$

where γ_L denotes the configuration of the last layer. Each of the coefficients $W(\gamma)$

can be expressed in terms of the triangular plaquettes:

$$W(\gamma) = \prod_{\triangleleft \in \gamma} J_{\sigma_2 \sigma_3}^{\sigma_1}. \quad (3.7)$$

3.3.2 Reduced trapezoidal circuit

In the shallow circuit regime, we can simplify the quantity of interest in Definition 6 by considering the lightcone of U_A , which consists of the trapezoidal region of size $|A| + 2L$. We will assume for the remainder of this section that $|A| + 2L \leq n$, so that the lightcone is a trapezoidal region. It is clear that the rest of the circuit commutes with U_A ; to see this, consider the unitary $U_B = U^\dagger U_A U$ which is constructed by pushing U_A through the circuit. Since the architecture consists of alternating layers of two-site gates, for each layer U_A must be pushed through neighboring gates which act on two additional qudits, so that U_B acts nontrivially on at most $|A| + 2L$ qudits. Since the random circuit is an ensemble of CPTP channels and therefore itself CPTP, we can show that the original condition is upper bounded by its value for the trapezoidal circuit.

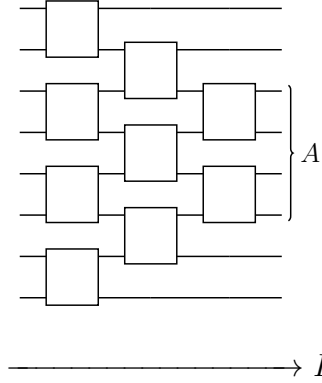


Figure 3-2: Trapezoidal circuit consisting of gates in the “lightcone” of the region A

Theorem 1 (Trapezoidal circuit reduction) *Suppose that $n \geq |A| + 2L$, and let μ_* denote the ensemble of unitaries generated by the trapezoidal region of gates in the lightcone of A . Then for any U_A which acts trivially on \overline{A} , we have*

$$\begin{aligned}
& \left\| \int d\mu_L U^{\otimes 2} \rho^{\otimes 2} (U^\dagger)^{\otimes 2} - \int d\mu_L (U_A U)^{\otimes 2} \rho^{\otimes 2} (U^\dagger U_A^\dagger)^{\otimes 2} \right\|_1 \\
& \leq \left\| \int d\mu_* U_*^{\otimes 2} \rho^{\otimes 2} (U_*^\dagger)^{\otimes 2} - \int d\mu_L (U_A U_*)^{\otimes 2} \rho^{\otimes 2} (U_*^\dagger U_A^\dagger)^{\otimes 2} \right\|_1.
\end{aligned}$$

Proof. Let μ_E denote the ensemble generated by the rest of the circuit which is not included in the trapezoidal region. Then for any $U \in \mu_L$, we can write $U = U_E U_*$ for some $U_E \in \mu_E$ and $U_* \in \mu_*$. We can then rewrite the left hand side in terms of the two parts of the circuit:

$$\begin{aligned}
& \left\| \int d\mu_L U^{\otimes 2} \rho^{\otimes 2} (U^\dagger)^{\otimes 2} - \int d\mu_L (U_A U)^{\otimes 2} \rho^{\otimes 2} (U^\dagger U_A^\dagger)^{\otimes 2} \right\|_1 \\
& = \left\| \int d\mu_* d\mu_E (U_E U_*)^{\otimes 2} \rho^{\otimes 2} (U_*^\dagger U_E^\dagger)^{\otimes 2} - \int d\mu_* d\mu_E (U_A U_E U_*)^{\otimes 2} \rho^{\otimes 2} (U_*^\dagger U_E^\dagger U_A^\dagger)^{\otimes 2} \right\|_1 \\
& = \left\| \int d\mu_E U_E^{\otimes 2} \left[\int d\mu_* U_*^{\otimes 2} \rho^{\otimes 2} (U_*^\dagger)^{\otimes 2} - \int d\mu_* (U_A U_*)^{\otimes 2} \rho^{\otimes 2} (U_*^\dagger U_A^\dagger)^{\otimes 2} \right] (U_E^\dagger)^{\otimes 2} \right\|_1 \\
& \leq \int d\mu_E \left\| U_E^{\otimes 2} \left[\int d\mu_* U_*^{\otimes 2} \rho^{\otimes 2} (U_*^\dagger)^{\otimes 2} - \int d\mu_* (U_A U_*)^{\otimes 2} \rho^{\otimes 2} (U_*^\dagger U_A^\dagger)^{\otimes 2} \right] (U_E^\dagger)^{\otimes 2} \right\|_1
\end{aligned}$$

Since the operator norm of a unitary U must satisfy $\|U\|_\infty = 1$, we can apply Hölder's inequality to obtain

$$\begin{aligned}
& \left\| \int d\mu_L U^{\otimes 2} \rho^{\otimes 2} (U^\dagger)^{\otimes 2} - \int d\mu_L (U_A U)^{\otimes 2} \rho^{\otimes 2} (U^\dagger U_A^\dagger)^{\otimes 2} \right\|_1 \\
& \leq \int d\mu_E \left\| \int d\mu_* U_*^{\otimes 2} \rho^{\otimes 2} (U_*^\dagger)^{\otimes 2} - \int d\mu_* (U_A U_*)^{\otimes 2} \rho^{\otimes 2} (U_*^\dagger U_A^\dagger)^{\otimes 2} \right\|_1 \\
& = \left\| \int d\mu_* U_*^{\otimes 2} \rho^{\otimes 2} (U_*^\dagger)^{\otimes 2} - \int d\mu_* (U_A U_*)^{\otimes 2} \rho^{\otimes 2} (U_*^\dagger U_A^\dagger)^{\otimes 2} \right\|_1. \tag{3.8}
\end{aligned}$$

We have thus reduced the definition to a condition explicitly independent of the total system size, as desired.

3.3.3 Combinatorial argument

We now consider the simplest case, in which we are given two copies of a product state as input. We observe that U_A must commute with terms that correspond to configurations whose last layer contain all identity or all swap in the region A . Our strategy is therefore to apply the triangle inequality to the right hand side of Eqs. 3.6 and 3.8, and bound the coefficients of the contributing configurations via a mapping between the set of configurations with a domain wall in A and the set of all configurations; we will then independently bound the sum of all configuration weights by adapting an anti-concentration result previously proved in [22].

We begin by establishing an upper bound for the quantity of interest in terms of the sum of all trapezoidal lattice configuration weights:

Theorem 2 (Correspondence counting bound) *Consider a trapezoidal circuit of depth L which forms the lightcone of a region A . Let Γ denote all trapezoidal lattice configurations, and $W(\gamma)$ denote the weight assigned to a given configuration γ . Then the right hand side of Eq. 3.8 can be bounded in terms of a sum over all configurations as follows:*

$$\left\| \int d\mu_* U_*^{\otimes 2} \rho^{\otimes 2} (U_*^\dagger)^{\otimes 2} - \int d\mu_* (U_A U_*)^{\otimes 2} \rho^{\otimes 2} (U_*^\dagger U_A^\dagger)^{\otimes 2} \right\|_1 \leq 2|A|q^{2|A|+4L} \left(\frac{2q}{q^2+1} \right)^{L-1} \sum_{\gamma \in \Gamma} W(\gamma).$$

Proof. Using the weight function, we can rewrite the quantity of interest as

$$\left\| \int d\mu_* U_*^{\otimes 2} \rho^{\otimes 2} (U_*^\dagger)^{\otimes 2} - \int d\mu_* (U_A U_*)^{\otimes 2} \rho^{\otimes 2} (U_*^\dagger U_A^\dagger)^{\otimes 2} \right\|_1 = \left\| \sum_{\gamma \in \Gamma} W(\gamma) \left[\gamma - U_A^{\otimes 2} \gamma (U_A^\dagger)^{\otimes 2} \right] \right\|_1.$$

Recall that for any U_A which acts trivially on \bar{A} and $\gamma = \sigma^{\otimes |A|} \otimes \gamma_{\bar{A}}$, we have $U_A^{\otimes 2} \gamma (U_A^\dagger)^{\otimes 2} = \gamma$. Then letting Γ^* denote trapezoidal lattice configurations with a domain wall in A , we can simplify this as

$$\begin{aligned}
& \left\| \sum_{\gamma \in \Gamma^*} W(\gamma) \left[\gamma - U_A^{\otimes 2} \gamma (U_A^\dagger)^{\otimes 2} \right] \right\|_1 \\
& \leq \sum_{\gamma \in \Gamma^*} W(\gamma) \left[\|\gamma\|_1 + \left\| U_A^{\otimes 2} \gamma (U_A^\dagger)^{\otimes 2} \right\|_1 \right] \\
& \leq 2 \sum_{\gamma \in \Gamma^*} W(\gamma) \|\gamma\|_1,
\end{aligned}$$

where we have again applied Hölder's inequality on U_A . Then since the single-qudit $\mathbb{1}$ and S each have q^2 eigenvalues with magnitude 1, we have

$$\left\| \int d\mu_* U_*^{\otimes 2} \rho^{\otimes 2} (U_*^\dagger)^{\otimes 2} - \int d\mu_* (U_A U_*)^{\otimes 2} \rho^{\otimes 2} (U_A^\dagger U_*^\dagger)^{\otimes 2} \right\|_1 \leq 2q^{2|A|+4L} \sum_{\gamma \in \Gamma^*} W(\gamma). \quad (3.9)$$

We now construct a mapping between $F : \Gamma^* \rightarrow \Gamma$ as follows: for any $\gamma_1 \in \Gamma^*$, we map it to the configuration $\gamma_0 \in \Gamma$ constructed by deleting the first domain wall that ends in A . Then consider the preimage $F^{-1} : \Gamma \rightarrow \Gamma^*$. Note that

$$\bigcup_{\gamma_0 \in \Gamma} F^{-1}(\gamma_0) = \Gamma^*, \quad (3.10)$$

so that all contributing configurations are counted by the union of preimages. We observe that for any $\gamma_0 \in \Gamma$, we can choose at most $|A|$ locations to add a domain wall, and at most two ways to propagate it backwards for each layer. Since domain walls are forbidden from annihilating in the direction of propagation, for any $\gamma_1 \in F^{-1}(\gamma_0)$, we must have

$$W(\gamma_1) = W(\gamma_0) \left(\frac{q}{q^2 + 1} \right)^{L-1}.$$

We can therefore bound the right hand side of Eq. 3.9 via

$$\begin{aligned}
2q^{2|A|+4L} \sum_{\gamma \in \Gamma^*} W(\gamma) & \leq 2q^{2|A|+4L} \sum_{\gamma_0 \in \Gamma} \sum_{\gamma \in F^{-1}(\gamma_0)} W(\gamma) \\
& \leq 2q^{2|A|+4L} \sum_{\gamma_0 \in \Gamma} |A| \left(\frac{2q}{q^2 + 1} \right)^{L-1} W(\gamma_0). \quad (3.11)
\end{aligned}$$

It is not immediately obvious that the right hand side of Eq. 3.11 should decrease in L ; however, since the second moment must approach that of the Haar distribution, it can be shown by applying the results of [22] that these weights are in fact exponentially small in the size of the lightcone.

Lemma 1 *For a trapezoidal circuit with depth L and terminal region A which is initialized to a product state input, the sum of lattice configuration weights is bounded by a constant times $1/q^{|A|+2L}(q^{|A|+2L} + 1)$, and satisfies*

$$\sum_{\gamma \in \Gamma} W(\gamma) \leq \frac{2z}{q^{|A|+2L}(q^{|A|+2L} + 1)} \left(1 + (e - 1)(|A| + 2L) \left(\frac{2q}{q^2 + 1} \right)^{L-1} \right) \exp \left[\frac{1}{1 - (2q/(q^2 + 1))^2} \right],$$

where we have defined $z := 1 + \left(\frac{q}{q^2 + 1} \right)^{L-1}$.

Proof. Consider the set Γ' which consists of all trapezoidal lattice configurations which contain only domain walls that exits through the side of the trapezoid, the set Γ_0 of all trapezoidal lattice configurations which do not contain a domain wall exiting through the side, as well as the set Γ_r which consists of all linear ring circuit configurations with dimension $|A| + 2L$ and depth L . For any trapezoidal configuration $\gamma \in \Gamma$, we can decompose γ into unique $\gamma' \in \Gamma'$ and $\gamma_0 \in \Gamma_0$, with $W(\gamma) = W(\gamma')W(\gamma_0)$. Thus we have that

$$\sum_{\gamma \in \Gamma} W(\gamma) \leq \sum_{\gamma' \in \Gamma'} W(\gamma') \sum_{\gamma_0 \in \Gamma_0} W(\gamma_0). \quad (3.12)$$

We first show that $\sum W(\gamma')$ is bounded even as L increases to infinity. Note that for each layer from $t = 1$ to L , we may have domain walls exiting from neither, one, or both sides of the trapezoid. Thus the sum of weights from exiting domain walls is upper bounded by

$$\begin{aligned}
\sum_{\gamma' \in \Gamma'} W(\gamma') &\leq \prod_{t=0}^{\infty} \left(1 + \left(\frac{2q}{q^2 + 1} \right)^t \right)^2 \\
&\leq \prod_{t=0}^{\infty} \exp \left[\left(\frac{2q}{q^2 + 1} \right)^{2t} \right] \\
&= \exp \left[\sum_{t=0}^{\infty} \left(\frac{2q}{q^2 + 1} \right)^{2t} \right] \\
&= \exp \left[\frac{1}{1 - (2q/(q^2 + 1))^2} \right]. \tag{3.13}
\end{aligned}$$

We now consider the relationship between $\sum W(\gamma_0)$ and $\sum W(\gamma_r)$. Note that due to closed boundary conditions, Γ_r contains only configurations with an even number of domain walls in each layer. Since domain wall parity is conserved, we can consider two separate cases; one in which $\gamma_0 \in \Gamma_0$ has an even number of domain walls, and one in which it has an odd number. We construct a mapping $F_r : \Gamma_0 \rightarrow \Gamma_r$ as follows: if $\gamma_0 \in \Gamma_0$ is even, it is mapped to the unique configuration $\gamma_r \in \Gamma_r$ with the rest of the circuit filled in such that no domain walls are added, and $W(\gamma_0) = W(\gamma_r)$. Otherwise, it is mapped to the unique configuration $\gamma_r \in \Gamma_r$ such that one domain wall is added along the upper side of the trapezoid. Thus each configuration $\gamma_r \in \Gamma_r$ is mapped to by at most one odd or one even configuration, and

$$\sum_{\gamma_0 \in \Gamma_0} W(\gamma_0) \leq \left[1 + \left(\frac{q}{q^2 + 1} \right)^{L-1} \right] \sum_{\gamma_r \in \Gamma_r} W(\gamma_r). \tag{3.14}$$

Finally, we can rewrite the norm of the output for the ring lattice and apply Eq. 3.3:

$$\begin{aligned}
&\frac{1}{q^{|A|+2L}} \left\| \int d\mu_r U_r^{\otimes 2} (|0\rangle\langle 0|)^{\otimes 2} (U_r^\dagger)^{\otimes 2} \right\|_1 \\
&= \frac{1}{q^{|A|+2L}} \text{Tr} \left| \int d\mu_r U_r^{\otimes 2} (|0\rangle\langle 0|)^{\otimes 2} (U_r^\dagger)^{\otimes 2} \right| \\
&= \frac{1}{q^{|A|+2L}} \sum_{i=0}^{q^{|A|+2L}-1} \int d\mu_r \langle i |^{\otimes 2} U_r^{\otimes 2} (|0\rangle\langle 0|)^{\otimes 2} (U_r^\dagger)^{\otimes 2} | i \rangle^{\otimes 2}
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{q^{|A|+2L}} \sum_{i=0}^{q^{|A|+2L}-1} E_{\mu_r}[p(|i\rangle)^2] = E_{\mu_r}[p(|0\rangle)^2] \\
&= \frac{1}{q^{|A|+2L}} \sum_{i=0}^{q^{|A|+2L}-1} \langle i|^{\otimes 2} \left(\sum_{\gamma_r \in \Gamma_r} W(\gamma_r) \gamma_r \right) |i\rangle^{\otimes 2} = \sum_{\gamma_r \in \Gamma_r} W(\gamma_r) \\
&\leq \frac{2}{q^{|A|+2L}(q^{|A|+2L} + 1)} \left(1 + (e-1)(|A| + 2L) \left(\frac{2q}{q^2 + 1} \right)^{L-1} \right). \tag{3.15}
\end{aligned}$$

Substituting into Eq. 3.12 yields the desired result.

We now apply the previous results to give an upper bound for local scrambling of a product state input by a random circuit.

Theorem 3 (Upper bound on product state scrambling) *Consider a linearly connected system which is initialized to a product state. In the regime in which the circuit depth and size of the region are small compared to the size of the entire system, a connected subsystem A is scrambled by a local random circuit in $O(\log |A|)$ depth.*

Proof. Substituting Theorems 1-2 and Lemma 1 into Definition 6, we have

$$\begin{aligned}
&\left\| \int d\mu_L U^{\otimes 2} (|0\rangle\langle 0|)^{\otimes 2} (U^\dagger)^{\otimes 2} - \int d\mu_L (U_A U)^{\otimes 2} (|0\rangle\langle 0|)^{\otimes 2} (U^\dagger U_A^\dagger)^{\otimes 2} \right\|_1 \\
&\leq \frac{4|A|szq^{2|A|+4L}}{q^{|A|+2L}(q^{|A|+2L} + 1)} \left(\frac{2q}{q^2 + 1} \right)^{L-1} \left(1 + (e-1)(|A| + 2L) \left(\frac{2q}{q^2 + 1} \right)^{L-1} \right) \\
&\leq 4|A|sz \left(\frac{2q}{q^2 + 1} \right)^{L-1} \left(1 + (e-1)(|A| + 2L) \left(\frac{2q}{q^2 + 1} \right)^{L-1} \right), \tag{3.16}
\end{aligned}$$

where for simplicity we have defined

$$z := 1 + \left(\frac{q}{q^2 + 1} \right)^{L-1}, \quad s := \exp \left[\frac{1}{1 - (2q/(q^2 + 1))^2} \right]. \tag{3.17}$$

Taking $q \geq 2$, for fixed $|A|$ and $L \gg 1/(\ln(q^2 + 1) - \ln 2q)$ the right hand side of Eq. 3.16 scales as

$$O \left(|A| \left(\frac{2q}{q^2 + 1} \right)^{L-1} \right).$$

Thus in the regime where the circuit is shallow relative to the size of the system, the depth required to locally scramble a product state input up to some $\epsilon > 0$ scales as $O(\log |A|/\epsilon)$.

3.4 Proof of lower bound for product state input

While it is relatively straightforward to show that the upper bound is independent of the total size of the system using Theorem 1, obtaining a tight lower bound is trickier. At the cost of a prefactor which linear in the size of accessible region, we provide a simple argument that the lower bound is independent of the total system size using techniques inspired by optimal transport theory.

3.4.1 Quantum Wasserstein distance of order 1

We first review some of the properties of the quantum W_1 norm proposed in [45], which is obtained as a generalization of the classical Hamming distance. Consider a system of n qudits with local dimension q , which has total state space \mathcal{H} . Let \mathcal{O}_n be the set of all Hermitian operators on \mathcal{H} , and $\mathcal{D}_n \subset \mathcal{O}_n$ denote the set of density matrices on \mathcal{H} . Then we define

Definition 7 (*Quantum W_1 norm*) For any Hermitian operator $X \in \mathcal{O}_n$, the quantum W_1 norm is defined as

$$\|X\|_{W_1} = \min\{t \geq 0 : X \in t\mathcal{B}_n\}.$$

Here \mathcal{B}_n is a closed, bounded, and convex subset of \mathcal{O}_n defined by

$$\mathcal{B}_n = \left\{ \sum_{i=1}^n p_i N_i : N_i \in \mathcal{N}_n^{(i)}, p \in \mathcal{M}_+([i]) \right\},$$

where \mathcal{N}_n is the set of all possible differences between two states that coincide after

tracing out one qudit:

$$\begin{aligned}\mathcal{N}_n^{(i)} &= \{\rho - \rho' : \rho, \rho' \in \mathcal{D}_n, \text{Tr}_i \rho = \text{Tr}_i \rho'\}, \\ \mathcal{N}_n &= \bigcup_{i=1}^n \mathcal{N}_n^{(i)}.\end{aligned}$$

It is clear that the above definition reduces to the classical Hamming distance when restricted to the set of computational basis states. In addition, we use the following properties of the W_1 norm, for which we provide proofs in Section B.2.1.

Lemma 2 (W_1 norm for local operators) *For any multi-index $I \subset [n]$ and $X \in \mathcal{O}_n$ such that $\text{Tr}_I X = 0$, we have that*

$$\|X\|_{W_1} \leq |I| \frac{q^2 - 1}{q^2} \|X\|_1. \quad (3.18)$$

Lemma 3 (Tensorization property of the W_1 norm) *The W_1 norm satisfies an additivity relation. For any $X \in \mathcal{O}_{m+n}$ with trace zero, we have that*

$$\|X\|_{W_1} \geq \|\text{Tr}_{1\dots m} X\|_{W_1} + \|\text{Tr}_{m+1\dots m+n} X\|_{W_1}. \quad (3.19)$$

Lemma 4 (Relation between W_1 and trace distance) *For any $X \in \mathcal{O}_n$ with trace zero, the W_1 distance is bounded via*

$$\frac{1}{2} \|X\|_1 \leq \|X\|_{W_1} \leq \frac{n}{2} \|X\|_1. \quad (3.20)$$

3.4.2 Circuit reduction for lower bound

Due to the symmetries of the random circuit, it is possible to convert to the W_1 norm at an intermediate step in order to bound the left hand side of Eq. 3.20 in terms of an expression independent of n . We then arrive at a circuit reduction similar to the case of the upper bound:

Theorem 4 (Lower bound on product state scrambling) *Suppose that $n \geq |A| + 2L$, and let μ_* denote the ensemble generated by the trapezoidal region of gates in the*

lightcone of A . Then for any unitary U_A which acts trivially on \overline{A} ,

$$\begin{aligned} & 2|A|\frac{q^2-1}{q^2}\left\|\int d\mu_L U^{\otimes 2}\rho^{\otimes 2}(U^\dagger)^{\otimes 2}-\int d\mu_L (U_A U)^{\otimes 2}\rho^{\otimes 2}(U^\dagger U_A^\dagger)^{\otimes 2}\right\|_1 \\ & \geq \frac{1}{2}\left\|\text{Tr}_{\overline{A}}\left(\int d\mu_* U_*^{\otimes 2}\rho^{\otimes 2}(U_*^\dagger)^{\otimes 2}-\int d\mu_* (U_A U_*)^{\otimes 2}\rho^{\otimes 2}(U_*^\dagger U_A^\dagger)^{\otimes 2}\right)\right\|_1. \end{aligned} \quad (3.21)$$

Proof. We define the operator

$$X = \int d\mu_L U^{\otimes 2}\rho^{\otimes 2}(U^\dagger)^{\otimes 2} - \int d\mu_L (U_A U)^{\otimes 2}\rho^{\otimes 2}(U^\dagger U_A^\dagger)^{\otimes 2}.$$

Note that X acts on $\mathcal{H} \otimes \mathcal{H} = \mathcal{H}^{\otimes 2}$. By linearity, we have that

$$\begin{aligned} \text{Tr}_A X &= \text{Tr}_A \left(\int d\mu_L U^{\otimes 2}\rho^{\otimes 2}(U^\dagger)^{\otimes 2} \right) - \text{Tr}_A \left(\int d\mu_L (U_A U)^{\otimes 2}\rho^{\otimes 2}(U^\dagger U_A^\dagger)^{\otimes 2} \right) \\ &= \text{Tr}_A \left(\int d\mu_L U^{\otimes 2}\rho^{\otimes 2}(U^\dagger)^{\otimes 2} \right) - \text{Tr}_A U_A^{\otimes 2} \left(\int d\mu_L U^{\otimes 2}\rho^{\otimes 2}(U^\dagger)^{\otimes 2} \right) (U_A^\dagger)^{\otimes 2} \\ &= \text{Tr}_A \left(\int d\mu_L U^{\otimes 2}\rho^{\otimes 2}(U^\dagger)^{\otimes 2} \right) - \text{Tr}_A \left(\int d\mu_L U^{\otimes 2}\rho^{\otimes 2}(U^\dagger)^{\otimes 2} \right) \\ &= 0. \end{aligned}$$

Then applying Eqs. 3.18, 3.19, and 3.20 in order yields a bound in terms of $\text{Tr}_{\overline{A}} X$:

$$2|A|\frac{q^2-1}{q^2}\|X\|_1 \geq \frac{1}{2}\|\text{Tr}_{\overline{A}} X\|_1. \quad (3.22)$$

Tracing out the rest of the system enables us to once again apply the trapezoidal circuit reduction from Section 3.3.2, as we have

$$\begin{aligned} \text{Tr}_{\overline{A}} X &= \text{Tr}_{\overline{A}} \int d\mu_E U_E^{\otimes 2} \left[\int d\mu_* U_*^{\otimes 2}\rho^{\otimes 2}(U_*^\dagger)^{\otimes 2} - \int d\mu_* (U_A U_*)^{\otimes 2}\rho^{\otimes 2}(U_*^\dagger U_A^\dagger)^{\otimes 2} \right] (U_E^\dagger)^{\otimes 2} \\ &= \text{Tr}_{\overline{A}} \left(\int d\mu_* U_*^{\otimes 2}\rho^{\otimes 2}(U_*^\dagger)^{\otimes 2} - \int d\mu_* (U_A U_*)^{\otimes 2}\rho^{\otimes 2}(U_*^\dagger U_A^\dagger)^{\otimes 2} \right). \end{aligned} \quad (3.23)$$

Plugging into Eq. 3.22 yields the desired result.

Note that in general, the left hand side of Eq. 3.21 depends on the choice of U_A ,

and for instance is zero when U_A is the identity. However, in the case in which U_A is nearly maximally entangling, which is true of most unitary matrices in the limit of large dimension [46], we expect a scaling of

$$\begin{aligned}\|\text{Tr}_{\bar{A}} X\|_1 &= \left\| \text{Tr}_{\bar{A}} \sum_{\gamma \in \Gamma^*} W(\gamma) \left[\gamma - U_A^{\otimes 2} \gamma (U_A^\dagger)^{\otimes 2} \right] \right\|_1 \\ &\approx g(q) \left\| \text{Tr}_{\bar{A}} \sum_{\gamma \in \Gamma^*} W(\gamma) \gamma \right\|_1 \\ &\gtrsim O(|A|e^{-cL}),\end{aligned}\tag{3.24}$$

where as in Section 3.3.3 we use Γ^* to denote trapezoidal lattice configurations with a domain wall in A , and $g(q)$ represents a polynomial function of q . Unfortunately, this relation is not tight, as we have “traded” a factor of $|A|$ when converting between the W_1 and 1-norm.

3.5 Numerical results

While we have not obtained tight lower bounds for the local scrambling depth, we can infer the strength of our upper bound and the true scaling of the quantity in Definition 6 from numerical simulations. We computed the exact quantities

$$B_0 = \left\| \int d\mu_L U^{\otimes 2} (|0\rangle\langle 0|)^{\otimes 2} (U^\dagger)^{\otimes 2} \right\|_1, \quad B_1 = q^{2|A|+4L} \sum_{\gamma \in \Gamma^*} W(\gamma) \tag{3.25}$$

for $q = 2, 4$ and $|A| = 4$, up to depth $L = 10$. Recall Γ^* is defined as the set of all configurations which have a domain wall in the region A and therefore contribute to the quantity in Definition 6, and W denotes the weights for a product state input. Note that here B_0 is computed for a ring architecture using $n = |A| + 2(t - 1)$ for the total size of the system at depth t ; however the mixed channel analysis still applies.

We observe that B_0 appears to decrease exponentially in the depth of the circuit, as expected. On the other hand, B_1 appears to decay more slowly and does not approach B_0 , suggesting that our analytic bound is fairly loose as it is an upper

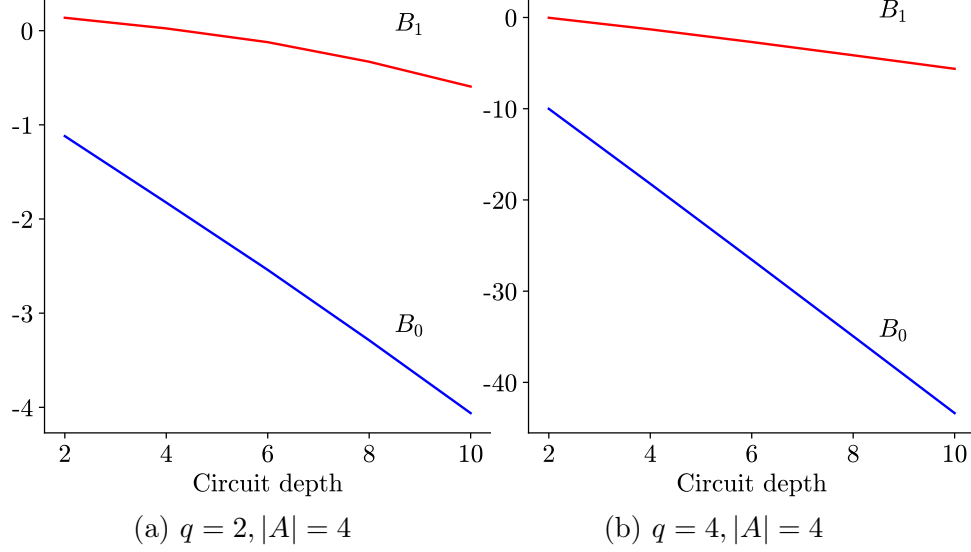


Figure 3-3: Comparison of the natural logarithm of numerically computed values for the upper bounds given in Eq. 3.25, with two different values of q

bound for B_1 . However, we recover the asymptotically exponential decay, suggesting that the log depth scaling is correct. See Figure 3-3 for a log plot of both bounds with circuit depth.

3.6 Discussion

We have shown that up to the second moment, a large system with no initial entanglement is locally scrambled by a random circuit in log depth, i.e. the depth required scales asymptotically with the logarithm of the size of the region of interest. Given analytical and numerical evidence from related work based on the statistical mechanics mapping, we conjecture that this result generalizes to more general states, as well as higher moments. We also expect that a similar relation holds for other architectures, such as 2D lattice geometries. We further note that in the regime in which the circuit is not shallow relative to the size of the system, it is not possible to separate the rest of the circuit, and the analysis becomes similar to that for the anti-concentration result in [22].

We emphasize that this notion of state scrambling should be distinguished from operator scrambling, which is more generally defined using the approximate design

condition and is expected to occur in linear depth in random circuits. One potential reason to expect the difference in scaling is the possibility of operators which entangle the two copies of the Hilbert space in the definition of a unitary design. We can consider the decomposition of any such operator into permutations on the copies of the Hilbert space, and unitary transformations on each of the copies. The same instance of a random circuit is applied to each copy, and thus cannot directly “undo” the entanglement between the different copies of the Hilbert space.

The formal definition of local scrambling presented in this paper can be applied to a variety of problems, offering potential directions for future work:

- Classical shadows are constructed using an ensemble of random unitary transformations, and can be used to efficiently reconstruct certain properties of a quantum state depending on the types of transformations performed [19]. In particular, the amount of long-range entanglement generated limits the measurement of long-range correlations in the system. Although strategies for the two limiting cases of measuring single-qudit and global properties are understood, recent work such as [28] has focused on interpolating between the two cases by applying finite-depth random circuits; importantly, the ensemble chosen to measure a local property only needs to locally scramble the information contained in the state. Although our results are not directly applicable to the rate of convergence of classical shadow tomography, which depends on a third-moment quantity, we conjecture that a similar circuit depth scaling may be achieved.
- It may be possible to extend techniques such as the statistical mechanics mapping to study scrambling in certain physical systems with symmetries or conserved quantities. In particular, the convergence of different coefficients can be analyzed for any subgroup of unitary operators which is equipped with a Weingarten calculus, such as those described in [47]. For the special case of a $U(1)$ gauge symmetry, the moments are identical and we obtain the same bounds for information scrambling, which we prove in Section A.3.1. On the

other hand, in the presence of a conservation law, the space of allowed unitary transformations can be decomposed into charge sectors. While it is known that in general actions on these charge sectors are described by the orthogonal and symplectic subgroups [37], it is not clear whether these models converge to the global moments within each charge sector.

In addition, state scrambling can be applied to study the properties of topological order. In particular, topological phases can be defined as the set of states which can be obtained by perturbing the system for a finite time, which is modeled by a shallow random quantum circuit. It is known that two topological phases cannot be distinguished by any single measurement on a state [29]. Our results suggest that when more than one copy of the state is provided, the ability to distinguish two phases is limited by the localization of the measurement. We discuss this application in more detail in Chapter 4.

Finally, although this work does not generalize directly to the diamond norm bound for channels, it may be possible to apply similar techniques to analyze a local form of scrambling on operators. In particular, it is expected that the “local” form of the approximate design property will be satisfied before the system is completely scrambled, requiring depth scaling linearly with the size of a local region.

Chapter 4

Classification of topological phases

In this chapter we turn to higher-dimensional systems, and consider random circuits as a model for the emergence of purely quantum phenomena. In particular, we focus on the application of random circuits to classify topological phases of matter, which arise from long-range patterns of entanglement in many-body systems. The application of circuit models to study quantum matter is motivated by fundamental as well as practical questions, including the simulation exotic phases of matter on quantum devices [48] as well as topological schemes for quantum error correction and fault-tolerant quantum computation [49]. Furthermore, the versatility of these models lends itself as a framework for analyzing a broad class of quantum systems.

We first provide an overview of topological phases in quantum circuit models, as well as different proposals for the detection and characterization of topological order, in Section 4.1. In addition, we describe bounds for the sensitivity of functions which depend on a finite order of moments in terms of the support of the function and the time scales under consideration. We then consider one proposal, the topological entanglement entropy (TEE), and provide intuition for the quantity in Section 4.2. Motivated by recent results on spurious contributions to the TEE, we show in Section 4.3 that it decays to zero for the trivial phase and is therefore in general a robust indicator of topological order. In addition, we include numerical evidence in Section 4.4 from simulations of a random spin chain model.

4.1 Topological phases in random circuit models

In condensed matter physics, topological order is a form of long-range order which does not correspond to a classical order parameter or explicit conserved symmetry. Macroscopically, it is conjectured that bosonic phases in two dimensions can be completely classified using the algebraic properties of their anyonic excitations, as well as the chiral central charge [50]. The power of these macroscopic descriptions for characterizing topological phenomena allows topological phases to be studied without an explicit reference Hamiltonian.

In addition, it is often useful to discuss the stability of physical properties and phases of matter [51, 52]. Topological phases are often associated with the phenomenon of topological ground state degeneracy, which is robust against local perturbations and can in principle be realized in zero-temperature phases of matter in the thermodynamic limit. It is the apparent stability of this structure and its support on a finite area in phase space which enables both non-trivial physical phenomena and quantum error correction protocols. These properties further motivate the application of shallow random circuits as a model for topological order. We thereby adopt the following definition:

Definition 8 (Topological phase) *Given a system of n qudits and corresponding Hilbert space \mathcal{H} , a topological phase is an equivalence class in \mathcal{H} defined by the region of state space which can be accessed by varying the Hamiltonian for a finite time, or equivalently by applying a constant-depth random circuit, for a given initial state.*

It is therefore possible to discuss the complexity of a topological phase in terms of the difficulty of preparing it. In particular, we define the trivial phase as follows:

Definition 9 (Trivial phase) *The trivial phase is the set of states which can be prepared by applying a constant-depth circuit on a product state input.*

States in the trivial phase are therefore characterized by a lack of long-range entanglement. We emphasize that these definitions and properties are implicitly dependent on the geometry of the system. Throughout this chapter we assume the 2D square

lattice architecture. To avoid confusion, we use d to denote the maximum number of gates applied between any nearest-neighbor pair.

Furthermore, we typically restrict to the case in which the input state exhibits translation invariance, which simplifies the analysis of topological orders and is relevant both for describing physical systems which are periodic in space [51], as well as for constructing scalable quantum error correcting codes. In two dimensions, it is known that any translation invariant stabilizer code with local generators and code distance linear in the system size can be decomposed into a finite number of copies of the toric code [53]. Below we discuss general bounds on the detection of these phases.

4.1.1 Bounds from quantum complexity

Given their importance in both fundamental physics and proposals for practical quantum computation, there is broad interest in developing methods for characterizing generic topological phases of matter. While in principle such phases are completely characterized by the microscopic behavior and ground state degeneracy of the parent Hamiltonian, it is often impractical to exactly solve such systems. We are therefore interested in quantities which are more straightforward to compute but serve as reliable indicators of topological order.

The definition of topological phases in terms of a random circuit model enables the application of complexity-theoretic techniques to consider potential measurement protocols. In particular, the expectation value of any analytic function over a topological phase can be written as a sum of moments over the circuit ensemble. For instance, it is known that topological phases cannot be distinguished by any linear function of the density matrix [29]. Formally, we have:

Lemma 5 (No-go theorem on measurement of topological phases) *Given generic system dynamics, there does not exist an observable O which is capable of distinguishing topological phases.*

Proof. We reproduce the proof in [29]. By definition, a topological phase is the set of all states which are generated by a random circuit of finite depth d . Denoting by

\mathcal{E}_0 the trivial phase and \mathcal{E}' as any other phase, we have that for any $d \geq 1$,

$$\begin{aligned}
\langle \rho \rangle_{\mathcal{E}'} &= \langle U | \psi \rangle \langle \psi | U^\dagger \rangle \\
&= \int d\mu_d U | \psi \rangle \langle \psi | U^\dagger \\
&= \frac{1}{q^n} \mathbb{1} \\
&= \langle \rho \rangle_{\mathcal{E}_0}
\end{aligned}$$

from a simple application of Eq. 2.2, where we have used the fact that the result must commute with all unitaries to express the integral as a multiple of the identity. We therefore have by linearity

$$\langle O \rangle_{\mathcal{E}'} = \langle O \rangle_{\mathcal{E}_0} = \frac{1}{q^n} \text{tr } O. \quad (4.1)$$

In addition, our result in Theorem 3 suggests that when more than one copy of the state is provided, the ability to distinguish two phases is limited by the localization of the measurement. Taking the converse of the scrambling statement, given fixed time scale d , the minimum required support of a distinguishing measurement on a single copy of the system is expected to scale exponentially with d . Furthermore, the approximate design property [21] suggests that in general the number of copies of the state required to make any distinguishing measurement is expected to scale linearly in d . Thus in general, access to a large number of qudits as well as exposure to higher-order moments improves robustness against perturbations.

4.2 Topological entanglement entropy

We now consider the topological entanglement entropy (TEE), a quantity first proposed in [54] as an indicator of topological order. Empirically, many systems of interest which are subject to local dynamics and exhibit a gapped spectrum are observed to exhibit area law scaling, in which the entanglement entropy of any region

R in the ground state scales with the size of the boundary. The presence of topological order removes additional degrees of freedom from the system, so that the total entanglement entropy is expected to be of the form

$$S = c[R] - \gamma + \dots \quad (4.2)$$

where $[R]$ is the area of the boundary, γ is a finite correction identified as the TEE, and the additional terms vanish in the limit as $|R| \rightarrow \infty$. In a topological quantum field theory with massive quasiparticle excitations, γ is given by

$$\gamma = \log D. \quad (4.3)$$

Here D is the total quantum dimension and is equal to

$$D = \sqrt{\sum_a d_a^2}, \quad (4.4)$$

where the sum runs over charge sectors of different quantum numbers, and d_a is the quantum dimension of a particle with charge a [54]. While this formula relies on knowledge about dynamics of the system, it is possible to approximate γ directly by explicitly computing the entropy of finite subregions and canceling out contributions from the boundaries. We focus on the annulus construction illustrated in Figure 4-1, for which the TEE is given in terms of the three subregions A, B, C via

$$2\gamma = S_{AB} + S_{BC} + S_C - S_{ABC} \quad (4.5)$$

$$= I(A : C|B). \quad (4.6)$$

This definition yields the same value as Eq. 4.2 in the limit as the size of the regions goes to infinity. The condition of positive TEE is therefore equivalent to the existence of non-trivial mutual information between two spatially separated regions, conditioned on the portion of the system in between. We make use of this definition in order to analyze the robustness of TEE as an indicator for the presence of topo-

logical order. While the TEE offers a succinct test for topological degrees of freedom,

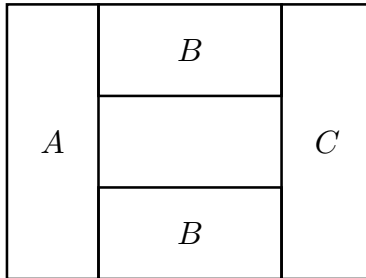


Figure 4-1: Illustration of annulus construction for TEE

it is not an invariant over topological phases, in the sense that it is known to differ for some pairs of states which are related by a constant-depth circuit [55, 56]. It is shown in [57] that these spurious contributions are always non-negative for states in same phase as a broad class of 2D reference states of interest. In addition, it is proved in [52] that when the state at the boundary is a stabilizer state, spurious contributions arise only when the state is in a nontrivial $G_1 \times G_2$ symmetry-protected topological (SPT) phase, though it is not known in general whether SPT phases are necessary for the presence of spurious TEE [58].

We show in Section 4.3 that spurious contributions in the trivial phase decay to zero on all but a fraction zero subset of the phase. To do so, we adapt a recent result in [59] which bounds the decay of CMI in matrix product density operators, and extend it to non-translation invariant states. We thus argue that the TEE is in general a robust indicator of topological order, as it is generically non-zero only on a non-trivial phase. Furthermore, we conjecture that this result extends more generally to spurious contributions in other topological phases.

4.2.1 Stabilizer state case

We first review a simpler problem, in which we restrict to stabilizer states which can be generated by a constant-depth Clifford circuit on a local 2D architecture. In this case, the entropy of any region A can be expressed in terms of a canonical set of

generators of the stabilizer group with respect to the region [60]:

$$S_A = n_A - |g_A|, \quad (4.7)$$

where n_A is the size of the region and $|g_A|$ is the rank of the subgroup produced when the stabilizer group is projected onto A , or equivalently the number of generators which are contained entirely within A . We therefore have for the annulus construction

$$\gamma = \frac{1}{2}|g_{NL}| = \frac{1}{2}(|g_{ABC}| + |g_B| - |g_{AB}| - |g_{BC}|), \quad (4.8)$$

so that γ is non-zero only when non-local structure is present and there are necessarily stabilizer generators supported on AC .

In the absence of any additional structure in the random Clifford circuit, we outline a heuristic argument for why $\frac{1}{2}|g_{NL}|$ should vanish in all but a vanishing subset of stabilizer states in the trivial phase as the size of the annulus region goes to infinity, similar to the result in [52]. First, for a product state input, the generators can always be written in a local form. Since each gate in the circuit is sampled from a finite set, for any lattice site s there is a finite probability p that all of the gates in the lightcone-area of s are the identity. As the size of the boundary of the region approaches infinity, with probability 1 there exist sites on each piece of the boundary so that the two ends are not entangled and the resulting state has no non-local generators.

4.2.2 Outline of trivial phase case

For the more general case of constant-depth Haar-random quantum circuits, the above argument fails as gates are sampled from a continuous space of transformations. Since the set of non-entangling gates is a lower-dimensional subspace of all possible unitary transformations, it comprises a fraction of size zero.

However, it is still possible to show that spurious contributions to the TEE decay to zero in the limit of large system size by analyzing the entanglement structure of the boundary. We build upon previous results in the literature which show that the

quantum CMI decays exponentially in certain matrix product density operators. In particular, by generalizing the transfer matrix of the boundary state to a Y-shaped channel, it is sufficient to show that the density operator can be constructed from a sequence of channels which have a “forgetful” component mapping any left input state to the same right output state with a fixed probability. Finally, we use typicality arguments about Haar-random unitaries in order to bound this probability.

4.3 Proof of decay in trivial phase

4.3.1 Tensor network representation

We focus on the $2D$ case of a square lattice with gates between each qudit and its four nearest neighbors. Consider the set of all states on a rectangular region R which can be obtained by acting on a product state $|0\rangle^{\otimes |R|}$ with a circuit of depth d . Since the entanglement entropy of R is not affected by a local change of basis, we need only analyze gates which affect the entanglement at the boundary. We can visualize

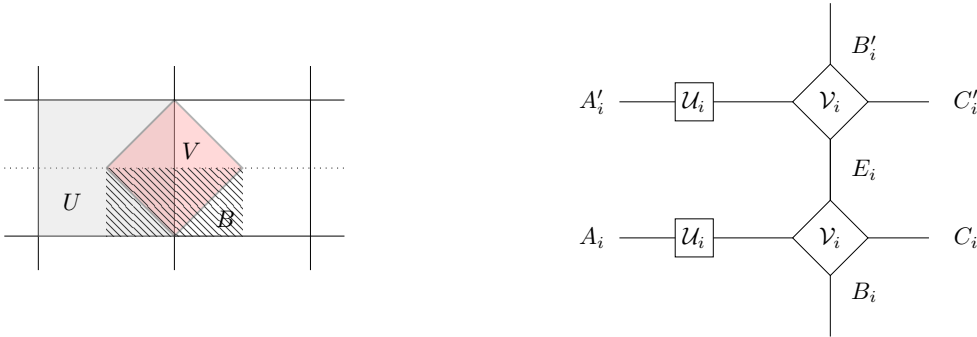


Figure 4-2: Illustration of decomposition of random circuit on coarse-grained sites of boundary region, where V_i is a polyhedral region of gates applied after the square prism U_i , with their projection onto the 2D surface on the left. The remaining polyhedral regions act only on the interior or exterior of R , and do not contribute to the entanglement entropy. On the right is the tensor network representation of the circuit, where \mathcal{U}_i and \mathcal{V}_i are the rotated tensors corresponding to $U_i(|0\rangle^{\otimes 16d^2})$ and V_i , respectively.

the ensemble of resulting states as a closed chain of length l which includes coarse-grained sites and gates representing polyhedral regions on the boundary as in Figure

4-2, where each of the new sites is a half-square of side length $4d$. Two horizontally neighboring sites are input into each U_i , which represents a square prism of base $4d \times 4d$. The right half of the output is denoted as A_i , and is input along with the left half of the next adjacent square, C_i , into the polyhedral region of gates V_i . Finally, the output of V_i is split vertically into the inside and outside half-squares B_i and E_i , representing the interior and exterior boundary state, respectively. The E_i legs are then traced out.

We adopt a transfer matrix approach to analyze this state, by considering the transfer channels, also called “Y-shaped” channels [59], \mathcal{T}_i which has the input leg A_i and output legs B_i and C_i :

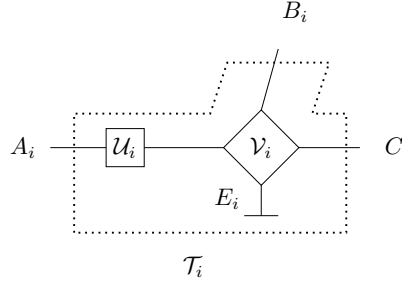


Figure 4-3: Illustration of the Y-shaped transfer channel \mathcal{T}_i , in which we have traced out the E_i output legs to obtain a channel with one input and two outputs.

Definition 10 (Transfer channel) *Given a chain of length l , let A_i denote the left legs associated with each site and B_i, C_i denote the right legs. Then define the transfer channel $\mathcal{T}_i : \mathcal{O}_{A_i} \rightarrow \mathcal{O}_{B_i} \otimes \mathcal{O}_{C_i}$ such that*

$$\rho_i = \rho_{A_1 B_1 \dots B_i C_i} = (\mathbb{1}_{A_1 B_1 \dots B_{i-1}} \otimes \mathcal{T}_i)(\rho_{i-1}). \quad (4.9)$$

We note that while the transfer matrices associated to each group of unitary transformations may not be explicitly CPTP, each of the \mathcal{T}_i must be in order to specify a valid state on any region of the boundary.

We can explicitly check that this holds for the product state input by writing out the resulting transformation:

Lemma 6 (CPTP property for trivial transfer channel) *Given product state input, the transfer channel associated to an instance of a random circuit with depth d on the boundary of a region is CPTP.*

Proof. First, we rewrite the effect of each polyhedral transformation in the transfer matrix fashion. Using the fact that $|\psi\rangle_{A_i} = U_i|0\rangle^{\otimes 16d^2}$ is the output of a unitary transformations within a square prism, we can write it in terms of the Schmidt decomposition

$$U_i|0\rangle^{\otimes 16d^2} = u_{i,a}|a_i a_i\rangle, \quad (4.10)$$

where $|\{|a_i\rangle\}| = q^{2d(d+1)}$ and $\sum |u_{i,a}|^2 = 1$. Note here we use the same label for the basis states on both the left and right half-squares which appear together in the Schmidt decomposition, though they do not necessarily refer to the same vectors in the q^{8d^2} -dimensional Hilbert space. Rotating the tensors gives

$$\mathcal{U}_i(|a_i\rangle\langle a'_i|) = u_{i,a}u_{i,a'}^*|a_i\rangle\langle a'_i|. \quad (4.11)$$

Similarly, we can consider a basis $\{|c_i\rangle\}$ for the immediately adjacent square to the right. We denote the effect of the unitary transformation V_i as

$$V_i|a_i c_i\rangle = \sum_{b_i} v_{i,a,c,b} |b_i b_i\rangle, \quad (4.12)$$

where we have that $|\{|b_i\rangle\}| = q^{4d^2}$ and $\sum_{b_i} |v_{i,a,c,b}|^2 = 1$. Rotating this matrix yields

$$\mathcal{V}_i(|a_i\rangle) = \sum_{c_i, b_i} v_{i,a,c,b} |c_i\rangle \otimes |b_i b_i\rangle = \sum_{c_i} |c_i\rangle \otimes V_i|a_i c_i\rangle, \quad (4.13)$$

which maps all unit norm vectors to a sum over unit norm vectors and is therefore manifestly completely positive. Furthermore, as the partial trace is always CPTP, summing over the output spectrum $u_{i+1,c}|c_i\rangle$ yields that the composition $\mathcal{T}_i = \text{Tr}_{E_i} \circ \mathcal{V}_i \circ \mathcal{U}_i$ is also CPTP, as desired.

4.3.2 Conditions for decay of quantum CMI

Definition 11 (Forgetful channel) *A transfer channel \mathcal{T} is forgetful if it has a forgetful component, i.e. there exists η with $0 < \eta \leq 1$ such that for any input ρ , we can write the action of \mathcal{T} as*

$$\mathcal{T}(\rho) = \eta \mathcal{F}(\rho) + (1 - \eta) \mathcal{T}'(\rho), \quad (4.14)$$

where \mathcal{T}' is also a transfer channel and \mathcal{F} is a channel which sends

$$\rho \mapsto \Phi(\rho) \otimes \omega, \quad (4.15)$$

so that any input state is mapped to the same state in one of the outputs.

Theorem 5 (CMI decay for forgetful channels [59]) *Given a chain of forgetful transfer channels $\mathcal{T}_i = \mathcal{T}$, where \mathcal{T} admits a decomposition $\eta \mathcal{F} + (1 - \eta) \mathcal{T}'$, we have*

$$I(A_1 : C_l | B_1 \dots B_l) = O((1 - \eta)^l). \quad (4.16)$$

Proof. We adapt the short proof from [59]. Let A denote A_1 and any additional part of the system to the left of A_1 , and B the boundary region between A_1 and C_i . Using the joint convexity of relative entropy and the fact that all channels are CPTP, it can be shown that

$$\begin{aligned} I(A : BC_i) &= D(\mathcal{T}(\rho_{i-1}) || \rho_A \otimes \mathcal{T}(\rho_{B_1 \dots B_{i-1} C_{i-1}})) \\ &\leq \eta D(\mathcal{F}(\rho_{i-1}) || \rho_A \otimes \mathcal{F}(\rho_{B_1 \dots B_{i-1} C_{i-1}})) + (1 - \eta) D(\mathcal{T}'(\rho_{i-1}) || \rho_A \otimes \mathcal{T}'(\rho_{B_1 \dots B_{i-1} C_{i-1}})) \\ &= \eta D(\rho_{AB} \otimes \omega || \rho_A \otimes \rho_B \otimes \omega) + (1 - \eta) D(\mathcal{T}'(\rho_{i-1}) || \rho_A \otimes \mathcal{T}'(\rho_{B_1 \dots B_{i-1} C_{i-1}})) \\ &\leq \eta D(\rho_{AB} || \rho_A \otimes \rho_B) + (1 - \eta) D(\rho_{i-1} || \rho_A \otimes \rho_{B_1 \dots B_{i-1} C_{i-1}}) \\ &= I(A : B) + (1 - \eta) (I(A : B_1 \dots B_{i-1} C_{i-1}) - I(A : B)). \end{aligned}$$

Moving $I(A : B)$ to the left hand side and applying the monotonicity of mutual

information yields

$$\begin{aligned}
I(A : C_i | B) &\leq (1 - \eta)(I(A : B_1 \dots B_{i-1} C_{i-1}) - I(A : B)) \\
&\leq (1 - \eta)(I(A : B_1 \dots B_{i-1} C_{i-1}) - I(A : B_1 \dots B_{i-1})) \\
&= (1 - \eta)I(A : C_{i-1} | B_1 \dots B_{i-1}).
\end{aligned}$$

Thus the CMI contracts by a factor of $(1 - \eta)$, giving the desired result.

We now extend this result to the case of non-translation invariant transfer channels, by considering a class of transfer channel chains such that each channel is independently η -forgetful with some probability p .

Corollary 1 (CMI decay for probably forgetful channels) *Given a chain of forgetful transfer channels \mathcal{T}_i such that for some η , with probability p each \mathcal{T}_i independently admits a decomposition $\eta\mathcal{F}_i + (1 - \eta)\mathcal{T}'_i$, we have*

$$I(A_1 : C_l | B_1 \dots B_l) \lesssim O((1 - \eta)^{lp}). \quad (4.17)$$

Proof. This follows from the contraction coefficient computed in Theorem 5 and the observation that we expect on average lp channels in the chain to be forgetful. Furthermore, we have that in the limit as $l \rightarrow \infty$, the probability that we have fewer than $l(p - \epsilon)$ forgetful channels is bounded via Chebyshev's inequality as

$$P \leq \frac{p(1 - p)}{l\epsilon^2} \quad (4.18)$$

for any $\epsilon < p$. Thus for large l we have vanishing probability of observing less than any constant fraction of lp forgetful channels, as well as resulting CMI with any constant factor larger than the contraction coefficient $(1 - \eta)^{lp}$ from Theorem 5.

4.3.3 Bounds on forgetful component

Corollary 2 (Trivial transfer channels are forgetful) *Given a chain of length l with input $|0\rangle^{\otimes 16d^2l}$ and $\rho_{ABC} = (\mathcal{T}_l \circ \dots \circ \mathcal{T}_1) \left((|0\rangle\langle 0|)^{\otimes 16d^2l} \right)$, for each i with $1 \leq i \leq l$*

there exists some $\eta_i > 0$ and decomposition of \mathcal{T}_i

$$\mathcal{T}_i = \eta_i \mathcal{F}_i + (1 - \eta_i) \mathcal{T}'_i, \quad (4.19)$$

where \mathcal{T}'_i is CPTP and \mathcal{F}_i is forgetful.

Proof. Using the decomposition given in the proof of Lemma 6, we have that

$$\mathcal{T}_i(|a_i\rangle\langle a_i|) = \text{Tr}_{E_i} \circ \mathcal{V}_i \circ \mathcal{U}_i(|a_i\rangle\langle a_i|) \quad (4.20)$$

$$= \sum_{c_i, c'_i} |c_i\rangle\langle c'_i| \otimes \text{Tr}_{E_i} \left(V_i |a_i c_i\rangle\langle a_i c'_i| V_i^\dagger \right) \quad (4.21)$$

$$= \sum_{c_i, c'_i} |c_i\rangle\langle c'_i| \otimes \sum_{e_i, b'_i} v_{i,a,c,e} v_{i,a,c',b'}^* |e\rangle\langle b'| \cdot \langle b'|e\rangle \quad (4.22)$$

$$= \sum_{c_i, c'_i} |c_i\rangle\langle c'_i| \otimes \sum_{e_i, b'_i} v_{i,a,c,e} v_{i,a,c',b'}^* |e\rangle\langle e^*| b'^* \rangle\langle b'|. \quad (4.23)$$

We can equivalently “factorize” this result to write

$$\mathcal{T}_i(|a_i\rangle) = \sum_{c_i} |c_i\rangle \otimes \sum_{b_i} v_{i,a,c,b} |b\rangle\langle b^*|. \quad (4.24)$$

Thus any unit norm vector for A is mapped to a superposition of the non-degenerate unit norm vectors in C tensored with some operator on B . Since we are guaranteed positivity by construction regardless of the values of $v_{i,a,c,b}$, we can find a forgetful decomposition for \mathcal{T}_i with $\eta = \min |v_{i,a,c,b}|^2$.

Furthermore, we can bound the values of these coefficients by considering the distribution of Schmidt coefficients for Haar-random bipartite states, as in [46]. Below, we provide values of p_0 and η_0 for the case $q = 2$.

Lemma 7 (Magnitude of forgetful component for $q = 2$) *Given a random circuit ensemble on qubits, and constant η_0 which satisfies*

$$\eta_0^{1/4d^2} \leq \frac{1}{2} - \sqrt{\frac{3}{20}}, \quad (4.25)$$

we are guaranteed that each \mathcal{T}_i is independently η_0 -forgetful with probability

$$p_0 = \left(1 - \frac{3}{20(1/2 - \eta_0^{1/4d^2})^2}\right)^{4d^2}. \quad (4.26)$$

Proof. Given a bipartite Haar-random state on two qubits, the distribution of the Schmidt coefficients λ_i satisfies [46]

$$E[\lambda] = \frac{1}{2}, \quad E[\lambda^2] = \frac{2}{5}.$$

Let $\lambda_m = \min \lambda_i$. We again apply Chebyshev's inequality to obtain

$$\begin{aligned} P[\lambda_m < \epsilon] &= P\left[\left|\lambda_i - \frac{1}{2}\right| \geq \frac{1}{2} - \epsilon\right] \\ &\leq \frac{3}{20(1/2 - \epsilon)^2}. \end{aligned}$$

Furthermore, for a series of Haar-random transformations, we have with mutually distinct Schmidt bases with probability 1 due to the continuous nature of the unitary group. For a generic instance of the circuit we thus have $\eta = \prod_U \lambda_m(U)$, so for a given η_0 we are guaranteed

$$\begin{aligned} p_0 &\geq \left(1 - P\left[\lambda_m < \eta_0^{1/|\{U\}|}\right]\right)^{|\{U\}|} \\ &= \left(1 - P\left[\lambda_m < \eta_0^{1/4d^2}\right]\right)^{4d^2} \\ &\geq \left(1 - \frac{3}{20(1/2 - \eta_0^{1/4d^2})^2}\right)^{4d^2}. \end{aligned}$$

4.3.4 Upper bound on decay in trivial phase

Theorem 6 (Decay of TEE in trivial phase) *Given a rectangular annulus region R with boundary dimension $l \gg 1$ and initial state $|0\rangle^{|R|}$, the spurious TEE decays as*

$$\gamma \lesssim O((1 - \eta_0)^{lp_0^4}), \quad (4.27)$$

where $p_0, \eta_0 > 0$ are functions of q and d .

Proof. We abstract away the geometry of the problem by combining the four total interior and exterior boundaries of region B of the annulus into a single chain of transfer channels with length l :

$$\mathcal{T}_i^* = \mathcal{T}_{\text{ext},i}^{(1)} \otimes \mathcal{T}_{\text{int},i}^{(1)} \otimes \mathcal{T}_{\text{ext},i}^{(2)} \otimes \mathcal{T}_{\text{int},i}^{(2)}.$$

The result follows from substituting Corollary 2 and the arguments in Lemma 7 into Corollary 1 in order to obtain the decay of long-range CMI in the trivial phase under the action of the \mathcal{T}_i^* . Note we are guaranteed probability p_0^4 that for some i all four of the transfer channels on each boundary are forgetful. Finally we apply Eq. 4.6 which yields that the TEE of the total state is proportional to that of the CMI, and therefore decays exponentially in l as desired.

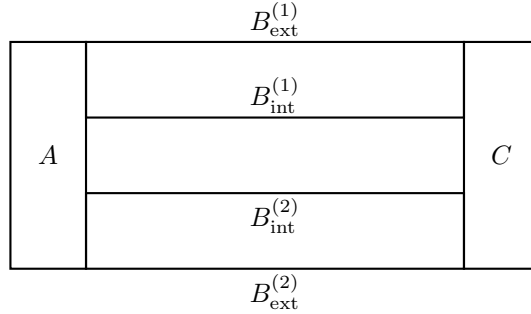


Figure 4-4: Illustration of annulus construction with four boundaries of B labeled

4.4 Numerical model

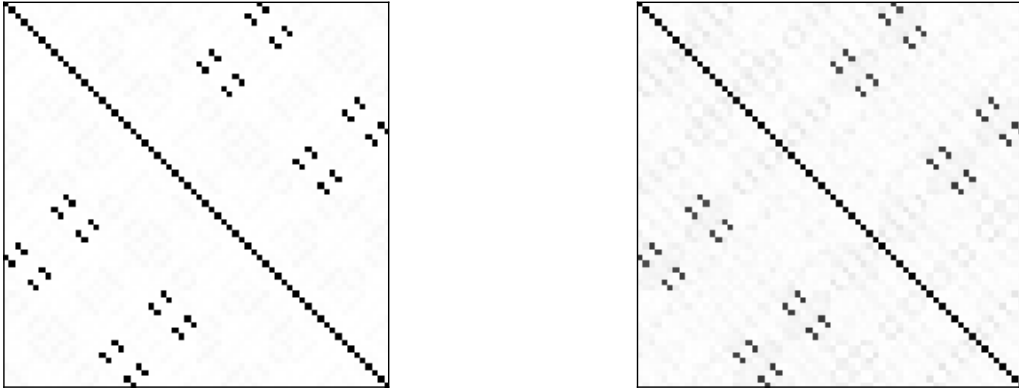
In addition to our analytic bound for the case $q = 2$, we provide numerical evidence for the decay of spurious TEE in the trivial phase through the simulation of a random spin chain model. Our numerical model is inspired by [52] and reproduces some of the properties of the coarse-grained tensor network representation of the boundary state described in Section 4.3.1. In particular, the state on the random spin chain sites is given by replacing $U|00\rangle$ with maximally entangled states

$$\Omega_{q'} = \frac{1}{\sqrt{q'}} \sum_{a=1}^{q'} |aa\rangle, \quad (4.28)$$

and applying the unitary transformation to neighboring left and right sites corresponding to the Stinespring dilation of the channel

$$\mathcal{W}_i = \frac{1}{(q')^2} \sum_{a=1}^{q'} (P_a \otimes P_a W_i) \cdot (P_a \otimes W_i^\dagger P_a), \quad (4.29)$$

where the P_a are generalized Pauli operators and W_i is a random unitary. Random noise proportional to $1 + i\epsilon H$ for some $\epsilon > 0$ and Hermitian H drawn from the Gaussian unitary ensemble is then applied at each site. The code used to produce the following results can be found at [61]. The resulting state has exact spurious value $\gamma = 1$ for all $l \geq 2$ when the noise is set to zero. The TEE is observed to decay for finite ϵ . We provide an example of the density matrix for $l = 3$ in both the original and noisy model in Figure 4-5. In addition, we investigate the distribution of γ in noisy



(a) Random spin chain model

(b) Noisy model with $1 + i\epsilon H$, $\epsilon = 0.5$

Figure 4-5: Density matrices of boundary state in original and noisy model

chains of different lengths using both the von Neumann and Rényi entropies, and find that for small numerical examples the spurious contributions decay with l with the distribution shown in Figure 4-6. Furthermore, we observe that in general the

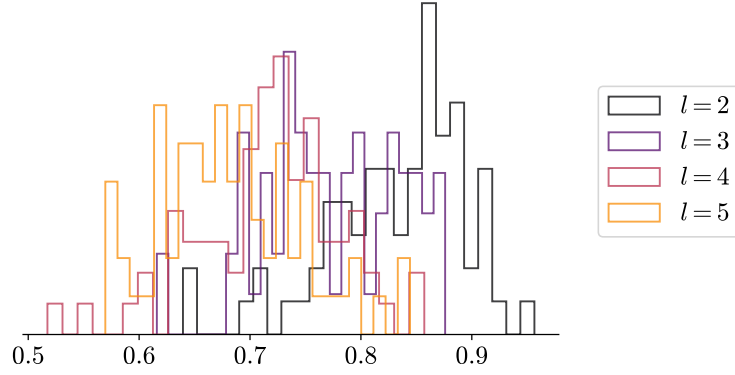
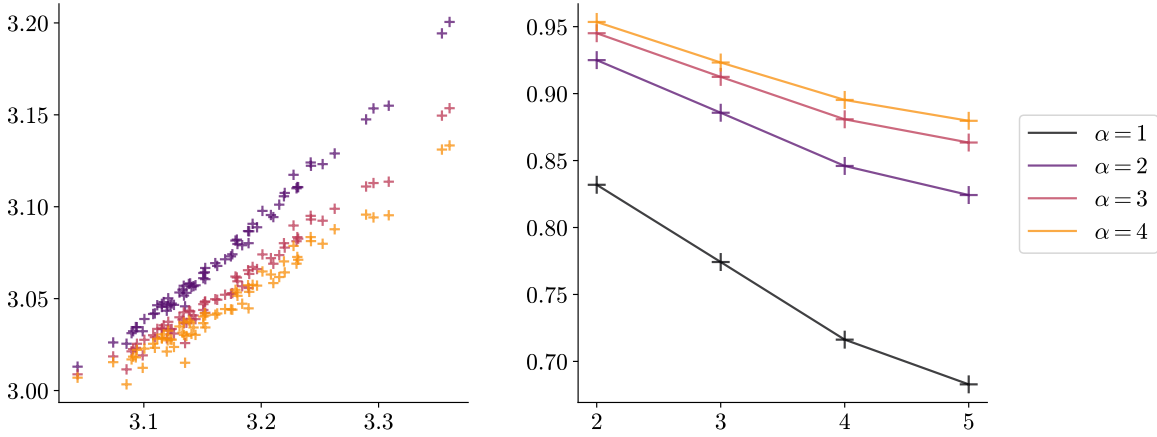


Figure 4-6: Distribution of γ computed exactly with von Neumann entropy for noisy chains with $l = 2, 3, 4, 5$ and $\varepsilon = 0.5$

behavior of the Rényi entropies mirror that of the von Neumann entropy, suggesting that it may be possible to extend our analysis to more computationally tractable proxies for the von Neumann TEE.



(a) Rényi entropies plotted against von Neumann entropy for noisy chains with $l = 2$

(b) Average values of γ_α computed with Rényi entropies for noisy chains with $l = 2, 3, 4, 5$, where $\alpha = 1$ denotes the von Neumann entropy

Figure 4-7: Comparison of Rényi and von Neumann entropies for chains with $\epsilon = 0.5$

4.5 Discussion

We have shown that under generic local deformations, spurious contributions to the TEE of the trivial phase decay exponentially in the size of the region. As the spurious TEE is known to be non-negative in models of nontrivial topological order, we argue that non-zero TEE is a necessary and sufficient condition for topological order in two-dimensional systems. We further conjecture that our results for spurious TEE generalize to other local circuit architectures as well as translation-invariant states which are in non-trivial topological phases, including the toric code.

In addition, the decay of TEE and CMI in different classes of systems is connected to broader issues in quantum complexity, notably questions such as when local parent Hamiltonians admit ground states with finite correlation length or trivial complexity, and vice versa. For instance, it is known that in 1D the exponential decay of CMI is both necessary and sufficient for a state to be approximated by a local Gibbs state corresponding to a parent Hamiltonian [59]. Furthermore, the preparation of non-trivial low energy states corresponding to certain classes of generic Hamiltonians has been recently proposed as a task which is classically intractable but quantumly easy [62], suggesting long-range order as a potential source of quantum advantage.

Notably, the results presented here do not apply in the case of structured circuits, in which the circuit may be restricted to commute with or remain invariant under the action of certain symmetries. Additional work is required to fully characterize the effects of these constraints and their interactions with SPT phases or other phenomena in systems of interest for condensed matter physics. Furthermore, we comment that the TEE cannot fully classify different topological phases, as for example it does not distinguish between models with Abelian and non-Abelian anyons. The development of robust diagnostic methods to more completely characterize topological order in the ground states of generic many-body systems therefore remains an open problem.

Finally, we comment that while the TEE has gained traction for its relative simplicity and versatility which enable analysis of its theoretical properties, it may be impractical to compute for large simulations, or to measure experimentally in a macro-

scopic system. However, we expect that our bound may be extended to the Rényi entropy analogues as well as other quantities which encode information about the entanglement entropy of the system, offering opportunities for further analysis.

Appendix A

Representations of the unitary group

In this appendix we provide a basic introduction to representation theory, before discussing results for the unitary group. In particular, we discuss the application of the Schur-Weyl duality to $U(N)$ to derive the Weingarten calculus, though we do not prove most of these results. For a more complete description of Lie groups and algebras and their representations, we suggest [63], which focuses on matrix Lie groups, or [64] for a more formal approach.

A.1 Overview of representation theory

Roughly speaking, the goal of representation theory is to relate objects from abstract algebra to linear transformations, which provide a tractable framework for analyzing their properties. More precisely:

Definition 12 (Group representation) *Given a vector space V , let $\text{End}(V)$ denote the set of linear maps defined on $V \rightarrow V$. Then a representation of a group G is a vector space V along with a homomorphism $R : G \rightarrow \text{End}(V)$.*

In order to fully classify the representations of a group, we can examine the class of all irreducible representations:

Definition 13 (Irreducible representation) *An irreducible representation or ir-rep (R, V) is one which does not contain a non-trivial subrepresentation $(R|_W, W)$*

with $W \subset V$ closed under the action of R . All unitary representations can be expressed as the direct sum of irreps.

Finite-dimensional unitary representations are directly relevant to the implementation of different sets of operations on quantum platforms. In particular, we are often interested in representations of Lie groups, which are a type of continuous group:

Definition 14 (Lie group) *A Lie group is a group $G = (G, \cdot)$ which is also a finite-dimensional differentiable manifold in the sense that for any $a, b \in G$ both $a \mapsto a^{-1}$ and $a, b \mapsto a \cdot b$ are differentiable maps.*

Lie groups often appear in physics as descriptions of continuous symmetries, such as rotations or dilations. It is useful to focus on matrix Lie groups, which can be described as follows:

Definition 15 (Matrix Lie group) *Let $GL_n(\mathbb{C})$ denote the group of invertible matrices of dimension n with complex entries. A matrix Lie group G is a closed subgroup of $GL_n(\mathbb{C})$.*

Due to their continuous nature, it is often difficult to explicitly compute properties of Lie groups. It is typically more straightforward to adopt a Lie algebra picture, which enables us to work with vector spaces directly rather than with differentiable manifolds.

Definition 16 (Lie algebra) *A finite-dimensional Lie algebra is a finite-dimensional vector space \mathfrak{g} along with a commutator operation $[\cdot, \cdot]$ which is a bilinear, skew symmetric map which satisfies the Jacobi identity that for any $a, b, c \in \mathfrak{g}$,*

$$[a, [b, c]] + [c, [a, b]] + [b, [c, a]] = 0. \quad (\text{A.1})$$

Definition 17 (Lie algebra representation) *A representation of a Lie algebra \mathfrak{g} on a vector space V is a map of Lie algebras*

$$r : \mathfrak{g} \rightarrow \mathfrak{gl}(V) \cong \text{End}(V).$$

We can then translate between Lie groups and algebras via a map which is motivated by the follow result:

Lemma 8 (One-parameter subgroup) *If A is a one-parameter subgroup of $GL_n(\mathbb{C})$, there exists a unique matrix $X \in M_n(\mathbb{C})$, up to a constant factor, such that $A = e^{tX}$.*

Definition 18 (Matrix Lie algebra) *For a matrix Lie group G , the corresponding Lie algebra \mathfrak{g} is given by the set of all X such that $e^{tX} \in G$.*

Here we have adopted the mathematical convention, though both of the results can be stated equivalently using the Wick rotated map e^{itX} , as is the physics convention. In other words, the Lie algebra is given by the tangent space of G at the identity. Any group operation can then be expressed in terms of the Lie algebra:

Lemma 9 (Group homomorphisms in the Lie algebra) *For any two matrix Lie groups G, H and homomorphism $\Phi : G \rightarrow H$, there exists a unique real-linear map $\phi : \mathfrak{g} \rightarrow \mathfrak{h}$ such that for all $X \in \mathfrak{g}$,*

$$\Phi(e^X) = e^{\phi(X)}. \quad (\text{A.2})$$

We can use this fact to relate representations of Lie algebras to those of the corresponding Lie group. In particular, for any simply connected Lie group, the representations of the Lie group are in one-to-one correspondence with those of the Lie algebra. While $U(N)$ and many of its subgroups are not simply connected, it can be shown that $SU(N)$ is simply connected for all $N \geq 1$, and its representation theory is well-known. These properties can be used to describe the algebraic structure of physical theories which obey certain gauge symmetries. We consider an example with $U(1)$ symmetry in Section A.3.1.

A.2 Schur-Weyl duality

In this section we describe a well-known result and its applications in quantum information theory. Given some finite-dimensional vector space V , consider the vector

space $V^{\otimes n}$. In the context of quantum circuits, V typically describes the state space of a single qudit and is of dimension $d = q^n$. Intuitively, for any $v = v_1 \otimes \dots \otimes v_n$ and permutation and linear transformations $\sigma \in S_n$, $g \in GL_q(\mathbb{C})$,

$$(\sigma \circ g^{\otimes n})(v) = (g^{\otimes n} \circ \sigma)(v).$$

This implies that there exists a basis in which σ and g are simultaneously diagonalized. The Schur-Weyl duality gives a stronger version of this statement, namely that this basis decomposes $V^{\otimes n}$ into irreps:

Theorem 7 (Schur-Weyl duality) *Any composite vector space $V^{\otimes n}$ admits a decomposition of the form*

$$V^{\otimes n} \cong \bigoplus_{\lambda} V_{\lambda} \otimes S_{\lambda} V, \quad (\text{A.3})$$

where the V_{λ} denote all irreps of S_n and $S_{\lambda} V$ is either an irrep of $GL_q(\mathbb{C})$ or zero.

Formally, this statement is equivalent to the condition that the span of S_n and $GL_q(\mathbb{C})$ are centralizers of each other in $\text{End}(V^{\otimes n})$, i.e. form the subset of all elements which leave the other fixed. This property holds when we restrict to unitary single-qudit operations, so that we obtain irreps corresponding to the action of a unitary transformation on each block. The Schur-Weyl duality thus provides a natural basis for describing systems which multipartite systems of independently and identically distributed states [65], and can be used to analyze a broad range of quantum protocols. Below, we discuss its application to explicitly computing coefficients in integrals over the unitary group.

A.3 Weingarten calculus

A key technique for the analysis of random circuit ensembles and quantum complexity is the Weingarten calculus, which was first derived explicitly in [66] using the Schur-Weyl duality. In order to compute statistics over the unitary group, we must first describe the equal-volume integration measure. Formally, the Haar measure is the

unique left- and right-invariant measure over $U(N)$, and therefore represents the uniform distribution over the group. Using the symmetries of the unitary group, it can be shown that

$$\int_{\text{Haar}} dU U_{i_1 j_1} \dots U_{i_k j_k} U_{l_1 m_1}^\dagger \dots U_{l_k m_k}^\dagger = \sum_{\sigma, \tau \in S_k} \delta(\vec{i}, \sigma(\vec{m})) \delta(\vec{j}, \tau(\vec{l})) Wg(\sigma^{-1} \tau, d), \quad (\text{A.4})$$

where we sum over permutation operators σ, τ . The Weingarten coefficients Wg are rational functions of q with degree at most $-k$. One method of computing these functions is by applying the Schur-Weyl duality and combinatorially constructing the irreps of the symmetric group, then solving the resulting system of equations, as we describe below:

First note that for any $V \in U(q)$, the twirling operators satisfy

$$\begin{aligned} V^{\otimes k} T_{\text{Haar}}^{(k)}(X) (V^\dagger)^{\otimes k} &= \int_{\text{Haar}} dU (VU)^{\otimes k} X (U^\dagger V^\dagger)^{\otimes k} \\ &= \int_{\text{Haar}} dU U^{\otimes k} X (U^\dagger)^{\otimes k} \\ &= T_{\text{Haar}}^{(k)}(X), \end{aligned} \quad (\text{A.5})$$

since $V^{\otimes k}$ can be absorbed into the integral due to the left- and right-invariance of the Haar measure. Since $T_{\text{Haar}}^{(k)}(X)$ commutes with any single-qudit unitary, it must consist of a linear combination of permutation operators [67]. Furthermore, for $X \in S_k$, we have that $T_{\text{Haar}}^{(k)}(X) = X$, thus yielding a complete system of equations which we may solve explicitly to compute the Weingarten calculus coefficients.

In addition, we note that a similar combinatorial formula exists for several subgroups of $U(N)$, including $U(N)/O(N)$ and $U(N)/Sp(N)$, which describe circuits which obey time-reversal symmetry [47, 68], in addition to other gauge symmetry groups [69]. As a simple example, the moments of $U(N)/U(1)$, which represents the group of unique transformations on systems which observe a $U(1)$ gauge symmetry [70], coincide with those of $U(N)$. As far as we know, this result does not appear in any previous literature, though it is relatively simple to show. We include a proof of this result in the following section.

A.3.1 $U(1)$ gauge symmetry

While the case of a $U(1)$ gauge symmetry is fairly straightforward, as far as we know, it has not been explicitly computed in the literature. The fact that the moments of the resulting subgroup is the same implies that circuits obeying such a symmetry exhibit the same qualitative behavior. We provide a proof of this fact below.

Lemma 10 *For any N , $U(N)/U(1) \cong SU(N)/Z(N)$.*

Proof. This is a well-known property [71] which follows from the definition of $SU(N)$ as the group of unitary matrices with determinant 1. Recall that any matrix $U \in U(N)$ can be written in the form $U = \det(U)U'$, where $|\det(U)| = 1$ and $\det(U') = 1$. Then there exists $\theta(U) \in \mathbb{R}$ such that $e^{i\theta(U)} = \det U$. We therefore have $U = e^{i\theta(U)}U'$, so that any unitary can be written as the product of an element of $U(1)$ and an element of $SU(N)$. Furthermore, enforcing $\det(U)$ shows that this decomposition is unique up to a factor of $e^{2\pi ij/N}$. We therefore have an isomorphism mapping $U(N) \rightarrow U(1) \times SU(N)/Z(N)$. Taking the quotient of both the domain and codomain gives the desired result.

Corollary 3 *For any N , the moments of $U(N)/U(1)$ coincide with those of $U(N)$.*

Proof. We apply the isomorphism from the above result to write integrals over $U(N)$ in terms of commuting integrals over $U(1)$, $SU(N)$, and $Z(N)$, so that the moments over $U(N)$ are given by

$$\begin{aligned} T_{\text{Haar}}^{(k)}(X) &= \int_{\text{Haar}} dU U^{\otimes k} X(U^\dagger)^{\otimes k} \\ &= \sum_{j \in Z(N)} \int_{U(1)} d\theta \int_{SU(N)} dU' (e^{i\theta} e^{2\pi ij/N} U')^{\otimes k} X(e^{-i\theta} e^{-2\pi ij/N} U'^\dagger)^{\otimes k} \\ &= \sum_{j \in Z(N)} \int_{U(1)} d\theta \int_{SU(N)} dU' (U')^{\otimes k} X(U'^\dagger)^{\otimes k}. \end{aligned}$$

We therefore have that the moments are independent of the integral over θ , and are thus equal for both $SU(N)$ and $U(N)/U(1) \cong SU(N) \times Z(N)$. Note that a more comprehensive treatment of integrals over $SU(N)$ can be found in [72].

Appendix B

Properties of the quantum W_1 norm

In this appendix we briefly review background on optimal transport theory and inspiration for the definition of the quantum Wasserstein distance of order 1 in [45]. We first provide an overview of optimal transport theory and discuss a classical W_1 distance derived from the Hamming distance on bit strings. We then provide proofs for some basic properties of the quantum W_1 norm.

B.1 Classical Wasserstein distance

B.1.1 Foundations of optimal transport

Optimal transport theory is concerned with optimization problems over distributions of resources, and is well-established in the classical setting. The central objective of this area is the Monge-Kantorovich problem, which for the discrete and finite-dimensional case can be given as follows [73]:

Definition 19 (Monge-Kantorovich problem) *Given any two discrete and finite-dimensional measures*

$$\alpha = \sum_{i=1}^n a_i \delta_{x_i}, \quad \beta = \sum_{j=1}^m b_j \delta_{y_j}, \quad (\text{B.1})$$

with $x_i, y_j \in \mathcal{X}$, as well as the cost matrix $C_{ij} = c(x_i, y_j)$, the minimization of trans-

portation cost between the two measures can be expressed as a linear program via

$$L_C = \min_P \sum_{ij} P_{ij} C_{ij}, \quad (\text{B.2})$$

where P is doubly stochastic and the solution describes the optimal coupling, or equivalently transport map, between α and β .

Optical offers a broad framework for statistical analysis, and can be applied to a wide range of problems, including in machine learning and economics theory [73].

Notably, when the cost matrix is chosen to be a power of a Hilbertian distance on the underlying event space \mathcal{X} , the transportation problem induces another distance on the space of measures via the optimal transport cost, which is known as the Wasserstein or W_p distance. Unlike metrics such as relative entropy, W_p satisfies the triangle inequality for all p and is therefore a true distance over measures.

In practice, p is often chosen to be 1 or 2 for ease of analysis. It can be shown that the W_2 distance induces a Riemannian structure on the space of continuous probability measures, providing additional insight on classical information geometry [74–76]. On the other hand, for discrete probability distributions, it is often useful to consider W_1 distances [73, 77]. In general, the W_1 metric is upper bounded by the W_2 , and in some settings they can be directly related to show that the structure of continuous probability spaces extends to the discrete case, or vice versa [73, 78].

B.1.2 Continuity bound for Shannon entropy

Continuity bounds for the Shannon entropy arise in applications such as characterizing the capacity region of classical interference channels. While the capacity may be difficult to compute exactly for an arbitrary channel, such bounds enable efficient estimation of a large family of channels. Notably, no dimension-independent bound can be given in terms of the relative entropy. More precisely:

Lemma 11 (*Difference of Shannon entropies*) *Given discrete probability distribu-*

tions $P, Q \in \mathcal{M}(\mathcal{X})$ and metric δ over $\mathcal{M}(\mathcal{X})$, consider bounds of the form

$$|H(P) - H(Q)| \leq c\delta(P, Q). \quad (\text{B.3})$$

For $\delta(P, Q) = D(P||Q)$, there is no constant c which satisfies Eq. B.3.

To see this, consider the case in which P is the uniform distribution over the sequence of integers between 1 and n . Let Q be the mixture of P and the uniform distribution over the integers between 1 and n^2 , with probability $(n+1)/2n^2$ for the integers from 1 to n . Note that in the limit $n \rightarrow \infty$, $D(P||Q)$ converges to $\log 2$, whereas the gap of entropies is unbounded and grows as $\Theta(\log n)$. Thus there is no constant c which satisfies Eq. B.3.

On the other hand, it is shown in [77] that choosing the Wasserstein distance of order 1 induced by the classical Hamming distance over bit strings does provide a bound for the difference of Shannon entropies. Defining

$$W_1(P, Q) = \inf_{\pi_{XY}} \mathbb{E}[d_H(X, Y)], \quad (\text{B.4})$$

where the infimum is taken over all couplings between the distributions X, Y , it can be shown that the difference of entropies is bounded by a convex function of the W_1 distance:

Theorem 8 (*Continuity of Shannon entropy*) *Given discrete probability distributions $P, Q \in \mathcal{M}(\mathcal{X})$, we have that*

$$|H(P) - H(Q)| \leq nF_{\mathcal{X}}\left(\frac{1}{n}W_1(P, Q)\right), \quad (\text{B.5})$$

$$F_{\mathcal{X}} = x \log(|\mathcal{X}| - 1) + H_2(x),$$

up to a constant factor which is determined by the base of the logarithm.

This result is due to the independence of each bit and the key result that this W_1 distance is additive in the length of the bit string, and is proved in [77].

B.2 Quantum Wasserstein distance of order 1

The proposal in [45] generalizes the analysis in [77] on bit strings to construct a quantum Wasserstein distance of order 1. This metric is based on the concept of “neighboring” states, which are equal after one qudit is discarded.

More precisely, consider a system of n qudits with local dimension q , with total state space \mathcal{H}^n . Let \mathcal{O}_n be the set of all Hermitian operators on \mathcal{H}^n , and $\mathcal{D}_n \subset \mathcal{O}_n$ denote the set of density matrices on \mathcal{H}^n . We then extend the Hamming distance as follows:

Definition 20 (*Neighboring states*) *Two states $\rho, \rho' \in \mathcal{D}_n$ are neighboring if their reduced density matrices over $n - 1$ qudits coincide, so that*

$$\mathrm{tr}_i \rho = \mathrm{tr}_i \rho', \quad 1 \leq i \leq n. \quad (\text{B.6})$$

In addition, we define the set of all possible differences between two neighboring states:

$$\begin{aligned} \mathcal{N}_n^{(i)} &= \{\rho - \rho' : \rho, \rho' \in \mathcal{D}_n, \mathrm{tr}_i \rho = \mathrm{tr}_i \rho'\}, \\ \mathcal{N}_n &= \bigcup_{i=1}^n \mathcal{N}_n^{(i)}. \end{aligned} \quad (\text{B.7})$$

Now consider the convex hull of \mathcal{N}_n , which we denote by \mathcal{B}_n :

$$\mathcal{B}_n = \left\{ \sum_{i=1}^n p_i N_i : N_i \in \mathcal{N}_n^{(i)}, p \in \mathcal{M}_+([i]) \right\}. \quad (\text{B.8})$$

Note that \mathcal{D}_n is closed and bounded in the Schatten 1-norm, so \mathcal{B}_n is also a closed, bounded, and convex subset of \mathcal{O}_n . Furthermore, $\mathcal{B}_n = -\mathcal{B}_n$ due to symmetry. This set can therefore be interpreted as a “unit ball”, which induces the desired metric:

Definition 21 (*Quantum W_1 norm*) *For any $X \in \mathcal{O}_n$, the quantum W_1 norm is defined via*

$$\|X\|_{W_1} = \min\{w \geq 0 : X \in w\mathcal{B}_n\}. \quad (\text{B.9})$$

The W_1 distance between two states on the system $\rho, \rho' \in \mathcal{D}_n$ is then given by

$$W_1(\rho, \rho') = \|\rho - \rho'\|_{W_1}. \quad (\text{B.10})$$

Note from the definition of \mathcal{B}_n that the quantum W_1 norm is non-negative, symmetric, and satisfies the triangle inequality, and is therefore a proper distance. In addition, it is clear that the above definition reduces to the classical Hamming distance when restricted to the set of computational basis states. In the remainder of this section, we will use W_1 to denote the quantity given in Def. 21 unless otherwise specified.

Similar to the classical case, the quantum W_1 norm also yields a Lipschitz continuity bound for the von Neumann entropy, which is proved in [45]:

Theorem 9 *For $\rho, \rho' \in \mathcal{D}_n$, the difference of von Neumann entropies is bounded via*

$$\begin{aligned} |S(\rho) - S(\rho')| &\leq g(\|\rho - \rho'\|_{W_1}) + \|\rho - \rho'\|_{W_1} \ln d^2 n, \\ g(t) &= (t + 1) \ln(t + 1) - t \ln t. \end{aligned} \quad (\text{B.11})$$

For a family of states ρ, ρ' with fixed d and $\epsilon > 0$, Theorem 9 implies that if

$$\|\rho - \rho'\|_{W_1} \leq \frac{\epsilon n}{\ln d^2 n}, \quad (\text{B.12})$$

then we have that asymptotically for large n ,

$$|S(\rho) - S(\rho')| \leq \epsilon n + O(\ln n). \quad (\text{B.13})$$

This result suggests that the entanglement structure of a quantum state is robust under perturbations up to $o(n/\ln n)$ in the W_1 distance [45], motivating the application of the W_1 norm to study the growth of information complexity in many-body systems such as in [79].

B.2.1 Properties

In this section we denote by \mathcal{O}_n^T the set of all traceless Hermitian operators on \mathcal{H}^n . We summarize some of the key properties of the quantum W_1 distance and its behavior under quantum operations, and reproduce the proofs given in [45].

Lemma 12 (Symmetries of the W_1 distance) *For any $X \in \mathcal{O}_n$ and quantum channel $\Phi : \mathcal{D}_n \rightarrow \mathcal{D}_n$ which can be decomposed into a permutation of qudits $\sigma \in S_n$ and a series of single-qudit channels Φ_i , we have*

$$\|\Phi(X)\|_{W_1} \leq \|X\|_{W_1}, \quad (\text{B.14})$$

with equality when the single-qudit channels are unitary.

Proof. First, note that from Definition 20, \mathcal{B}_n is the convex hull of a set which is symmetric over all qudits and local changes of basis, and mapped to itself under single-qudit channels. It follows that the set \mathcal{B}_n is also invariant under permutations and single-qudit operations. Thus for $w = \|X\|_{W_1}$,

$$\Phi(X) \in w\mathcal{B}_n,$$

so we have $\|\Phi(X)\|_{W_1} \leq w$. Furthermore, when Φ is the composition of a permutation and a series of unitaries, we have $\Phi^{-1}(\mathcal{B}_n) = \mathcal{B}_n$, so that the W_1 norm is invariant.

Lemma 13 (Relation to trace distance, Lemma 4 in main text) *For any $X \in \mathcal{O}_n^T$, the W_1 norm of X is bounded via*

$$\frac{1}{2}\|X\|_1 \leq \|X\|_{W_1} \leq \frac{n}{2}\|X\|_1. \quad (\text{B.15})$$

Proof. Note that we can equivalently write the W_1 norm as

$$\|X\|_{W_1} = \frac{1}{2} \min \left(\sum_{i=1}^n \|X^{(i)}\|_1 : X^{(i)} \in \mathcal{O}_n^T, \text{Tr}_i X^{(i)} = 0, X = \sum_{i=1}^n X^{(i)} \right), \quad (\text{B.16})$$

where we have obtained an extra factor of $1/2$ from the symmetry of \mathcal{B}_n . From the triangle inequality,

$$\|X\|_1 \leq \sum_{i=1}^n \|X^{(i)}\|_1,$$

so that we obtain the lower bound

$$\frac{1}{2}\|X\|_1 \leq \|X\|_{W_1}.$$

On the other hand, for any traceless X we can consider a decomposition $X = X^+ - X^-$, where X^+, X^- are positive semi-definite operators with orthogonal support and $\text{Tr } X^\pm = \|X\|_1/2$. We can pick in Eq. B.16

$$X^{(i)} = \frac{2}{\|X\|_1} (\text{Tr}_{1\dots i-1} X^+ \otimes \text{Tr}_{i\dots n} X^- - \text{Tr}_{1\dots i} X^+ \otimes \text{Tr}_{i+1\dots n} X^-),$$

so that for each i we are guaranteed

$$\|X^{(i)}\|_1 \leq \|X\|_1.$$

It therefore follows that

$$\|X\|_{W_1} \leq \frac{1}{2} \sum_{i=1}^n \|X^{(i)}\|_1 \leq \frac{n}{2} \|X\|_1. \quad (\text{B.17})$$

Lemma 14 (Tensorization, Lemma 3 in main text) *For any $X \in \mathcal{O}_{m+n}^T$,*

$$\|X\|_{W_1} \geq \|\text{tr}_{1\dots m} X\|_{W_1} + \|\text{tr}_{m+1\dots m+n} X\|_{W_1}. \quad (\text{B.18})$$

Proof. Consider X and $X^{(i)} \in \mathcal{O}_{m+n}^T$ such that

$$\text{Tr}_i X^{(i)} = 0, \quad X = \sum_{i=1}^{m+n} X^{(i)}.$$

We can then write the partial traces of X as

$$\begin{aligned}\mathrm{Tr}_{1\dots m} X &= \sum_{i=m+1}^{m+n} \mathrm{Tr}_{1\dots m} X^{(i)}, \\ \mathrm{Tr}_{m+1\dots m+n} X &= \sum_{i=1}^m \mathrm{Tr}_{m+1\dots m+n} X^{(i)}.\end{aligned}$$

Then applying Eq. B.17 yields

$$\|\mathrm{Tr}_{1\dots m} X\|_{W_1} + \|\mathrm{Tr}_{m+1\dots m+n} X\|_{W_1} \leq \frac{1}{2} \sum_{i=1}^m \|\mathrm{Tr}_{m+1\dots m+n} X^{(i)}\|_1 + \frac{1}{2} \sum_{i=m+1}^{m+n} \|\mathrm{Tr}_{1\dots m} X^{(i)}\|_1.$$

Since this holds for all possible choices of $X^{(i)}$, we can apply the contractivity of the 1-norm under partial traces to obtain the result

$$\|\mathrm{Tr}_{1\dots m} X\|_{W_1} + \|\mathrm{Tr}_{m+1\dots m+n} X\|_{W_1} \leq \min \frac{1}{2} \sum_{i=1}^n \|X^{(i)}\|_1 = \|X\|_{W_1}.$$

Lemma 15 (Local operations, Lemma 2 in main text) *For any multi-index $I \subset [n]$ and $X \in \mathcal{O}_n$ such that $\mathrm{tr}_I X = 0$, the W_1 norm satisfies*

$$\|X\|_{W_1} \leq |I| \frac{q^2 - 1}{q^2} \|X\|_1. \quad (\text{B.19})$$

Proof. First, we note that for any $X \in \mathcal{O}_n$, the completely depolarizing channel on \mathcal{D}_n obeys the relation

$$\left\| X - \frac{\mathbb{1}}{q} \otimes \mathrm{Tr}_1 X \right\|_1 \leq 2 \frac{q^2 - 1}{q^2} \|X\|_1, \quad (\text{B.20})$$

which can be seen from the standard decomposition of the depolarizing channel as a sum of generalized Pauli terms:

$$\left\| X - \frac{\mathbb{1}}{q} \otimes \mathrm{Tr}_1 X \right\|_1 = \left\| \frac{d^2 - 1}{d^2} X - \frac{1}{d^2} \sum_{i=2}^{d^2} (P_i \otimes \mathbb{1}^{\otimes(n-1)}) X (P_i^\dagger \otimes \mathbb{1}^{\otimes(n-1)}) \right\|_1.$$

Without loss of generality, let $I = [k]$ for $k \leq n$ and

$$X^{(i)} = \frac{1}{q^{i-1}} \mathbb{1}^{\otimes(i-1)} \otimes \text{Tr}_{1\dots i-1} X - \frac{1}{q^i} \mathbb{1}^{\otimes i} \otimes \text{Tr}_{1\dots i} X,$$

so that we have

$$\text{Tr}_i X^{(i)} = 0, \quad X = \sum_{i=1}^k X^{(i)}.$$

Then the result follows from substituting Eq. B.20 into Eq. B.16:

$$\begin{aligned} \|X\|_{W_1} &\leq \frac{1}{2} \sum_{i=1}^n \|X^{(i)}\|_1 \\ &\leq \frac{1}{2} \sum_{i=1}^k \|X^{(i)}\|_1 \\ &\leq \|I\| \frac{q^2 - 1}{q^2} \|X\|_1. \end{aligned}$$

Bibliography

- [1] L. Cui, “Local information scrambling in random quantum circuits,” 2021. unpublished manuscript.
- [2] L. Cui, “Quantum complexity frameworks from optimal transport,” 2022. unpublished manuscript.
- [3] T. Cover and J. Thomas, *Elements of information theory*. Wiley, 2012.
- [4] J. M. Deutsch, “Quantum statistical mechanics in a closed system,” *Physical Review A*, vol. 43, pp. 2046–2049, Feb 1991.
- [5] M. Srednicki, “Chaos and quantum thermalization,” *Physical Review E*, vol. 50, no. 2, p. 888, 1994.
- [6] R. Nandkishore and D. A. Huse, “Many-body localization and thermalization in quantum statistical mechanics,” *Annual Review of Condensed Matter Physics*, vol. 6, no. 1, pp. 15–38, 2015.
- [7] L. Savary and L. Balents, “Quantum spin liquids: a review,” *Reports on Progress in Physics*, vol. 80, no. 1, p. 016502, 2016.
- [8] P. Hayden and J. Preskill, “Black holes as mirrors: quantum information in random subsystems,” *Journal of High Energy Physics*, vol. 2007, p. 120–120, Sep 2007.
- [9] N. Lashkari, D. Stanford, M. Hastings, T. Osborne, and P. Hayden, “Towards the fast scrambling conjecture,” *Journal of High Energy Physics*, vol. 2013, no. 4, p. 22, 2013.
- [10] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian, and A. Tajdini, “Replica wormholes and the entropy of Hawking radiation,” *Journal of High Energy Physics*, vol. 2020, no. 5, 2020.
- [11] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian, and A. Tajdini, “The entropy of Hawking radiation,” *Reviews of Modern Physics*, vol. 93, no. 3, 2021.
- [12] R. P. Feynman, “Simulating physics with computers,” *International Journal of Theoretical Physics*, vol. 21, no. 6, pp. 467–488, 1982.

- [13] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell, *et al.*, “Quantum supremacy using a programmable superconducting processor,” *Nature*, vol. 574, no. 7779, pp. 505–510, 2019.
- [14] J. Emerson, Y. S. Weinstein, M. Saraceno, S. Lloyd, and D. G. Cory, “Pseudo-random unitary operators for quantum information processing,” *Science*, vol. 302, no. 5653, pp. 2098–2100, 2003.
- [15] A. Nahum, S. Vijay, and J. Haah, “Operator spreading in random unitary circuits,” *Physical Review X*, vol. 8, no. 2, p. 021014, 2018.
- [16] J. Choi, A. L. Shaw, I. S. Madjarov, X. Xie, J. P. Covey, J. S. Cotler, D. K. Mark, H.-Y. Huang, A. Kale, H. Pichler, *et al.*, “Emergent randomness and benchmarking from many-body quantum chaos,”
- [17] J. S. Cotler, D. K. Mark, H.-Y. Huang, F. Hernandez, J. Choi, A. L. Shaw, M. Endres, and S. Choi, “Emergent quantum state designs from individual many-body wave functions,” *PRX Quantum*, vol. 4, no. 1, p. 010311, 2023.
- [18] M. P. Fisher, V. Khemani, A. Nahum, and S. Vijay, “Random quantum circuits,” *Annual Review of Condensed Matter Physics*, vol. 14, pp. 335–379, 2023.
- [19] H.-Y. Huang, R. Kueng, and J. Preskill, “Predicting many properties of a quantum system from very few measurements,” *Nature Physics*, vol. 16, p. 1050–1057, Jun 2020.
- [20] L. Lewis, H.-Y. Huang, V. T. Tran, S. Lehner, R. Kueng, and J. Preskill, “Improved machine learning algorithm for predicting ground state properties,” 2023. arXiv:2301.13169 [quant-ph].
- [21] F. G. Brandão, W. Chemsyany, N. Hunter-Jones, R. Kueng, and J. Preskill, “Models of quantum complexity growth,” *PRX Quantum*, vol. 2, no. 3, p. 030316, 2021.
- [22] A. M. Dalzell, N. Hunter-Jones, and F. G. Brandão, “Random quantum circuits anticoncentrate in log depth,” *PRX Quantum*, vol. 3, no. 1, p. 010333, 2022.
- [23] J. C. Napp, R. L. La Placa, A. M. Dalzell, F. G. Brandao, and A. W. Harrow, “Efficient classical simulation of random shallow 2D quantum circuits,” *Physical Review X*, vol. 12, no. 2, p. 021021, 2022.
- [24] D. Aharonov, X. Gao, Z. Landau, Y. Liu, and U. Vazirani, “A polynomial-time classical algorithm for noisy random circuit sampling,” 2022. arXiv:2211.03999 [quant-ph].
- [25] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, “Local random quantum circuits are approximate polynomial-designs,” *Communications in Mathematical Physics*, vol. 346, no. 2, pp. 397–434, 2016.

- [26] N. Hunter-Jones, “Unitary designs from statistical mechanics in random quantum circuits,” 2019. arXiv:1905.12053 [quant-ph].
- [27] J. Haferkamp, P. Faist, N. B. Kothakonda, J. Eisert, and N. Yunger Halpern, “Linear growth of quantum circuit complexity,” *Nature Physics*, vol. 18, no. 5, pp. 528–532, 2022.
- [28] H.-Y. Hu, S. Choi, and Y.-Z. You, “Classical shadow tomography with locally scrambled quantum dynamics,” *Physical Review Research*, vol. 5, no. 2, p. 023027, 2023.
- [29] H.-Y. Huang, R. Kueng, G. Torlai, V. V. Albert, and J. Preskill, “Provably efficient machine learning for quantum many-body problems,” *Science*, vol. 377, no. 6613, p. eabk3333, 2022.
- [30] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, “Symmetric informationally complete quantum measurements,” *Journal of Mathematical Physics*, vol. 45, no. 6, pp. 2171–2180, 2004.
- [31] A. Ambainis and J. Emerson, “Quantum t-designs: t-wise independence in the quantum world,” in *Twenty-Second Annual IEEE Conference on Computational Complexity (CCC’07)*, pp. 129–140, IEEE, 2007.
- [32] Z. Webb, “The Clifford group forms a unitary 3-design,” 2015. arXiv:1510.02769 [quant-ph].
- [33] F. Barratt, U. Agrawal, S. Gopalakrishnan, D. A. Huse, R. Vasseur, and A. C. Potter, “Field theory of charge sharpening in symmetric monitored quantum circuits,” *Physical review letters*, vol. 129, no. 12, p. 120604, 2022.
- [34] V. Khemani, A. Vishwanath, and D. A. Huse, “Operator spreading and the emergence of dissipative hydrodynamics under unitary evolution with conservation laws,” *Physical Review X*, vol. 8, p. 031057, Sep 2018.
- [35] T. Rakovszky, F. Pollmann, and C. Von Keyserlingk, “Diffusive hydrodynamics of out-of-time-ordered correlators with charge conservation,” *Physical Review X*, vol. 8, no. 3, p. 031058, 2018.
- [36] I. Marvian, “Restrictions on realizable unitary operations imposed by symmetry and locality,” *Nature Physics*, vol. 18, no. 3, pp. 283–289, 2022.
- [37] I. Marvian, “(Non-) universality in symmetric quantum circuits: Why Abelian symmetries are special,” 2023. arXiv:2302.12466 [quant-ph].
- [38] M. M. Wilde, “From classical to quantum Shannon theory,” 2011. arXiv:1106.1445 [quant-ph].
- [39] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, “Matrix product state representations,” 2006. arXiv:quant-ph/0608197.

- [40] M. B. Hastings, “An area law for one-dimensional quantum systems,” *Journal of statistical mechanics: theory and experiment*, vol. 2007, no. 08, p. P08024, 2007.
- [41] H. Gharibyan, M. Hanada, S. H. Shenker, and M. Tezuka, “Onset of random matrix behavior in scrambling systems,” *Journal of High Energy Physics*, vol. 2018, no. 7, pp. 1–62, 2018.
- [42] M. A. Nielsen, “A geometric approach to quantum circuit lower bounds,” 2005. arXiv:quant-ph/0502070.
- [43] M. A. Nielsen, M. R. Dowling, M. Gu, and A. C. Doherty, “Quantum computation as geometry,” *Science*, vol. 311, no. 5764, pp. 1133–1135, 2006.
- [44] M. A. Nielsen, M. R. Dowling, M. Gu, and A. C. Doherty, “Optimal control, geometry, and quantum computing,” *Physical Review A*, vol. 73, p. 062323, Jun 2006.
- [45] G. D. Palma, M. Marvian, D. Trevisan, and S. Lloyd, “The quantum Wasserstein distance of order 1,” *IEEE Transactions on Information Theory*, vol. 67, no. 10, pp. 6627–6643, 2021.
- [46] O. C. Dahlsten, C. Lupo, S. Mancini, and A. Serafini, “Entanglement typicality,” *Journal of Physics A: Mathematical and Theoretical*, vol. 47, no. 36, p. 363001, 2014.
- [47] N. Hunter-Jones, “Operator growth in random quantum circuits with symmetry,” 2018. arXiv:1812.08219 [quant-ph].
- [48] “Non-abelian braiding of graph vertices in a superconducting processor,” *Nature*, pp. 1–6, 2023.
- [49] A. Y. Kitaev, “Fault-tolerant quantum computation by anyons,” *Annals of physics*, vol. 303, no. 1, pp. 2–30, 2003.
- [50] Z.-C. Gu, Z. Wang, and X.-G. Wen, “Classification of two-dimensional fermionic and bosonic topological orders,” *Physical Review B*, vol. 91, no. 12, p. 125149, 2015.
- [51] B. Zeng, X. Chen, D.-L. Zhou, X.-G. Wen, *et al.*, *Quantum information meets quantum matter*. Springer, 2019.
- [52] K. Kato and F. G. Brandão, “Toy model of boundary states with spurious topological entanglement entropy,” *Physical Review Research*, vol. 2, no. 3, p. 032005, 2020.
- [53] J. Haah, “Classification of translation invariant topological pauli stabilizer codes for prime dimensional qudits on two-dimensional lattices,” *Journal of Mathematical Physics*, vol. 62, no. 1, p. 012201, 2021.

- [54] A. Kitaev and J. Preskill, “Topological entanglement entropy,” *Physical Review Letters*, vol. 96, no. 11, p. 110404, 2006.
- [55] L. Zou and J. Haah, “Spurious long-range entanglement and replica correlation length,” *Physical Review B*, vol. 94, no. 7, p. 075151, 2016.
- [56] D. J. Williamson, A. Dua, and M. Cheng, “Spurious topological entanglement entropy from subsystem symmetries,” *Physical Review Letters*, vol. 122, no. 14, p. 140506, 2019.
- [57] I. H. Kim, M. Levin, T.-C. Lin, D. Ranard, and B. Shi, “Universal lower bound on topological entanglement entropy,” 2023. arXiv:2302.00689 [quant-ph].
- [58] K. Kato, 2022. private communication.
- [59] C.-F. Chen, K. Kato, and F. G. Brandão, “Matrix product density operators: When do they have a local parent Hamiltonian?,” 2020. arXiv:2010.14682 [quant-ph].
- [60] D. Fattal, T. S. Cubitt, Y. Yamamoto, S. Bravyi, and I. L. Chuang, “Entanglement in the stabilizer formalism,” *arXiv preprint quant-ph/0406168*, 2004.
- [61] L. Cui, “Topological entropy,” 2022. Github repository.
- [62] A. M. Dalzell, M. Berta, F. G. Brandão, J. A. Tropp, *et al.*, “Sparse random hamiltonians are quantumly easy,” *arXiv preprint arXiv:2302.03394*, 2023.
- [63] B. C. Hall and B. C. Hall, *Lie groups, Lie algebras, and representations*. Springer, 2013.
- [64] W. Fulton and J. Harris, *Representation theory: a first course*, vol. 129. Springer Science & Business Media, 2013.
- [65] A. W. Harrow, “Applications of coherent classical communication and the schur transform to quantum information theory,” *arXiv preprint quant-ph/0512255*, 2005.
- [66] B. Collins, “Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral, and free probability,” *International Mathematics Research Notices*, vol. 2003, no. 17, pp. 953–982, 2003.
- [67] A. Kitaev, “Averaging over the unitary group.” unpublished manuscript.
- [68] F. J. Dyson, “The threefold way. algebraic structure of symmetry groups and ensembles in quantum mechanics,” *Journal of Mathematical Physics*, vol. 3, no. 6, pp. 1199–1215, 1962.
- [69] S. Matsumoto, “Weingarten calculus for matrix ensembles associated with compact symmetric spaces,” 2013. arXiv:1301.5401 [math.PR].

- [70] A. Moghimnejad and S. Parvizi, “Circuit complexity in $U(1)$ gauge theory,” *Modern Physics Letters A*, vol. 36, no. 34, p. 2150240, 2021.
- [71] AG learner, “Showing $U(n)/Z(U(n)) = SU(n)/Z(SU(n))$.” Mathematics Stack Exchange. (version: 2017-04-13).
- [72] O. Borisenko, S. Voloshyn, and V. Chelnokov, “ $SU(N)$ polynomial integrals and some applications,” *Reports on Mathematical Physics*, vol. 85, no. 1, pp. 129–145, 2020.
- [73] G. Peyré and M. Cuturi, “Computational optimal transport,” 2018. arXiv:1803.00567 [stat.ML].
- [74] S.-i. Amari, R. Karakida, and M. Oizumi, “Information geometry connecting Wasserstein distance and Kullback–Leibler divergence via the entropy-relaxed transportation problem,” *Information Geometry*, vol. 1, no. 1, pp. 13–37, 2018.
- [75] I. Gentil, C. Léonard, L. Ripani, and L. Tamanini, “An entropic interpolation proof of the HWI inequality,” *Stochastic Processes and their Applications*, vol. 130, no. 2, pp. 907–923, 2020.
- [76] F. Otto and C. Villani, “Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality,” *Journal of Functional Analysis*, vol. 173, no. 2, pp. 361–400, 2000.
- [77] Y. Polyanskiy and Y. Wu, “Wasserstein continuity of entropy and outer bounds for interference channels,” *IEEE Transactions on Information Theory*, vol. 62, no. 7, pp. 3992–4002, 2016.
- [78] N. Datta and C. Rouzé, “Relating relative entropy, optimal transport and Fisher information: a quantum HWI inequality,” in *Annales Henri Poincaré*, vol. 21, pp. 2115–2150, Springer, 2020.
- [79] L. Li, K. Bu, D. E. Koh, A. Jaffe, and S. Lloyd, “Wasserstein complexity of quantum circuits,” 2022. arXiv:2208.06306 [quant-ph].