MIT Open Access Articles

Entanglement, quantum randomness, and complexity beyond scrambling

Entanglement, quantum randomness, and complexity beyond scrambling

Zi-Wen Liu,[a,b,1] **Seth Lloyd,**[c,b] **Elton Zhu**[a,b] **and Huangjun Zhu**[d,e,f,g,h]

[a] *Center for Theoretical Physics, Massachusetts Institute of Technology,*
Cambridge, MA 02139, U.S.A.

[b] *Department of Physics, Massachusetts Institute of Technology,*
Cambridge, MA 02139, U.S.A.

[c] *Department of Mechanical Engineering, Massachusetts Institute of Technology,*
Cambridge, MA 02139, U.S.A.

[d] *Institute for Theoretical Physics, University of Cologne,*
50937 Cologne, Germany

[e] *Department of Physics and Center for Field Theory and Particle Physics, Fudan University,*
Shanghai 200433, China

[f] *Institute for Nanoelectronic Devices and Quantum Computing, Fudan University,*
Shanghai 200433, China

[g] *State Key Laboratory of Surface Physics, Fudan University,*
Shanghai 200433, China

[h] *Collaborative Innovation Center of Advanced Microstructures,*
Nanjing 210093, China

E-mail: zwliu@mit.edu, slloyd@mit.edu, eltonzhu@mit.edu,
zhuhuangjun@fudan.edu.cn

ABSTRACT: Scrambling is a process by which the state of a quantum system is effectively randomized due to the global entanglement that "hides" initially localized quantum information. Closely related notions include quantum chaos and thermalization. Such phenomena play key roles in the study of quantum gravity, many-body physics, quantum statistical mechanics, quantum information etc. Scrambling can exhibit different complexities depending on the degree of randomness it produces. For example, notice that the complete randomization implies scrambling, but the converse does not hold; in fact, there is a significant complexity gap between them. In this work, we lay the mathematical foundations of studying randomness complexities beyond scrambling by entanglement properties. We do so by analyzing the generalized (in particular Rényi) entanglement entropies of designs, i.e. ensembles of unitary channels or pure states that mimic the uniformly random distribution (given by the Haar measure) up to certain moments. A main

---

[1]Corresponding author.

collective conclusion is that the Rényi entanglement entropies averaged over designs of the same order are almost maximal. This links the orders of entropy and design, and therefore suggests Rényi entanglement entropies as diagnostics of the randomness complexity of corresponding designs. Such complexities form a hierarchy between information scrambling and Haar randomness. As a strong separation result, we prove the existence of (state) 2-designs such that the Rényi entanglement entropies of higher orders can be bounded away from the maximum. However, we also show that the min entanglement entropy is maximized by designs of order only logarithmic in the dimension of the system. In other words, logarithmic-designs already achieve the complexity of Haar in terms of entanglement, which we also call max-scrambling. This result leads to a generalization of the fast scrambling conjecture, that max-scrambling can be achieved by physical dynamics in time roughly linear in the number of degrees of freedom. This paper is an extended version of *Phys. Rev. Lett.* **120** (2018) 130502 [1].

Contents

1 Introduction

Scrambling describes a property of the dynamics of isolated quantum systems, in which initially localized quantum information spreads out over the whole system, thereby becoming inaccessible to local observers. The notion of scrambling originates from the study of black holes in quantum gravity [2–4]. The thermal nature of the Hawking radiation [5–7] indicates that the state of any matter and information falling into the black hole has been scrambled and so gets lost from the perspective of an external observer. In particular, the "fast scrambling conjecture" [3] states that the fastest scramblers take time logarithmic in the system size to scramble information, and that black holes are the fastest scramblers.

Scrambling and similar notions play important roles in other areas of physics as well. For example, scrambling is closely related to many-body localization and quantum thermalization (see [8] for a recent review): quantum systems that exhibit localization clearly do not scramble or thermalize, since local quantum information may fail to spread, and so remains accessible to certain local measurements. By contrast, a many-body system that undergoes scrambling evolves to states that appear random with respect to local measurements: here, the notion of scrambling can be seen as a form of thermalization at infinite temperature. Quantum chaos is also a close relative of scrambling. Under chaotic dynamics, initially local operators grow to overlap with the whole system (the butterfly effect). That is, chaotic quantum systems are scramblers [9]. In particular, the behaviors of the so-called out-of-time-order (OTO) correlators can probe the growth of local perturbations.

Their role as diagnostics of chaos has led to the active application of OTO correlators to the study of scrambling [9–18] and many-body localization [19–21].

This work is mainly motivated by two key features of scrambling. First, scrambling of quantum information and the growth of entanglement go hand in hand: information initially present in local perturbations ends up being irretrievable by local or simple measurements even though closed-system (unitary) evolutions do not actually erase any information, since it gets encoded in global entanglement. Entanglement captures the nonclassical essence of scrambling, and could be a natural and powerful probe of scrambling properties. Second, scrambling is intimately connected to the generation of randomness. Loosely speaking, scrambling and chaos describe the phenomenon that the system is effectively randomized. Indeed, the effects of information scrambling such as local indistinguishability [22] and the decay of OTO correlators [9] can be achieved by random dynamics given by a random unitary channel drawn from the group-invariant Haar measure. A key idea of the seminal Hayden-Preskill work [2] is to use random dynamics to model the scrambling behaviors of black holes. However, such observations are essentially one-way: scrambling do not necessarily imply full randomness. As we shall further clarify, there is in fact a large gap of complexity between information scrambling and complete randomness. The notion of "scrambling" needs to be refined since it can correspond to vastly different randomness complexities.

The major goal of this paper is to connect these two features and lay the mathematical foundations of diagnosing the randomness complexities associated with scrambling by entanglement. This is achieved by studying the interplay between the degrees of entanglement and quantum randomness. Note that studies along this line are also of great interest to many areas in quantum information. A basic result in this direction is that the expected entanglement entropy of a Haar random pure state is almost maximal, which is usually known as the Page's theorem [23–26]. However, this result is not tight in the sense that there is a large gap between the complexities of the Haar randomness and entanglement entropy conditions: the complexity of the Haar measure (given by e.g. the optimal depth of local circuits that approximate it) grows exponentially in the number of qubits [27], while the near-maximal entanglement entropy only needs finite moments of the Haar measure, which have only polynomial complexity and can be efficiently implemented [28–31]. This also illustrates the separation between the loss of local information or information scrambling and Haar randomness as large entanglement entropy indicates that local information is spread out (which will be discussed in more detail later). The regime in between information loss and complete randomness is not well understood in the contexts of both the dynamical behaviors of scrambling or chaos, and the kinematic entanglement properties.

To fill this gap, we consider more stringent entanglement measures and pseudorandom ensembles of quantum states and processes. In particular, we analyze the generalized entanglement entropies of pseudorandom ensembles of pure states and unitary channels known as designs, both parametrized by an order index. Generalized entanglement entropies of order α are entropic functions of the α-th power of the reduced density matrix. The higher the order of the generalized entropy, the more sensitive that entropy is to nonuniformity (such as sharp peaks) in the spectrum of the density matrix and so the

harder it is to maximize. (A particular family known as the Rényi entropy is most ideal for our purpose.) An α-design is an ensemble of pure states or unitary operators whose first α moments are indistinguishable from the Haar random states or unitaries. The higher the order of the design, the better it emulates the completely random Haar distribution. We establish a strong connection between the order of the generalized entanglement entropies and the order of designs, in both the random unitary channel and random state settings. (We note that a recent paper [18] establishes a related connection between 2α-point OTO correlators and α-designs via frame potentials.) Our analysis indicates that α-designs induce almost maximal Rényi-α entanglement entropies, thereby tightening (in a complexity-theoretical sense) known results relating entanglement entropy and quantum randomness, such as Page's theorem for random states and similar results for random unitaries by Hosur/Qi/Roberts/Yoshida [9]. This result reveals a fine-grained hierarchy of randomness complexities between information and Haar scrambling defined relative to the moments of the Haar measure, and suggests Rényi entanglement entropies of the corresponding order as useful diagnostics. For example, if the Rényi-α entanglement entropy for some way of partitioning the system does not meet the maximality condition, then one can argue that the system has not reached the complexity of α-designs. Since our characterization of such complexities of designs rely on entropy, we also refer to the joint notions as "entropic scrambling/randomness complexities".

Interestingly, there cannot be infinitely many different orders of designs that can be separated by Rényi entanglement entropies. This is seen by analyzing the min entanglement entropy, i.e. the infinite order limit of Rényi entropy, which only depends on the largest eigenvalue and lower bounds all Rényi entropies. Large min entanglement entropy indicates that the entanglement spectrum is almost completely uniform, and therefore the local information is totally lost and the system looks completely random even if one has access to the whole reduced density matrix. That is, the system essentially becomes indistinguishable from being Haar random by entanglement. This corresponds to a strong form of information scrambling, which we call "max-scrambling". We show that the min entanglement entropy (and therefore all Rényi entanglement entropies) becomes almost maximal, for designs of an order that is only logarithmic in the dimension of the system. In terms of entanglement properties, there can be at most logarithmic "nontrivial" orders of designs or moments of the Haar measure. Designs of higher orders all behave like completely random and are essentially the same. This result leads to a strong estimate of the shortest max-scrambling time, which generalizes the fast scrambling conjecture, that max-scrambling can be achieved by physical dynamics in time roughly linear in the number of degrees of freedom.

Now we summarize the mathematical techniques and results more specifically. We first focus on the intrinsic scrambling and randomness properties of physical processes, which are represented by unitary channels. We map unitary channels to a dual state via the Choi isomorphism, and study the entanglement associated with this dual state. As in [9], we partition the input register of the Choi state into two parts, A and B, and the output register into C and D. Our results rely on the calculation of average $\text{tr}\{\rho_{AC}^{\alpha}\}$, the defining element of order-α entanglement entropies between AC and BD of the Choi

state. We mainly employ tools from combinatorics and Weingarten calculus to compute the Haar integrals of $\mathrm{tr}\{\rho_{AC}^\alpha\}$ in various cases, which are equal to the average over unitary α-designs due to their defining properties. The convexity of Rényi entropies in the trace term allows us to use these results to lower bound the Rényi entanglement entropies by Jensen's inequality. The asymptotic result is that the Rényi-α entanglement entropies for equal partitions averaged over unitary α-designs are almost maximal, or more precisely, at most smaller than the maximal value by a constant that is independent of the dimension and the order. In other words, the difference is vanishingly small. This conclusion relies on a lemma on the number of cycles associated with permutations. In other words, a random unitary sampled from a unitary α-design is very likely to exhibit nearly maximal Rényi-α entanglement entropies, which supports the idea of using Rényi-α entanglement entropies as witnesses of the complexity of α-designs. For finite dimensions, we also derive explicit bounds on the α-design-averaged Rényi-α entanglement entropy using modern tools developed for Haar integrals. It is natural to ask how robust the above results are against small deviations from exact unitary designs. We derive error bounds for two common but slightly different ways to define approximate unitary designs. The extreme cases are actually quite interesting. In particular, we find that finite-order designs are sufficient to maximize the entanglement entropy given by the Rényi entropy of infinite order, namely the min entropy. As mentioned above, we show that, rather surprisingly, unitary designs of an order that scales logarithmically in the dimension of the unitary induce min entanglement entropy that is at most a constant away from the maximum, which implies that they are already indistinguishable from Haar by the entanglement spectrum alone.

Then we study the mathematically more straightforward and more well-known problem of entanglement in random states. The main results are very analogous to those in the random unitary setting, but the derivations are simpler since there are only two subsystems involved. Most importantly, we show that (projective) α-designs exhibit almost maximal Rényi-α entanglement entropies, which can be regarded as a collection of tight Page's theorems. And similarly, designs of logarithmic order maximize the min entanglement entropy. In addition, we are able to obtain the following separation result which is not there yet in the unitary setting. We show by representation theory that there exist 2-designs whose Rényi entanglement entropies of higher orders are bounded away from the maximum. The existence of such 2-designs can be regarded as the indicator of a separation between the complexity of 2-designs and those of higher orders as diagnosed by Rényi entanglement entropies. The paper also includes several other results related to e.g. Rényi entropies, designs, and Weingarten calculus, which may be of independent interest. These mathematical results may find applications in many other relevant areas, such as quantum cryptography and quantum computing.

The paper is organized as follows. In section 2, we formally define the central concepts of this paper — the generalized quantum entropies, and projective and unitary designs. In sections 3 and 4, we study the Choi model of unitary channels and pure states respectively. We conclude in section 5 with open problems and some discussions on the connections and possible extensions of our results to several other topics. The appendix contains several peripheral results and technical tools. See e.g. [32] for a comprehensive introduction of

standard and soft notations of asymptotics (e.g. big-O and soft big-O) that will be used throughout this paper. This paper provides the technical details of the results in [1].

2 Preliminaries

The theme of this paper is to establish connections between generalized quantum entropies and quantum designs, which we shall formally introduce in this section.

2.1 Generalized quantum entropies

2.1.1 Definitions of unified and Rényi entropies

Some parametrized generalizations of the Shannon and von Neumann entropy, most importantly the Rényi and Tsallis entropies, are found to be useful in both classical and quantum regimes. Here we focus on the entropies defined on a quantum state represented by density matrix ρ living in a finite-dimensional Hilbert space. A unified definition of generalized quantum entropies is given in [33, 34]:

Definition 1 (Quantum unified entropies). The quantum unified (α, s)-entropy of a density matrix ρ is defined as

$$S_s^{(\alpha)}(\rho) = \frac{1}{s(1-\alpha)} \left[(\mathrm{tr}\{\rho^\alpha\})^s - 1 \right]. \tag{2.1}$$

The two parameters α and s are respectively referred to as the order and the family of an entropy. In this paper, we mostly care about the cases where α is a positive integer and s is a nonnegative integer.

The $\mathrm{tr}\{\rho^\alpha\}$ element plays a key role in this paper. Entropies specified by a certain order α are collectively called α entropies. The $\alpha \to 1$ limit gives the von Neumann entropy. By fixing s, one obtains a family of entropies parametrized by order α. We define the following function to be the characteristic function of an entropy:

$$f_s^{(\alpha)}(x) = -\frac{x^s - 1}{s(1-\alpha)}, \tag{2.2}$$

which is obtained by treating $\mathrm{tr}\{\rho^\alpha\}$ as the argument x. The convexity of characteristic functions is important to many of our results.

The most representative families of quantum entropies are Rényi (the limiting case $s \to 0$) and Tsallis ($s = 1$) entropies. In this work, we shall mostly focus on the Rényi entropies:

Definition 2 (Quantum Rényi entropies). The quantum Rényi-α entropy of a density matrix ρ is defined as

$$S_R^{(\alpha)}(\rho) = \frac{1}{1-\alpha} \log \mathrm{tr}\{\rho^\alpha\}. \tag{2.3}$$

For $\alpha = 0, 1, \infty$, $S_R^{(\alpha)}$ is singular and defined by taking a limit. $S_R^{(0)}(\rho) = \log \mathrm{rank}(\rho)$ is called the max/Hartley entropy; $S_R^{(1)} = -\mathrm{tr}\rho \log \rho$ is just the von Neumann entropy. The $s \to \infty$ limit, which is called the min entropy, is particularly important for our study:
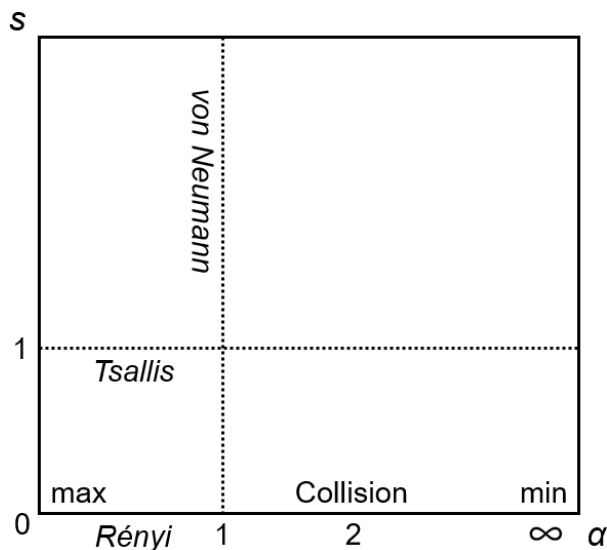
Figure 1. Unified (α, s)-entropies, $\alpha > 0, s \geq 0$. Italicized names refer to the whole line.

Definition 3 (Quantum min entropy). The quantum min entropy of a density matrix ρ is defined as

$$S_{\min}(\rho) = -\log \|\rho\| = -\log \lambda_{\max}(\rho), \tag{2.4}$$

where $\|\rho\|$ denotes the operator norm of ρ, and $\lambda_{\max}(\rho)$ is the largest eigenvalue of ρ.

Other Rényi entropies are well defined by eq. (2.3). The $\alpha = 2$ case $S_R^{(2)}(\rho) = -\log \mathrm{tr}\{\rho^2\}$, also called the second Rényi entropy or collision entropy for classical probability distributions, is also a widely used and highly relevant quantity. In the context of scrambling, a key result of [9] is that the Rényi-2 entanglement entropy is directly related to the 4-point OTO correlators, which has become a widely concerned quantity in recent years as a probe of chaos. Also notice that $S_R^{(2)}$ is directly related to the quantum purity $\mathrm{tr}\{\rho^2\}$ (recall that less pure subsystems dictate entanglement), and is thus frequently employed in the study of entanglement [35, 36].

Figure 1 summarizes the important generalized entropies in the relevant regime.

2.1.2 Important features of Rényi entropies

We are particularly interested in the family of Rényi entropies since they have several desirable features that play important roles in our arguments throughout.

The following properties of each Rényi entropy are important for our purposes:

1. They have the same maximal value n for systems of n qubits (attained by the uniform spectrum). This allows meaningful comparisons with the maximal value and between different orders;

2. They are additive on product states, i.e., $S_R^{(\alpha)}(\rho \otimes \sigma) = S_R^{(\alpha)}(\rho) + S_R^{(\alpha)}(\sigma)$ for all α and density matrices ρ, σ. Otherwise it is not natural to define extensive quantities such as mutual information and tripartite information;

3. Their characteristic functions $f_R^{(\alpha)}$ are convex, i.e., $S_R^{(\alpha)}(\rho)$ is convex in $\text{tr}\{\rho^\alpha\}$. This allows us to use Jensen's inequality to lower bound the design-averaged values by Haar integrals.

These properties are all straightforward to verify. These properties do not simultaneously hold for other families. For example, it is easy to see that the first two fail for Tsallis entropies. Later we shall further explain why these properties are desirable in explicit contexts. However, we note that the calculations are essentially only about the trace term, so it is straightforward to obtain results for all families if one wishes.

In this work, we are particularly interested in the regimes where certain Rényi entropies are nearly maximal. The following "cutoff" phenomenon concerning the maximality is an important foundation of our scheme of characterizing the complexity of scrambling by Rényi entropies. First, notice that the unified entropy of a certain family, such as the Rényi entropy, is monotonically nonincreasing in the order: $S_R^{(\alpha)} \geq S_R^{(\beta)}$ if $\alpha < \beta$. (In particular, the min entropy sets a lower bound on all Rényi entropies: $S_{\min} \leq S_R^{(\alpha)}$ for all α.) So if the Rényi entropy of some order is almost maximal, then those of lower orders are all almost maximal. Moreover, asymptotically, the values of Rényi entropies of different orders can be well separated, and for each order there exist inputs that attain almost maximal Rényi entropy of this order but those of all higher orders are small. As will become clearer later, this allows for the possibility of distinguishing between different complexities by the asymptotic maximality of Rényi entropies of certain orders. This feature can be illustrated by the following simple example. Given some order $\tilde{\alpha}$. Consider a density operator in the d-dimensional Hilbert space which has one large eigenvalue $1/d^{\frac{\tilde{\alpha}-1}{\tilde{\alpha}}}$, and the rest of the spectrum is uniform/degenerate. That is, the spectrum reads

$$\boldsymbol{\lambda} = \left(\frac{1}{d^{\frac{\tilde{\alpha}-1}{\tilde{\alpha}}}}, \underbrace{\frac{1 - \frac{1}{d^{\frac{\tilde{\alpha}-1}{\tilde{\alpha}}}}}{d-1}, \cdots, \frac{1 - \frac{1}{d^{\frac{\tilde{\alpha}-1}{\tilde{\alpha}}}}}{d-1}}_{d-1} \right). \tag{2.5}$$

The Rényi-$\tilde{\alpha}$ entropy (and thus all lower orders) is insensitive to this single peak:

$$S_R^{(\tilde{\alpha})}(\boldsymbol{\lambda}) = \frac{1}{1-\tilde{\alpha}} \log \left(\frac{1}{d^{\tilde{\alpha}-1}} + (d-1) \left(\frac{1 - \frac{1}{d^{\frac{\tilde{\alpha}-1}{\tilde{\alpha}}}}}{d-1} \right)^{\tilde{\alpha}} \right) \tag{2.6}$$

$$= \log d - \frac{1}{\tilde{\alpha}-1} \log \left(1 + \frac{d^{\tilde{\alpha}-1} \left(1 - \frac{1}{d^{\frac{\tilde{\alpha}-1}{\tilde{\alpha}}}} \right)^{\tilde{\alpha}}}{(d-1)^{\tilde{\alpha}-1}} \right) \tag{2.7}$$

$$\geq \log d - 1, \tag{2.8}$$

that is, $S_R^{(\tilde{\alpha})}(\boldsymbol{\lambda})$ is almost maximal, up to a small residual constant. However, the Rényi

entropies of higher orders can detect this peak and become small. For $\beta > \tilde{\alpha}$,

$$S_R^{(\beta)}(\boldsymbol{\lambda}) = \frac{1}{1-\beta} \log \left(\frac{1}{d^{\frac{\beta(\tilde{\alpha}-1)}{\tilde{\alpha}}}} + (d-1) \left(\frac{1 - \frac{1}{d^{\frac{\tilde{\alpha}-1}{\tilde{\alpha}}}}}{d-1} \right)^\beta \right) \approx \frac{\tilde{\alpha}\beta - \beta}{\tilde{\alpha}\beta - \tilde{\alpha}} \log d, \qquad (2.9)$$

which is $\Theta(\log d)$ (linear in the number of qubits) smaller than the maximal value $\log d$. In fact, $\boldsymbol{\lambda}$ produces $\Theta(\log d)$ gaps between all higher orders. The extreme case min entropy only cares about the largest eigenvalue by definition:

$$S_{\min}(\boldsymbol{\lambda}) = -\log \lambda_{\max} = \frac{\tilde{\alpha}-1}{\tilde{\alpha}} \log d, \qquad (2.10)$$

which is small for all finite $\tilde{\alpha}$. That is, the slope of $S_R^{(\beta)}(\boldsymbol{\lambda})$ in $\log d$ decreases with β. It equals one for $\beta = \tilde{\alpha}$, and approaches $\frac{\tilde{\alpha}-1}{\tilde{\alpha}}$ in the $\beta \to \infty$ limit. So there can be an asymptotic separation between Rényi entropies of any orders. The intuition is simply that promoting the power of eigenvalues essentially amplifies the nonuniformity of the spectrum. We shall construct a similar separation for certain 2-designs, which indicates that Rényi entropies can distinguish low-degree pseudorandom states from truly random states.

In our calculations we often assume equal partitions for simplicity. Since the subsystems contain half the total degrees of freedom, the equal partitions admit the largest possible entanglement entropy. Also, the following simple argument ensures that as long as the (Rényi) entanglement entropies between all equal partitions are close to the maximum, then that between generic partitions must be close to the maximum as well. Notice that the quantum Rényi divergence/relative entropy (either the non-sandwiched or sandwiched/non-commutative version, see e.g. [37] for definitions) between ρ and the maximally mixed state yields the gap between the Rényi entropy of ρ and the maximum:

$$D_R^{(\alpha)}(\rho \| I/d) = \frac{1}{\alpha - 1} \log(\text{tr}\{\rho^\alpha\} d^{\alpha-1}) = \log d - S_R^{(\alpha)}(\rho). \qquad (2.11)$$

For sandwiched Rényi divergence with $\alpha \geq 1/2$ (which covers the parameter range of interest in this paper), it is shown in [38, 39] that the data processing inequality holds, which implies that the divergence is monotonically nonincreasing under partial trace. So the gap can only be smaller when we look at smaller subsystems.

In the appendix, we derive more properties of Rényi entropies, including inequalities relating different orders of Rényi entropies (appendix A), and a weaker form of subadditivity (appendix B). The above discussions on Rényi entropies are more or less tailored to our needs. We refer the interested readers to [37] for a more comprehensive discussion of the motivations and properties of quantum Rényi entropies and divergences. We also note that a close variant of the quantum Rényi entropy known as the "modular entropy", given by $\tilde{S}_R^{(\alpha)}(\rho) = \frac{1}{\alpha^2} \partial_\alpha(\frac{\alpha-1}{\alpha} S_R^{(\alpha)}(\rho))$, is found to be meaningful in the context of holography and admits a natural thermodynamic interpretation [40, 41].

2.2 Designs

In quantum information theory, the notion of t-designs characterizes distributions of pure states or unitary channels that mimic the uniform distribution up to the first t moments,

and so can be considered as good approximations to Haar randomness. Analogous classical notions such as t-wise independence and t-universal hash functions are also found to be very useful in computer science and combinatorics. We shall formally introduce the definitions of state and unitary designs relevant to this work in the following.

2.2.1 Complex projective designs

Complex projective t-designs, which we may call "t-designs" for short throughout the paper, are distributions of vectors on the complex unit sphere that are good approximations to the uniform distribution, or pseudorandom, in the sense that they reproduce the first t moments of the uniform distribution [42–44]. They are of interest in many research areas, such as approximation theory, experimental designs, signal processing, and quantum information. There are many equivalent definitions of exact designs (see [44] for an introduction). Here we mention a few that are directly relevant to the current study.

The canonical definition based on polynomials of vector entries will be directly used in deriving our results. Define $\text{Hom}_{(t,t)}(\mathbb{C}^d)$ as the space of polynomials homogeneous of degree t both in the coordinates of vectors in $\mathbb{C}^d$ and in their complex conjugates.

Definition 4 (t-designs by polynomials)**.** An ensemble ν of pure state vectors in dimension d is a (complex projective) t-design if

$$\mathbb{E}_\nu \, p(\psi) = \int p(\psi) \mathrm{d}\psi \quad \forall p \in \text{Hom}_{(t,t)}(\mathbb{C}^d), \tag{2.12}$$

where the integral is taken with respect to the (normalized) uniform measure on the complex unit sphere in $\mathbb{C}^d$.

The second definition, based on the frame operator, is also widely used. Let $\text{Sym}_t(\mathbb{C}^d)$ be the t-partite symmetric subspace of $(\mathbb{C}^d)^{\otimes t}$ with corresponding projector $P_{[t]}$. The dimension of $\text{Sym}_t(\mathbb{C}^d)$ reads

$$D_{[t]} = \binom{d+t-1}{t}. \tag{2.13}$$

Definition 5 (t-designs by frame)**.** The t-th frame operator of ν is defined as

$$\mathcal{F}_t(\nu) := D_{[t]} \, \mathbb{E}_\nu (|\psi\rangle\langle\psi|)^{\otimes t}, \tag{2.14}$$

and the t-th frame potential is

$$\Phi_t(\nu) := \text{tr}\left\{ \mathcal{F}_t(\nu)^2 \right\}. \tag{2.15}$$

The ensemble ν is a t-design if and only if $\mathcal{F}_t(\nu) = P_{[t]}$ or, equivalently, if $\Phi_t(\nu) = D_{[t]}$ [44].

The above definitions for exact designs are equivalent. However, they lead to slightly different ways to define approximate designs by directly considering the deviations from equality, which essentially represent different norms. We shall discuss the approximate designs in more detail later for error analysis.

2.2.2 Unitary designs

In analogy to complex projective t-designs, unitary t-designs are distributions on the unitary group that are good approximations to the Haar measure, in the sense that they reproduce the Haar measure up to the first t moments [44–48]. They also play key roles in many research areas, such as randomized benchmarking, data hiding, and decoupling. As in the case of state designs there are also many equivalent definitions of exact unitary designs (see [44]). Similarly, we formally define unitary designs by polynomials and frame operators/potentials.

Let $\mathrm{Hom}_{(t,t)}(\mathrm{U}(d))$ be the space of polynomials homogeneous of degree t both in the matrix elements of $U \in \mathrm{U}(d)$ and in their complex conjugates.

Definition 6 (Unitary t-designs by polynomials)**.** An ensemble ν of unitary operators in dimension d is a *unitary t-design* if

$$\mathbb{E}_\nu \, p(U) = \int \mathrm{d}U p(U) \quad \forall p \in \mathrm{Hom}_{(t,t)}(\mathrm{U}(d)), \tag{2.16}$$

where the integral is taken over the normalized Haar measure on $\mathrm{U}(d)$.

Definition 7 (Unitary t-designs by frame)**.** The t-th frame operator of ν is defined as

$$\mathcal{F}_t(\nu) := \mathbb{E}_\nu \left[U^{\otimes t} \otimes U^{\dagger \otimes t} \right], \tag{2.17}$$

and the t-th frame potential is

$$\Phi_t(\nu) := \mathrm{tr} \left\{ \mathcal{F}_t(\nu)^2 \right\} \tag{2.18}$$

The ensemble ν is a unitary t-design if and only if $\mathcal{F}_t(\nu) = \mathcal{F}_t(\mathrm{U}(d))$, where $\mathcal{F}_t(\mathrm{U}(d))$ is the tth frame operator of the unitary group $\mathrm{U}(d)$ with Haar measure [44]. In addition,

$$\Phi_t(\nu) \geq \gamma(t,d) := \int \mathrm{d}U |\, \mathrm{tr}\{U\}|^{2t}, \tag{2.19}$$

and the lower bound is saturated if and only if ν is a unitary t-design [44, 48, 49]. When $t \leq d$, which is the case we are mostly interested in,

$$\gamma(t,d) = t!. \tag{2.20}$$

Again, the definitions are equivalent for exact unitary designs, but lead to different ways to define approximate unitary designs, which we shall look into later.

3 Generalized entanglement entropies and random unitary channels

Unitary channels describe the evolutions of closed quantum systems. Here we study the entanglement and scrambling properties of random unitary channels, which directly motivates this work. As suggested by [9], we employ the Choi isomorphism to map a unitary channel to a dual state, and study scrambling by the relevant entanglement properties of this state. In this section, we first briefly introduce the Choi state model, and then present the explicit calculations of generalized entanglement entropies averaged over unitary designs. The results lead to an entropic notion of scrambling or randomness complexities, which we shall discuss in depth.

3.1 Model: entanglement in the Choi state

Ref. [9] proposed that one can use the negativity of the tripartite information associated with the Choi state of a unitary channel to probe information scrambling. The negative tripartite information is actually a measure of global entanglement that quantifies the degree to which local information in the input to the channel becomes non-local in the output. We first introduce the definitions and motivations of this formalism to set the stage.

The Choi isomorphism (more generally, the channel-state duality) is widely used in quantum information theory to study quantum channels as states. It says that a unitary operator U acting on a d-dimensional Hilbert space $U = \sum_{i,j=0}^{d-1} U_{ij} |i\rangle\langle j|$ is dual to the pure state

$$|U\rangle = \frac{1}{\sqrt{d}} \sum_{i,j=0}^{d-1} U_{ji} |i\rangle_{\text{in}} \otimes |j\rangle_{\text{out}}, \tag{3.1}$$

which is called the Choi state of U. Now consider arbitrary bipartitions of the input register into A and B, and the output register into C and D. Let d_A, d_B, d_C, d_D be the dimensions of subregions A, B, C, D respectively ($d_A d_B = d_C d_D = d$). One expects that, in a system that scrambles information, any measurement on local regions of the output cannot reveal much information about local perturbations applied to the input. In other words, the mutual information between local regions of the input and output $I(A:C)$ and $I(A:D)$ should be small. This suggests that the negative tripartite information

$$- I_3(A:C:D) := I(A:CD) - I(A:C) - I(A:D) \tag{3.2}$$

can diagnose scrambling, since it essentially measures the amount of information of A hidden nonlocally over the whole output register. Here $I(A:C) = S(A) + S(C) - S(AC)$ is the mutual information, which measures the total correlation between A and C. Since the input and output are maximally mixed due to unitarity, the four subregions are all maximally mixed. For example, here $I(A:C)$ is reduced to $\log d_A d_C - S(AC)$, so we only need to analyze the entanglement entropy $S(AC)$. Note that $-I_3$ can be reduced to the conditional mutual information $I(A:B|C)$ [50], which is a quantity of great interest in quantum information theory.

The Haar-averaged (completely random) values of the terms in the von Neumann $-I_3$ was computed in [9], as a baseline for scrambling. However, it is clear that a pseudorandom ensemble (such as 2-designs) can already reach these roof values [18], which indicates that information scrambling only corresponds to randomness of low complexity in contrast to Haar. We are going to generalize the above quantities in the Choi state model using generalized entropies $S_s^{(\alpha)}$. Since the maximally mixed states have maximal generalized entropies, we only need to analyze $S_s^{(\alpha)}(AC)$.

3.2 Relevant reduced density matrices of the Choi state

To calculate the generalized entanglement entropies, we first need to derive the moments of the reduced density matrix of AC and the expression for their traces.

By using individual indices for different subregions, we rewrite the Choi state in eq. (3.1) as

$$|U\rangle = \frac{1}{\sqrt{d}} \sum_{klmo} U_{mo,kl} |kl\rangle_{AB} \otimes |mo\rangle_{CD}, \tag{3.3}$$

where k, l, m, o are respectively indices for A, B, C, D. The corresponding density matrix is then

$$\rho_{ABCD} = |U\rangle\langle U| = \frac{1}{d} \sum_{\substack{klmo \\ k'l'm'o'}} U_{mo,kl} U^*_{m'o',k'l'} |kl\rangle_{AB}\langle k'l'| \otimes |mo\rangle_{CD}\langle m'o'|. \tag{3.4}$$

By tracing out BD, we obtain the reduced density matrix of AC:

$$\rho_{AC} = \frac{1}{d} \sum_{\substack{klmo \\ k'm'}} U_{mo,kl} U^*_{m'o,k'l} |k\rangle_A\langle k'| \otimes |m\rangle_C\langle m'|. \tag{3.5}$$

The entropy of ρ_{AC} measures the entanglement between AC and BD. In order to compute the generalized α entanglement entropies, we need to raise ρ_{AC} to the power α:

$$\rho_{AC}^\alpha = \frac{1}{d^\alpha} \sum_{\text{all indices}} U_{m_1 o_1, k_1 l_1} U^*_{m_2 o_1, k_2 l_1} U_{m_2 o_2, k_2 l_2} U^*_{m_3 o_2, k_3 l_2} \cdots$$
$$U_{m_\alpha o_\alpha, k_\alpha l_\alpha} U^*_{m_{\alpha+1} o_\alpha, k_{\alpha+1} l_\alpha} |k_1\rangle\langle k_{\alpha+1}| \otimes |m_1\rangle\langle m_{\alpha+1}|. \tag{3.6}$$

Therefore,

$$\text{tr}\{\rho_{AC}^\alpha\} = \frac{1}{d^\alpha} \sum_{\text{all indices}} U_{m_1 o_1, k_1 l_1} U^*_{m_2 o_1, k_2 l_1} U_{m_2 o_2, k_2 l_2} U^*_{m_3 o_2, k_3 l_2} \cdots U_{m_\alpha o_\alpha, k_\alpha l_\alpha} U^*_{m_1 o_\alpha, k_1 l_\alpha}. \tag{3.7}$$

This result can also take more concise operator forms:

$$\text{tr}\{\rho_{AC}^\alpha\} = \frac{1}{d^\alpha} \text{tr}\left\{(U \otimes U^*)^{\otimes\alpha} X_\alpha\right\} = \frac{1}{d^\alpha} \text{tr}\left\{(U \otimes U^\dagger)^{\otimes\alpha} Y_\alpha\right\}, \tag{3.8}$$

where

$$X_\alpha := \sum_{\text{all indices}} |m_1 o_1\rangle\langle k_1 l_1| \otimes |m_2 o_1\rangle\langle k_2 l_1| \otimes |m_2 o_2\rangle\langle k_2 l_2| \otimes |m_3 o_2\rangle\langle k_3 l_2| \otimes \cdots$$
$$\otimes |m_\alpha o_\alpha\rangle\langle k_\alpha l_\alpha| \otimes |m_1 o_\alpha\rangle\langle k_1 l_\alpha|, \tag{3.9}$$

$$Y_\alpha := \sum_{\text{all indices}} |m_1 o_1\rangle\langle k_1 l_1| \otimes |k_2 l_1\rangle\langle m_2 o_1| \otimes |m_2 o_2\rangle\langle k_2 l_2| \otimes |k_3 l_2\rangle\langle m_3 o_2| \otimes \cdots$$
$$\otimes |m_\alpha o_\alpha\rangle\langle k_\alpha l_\alpha| \otimes |k_1 l_\alpha\rangle\langle m_1 o_\alpha| = X_\alpha^{\Gamma_{\text{even}}}, \tag{3.10}$$

where Γ_{even} denotes partial transpose on even parties. Notice that $Y_\alpha Y_\alpha^\dagger = I$ so Y_α is unitary.

Other density matrices can be derived in a similar way. Again note that the input and output are maximally entangled due to unitarity, so all four individual subregions are maximally mixed.

3.3 Haar random unitaries

3.3.1 General trace formula

We first employ tools from random matrix theory, combinatorics, and in particular Weingarten calculus, to compute the Haar integrals of the trace term in generalized entanglement entropies.

It is known that the Haar-averaged value of each monomial of degree α can be written in the following form [51]:

$$\int \mathrm{d}U\, U_{i_1 j_1} \cdots U_{i_\alpha j_\alpha} U^*_{i'_1 j'_1} \cdots U^*_{i'_\alpha j'_\alpha} = \sum_{\sigma,\gamma \in S_\alpha} \delta_{i_1 i'_{\sigma(1)}} \delta_{j_1 j'_{\gamma(1)}} \cdots \delta_{i_\alpha i'_{\sigma(\alpha)}} \delta_{j_\alpha j'_{\gamma(\alpha)}} \mathrm{Wg}(d, \sigma\gamma^{-1}),$$
(3.11)

where S_α is the symmetric group of α symbols, and

$$\mathrm{Wg}(d, \sigma) = \frac{1}{(\alpha!)^2} \sum_{\lambda \vdash \alpha} \frac{\chi^\lambda(1)^2 \chi^\lambda(\sigma)}{s_{\lambda,d}(1, \cdots, 1)}$$
(3.12)

are called Weingarten functions of $\mathrm{U}(d)$. Here $\lambda \vdash \alpha$ means λ is a partition of α, χ^λ is the corresponding character of S_α, and s_λ is the corresponding Schur function/polynomial. Notice that $s_{\lambda,d}(1, \cdots, 1)$ is simply the dimension of the irrep of $\mathrm{U}(d)$ associated with λ. The Weingarten function can be derived by various tools in representation theory, such as Schur-Weyl duality [52, 53] and Jucys-Murphy elements [54]. Therefore, we obtain the following general result:

Theorem 1.
$$\int \mathrm{d}U \,\mathrm{tr}\,\{\rho_{AC}^\alpha\} = \frac{1}{d^\alpha} \sum_{\text{all indices}} \int \mathrm{d}U\, U_{m_1 o_1, k_1 l_1} U^*_{m_2 o_1, k_2 l_1} U_{m_2 o_2, k_2 l_2} U^*_{m_3 o_2, k_3 l_2}$$
(3.13)
$$\cdots U_{m_\alpha o_\alpha, k_\alpha l_\alpha} U^*_{m_1 o_\alpha, k_1 l_\alpha}$$
$$= \frac{1}{d^\alpha} \sum_{\sigma,\gamma \in S_\alpha} d_A^{\xi(\sigma\tau)} d_B^{\xi(\sigma)} d_C^{\xi(\gamma\tau)} d_D^{\xi(\gamma)} \mathrm{Wg}(d, \sigma\gamma^{-1}),$$
(3.14)

where $\xi(\sigma)$ is the number of disjoint cycles associated with σ,[1] and $\tau := (1\ 2\ \cdots\ \alpha)$ is the 1-shift (canonical full cycle).

One can easily recover the $\alpha = 2$ results given in [9] from eq. (3.14) as follows. The Weingarten functions for $\sigma \in S_2$ are

$$\mathrm{Wg}(d, \sigma) = \begin{cases} \frac{1}{d^2-1} & \sigma = I, \\ -\frac{1}{d(d^2-1)} & \sigma = (1\ 2). \end{cases}$$
(3.15)

---

[1]Every member of the symmetric group can be uniquely decomposed into a product of disjoint cycles (up to relabeling).

There are 4 terms corresponding to two different Weingarten functions:

$\sigma\gamma$	σ	γ	$\xi(\sigma\tau)$	$\xi(\sigma)$	$\xi(\gamma\tau)$	$\xi(\gamma)$
I	I	I	1	2	1	2
	(1 2)	(1 2)	2	1	2	1
(1 2)	I	(1 2)	1	2	2	1
	(1 2)	I	2	1	1	2

Plugging them into eq. (3.14) yields

$$\int \mathrm{d}U \mathrm{tr}\{\rho_{AC}^2\} = \frac{1}{d^2}\left(\frac{1}{d^2-1}\left(d_A d_B^2 d_C d_D^2 + d_A^2 d_B d_C^2 d_D\right) - \frac{1}{d(d^2-1)}\left(d_A d_B^2 d_C^2 d_D + d_A^2 d_B d_C d_D^2\right)\right) \tag{3.16}$$

$$\approx d_A^{-1}d_C^{-1} + d_B^{-1}d_D^{-1} - d^{-1}d_A^{-1}d_D^{-1} - d^{-1}d_B^{-1}d_C^{-1}, \tag{3.17}$$

which confirms eq. (66) of [9]. A series of results of [9] such as an $O(1)$ gap between the Haar-averaged and maximal Rényi-2 entanglement entropies are obtained based on this formula.

More generally, we have

$$\int \mathrm{d}U \left(\mathrm{tr}\{\rho_{AC}^\alpha\}\right)^s = \frac{1}{d^{s\alpha}} \sum_{\sigma,\gamma \in S_{s\alpha}} d_A^{\xi(\sigma\tau_{\alpha,s})} d_B^{\xi(\sigma)} d_C^{\xi(\gamma\tau_{\alpha,s})} d_D^{\xi(\gamma)} \mathrm{Wg}(d, \sigma\gamma^{-1}), \tag{3.18}$$

where $\tau_{\alpha,s} := \prod_{r=0}^{s-1}(\alpha r + 1\ \alpha r + 2\ \cdots\ \alpha(r+1))$ is the product of canonical full cycles on each of the s blocks with α symbols.

3.3.2 Large d limit asymptotics

We now analyze the asymptotic behaviors of generalized entanglement entropies in the $d \to \infty$ limit to provide a big picture. Later we shall introduce some non-asymptotic bounds that hold for general d. To simplify the analysis, we consider equal partitions $d_A = d_B = d_C = d_D = \sqrt{d}$ here, which delivers the main idea.

Trace. We first introduce a series of useful combinatorics lemmas, which play critical roles in the behavior of generalized entanglement entropies (in particular Rényi). These results are known in the contexts of random matrix theory and free probability theory. We refer the readers to [55] (cf. references therein) for a summary of related results or [56] for a textbook on the subject.

Lemma 2 (Cycle Lemma). $\xi(\sigma) + \xi(\sigma\tau) \leq \alpha + 1$ *for all* $\sigma \in S_\alpha$, *where* ξ *counts the number of disjoint cycles.*

This result can be obtained by combining Lemmas A.1 and A.4 of [55]. See appendix C for our proof by induction.

Lemma 3. *Let* $g(\alpha)$ *be the number of* $\sigma \in S_\alpha$ *that saturate the inequality in Lemma 2. Then* $g(\alpha) = \mathrm{Cat}_\alpha := 2\alpha!/\alpha!(\alpha+1)! = \frac{1}{\alpha+1}\binom{2\alpha}{\alpha}$, *i.e., the* α-*th Catalan number.*

This result follows from Lemmas A.4 and A.5 of [55]. Such permutations lie on the geodesic from identity to τ. The above lemmas guarantee that the gap between the Haar-averaged Rényi entropies and the maximum value is independent of the system size, as will become clear shortly. We note that Catalan numbers frequently occur in counting problems. The first few Catalan numbers are $1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, 16796, \cdots$. Some useful bounds on the Catalan numbers are derived in appendix D.

Corollary 4. $\xi(\sigma) + \xi(\sigma\tau_{\alpha,s}) \le s\alpha + s$ for all $\sigma \in S_{s\alpha}$. The number of $\sigma \in S_{s\alpha}$ that saturate the inequality is $g(\alpha, s) = g(\alpha)^s = \mathrm{Cat}_\alpha^s = \frac{1}{(\alpha+1)^s}\binom{2\alpha}{\alpha}^s$.

We also need the large d asymptotic behaviors of the Weingarten function:

Lemma 5 (Asymptotics of Wg [52, 57])**.** *Given $\sigma \in S_\alpha$ with cycle decomposition $\sigma = C_1 \cdots C_k$. Let $|\sigma|$ be the minimal number of factors needed to write σ as a product of transpositions. The Möbius function of σ is defined by*

$$\mathrm{Moeb}(\sigma) := \prod_{i=1}^{k}(-1)^{|C_i|}\mathrm{Cat}_{|C_i|}, \tag{3.19}$$

where Cat_n is the n-th Catalan number (defined in Lemma 3). (Note that $|C_i|$ here is often replaced by $|C_i| - 1$ in literature, where $|\cdot|$ means the length of the cycle.) Then, in the large d limit, the Weingarten function has the asymptotic behavior

$$d^{\alpha+|\sigma|}\mathrm{Wg}(d, \sigma) = \mathrm{Moeb}(\sigma) + O(d^{-2}). \tag{3.20}$$

Corollary 6. *We mainly need to distinguish the following two cases:*

- *$\sigma = I$: $|\sigma| = 0$ and $\mathrm{Moeb}(\sigma) = 1$, thus $\mathrm{Wg}(d, I) = d^{-\alpha} + O(d^{-(\alpha+2)})$;*

- *$\sigma \ne I$: $|\sigma| \ge 1$, thus $\mathrm{Wg}(d, \sigma) = O(d^{-(\alpha+|\sigma|)}) = O(d^{-(\alpha+1)})$.*

Some bounds on the Möbius function are derived in appendix E.

Now we are equipped to derive the asymptotic behaviors of the Haar-averaged traces, $\int \mathrm{d}U\,\mathrm{tr}\,\{\rho_{AC}^\alpha\}$:

Theorem 7. *For equal partitions ($d_A = d_B = d_C = d_D = \sqrt{d}$), in the large d limit,*

$$\int \mathrm{d}U\,\mathrm{tr}\,\{\rho_{AC}^\alpha\} = \mathrm{Cat}_\alpha d^{1-\alpha}(1 + O(d^{-1})), \tag{3.21}$$

$$\int \mathrm{d}U\,(\mathrm{tr}\,\{\rho_{AC}^\alpha\})^s = \mathrm{Cat}_\alpha^s d^{(1-\alpha)s}(1 + O(d^{-1})), \tag{3.22}$$

Proof. Starting from eq. (3.14), Theorem 1

$$\int \mathrm{d}U \,\mathrm{tr}\,\{\rho_{AC}^{\alpha}\}$$

$$=\frac{1}{d^{\alpha}}\sum_{\sigma}(d_A d_C)^{\xi(\sigma\tau)}(d_B d_D)^{\xi(\sigma)}\mathrm{Wg}(d,I) + \frac{1}{d^{\alpha}}\sum_{\sigma\neq\gamma}d_A^{\xi(\sigma\tau)}d_B^{\xi(\sigma)}d_C^{\xi(\gamma\tau)}d_D^{\xi(\gamma)}\mathrm{Wg}(d,\sigma\gamma^{-1})$$

(3.23)

$$=\frac{1}{d^{\alpha}}\sum_{\sigma}d^{\xi(\sigma\tau)+\xi(\sigma)}\mathrm{Wg}(d,I) + \frac{1}{d^{\alpha}}\sum_{\sigma}d^{(\xi(\sigma\tau)+\xi(\sigma))/2}\sum_{\gamma\neq\sigma}d^{(\xi(\gamma\tau)+\xi(\gamma))/2}\mathrm{Wg}(d,\sigma\gamma^{-1})$$ (3.24)

$$=\sum_{\sigma}d^{\xi(\sigma\tau)+\xi(\sigma)}(d^{-2\alpha}+O(d^{-(2\alpha+2)})) + \sum_{\sigma}d^{(\xi(\sigma\tau)+\xi(\sigma))/2}\sum_{\gamma\neq\sigma}d^{(\xi(\gamma\tau)+\xi(\gamma))/2}O(d^{-(2\alpha+1)})$$

(3.25)

$$=\mathrm{Cat}_{\alpha}d^{1-\alpha}(1+O(d^{-1})),$$ (3.26)

where the second line follows from the equal bipartition assumption, the third line follows from Lemma 5 and Corollary 6, and the fourth line follows from Lemmas 2, 3 and some simple scaling analysis. Similarly, the asymptotic behavior of $\int \mathrm{d}U\,(\mathrm{tr}\,\{\rho_{AC}^{\alpha}\})^s$ follows from Corollary 4. $\qquad\square$

$s > 0$ entropies. The calculations of $s > 0$ entropies (e.g. Tsallis) are straightforward, since the term $(\mathrm{tr}\{\rho^{\alpha}\})^s$ linearly appears in the definition. By Theorem 7, for positive integers α, s:

$$\int \mathrm{d}U S_s^{(\alpha)}(\rho_{AC}) = \frac{1}{s(1-\alpha)}\left(\int \mathrm{d}U(\mathrm{tr}\,\{\rho_{AC}^{\alpha}\})^s - 1\right) = \frac{1 - \mathrm{Cat}_{\alpha}^s d^{(1-\alpha)s}(1+O(d^{-1}))}{s(\alpha-1)}.$$

(3.27)

Notice that the maximum value of $S_s^{(\alpha)}$ for a d-dimensional state is (achieved by the maximally mixed state I/d)

$$S_s^{(\alpha)}(I/d) = \frac{1 - d^{(1-\alpha)s}}{s(\alpha-1)}.$$ (3.28)

So we see a gap between the Haar-averaged and the maximal value:

$$\bar{\Delta}S_s^{(\alpha)} := S_s^{(\alpha)}(I/d) - \int \mathrm{d}U S_s^{(\alpha)}(\rho_{AC}) = \frac{\mathrm{Cat}_{\alpha}^s(1+O(d^{-1})) - 1}{s(\alpha-1)}d^{(1-\alpha)s},$$ (3.29)

which is vanishingly small in d.

As mentioned above, $s > 0$ entropies are less ideal than Rényi entropies for our study since they do not exhibit the three nice properties. Here we elaborate on the resulting problems one by one:

1. We see from eq. (3.28) that the roof (maximally mixed) values of $s > 0$ entropies vary with the order α. Therefore, it does not make much sense to compare $s > 0$ entropies of different orders α or with the roof value, on which our entropic characterization of scrambling and randomness complexities and several other arguments rely.

2. The $s > 0$ entropies are not even additive on maximally mixed states. So the derived quantities of mutual information and tripartite information in terms of $s > 0$ entropies do not make good sense. Recall that all partitions are in the maximally mixed state $I/\sqrt{d}$. However, the generalized mutual information $\int dU I_s^{(\alpha)}(A : C)$ given by $I_s^{(\alpha)}(A : C) := S_s^{(\alpha)}(A) + S_s^{(\alpha)}(C) - S_s^{(\alpha)}(AC)$ is not directly given by $\bar{\Delta}S_s^{(\alpha)}$. Define

$$\delta_s^{(\alpha)} := 2S_s^{(\alpha)}(I/\sqrt{d}) - S_s^{(\alpha)}(I/d) = \frac{1}{s(\alpha - 1)}(1 - d^{(1-\alpha)s/2+1} + d^{(1-\alpha)s}), \quad (3.30)$$

then

$$\int dU I_s^{(\alpha)}(A : C) = 2S_s^{(\alpha)}(I/\sqrt{d}) - \int dU S_s^{(\alpha)}(\rho_{AC}) = \delta_s^{(\alpha)} + \bar{\Delta}S_s^{(\alpha)}, \quad (3.31)$$

which is dominated by the irrelevant $\delta_s^{(\alpha)}$ ($\bar{\Delta}S_s^{(\alpha)}$ is vanishingly small).

3. The characteristic function for $s \geq 1$ entropies are not convex (linear for Tsallis). Although Theorem 7 enables us to directly calculate the Haar-averaged $s > 1$ entanglement entropies, the nonconvexity prevents us from using Jensen's inequality to lower bound their design-averaged values.

Rényi entropy. Now we analyze the behaviors of the Rényi entropies, the $s \to 0$ limit. Compared to $s > 0$ entropies, the calculations of Rényi are trickier because of the logarithm, which nevertheless directly leads to the desirable properties — constant roof value, additivity, and convexity. We are able to establish the following result:

Theorem 8. *In the large d limit,*

$$\int dU S_R^{(\alpha)}(\rho_{AC}) \geq \log d - O(1), \quad (3.32)$$

Proof. By definition,

$$\int dU S_R^{(\alpha)}(\rho_{AC}) = \int dU f_R^{(\alpha)}(\text{tr}\{\rho_{AC}^\alpha\}), \quad (3.33)$$

where

$$f_R^{(\alpha)}(x) = \frac{1}{1-\alpha} \log x \quad (3.34)$$

is the characteristic function. Since

$$\frac{d^2 f_R^{(\alpha)}(x)}{dx^2} = \frac{1}{(\alpha - 1)\ln 2} \frac{1}{x^2} \geq 0 \quad (3.35)$$

when $\alpha > 1$, $f_R^{(\alpha)}(x)$ is convex. So

$$\int dU S_R^{(\alpha)}(\rho_{AC}) \geq f_R^{(\alpha)}\left(\int dU \text{tr}\{\rho_{AC}^\alpha\}\right) \quad (3.36)$$

by Jensen's inequality. We note that this Jensen's lower bound due to convexity ($\mathbb{E}\, f_R \geq f_R\,\mathbb{E}$) will be repeatedly used to establish bounds for Rényi entropies. Then according to eq. (3.21),

$$f_R^{(\alpha)}\left(\int dU\,\mathrm{tr}\,\{\rho_{AC}^{\alpha}\}\right) = \frac{1}{1-\alpha}\log\left(\int dU\,\mathrm{tr}\,\{\rho_{AC}^{\alpha}\}\right) = \frac{1}{1-\alpha}\log\left(\mathrm{Cat}_{\alpha}\,d^{1-\alpha}(1+O(d^{-1}))\right)$$

$$= \log d - \frac{1}{\alpha-1}\log\mathrm{Cat}_{\alpha} + O(d^{-1}) \geq \log d - \frac{2\alpha}{\alpha-1} + O(d^{-1}). \tag{3.37}$$

Notice that the Cycle Lemma guarantees that the leading correction term (the second term) is independent of d asymptotically. In fact

$$\frac{1}{\alpha-1}\log\mathrm{Cat}_{\alpha} \leq \frac{2\alpha}{\alpha-1} \leq 4 \quad \forall \alpha \geq 2. \tag{3.38}$$

In conclusion, in the limit of large d, we have

$$\int dU\, S_R^{(\alpha)}(\rho_{AC}) \geq f_R^{(\alpha)}\left(\int dU\,\mathrm{tr}\,\{\rho_{AC}^{\alpha}\}\right) = \log d - O(1). \tag{3.39}$$

$\square$

So the gap between the Haar-averaged and maximal value of S_R (the "residual entropy") is

$$\bar{\Delta}S_R^{(\alpha)} := \log d - \int dU\, S_R^{(\alpha)}(\rho_{AC}) \leq O(1). \tag{3.40}$$

That is, the average Rényi entanglement entropies of the Haar measure are only bounded by a constant from the maximum. Recall the discussion in section 2.1.2: this $O(1)$ gap holds for non-equal partitions as well. The result implies that a Haar random unitary typically has almost maximal Rényi entanglement entropies for any partition. Rigorous probabilistic arguments require more careful analysis using concentration inequalities, which we leave for future work.

Now consider the Rényi mutual information and tripartite information based on the entanglement entropy results. First, we can directly obtain

$$\int dU\, I_R^{(\alpha)}(A:C) = \log d - \int dU\, S_R^{(\alpha)}(\rho_{AC}) \leq O(1), \tag{3.41}$$

which is equal to $\bar{\Delta}S_R^{(\alpha)}$ by additivity. The results hold similarly for AD. That is, the Rényi mutual information between any two local regions of the input and output is vanishingly small compared to the system size. On the other hand, for any partition size, notice that

$$I_R^{(\alpha)}(A:CD) = S_R^{(\alpha)}(A) + S_R^{(\alpha)}(CD) - S_R^{(\alpha)}(ACD) \tag{3.42}$$

$$= S_R^{(\alpha)}(A) + S_R^{(\alpha)}(CD) - S_R^{(\alpha)}(B) \tag{3.43}$$

$$= \log d_A + \log d - \log d_B \tag{3.44}$$

$$= 2\log d_A, \tag{3.45}$$

where the second line follows from $S_R^{(\alpha)}(ACD) = S_R^{(\alpha)}(B)$ since the whole Choi state is pure, the third line follows from the fact that the three subregions involved are maximally mixed, and the fourth line follows from that $d_A d_B = d$. Under the equal partition assumption, $I_R^{(\alpha)}(A:CD) = \log d$. This is consistent with the fact that all information of A is kept in the whole output CD because of unitarity. As a result:

$$-I_{3R}^{(\alpha)}(A:C:D) := I_R^{(\alpha)}(A:CD) - I_R^{(\alpha)}(A:C) - I_R^{(\alpha)}(A:D) \geq \log d - O(1), \quad (3.46)$$

by plugging in all relevant terms. So the negative Rényi tripartite information of Haar scrambling is indeed close to the maximum. However, we note that the Rényi-α entropy is not subadditive except for $\alpha = 1$, thus $-I_{3R}^{(\alpha)}(A:C:D)$ is not necessarily nonnegative. A weaker form of subadditivity of Rényi entropies is given in appendix B.

3.3.3 Non-asymptotic bounds

Here we prove some explicit bounds on the Haar-averaged trace, Rényi entropies, and in particular the min entropy, in the non-asymptotic regime. These bounds sharpen the asymptotic results. Many useful lemmas are proved in the appendices.

Trace and Rényi entropies. We directly put the results of trace and Rényi entropies together. We need the following refined cycle lemma:

Lemma 9. *Suppose* $q := \alpha^3/(32d_B^2) < 1$, *and* $d_A \leq d_B$. *Then*

$$\sum_{\sigma \in S_\alpha} d_A^{\xi(\sigma\tau)} d_B^{\xi(\sigma)} \leq h(q)\mathrm{Cat}_\alpha d_A d_B^\alpha \leq \frac{4^\alpha h(q)}{\sqrt{\pi}\alpha^{3/2}} d_A d_B^\alpha, \quad (3.47)$$

where $h(q) = 1 + 2q/[3(1-q)]$.

Proof. Define $c_{\delta,\alpha}$ as the number of permutations in S_α with genus δ, that is,

$$c_{\delta,\alpha} := |\{\sigma \in S_\alpha | \xi(\sigma) + \xi(\sigma\tau) = \alpha + 1 - 2\delta\}|. \quad (3.48)$$

Note that $c_{0,\alpha}$ is the Catalan number Cat_α by Lemma 3. Then

$$c_{\delta,\alpha} \leq \frac{2}{3}\left(\frac{\alpha^3}{32}\right)^\delta \mathrm{Cat}_\alpha, \quad (3.49)$$

according to Lemma 39 in appendix G. As a consequence of this inequality and the assumption $d_A \leq d_B$,

$$\sum_{\sigma \in S_\alpha} d_A^{\xi(\sigma\tau)} d_B^{\xi(\sigma)} \leq \sum_{\sigma \in S_\alpha} d_A d_B^{\xi(\sigma)+\xi(\sigma\tau)-1} = \sum_{\delta=0}^{(n-1)/2} c_{\delta,\alpha} d_A d_B^{\alpha-2\delta} = c_{0,\alpha} d_A d_B^\alpha \sum_{\delta=0}^{(n-1)/2} \frac{c_{\delta,\alpha}}{\mathrm{Cat}_\alpha} d_B^{-2\delta}$$

$$\leq c_{0,\alpha} d_A d_B^\alpha \left[1 + \frac{2}{3}\sum_{\delta=1}^{(n-1)/2}\left(\frac{\alpha^3}{32d_B^2}\right)^\delta\right] \leq \mathrm{Cat}_\alpha d_A d_B^\alpha \left[1 + \frac{2}{3}\sum_{\delta=1}^{\infty} q^\delta\right] = \mathrm{Cat}_\alpha d_A d_B^\alpha \left[1 + \frac{2q}{3(1-q)}\right]$$

$$= h(q)\mathrm{Cat}_\alpha d_A d_B^\alpha \leq \frac{4^\alpha h(q)}{\sqrt{\pi}\alpha^{3/2}} d_A d_B^\alpha, \quad (3.50)$$

where the last inequality follows from Lemma 29 in appendix D, which sets an upper bound on the Catalan numbers. $\square$

In the following we still assume equal partitions so that $d_{AC} = d$ for simplicity. Recall that for generic partitions the residual entropy cannot be larger anyway. By Lemma 9, we can obtain the following non-asymptotic bounds for the Haar integrals of traces and Rényi entanglement entropies:

Theorem 10. *Suppose $d > \sqrt{6}\alpha^{7/4}$. Then*

$$\int dU \mathrm{tr}\{\rho_{AC}^\alpha\} \leq \frac{a_\alpha \mathrm{Cat}_\alpha d^{1-\alpha}}{8} \left(1 + \frac{2q}{3(1-q)}\right)\left(7 + \cosh\frac{2\alpha(\alpha-1)}{d}\right), \qquad (3.51)$$

$$\int dU S_R^{(\alpha)}(\rho_{AC}) \geq \log d - \frac{\log \mathrm{Cat}_\alpha}{\alpha-1} - \frac{\log\left[\frac{a_\alpha}{8}\left(1+\frac{2q}{3(1-q)}\right)\left(7+\cosh\frac{2\alpha(\alpha-1)}{d}\right)\right]}{\alpha-1}, \quad (3.52)$$

where $a_\alpha := \frac{1}{1 - \frac{6\alpha^{7/2}}{d^2}}$.

Proof. By eq. (3.14) (Theorem 1):

$$\int dU \mathrm{tr}\{\rho_{AC}^\alpha\} = \sum_{\sigma,\gamma\in S_\alpha} d^{[\xi(\sigma\tau)+\xi(\sigma)+\xi(\gamma\tau)+\xi(\gamma)]/2}\mathrm{Wg}(d,\sigma\gamma^{-1}) \qquad (3.53)$$

$$= \sum_{\zeta\in S_\alpha}\left[\sum_{\gamma\in S_\alpha} d^{[\xi(\zeta\gamma\tau)+\xi(\zeta\gamma)+\xi(\gamma\tau)+\xi(\gamma)]/2}\mathrm{Wg}(d,\zeta)\right] \qquad (3.54)$$

$$\leq \sum_{\zeta\in A_\alpha}\left[\sum_{\gamma\in S_\alpha} d^{[\xi(\zeta\gamma\tau)+\xi(\zeta\gamma)+\xi(\gamma\tau)+\xi(\gamma)]/2}\mathrm{Wg}(d,\zeta)\right] \qquad (3.55)$$

$$\leq \sum_{\zeta\in A_\alpha}\sum_{\gamma\in S_\alpha} d^{\xi(\gamma\tau)+\xi(\gamma)}\mathrm{Wg}(d,\zeta) \qquad (3.56)$$

$$\leq \mathrm{Cat}_\alpha d^{\alpha+1}\left(1+\frac{2q}{3(1-q)}\right)\sum_{\zeta\in A_\alpha}\mathrm{Wg}(d,\zeta) \qquad (3.57)$$

$$\leq \frac{a_\alpha\mathrm{Cat}_\alpha d}{8}\left(1+\frac{2q}{3(1-q)}\right)\left(7+\cosh\frac{2\alpha(\alpha-1)}{d}\right), \qquad (3.58)$$

where A_α is the set of even permutations, i.e. the alternating group. The first inequality follows from the fact that $\mathrm{Wg}(d,\zeta)$ is negative when ζ is an odd permutation, the second inequality follows from the Cauchy-Schwarz inequality, noting that $\sum_{\gamma\in S_\alpha} d^{\xi(\zeta\gamma\tau)+\xi(\zeta\gamma)} = \sum_{\gamma\in S_\alpha} d^{\xi(\gamma\tau)+\xi(\gamma)}$, the third inequality follows from Lemma 9, and the last inequality follows from Lemma 35 in appendix F. By plugging eq. (3.58) into eq. (3.14), we immediately obtain the trace result eq. (3.51). The Rényi result eq. (3.52) then follows from Jensen's inequality. ∎

We see that the leading terms indeed match the asymptotic results. The overall observation is similar: the Haar integrals of Rényi entanglement entropies are very close to the maximum for sufficiently large d.

To gain intuition, we compute $\int dU \text{tr}\{\rho_{AC}^\alpha\}$ for $\alpha = 2, 3$ based on the explicit formulas for Weingarten functions. When $\sigma \in S_2$,

$$\text{Wg}(d, \sigma) = \begin{cases} \frac{1}{d^2-1} & \sigma = I, \\ -\frac{1}{d(d^2-1)} & \sigma = (1\ 2). \end{cases} \tag{3.59}$$

When $\sigma \in S_3$,

$$\text{Wg}(d, \sigma) = \frac{1}{d(d^2-1)(d^2-4)} \begin{cases} d^2 - 2 & \sigma = I, \\ -d & \sigma = (1\ 2), \\ 2 & \sigma = (1\ 2\ 3). \end{cases} \tag{3.60}$$

Therefore,

$$\int dU \text{tr}\{\rho_{AC}^2\} = \frac{2}{d+1} \leq \frac{2}{d}, \tag{3.61}$$

$$\int dU \text{tr}\{\rho_{AC}^3\} = \frac{5d^3 - 7d^2 - 6d + 2}{d^2(d+1)(d^2-4)} \leq \frac{5}{d^2}. \tag{3.62}$$

Min entropy. The results so far only directly apply to positive integer α. The min entanglement entropy, which corresponds to the special limit $\alpha \to \infty$, plays a crucial role in our framework of scrambling complexities. Now we examine the Haar integral of the min entanglement entropy.

Theorem 11.

$$\int dU \|\rho_{AC}\| \leq \frac{m_d}{d}, \tag{3.63}$$

$$\int dU S_{\min}(\rho_{AC}) \geq \log d - \log m_d, \tag{3.64}$$

where $m_d := \min\{7,\ 4(8\sqrt{d})^{1/\sqrt{d}}\}$.

Proof. Suppose $d \leq 70$. Then we have

$$d \int dU \|\rho_{AC}\| \leq d \left(\int dU \text{tr}\{\rho_{AC}^3\} \right)^{1/3} \leq d \left(\frac{5d^3 - 7d^2 - 6d + 2}{d^2(d+1)(d^2-4)} \right)^{1/3} \leq 7. \tag{3.65}$$

Now suppose $d \geq 50$. Let $\alpha = \lceil \sqrt{d}/2 \rceil$. Then

$$\frac{6\alpha^{7/2}}{d^2} \leq \frac{2}{5}, \quad q = \frac{\alpha^3}{32d^2} \leq \frac{1}{210}, \quad \frac{2\alpha(\alpha-1)}{d} \leq \frac{2}{3}, \tag{3.66}$$

so that

$$a_\alpha \leq \frac{5}{3}, \quad h(q) \leq \frac{301}{300}, \quad \frac{1}{8}\left(7 + \cosh\frac{2\alpha(\alpha-1)}{d}\right) \leq \frac{103}{100}. \tag{3.67}$$

Consequently,

$$\int dU \text{tr}\{\rho_{AC}^\alpha\} \leq \left[\frac{a_\alpha}{8\sqrt{\pi}}\left(1 + \frac{2q}{3(1-q)}\right)\left(7 + \cosh\frac{2\alpha(\alpha-1)}{d}\right)\right]\frac{4^\alpha d^{1-\alpha}}{\alpha^{3/2}} \leq \frac{4^\alpha d^{1-\alpha}}{\alpha^{3/2}}, \tag{3.68}$$

and thus

$$d \int \mathrm{d}U \|\rho_{AC}\| \leq d \left(\int \mathrm{d}U \, \mathrm{tr}\{\rho_{AC}^{\alpha}\} \right)^{1/\alpha} \leq 4 \left(\frac{d}{\alpha^{3/2}} \right)^{1/\alpha} \leq 4 \left(\frac{d}{(\sqrt{d}/2)^{3/2}} \right)^{2/\sqrt{d}} = 4(8\sqrt{d})^{1/\sqrt{d}}.$$

(3.69)

The proof of eq. (3.63) is completed by observing that $4(8\sqrt{d})^{1/\sqrt{d}} > 7$ when $50 \leq d \leq 52$ and $4(8\sqrt{d})^{1/\sqrt{d}} < 7$ when $53 \leq d \leq 70$. Eq. (3.64) then follows from the convexity of $-\log$. We note that slightly lower m_d can in principle be obtained by computing to higher orders in eq. (3.65), which is nevertheless not important for the main idea. □

As d gets large, m_d approaches the limit 4, and $\log m_d$ approaches the limit 2. As an implication of Lemma 28 in appendix B, Theorem 11 with d replaced by d_{AC} also holds when the four subregions have different dimensions, as long as $d_A, d_C \leq \sqrt{d}$ and $d_{AC} \leq d$. The same remark also applies to Theorem 17 below.

Note that the above results essentially confirm the conjecture in [58] that a Haar random unitary U and its reshuffled matrix U^R are asymptotically free, and the conjecture in [59] (based on extensive numerical evidence) that U^R converges to the Ginibre ensemble (of random non-Hermitian matrices) so that their moments will be asymptotically given by the Catalan numbers and the distribution of their spectra will be described by the famous Marchenko-Pastur distribution.

3.4 Unitary designs and their approximates

3.4.1 Average over unitary designs

Now we state a key observation: the Haar integral of $\mathrm{tr}\{\rho_{AC}^{\alpha}\}$, the defining term of α entropies, only uses the first α moments of the Haar measure. In other words, pseudorandom unitary α-designs are indistinguishable from Haar random by $\mathrm{tr}\{\rho_{AC}^{\alpha}\}$. More explicitly, let ν_α be a unitary α-design ensemble, then we have

$$\int \mathrm{d}U U_{i_1 j_1} \cdots U_{i_\alpha j_\alpha} U_{i_1' j_1'}^* \cdots U_{i_\alpha' j_\alpha'}^* = \mathbb{E}_{\nu_\alpha} \left[U_{i_1 j_1} \cdots U_{i_\alpha j_\alpha} U_{i_1' j_1'}^* \cdots U_{i_\alpha' j_\alpha'}^* \right]$$

(3.70)

by definition. Therefore, all Haar integrals of $\mathrm{tr}\{\rho_{AC}^{\alpha}\}$ from section 3.3 (those derived from eqs. (3.14) and (3.21)) directly carry over to α-designs.

This observation is the essential basis for the order correspondence results and in turn the idea that α entropies can generically diagnose whether a scrambler is locally indistinguishable from random dynamics as powerful as α-designs. The Haar-averaged Tsallis-α entropies ($s = 1$) are exactly saturated by α-designs due to the linearity in $\mathrm{tr}\{\rho^{\alpha}\}$. However, as mentioned, we cannot make analogous arguments for $s > 1$: the exact saturation requires f_s and the Haar integral to commute asymptotically, which is not known to hold; and the lower bound following from Jensen's inequality does not hold since $f_{s>1}$ becomes concave. In contrast, the Rényi entropies can be lower bounded because of the convexity. Due to the importance of the Rényi entropies, we state this result as a theorem:

Lemma 12.
$$\mathbb{E}_{\nu_\alpha} \left[S_R^{(\alpha)}(\rho_{AC}) \right] \geq f_R^{(\alpha)} \left(\int \mathrm{d}U \, \mathrm{tr}\{\rho_{AC}^{\alpha}\} \right) = \frac{1}{1-\alpha} \log \left(\int \mathrm{d}U \, \mathrm{tr}\{\rho_{AC}^{\alpha}\} \right).$$

(3.71)

Proof.

$$\mathbb{E}_{\nu_\alpha}\left[S_R^{(\alpha)}\left(\rho_{AC}\right)\right] = \mathbb{E}_{\nu_\alpha}\left[f_R^{(\alpha)}\left(\mathrm{tr}\left\{\rho_{AC}^\alpha\right\}\right)\right] \geq f_R^{(\alpha)}(\mathbb{E}_{\nu_\alpha}[\mathrm{tr}\{\rho_{AC}^\alpha\}]) = f_R^{(\alpha)}\left(\int dU\,\mathrm{tr}\left\{\rho_{AC}^\alpha\right\}\right),$$

(3.72)

where the inequality follows from Jensen's inequality, and the last equality follows from the fact that ν_α is an α-design. □

The lemma enables us to use the Haar integrals of traces to lower bound the design-averaged Rényi entanglement entropies in all dimensions. By combining this lemma and Theorem 7, we directly see that the $O(1)$ upper bound on the residual Rényi-α entropy still holds:

Theorem 13. *In the large d limit,*

$$\mathbb{E}_{\nu_\alpha}\left[S_R^{(\alpha)}\left(\rho_{AC}\right)\right] \geq \log d - O(1).$$

(3.73)

This is a key result of this work. We conclude that Rényi-α entanglement entropies are very likely to be almost maximal when sampling from unitary α-designs, as well as from the Haar measure. This result establishes the correspondence between the order of Rényi entropy and the order of designs, and lays the basis for the notion of entropic scrambling complexities. The non-asymptotic bound in Theorem 10 carries over in a similar fashion:

Theorem 14. *Suppose $d > \sqrt{6}\alpha^{7/4}$. Then*

$$\mathbb{E}_{\nu_\alpha}\left[S_R^{(\alpha)}\left(\rho_{AC}\right)\right] \geq \log d - \frac{\log \mathrm{Cat}_\alpha}{\alpha - 1} - \frac{\log\left[\frac{a_\alpha}{8}\left(1 + \frac{2q}{3(1-q)}\right)\left(7 + \cosh\frac{2\alpha(\alpha-1)}{d}\right)\right]}{\alpha - 1}, \quad (3.74)$$

where $a_\alpha := \frac{1}{1 - \frac{6\alpha^{7/2}}{d^2}}$.

Later we analyze the min entanglement entropy of designs in particular, which leads to another main result.

3.4.2 Error analysis: approximate unitary designs

The above analysis is based on exact unitary designs, but in most contexts we need to deal with the approximate versions of them. How robust or sensitive are these results under small deviations from exact unitary designs? One would expect ensembles that are very close to exact unitary α-designs to maintain near-maximal Rényi-α entanglement entropies. A subtlety is that different ways of measuring the deviation may lead to inequivalent definitions of approximate unitary designs, in contrast to the exact case. Here we discuss the deviation bounds for two commonly used definitions of approximate unitary designs, based on polynomials and frame operators respectively. This error analysis will be directly useful for e.g. relating the entropic scrambling complexities to circuit depth.

First, the canonical definition of unitary designs by polynomials leads to the following measure of deviation:

Definition 8 (m-approximate unitary designs [29])**.** An ensemble ν is an ϵ-m-approximate unitary t-design ("m" represents monomial) if

$$\left| \int \mathrm{d}U q^{(k)}(U) - \mathbb{E}_\nu \left[q^{(k)}(U) \right] \right| \le \epsilon \quad \forall q^{(k)}, k \le t, \tag{3.75}$$

where $q^{(k)}(U) = U_{i_1 j_1} \cdots U_{i_k j_k} U^*_{m_1 n_1} \cdots U^*_{m_k n_k}$ is a monomial of degree k both in the entries of U and in their complex conjugates.

Note that the bound is on each monomial with unit constant factor, otherwise the difference can be arbitrarily amplified by including more terms or changing the constant.

Theorem 15. *Let ω_α be an ϵ-m-approximate unitary α-design. Then*

$$\mathbb{E}_{\omega_\alpha}[\mathrm{tr}\,\{\rho^\alpha_{AC}\}] \le \int \mathrm{d}U \mathrm{tr}\,\{\rho^\alpha_{AC}\} + d^\alpha \epsilon, \tag{3.76}$$

$$\mathbb{E}_{\omega_\alpha} \left[S_R^{(\alpha)} (\rho_{AC}) \right] \ge \frac{1}{1-\alpha} \log \left(\int \mathrm{d}U \mathrm{tr}\,\{\rho^\alpha_{AC}\} + d^\alpha \epsilon \right). \tag{3.77}$$

In the large d limit,

$$\mathbb{E}_{\omega_\alpha}[S_R^{(\alpha)} (\rho_{AC})] \ge \log d - O(1) - \frac{1}{(\alpha-1)\mathrm{Cat}_\alpha \ln 2} d^{2\alpha-1} \epsilon \left(1 + O\left(d^{-1}\right) \right). \tag{3.78}$$

Proof.

$$\mathbb{E}_{\omega_\alpha}[\mathrm{tr}\,\{\rho^\alpha_{AC}\}] - \int \mathrm{d}U \mathrm{tr}\,\{\rho^\alpha_{AC}\} \le \left| \int \mathrm{d}U \mathrm{tr}\,\{\rho^\alpha_{AC}\} - \mathbb{E}_{\omega_\alpha}[\mathrm{tr}\,\{\rho^\alpha_{AC}\}] \right| \le \frac{1}{d^\alpha} d^{2\alpha} \epsilon = d^\alpha \epsilon \tag{3.79}$$

by triangle inequality, since $\mathrm{tr}\,\{\rho^\alpha_{AC}\}$ is the sum of $d^{2\alpha}$ monomials according to eq. (3.7). Then

$$\mathbb{E}_{\omega_\alpha} \left[S_R^{(\alpha)} (\rho_{AC}) \right] \ge f_R^{(\alpha)} \left(\mathbb{E}_{\omega_\alpha} [\mathrm{tr}\,\{\rho^\alpha_{AC}\}] \right)$$

$$= \frac{1}{1-\alpha} \log \mathbb{E}_{\omega_\alpha}[\mathrm{tr}\{\rho^\alpha_{AC}\}] \ge \frac{1}{1-\alpha} \log \left(\int \mathrm{d}U \mathrm{tr}\,\{\rho^\alpha_{AC}\} + d^\alpha \epsilon \right), \tag{3.80}$$

where the first inequality follows from Jensen's inequality, and the second inequality follows from eq. (3.79) and the fact that $-\log$ is monotonically decreasing. We can then use the $\int \mathrm{d}U \mathrm{tr}\,\{\rho^\alpha_{AC}\}$ results to analyze the perturbation.

Most importantly, in the large d limit,

$$\frac{1}{1-\alpha} \log \left(\int \mathrm{d}U \mathrm{tr}\,\{\rho^\alpha_{AC}\} \right) - \mathbb{E}_{\omega_\alpha} \left[S_R^{(\alpha)} (\rho_{AC}) \right] \le -\frac{1}{1-\alpha} \log \left(1 + \frac{d^\alpha \epsilon}{\int \mathrm{d}U \mathrm{tr}\,\{\rho^\alpha_{AC}\}} \right) \tag{3.81}$$

$$\le \frac{1}{\alpha-1} \log \left(1 + \frac{1}{\mathrm{Cat}_\alpha} d^{2\alpha-1} \epsilon \left(1 + O\left(d^{-1}\right) \right) \right) \tag{3.82}$$

$$\le \frac{1}{(\alpha-1)\mathrm{Cat}_\alpha \ln 2} d^{2\alpha-1} \epsilon \left(1 + O\left(d^{-1}\right) \right), \tag{3.83}$$

where the second line follows from eq. (3.26) and the following analysis, and the third line follows from the inequality that $\ln(1+x) \le x$ when $x > -1$. Then we directly obtain eq. (3.78), which says that the error in $S_R^{(\alpha)}(\rho_{AC})$ scales at most as $O(d^{2\alpha-1}\epsilon)$. $\qquad\square$

Recall the other definition of exact designs by frame operators. The deviation of an ensemble from a unitary t-design can also be quantified by a suitable norm of the deviation operator

$$\Delta_t(\nu) := \mathcal{F}_t(\nu) - \mathcal{F}_t(\mathrm{U}(d)). \tag{3.84}$$

The operator norm and trace norm of $\Delta_t(\nu)$ are two common figures of merit. The latter choice is more convenient for the current study:

Definition 9 (FO-approximate unitary designs). Ensemble ν is a λ-FO-approximate unitary t-design (FO represents frame operator) if

$$\|\Delta_t(\nu)\|_1 \leq \lambda. \tag{3.85}$$

Note that this definition is very similar to the quantum tensor product expander (TPE) [61]. TPEs conventionally use the operator norm, and the deviation operators relate to each other by partial transposes (like operators X, Y in eqs. (3.9), (3.10)).

Here we can directly use the operator form of local density operators derived earlier to do an error analysis of FO-approximate unitary designs. Let ω_α be a λ-FO-approximate unitary α-design. We define $\tilde{\Delta}_\alpha$, and explicitly write out Δ_α:

$$\tilde{\Delta}_\alpha(\omega_\alpha) = \mathbb{E}_{\omega_\alpha}[(U \otimes U^\dagger)^{\otimes \alpha}] - \int \mathrm{d}U (U \otimes U^\dagger)^{\otimes \alpha}, \tag{3.86}$$

$$\Delta_\alpha(\omega_\alpha) = \mathbb{E}_{\omega_\alpha}[U^{\otimes \alpha} \otimes U^{\dagger \otimes \alpha}] - \int \mathrm{d}U U^{\otimes \alpha} \otimes U^{\dagger \otimes \alpha}. \tag{3.87}$$

Theorem 16. *Let ω_α be a λ-FO-approximate unitary α-design. Then*

$$\mathbb{E}_{\omega_\alpha}[\mathrm{tr}\,\{\rho_{AC}^\alpha\}] \leq \int \mathrm{d}U \mathrm{tr}\,\{\rho_{AC}^\alpha\} + \frac{1}{d^\alpha}\lambda, \tag{3.88}$$

$$\mathbb{E}_{\omega_\alpha}[S_R^{(\alpha)}(\rho_{AC})] \geq \frac{1}{1-\alpha} \log\left(\int \mathrm{d}U \mathrm{tr}\,\{\rho_{AC}^\alpha\} + \frac{1}{d^\alpha}\lambda\right). \tag{3.89}$$

In the large d limit,

$$\mathbb{E}_{\omega_\alpha}[S_R^{(\alpha)}(\rho_{AC})] \geq \log d - O(1) - \frac{1}{(\alpha-1)\mathrm{Cat}_\alpha \ln 2} d^{-1}\lambda \left(1 + O\left(d^{-1}\right)\right). \tag{3.90}$$

Proof.

$$\mathbb{E}_{\omega_\alpha}[\mathrm{tr}\,\{\rho_{AC}^\alpha\}] - \int \mathrm{d}U \mathrm{tr}\,\{\rho_{AC}^\alpha\} = \frac{1}{d^\alpha}\mathrm{tr}\{\tilde{\Delta}_\alpha(\omega_\alpha)Y_\alpha\} \leq \frac{1}{d^\alpha}\left\|\tilde{\Delta}_\alpha(\omega_\alpha)\right\|_1 \|Y_\alpha\|$$

$$= \frac{1}{d^\alpha}\left\|\tilde{\Delta}_\alpha(\omega_\alpha)\right\|_1 = \frac{1}{d^\alpha}\|\Delta_\alpha(\omega_\alpha)\|_1 \leq \frac{1}{d^\alpha}\lambda, \tag{3.91}$$

where the first inequality follows from Hölder's inequality, and the second line follows from the unitarity of Y_α defined by eq. (3.10). The large d limit calculation simply resembles the above. $\qquad\square$

The essential difference between the m- and FO-approximate unitary designs is that the deviation is measured by different norms [31]. Letting $\epsilon, \lambda = 0$ recovers equivalent definitions of exact designs. However, we can see from the asymptotic error bounds that

they pose constraints of different strengths. The ϵ-m-approximation condition is quite loose, in the sense that the deviation ϵ needs to be vanishingly small to guarantee that the residual entropy remains small. Or one could say that the Rényi entanglement entropy results can be very sensitive to this type of error. In contrast, the λ-FO-approximation condition is more stringent and suitable: the residual entropy remains $O(1)$ as long as $\lambda \leq O(d)$, which implies that the FO-approximation may be a more suitable scheme.

3.5 Hierarchy of entropic scrambling complexities

3.5.1 Scrambling complexities by Rényi entanglement entropy

As motivated in the introduction, we expect that there is a hierarchy of scrambling complexities that lie in between information scramblers and Haar random unitaries, with different levels of the hierarchy indexed by the order of unitary designs needed to mimic the scrambler. Our results in the above link the randomness complexity of designs and the maximality of Rényi entanglement entropies of the corresponding order. This suggests that we can use the generic maximality of Rényi-α entanglement entropy as i) a necessary indicator of the resemblance to an α-design, and ii) a diagnostic of the entanglement complexity of α-designs, or "α-scrambling". The basic logic is that if a supposedly random unitary dynamics does not produce nearly maximal Rényi-α entanglement entropy in all valid partitions, as α-designs must do, then it is simply not close to any unitary α-design. This strategy is not directly relevant to testing designs at the global level, but it can probe the typical behaviors of entanglement between local regions of designs. Recall that Rényi entropy is monotonically nonincreasing in the order, and all orders share the same roof value. So α-scrambling necessarily implies α'-scrambling, for $\alpha \geq \alpha'$. In scrambling dynamics, the Rényi-α entanglement entropy is expected to grow slower and saturate the maximum at a later time than Rényi-α' in general.

3.5.2 Extreme orders: min- and max-scrambling

Now we discuss the 1- and ∞-scrambling more carefully, which respectively correspond to the weakest and strongest entropic scrambling complexities, given by the low and high ends of Rényi entropies.

Recall that $\alpha \to 1$ gives the von Neumann entropy, which probes information scrambling. First notice that unitary 1-designs do not necessarily create nontrivial entanglement or scramble quantum information. For example, the ensemble of tensor product of Pauli operators acting on each qubit

$$\{P_1 \otimes P_2 \otimes \cdots \otimes P_n\}, \quad P = I, \sigma_x, \sigma_y, \sigma_z \tag{3.92}$$

forms a unitary 1-design [18]. However, this local Pauli ensemble clearly does not scramble in any sense, since it cannot create entanglement among qubits (so local operators do not grow). So any entanglement entropy will be zero. On the other hand, unitary 2-designs are sufficient to maximize Rényi-2 entropies, which lower bounds the corresponding von Neumann entropies. It is shown in [50] that there actually exists a clear gap between them. So information scrambling is strictly weaker than 2-scrambling, but on the other

hand strictly stronger than 1-designs. More precise characterizations may depend on the specific signatures of min-scrambling one is using, and require more careful analysis of designs and generalized entropies in the non-integer order regime, which remains largely unclear and is left for future work.

The other end of the spectrum is $\alpha \to \infty$, which leads to the min entropy $S_{\min}(\rho) = -\log \|\rho\| = -\log \lambda_{\max}(\rho)$. Large min entanglement entropy directly indicates that the spectrum of the reduced density matrix is almost completely uniform, since it only cares about the largest eigenvalue. As the example of $\boldsymbol{\lambda}$ in section 2.1 shows, the min entropy is extremely sensitive to even one small peak in the entanglement spectrum. So it can be regarded as the "harshest" entropy measure and the strongest entropic diagnostic of scrambling: if the min entanglement entropy is almost maximal, then the system must be very close to maximally entangled in any sense and we cannot effectively distinguish the scrambler from Haar random by any Rényi entanglement entropy. This corresponds to the highest entropic scrambling complexity in our framework and thus we call this "max-scrambling". We shall see in a moment that designs of sufficiently high orders are simply indistinguishable from the Haar measure (also in the random state setting) by studying the min entanglement entropy of designs, which implies that max-scrambling is not an infinitely strong condition.

3.5.3 Nontrivial moments and fast max-scrambling

Given the definition of max-scrambling by the min entanglement entropy, one may wonder if the full Haar measure is needed to achieve this strongest form of entropic scrambling. Here we answer this question in the negative: for a given dimension, only a finite number of moments (which scales logarithmically in the dimension) are needed to maximize the min entanglement entropy, which we call nontrivial moments.

First we note that the same Haar-averaged min entanglement entropy results in Theorem 11 hold if the average is taken over a unitary α-design with $\alpha \geq \lceil \sqrt{d}/2 \rceil$. The conclusion is clear from the proof when $d \geq 50$. When $17 \leq d \leq 49$, $\lceil \sqrt{d}/2 \rceil \geq 3$, so the conclusion also follows from the proof. The conclusion is obvious when $d \leq 7$. It remains to consider the case $8 \leq d \leq 16$, which means $\lceil \sqrt{d}/2 \rceil = 2$. Therefore, eq. (3.61) applies, so that

$$d \int \mathrm{d}U \|\rho_{AC}\| \leq d \left(\int \mathrm{d}U \operatorname{tr}\{\rho_{AC}^2\} \right)^{1/2} \leq \sqrt{2d} < 7. \tag{3.93}$$

Therefore, eqs. (3.63) and (3.64) hold.

We can further show that, in fact, a unitary $O(\log d)$-design is enough to achieve nearly maximal min entanglement entropy:

Theorem 17. *Let ν_α be a unitary α-design, where $1 \leq \alpha = \lceil \log d/a \rceil \leq \sqrt{d}/2$ and $a > 0$; then*

$$d \, \mathbb{E}_{\nu_\alpha} \|\rho_{AC}\| \leq 2^{2+a}, \tag{3.94}$$

$$\mathbb{E}_{\nu_\alpha} S_{\min}(\rho_{AC}) \geq \log d - 2 - a. \tag{3.95}$$

In particular, if $\alpha \geq \lceil \log d \rceil$, then

$$d\, \mathbb{E}_{\nu_\alpha} \|\rho_{AC}\| \leq 8, \tag{3.96}$$

$$\mathbb{E}_{\nu_\alpha} S_{\min}(\rho_{AC}) \geq \log d - 3. \tag{3.97}$$

Proof. If $1 \leq \alpha = \lceil \log d / a \rceil \leq \sqrt{d}/2$, then one can show that eq. (3.68) holds as in the proof of Theorem 11 even without additional restrictions. Therefore,

$$d\, \mathbb{E}_{\nu_\alpha} \|\rho_{AC}\| \leq d\, (\mathbb{E}_{\nu_\alpha}[\mathrm{tr}\{\rho_{AC}^\alpha\}])^{1/\alpha} \leq 4\Big(\frac{d}{\alpha^{3/2}}\Big)^{1/\alpha} \leq 4d^{1/\alpha} \leq 4d^{a/\log d} = 2^{2+a}, \tag{3.98}$$

which confirms eq. (3.94) and implies eq. (3.95).

Now suppose $\alpha = \lceil \log d \rceil$. If $\alpha \geq \lceil \sqrt{d}/2 \rceil$, then eqs. (3.96) and (3.97) hold by Theorem 11 and the above analysis for unitary $\lceil \sqrt{d}/2 \rceil$-designs. Otherwise, if $\alpha \leq \sqrt{d}/2$, the two equations follow from eqs. (3.94) and (3.95) with $a = 1$. The same conclusion also holds when $\alpha \geq \lceil \log d \rceil$. $\qquad\square$

This result is crucial to the understanding and characterization of max-scrambling. In particular, the observation that log-designs can already achieve max-scrambling leads to an interesting argument about max-scrambling in physical dynamics. The studies of the dynamical scrambling behaviors of physical systems primarily care about the amount of time needed for the system to scramble under certain constraints. The fast scrambling conjecture [3] is the standard general argument about the limitation on this scrambling time, roughly saying that the fastest min-scramblers take $O(\log n)$ time, where $n \sim \log d$ is the number of degrees of freedom (and black holes, as in reason the most complex physical system and the fastest quantum information processor in nature, should achieve this bound).

Here we may ask similar questions for the complexities beyond min-scrambling: how fast can physical dynamics achieve certain scrambling complexities, in particular, max-scrambling? To make the assumption of "physical" more explicit, one typically requires the Hamiltonian governing the evolution to be local (meaning that each interaction term involves at most a finite number of degrees of freedom) and time-independent. Ref. [31] introduces the notion of design Hamiltonian, and conjectures that there are physical Hamiltonians that approximate unitary α-designs in time that scales roughly as $O(\alpha \log n)$. Note that the approximation scheme and error dependence will be important in translating it to the language of scrambling complexities. For example, for m-approximation error ϵ, an $\omega(\log \log(1/\epsilon))$ dependence is sufficient to dominate $\log n$ by the previous error analysis. Based on the above nontrivial moments result and the design Hamiltonian conjecture, the fastest max-scrambling time scales roughly as $O(n \log n)$. To absorb the non-primary effects, we state the conjecture using soft notations (absorbing polylogarithmic factors) as follows:

Conjecture 1 (Fast max-scrambling conjecture)**.** Max-scrambling can be achieved by physical dynamics in $\tilde{O}(n)$ time, i.e. in time roughly linear in the number of degrees of freedom.

To better formalize and study this fast max-scrambling conjecture, it would be important to further investigate the error dependency. Fast scrambling is an active research topic that has led to many key developments in quantum gravity and quantum many-body physics in recent years, such as the SYK model [62–64]. It could be interesting to generalize the studies about fast scrambling to this strong notion of max-scrambling.

3.5.4 On the gaps between entropic scrambling complexities

A further question then arises as to whether the entropic scrambling complexities form a strict hierarchy, i.e., whether different complexities are gapped.

A straightforward but strong definition of a separation between α- and α'-scrambling $(\alpha' < \alpha)$ is the following: there exist scramblers such that the associated Rényi-α' entropies are always near maximal, but some Rényi-α entropies can be bounded away from maximal. Such separations are in principle possible according to the properties of Rényi entropies (recall $\boldsymbol{\lambda}$). However, by the nontrivial moments result, we already know that $O(\log d)$ and higher complexities are not truly separated.

We tried several approaches to establish general separations in the Choi model, with limited success. In particular, we attempted to generalize the partially scrambling unitary model [50], and attempted to extend the gap results in the random state setting (next section) to random unitaries. The partially scrambling unitary model is used in [50] to prove a large separation between von Neumann and Rényi-2 tripartite information in the Choi state setting. By contrast, as we analyze in appendix H, this model is not likely to provide similar separations among generalized entropies. The analysis nevertheless reveals a rather interesting tradeoff between sensitivity and robustness between Rényi and $s > 0$ entropies. However, we are able to establish gaps using projective designs in the random state setting (see next section), but the results cannot be directly generalized to unitary designs. The reasons will be explained in more detail in the next section. We leave the gap problem in the Choi model open for the moment. We note, however, that the absence of strict separations of this type is not indicating that the behaviors of Rényi entropies (of sublogarithmic orders) are not separated in physical scenarios. We may still expect, for example, that the higher orders grow slower than lower orders, so that they still separate different complexities.

3.5.5 Relating to other complexities

It would be interesting to relate the entropic scrambling complexities to other traditional types of complexities, such as circuit complexity. For example, consider the local random circuit model. It is shown in [28, 29] that $O(\alpha^9 n[\alpha n + \log(1/\epsilon)])$ Haar random local gates are sufficient to form an ϵ-m-approximate α-design of n qubits. By the error analysis result, one can easily see that the minimum number of gates/circuit depth needed to maximize Rényi-α entropies scales polynomially in α and n: let $\epsilon = 2^{-3\alpha n}$ so that $\log(1/\epsilon) = 3\alpha n$, then the number of gates scales as $O(\alpha^{10} n^2)$, but meanwhile the deviation ϵ is sufficiently small such that the error in $S_R^{(\alpha)}(\rho_{AC})$ is vanishingly small, which indicates that such circuit is a good α-scrambler. That is, the entropic scrambling complexity and the random circuit complexity (minimum number of random gates) can be polynomially related. We note

that the $O(\alpha^{10} n^2)$ scaling can be improved to $O(\alpha n^2)$ for $\alpha = o(\sqrt{n})$ [31]. Moreover, the fast design and max-scrambling conjectures discussed in the last part can be regarded as connections to time complexity (in the physical sense).

4 Generalized entanglement entropies and random states

The previous section focused on Choi states, which are representations of the corresponding unitary channels. Here we consider a more straightforward problem — the entanglement in random and pseudorandom states — to generalize the connections between generalized entropies and designs. Note that the Page-like results, that a truly random state should typically be highly entangled, have been playing important roles in many fields including quantum gravity, quantum statistical mechanics, and quantum information theory for a long time. In this pure state setting, we obtain analogous main results that designs maximize corresponding Rényi entanglement entropies, closing the complexity gap in the Page's theorem, and that there are at most logarithmic nontrivial moments. These results suggest a similar hierarchy of entropic randomness complexities of states, which we call Page complexities. In addition, we are able to get solid results on the gap problem. We shall follow similar steps as in the random unitary setting, but with more focus on the different aspects. The presentation of similar arguments and derivations is going to be more compact.

4.1 Setting

The mathematical setting is as follows. Consider a bipartite system with Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, where $\mathcal{H}_A, \mathcal{H}_B$ have dimensions d_A, d_B, respectively, assuming $d_A \leq d_B$. We essentially need to compute the generalized entropies of the reduced density operator ρ_A. From here on we use $\mathbb{E}$ to denote the average over states drawn uniformly from the unit sphere in $\mathcal{H}$. Note that this uniform distribution on pure states is equivalent to the distribution generated by a Haar random unitary acting on some fixed fiducial state, so the induced uniformly random pure state is also called a Haar random state.

More explicitly, the Page's theorem (originally conjectured by Page in [23], proved in [24–26]) states that the average entanglement entropy of each reduced state is given by

$$\mathbb{E}\, S(\rho_A) = \mathbb{E}\, S(\rho_B) = \frac{1}{\ln 2} \left(\sum_{j=d_B+1}^{d_A d_B} - \frac{d_A - 1}{2 d_B} \right) > \log d_A - \frac{1}{2 \ln 2} \frac{d_A}{d_B} \geq \log d_A - \frac{1}{2 \ln 2}. \quad (4.1)$$

The gap between the average entropy and the maximum $\log d_A$ is bounded by the dimension-independent constant $1/(2 \ln 2)$. Similar observations were even earlier made by Lubkin [65] and Lloyd/Pagels [66]. In particular, [66] derived the distribution of the local eigenvalues of a random state, which may imply this result. Also see e.g. [67, 68] for further studies of this phenomenon. In the following we shall strengthen this result by proving the gap between the average Rényi-α entropy of each reduced state and the roof value $\log d_A$ is also bounded by a constant that is independent of the dimensions d_A, d_B and the order α.

4.2 Haar random states

Similarly, we first derive the integrals of the trace term and generalized entanglement entropies over the uniform measure.

4.2.1 General trace formula

Suppose $|\psi\rangle$ is drawn uniformly from the unit sphere in $\mathcal{H}$. The analytical formula for the average of the α-moment $\mathrm{tr}\{\rho_A^\alpha\}$, where ρ_A is the reduced density matrix of $|\psi\rangle$ for system A, is derived as follows. Expand $|\psi\rangle$ in the standard product basis $|\psi\rangle = \sum_{jk}\psi_{jk}|jk\rangle$, where $j = 1, 2, \ldots, d_A$ label the basis elements for $\mathcal{H}_A$, and $k = 1, 2, \ldots, d_B$ label the basis elements for $\mathcal{H}_B$. Then

$$\rho_A = \sum_{j_1,j_2,k} \psi_{j_1 k}\psi_{j_2,k}^* |j_1\rangle\langle j_2|. \tag{4.2}$$

The general result on the Haar-averaged trace is as follows:

Theorem 18.

$$\mathbb{E}\,\mathrm{tr}\{\rho_A^\alpha\} = \frac{1}{\alpha! D_{[\alpha]}} \sum_{\sigma \in S_\alpha} d_A^{\xi(\sigma\tau)} d_B^{\xi(\sigma)}, \tag{4.3}$$

where

$$D_{[\alpha]} = \binom{d_A d_B + \alpha - 1}{\alpha} = \frac{d_A d_B (d_A d_B + 1)\cdots(d_A d_B + \alpha - 1)}{\alpha!} \tag{4.4}$$

is the dimension of the symmetric subspace of $\mathcal{H}^{\otimes\alpha}$.

Proof. By eq. (4.2),

$$\mathrm{tr}\{\rho_A^\alpha\} = \sum_{\text{all indices}} \psi_{j_1,k_1}\psi_{j_2,k_1}^*\psi_{j_2,k_2}\psi_{j_3,k_2}^*\cdots\psi_{j_\alpha,k_\alpha}\psi_{j_1,k_\alpha}^* = \mathrm{tr}\left[(|\psi\rangle\langle\psi|)^\alpha Q_\alpha\right], \tag{4.5}$$

where

$$Q_\alpha = \sum_{\text{all indices}} |j_2 k_1\rangle\langle j_1 k_1| \otimes |j_3 k_2\rangle\langle j_2 k_2| \otimes \cdots \otimes |j_1 k_\alpha\rangle\langle j_\alpha k_\alpha|. \tag{4.6}$$

Therefore,

$$\mathbb{E}\,\mathrm{tr}\{\rho_A^\alpha\} = \frac{1}{D_{[\alpha]}} \mathrm{tr}\{P_{[\alpha]} Q_\alpha\}, \tag{4.7}$$

where $P_{[\alpha]}$ is the projector onto the symmetric subspace of $\mathcal{H}^{\otimes\alpha}$, and $D_{[\alpha]}$ is its dimension. Recall that the symmetric group S_α acts on $\mathcal{H}^{\otimes\alpha}$ by permuting the tensor factors, and $P_{[\alpha]}$ can be expressed as follows

$$P_{[\alpha]} = \frac{1}{\alpha!} \sum_{\sigma \in S_\alpha} U_\sigma, \tag{4.8}$$

where U_σ is the unitary operator associated with the permutation σ. Simple analysis shows that

$$\mathrm{tr}\{U_\sigma Q_\alpha\} = d_A^{\xi(\sigma\tau)} d_B^{\xi(\sigma)}. \tag{4.9}$$

Consequently,

$$\mathbb{E}\,\mathrm{tr}\{\rho_A^\alpha\} = \frac{1}{\alpha! D_{[\alpha]}} \sum_{\sigma \in S_\alpha} d_A^{\xi(\sigma\tau)} d_B^{\xi(\sigma)}. \tag{4.10}$$

$\square$

We noticed that similar results have been derived and rederived several times [65, 69–72]. Compared to known approaches, our approach seems simpler; in addition, it admits easy generalization to states drawn from (approximate) complex projective designs, which is not obvious for other approaches of which we are aware.

To get an intuitive understanding of eq. (4.3), it is worth taking a closer look at several concrete examples. When $\alpha = 2$, we reproduce a formula derived by Lubkin [65]:

$$\mathbb{E}\,\mathrm{tr}\{\rho_A^2\} = \frac{d_A + d_B}{d_A d_B + 1}. \tag{4.11}$$

From this equation we can derive a nearly-tight lower bound for the average Rényi-2 entanglement entropy,

$$\mathbb{E}\,S_R^{(2)}(\rho_A) \geq \log \frac{d_A d_B + 1}{d_A + d_B} > \log d_A - \log \frac{d_A + d_B}{d_B} \geq \log d_A - 1. \tag{4.12}$$

When $d_A = d_B$, the averages of the first few moments are given by

$$\mathbb{E}\,\mathrm{tr}\{\rho_A^2\} = \frac{2d_A}{d_A^2 + 1} \leq \frac{2}{d_A}, \tag{4.13}$$

$$\mathbb{E}\,\mathrm{tr}\{\rho_A^3\} = \frac{5d_A^2 + 1}{(d_A^2 + 1)(d_A^2 + 2)} \leq \frac{5}{d_A^2}, \tag{4.14}$$

$$\mathbb{E}\,\mathrm{tr}\{\rho_A^4\} = \frac{14d_A^3 + 10d_A}{(d_A^2 + 1)(d_A^2 + 2)(d_A^2 + 3)} \leq \frac{14}{d_A^3}, \tag{4.15}$$

which imply that

$$\mathbb{E}\,S_R^{(2)}(\rho_A) \geq \log d_A - 1, \tag{4.16}$$

$$\mathbb{E}\,S_R^{(3)}(\rho_A) \geq \log d_A - \frac{\log 5}{2}, \tag{4.17}$$

$$\mathbb{E}\,S_R^{(4)}(\rho_A) \geq \log d_A - \frac{\log 14}{3}. \tag{4.18}$$

Note that the gap of each Rényi entropy from the maximum is tied with the corresponding Catalan number. This is not a coincidence.

4.2.2 Large d limit

When $d_A = d_B \to \infty$, the asymptotic results go as follows:

Theorem 19. *In the limit of large d_A,*

$$\mathbb{E}\,\mathrm{tr}\{\rho_A^\alpha\} = \mathrm{Cat}_\alpha d_A^{-\alpha+1}(1 + O(d_A^{-2})). \tag{4.19}$$

$$\mathbb{E}\,S_R^{(\alpha)}(\rho_A) \geq \log d_A - \frac{2\alpha}{\alpha - 1} + O(d_A^{-2}) \geq \log d_A - O(1). \tag{4.20}$$

Proof. The trace result also follows from the Cycle Lemma:

$$\mathbb{E}\,\mathrm{tr}\{\rho_A^\alpha\} = \frac{1}{\alpha! D_{[\alpha]}} \sum_{\sigma \in S_\alpha} d_A^{\xi(\sigma\tau)+\xi(\sigma)} = \frac{\mathrm{Cat}_\alpha d_A^{\alpha+1} + O(d_A^{\alpha-1})}{d_A^{2\alpha} + O(d_A^{2\alpha-2})} = \mathrm{Cat}_\alpha d_A^{-\alpha+1}(1 + O(d_A^{-2})). \tag{4.21}$$

Therefore,

$$\mathbb{E} S_R^{(\alpha)}(\rho_A) \geq \frac{\log \mathrm{Cat}_\alpha d_A^{-\alpha+1}}{1-\alpha} + O(d_A^{-2}) = \log d_A - \frac{\log \mathrm{Cat}_\alpha}{\alpha-1} + O(d_A^{-2}) \geq \log d_A - \frac{2\alpha}{\alpha-1} + O(d_A^{-2}). \tag{4.22}$$

So the residual Rényi entropy is $O(1)$. $\qquad\square$

This theorem suggests that the gap between the average Rényi-α entropy and the maximum $\log d_A$ is bounded by a constant asymptotically.

4.2.3 Non-asymptotic bounds

The following bounds hold for any $d_A \leq d_B$:

Lemma 20. *Let* $q := \alpha^3/(32 d_B^2) < 1, h(q) := 1 + 2q/[3(1-q)]$. *Then*

$$\mathbb{E} \,\mathrm{tr}\{\rho_A^\alpha\} \leq h(q)\mathrm{Cat}_\alpha d_A^{1-\alpha} \leq \frac{4^\alpha h(q)}{\sqrt{\pi}\alpha^{3/2}} d_A^{1-\alpha}, \tag{4.23}$$

$$\mathbb{E} S_R^{(\alpha)}(\rho_A) \geq \log d_A - \frac{2\alpha - \frac{3}{2}\log\alpha + \log h(q) - \frac{1}{2}\log\pi}{\alpha-1}. \tag{4.24}$$

Proof. According to Lemma 9,

$$\mathbb{E} \,\mathrm{tr}\{\rho_A^\alpha\} = \frac{1}{\alpha! D_{[\alpha]}} \sum_{\sigma \in S_\alpha} d_A^{\xi(\sigma\tau)} d_B^{\xi(\sigma)} \leq \frac{h(q)\mathrm{Cat}_\alpha d_A d_B^\alpha}{d_A^\alpha d_B^\alpha} \leq h(q)\mathrm{Cat}_\alpha d_A^{1-\alpha} \leq \frac{4^\alpha h(q)}{\sqrt{\pi}\alpha^{3/2}} d_A^{1-\alpha}, \tag{4.25}$$

which in turn implies that

$$\mathbb{E} S_R^{(\alpha)}(\rho_A) \geq \frac{1}{1-\alpha} \log \mathbb{E} \,\mathrm{tr}\{\rho_A^\alpha\} \geq \frac{1}{1-\alpha} \log \frac{4^\alpha h(q)}{\sqrt{\pi}\alpha^{3/2}} d_A^{1-\alpha}, \tag{4.26}$$

$$= \log d_A - \frac{2\alpha - \frac{3}{2}\log\alpha + \log h(q) - \frac{1}{2}\log\pi}{\alpha-1}. \tag{4.27}$$

$\qquad\square$

In fact we can show that the gap is at most 2:

Theorem 21. *For all* $d_A \leq d_B$ *and* $\alpha \geq 0$,

$$\mathbb{E} S_R^{(\alpha)}(\rho_A) \geq \log d_A - 2. \tag{4.28}$$

Proof. Recall that Rényi-α entropy is nonincreasing with α, so to establish the theorem, it suffices to prove the lower bound for the min entropy. For all α,

$$S_R^{(\alpha)}(\rho_A) \geq S_{\min}(\rho_A) = -\mathbb{E} \log \|\rho_A\| \geq -\log \mathbb{E} \|\rho_A\| \geq -\log(\mathbb{E} \|\rho_A\|^\beta)^{1/\beta}$$
$$\geq -\log(4 d_A^{-1}) = \log d_A - 2, \tag{4.29}$$

where the second line follows from Lemma 22 below, by taking $0 < \beta \leq \lfloor (29 d_B^2)^{1/3} \rfloor$. $\qquad\square$

Lemma 22. *For all* $d_A \leq d_B$ *and* $0 < \alpha \leq \lfloor (29 d_B^2)^{1/3} \rfloor$,

$$(\mathbb{E} \|\rho_A\|^\alpha)^{1/\alpha} \leq 4 d_A^{-1}. \tag{4.30}$$

Proof. The conclusion is obvious when $d_A \leq 4$. When $d_B \geq d_A \geq 5$, note that $(\mathbb{E} \|\rho_A\|^\alpha)^{1/\alpha}$ is nondecreasing with α for $\alpha > 0$, so it suffices to prove the lemma in the case $\alpha = \lfloor 29 d_B^{2/3} \rfloor$. Then $\alpha^{3/2} \geq d_B \geq d_A$ and $0.6 < q = \alpha^3/(32 d_B^2) \leq 29/32$. According to Lemma 20,

$$\mathbb{E} \|\rho_A\|^\alpha \leq \mathbb{E} \operatorname{tr}\{\rho_A^\alpha\} \leq \frac{4^\alpha h(q)}{\sqrt{\pi} \alpha^{3/2}} d_A^{1-\alpha} = \left[\frac{d_A}{\sqrt{32\pi q} d_B} \left(1 + \frac{2q}{3(1-q)} \right) \right] 4^\alpha d_A^{-\alpha}$$

$$\leq \frac{3-q}{3(1-q)\sqrt{32\pi q}} 4^\alpha d_A^{-\alpha} \leq 4^\alpha d_A^{-\alpha}, \tag{4.31}$$

which implies the lemma. Here the last inequality follows from the observation that $f(q) := (3-q)/[3(1-q)\sqrt{32\pi q}] < 1$ for $0.6 < q \leq 29/32$. This fact can be verified immediately if we notice that the derivative $f'(q)$ has a unique zero at $q_0 = 4 - \sqrt{13}$ in the interval $0 < q < 1$ and that $f(q)$ is monotonically decreasing for $0 < q < q_0$ and monotonically increasing for $q_0 < q < 1$. $\square$

We also obtain the following bound, which improves Theorem 21 when $d_A \ll d_B$:

Theorem 23. *For all $d_A \leq d_B$ and $\alpha \geq 0$,*

$$\mathbb{E} S_R^{(\alpha)}(\rho_A) \geq \log d_A - 2\log \left(1 + \sqrt{\frac{d_A}{d_B}} \right) - \log c \geq \log d_A - \frac{2}{\ln 2} \sqrt{\frac{d_A}{d_B}} - \log c, \tag{4.32}$$

where $c = 1$ if $\mathcal{H}$ is real and $c = 2$ if $\mathcal{H}$ is complex.

Proof. The proof goes similarly as Theorem 21. For all $\alpha \geq 0$,

$$S_R^{(\alpha)}(\rho_A) \geq S_{\min}(\rho_A) = -\mathbb{E} \log \|\rho_A\| \geq -\log \mathbb{E} \|\rho_A\| \geq -\log \left(\mathbb{E} \sqrt{\|\rho_A\|} \right)^2$$

$$\geq \log d_A - 2\log \left(1 + \sqrt{\frac{d_A}{d_B}} \right) - \log c \geq \log d_A - \frac{2}{\ln 2} \sqrt{\frac{d_A}{d_B}} - \log c, \tag{4.33}$$

where $c = 1$ if $\mathcal{H}$ is real and $c = 2$ if $\mathcal{H}$ is complex. The second line follows from Lemma 24 stated below. $\square$

Lemma 24.

$$\mathbb{E} \sqrt{\|\rho_A\|} \leq \sqrt{c} \left(\frac{1}{\sqrt{d_A}} + \frac{1}{\sqrt{d_B}} \right), \tag{4.34}$$

where $c = 1$ if $\mathcal{H}$ is real and $c = 2$ if $\mathcal{H}$ is complex.

The proof of this lemma is rather complicated, so we leave it in appendix I. We believe that the constant c in Theorem 23 and Lemma 24 can be set to 1 in both real and complex cases. We note that Hayden and Winter had a similar result [73], but they are not so explicit about the constant and the dimensions for which their result is applicable.

4.3 State designs and their approximates

4.3.1 Average over designs and tight Page's theorems

Recall Page's theorem, which states that Haar-averaged von Neumann entanglement entropies of small subsystems are almost maximal. This theorem is not tight from the perspectives of both entropy and randomness: by the results above, the Haar-averaged Rényi entanglement entropies of higher orders are generically close to maximum as well, and the complete randomness is an overkill to maximize the entanglement entropies in terms of randomness complexity. Our results imply that Page's theorem can be strengthened from both sides. Similar to the random unitary setting, since $\mathbb{E}\operatorname{tr}\{\rho_A^\alpha\}$ only uses α moments of the uniform measure, all bounds on $\mathbb{E}\operatorname{tr}\{\rho_A^\alpha\}$ and $\mathbb{E}\,S_R^{(\alpha)}(\rho_A)$ from the last part still hold if the average is over α-designs. So we arrive at the following bounds that can be regarded as tight Page's theorems for each order α, by Theorems 21, 23:

Theorem 25 (Tight Page's theorems). *Let ν_α be an α-design. Then*

$$\mathbb{E}_{\nu_\alpha}\operatorname{tr}\{\rho_A^\alpha\} = \mathbb{E}\operatorname{tr}\{\rho_A^\alpha\}, \tag{4.35}$$

$$\mathbb{E}_{\nu_\alpha}\,S_R^{(\alpha)}(\rho_A) \ge f_R^{(\alpha)}\left(\mathbb{E}\operatorname{tr}\{\rho_A^\alpha\}\right) = \frac{1}{1-\alpha}\log\mathbb{E}\operatorname{tr}\{\rho_A^\alpha\}. \tag{4.36}$$

For all $d_A \le d_B$ and all $\alpha \ge 0$, the following bounds hold:

$$\mathbb{E}_{\nu_\alpha}\,S_R^{(\alpha)}(\rho_A) \ge \log d_A - 2, \tag{4.37}$$

and

$$\mathbb{E}_{\nu_\alpha}\,S_R^{(\alpha)}(\rho_A) \ge \log d_A - 2\log\left(1 + \sqrt{\frac{d_A}{d_B}}\right) - \log c \ge \log d_A - \frac{2}{\ln 2}\sqrt{\frac{d_A}{d_B}} - \log c, \tag{4.38}$$

where $c = 1$ if $\mathcal{H}$ is real and $c = 2$ if $\mathcal{H}$ is complex.

Obviously $\mathbb{E}_{\nu_\alpha}\,S_R^{(\alpha)}(\rho_A) \ge \log d_A - O(1)$ also hold in the limit of large d_A.

4.3.2 Approximate designs

Here we directly consider the more relevant notion of approximate α-designs given by deviation in frame operators. This error analysis is important for characterizing the randomness complexity by Rényi entropies, as will be explained later.

Given an ensemble ν of quantum states, define

$$\Delta_\alpha(\nu) := D_{[\alpha]}\,\mathbb{E}_\nu(|\psi\rangle\langle\psi|)^{\otimes t} - P_{[\alpha]}. \tag{4.39}$$

Definition 10 (FO-approximate designs). An ensemble ν is an λ-approximate α-design if

$$\|\Delta_\alpha(\nu)\|_1 \le \lambda. \tag{4.40}$$

Theorem 26. *Let ω_α be an λ-FO-approximate α-design with $\alpha \geq 2$. Then*

$$\mathbb{E}_{\omega_\alpha} \operatorname{tr}\{\rho_A^\alpha\} \leq \mathbb{E} \operatorname{tr}\{\rho_A^\alpha\} + \frac{\lambda}{D_{[\alpha]}}, \tag{4.41}$$

$$\mathbb{E}_{\omega_\alpha} S_R^{(\alpha)}(\rho_A) \geq \frac{1}{1-\alpha} \log\left(\mathbb{E} \operatorname{tr}\{\rho_A^\alpha\} + \frac{\lambda}{D_{[\alpha]}}\right). \tag{4.42}$$

In the large d_A limit,

$$\mathbb{E}_{\omega_\alpha} S_R^{(\alpha)}(\rho_A) \geq \log d_A - O(1) - \frac{1}{(\alpha-1)\mathrm{Cat}_\alpha \ln 2} \frac{d_A^{\alpha-1}\lambda}{D_{[\alpha]}}(1 + O(d_A^{-2})). \tag{4.43}$$

Proof. According to the same argument that leads to eq. (4.7),

$$\mathbb{E}_{\omega_\alpha} \operatorname{tr}\{\rho_A^\alpha\} = \operatorname{tr}\left\{\mathbb{E}_\nu(|\psi\rangle\langle\psi|)^{\otimes t}Q_\alpha\right\} = \frac{1}{D_{[\alpha]}} \operatorname{tr}\left\{(P_{[\alpha]} + \Delta_\alpha(\nu))Q_\alpha\right\},$$

$$= \mathbb{E} \operatorname{tr}\{\rho_A^\alpha\} + \frac{1}{D_{[\alpha]}} \operatorname{tr}\left\{\Delta_\alpha(\nu)Q_\alpha\right\} \leq \mathbb{E} \operatorname{tr}\{\rho_A^\alpha\} + \frac{1}{D_{[\alpha]}} \|\Delta_\alpha(\nu)\|_1 \|Q_\alpha\|$$

$$\leq \mathbb{E} \operatorname{tr}\{\rho_A^\alpha\} + \frac{\lambda}{D_{[\alpha]}}, \tag{4.44}$$

where the last inequality follows from the assumption $\|\Delta_\alpha(\nu)\|_1 \leq \lambda$ and the fact that $\|Q_\alpha\| = 1$, since Q_α is unitary. $\qquad\square$

We see that the residual entropy remains $O(1)$ as long as $\lambda/D_{[\alpha]} = O(d_A^{1-\alpha})$.

4.4 Hierarchy of Page complexities

4.4.1 Page complexities by Rényi entanglement entropy

Like the unitary case, our analysis of Rényi entanglement entropies lead to an entropic notion of randomness complexities: the complexity of α-designs can be witnessed by whether the average Rényi-α entanglement entropies are close enough to the maximum. Here we call them Page complexities as the foundation of this framework is the hierarchy of tight Page's theorems.

Here we provide an illustrating example based on the Clifford group. As an application of Lemma 26, let us consider the average Rényi entanglement entropy of Clifford orbits for a multiqubit system. For simplicity we assume $d_B = d_A \gg \alpha$, so that

$$\mathbb{E} \operatorname{tr}\{\rho_A^\alpha\} \approx \mathrm{Cat}_\alpha d_A^{1-\alpha}, \quad \mathbb{E} S_R^{(\alpha)}(\rho_A) \gtrsim \log d_A - \frac{\log \mathrm{Cat}_\alpha}{\alpha - 1}. \tag{4.45}$$

Recall that the Clifford group is a unitary 3-design [74, 75], so any orbit of the Clifford group forms a 3-design. Consequently, the average Rényi-α entanglement entropy for $\alpha \leq 3$ of any Clifford orbit is close to the maximum,

$$\mathbb{E}_{\mathrm{orb}(\psi)} \operatorname{tr}\{\rho_A^\alpha\} \approx \mathrm{Cat}_\alpha d_A^{1-\alpha}, \quad \mathbb{E}_{\mathrm{orb}(\psi)} S_R^{(\alpha)}(\rho_A) \gtrsim \log d_A - \frac{\log \mathrm{Cat}_\alpha}{\alpha - 1}, \tag{4.46}$$

for any ψ, where $\mathrm{orb}(\psi)$ denotes the Clifford orbit generated from ψ.

However, the Clifford group is not a 4-design, and Clifford orbits are in general not 4-designs [74–76]. If ψ is a stabilizer state, then $\|\Delta_4(\mathrm{orb}(\psi))\|_1 \approx d_A^6/12$ according to [44]. In this case the bounds for the fourth moment and Rényi-4 entropy provided by Theorem 26 is not very informative, note that $\mathbb{E}\,\mathrm{tr}\{\rho_A^4\} \approx 14d_A^{-3}$ and $D_{[4]} \approx (d_A d_B)^4/24 = d_A^8/24$. For a typical Clifford orbit, by contrast, $\|\Delta_\alpha(\nu)\|_1 \approx d_A^2$ is much smaller [44]. Now Theorem 26 implies that

$$\mathbb{E}_{\mathrm{orb}(\psi)}\,\mathrm{tr}\{\rho_A^4\} \leq \mathbb{E}\,\mathrm{tr}\{\rho_A^4\} + \frac{\|\Delta_\alpha(\mathrm{orb}(\psi))\|_1}{D_{[\alpha]}} \approx 14d_A^{-3} + 24d_A^{-6} \approx \mathbb{E}\,\mathrm{tr}\{\rho_A^4\}. \qquad (4.47)$$

Therefore, eq. (4.46) also holds for typical Clifford orbits when $\alpha = 4$. In our language, a Clifford orbit is very likely to have the Page complexity of 4-designs, although it is not really a 4-design in general. This is a rather nontrivial example indicating that the Page complexity is a necessary but not sufficient condition for certifying designs.

4.4.2 Nontrivial moments

Again, the min entanglement entropy witnesses the strongest Page complexity: if the average min entanglement entropies are always close to the maximum, then we simply cannot distinguish the ensemble from the completely random ensemble by the entanglement spectrum. The following theorem indicates that designs of order $O(\log d_A)$ maximize the min entanglement entropy and therefore achieve the max-Page complexity:

Theorem 27. *Suppose $|\psi\rangle$ is drawn from an α-design in a bipartite Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ of dimension $d_A \times d_B$, where $\alpha = \lceil(\log d_A)/a\rceil \leq (16d_B^2)^{1/3}$ with $0 < a \leq 1$. Let ρ_A be the reduced state of subsystem A. Then*

$$\mathbb{E}\,\|\rho_A\| \leq \frac{2^{2+a}}{d_A}, \qquad (4.48)$$

$$\mathbb{E}\,S_{\min}(\rho_A) \geq \log d_A - 2 - a. \qquad (4.49)$$

In particular, $\mathbb{E}\,\|\rho_A\| \leq 8/d_A$ and $\mathbb{E}\,S_{\min}(\rho_A) \geq \log d_A - 3$ if $\alpha = \lceil\log d_A\rceil$.

Proof. According to Lemma 20,

$$\mathbb{E}\,\mathrm{tr}\{\rho_A^\alpha\} \leq \frac{4^\alpha h(q)}{\sqrt{\pi}\alpha^{3/2}}d_A^{1-\alpha} \leq \frac{5 \times 4^\alpha}{3\sqrt{\pi}\alpha^{3/2}}d_A^{1-\alpha} \leq 4^\alpha d_A^{1-\alpha}, \qquad (4.50)$$

where the first inequality follows from the fact that $q = \alpha^3/(32d_B^2) \leq 1/2$ and $h(q) \leq 5/3$ given that $\alpha^3 \leq 16d_B^2$ by assumption. Consequently,

$$\mathbb{E}\,\|\rho_A\| \leq (\mathbb{E}\,\|\rho_A\|^\alpha)^{1/\alpha} \leq [\mathbb{E}\,\mathrm{tr}\{\rho_A^\alpha\}]^{1/\alpha} \leq d_A^{1/\alpha}\frac{4}{d_A} \leq d_A^{a/\log d_A}\frac{4}{d_A} = \frac{2^{2+a}}{d_A}, \qquad (4.51)$$

$$\mathbb{E}\,S_{\min}(\rho_A) \geq -\log\mathbb{E}\,\|\rho_A\| \geq -\log\frac{2^{2+a}}{d_A} \geq \log d_A - 2 - a. \qquad (4.52)$$

In the case, $a = 1$ and $\alpha = \lceil\log d_A\rceil$, the inequality $\alpha \leq (16d_A^2)^{1/3} \leq (16d_B^2)^{1/3}$ holds automatically; therefore, $\mathbb{E}\,\|\rho_A\| \leq 8/d_A$ and $\mathbb{E}\,S_{\min}(\rho_A) \geq \log d_A - 3$. $\qquad\square$

So again the hierarchy of distinguishable Page complexities can only extend to logarithmic designs.

4.4.3 Gaps between Page complexities

Following the definition of gaps between the entropic scrambling complexities, one may wonder here whether there exist α-designs such that Rényi entanglement entropies of orders larger than α are bounded away from the maximum, which we call "gap α-designs". In this random state setting, we are able to construct a family of gap 2-designs and so establish a strict gap between the second and α-th Page complexities with all $\alpha \geq 3$. Our construction is based on the orbits of a special subgroup of the unitary group on $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. As mentioned before, any orbit of a unitary 2-design is a complex projective 2-design. What is interesting, our construction of projective 2-designs does not require unitary 2-designs. In this way, we also provide a novel recipe for constructing projective 2-designs, which is particularly useful when the dimension is not a prime power.

Consider the group $G := U_A \otimes U_B$, where U_A, U_B are the unitary groups on $\mathcal{H}_A, \mathcal{H}_B$, respectively. It is irreducible, but does not form a 2-design. Simple analysis shows that G has four irreducible components on $\mathcal{H}^{\otimes 2}$, with dimensions $d_A d_B (d_A \pm 1)(d_B \pm 1)/4$, respectively. The symmetric subspace of $\mathcal{H}^{\otimes 2}$ contains two irreducible components with dimensions $d_A d_B (d_A + 1)(d_B + 1)/4$ and $d_A d_B (d_A - 1)(d_B - 1)/4$. By a similar continuity argument as employed in [44], there must exist an orbit of G that forms a 2-design. Let $|\psi\rangle$ be a fiducial vector of a 2-design with reduced state ρ_A for subsystem A. Then $\mathrm{tr}\{\rho_A^2\}$ is necessarily equal to the average over the uniform ensemble, that is,

$$\mathrm{tr}\{\rho_A^2\} = \frac{d_A + d_B}{d_A d_B + 1}. \tag{4.53}$$

It turns out that this condition is also sufficient. To see this, note that the condition must be invariant under local unitary transformations and thus only depends on a symmetric polynomial of the eigenvalues of ρ_A of degree 2, which is necessarily a function of $\mathrm{tr}\{\rho_A^2\}$ given the normalization condition $\mathrm{tr}\{\rho_A\} = 1$. It is worth pointing out that the same conclusion also holds if U_A, U_B are replaced by groups that form unitary 2-designs on $\mathcal{H}_A, \mathcal{H}_B$, respectively.

The following spectrum of ρ_A with one large eigenvalue is a solution of eq. (4.53):

$$\lambda_1 = \frac{d_A d_B + 1 + (d_A - 1)\sqrt{(d_A + 1)(d_A d_B + 1)}}{d_A(d_A d_B + 1)}, \tag{4.54}$$

$$\lambda_2 = \cdots = \lambda_{d_A} = \frac{d_A d_B + 1 - \sqrt{(d_A + 1)(d_A d_B + 1)}}{d_A(d_A d_B + 1)}. \tag{4.55}$$

If $d_B \geq d_A^2$, then

$$\lambda_1 \leq \frac{d_A^3 + 1 + (d_A - 1)\sqrt{(d_A + 1)(d_A^3 + 1)}}{d_A(d_A^3 + 1)} < \frac{2}{d_A}. \tag{4.56}$$

Therefore, $S_{\min}(\rho_A) \geq \log d_A - 1$, and the gap of all Rényi entropies from the maximum is bounded. The case in which the ratio d_B/d_A is bounded by a constant, say r, has very similar features to the $\alpha = 2$ single-peak spectrum discussed in section 2.1. We have

$$\lambda_1 \geq \frac{\sqrt{(d_A + 1)(d_A d_B + 1)}}{d_A d_B + 1} \geq d_B^{-1/2} \geq (r d_A)^{-1/2}. \tag{4.57}$$

Consequently,

$$S_R^{(\alpha)}(\rho_A) \leq \frac{1}{1-\alpha} \log \lambda_1^\alpha \leq \frac{1}{1-\alpha} \log(rd_A)^{-\alpha/2} = \frac{\alpha}{2(\alpha-1)}(\log d_A + \log r). \qquad (4.58)$$

As d_A increases, the gap of $S_R^{(\alpha)}(\rho_A)$ from the maximum is unbounded whenever $\alpha > 2$.

We note that such construction cannot be directly generalized to establish gaps in the Choi setting. As mentioned, any orbit of a unitary t-design is a complex projective t-design, but to construct a projective t-design, a unitary t-design is not required. Here the complex projective 2-design is constructed using a group that is a tensor product. However, such a group can never be a unitary 2-design. Also, in the Choi setting, four parties are involved, and it is not easy to ensure unitarity using the idea for constructing projective designs. New approaches are necessary for such a construction.

5 Concluding remarks

5.1 Summary and open problems

This paper explores the complexity of scrambling by connecting it to the degrees of quantum randomness via entanglement properties. In particular, we study the entanglement of state and unitary designs to lay the mathematical foundations for using Rényi and other generalized entanglement to probe the randomness complexities corresponding to designs, which we introduce as entropic scrambling complexities (or Page complexities in the state setting). These complexities form a hierarchy that spans in between the most basic notions of scrambling and the max-scrambling which mimics the entanglement properties of Haar. In summary, our results mainly establish the following key features of entropic scrambling complexities:

1. α-designs and close approximations induce almost maximal Rényi-α entanglement entropies. This basic result links the maximality of Rényi entanglement entropies and the design complexity of corresponding orders.

2. $O(\log d)$-designs are sufficient to maximize the min entanglement entropy, which means that they achieve the highest entropic scrambling complexity, namely max-scrambling. So all higher complexities collapse in the sense that they are simply indistinguishable from Haar scrambling by Rényi entanglement entropies.

3. We show that there exist projective 2-designs with non-maximal Rényi-3 (and therefore higher order) entanglement entropies. This establishes a strict separation between the order-2 complexity and higher levels, at least in the state setting.

The known structure of the entropic scrambling complexities based on our results is illustrated in figure 2. In summary, this study reveals the fine-grained complexity structure of the regime beyond information scrambling, and introduces a set of tools for studying it. We also hope that this work initiates further research into this significant but relatively unknown regime.
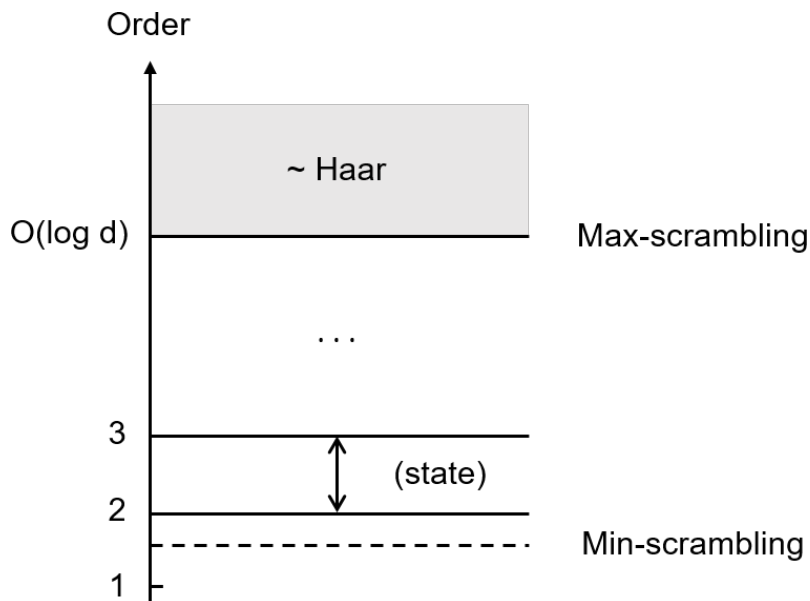
Figure 2. The hierarchy of entropic scrambling complexities. Each order is given by near-maximality of corresponding Rényi entanglement entropies, which diagnose the complexity of corresponding designs. The highest such complexity, corresponding to the notion of max-scrambling, is achieved at an order that roughly scales logarithmically in the dimension of the system. The weakest form of scrambling, or min-scrambling, is weaker than order-2 but stronger than order-1. Order-2 is strictly separated from higher orders in the state setting.

There are several technical open problems, especially in the setting of unitary channels. For example, we are not yet able to give a construction that opens a strict gap between the entropic scrambling complexities. Although we prove such gaps for projective 2-designs in the state setting, the similar techniques do not directly generalize to unitary channels. Moreover, due to the lack of subadditivity, we know that the negative tripartite information in terms of Rényi entropies are not necessarily nonnegative. It is worth looking into when this situation occurs, and further considering the meanings of such derived quantities. Furthermore, this paper mostly concerns the expected values. It would be important to further analyze the variances and derive probabilistic bounds using concentration inequalities, in order to talk about "typical" behaviors in a more rigorous sense.

5.2 Outlook

There are many interesting extensions to make. For example, our results suggest that Rényi entanglement entropies could be powerful tools to further advance the study of quantum randomness and pseudorandomness. A particularly interesting insight is that Rényi entropies of non-integer orders are naturally defined, which indicates that they can be helpful for characterizing and understanding the mysterious notion of designs of non-integer orders. This problem is of interest in quantum information, and as explained earlier, is key to a more precise characterization of the min-scrambling complexity. For example, it is reasonable to require that α-designs (where α can be non-integer) by any definition

must induce nearly maximal Rényi-α entanglement entropies. Then it is straightforward to see by eq. (4.58) that our gap 2-design induces small Rényi-$(2+\epsilon)$ entanglement entropy for any $\epsilon > 0$, and so cannot be a $(2+\epsilon)$-design. However, the attempts in properly defining non-integer designs and constructing such examples have mostly been negative so far. We tried a few possible ways to construct random ensembles such that the maximal-nonmaximal "phase transition" of Rényi entropy occurs at some non-integer order which do not work well. We also mention that the definition of designs by frame potential could be rather directly generalized to non-integer orders, but such generalization also suffers from fundamental problems.[2] We hope to give more well-behaved definitions or constructions of non-integer designs, or find more fundamental reasons that they are not meaningful notions — either of which is very interesting.

Also given that the entanglement properties of random states and channels play important roles in many areas in quantum information, including entanglement theory, quantum computing, and quantum cryptography, we expect the techniques and results here to find more interesting applications and advance the study of these fields. It is worth mentioning that the recent study of pseudorandom quantum states and unitaries from the perspective of computational indistinguishability [77], which represents a different notion of quantum pseudorandomness that is more directly related to the practical requirements for cryptographic security. It would be interesting to explore the role of entanglement in such computational quantum pseudorandomness, and find connections to our framework.

The current work focuses mostly on the kinematic or mathematical properties of unitary channels and states, which constitute a framework for further exploring the post-scrambling physics. For example, it would be interesting to study the dynamical behaviors of Rényi entanglement entropies and randomness, and in particular investigate fast max-scrambling, in specific many-body or holographic systems. By doing so we may extend existing studies of entanglement growth such as "entanglement tsunami" [78, 79]), and eventually understand the whole process of scrambling and especially its relation to randomness and complexity generation. In general, the study of randomness complexities may also shed new light on the fruitful idea of modeling complex systems (especially black holes [2]) by random states or dynamics. A recent study [80] on (a 1d variant of) the strongly chaotic SYK model (which has drawn considerable interest as a solvable toy model of quantum black holes and holography) shows that, after a quench, there is a "prethermal" regime where light modes rapidly scramble, but the Rényi entanglement entropies do not reach thermal values, which confirms our expectation that the randomness complexity of the system is still low. However, the late-time behaviors remain unclear. Another recent work [81] studies the Rényi entanglement entropies of random dynamics generated by Hamiltonians drawn from the Gaussian unitary ensemble (GUE). In general, the Rényi entanglement entropies are useful and analyzable quantities in the study of scrambling and chaos, and our work strengthens the motivation by connecting them to different randomness complexities.

We also hope to establish more solid connections between the randomness complexities and the conventional ones, such as computational, gate and Kolmogorov complexities,

---

[2]Learned from communications with Yoshifumi Nakata.

which play active roles in recent studies of holographic duality and black holes [82–84], and are of independent interest. Note that an interesting recent paper [84] directly concerns the evolution of complexity in generic physical dynamics. Here the complexity roughly means the computational/gate complexity, which is rather difficult to rigorously analyze. We feel that it is fruitful to combine their framework and viewpoints with ours.

Moreover, the notion of scrambling and randomness discussed here is associated with the entire Hilbert space. It would be nice to extend the techniques and results to the finite temperature regime or systems with conserved quantities, so as to apply our ideas in more physical scenarios and in general the study of quantum thermalization and many-body localization. We also hope to solidify the connections to several other relevant topics, including random tensor network holography [85] and OTO correlators. In summary, we believe that further research along the lines of research mentioned in this section could be essential to our understanding of quantum chaos, quantum statistical mechanics, quantum many-body physics, and quantum gravity.

A Inequalities relating Rényi entropies of different orders

First, we present a series of inequalities relating Rényi entropies of different orders. It is well known that the Rényi entropy is monotonically nonincreasing with the parameter α, that is $S_R^{(\alpha)}(\rho) \geq S_R^{(\beta)}(\rho)$ whenever $\beta \geq \alpha$. On the other hand, $S_R^{(\alpha)}(\rho)$ can also be used to construct a lower bound for $S_R^{(\beta)}(\rho)$ when $\beta \geq \alpha \geq 1$ as shown below,

$$
\begin{aligned}
S_R^{(\beta)}(\rho) &= -\frac{1}{\beta-1}\log\operatorname{tr}\{\rho^\beta\} = -\frac{\beta}{\beta-1}\log(\operatorname{tr}\{\rho^\beta\})^{1/\beta} \geq -\frac{\beta}{\beta-1}\log(\operatorname{tr}\{\rho^\alpha\})^{1/\alpha} \\
&= \frac{\beta}{\beta-1}\frac{\alpha-1}{\alpha}S_R^{(\alpha)}(\rho).
\end{aligned}
\tag{A.1}
$$

In particular, this equation yields a lower bound for the min entropy

$$
S_{\min}(\rho) \geq \frac{\alpha-1}{\alpha}S_R^{(\alpha)}(\rho) = S_R^{(\alpha)}(\rho) - \frac{S_R^{(\alpha)}(\rho)}{\alpha}.
\tag{A.2}
$$

When $\alpha \geq \log d$, we have

$$
S_R^{(\alpha)}(\rho) - 1 \leq S_{\min}(\rho) \leq S_R^{(\alpha)}(\rho),
\tag{A.3}
$$

so the difference between $S_R^{(\alpha)}(\rho)$ and $S_{\min}(\rho)$ is less than 1. When $\beta = \alpha + 1$, we have $S_R^{(\alpha+1)}(\rho) \geq \frac{\alpha^2-1}{\alpha^2} S_R^{(\alpha)}(\rho)$, so the difference between $S_R^{(\alpha+1)}(\rho)$ and $S_R^{(\alpha)}(\rho)$ is upper bounded by $S_R^{(\alpha)}(\rho)/\alpha^2$.

Next we derive another lower bound for $S_R^{(\beta)}(\rho)$ in terms of $S_R^{(\alpha)}(\rho)$ and the min entropy in the case $\beta \geq \alpha \geq 1$. The following equation

$$\text{tr}\{\rho^\beta\} = \text{tr}\left(\rho^\alpha \rho^{\beta-\alpha}\right) \leq \text{tr}\{\rho^\alpha\}\|\rho\|^{\beta-\alpha} \tag{A.4}$$

implies that

$$S_R^{(\beta)}(\rho) \geq \frac{1}{\beta-1}[(\alpha-1)S_R^{(\alpha)}(\rho) + (\beta-\alpha)S_{\min}(\rho)]. \tag{A.5}$$

In particular, any Rényi β-entropy with $\beta \geq 2$ is lower bounded by a convex combination of Rényi 2-entropy and the min entropy,

$$S_R^{(\beta)}(\rho) \geq \frac{1}{\beta-1}[S_R^{(2)}(\rho) + (\beta-2)S_{\min}(\rho)]. \tag{A.6}$$

B Weak subadditivity of the Rényi entropies

It is known that Rényi-α entropy is not subadditive except for the special case $\alpha = 1$. The following lemma yields a weaker form of subadditivity:

Lemma 28. *Let ρ_{AB} be any bipartite state on the product Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ with dimension $d_A \times d_B$. Let ρ_A, ρ_B be the two reduced states. Then*

$$\rho_{AB} \succ \rho_A \otimes \frac{I}{d_B}, \tag{B.1}$$

$$S_R^{(\alpha)}(\rho_{AB}) \leq S_R^{(\alpha)}(\rho_A) + \log d_B, \tag{B.2}$$

$$\log(d_A d_B) - S_R^{(\alpha)}(\rho_{AB}) \geq \log d_A - S_R^{(\alpha)}(\rho_A). \tag{B.3}$$

The first inequality in Lemma 28 means that the spectrum of ρ_{AB} majorizes that of $\rho_A \otimes \frac{I}{d_B}$. The second and third inequalities are immediate consequences of the first one, which are equivalent. The second one can be seen as a weaker form of subadditivity, while the third one means that the gap of Rényi entropy of a joint state from the maximum is no smaller than the corresponding gap for each reduced state, which has already been discussed in a slightly different way.

Proof. Let $|j\rangle$ for $j = 1, 2, \ldots, d_B$ be an orthonormal basis for $\mathcal{H}_B$ and $P_j = |j\rangle\langle j|$ be the corresponding projectors. Let

$$\sigma = \sum_j (I \otimes P_j)\rho_{AB}(I \otimes P_j) = \sum_j \rho_j \otimes P_j, \tag{B.4}$$

where ρ_j are subnormalized states that sum up to ρ_A. Define

$$\sigma_k = \sum_j \rho_j \otimes P_{j+k}, \quad k = 1, 2, \cdots, d_B, \tag{B.5}$$
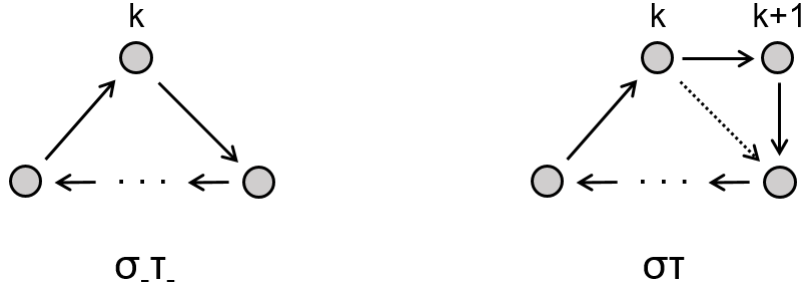
Figure 3. Comparison between $\sigma_-\tau_-$ and $\sigma\tau$, when $k+1$ is a 1-cycle in σ. Dashed arrows represent the mappings that are in $\sigma_-\tau_-$ but no longer there in $\sigma\tau$, and identical cycles are not shown. We see that in $\sigma\tau$ the element $k+1$ is inserted in the cycle, but the total number of cycles does not change.

where the addition in the indices is modulo d_B; note that $\sigma_0 = \sigma$. Then all σ_k have the same spectrum, which is majorized by ρ_{AB}, that is, $\rho_{AB} \succ \sigma_k$. Consequently,

$$\rho_{AB} \succ \frac{1}{d_B} \sum_{k=0}^{d_B-1} \sigma_k = \rho_A \otimes \frac{I}{d_B}. \tag{B.6}$$

Since Rényi α-entropy is Schur concave for $0 \leq \alpha \leq \infty$, it follows that

$$S_R^{(\alpha)}(\rho_{AB}) \leq S_R^{(\alpha)}\left(\rho_A \otimes \frac{I}{d_B}\right) = S_R^{(\alpha)}(\rho_A) + \log d_B \tag{B.7}$$

which confirms the second inequality in Lemma 28 and implies the third inequality.

$\square$

C Proof of the Cycle Lemma

We include here an intuitive proof of Lemma 2 (the Cycle Lemma), which plays a key role in our study, by induction. The intuition is that any element $\sigma \in S_\alpha$ can be viewed as a local deformation of some element $\sigma_- \in S_{\alpha-1}$, such that $\xi(\sigma) + \xi(\sigma\tau)$ can only increase by at most 1. We formalize the argument below.

Suppose the statement is true for $\alpha = k$. That is, $\xi(\sigma) + \xi(\sigma\tau) \leq k+1$ for all $\sigma \in S_k$. Now for some $\sigma \in S_{k+1}$, look at element $k+1$. There are two possibilities:

1. $k+1$ appears in a 1-cycle (is mapped to itself): $\sigma[k+1] = k+1$. So $\sigma = \sigma_-(k+1)$, for some $\sigma_- \in S_k$.

   $\xi(\sigma)$: we directly see $\xi(\sigma) = \xi(\sigma_-) + 1$.

   $\xi(\sigma\tau)$: write $\tau = (1\ 2\ \cdots\ k+1) = \tau_-(k\ k+1)$, where $\tau_- = (1\ 2\ \cdots\ k)$. Then $\sigma\tau = \sigma_-(k+1)\tau_-(k\ k+1) = \sigma_-\tau_-(k\ k+1)$, with $\sigma_-\tau_- \in S_k$. Now compare the action of $\sigma_-\tau_-$ and $\sigma\tau$ on individual elements. The only differences is $\sigma\tau[k] = k+1$ but $\sigma_-\tau_-[k] = \sigma_-\tau_-[k]$, and in addition $\sigma\tau[k+1] = \sigma_-\tau_-[k]$. So $\sigma\tau$ simply increases the length of a cycle in $\sigma_-\tau_-$ by one, and does nothing to other cycles. This is illustrated in figure 3. So $\xi(\sigma\tau) = \xi(\sigma_-\tau_-)$.

   From the induction hypothesis, $\xi(\sigma_-) + \xi(\sigma_-\tau_-) \leq k+1$, so $\xi(\sigma) + \xi(\sigma\tau) = \xi(\sigma_-) + 1 + \xi(\sigma_-\tau_-) \leq k+2$. Check.

2. $k + 1$ appears in a cycle of length > 1: $\sigma[a] = k + 1$, $\sigma[k + 1] = b$ for some elements $a, b \in \{1, \dots, k\}$. Define $\sigma'_- \in S_k$ by $\sigma'_-[i] = \sigma[i]$ for $i \in \{1, \dots, k\} \backslash \{a\}$ and $\sigma'_-[a] = b$.

   $\xi(\sigma)$: clearly $\xi(\sigma) = \xi(\sigma'_-)$.

   $\xi(\sigma\tau)$: compare the action of $\sigma'_-\tau_-$ and $\sigma\tau$ on individual elements. Depending on the value of a, there are two cases:

   (a) $a \neq 1$. The differences are: $\sigma'_-\tau_-[a - 1] = b$ and $\sigma'_-\tau_-[k] = \sigma[1]$, but $\sigma\tau[a - 1] = k + 1$, $\sigma\tau[k] = b$, and in addition $\sigma\tau[k + 1] = \sigma[1]$. They act identically on all other elements. There are two possible effects (see figure 4 for illustration):

      i. In $\sigma'_-\tau_-$, $\{a - 1, \, b\}$ and $\{k, \, \sigma[1]\}$ belong to the same cycle. Then $\sigma\tau$ breaks this cycle into two disjoint ones involving $\{a - 1, \, k + 1, \, \sigma[1]\}$ and $\{k, \, b\}$ respectively. So $\xi(\sigma\tau) = \xi(\sigma'_-\tau_-) + 1$;

      ii. In $\sigma'_-\tau_-$, $\{a - 1, \, b\}$ and $\{k, \, \sigma[1]\}$ belong to two disjoint cycles. Then $\sigma\tau$ glues these two cycles together into one. So $\xi(\sigma\tau) = \xi(\sigma'_-\tau_-) - 1$.

   (b) $a = 1$. Then $\sigma'_-\tau_-$ and $\sigma\tau$ act identically on $\{1, \dots, k\}$ and in addition $\sigma\tau[k + 1] = k + 1$. So $\xi(\sigma\tau) = \xi(\sigma'_-\tau_-) + 1$.

In conclusion, $\xi(\sigma\tau)$ can only increase by one or decrease by one as compared to $\xi(\sigma'_-\tau_-)$, so $\xi(\sigma) + \xi(\sigma\tau) = \xi(\sigma_-) + \xi(\sigma'_-\tau_-) \pm 1 \leq k + 2$ in either case. Check.

Lastly, consider $k = 1$. The only element of S_1 is (1), and $\xi((1)) + \xi((1)(1)) = 2 \leq k + 1$, so the statement trivially holds. This completes our proof.

D Bounds on the Catalan numbers

It is well known that the Catalan number $\text{Cat}_k = (2k)!/[k!(k + 1)!]$ is approximated by $4^k/\sqrt{\pi}k^{3/2}$ when k is large. To make this statement more precise, here we provide both lower and upper bounds for Cat_k.

Lemma 29. *The Catalan number Cat_k satisfies*

$$\frac{4^k}{\sqrt{\pi}(k + 1)^{3/2}} < \text{Cat}_k < \frac{4^k}{\sqrt{\pi}k^{3/2}} \quad \forall k \geq 1, \tag{D.1}$$
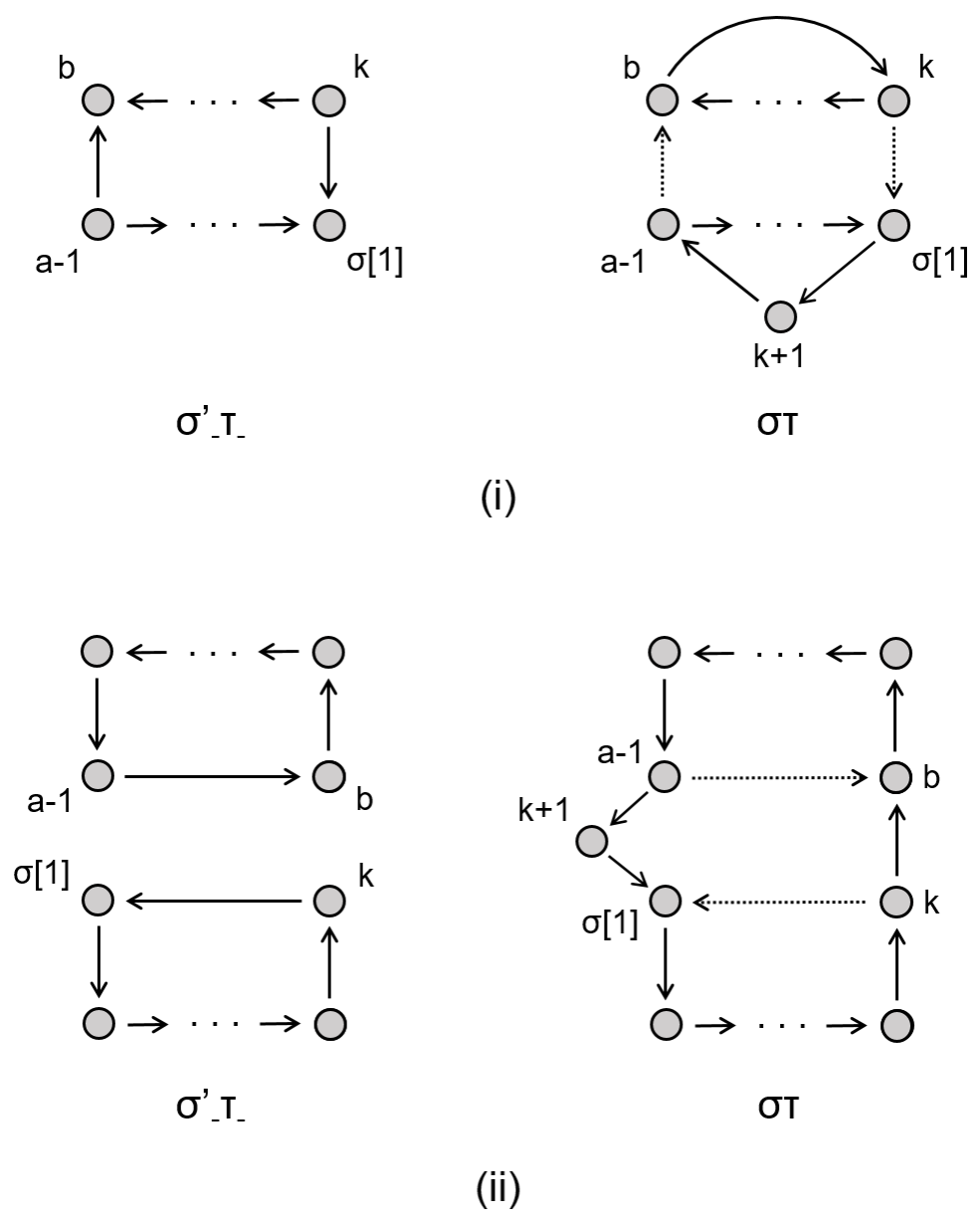
where k is not necessarily an integer.

Proof. The basis of our proof is the following Stirling approximation formula

$$\sqrt{2\pi}k^{k+\frac{1}{2}}e^{-k} \leq k! \leq \sqrt{2\pi}k^{k+\frac{1}{2}}e^{-k}e^{\frac{1}{12k}}. \tag{D.2}$$

As an implication,

$$\text{Cat}_k \leq \frac{\sqrt{2\pi}(2k)^{2k+\frac{1}{2}}e^{-2k}e^{\frac{1}{24k}}}{\sqrt{2\pi}k^{k+\frac{1}{2}}e^{-k}\sqrt{2\pi}(k + 1)^{k+\frac{3}{2}}e^{-k-1}} = \frac{2^{2k+\frac{1}{2}}k^k e^{1+\frac{1}{24k}}}{\sqrt{2\pi}(k + 1)^{k+\frac{3}{2}}}$$

$$= \frac{2^{2k}e^{1+\frac{1}{24k}}}{\sqrt{\pi}k^{\frac{3}{2}}(1 + \frac{1}{k})^{k+\frac{3}{2}}} < \frac{2^{2k}e^{\frac{1}{24k}}}{\sqrt{\pi}k^{\frac{3}{2}}(1 + \frac{1}{k})} < \frac{4^k}{\sqrt{\pi}k^{\frac{3}{2}}}. \tag{D.3}$$

Figure 4. Comparison between $\sigma'_-\tau_-$ and $\sigma\tau$, when $k+1$ is in a cycle of length >1 in σ. Dashed arrows represent the mappings that are in $\sigma'_-\tau_-$ but no longer there in $\sigma\tau$, and identical cycles are not shown. There are two possible cases: (i) The relevant elements $a-1, b, \sigma[1], k$ belong to the same cycle in $\sigma'_-\tau_-$. In $\sigma\tau$, this cycle is broken into two, so $\sigma\tau$ has one more cycle than $\sigma'_-\tau_-$; (ii) $\{a-1,\ b\}$ and $\{k,\ \sigma[1]\}$ belong to two cycles in $\sigma'_-\tau_-$. In $\sigma\tau$, these two cycles are combined as one with element $k+1$ inserted, so $\sigma\tau$ has one less cycle than $\sigma'_-\tau_-$.

Here the second inequality follows from the inequality

$$\left(1+\frac{1}{k}\right)^{k+\frac{1}{2}} > e, \tag{D.4}$$

note that the left hand side is monotonically decreasing with k and approaches e in the limit $k \to \infty$.

On the other hand,

$$\mathrm{Cat}_k \geq \frac{\sqrt{2\pi}(2k)^{2k+\frac{1}{2}}e^{-2k}}{\sqrt{2\pi}k^{k+\frac{1}{2}}e^{-k}e^{\frac{1}{12k}}\sqrt{2\pi}(k+1)^{k+\frac{3}{2}}e^{-k-1}e^{\frac{1}{12(k+1)}}} = \frac{2^{2k}k^k e}{\sqrt{\pi}(k+1)^{k+\frac{3}{2}}e^{\frac{1}{12k}+\frac{1}{12(k+1)}}}$$

$$= \frac{4^k e}{\sqrt{\pi}(k+1)^{\frac{3}{2}}(1+\frac{1}{k})^k e^{\frac{1}{12k}+\frac{1}{12(k+1)}}} > \frac{4^k}{\sqrt{\pi}(k+1)^{\frac{3}{2}}}. \tag{D.5}$$

Here the last inequality follows from the inequality

$$\left(1+\frac{1}{k}\right)^k e^{\frac{1}{12k}+\frac{1}{12(k+1)}} < e. \tag{D.6}$$

To confirm this claim, we shall prove the equivalent inequality

$$f(k) := \ln\left[\left(1+\frac{1}{k}\right)^k e^{\frac{1}{12k}+\frac{1}{12(k+1)}}\right] < 1. \tag{D.7}$$

The first and second derivatives of $f(k)$ read

$$f'(k) = \ln\left(\frac{k+1}{k}\right) - \frac{1}{k+1} - \frac{1}{12k^2} - \frac{1}{12(k+1)^2},$$

$$f''(k) = -\frac{1}{k(k+1)} + \frac{1}{(k+1)^2} + \frac{1}{6k^3} + \frac{1}{6(k+1)^3} = -\frac{1}{6k^3(k+1)^3}(4k^3+3k^2-3k-1) < 0. \tag{D.8}$$

Since $f''(k)$ is negative, $f'(k)$ is monotonically decreasing, which implies that $f'(k) > 0$ given that $\lim_{k\to\infty} f'(k) = 0$. Consequently, $f(k)$ is monotonically increasing, which confirms our claim $f(k) < 1$ given that $\lim_{k\to\infty} f(k) = 1$. $\qquad\square$

The following two corollaries are easy consequences of Lemma 29, though it is straightforward to prove them directly.

Corollary 30. $\mathrm{Cat}_k \leq \mathrm{Cat}_{k+1}$ *for any positive integer* k.

Proof. The corollary holds for $k = 0, 1$ by direct calculation. When $k \geq 2$, Lemma 29 implies that

$$\frac{\mathrm{Cat}_{k+1}}{\mathrm{Cat}_k} \geq \frac{4k^{3/2}}{(k+2)^{3/2}} \geq \frac{4}{2^{3/2}} = \sqrt{2} > 1, \tag{D.9}$$

which confirms the corollary. $\qquad\square$

Corollary 31. $\mathrm{Cat}_j\mathrm{Cat}_k < \mathrm{Cat}_{j+k}$ *for arbitrary positive integers* j, k.

Proof. The corollary holds when $j = 1$ or $k = 1$ according to Corollary 30, given that $\mathrm{Cat}_1 = 1$. When $j, k \geq 2$, Lemma 29 implies that

$$\frac{\mathrm{Cat}_j\mathrm{Cat}_k}{\mathrm{Cat}_{j+k}} < \frac{(j+k+1)^{3/2}}{\sqrt{\pi}j^{3/2}k^{3/2}} < 1. \tag{D.10}$$

$\qquad\square$

E Bounds on the Möbius function

Recall the definition of the Möbius function,

$$\text{Moeb}(\sigma) := \prod_{j=1}^{k} (-1)^{|C_j|} \text{Cat}_{|C_j|} = (-1)^{|\sigma|} \prod_{j=1}^{k} \text{Cat}_{|C_j|}. \tag{E.1}$$

Lemma 32.

$$1 \leq |\text{Moeb}(\sigma)| \leq \text{Cat}_{|\sigma|} < \frac{4^{|\sigma|}}{\sqrt{\pi} |\sigma|^{3/2}} \quad \forall |\sigma| \geq 1. \tag{E.2}$$

The lower bound is saturated iff σ is the identity or a product of disjoint transpositions. The upper bound $|\text{Moeb}(\sigma)| \leq \text{Cat}_{|\sigma|}$ is saturated iff σ is a cycle of length $|\sigma| + 1$.

Proof. The lemma holds when σ is the identity. Otherwise, suppose σ has disjoint cycle decomposition $\sigma = C_1 C_2 \cdots C_k$, where C_j for $1 \leq j \leq k$ are nontrivial cycles. Then

$$|\text{Moeb}(\sigma)| = \prod_{j=1}^{k} \text{Cat}_{|C_j|} \geq 1 \tag{E.3}$$

given that $\text{Cat}_{|C_j|} \geq 1$ for all j. The inequality is saturated iff $|C_j| = 1$ for all j, that is, σ is a product of disjoint transpositions. On the other hand,

$$|\text{Moeb}(\sigma)| = \prod_{j=1}^{k} \text{Cat}_{|C_j|} \leq \text{Cat}_{\sum_j |C_j|} = \text{Cat}_{|\sigma|} < \frac{4^{|\sigma|}}{\sqrt{\pi} |\sigma|^{3/2}}, \tag{E.4}$$

where the two inequalities follow from Corollary 31 and Lemma 29, respectively. The first inequality is saturated when $k = 1$, but is strict whenever $k \geq 2$. So the upper bound $|\text{Moeb}(\sigma)| \leq \text{Cat}_{|\sigma|}$ is saturated iff σ is a cycle of length $|\sigma| + 1$. $\qquad\square$

F Bounds on the Weingarten function

The following theorem is reproduced from [86],

Theorem 33. *When $d > \sqrt{6} k^{7/4}$, any $\sigma \in S_k$ satisfies*

$$\frac{1}{1 - \frac{k-1}{d^2}} \leq \frac{d^{k+|\sigma|} \text{Wg}(\sigma, d)}{\text{Moeb}(\sigma)} \leq a_k := \frac{1}{1 - \frac{6k^{7/2}}{d^2}}. \tag{F.1}$$

The following lemma is an immediate consequence of Theorem 33 and Lemma 32.

Lemma 34. *When $d > \sqrt{6} k^{7/4}$, any $\sigma \in S_k$ satisfies*

$$d^k |\text{Wg}(\sigma, d)| \leq \begin{cases} a_k (\frac{1}{d})^{|\sigma|} & |\sigma| = 0, 1, \\ \min \left\{ \frac{a_k (\frac{4}{d})^{|\sigma|}}{\sqrt{\pi} |\sigma|^{3/2}}, \frac{a_k (\frac{4}{d})^{|\sigma|}}{8} \right\} & |\sigma| \geq 2, \end{cases} \tag{F.2}$$

where a_k is defined in Theorem 33.

Lemma 35. *Suppose $d > \sqrt{6}k^{7/4}$; then*

$$\sum_{\sigma \in A_k} d^k \mathrm{Wg}(\sigma, d) \le \frac{a_k}{8} \left[7 + \cosh \frac{2k(k-1)}{d} \right]. \tag{F.3}$$

Proof. According to Lemma 34,

$$\sum_{\sigma \in A_k} d^k \mathrm{Wg}(\sigma, d) \le a_k + \sum_{\sigma \in A_k \, |\sigma| \ge 2} d^k \mathrm{Wg}(\sigma, d) \le a_k + \frac{a_k}{8} \sum_{\sigma \in A_k \, |\sigma| \ge 2} \left(\frac{4}{d}\right)^{|\sigma|}$$

$$\le \frac{7a_k}{8} + \frac{a_k}{8} \sum_{\sigma \in A_k} \left(\frac{4}{d}\right)^{|\sigma|} = \frac{7a_k}{8} + \frac{a_k}{8} \sum_{\sigma \in A_k} \left(\frac{4}{d}\right)^{k-\xi(\sigma)}$$

$$= \frac{7a_k}{8} + \frac{a_k}{8} \left(\frac{4}{d}\right)^k \sum_{\sigma \in A_k} \left(\frac{d}{4}\right)^{\xi(\sigma)}$$

$$= \frac{7a_k}{8} + \frac{a_k}{16} \left(\frac{4}{d}\right)^k \left[\sum_{\sigma \in S_k} \left(\frac{d}{4}\right)^{\xi(\sigma)} + \sum_{\sigma \in S_k} \left(-\frac{d}{4}\right)^{\xi(\sigma)} \right]$$

$$= \frac{7a_k}{8} + \frac{a_k}{16} \left(\frac{4}{d}\right)^k \left[\prod_{j=0}^{k-1} \left(\frac{d}{4} + j\right) + \prod_{j=0}^{k-1} \left(\frac{d}{4} - j\right) \right]$$

$$= \frac{7a_k}{8} + \frac{a_k}{16} \left[\prod_{j=0}^{k-1} \left(1 + \frac{4j}{d}\right) + \prod_{j=0}^{k-1} \left(1 - \frac{4j}{d}\right) \right]$$

$$\le \frac{7a_k}{8} + \frac{a_k}{16} \left[\prod_{j=0}^{k-1} \mathrm{e}^{4j/d} + \prod_{j=0}^{k-1} \mathrm{e}^{-4j/d} \right] = \frac{7a_k}{8} + \frac{a_k}{16} \left[\mathrm{e}^{\sum_{j=0}^{k-1} 4j/d} + \mathrm{e}^{-\sum_{j=0}^{k-1} 4j/d} \right]$$

$$= \frac{7a_k}{8} + \frac{a_k}{16} \left[\mathrm{e}^{2k(k-1)/d} + \mathrm{e}^{-2k(k-1)/d} \right] = \frac{a_k}{8} \left[7 + \cosh \frac{2k(k-1)}{d} \right]. \tag{F.4}$$

$\square$

G   Bounds on the number of permutations with a given genus

In this appendix, we provide an easy-to-use upper bound for the number of permutations with a given genus (Lemma 39 below), which plays a crucial role in understanding Rényi entanglement entropies of Haar random states as well as states drawn from designs.

The basis of our endeavor is the following theorem due to Goupil and Schaeffer [87].

Theorem 36. *The number of permutations in the symmetric group S_n with genus g is given by*

$$c_{g,n} = \frac{(n+1)_{2g}}{(n+1)2^{2g}} \sum_{g_1 + g_2 = g} \sum_{\substack{0 \le \ell_1 \le g_1 \\ 0 \le \ell_2 \le g_2}} a_{g_1, \ell_1} a_{g_2, \ell_2} (n+1-2g)_{\ell_1 + \ell_2} \binom{2n - 2g - \ell_1 - \ell_2}{n - 2g_1 - \ell_1}, \tag{G.1}$$

where $(n)_k := n(n-1)\cdots, (n-k+1)$, $a_{0,0} = 1$, $a_{g,0} = 0$ for $g \ge 1$, and

$$a_{g,\ell} = \sum_{\substack{\gamma \vdash g, \, \ell(\gamma) = \ell \\ \gamma = 1^{c_1} 2^{c_2} g^{c_g}}} \frac{1}{\prod_{j=1}^{g} c_j! (2j+1)^{c_j}} \quad 0 < \ell \le g. \tag{G.2}$$

Here the summation runs over all partition γ of g, the expression $\gamma = 1^{c_1} 2^{c_2} g^{c_g}$ means that γ has c_j parts equal to j, and $\ell(\gamma) = \sum_j c_j$ denotes the number of parts of γ.

In addition, we need two auxiliary lemmas.

Lemma 37. $a_{g,\ell} \leq 2^{-\ell}$ for all $0 \leq \ell \leq g$.

Proof. By definition, the lemma holds when $g = 0$, or $g \geq 1$ and $\ell = 0$. Now suppose $0 < \ell \leq g$; then

$$
a_{g,\ell} = \sum_{\substack{\gamma \vdash g, \, \ell(\gamma) = \ell \\ \gamma = 1^{c_1} 2^{c_2} g^{c_g}}} \frac{1}{\prod_{j=1}^{g} c_j!(2j+1)^{c_j}} = \sum_{\substack{\gamma \vdash g, \, \ell(\gamma) = \ell \\ \gamma = 1^{c_1} 2^{c_2} g^{c_g}}} \left[\frac{1}{\prod_{j=1}^{g} c_j! j^{c_j}} \prod_{j=1}^{g} \left(\frac{j}{2j+1} \right)^{c_j} \right]
$$

$$
\leq \sum_{\substack{\gamma \vdash g, \, \ell(\gamma) = \ell \\ \gamma = 1^{c_1} 2^{c_2} g^{c_g}}} \left[\frac{1}{\prod_{j=1}^{g} c_j! j^{c_j}} \prod_{j=1}^{g} \left(\frac{1}{2} \right)^{c_j} \right] = \sum_{\substack{\gamma \vdash g, \, \ell(\gamma) = \ell \\ \gamma = 1^{c_1} 2^{c_2} g^{c_g}}} \left[\frac{1}{\prod_{j=1}^{g} c_j! j^{c_j}} 2^{-\sum_{j=1}^{g} c_j} \right]
$$

$$
= \sum_{\substack{\gamma \vdash g, \, \ell(\gamma) = \ell \\ \gamma = 1^{c_1} 2^{c_2} g^{c_g}}} \frac{1}{\prod_{j=1}^{g} c_j! j^{c_j}} 2^{-\ell(\gamma)} = 2^{-\ell} \sum_{\substack{\gamma \vdash g, \, \ell(\gamma) = \ell \\ \gamma = 1^{c_1} 2^{c_2} g^{c_g}}} \frac{1}{\prod_{j=1}^{g} c_j! j^{c_j}} \leq 2^{-\ell}. \tag{G.3}
$$

Here the last inequality can be derived as follows. Note that $\prod_{j=1}^{g} c_j! j^{c_j}$ is the order of the centralizer in S_g of each element in the conjugacy class labeled by the partition γ. Therefore, $g! / \prod_{j=1}^{g} c_j! j^{c_j}$ is the number of elements in this conjugacy class, so that

$$
\sum_{\substack{\gamma \vdash g \\ \gamma = 1^{c_1} 2^{c_2} g^{c_g}}} \frac{g!}{\prod_{j=1}^{g} c_j! j^{c_j}} = g!, \tag{G.4}
$$

which amounts to the identity

$$
\sum_{\substack{\gamma \vdash g \\ \gamma = 1^{c_1} 2^{c_2} g^{c_g}}} \frac{1}{\prod_{j=1}^{g} c_j! j^{c_j}} = 1. \tag{G.5}
$$

As an immediate consequence,

$$
\sum_{\substack{\gamma \vdash g, \, \ell(\gamma) = \ell \\ \gamma = 1^{c_1} 2^{c_2} g^{c_g}}} \frac{1}{\prod_{j=1}^{g} c_j! j^{c_j}} \leq 1. \tag{G.6}
$$

$\square$

Lemma 38. *Suppose j, k, n are nonnegative integers satisfying $j \leq n$, $k < 2n$, and $k \leq n + j$. Then*

$$
\binom{2n - k}{n - j} \leq 2^{-k} \sqrt{\frac{n}{n - \lfloor k/2 \rfloor}} \binom{2n}{n}. \tag{G.7}
$$

Proof. Straightforward calculation shows that

$$\binom{2n-k}{n-j} \leq \binom{2n-k}{n-\lfloor k/2 \rfloor} = \binom{2n-k}{n-\lceil k/2 \rceil}. \tag{G.8}$$

So without loss of generality, we may assume that $j = \lfloor k/2 \rfloor$. Then

$$\frac{\binom{2n}{n}}{\binom{2n-k}{n-j}} = \frac{(2n)!(n-j)!(n+j-k)!}{(2n-k)!n!n!} = \frac{2n(2n-1)\cdots(2n-k+1)}{[n(n-1)\cdots n-j+1][n(n-1)\cdots(n+j-k+1)]}$$

$$= \frac{2^k n(n-\frac{1}{2})\cdots(n-\frac{k}{2}+\frac{1}{2})}{[n(n-1)\cdots n-j+1][n(n-1)\cdots(n+j-k+1)]} = 2^k f, \tag{G.9}$$

where

$$f = \frac{(n-\frac{1}{2})(n-\frac{3}{2})\cdots(n-j+\frac{1}{2})}{n(n-1)\cdots(n-j+1)}. \tag{G.10}$$

The square of f can be bounded from below as follows,

$$f^2 = \frac{(n-\frac{1}{2})^2(n-\frac{3}{2})^2\cdots(n-j+\frac{1}{2})^2}{n^2(n-1)^2\cdots(n-j+1)^2}$$

$$= \frac{1}{n} \times \frac{(n-\frac{1}{2})^2}{n(n-1)} \times \cdots \times \frac{(n-j+\frac{3}{2})^2}{(n-j+2)(n-j+1)} \times \frac{(n-j+\frac{1}{2})^2}{n-j+1}$$

$$\geq \frac{1}{n} \times \frac{(n-j+\frac{1}{2})^2}{n-j+1} \geq \frac{n-j}{n} = \frac{n-\lfloor \frac{k}{2} \rfloor}{n}. \tag{G.11}$$

Therefore $f \geq \sqrt{\frac{n-\lfloor k/2 \rfloor}{n}}$, from which the lemma follows. $\qquad\square$

Lemma 39.

$$\frac{c_{g,n}}{c_{0,n}} \leq \frac{(g+1)n^{3g}}{2^{6g}}, \quad \frac{c_{g,n}}{c_{0,n}} \leq \frac{2}{3}\left(\frac{n^3}{32}\right)^g \quad \forall 1 \leq g \leq \frac{n-1}{2}. \tag{G.12}$$

Proof. Recall that $c_{0,n} = c_n = (2n)!/[n!(n+1)!]$. The values of $c_{1,n}, c_{2,n}$ can be computed explicitly according to Theorem 36, with the result

$$c_{1,n} = \frac{n(n-1)}{6}\binom{2n-3}{n} = \frac{(2n-3)!}{6(n-2)!(n-3)!}, \tag{G.13}$$

$$c_{2,n} = \frac{(2n-5)!(5n^2-7n+6)}{720(n-3)!(n-5)!}. \tag{G.14}$$

The coefficients $a_{g,\ell}$ necessary for deriving this result are given by

$$a_{00} = 1, \quad a_{1,1} = \frac{1}{3}, \quad a_{2,1} = \frac{1}{5}, \quad a_{2,2} = \frac{1}{18}. \tag{G.15}$$

As a consequence,

$$\frac{c_{1,n}}{c_{0,n}} = \frac{n(n+1)(n-1)(n-2)}{24(2n-1)} \leq \frac{n^3}{48}, \tag{G.16}$$

$$\frac{c_{2,n}}{c_{0,n}} = \frac{(n+1)n(n-1)(n-2)(n-3)(n-4)(5n^2-7n+6)}{5760(2n-1)(2n-3)} \leq \frac{n^6}{4608}. \tag{G.17}$$

Therefore, Lemma 39 holds when $g = 1, 2$. Now suppose $g \geq 3$, so that $n \geq 7$. According to Theorem 36, we have

$$
\begin{aligned}
\frac{c_{g,n}}{c_{0,n}} &= \frac{(n+1)_{2g}}{2^{2g}} \sum_{g_1+g_2=g} \sum_{\substack{0 \leq \ell_1 \leq g_1 \\ 0 \leq \ell_2 \leq g_2}} a_{g_1,\ell_1} a_{g_2,\ell_2} (n+1-2g)_{\ell_1+\ell_2} \frac{\binom{2n-2g-\ell_1-\ell_2}{n-2g_1-\ell_1}}{\binom{2n}{n}} \\
&\leq \frac{(n+1)_{2g}}{2^{2g}} \sum_{g_1+g_2=g} \sum_{\substack{0 \leq \ell_1 \leq g_1 \\ 0 \leq \ell_2 \leq g_2}} a_{g_1,\ell_1} a_{g_2,\ell_2} (n+1-2g)_{\ell_1+\ell_2} \times 2^{-(2g+\ell_1+\ell_2)} \sqrt{\frac{n}{n-g-\lfloor (\ell_1+\ell_2)/2 \rfloor}} \\
&\leq \frac{(n+1)_{2g}}{2^{4g}} \sum_{g_1+g_2=g} \sum_{\substack{0 \leq \ell_1 \leq g_1 \\ 0 \leq \ell_2 \leq g_2}} 2^{-(\ell_1+\ell_2)} a_{g_1,\ell_1} a_{g_2,\ell_2} \frac{n(n+1-2g)_{\ell_1+\ell_2}}{n-g-\lfloor (\ell_1+\ell_2)/2 \rfloor} \\
&\leq \frac{(n+1)_{2g}}{2^{4g}} \sum_{g_1+g_2=g} \sum_{\substack{0 \leq \ell_1 \leq g_1 \\ 0 \leq \ell_2 \leq g_2}} 4^{-(\ell_1+\ell_2)} n^{\ell_1+\ell_2} \frac{\max\{0, n+2-2g-(\ell_1+\ell_2)\}}{n-g-\lfloor (\ell_1+\ell_2)/2 \rfloor}.
\end{aligned}
\tag{G.18}
$$

Here the first inequality follows from Lemma 38, and the last one from Lemma 37 and the fact that $a_{g,0} = 0$ for $g > 0$. The fraction at the end of the above equation is no larger than 1 given that $g \geq 3$. Therefore,

$$
\begin{aligned}
\frac{c_{g,n}}{c_{0,n}} &\leq \frac{(n+1)_{2g}}{2^{4g}} \sum_{g_1+g_2=g} \sum_{\substack{0 \leq \ell_1 \leq g_1 \\ 0 \leq \ell_2 \leq g_2}} \left(\frac{n}{4}\right)^{-(\ell_1+\ell_2)} = \frac{(n+1)_{2g}}{2^{4g}} \sum_{g_1+g_2=g} \frac{[\left(\frac{n}{4}\right)^{g_1+1}-1]}{\frac{n}{4}-1} \frac{[\left(\frac{n}{4}\right)^{g_2+1}-1]}{\frac{n}{4}-1} \\
&\leq \frac{(n+1)_{2g}}{2^{4g}} \frac{1}{\left(\frac{n}{4}-1\right)^2} \sum_{g_1+g_2=g} \left(\frac{n}{4}\right)^{g+2} = \frac{(n+1)_{2g}}{2^{4g}} \frac{(g+1)\left(\frac{n}{4}\right)^{g+2}}{\left(\frac{n}{4}-1\right)^2} \\
&= \frac{(g+1)n^{g+2}(n+1)_{2g}}{2^{6g}(n-4)^2} \leq \frac{(g+1)n^{3g-3}(n+1)(n-1)(n-2)(n-3)(n-4)}{2^{6g}(n-4)^2} \\
&\leq \frac{(g+1)n^{3g}}{2^{6g}}.
\end{aligned}
\tag{G.19}
$$

This result confirms the first inequality in Lemma 39 in the remaining case $g \geq 3$, which in turn implies the second inequality in the lemma. $\qquad\square$

H Partially scrambling unitary

Here we analyze the partially scrambling unitary model proposed in [50], which can lead to a large separation between von Neumann and Rényi-2 entanglement entropies and tripartite information in the Choi state setting. More explicitly, let $\tilde{U}$ be a unitary that perfectly scrambles on almost the whole space besides a small subspace. Then, on the one hand, $\tilde{U}$ still has nearly maximal $-I_3$ due to continuity; while on the other hand, $-I_3^{(2)}$ can be gapped from maximum by $\Theta(\log d)$. However, we find that this model is not likely to provide strict separations between Rényi entropies of order ≥ 2.

The generalized partially scrambling unitary is defined as follows. Given α, define

$$
\tilde{U}|mo\rangle = \begin{cases} U_S|mo\rangle & 0 \leq m, o < D \\ |mo\rangle & \text{otherwise} \end{cases}
\tag{H.1}
$$

where U_S is α-scrambling, and $D \le \sqrt{d}$ controls the size of this α-scrambling subspace (labeled by subscript S). Then the Choi state of $\tilde{U}$ is

$$|\tilde{U}\rangle = \frac{D}{\sqrt{d}}|U_S\rangle_{A_S B_S C_S D_S} + \frac{1}{\sqrt{d}} \sum_{D \le m < \sqrt{d} \wedge D \le o < \sqrt{d}} |mo\rangle_{AB} \otimes |mo\rangle_{CD}. \qquad \text{(H.2)}$$

The question is whether there exists some D that can lead to separations between higher Rényi entropies associated with this Choi state, say α and α', $\alpha' > \alpha \ge 2$. To establish such separations, we need to show a large ($\Theta(\log d)$) gap between Rényi-α' entropies and the maximum for some small D, as well as upper bound the difference between Rényi-α entropies and the maximum by continuity. The gap side can work out by directly generalizing the corresponding calculation in [50]: let $\beta = \log(\sqrt{d} - D)/\log\sqrt{d}$. Then $\log d - S_R^{(\alpha')}(\text{tr}_{BD}|\tilde{U}\rangle\langle\tilde{U}|) = \Theta(\log d)$ as long as β is a positive constant. However, we find that the continuity bound for unified entropies can only give trivial results on the continuity side:

Lemma 40 (Generalized Fannes' inequality [34]). *Let ρ and ρ' be density operators in Hilbert space of dimension d. Denote $\epsilon = D_{\text{tr}}(\rho, \rho')$. For $\alpha > 1$ and $s \ge 0$:*

$$|S_s^{(\alpha)}(\rho) - S_s^{(\alpha)}(\rho')| \le \chi_s[\epsilon^\alpha \log_\alpha(d-1) + H^{(\alpha)}(\epsilon, 1 - \epsilon)], \qquad \text{(H.3)}$$

where $\chi_s = 1$ for $s \ge 1$, and $\chi_s = d^{2(\alpha-1)}$ for $s = 0$. $H^{(\alpha)}$ denotes the α binary entropy.

It can be seen that this generalized Fannes' bound for Rényi entropies grows with the dimension d for $\alpha > 1$, which indicates that even a tiny non-scrambling subspace may perturb the Rényi entropies drastically. Indeed, some simple scaling analysis can confirm that this bound is trivial even for the Rényi-2 entropy. Notice that $\epsilon = D_{\text{tr}}(\text{tr}_{BD}|\tilde{U}\rangle\langle\tilde{U}|, I) \le O(D/\log\sqrt{d}) = O(d^{(\beta-1)/2})$. Then it must hold that $2(\alpha-1) + \alpha(\beta-1)/2 < 0$ so that $\log d - S_R^{(\alpha)}(\text{tr}_{BD}|\tilde{U}\rangle\langle\tilde{U}|) = o(\log d)$. This gives $\beta < -3 + 4/\alpha$, which has no overlap with the $\beta > 0$ solution on the gap side when $\alpha \ge 2$. Equivalently, by plugging in $\beta > 0$ we can solve that the desired separation can exist when $\alpha' < 2$. Summarizing, in order to have a nontrivial bound on Rényi entropies D needs to be $o(1)$, which is meaningless. This is hardly surprising: one expects that Rényi entropies are very sensitive, especially in the near-maximum regime, due to the logarithm. In fact, we are able to obtain a large gap on the α' side basically because of such exponential sensitivity. Suppose we consider $s > 0$ entropies instead. Then the continuity bound is strong since $\chi_s = 1$, but it becomes hard to find a gap on the other side. There is a fundamental tradeoff between sensitivity and robustness in these unified entropies. In conclusion, we believe that partially scrambling unitaries are not likely to produce separations between generalized entropies in the Choi model.

I Proof of Lemma 24

To prove Lemma 24, we need to introduce several auxiliary concepts and lemmas. An $m \times s$ matrix G is a (standard) Gaussian random matrix if the entries of G are i.i.d. standard Gaussian random variables (with mean 0 and variance 1). It is a complex Gaussian random matrix if its real part and imaginary part are independent Gaussian random matrices.

Lemma 41. *Suppose G is a standard $m \times s$ real Gaussian random matrix. Then*

$$\mathbb{E}\,\|G\| \leq \frac{\sqrt{2}\Gamma\left(\frac{m+1}{2}\right)}{\Gamma\left(\frac{m}{2}\right)} + \frac{\sqrt{2}\Gamma\left(\frac{s+1}{2}\right)}{\Gamma\left(\frac{s}{2}\right)} \leq \sqrt{m} + \sqrt{s}. \tag{I.1}$$

Usually this lemma is stated without the intermediate term, as it appears in [88]. However, the first inequality is essential to achieve our goal. Fortunately, this inequality is already implied by the proof in [88]. Note that $\sqrt{2}\Gamma\left(\frac{m+1}{2}\right)/\Gamma\left(\frac{m}{2}\right)$ is the average norm of a vector composed of m iid standard Gaussian random variables, while $\sqrt{m}$ is the root mean square norm. This observation implies the second inequality in the lemma, which is nearly tight when m, s are large.

Lemma 42. *Suppose G is a standard $m \times s$ complex Gaussian random matrix. Then*

$$\mathbb{E}\,\|G\| \leq \frac{2\sqrt{2}\Gamma\left(\frac{m+1}{2}\right)}{\Gamma\left(\frac{m}{2}\right)} + \frac{2\sqrt{2}\Gamma\left(\frac{s+1}{2}\right)}{\Gamma\left(\frac{s}{2}\right)} \leq 2\sqrt{m} + 2\sqrt{s}. \tag{I.2}$$

This lemma is an immediate consequence of the triangle inequality and Lemma 41 applied to the real and imaginary parts of G.

Lemma 43.

$$\mathbb{E}\,\|\rho_A\|^a = \frac{\Gamma(k)\,\mathbb{E}\,\|G\|^{2a}}{2^a\Gamma(k+a)} \quad \forall a \geq 0, \tag{I.3}$$

where G is a complex (real) Gaussian random matrix of size $d_A \times d_B$ and $k = d_A d_B$ ($k = d_A d_B/2$ in the real case).

Proof. It is well known that $G/\|G\|_2$ considered as a unit vector in $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ is distributed uniformly. In addition, the spectrum of $G/\|G\|_2$ is independent of the Frobenius norm $\|G\|_2 = \sqrt{\mathrm{tr}\{GG^\dagger\}}$. Therefore,

$$\mathbb{E}\,\|G\|^{2a} = \mathbb{E}[\mathrm{tr}\{GG^\dagger\}]^a\,\mathbb{E}\left\|\frac{G}{\|G\|_2}\right\|^{2a} = \mathbb{E}[\mathrm{tr}\{GG^\dagger\}]^a\,\mathbb{E}\,\|\rho_A\|^a = \frac{2^a\Gamma(k+a)}{\Gamma(k)}\,\mathbb{E}\,\|\rho_A\|^a, \tag{I.4}$$

from which the lemma follows. Here the last equality in the above equation follows from the fact that $\mathrm{tr}\{GG^\dagger\}$ obeys χ^2-distribution with $2k$-degrees of freedom and pdf.

$$f(x) = \frac{x^{k-1}\mathrm{e}^{-x/2}}{2^k\Gamma(k)}, \tag{I.5}$$

which satisfies

$$\int x^a f(x)\mathrm{d}x = \frac{2^a\Gamma(k+a)}{\Gamma(k)} \quad \forall a \geq 0. \tag{I.6}$$

□

Proof. According to Lemmas 43 and 41, in the real case, we have

$$\mathbb{E}\sqrt{\|\rho_A\|} = \frac{\Gamma\left(\frac{d_A d_B}{2}\right)\mathbb{E}\,\|G\|}{\sqrt{2}\Gamma(\frac{d_A d_B+1}{2})} \leq \frac{\Gamma\left(\frac{d_A d_B}{2}\right)}{\sqrt{2}\Gamma(\frac{d_A d_B+1}{2})}\left(\frac{\sqrt{2}\Gamma\left(\frac{d_A+1}{2}\right)}{\Gamma\left(\frac{d_A}{2}\right)} + \frac{\sqrt{2}\Gamma\left(\frac{d_B+1}{2}\right)}{\Gamma\left(\frac{d_B}{2}\right)}\right)$$

$$= \frac{\gamma(d_B)}{\gamma(d_A d_B)}\frac{1}{\sqrt{d_A}} + \frac{\gamma(d_A)}{\gamma(d_A d_B)}\frac{1}{\sqrt{d_B}} \leq \frac{1}{\sqrt{d_A}} + \frac{1}{\sqrt{d_B}}, \tag{I.7}$$

where $\gamma(m) := \Gamma(\frac{m+1}{2})/(\sqrt{m}\Gamma(\frac{m}{2}))$, and the last inequality follows from the fact that $\gamma(m)$ is monotonic increasing with m for $m \geq 1$. This conclusion is intuitive if we observe that $\gamma(m)$ is equal to the ratio of the mean length over the root mean square length of a standard Gaussian random vector with m components. To derive an analytical proof, we can compute the log-derivative of $\gamma(m)$ with respect to m, note that the definition of $\gamma(m)$ can be extended to positive real numbers. Straightforward calculations shows that

$$\frac{\mathrm{d}\ln\gamma(m)}{\mathrm{d}m} = \frac{1}{2}\left[\psi^{(0)}\left(\frac{m+1}{2}\right) - \psi^{(0)}\left(\frac{m}{2}\right) - \frac{1}{m}\right] \geq \frac{1}{4}\left[\psi^{(0)}\left(\frac{m+2}{2}\right) - \psi^{(0)}\left(\frac{m}{2}\right) - \frac{2}{m}\right] = 0. \tag{I.8}$$

Here $\psi^{(0)}$ denotes the digamma function (instead of a ket), the inequality follows from the concavity of $\psi^{(0)}$, and the last equality follows from the identity $\psi^{(0)}(x+1) = \psi^{(0)}(x) + \frac{1}{x}$.

In the complex case, Lemmas 43 and 42 imply that

$$\mathbb{E}\sqrt{\|\rho_A\|} \leq \sqrt{2}\left(\frac{\gamma(d_B)}{\gamma(2d_A d_B)}\frac{1}{\sqrt{d_A}} + \frac{\gamma(d_A)}{\gamma(2d_A d_B)}\frac{1}{\sqrt{d_B}}\right) \leq \sqrt{2}\left(\frac{1}{\sqrt{d_A}} + \frac{1}{\sqrt{d_B}}\right), \tag{I.9}$$

where the second inequality follows from the monotonicity of $\gamma(\cdot)$, as in the real case. $\qquad\square$

Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

[1] Z.-W. Liu, S. Lloyd, E.Y. Zhu and H. Zhu, *Generalized Entanglement Entropies of Quantum Designs*, *Phys. Rev. Lett.* **120** (2018) 130502 [arXiv:1709.04313] [INSPIRE].

[2] P. Hayden and J. Preskill, *Black holes as mirrors: Quantum information in random subsystems*, *JHEP* **09** (2007) 120 [arXiv:0708.4025] [INSPIRE].

[3] Y. Sekino and L. Susskind, *Fast Scramblers*, *JHEP* **10** (2008) 065 [arXiv:0808.2096] [INSPIRE].

[4] L. Susskind, *Addendum to Fast Scramblers*, arXiv:1101.6048 [INSPIRE].

[5] S.W. Hawking, *Black hole explosions*, *Nature* **248** (1974) 30 [INSPIRE].

[6] S.W. Hawking, *Particle Creation by Black Holes*, *Commun. Math. Phys.* **43** (1975) 199 [*Erratum ibid.* **46** (1976) 206] [INSPIRE].

[7] S.W. Hawking, *Black Holes and Thermodynamics*, *Phys. Rev.* **D 13** (1976) 191 [INSPIRE].

[8] R. Nandkishore and D.A. Huse, *Many-body localization and thermalization in quantum statistical mechanics*, *Ann. Rev. Condensed Matter Phys.* **6** (2015) 15.

[9] P. Hosur, X.-L. Qi, D.A. Roberts and B. Yoshida, *Chaos in quantum channels*, *JHEP* **02** (2016) 004 [arXiv:1511.04021] [INSPIRE].

[10] S.H. Shenker and D. Stanford, *Black holes and the butterfly effect*, *JHEP* **03** (2014) 067 [arXiv:1306.0622] [INSPIRE].

[11] S.H. Shenker and D. Stanford, *Multiple Shocks*, *JHEP* **12** (2014) 046 [arXiv:1312.3296] [INSPIRE].

[12] D.A. Roberts, D. Stanford and L. Susskind, *Localized shocks*, *JHEP* **03** (2015) 051 [arXiv:1409.8180] [ɪɴSPIRE].

[13] S.H. Shenker and D. Stanford, *Stringy effects in scrambling*, *JHEP* **05** (2015) 132 [arXiv:1412.6087] [ɪɴSPIRE].

[14] D.A. Roberts and D. Stanford, *Two-dimensional conformal field theory and the butterfly effect*, *Phys. Rev. Lett.* **115** (2015) 131603 [arXiv:1412.5123] [ɪɴSPIRE].

[15] J. Maldacena, S.H. Shenker and D. Stanford, *A bound on chaos*, *JHEP* **08** (2016) 106 [arXiv:1503.01409] [ɪɴSPIRE].

[16] P. Caputa, T. Numasawa and A. Veliz-Osorio, *Out-of-time-ordered correlators and purity in rational conformal field theories*, *PTEP* **2016** (2016) 113B06.

[17] B. Swingle, G. Bentsen, M. Schleier-Smith and P. Hayden, *Measuring the scrambling of quantum information*, *Phys. Rev.* **A 94** (2016) 040302 [arXiv:1602.06271] [ɪɴSPIRE].

[18] D.A. Roberts and B. Yoshida, *Chaos and complexity by design*, *JHEP* **04** (2017) 121 [arXiv:1610.04903] [ɪɴSPIRE].

[19] Y. Huang, Y.-L. Zhang and X. Chen, *Out-of-time-ordered correlators in many-body localized systems*, *Annalen Phys.* **529** (2016) 1600318.

[20] R. Fan, P. Zhang, H. Shen and H. Zhai, *Out-of-time-order correlation for many-body localization*, *Science Bulletin* **62** (2017) 707.

[21] X. Chen, T. Zhou, D.A. Huse and E. Fradkin, *Out-of-time-order correlations in many-body localized and thermal phases*, *Annalen Phys.* **529** (2016) 1600332.

[22] N. Lashkari, D. Stanford, M. Hastings, T. Osborne and P. Hayden, *Towards the Fast Scrambling Conjecture*, *JHEP* **04** (2013) 022 [arXiv:1111.6580] [ɪɴSPIRE].

[23] D.N. Page, *Average entropy of a subsystem*, *Phys. Rev. Lett.* **71** (1993) 1291 [gr-qc/9305007] [ɪɴSPIRE].

[24] S.K. Foong and S. Kanno, *Proof of Page's conjecture on the average entropy of a subsystem*, *Phys. Rev. Lett.* **72** (1994) 1148.

[25] J. Sánchez-Ruiz, *Simple proof of Page's conjecture on the average entropy of a subsystem*, *Phys. Rev.* **E 52** (1995) 5653.

[26] S. Sen, *Average entropy of a subsystem*, *Phys. Rev. Lett.* **77** (1996) 1 [hep-th/9601132] [ɪɴSPIRE].

[27] E. Knill, *Approximation by quantum circuits*, quant-ph/9508006 [ɪɴSPIRE].

[28] F.G.S.L. Brandão, A.W. Harrow and M. Horodecki, *Local Random Quantum Circuits are Approximate Polynomial-Designs*, *Commun. Math. Phys.* **346** (2016) 397 [ɪɴSPIRE].

[29] F.G.S.L. Brandão, A.W. Harrow and M. Horodecki, *Efficient quantum pseudorandomness*, *Phys. Rev. Lett.* **116** (2016) 170502.

[30] G. Tóth and J.J. García-Ripoll, *Efficient algorithm for multiqudit twirling for ensemble quantum computation*, *Phys. Rev.* **A 75** (2007) 042311.

[31] Y. Nakata, C. Hirche, M. Koashi and A. Winter, *Efficient quantum pseudorandomness with nearly time-independent Hamiltonian dynamics*, *Phys. Rev.* **X 7** (2017) 021006.

[32] Z.-W. Liu, C. Perry, Y. Zhu, D.E. Koh and S. Aaronson, *Doubly infinite separation of quantum information and communication*, *Phys. Rev.* **A 93** (2016) 012347.

[33] X. Hu and Z. Ye, *Generalized quantum entropy*, *J. Math. Phys.* **47** (2006) 023502.

[34] A.E. Rastegin, *Some general properties of unified entropies*, *J. Stat. Phys.* **143** (2011) 1120.

[35] R. Horodecki, P. Horodecki, M. Horodecki and K. Horodecki, *Quantum entanglement*, *Rev. Mod. Phys.* **81** (2009) 865 [quant-ph/0702225] [INSPIRE].

[36] R. Islam et al., *Measuring entanglement entropy in a quantum many-body system*, *Nature* **528** (2015) 77.

[37] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr and M. Tomamichel, *On quantum Rényi entropies: A new generalization and some properties*, *J. Math. Phys.* **54** (2013) 122203.

[38] F. Leditzky, C. Rouzé and N. Datta, *Data processing for the sandwiched Rényi divergence: a condition for equality*, *Lett. Math. Phys.* **107** (2017) 61 [arXiv:1604.02119] [INSPIRE].

[39] S. Beigi, *Sandwiched Rényi divergence satisfies data processing inequality*, *J. Math. Phys.* **54** (2013) 122202.

[40] X. Dong, *The gravity dual of Rényi entropy*, *Nature Commun.* **7** (2016) 12472.

[41] M. Rangamani and T. Takayanagi, *Holographic Entanglement Entropy*, Springer, (2017).

[42] S.G. Hoggar, *t-designs in projective spaces*, *Eur. J. Combinator.* **3** (1982) 233.

[43] J.M. Renes, R. Blume-Kohout, A.J. Scott and C.M. Caves, *Symmetric informationally complete quantum measurements*, *J. Math. Phys.* **45** (2004) 2171.

[44] H. Zhu, R. Kueng, M. Grassl and D. Gross, *The Clifford group fails gracefully to be a unitary 4-design* [arXiv:1609.08172].

[45] D.P. DiVincenzo, D.W. Leung and B.M. Terhal, *Quantum data hiding*, *IEEE Trans. Inf. Theory* **48** (2002) 580.

[46] C. Dankert, *Efficient Simulation of Random Quantum States and Operators*, MSc Thesis, University of Waterloo, (2005).

[47] C. Dankert, R. Cleve, J. Emerson and E. Livine, *Exact and approximate unitary 2-designs and their application to fidelity estimation*, *Phys. Rev.* **A 80** (2009) 012304.

[48] D. Gross, K. Audenaert and J. Eisert, *Evenly distributed unitaries: On the structure of unitary designs*, *J. Math. Phys.* **48** (2007) 052104.

[49] A. Roy and A.J. Scott, *Unitary designs and codes*, *Des. Codes Cryptogr.* **53** (2009) 13.

[50] D. Ding, P. Hayden and M. Walter, *Conditional Mutual Information of Bipartite Unitaries and Scrambling*, *JHEP* **12** (2016) 145 [arXiv:1608.04750] [INSPIRE].

[51] D. Weingarten, *Asymptotic Behavior of Group Integrals in the Limit of Infinite Rank*, *J. Math. Phys.* **19** (1978) 999 [INSPIRE].

[52] B. Collins, *Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral, and free probability*, *Int. Math. Res. Not.* **2003** (2003) 953.

[53] B. Collins and S. Matsumoto, *On some properties of orthogonal Weingarten functions*, *J. Math. Phys.* **50** (2009) 113516.

[54] P. Zinn-Justin, *Jucys-Murphy elements and Weingarten matrices*, *Lett. Math. Phys.* **91** (2010) 119.

[55] C. Lancien, *k-extendibility of high-dimensional bipartite quantum states*, *Random Matrices: Theory and Applications* **05** (2016) 1650011.

[56] A. Nica and R. Speicher, *Lectures on the Combinatorics of Free Probability.* Cambridge University Press, (2006).

[57] B. Collins and P. Śniady, *Integration with respect to the Haar measure on unitary, orthogonal and symplectic group*, [*Commun. Math. Phys.* **264** (2006) 773](#).

[58] A. Mandarino, T. Linowski and K. Życzkowski, *Bipartite unitary gates and billiard dynamics in the Weyl chamber*, [arXiv:1710.10983](#).

[59] M. Musz, M. Kuś and K. Życzkowski, *Unitary quantum gates, perfect entanglers, and unistochastic maps*, [*Phys. Rev.* **A 87** (2013) 022111](#).

[60] F.G.S.L. Brandão, A.W. Harrow and M. Horodecki, *Efficient quantum pseudorandomness*, [*Phys. Rev. Lett.* **116** (2016) 170502](#).

[61] M.B. Hastings and A.W. Harrow, *Classical and quantum tensor product expanders*, *Quant. Inf. Comput.* **9** (2009) 336.

[62] S. Sachdev and J. Ye, *Gapless spin fluid ground state in a random, quantum Heisenberg magnet*, [*Phys. Rev. Lett.* **70** (1993) 3339](#) [[cond-mat/9212030](#)] [[INSPIRE](#)].

[63] A. Kitaev, http://online.kitp.ucsb.edu/online/entangled15/kitaev/.

[64] http://online.kitp.ucsb.edu/online/entangled15/kitaev2/.

[65] E. Lubkin, *Entropy of an n-system from its correlation with a k-reservoir*, [*J. Math. Phys.* **19** (1978) 1028](#).

[66] S. Lloyd and H. Pagels, *Complexity as thermodynamic depth*, [*Annals Phys.* **188** (1988) 186](#) [[INSPIRE](#)].

[67] P. Hayden, D.W. Leung and A. Winter, *Aspects of generic entanglement*, [*Commun. Math. Phys.* **265** (2006) 95](#).

[68] G. Smith and D. Leung, *Typical entanglement of stabilizer states*, [*Phys. Rev.* **A 74** (2006) 062314](#).

[69] K. Życzkowski and H.-J. Sommers, *Induced measures in the space of mixed quantum states*, [*J. Phys.* **A 34** (2001) 7111](#).

[70] L.C. Malacarne, R.S. Mendes and E.K. Lenzi, *Average entropy of a subsystem from its average Tsallis entropy*, [*Phys. Rev.* **E 65** (2002) 046131](#) [[INSPIRE](#)].

[71] B. Collins and I. Nechita, *Random quantum channels I: Graphical calculus and the Bell state phenomenon*, [*Commun. Math. Phys.* **297** (2010) 345](#).

[72] B. Collins and I. Nechita, *Gaussianization and eigenvalue statistics for random quantum channels (III)*, [*Annals Appl. Probab.* **21** (2011) 1136](#).

[73] P. Hayden and A. Winter, *Counterexamples to the maximal p-norm multiplicativity conjecture for all p > 1*, [*Commun. Math. Phys.* **284** (2008) 263](#).

[74] H. Zhu, *Multiqubit Clifford groups are unitary 3-designs*, [*Phys. Rev.* **A 96** (2017) 062336](#).

[75] Z. Webb, *The Clifford group forms a unitary 3-design*, *Quant. Inf. Comput.* **16** (2016) 1379.

[76] R. Kueng and D. Gross, *Qubit stabilizer states are complex projective 3-designs*, [arXiv:1510.02767](#).

[77] Z. Ji, Y.-K. Liu and F. Song, *Pseudorandom States, Non-Cloning Theorems and Quantum Money*, [arXiv:1711.00385](#).

[78] H. Liu and S.J. Suh, *Entanglement Tsunami: Universal Scaling in Holographic Thermalization*, *Phys. Rev. Lett.* **112** (2014) 011601 [arXiv:1305.7244] [INSPIRE].

[79] H. Liu and S.J. Suh, *Entanglement growth during thermalization in holographic systems*, *Phys. Rev.* **D 89** (2014) 066012 [arXiv:1311.1200] [INSPIRE].

[80] Y. Gu, A. Lucas and X.-L. Qi, *Spread of entanglement in a Sachdev-Ye-Kitaev chain*, *JHEP* **09** (2017) 120 [arXiv:1708.00871] [INSPIRE].

[81] Y.-Z. You and Y. Gu, *Entanglement Features of Random Hamiltonian Dynamics*, arXiv:1803.10425 [INSPIRE].

[82] A.R. Brown, D.A. Roberts, L. Susskind, B. Swingle and Y. Zhao, *Complexity, action and black holes*, *Phys. Rev.* **D 93** (2016) 086006 [arXiv:1512.04993] [INSPIRE].

[83] A.R. Brown, D.A. Roberts, L. Susskind, B. Swingle and Y. Zhao, *Holographic Complexity Equals Bulk Action?*, *Phys. Rev. Lett.* **116** (2016) 191301 [arXiv:1509.07876] [INSPIRE].

[84] A.R. Brown and L. Susskind, *Second law of quantum complexity*, *Phys. Rev.* **D 97** (2018) 086015 [arXiv:1701.01107] [INSPIRE].

[85] P. Hayden, S. Nezami, X.-L. Qi, N. Thomas, M. Walter and Z. Yang, *Holographic duality from random tensor networks*, *JHEP* **11** (2016) 009 [arXiv:1601.01694] [INSPIRE].

[86] B. Collins and S. Matsumoto, *Weingarten calculus via orthogonality relations: new applications*, *ALEA, Lat. Am. J. Probab. Math. Stat.* **14** (2017) 631.

[87] A. Goupil and G. Schaeffer, *Factoring n-cycles and counting maps of given genus*, *Eur. J. Combinatorics* **19** (1998) 819.

[88] K. Davidson and S. Szarek, *Local Operator Theory, Random Matrices and Banach Spaces*, North-Holland, Amsterdam, (2001).